

架構

AWS Well-Architected Framework



AWS Well-Architected Framework: 架構

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon 的商標和商業外觀不得用於任何非 Amazon 的產品或服務，也不能以任何可能造成客戶混淆、任何貶低或使 Amazon 名譽受損的方式使用 Amazon 的商標和商業外觀。所有其他非 Amazon 擁有的商標均為其各自擁有者的財產，這些擁有者可能隸屬於 Amazon，或與 Amazon 有合作關係，或由 Amazon 贊助。

Table of Contents

摘要和介紹	1
簡介	1
定義	2
論架構	3
一般設計原則	4
架構的六大支柱	6
操作效能	6
設計原則	6
定義	7
最佳實務	8
資源	14
安全	15
設計原則	15
定義	16
最佳實務	16
資源	24
可靠性	25
設計原則	25
定義	26
最佳實務	26
資源	30
效能效率	31
設計原則	31
定義	32
最佳實務	32
資源	36
成本最佳化	37
設計原則	37
定義	38
最佳實務	38
資源	43
永續性	43
設計原則	44
定義	45

最佳實務	45
資源	50
審查程序	51
結論	53
貢獻者	54
深入閱讀	55
文件修訂	56
附錄：問題與最佳實務	59
操作效能	59
組織	59
準備	105
操作	163
演進	200
安全	216
安全基礎	216
身分與存取管理	236
偵測	282
基礎設施保護	294
資料保護	315
事件反應	343
應用程式安全	362
可靠性	381
基礎	381
工作負載架構	416
變更管理	455
故障管理	493
效能效率	576
架構選擇	576
運算與硬體	589
資料管理	604
聯網與內容交付	624
程序和文化	647
成本最佳化	660
實作雲端財務管理	660
了解支出和用量	679
具有經濟效益的資源	714

管理需求與供應資源	747
隨時間優化	758
永續性	764
區域選擇	765
因應需求	767
軟體和架構	779
資料	789
硬體和服務	805
程序和文化	813
注意	823
AWS 詞彙表	824

AWS Well-Architected Framework

發布日期：2024 年 11 月 6 日 ([文件修訂](#))

AWS Well-Architected Framework 可協助您了解在 AWS 上建置系統時所做決策的優缺點。透過使用架構，您將了解在雲端設計和操作可靠、安全、有效率、經濟實惠且永續的系統的架構最佳實務。

簡介

AWS Well-Architected Framework 可協助您了解在 AWS 上建置系統時所做決策的優缺點。透過此架構，您將了解架構的最佳實務，以便在 AWS 雲端設計和操作安全、可靠、有效率、經濟實惠且永續的工作負載。其可讓您根據最佳實務以一致的方式來衡量架構，並識別需要改善的區域。審核架構的程序是與架構決策相關的建設性討論，而不是稽核機制。我們相信，擁有架構良好的系統可大幅提高企業成功的可能性。

AWS 解決方案架構師對於橫跨廣泛的各種垂直業務和使用案例建構解決方案，已累積多年經驗。我們已協助設計及審查數千套客戶在 AWS 上的架構。從這些經驗當中，我們已識別在雲端建構系統的最佳實務和核心策略。

AWS Well-Architected Framework 記錄一組基本問題，有助於您了解特定架構是否妥善符合雲端最佳實務的條件。該架構提供一致的方針，可依照您預計自現代雲端系統可獲得的品質來評估系統，並能得知欲達到此等品質會需要的修補措施。AWS 持續在演進當中，我們也不斷地從與客戶一同工作之中學到更多，因此架構完善的定義會始終精進下去。

本架構適用於擔任技術職務的人員，例如技術長 (CTO)、架構師、開發人員和營運團隊成員。內容說明設計及操作雲端工作負載時運用的 AWS 最佳實務和策略，並提供連結，可取得進一步實作的詳細資訊，和架構模式。如需詳細資訊，請參閱 [AWS Well-Architected 首頁](#)。

AWS 也免費提供您審查工作負載的服務。[AWS Well-Architected Tool](#) (AWS WA Tool) 是雲端服務，可為您提供一致的程序以審查和測量使用 AWS Well-Architected Framework 的架構。AWS WA Tool 提供的建議可讓工作負載更可靠、安全、有效率且經濟實惠。

為協助您套用最佳實務，我們已建立 [AWS Well-Architected 實驗室](#)，它可為您提供程式碼與文件儲存庫，給您實作最佳實務的實際經驗。我們也與精選 AWS 合作夥伴網路 (APN) 合作夥伴組成團隊並肩合作，而這些合作夥伴是 [AWS Well-Architected 合作夥伴計畫](#) 的成員。這些 AWS 合作夥伴對 AWS 擁有深入的知識，能協助您審查和改進工作負載。

定義

AWS 的專家每一天都在輔助客戶進行系統架構，善用雲端的最佳實務。當您的設計演進時，有我們一同進行架構上的權衡。您將這些系統部署至即時環境後，我們可得知這些系統的效能狀況，以及這些權衡形成的後果。

我們便是基於得到的專業知識建立起 AWS Well-Architected Framework，其提供一套一致的最佳實務，供客戶和合作夥伴評估架構；並提供一份問題，您可用來評估架構與 AWS 最佳實務的吻合程度。

AWS Well-Architected Framework 以六個支柱為基礎：卓越營運、安全性、可靠性、效能達成效率、成本最佳化和永續性。

表 1. AWS Well-Architected Framework 的支柱

\$.Name	描述
卓越營運	可有效支援開發和執行工作負載、深入了解其營運狀況，以及持續改善支援流程和程序以產生商業價值的能力。
安全	安全支柱描述如何利用雲端技術來保護資料、系統和資產，從而改善您的安全狀態。
可靠性	可靠性支柱包括工作負載如預期般正確、一致地執行其預期功能的能力。包括在整個生命週期中執行及測試工作負載。本白皮書深入說明在 AWS 上實作可靠工作負載的相關事項，提供最佳實務指引。
效能效率	能夠有效率地使用運算資源，以滿足系統需求，並隨著需求變更與技術發展來保持該效率需求。
成本最佳化	能夠以最低的價格執行系統來提供商業價值。
永續性	能夠透過獲取所佈建資源的最大效益，並將所需的總資源數降至最低，而減少工作負載所有組件的能源消耗、提高效率，最終持續改善永續性影響。

在 AWS Well-Architected Framework 中，我們會使用下列術語：

- 元件：程式碼、組態及 AWS 資源，這些會一起交付以滿足需求。一個元件往往是技術擁有的單元，並自其他元件所解偶。
- 術語工作負載是指一組一起提供業務價值的元件。工作負載通常是商業和技術領導人溝通所談及的最細節的內容。
- 我們心目中的架構是指工作負載之中元件一同運作的方式。元件通訊與互動的方式往往成為架構圖的焦點。
- 里程碑標示架構於產品生命週期之中演進的重要改變 (設計、實作、測試、上線，投入生產)。
- 在組織內，技術組合是業務運作所需工作負載的集合。
- 工作量是將任務針對實作所需的時間、工作和複雜性進行分類。每個組織都需要考慮團隊的大小和專業知識，以及工作負載的複雜性，以取得其他內容，將組織的工作量適當地分類。
 - 高：工作可能需要數週或數個月。這可以分成多個案例、版本和任務。
 - 中：工作可能需要數天或數週。這可以分成多個版本和任務。
 - 低：工作可能需要數小時或數天。這可以分成多個任務。

建立工作負載的架構時，您可依照業務環境，在各支柱之間作出權衡。這些業務決策可以讓您了解工程設計的優先順序。您可以最佳化以開發環境中的可靠性作為代價改善永續性影響並降低成本，或者針對關鍵任務解決方案，以較高成本和永續性影響達到可靠性的最佳化。在電子商務解決方案中，效能能影響營收和客戶購買的傾向。安全和卓越營運通常不會因其他要件而被犧牲。

論架構

在內部部署環境中，客戶經常具備負責技術架構的集中團隊 (而技術架構將疊覆在其他產品或功能團隊上) 以確認其過程遵照最佳實務。技術架構團隊通常包含一組角色，例如技術架構師 (基礎設施)、解決方案架構師 (軟體)、資料架構師、聯網架構師和安全架構師。這些團隊經常使用 [TOGAF](#) 或 [Zachman 框架](#) 作為企業架構能力的部分。

在 AWS 上，我們偏好將能力分散至團隊中，不以集中團隊具備該項能力。選擇將決策權分散有其風險存在，例如為了確認團隊符合內部標準之際。我們以兩種方式降低這類風險。首先，我們演練 (做事方式、程序、標準，及可接受的規範)，目的是讓各個團隊具備該項能力，並且聘用專家，確認該團隊提高所需符合標準的標竿。第二，我們實作機制來實施自動化檢查，確認其符合標準。

 Jeff Bezos 說道「立意良好是不夠的，需要以良好的機制才能有所實現」。

這相當於將人為的盡力取代為機制，其能夠檢查是否遵循規則或程序 (經常為自動化形式)。這種分散式的作法受到 [Amazon 領導方針](#) 的支援，遍及所有角色培養一種返向工作從客戶需求出發的文化。返向工作是我們創新程序的基礎部分。我們從客戶及其期望著手，根據之定義並主導我們的工作方向。以客戶為尊的團隊會因應客戶的需要建置產品。

對架構而言，這表示我們期望每個團隊皆有能力建立架構，並且遵照最佳實務。為協助新團隊獲得這些能力，或讓現有團隊提高標竿，我們促成與首席工程師的虛擬社群聯繫，委請審查團隊的設計，並協助團隊了解有哪些 AWS 最佳實務。首席工程設計社群使得最佳實務成為可見並可取得。例如，他們的一種作法是藉由午餐會報，專講將最佳實務套用至實際範例。這些會報經過錄製，可作為新進團隊成員的到任參考資料。

AWS 最佳實務源自我們以網際網路規模執行數千套系統所累積的經驗。我們偏好以資料定義最佳實務，不過也會起用主題專家，例如首席工程師進行訂定。首席工程師會在看見新的最佳實務出現時，採取社群方式進行工作，以確認團隊會遵守這些實務。假以時日，這些最佳實務會正式列入我們內部的審查程序，同時成為落實合規的機制。Well-Architected 架構是我們內部審查程序面向客戶的實作版，透過我們遍及領域的角色例如「解決方案架構」和內部工程設計團隊，將首席工程設計思維予以編撰。Well-Architected 架構是可擴展的機制，讓您能夠善用這些學習成果帶來的優勢。

依循這種對於架構的責任採取分散形式的首席工程設計社群作法，我們相信 Well-Architected 企業架構能因應客戶的需要而成形。技術領導者 (例如技術長或開發經理) 遍及您所有工作負載執行 Well-Architected 審查，讓您更了解技術組合所具的風險。採行此方式之下，您可看出遍及團隊的主題，您的組織能以機制、培訓或午餐會報妥善顧及，如此一來首席工程師可向多個團隊分享對於特定領域的想法。

一般設計原則

Well-Architected 架構會確定一組一般設計原則，以促進在雲端進行良好的設計：

- 停止猜測您的容量需求：如果您在部署工作負載時做出糟糕的容量決定，可能最後變成坐擁昂貴的閒置資源，或處理容量有限的效能影響。而利用雲端運算，這些問題都會消失。您可依照需要使用大小不拘的容量，並自動擴展和縮減。
- 生產規模測試系統：在雲端，您可隨需建立生產規模的測試環境、完成測試，再將資源除役。因為您只為執行中的測試環境付費，所以能以與內部部署測試相較之下相當微小比例的成本來模擬即時環境。
- 考量架構試驗的自動化：自動化可讓您用低成本建立並複製工作負載，避免產生人工開支。您可追蹤自動化的變更，稽核其影響，並可視需要還原為先前參數。
- 考量演進的架構：在傳統環境中，架構上的決策往往實作成為靜態的一次性活動，其生命週期當中只有系統的少數主要版本。隨著業務及其環境持續改變，這些初始決定可能妨礙系統，無法符合不斷改

變的業務要求。在雲端，按需自動化與測試的能力，可降低因設計變更而形成衝擊的風險。如此可讓系統隨時間演進，因此企業能以標準實務的形式享有創新的優勢。

- 使用資料來驅動架構：在雲端，您可收集架構上的選擇對於工作負載的行為有何影響的資料。如此可讓您為如何提升工作負載，做出以事實為根據的決策。您的雲端基礎設施為程式碼，因此可隨時間利用該資料得知基礎設施的適當選擇及提升。
- 透過演練日進行改進：為測試您的架構與程序的執行情況，可定期排定演練日，以模擬生產中的活動。如此可協助您了解何處有改善空間，並能協助發展組織處理活動的經驗。

架構的六大支柱

建立軟體系統很像是在興建大樓。若基礎不牢固，結構問題可能會逐漸影響建築物的完整性和功能。建構技術解決方案時，若您忽略卓越營運、安全性、可靠性、效能效率、成本最佳化和永續性這六大支柱，那麼建置滿足您期望與需求的系統將成為一項挑戰。將這些支柱納入您的架構，可協助您產出穩定又高效的系統。如此可允許您聚焦在設計的其他面向，例如功能要求。

支柱

- [操作效能](#)
- [安全](#)
- [可靠性](#)
- [效能效率](#)
- [成本最佳化](#)
- [永續性](#)

操作效能

卓越營運 (OE) 是一項承諾，代表致力於妥善設計軟體，同時持續提供絕佳的客戶體驗。卓越營運支柱包括組織團隊、設計工作負載、大規模運作工作負載及促使其進化的最佳實務。

卓越營運支柱概述了設計原則、最佳實務和相關問題。您可以在[卓越營運支柱白皮書](#)中找到實作指引。

主題

- [設計原則](#)
- [定義](#)
- [最佳實務](#)
- [資源](#)

設計原則

下列設計原則有助於實現雲端中的卓越營運：

- 圍繞業務成果組織團隊：團隊實現業務成果的能力來自於領導力願景、高效運營以及與業務保持一致的運營模式。領導力應該充分投入，並致力於使用合適的雲端作業模式進行 CloudOps 轉型，以鼓勵

團隊以最有效的方式營運並達成業務成果。正確的營運模式會利用人員、流程和技術能力來擴展、最佳化生產力，並透過敏捷性、回應性和適應性來實現差異化。組織的長期願景轉化為目標，並在整個企業內傳達給雲端服務的利益相關者和消費者。目標和營運 KPI 在各個層面都一致。這種做法維持了實作下列設計原則所帶來的長期價值。

- 實作可觀測性以獲得可採取行動的見解：全面了解工作負載的行為、效能、可靠性、成本和運作狀態。建立關鍵績效指標 (KPI)，並利用可觀測性遙測來做出明智的決策，並在業務成果有風險時迅速採取行動。根據可採取行動的可觀測性資料，主動改善效能、可靠性和成本。
- 盡可能安全地自動化：在雲端，您可以在整個環境中套用與您應用程式程式碼所用相同的工程原則。可以將整個工作負載及其操作 (應用程式、基礎設施、組態和程序) 定義為程式碼，然後進行更新。然後，可以透過回應事件來啟動工作負載操作，從而將其自動化。在雲端中，可以透過設定防護機制來採用自動化安全性，包括速率控制、錯誤閾值和核准。透過有效的自動化，可以實現對事件的一致回應，限制人為錯誤並減少操作員的辛勞。
- 進行頻繁、細微和可逆的變更：設計可擴展且鬆散耦合的工作負載以允許定期更新元件。自動化部署技術加上較細微的增量變更可縮減影響範圍，並在發生故障時更快地反轉情況。這增加了信心，可以為您的工作負載提供有益的變化，同時保持品質並快速適應市場情況的變化。
- 經常改進營運程序：隨著工作負載的進化，適當地發展您的營運。在使用營運程序時，尋找機會予以改善。定期審查並驗證所有程序是否有效以及團隊是否熟悉這些程序。如果發現差距，請相應地更新程序。向所有利益相關者和團隊傳達程序更新。將營運遊戲化以分享最佳實務並教導團隊。
- 預料失敗：透過推動故障情境來了解工作負載的風險狀況及其對業務成果的影響，從而最大程度提高營運成功率。針對這些模擬失敗，測試程序的有效性和團隊的回應。制定明智的決策，以管理您的測試所識別的開放式風險。
- 從所有營運事件和指標中學習：透過從所有營運事件和失敗中學習的經驗來推動改進。跨團隊及在整個組織中分享獲得的經驗。學習應強調有關營運如何為業務成果做出貢獻的資料和軼事。
- 使用受管服務：盡可能地使用 AWS 受管服務以降低營運負擔。圍繞與這些服務的互動建置營運程序。

定義

雲端有四種最佳實務領域可實現卓越營運：

- 組織
- 準備
- 營運
- 演進

組織的領導階層定義業務目標。您的組織必須了解需求和優先順序，並使用這些來組織和執行工作以支援業務成果的實現。您的工作負載必須提供支援工作負載所需的資訊。透過自動化重複程序，實作服務以實現工作負載的整合、部署和交付將為生產帶來更多有利的變更。

工作負載的操作本質上就可能存在著風險。了解這些風險，並做出明智的決策才能進入生產階段。您的團隊必須能夠支援您的工作負載。從所需業務成果衍生的業務和營運指標，將允許您了解工作負載的運作狀態、營運活動，並回應事件。您的優先事項會隨著業務需求和業務環境的變化而改變。運用這些方面做為回饋迴圈，以持續改善貴組織和工作負載的運作。

最佳實務

Note

所有卓越營運問題都有 OPS 字首作為支柱的速記。

主題

- [組織](#)
- [準備](#)
- [操作](#)
- [演進](#)

組織

您的團隊必須對您的整個工作負載以及團隊成員在其中的作用達成共識，並且擁有共同的業務目標，以便設定能實現業務成功的優先事項。明確定義的優先事項將實現工作的最大收益。評估內部與外部客戶需求，並讓關鍵利益相關者 (包括業務、開發和營運團隊) 參與進來，以確定工作的重點領域。評估客戶需求將驗證您對實現業務成果所需的支援有透徹的了解。確保您了解由貴組織管控所定義的、可能要求或強調特定重點的準則或義務以及外部因素，例如法規合規要求和產業標準。確認您是否設有識別內部管控和外部合規要求變更的機制。如果未識別要求，請確保您已對此決定進行盡職調查。定期審查您的優先事項，以便在需求變更時更新優先事項。

評估對業務的威脅 (例如，業務風險和負債以及資訊安全威脅)，並將此資訊保存在風險登記表內。評估風險的影響，以及利益衝突或替代方法之間的權衡。例如，新功能加速上市可能是成本最佳化所強調的重點，或您可為非關聯式資料選擇關聯式資料庫，以簡化系統遷移工作而無須重構。管理收益和風險，以便在確定工作重點時做出明智的決定。某些風險或選擇可能在一段時間內是可以接受的，相關風險可能得以減輕，也可能出現無法接受風險存在的事實，在此情況下，您將需要採取動作來解決風險。

您的團隊必須了解其在達成業務成果中所扮演的角色。團隊必須了解自己在促成其他團隊成功的過程中所扮演的角色、其他團隊在促進其成功的過程中所扮演的角色，以及擁有共同目標。了解責任、擁有權、決策方式，以及誰有權制定決策，將有助於找到工作重點，並充分發揮團隊的優勢。團隊的需求將由其所支援的客戶、組織、團隊組成，以及工作負載的特性形塑而成。合理來說，無法要求單一操作模式支援貴組織中的所有團隊及其工作負載。

驗證每個應用程式、工作負載、平台和基礎設施元件都有已識別擁有者，而且每個流程和程序都有負責其定義的已識別擁有者，以及負責其執行的擁有者。

透過了解每個元件、流程和程序的商業價值、為何部署這些資源或為何執行活動，以及該擁有權為何存在，有助於團隊成員採取適當動作。明確界定團隊成員的責任，以便他們可以採取適當行動，並具有識別責任和擁有權的機制。擁有請求增加、變更和例外狀況的機制，這樣您就不會限制創新。定義團隊之間的協議，說明他們如何協同合作以互相支援並實現業務成果。

為您的團隊成員提供支援，讓他們能夠更有效地採取動作以及支援業務成果。參與的高階領導層應設定期望並衡量成功。資深領導階層應是採用最佳實務和組織演進的發起者、倡導者和推動者。當成果出現風險時，讓團隊成員採取行動，將影響降到最低，同時鼓勵他們在遇到風險時，向決策者和利益相關者呈報，以便處理問題並避免事故。針對已知風險和計劃事件進行及時、明確且可採取動作的溝通，讓團隊成員能夠及時採取適當的動作。

鼓勵試驗以加速學習，讓團隊成員保持興趣並積極參與。團隊必須發展自己的技能集，以採用新技術，並支援需求和責任的變更。提供專門的結構化時間用於學習，以支援並鼓勵這一舉措。驗證團隊成員擁有可取得成功並進行擴展的資源 (包括工具和團隊成員)，以協助達成您的業務成果。利用跨組織的多樣性，尋求多種獨特的觀點。使用此觀點來增加創新、挑戰假設，並降低確認偏差的風險。在團隊中增加包容性、多樣性和可及性，以獲得有益的觀點。

若有適用於貴組織的外部法規或合規要求，則您應使用 [AWS 雲端合規](#) 提供的資源來協助教育您的團隊，以便他們可以判斷對您的優先事項的影響。Well-Architected 架構強調學習、衡量和改善。它為您提供可評估架構並實作將隨時間擴展之設計的一致方法。AWS 提供 AWS Well-Architected Tool 以協助您在開發前審核您的方法、生產前的工作負載狀態以及生產中工作負載的狀態。您可以將工作負載與最新的 AWS 架構最佳實務做比較、監控工作負載的整體狀態，以及深入了解潛在風險。AWS Trusted Advisor 是一款可存取核心檢查集的工具，這些檢查提出了優化建議，可能有助您確定優先事項。商業和企業支援客戶可存取針對安全性、可靠性、效能、成本優化以及永續性的其他檢查，從而進一步協助確定他們的優先事項。

AWS 可以協助您教育您的團隊有關 AWS 及其服務的知識，從而增進他們對自己的選擇會如何影響工作負載的了解。使用 AWS 支援 (AWS 知識中心、AWS 論壇和 AWS 支援中心) 和 AWS 文件中的資源來教育您的團隊。請透過 AWS 支援 Center 聯絡 AWS 支援，尋求 AWS 相關問題的協助。AWS 還分享了我們透過在 Amazon 建置者資料中心的 AWS 操作中學到的最佳實務和模式。您可透過 AWS 部

落格和官方 AWS 播客獲得其他各種實用資訊。AWS Training and Certification 透過 AWS 基礎原理中的自主進度數位課程提供一些培訓。您還可以報名參加講師指導下的培訓，以進一步協助開發團隊的 AWS 技能。

使用能集中管控跨帳戶環境的工具或服務，例如 AWS Organizations，以便協助您管理操作模式。AWS Control Tower 等服務會擴大此管理功能，允許您定義帳戶設定的藍圖 (支援您的操作模式)、使用 AWS Organizations 套用持續管控，以及自動化新帳戶的佈建作業。AWS Managed Services、AWS Managed Services 合作夥伴等受管服務供應商，或 AWS 合作夥伴網路中的受管服務供應商，都會提供實作雲端環境的專業知識，並支援您的安全和合規要求及業務目標。將受管服務加入操作模式後，便可節省時間和資源，讓您的內部團隊精簡並專注於將使您的企業脫穎而出的策略性成果，而非開發新技能和功能。

下列問題著重於卓越營運方面的這些考量。(如需卓越營運問題清單和最佳實務，請參閱[附錄](#))。

OPS 1：如何決定您的優先事項？

每個人都必須了解自己在實現商業成功中的角色。擁有共同目標以設定資源優先順序。這會充分發揮您所做努力的優勢。

OPS 2：如何建構組織以支援業務成果？

您的團隊必須了解其在達成業務成果中所扮演的角色。團隊必須了解自己在促成其他團隊成功的過程中所扮演的角色、其他團隊在促進其成功的過程中所扮演的角色，以及擁有共同目標。了解責任、擁有權、決策方式，以及誰有權制定決策，將有助於找到工作重點，並充分發揮團隊的優勢。

OPS 3：您的組織文化如何支援您的業務成果？

為您的團隊成員提供支援，讓他們能夠更有效地採取動作以及支援業務成果。

您可能會發現，您在某個時間點會想要強調一小部分的優先事項。長期利用平衡的方法，以驗證開發所需的功能和管理風險。定期審查優先事項，並隨需求的變更進行更新。如果責任和擁有權未定義或未知，則您會面臨風險，不僅無法及時執行必要的動作，在解決這些需求時還會出現冗餘和可能相互衝突的工作。組織文化對團隊成員工作滿意度和留任率有直接影響。讓團隊成員參與其中並習得能力，以實現業務成功。必需要經由試驗才能實現創新，並讓想法轉化為成果。辨識不想要的結果是代表這是一場成功的試驗，因為可以判斷出不會通往成功途徑。

準備

要為卓越營運做好準備，您必須了解您的工作負載及其預期行為。然後，您就能將其設計出來，以了解它們的狀態並建置可提供支援的程序。

設計您的工作負載，使其提供必要資訊，讓您了解所有元件的內部狀態 (例如，指標、日誌、事件和追蹤)，以支援可觀測性和調查問題。可觀測性不僅是單純的監控，還可根據系統的外部輸出全面了解系統的內部運作狀況。基於指標、日誌和追蹤，可觀測性為系統行為和動態提供深刻的見解。透過有效的可觀測性，團隊可以辨別模式、異常情況和趨勢，讓他們能夠主動解決潛在問題並維持最佳的系統運作狀態。確定關鍵績效指標 (KPI) 對於確保監控活動與業務目標之間的一致性至關重要。這種一致性可確保團隊使用真正重要的指標來做出資料驅動的決策，從而最佳化系統效能和業務成果。此外，可觀測性使企業能夠主動出擊，而不是被動應對。團隊可以了解系統內的因果關係，預測和預防問題，而不僅僅是對問題做出反應。隨著工作負載的演進，務必重新檢視並改進可觀測性策略，以保持相關性和有效性。

採用的方法需能夠改善變更進入生產環境的流程，並實現重構、快速品質意見回饋及錯誤修復。這會加快有助益的變更進入生產環境的速度、限制部署問題，並快速識別和修復部署活動所導致或在您的環境中所發現的問題。

採用可快速提供品質意見回饋，並從成果不盡理想的改變中快速復原的方法。使用這些實務可緩解部署變更所帶來問題的影響。為變更失敗做好規劃，以便在必要時能夠快速回應，同時測試並驗證所做變更。了解環境中的計劃內活動，以便管理會影響計劃內活動的變更風險。強調頻繁、細微、可逆的變更，以限制變更範圍。透過回復變更，可以更快地進行疑難排解和修復。這也表示您從有價值變更中受益的頻率會提高。

評估工作負載、流程、程序及人員的營運準備度，以了解與工作負載相關的營運風險。使用一致的程序 (包括手動或自動檢查清單) 來獲悉工作負載或變更執行就緒的時間。這樣也有助於尋找您必須制定計畫以解決問題的任何領域。具備可記錄例行活動的執行手冊，以及可指引問題解決程序的程序手冊。了解收益和風險，以做出明智決策，讓變更順利進入生產環境。

AWS 允許您以程式碼檢視您的整個工作負載 (應用程式、基礎設施、原則、管控和營運)。這表示您可以將用於應用程式程式碼的相同工程規則套用到堆疊的每個元素，並在團隊或組織之間分享這些元素，以擴大開發工作的優勢。在雲端以程式碼執行營運，並利用安全進行試驗的能力，開發工作負載、營運程序以及實務失敗案例。使用 CloudFormation 允許您擁有一致的範本化沙盒開發、測試和生產環境，同時還能提高營運控制等級。

下列問題著重於卓越營運方面的這些考量。

OPS 4：如何在工作負載中實作可觀測性？

在工作負載中實作可觀測性，以便了解其狀態，並根據業務需求做出資料驅動的決策。

OPS 5：您如何減少缺陷、幫助輕鬆修復，以及改善生產流程？

採用可改進生產變更流程的方法，從而重構快速的品質回饋及錯誤修復。這些方法會加快有助益的改變發揮作用的速度、限制部署問題，並快速識別和修復部署活動造成的問題。

OPS 6：您如何緩解部署風險？

採用可快速提供品質意見回饋，並從成果不盡理想的改變中快速復原的方法。使用這些實務可緩解部署變更所帶來問題的影響。

OPS 7：您如何知道自己準備好支援工作負載？

評估工作負載、流程和程序及人員的營運準備度，了解工作負載相關營運風險。

對以程式碼形式實作營運活動進行投資，從而最大程度地提高營運人員的生產力，將錯誤率降至最低以及實現自動回應。使用「事前剖析」可預測失敗並適時建立程序。依照一致的標記策略，使用資源標籤和 AWS Resource Groups 來套用中繼資料，以識別您的資源。標記您的資源，以用於組織、成本會計、存取控制，以及將自動執行營運活動設為目標。採用可利用雲端彈性的部署實務，以促進開發活動和系統的預部署，進而加快實作速度。變更您用於評估工作負載的檢查清單時，請計劃如何處理不再合規的即時系統。

操作

可觀測性讓您能夠專注於有意義的資料，並了解工作負載的互動和結果。透過專注於基本洞察並消除不必要的資料，可以維持一種簡單的方法來了解工作負載效能。不僅要收集資料，還要正確解譯資料，這至關重要。定義明確的基準，設定適當的警示閾值，並主動監控任何偏差。關鍵指標的變化，特別是與其他資料相關時，可以查明特定的問題區域。有了可觀測性，您就具備更優異的預測能力，並且能應付潛在的挑戰，進而確保工作負載順利運行並滿足業務需求。

我們可根據業務和客戶成果的實現情況，衡量是否成功運作工作負載。定義預期成果，確定如何衡量成功，並識別可用於這些計算的指標，以確定工作負載和營運是否成功。運作狀態包括工作負載的運作狀態以及為支援工作負載而執行的營運活動的運作狀態和成功情況 (例如，部署和事故回應)。建立指標基準以便進行改善、調查和介入；收集並分析指標；然後，驗證您對營運成功及其隨著時間的變化情況的理解。使用收集的指標來確定您是否滿足客戶和業務需求，並識別有待改善的領域。

要實現卓越營運，必須高效且有效地管理營運事件。這適用於計劃和非計劃中的營運事件。使用已建立的執行手冊處理眾所周知的事件，並使用程序手冊協助調查和解決問題。根據事件對業務和客戶的影響來確定回應事件的優先順序。驗證如因回應事件而發出提醒，則將由明確識別的擁有者執行關聯程序。事先定義解決事件所需的人員，並納入向上呈報程序，以在必要時根據緊迫性和影響力，在其中新增額外的參與人員。識別並邀請具有權限的個人來決定行動方案，該方案將受到先前未解決的事件回應的業務影響。

透過針對目標受眾 (例如，客戶、業務、開發人員、營運) 量身定制的儀表板和通知來傳達工作負載的運行狀態，以便他們能採取適當的動作，進而管理他們的期望並在恢復正常營運時得到通知。

在 AWS 中，您可以產生從工作負載或以原生方式從 AWS 收集的指標的儀表板視圖。您可以利用 CloudWatch 或第三方應用程式來彙總和呈現營運活動的業務、工作負載和營運層級檢視。AWS 透過 AWS X-Ray、CloudWatch、CloudTrail 以及 VPC Flow Logs 等日誌記錄功能來提供工作負載洞見，以識別工作負載問題，支援根本原因分析和修復。

下列問題著重於卓越營運方面的這些考量。

OPS 8：如何在組織中利用工作負載可觀測性？

利用可觀測性確保最佳的工作負載運作狀況。利用相關指標、日誌和追蹤，全面掌握工作負載效能並有效解決問題。

OPS 9：您如何了解營運狀況？

定義、擷取和分析營運指標，掌握營運事件，以便採取適當行動。

OPS 10：您如何管理工作負載和營運事件？

準備和驗證回應事件的程序，大幅降低工作負載中斷情形。

您收集的所有指標都應該符合業務需求及其支援的結果。開發針對已充分了解之事件的指令碼式回應，並自動化其效能以回應事件辨識。

演進

學習、分享和持續改善以維持卓越營運。將工作週期用於進行幾乎持續的逐漸改善。針對所有影響客戶的事件執行事件後分析。確定促成因素和預防措施，以限制或防止再次發生。適當地與受影響的社區溝通促成因素。定期評估改進機會 (例如，功能請求、問題修復和合規要求) 並確定其優先順序，包括工作負載和營運程序。

在您的程序中納入回饋迴圈，以快速識別有待改善的領域並從正在執行的營運中獲得經驗。

在遊戲日內，可跨團隊分享獲得的經驗，進而分享這些經驗的益處。分析經驗教訓中的趨勢，並對運營指標執行跨團隊回顧性分析，以確定改進的機會和方法。實作旨在帶來改善的變更，並評估結果以判斷是否成功。

在 AWS 中，您可以將日誌資料匯出至 Amazon S3 或直接將日誌傳送至 Amazon S3，以便長期儲存。您可以使用 AWS Glue，在 Amazon S3 中探索和準備日誌資料以進行分析，並將相關聯的中繼資料儲存在 AWS Glue Data Catalog 中。Amazon Athena，透過與 AWS Glue 的原生整合，可用來分析日誌資料，並使用標準 SQL 進行查詢。您可以使用 Amazon Quick 這類商業智慧工具來視覺化、探索和分析資料。探索可能推動改善的感興趣趨勢和事件。

下列問題著重於卓越營運方面的這些考量。

OPS 11：您如何改善營運？

投入時間和資源，盡量持續逐漸改善，以加強營運的效果和效率。

成功的營運演進基於：頻繁、細微的改善；提供安全的環境和時間來試驗、開發和測試改善；鼓勵營造從失敗中學習的環境。隨著營運控制等級的提高，對沙盒、開發、測試和生產環境的營運支援可促進開發，並提高將變更部署至生產中後取得成功結果的可預測性。

資源

請參閱下列資源，進一步了解我們卓越營運的最佳實務。

文件

- [DevOps 和 AWS](#)

白皮書

- [卓越運作支柱](#)

影片

- [Amazon 的 DevOps](#)

安全

安全支柱包含能夠保護資料、系統和資產，以利用雲端技術來改善安全性。

安全支柱概述了設計原則、最佳實務和相關問題。可以在[安全支柱白皮書](#)中找到實作指引。

主題

- [設計原則](#)
- [定義](#)
- [最佳實務](#)
- [資源](#)

設計原則

雲端中有一些原則能協助您強化工作負載的安全：

- 實作堅實的身分識別基礎：實作最低權限原則，並對於每個與 AWS 資源的互動強制執行職責與適當的授權分離。集中進行身分管理，旨在消除對長期靜態憑證的倚賴。
- 維持可追蹤性：即時監控、提醒和稽核您環境中發生的動作和變更。將日誌和指標收集與系統進行整合，以自動調查並採取動作。
- 在所有層級套用安全：透過多個安全控制，套用深度防禦方法。套用至所有層級 (例如，網路邊緣、VPC、負載平衡、每個執行個體和運算服務、作業系統、應用程式和程式碼)。
- 將安全最佳實務自動化：將基於軟體的安全機制自動化，以提高您安全、快速和以具成本效益的方式擴展的能力。建立安全架構 (包括實作控制) 在版本控制的範本中作為程式碼定義和管理。
- 保護傳輸中和靜態資料：將您的資料分為不同的敏感性等級，並使用適當的機制，例如加密、記號化及存取控制。

- 讓人員遠離資料：使用機制和工具，來降低或消除對直接存取或手動處理資料的需要。在處理敏感資料時，這降低了處理不當或修改以及人為錯誤的風險。
- 為安全事件作準備：為事故做好萬全準備，建立與您組織的要求吻合的事件管理和調查政策與程序。執行失敗回應模擬和使用工具與自動化，以提高偵測、調查和復原的速度。

定義

雲端安全有七個最佳實務領域：

- 安全基礎
- 身分與存取管理
- 偵測
- 基礎設施保護
- 資料保護
- 事件回應
- 應用程式安全

在架構任何工作負載之前，您需要採取影響安全性的實務。您會希望控制誰可以做什麼。另外，您需要能夠識別安全事故、保護系統和服務，並透過資料保護維持資料的保密與完整。您應當具備界定完善且經過演練的程序，以因應安全事故。這些工具和技術之所以重要，因為能支援諸多目的，例如防止財務損失或遵循法規義務。

AWS 共同責任模型可幫助採用雲端的組織能夠達成安全與合規目標。AWS 能實體上地保護支援本公司雲端服務的基礎設施，好讓作為 AWS 戶的您專心使用服務以達成目標。AWS 雲端還提供對安全資料更好的存取，並有自動方式可回應安全事件。

最佳實務

主題

- [安全基礎](#)
- [身分與存取管理](#)
- [偵測](#)
- [基礎設施保護](#)
- [資料保護](#)

- [事件反應](#)
- [應用程式安全](#)

安全基礎

下列問題著重於這些安全方面的考量。(如需安全問題和最佳實務的清單，請參閱[附錄](#)。)

SEC 1：如何安全地操作工作負載？

若要安全地操作工作負載，您必須將整體的最佳實務套用到每個安全區域。採用您在組織和工作負載層級於卓越營運中定義的要求和程序，並將其應用於所有區域。

透過 AWS 和產業建議與威脅情報持續取得最新資訊，可協助您發展威脅模型和控制目標。將安全程序、測試和驗證自動化可讓您擴展安全操作。

在 AWS 中，建議根據不同的功能和合規或資料敏感性等要求，依帳戶分隔不同的工作負載。

身分與存取管理

Identity and Access Management 是資訊安全計畫的關鍵部分，可確保只有經過授權和身分驗證的使用者和元件，才能以您想要的方式存取您的資源。例如，您應定義主體 (即為可在您的帳戶內執行動作的帳戶、使用者、角色和服務)，建立與這些主體一致的政策，並實作強勢憑證管理。這些權限管理元素構成身份驗證與授權的核心。

在 AWS 中，權限管理主要由 AWS Identity and Access Management (IAM) 服務支援，它讓您可以控制對 AWS 服務和資源的使用者和程式設計存取。您應該套用精細的政策，將權限分配給使用者、群組、角色或資源。您還可以要求使用強式密碼，例如要求複雜性等級、避免重複使用以及強制執行多重要素驗證 (MFA)。您可以將聯合身分驗證與現有目錄服務一起使用。對於要求系統有權存取 AWS 的工作負載，IAM 可以透過角色、執行個體描述檔、聯合身分和臨時登入資料來實現安全存取。

下列問題著重於安全方面的這些考量。

SEC 2：如何管理人員和機器的身分？

處理安全的 AWS 工作負載時，您需要管理兩種身分類型。了解您需要管理和授予存取權的身分類型，有助於確認正確的身分在適當的條件下存取正確的資源。

SEC 2：如何管理人員和機器的身分？

人員身分：管理員、開發人員、操作員及終端使用者需要身分才能存取 AWS 環境和應用程式。這些是貴組織的成員或與您協作的外部使用者，以及透過 Web 瀏覽器、用戶端應用程式或互動式命令列工具與 AWS 資源互動的外部使用者。

機器身分：您的服務應用程式、操作工具和工作負載需要身分，才能向 AWS 服務發出請求，例如讀取資料。這些身分包括在 AWS 環境 (例如 Amazon EC2 執行個體或 AWS Lambda 函數) 中執行的機器。您也可以為需要存取權的外部各方管理其機器身分。此外，您可能還有 AWS 外部機器需要存取 AWS 環境。

SEC 3：如何管理人員和機器的許可？

管理許可，以控制對需要存取 AWS 和工作負載的人員和機器身分的存取。許可控制誰可以在何種條件下存取哪些內容。

登入資料不得在任何使用者或系統之間共用。應使用最低權限的方法以及最佳實務 (包括密碼要求和強制執行 MFA) 來授予使用者存取權限。包括對 AWS 服務的 API 呼叫在內的程式設計存取應使用臨時和有限權限的登入資料 (例如由 AWS Security Token Service 發出的登入資料) 執行。

若使用者想要與 AWS 管理主控台 之外的 AWS 互動，則需要程式化存取權。授與程式化存取權的方式取決於存取 AWS 的使用者類型。

若要授予使用者程式設計存取權，請選擇下列其中一個選項。

哪個使用者需要程式設計存取權？	到	根據
IAM	(建議事項) 將主控台憑證用作臨時憑證來簽署對 AWS CLI、AWS SDK 或 AWS API 的程式設計請求。	請依照您要使用的介面所提供的指示操作。 關於 AWS CLI，請參閱《AWS Command Line Interface 使用者指南》中的登入以進行 AWS 本機開發。

哪個使用者需要程式設計存取權？	到	根據
		如需 AWS SDK，請參閱《AWS SDK 和工具參考指南》中的登入以進行 AWS 本機開發。
人力資源身分 (IAM Identity Center 中管理的使用者)	使用臨時憑證簽署對 AWS CLI、AWS SDK 或 AWS API 的程式設計請求。	請依照您要使用的介面所提供的指示操作。 - 關於 AWS CLI，請參閱《AWS Command Line Interface 使用者指南》中的設定 AWS CLI 來使用 AWS IAM Identity Center。 - 關於 AWS SDK、工具和 AWS API，請參閱《AWSSDK 和工具參考指南》中的IAM Identity Center 驗證。
IAM	使用臨時憑證簽署對 AWS CLI、AWS SDK 或 AWS API 的程式設計請求。	請遵循《IAM 使用者指南》中 使用臨時憑證搭配 AWS 資源 中的指示。

哪個使用者需要程式設計存取權？	到	根據
IAM	(不建議使用) 使用長期憑證簽署 AWS CLI、AWS SDK 或 AWS API 的程式設計要求。	請依照您要使用的介面所提供的指示操作。 - 關於 AWS CLI，請參閱《AWS Command Line Interface 使用者指南》中的使用 IAM 使用者憑證進行驗證。 - 關於 AWS SDK 和工具，請參閱《AWSSDK 和工具參考指南》中的使用長期憑證進行驗證。 - 關於 AWS API，請參閱《IAM 使用者指南》中的管理 IAM 使用者的存取金鑰。

AWS 提供了可以幫助您進行身分和存取管理的資源。若要協助學習最佳實務，請探索我們的實作實驗室，了解如何[管理憑證和驗證](#)、[控制人員存取](#)以及[控制程式設計式存取](#)。

偵測

您可以使用偵測控制來識別潛在的安全威脅或事故。它們是管控框架的重要組成部分，可用於支援品質流程、法律或合規義務以及用於威脅識別和回應工作。偵測控制有不同的類型。例如，建立資產及其詳細屬性的詳細目錄可促進更有效的決策 (和生命週期控制)，以幫助建立營運基準。您還可以使用內部稽核，即檢查與資訊系統相關的控制，以確認實務符合政策和要求，並確保已根據定義的條件設定正確的自動提醒通知。這些控制是重要的反應式因素，可以幫助您的組織識別和了解異常活動的範圍。

在 AWS 中，您可以透過處理日誌、事件和監控來實作偵測控制，以進行稽核、自動分析和警示。CloudTrail 日誌、AWS API 呼叫和 CloudWatch 可監控指標並發出警示，AWS Config 可提供組態歷史。Amazon GuardDuty 是受管威脅偵測服務，可持續監控惡意或未經授權的行為，協助您保護 AWS 帳戶和工作負載。也提供服務層級日誌。例如，您可以使用 Amazon Simple Storage Service (Amazon S3) 記錄存取請求。

下列問題著重於這些安全方面的考量。

SEC 4：您如何偵測和調查安全事件？

從日誌和指標中擷取並分析事件，以獲得可見性。請針對安全事件和潛在威脅採取行動，以協助保護工作負載。

日誌管理對 Well-Architected 工作負載至關重要，原因包括安全/鑑識，以及法規或法律要求等。分析日誌並對其進行回應，以便可以識別潛在的安全事故，這一點至關重要。AWS 提供了讓您能夠定義資料保留生命週期或定義將在何處儲存、存檔或最終刪除資料的功能，從而使日誌管理更易於實作。這使得可預測和可靠的資料處理更加簡單，且更具成本效益。

基礎設施保護

基礎設施保護包括符合最佳實務和組織或監管義務所必需的控制方法，例如深度防禦。這些方法的使用對於雲端或內部部署成功持續營運至關重要。

在 AWS 中，您可以透過使用 AWS 原生技術或透過 AWS Marketplace 獲得的合作夥伴產品和服務，來實作有狀態和無狀態封包檢查。您應該使用 Amazon Virtual Private Cloud (Amazon VPC) 建立一個私有、安全且可擴展的環境，您可以在其中定義拓撲，包括閘道、路由表以及公有和私有子網路。

下列問題著重於安全方面的這些考量。

SEC 5：如何保護您的網路資源？

任何具有某種網路連線能力的工作負載，無論是網際網路或私有網路，都需要多層防禦來協助保護其不受外部和內部網路威脅的影響。

SEC 6：您如何保護運算資源？

工作負載中的運算資源需要多層防禦，以協助防範外部和內部威脅。運算資源包括 EC2 執行個體、容器、AWS Lambda 函數、資料庫服務、IoT 裝置等。

不管是何種類型的環境，建議使用多層防禦。就基礎設施保護而言，許多概念和方法在雲端和內部部署均有效。加強邊界保護、監控入口和出口以及全面的日誌記錄、監控和提醒，對於有效的資訊安全計畫均很重要。

AWS 客戶能夠量身訂製或強化 Amazon Elastic Compute Cloud (Amazon EC2)、Amazon Elastic Container Service (Amazon ECS) 容器或 AWS Elastic Beanstalk 執行個體的組態，並將組態持久保留到一個不變的 Amazon Machine Image (AMI)。然後，無論是由 Auto Scaling 啟動還是手動啟動，使用此 AMI 啟動的所有新虛擬伺服器 (執行個體) 都將獲得此強化組態。

資料保護

在設計任何系統之前，應建立影響安全性的基礎實務。例如，資料分類可基於敏感層級將組織的資料分類，加密則能對未經授權的存取將資料呈現為無法辨識，以保護資料。這些工具和技術之所以重要，因為能支援諸多目的，例如防止財務損失或遵循法規義務。

在 AWS 中，以下實務有助於保護資料：

- 作為 AWS 客戶，您可完全控管資料。
- AWS 讓您可以更輕鬆地加密資料和管理金鑰，包括常規的金鑰輪換。這些可以透過 AWS 輕鬆地自動化或由您手動維護。
- 提供了包含重要內容 (例如檔案存取和變更) 的詳細日誌記錄。
- AWS 設計的儲存系統具有卓越彈性。例如，Amazon S3 Standard、S3 Standard-IA、S3 One Zone-IA 和 Amazon Glacier 都在給定年份內提供 99.99999999% 的物件耐用性。此耐用性等級相當於 0.000000001% 物件年平均預期損失率。
- 版本控制可以作為更大的資料生命週期管理過程的一部分，可以防止意外的覆寫、刪除和類似損害。
- AWS 永遠不會主動移動區域之間的資料。除非您明確使用相關功能或利用提供相關功能的服務，否則放置在某個區域中的內容將保留在該區域中。

下列問題著重於安全方面的這些考量。

SEC 7：您如何分類資料？

資料分類可讓您根據關鍵性和敏感度將資料分類，協助判定適當的保護和保留控制。

SEC 8：您如何保護靜態資料？

實作多重控制來保護靜態資料，以降低未經授權存取或處理不當的風險。

SEC 9：您如何保護傳輸中資料？

實作多重控制來保護傳輸中的資料，以降低未經授權存取或遺失的風險。

AWS 提供多種加密靜態資料和傳輸中資料的方法。我們將功能內建到我們的服務中，讓您可以更輕鬆地加密資料。例如，我們為 Amazon S3 實作了伺服器端加密 (SSE)，讓您可以更輕鬆地以加密形式儲存資料。您還可以安排由 Elastic Load Balancing (ELB) 處理整個 HTTPS 加密和解密過程 (通常稱為 SSL 終止)。

事件反應

即使採用了非常成熟的預防和偵測控制，您的組織仍應建立適當的流程，來回應和緩和安全事故的潛在影響。工作負載的架構嚴重影響團隊在事故期間有效執行、隔離或控制系統，以及將營運恢復到已知良好狀態的能力。在發生安全事故之前布置好工具和存取權限，然後在演練日期間例行練習事故回應，將幫助您確認架構可以適應即時調查和復原。

在 AWS 中，以下實務有助於有效地回應事故：

- 提供了包含重要內容的詳細日誌記錄，例如檔案存取和變更。
- 可以自動處理事件並啟動工具，以透過使用 AWS API 來自動執行回應。
- 您可以使用 AWS CloudFormation 預先佈建工具和「無塵室」。這樣一來，您就可以在安全、隔離的環境中進行鑑識。

下列問題著重於這些安全方面的考量。

SEC 10：您如何預估、回應事件以及從事件中復原？

準備對於及時且有效的調查、回應事件以及從事件中復原至關重要，有助於將對組織的干擾降到最低。

確認您有一種方法可以快速授予安全團隊存取權限，並自動隔離執行個體以及為鑑識收集資料和狀態。

應用程式安全

應用程式安全 (AppSec) 介紹如何為開發工作負載的安全屬性進行設計、建置與測試的整體過程。您應該安排組織成員接受適當訓練，了解您的建置與發佈基礎結構的安全屬性，以及應用自動化來識別出安全問題。

採用應用程式安全測試做為軟體開發生命週期 (SDLC) 與發佈後程序中的常規部分，有助於驗證您可建立一套結構化機制，專門識別、修正應用程式安全問題，以及防止這些問題進入您的生產環境。

您的應用程式開發方法應該在設計、建置與操作工作負載期間納入安全控制。過程當中，可以調整程序，達到持續減少缺陷和最低技術負債。例如，在設計階段中應用威脅建模有助於提早發現設計瑕疵，修正更加簡單，成本節省更多，而不需要等到日後才能進行緩解。

解決瑕疵的成本與複雜性通常比過去在 SDLC 中來得低。最簡單的解決問題方法就是別讓問題發生，因此從使用威脅模型開始，有助於您專注在設計階段的正確成果。隨著 AppSec 計畫逐漸成熟，您可以自動化方式提高測試量，改善意見回饋對建置人員的準確性 (保真度)，同時縮短安全審核所需要的時間。這些動作全都可以改善所建置軟體的品質，並且加快功能進入生產階段。

這些實作準則著重於四個領域：組織與文化、管道的安全性、管道中的安全性以及相依性管理。每個區域都會提供一組可實作原則，並針對設計、建置與操作工作負載的做法提供端對端檢視。

在 AWS 中，您可以使用許多方法來解決應用程式安全計畫問題。當中有一些方法依賴技術，而其他方法則著重在應用程式安全計畫的人員和組織層面。

下列問題著重於這些應用程式安全方面的考量。

SEC 11：如何在橫跨應用程式設計、開發和部署的整個生命週期內融入安全屬性並進行驗證？

人員培訓、使用自動化測試、了解相依性，以及驗證工具和應用程式的安全屬性，有助於減少生產工作負載中發生安全問題的機率。

資源

請參閱以下資源，進一步了解我們的安全最佳實務。

文件

- [AWS 雲端安全](#)
- [AWS 合規](#)
- [AWS 安全部落格](#)
- [AWS 安全成熟度模型](#)

白皮書

- [安全支柱](#)
- [AWS 安全概觀](#)
- [AWS 風險與合規](#)

影片

- [AWS 聯盟的安全狀態](#)
- [共同責任概觀](#)

可靠性

可靠性支柱包括工作負載如預期般正確、一致地執行其預期功能的能力。包括在整個生命週期中執行及測試工作負載。本白皮書深入說明在 AWS 上實作可靠工作負載的相關事項，提供最佳實務指引。

可靠性支柱概述了設計原則、最佳實務和相關問題。可以在[可靠性支柱白皮書](#)中找到實作指引。

主題

- [設計原則](#)
- [定義](#)
- [最佳實務](#)
- [資源](#)

設計原則

雲端可靠性有五個基本的設計原則：

- 自動從故障中復原：透過監控工作負載的關鍵績效指標 (KPI)，您可在達到閾值時啟動自動化。這些 KPI 應為業務價值的衡量指標，而非服務營運的技術方面。如此一來，即可自動通知和追蹤失敗，以及自動化可解決或修復失敗的復原程序。藉助更複雜的自動化功能，您可以在發生失敗前進行預測和修補。
- 測試復原程序：在內部部署環境中，經常執行測試以證明工作負載可在特定情況下正常工作。測試通常不可用於驗證復原策略。在雲端，您可測試工作負載會發生哪些失敗情境，同時可驗證復原程序。

您可使用自動化來模擬不同的失敗情境或重新建立會導致之前失敗的情境。此方法會在實際的失敗情境發生前公開您可以測試和修正的失敗路徑，從而降低風險。

- 水平擴展，以增加彙總工作負載的可用性：使用多個小資源取代一個大資源，以降低整體工作負載上發生單一失敗時造成的影響。將請求分散到多個較小的資源，以確認其不會有共同的失敗點。
- 停止猜測容量：內部部署工作負載失敗的一個常見原因是資源飽和，即當對工作負載的需求超出該工作負載的容量時發生的情況 (這通常為阻斷服務攻擊的目標)。在雲端，您可以監控需求和工作負載利用率，並自動新增或刪除資源，以保持可滿足需求的更有效水平，而不會過度佈建或佈建不足。仍然存在限制，但是某些配額可以控制，而其他限制則可管理 (請參閱管理服務配額和限制)。
- 透過自動化管理變更：應透過自動化來執行對基礎架構的變更。必須管理的變更包括之後可以追蹤和審查的自動化變更。

定義

雲端可靠性有四個最佳實務領域：

- 基礎
- 工作負載架構
- 變更管理
- 故障管理

若要實現可靠性，您必須先從基礎開始，即服務配額和網路拓撲能適應工作負載的環境。分散式系統的工作負載架構在設計上必須能防止失敗並減輕失敗的影響。工作負載必須處理需求或要求的變更，且在設計上須能偵測失敗並自動進行自我修復。

最佳實務

主題

- [基礎](#)
- [工作負載架構](#)
- [變更管理](#)
- [故障管理](#)

基礎

基礎要求是其範圍超過單一工作負載或專案的要求。在建立任何系統架構之前，應確立會影響可靠性的基本要求。例如，您必須為資料中心提供足夠的網路頻寬。

藉助 AWS，這些基礎需求中的大多數已予以納入或可以按需要進行處理。設計的雲端近乎無限，因此 AWS 有責任滿足足夠的聯網和運算容量的要求，允許您根據需要自由變更資源大小和分配。

下列問題著重於可靠性方面的這些考量。(如需可靠性問題清單和最佳實務，請參閱[附錄](#))。

REL 1：您如何管理服務配額和限制？

雲端型工作負載架構具有服務配額 (也稱為服務限制)。這些配額的存在是為了防止意外佈建超出所需數量的資源，並限制 API 操作的請求率，以保護服務免於遭到濫用。另外還有一些資源限制，例如可將位元下推到光纖纜線的速率，或實體磁碟的儲存量。

REL 2：如何規劃您的網路拓撲？

工作負載通常存在於多個環境中。其中包括多個雲端環境 (可公開存取與私有)，也可能包含您現有的資料中心基礎設施。這些計畫必須包含網路考量事項，例如系統內部與系統間連線能力、公有 IP 位址管理、私有 IP 位址管理以及網域名稱解析。

工作負載架構

可靠的工作負載始於對軟體和基礎設施的前期設計決策。您的架構選擇會對所有 Well-Architected 支柱的工作負載行為產生影響。為求可靠性，您必須依循特定模式。

藉助 AWS，工作負載開發人員可以選擇要使用的語言和技術。AWS SDK 透過為 AWS 服務提供特定語言 API，讓編碼不再如此複雜。這些開發套件加上各種語言選項，允許開發人員實作本文列出的可靠性最佳實務。開發人員也可在 [Amazon 建置者資料中心](#) 中閱讀並學習 Amazon 如何構建和操作軟體。

下列問題著重於可靠性方面的這些考量。

REL 3：如何設計您的工作負載服務架構？

使用服務導向架構 (SOA) 或微型服務架構，建置擴展性與可靠性高的工作負載。服務導向架構 (SOA) 是透過服務界面讓軟體元件可重複使用的做法。微型服務架構則進一步讓元件變得更小、更簡單。

REL 4：如何在分散式系統中設計防止失敗的互動？

分散式系統仰賴通訊網路將伺服器或服務等元件互相連線。儘管這些網路中出現資料遺失或延遲，但工作負載仍必須可靠地運作。分散式系統的元件必須以不會對其他元件或工作負載造成負面影響的方式運作。這些最佳實務可防止故障，並改善平均故障間隔時間 (MTBF)。

REL 5：如何設計分散式系統中的互動以緩解或承受故障？

分散式系統倚賴通訊網路來互連元件 (例如，伺服器或服務)。即使這些網路上的資料遺失或延遲，您的工作負載仍必須可靠運作。分散式系統的元件必須以不會對其他元件或工作負載造成負面影響的方式運作。這些最佳實務讓工作負載能夠承受壓力或故障，更快速地從其中復原，並減輕這類受損的影響。最終縮短平均復原時間 (MTTR)。

變更管理

必須預期並因應工作負載或其環境的變更，才能實現可靠的工作負載操作。變更包括對工作負載強加的變更，例如需求峰值，以及內部的變更，例如功能部署和安全性修補程式。

您可以使用 AWS 監控工作負載的行為，並自動化對 KPI 的回應。例如，隨著工作負載的使用者增加，您的工作負載可能會新增其他伺服器。您可以控制有權作出工作負載變更的人員，並稽核這些變更的歷史紀錄。

下列問題著重於可靠性方面的這些考量。

REL 6：如何監控工作負載資源？

日誌和指標是可深入洞察工作負載運作狀態的強大工具。您可以設定工作負載以監控日誌和指標，並在超過閾值或發生重大事件時傳送通知。監控可讓您的工作負載識別何時會超過低效能閾值或發生故障，以便自動復原來回應。

REL 7：如何設計工作負載以適應需求變更？

可擴展的工作負載提供可自動新增或移除資源的彈性，讓資源能夠在任何特定時間點充分滿足目前的需求。

REL 8：您如何實作變更？

變更須在受控的情況下，才能部署新功能，並確認工作負載和運作環境執行已知的軟體，且能夠以可預測的方式修補或取代。如果這些變更不受控制，那麼就很難預測這些變更的影響，也很難解決由於這些變更而產生的問題。

當您建立工作負載架構以根據需求的變更自動新增和刪除資源時，其不僅可以提高可靠性，而且還能驗證企業成功不會成為負擔。在適當監控下，當 KPI 偏離預期規範時，您的團隊將會自動收到提醒。自動日誌記錄對環境的變更，允許您進行稽核並快速識別可能影響可靠性的動作。對變更管理的控制將確保您能執行交付所需可靠性的規則。

故障管理

在任何合理複雜的系統中，均有可能會發生失敗。為達可靠性要求，您的工作負載應在發生失敗時察覺此情況，並採取行動以免影響可用性。工作負載必須能夠承受失敗並自動修復問題。

藉助 AWS，您可以利用自動化對監控資料作出反應。例如，當特定指標超過閾值時，您可以啟動可修補問題的自動化動作。此外，您無需嘗試診斷和修正生產環境中的失敗資源，而是可以用新的資源取代它，並對失敗的額外資源執行分析。由於雲端可讓您以低成本建立整個系統的臨時版本，因此您可以使用自動化測試來驗證完整的復原程序。

下列問題著重於可靠性方面的這些考量。

REL 9：您如何備份資料？

備份資料、應用程式和組態以滿足您對復原時間點目標 (RTO) 和復原點目標 (RPO) 的要求。

REL 10：如何使用故障隔離來保護您的工作負載？

故障隔離會將元件或系統故障的影響侷限在定義的界限內。藉由適當的隔離，界限外部的元件就不會受到故障影響。只要能跨多個故障隔離界限執行工作負載，就可更靈活地應對故障。

REL 11：如何設計工作負載以承受元件失敗？

須架構具高可用性和低平均復原時間 (MTTR) 需求的工作負載以實現彈性。

REL 12：如何測試可靠性？

在將工作負載設計為可彈性應對生產壓力之後，進行測試是確認其依設計運作並提供預期之彈性的唯一方法。

REL 13：您如何規劃災難復原 (DR)？

備妥備份和冗餘工作負載元件是 DR 策略的開始。[RTO](#) 和 [RPO](#) 是您還原工作負載的目標。根據業務需求設定這些目標。實作策略以滿足這些目標，考量工作負載資源和資料的位置和功能。發生中斷的可能性和復原成本也是重要因素，可反映為工作負載提供災難復原的商業價值。

定期備份資料並測試備份檔案，從而確認您可以從邏輯和實際錯誤復原。管理失敗的關鍵是對導致失敗的工作負載頻繁進行自動化測試，然後觀察其可如何復原。定期執行此操作，並確認在出現重大工作負載變更後也能啟動此類測試。主動追蹤 KPI、復原時間點目標 (RTO) 和復原點目標 (RPO)，以評估工作負載的彈性 (尤其是在失敗測試情境下)。追蹤 KPI 將能助您識別和減輕單一失敗點。其目標是徹底測試您的工作負載復原程序，以便您確信即使面對持續問題，您也可以復原所有資料並繼續為客戶提供服務。應與執行正常生產程序一樣執行復原程序。

資源

請參閱以下資源，進一步了解我們的可靠性最佳實務。

文件

- [AWS 文件](#)

- [AWS 全球基礎設施](#)
- [AWS Auto Scaling：擴展計畫的運作方式](#)
- [什麼是 AWS Backup？](#)

白皮書

- [可靠性支柱：AWS Well-Architected](#)
- [實作 AWS 上的微型服務](#)

效能效率

效能效率支柱包括能夠有效率地使用雲端資源，以滿足效能需求，並隨著需求變更與技術發展來保持該效率需求。

效能達成效率支柱概述了設計原則、最佳實務和相關問題。可以在[效能達成效率支柱白皮書](#)中找到實作指引。

主題

- [設計原則](#)
- [定義](#)
- [最佳實務](#)
- [資源](#)

設計原則

雲端有五個設計原則來維持效能達成效率：

- 讓進階技術變得更普及：將複雜的任務委派給雲端廠商，讓團隊更順暢地實作進階技術。與其要求 IT 團隊了解新技術的託管和執行方式，不如考慮使用技術即服務。例如，沒有任何 SQL 資料庫、媒體轉碼和機器學習都是需要專業知識的技術。在雲端，這些技術成為團隊可以使用的服務，讓團隊能夠專注於產品開發，而非資源佈建及管理。
- 幾分鐘內即可全球化：在全球多個 AWS 區域中部署工作負載可讓您以最低成本為客戶提供更低的延遲和更好的體驗。
- 使用無伺服器架構：採用無伺服器架構，您便無需執行和維護實體伺服器來完成傳統運算活動。例如，無伺服器儲存服務可以充當靜態網站 (因此無需 Web 伺服器)，而事件服務可以為您託管程式

碼。如此一來，即可減輕管理實體伺服器的營運負擔，而且由於這些受管服務是在雲端規模上運行，因此還可以降低交易成本。

- 提高試驗頻率：使用虛擬及可自動化的資源，您可以使用不同類型的執行個體、儲存設備或組態，迅速完成比較測試。
- 考慮機械同感：了解雲端服務的使用方式，並一律使用符合工作負載目標的技術方法。例如，在您選擇資料庫或儲存方法時，請考慮資料存取模式。

定義

維持雲端效能達成效率的最佳實務有五個領域：

- 架構選取
- 運算與硬體
- 資料管理
- 聯網與內容交付
- 程序和文化

採取資料驅動的方法來建置高效能架構。從高階設計到選取和設定資源類型，收集架構各方面的資料。

定期檢閱您的選擇可驗證您是否利用不斷發展的 AWS 雲端。監控可確保您能察覺預期效能發生的任何偏差情形。在架構中做出權衡以改進效能，例如使用壓縮或快取，或放寬一致性要求。

最佳實務

主題

- [架構選擇](#)
- [運算與硬體](#)
- [資料管理](#)
- [聯網與內容交付](#)
- [程序和文化](#)

架構選擇

適用於特定工作負載的最佳解決方案各不相同，而解決方案通常會結合多種方法。Well-Architected 工作負載會使用多種解決方案，並採用不同的功能以提升效能。

AWS 資源提供多種類型和組態，讓您更輕鬆地找到最符合您需求的方法。您還可以發現使用內部部署基礎設施不易實現的選項。例如，諸如 Amazon DynamoDB 的受管服務提供完全受管的無SQL資料庫，在任何規模下都具有單位數毫秒延遲。

下列問題著重於效能達成效率方面的這些考量。(如需效能達成效率問題和最佳實務的清單，請參閱[附錄](#)。)

PERF 1：如何為工作負載選取適當的雲端資源和架構模式？

欲讓工作負載達到更有效的效能通常需要採用多種方法。Well-Architected 系統會使用多重解決方案和功能以提升效能。

運算與硬體

特定工作負載的最佳運算選擇會根據應用程式設計、使用模式和組態設定而有所不同。架構會針對不同元件使用不同運算選擇，並採用不同功能以提升效能。若選錯運算資源，可能使架構的效能達成效率降低。

在中 AWS，運算有三種形式：執行個體、容器和函數：

- 執行個體是虛擬化伺服器，可讓您使用按鈕或API呼叫來變更其功能。由於在雲端中，資源決策不是固定的，您可以使用不同的伺服器類型進行試驗。在 AWS 中，這些虛擬伺服器執行個體具有不同的系列和大小，且提供廣泛的功能，包括固態硬碟（SSDs）和圖形處理單元（）GPUs。
- 容器是一種作業系統虛擬化方法，可讓您在資源隔離的程序中執行應用程式及其相依性。如果您需要控制運算環境的安裝、組態和管理，EC2則可以使用容器或 Amazon 的無 AWS Fargate 伺服器運算。您也可以從多個容器協調平台中選擇：Amazon Elastic Container Service（ECS）或 Amazon Elastic Kubernetes Service（EKS）。
- 函數則從您想套用的程式碼中將執行環境抽象化。例如，AWS Lambda 允許您在不執行執行個體的情況下執行程式碼。

下列問題著重於效能達成效率方面的這些考量。

PERF 2：如何在工作負載中選取和使用運算資源？

工作負載的更高效解決方案會根據應用程式設計、使用模式和組態設定而有所不同。架構可針對不同元件使用不同運算解決方案並開啟不同功能，以提升效能。為架構選錯運算解決方案，可能使效能達成效率降低。

資料管理

特定系統的最佳資料管理解決方案會根據資料類型（區塊、檔案或物件）、存取模式（隨機或循序）、必要的輸送量、存取頻率（線上、離線、封存）、更新頻率（WORM、動態），以及可用性和耐久性限制而有所不同。Well-Architected 工作負載會使用專用資料存放區，這些存放區採用不同的功能以提升效能。

在中 AWS，儲存有三種形式：物件、區塊和檔案：

- 物件儲存提供可擴展且耐用的平台，以利從任何網際網路位置存取資料，例如使用者產生的內容、作用中存檔、無伺服器運算、大數據儲存或備份與復原。Amazon Simple Storage Service (Amazon S3) 是一項物件儲存服務，提供領先業界的可擴展性、資料可用性、安全性和效能。Amazon S3 旨在提供 99.999999999% 的耐久性，並為全球公司存放數百萬個應用程式的資料。
- 區塊式儲存為每個虛擬主機提供高可用性、一致、低延遲的區塊式儲存，類似於直接連接的儲存（DAS）或儲存區域網路（SAN）。Amazon Elastic Block Store（Amazon EBS）專為需要 EC2 執行個體可存取持久性儲存的工作負載而設計，可協助您調整具有適當儲存容量、效能和成本的應用程式。
- 檔案儲存可讓您跨多個系統存取共用檔案系統。Amazon Elastic File System（Amazon EFS）等檔案儲存解決方案非常適合使用案例，例如大型內容儲存庫、開發環境、媒體存放區或使用者主目錄。Amazon FSx 讓啟動和執行熱門檔案系統更有效率且符合成本效益，因此您可以利用廣泛使用的開放原始碼和商業授權檔案系統的豐富功能集和快速效能。

下列問題著重於效能達成效率方面的這些考量。

PERF 3：如何在工作負載中儲存、管理和存取資料？

系統更有效率的儲存解決方案會根據存取操作類型（區塊、檔案或物件）、存取模式（隨機或循序）、必要的輸送量、存取頻率（線上、離線、封存）、更新頻率（WORM、動態），以及可用性和耐久性限制而有所不同。Well-Architected 系統使用多重儲存解決方案，並開啟不同功能以提升效能並有效使用資源。

聯網與內容交付

工作負載的最佳聯網解決方案會根據延遲、輸送量需求、抖動和頻寬而有所不同。實體限制（例如使用者或內部部署資源）會決定位置選項。這些限制可能隨著邊緣節點或資源位置而有所差異。

在上 AWS，聯網是虛擬化的，並提供多種不同的類型和組態。這可讓您更輕鬆地符合聯網需求。AWS 提供產品功能（例如增強型網路、Amazon EC2 網路最佳化執行個體、Amazon S3 傳輸加速和動態 Amazon CloudFront）來最佳化網路流量。AWS 也提供聯網功能（例如 Amazon Route 53 延遲路由、Amazon VPC 端點 AWS Direct Connect 和 AWS Global Accelerator），以減少網路距離或抖動。

下列問題著重於效能達成效率方面的這些考量。

PERF 4：如何在工作負載中選取和設定聯網資源？

此問題包括在雲端中設計、設定和操作高效聯網和內容交付解決方案的指引和最佳實務。

程序和文化

在架構工作負載時，您可以採取一些原則和實務來協助您更有效率地執行高效能雲端工作負載。為了培養高效能雲端工作負載的文化，請考慮下列重要原則和實務。

打造這類文化時，請考慮以下重要原則：

- **基礎設施為程式碼：**使用 AWS CloudFormation 範本等方法將基礎設施定義為程式碼。使用範本可讓您將基礎設施與應用程式程式碼和組態一起置於原始檔控制中。這可讓您在基礎設施中套用開發軟體時所使用的相同做法，進而快速進行迭代。
- **部署管道：**使用持續整合/持續部署 (CI/CD) 管道 (例如，原始程式碼儲存庫、建置系統、部署和測試自動化) 來部署您的基礎架構。這樣您就可以在反覆執行的過程中，採用可重複、一致且低成本的方式進行部署。
- **定義明確的指標：**設定和監控指標，以擷取關鍵績效指標 (KPIs)。我們建議您同時使用技術和業務指標。對於網站或行動應用程式，關鍵指標正在擷取 time-to-first-byte 或轉譯。其他一般適用的指標包括執行緒計數、垃圾回收率和等待狀態。業務指標 (例如每個請求的彙總累計成本) 會提示您降低成本的方法。仔細考慮您計劃如何解釋指標。例如，您可以選擇最大值或第 99 個百分位數，而非平均值。
- **自動執行效能測試：**在部署程序中，在成功通過快速執行測試之後，會自動啟動效能測試。自動化應建立一個新的環境，設定如測試資料之類的初始條件，然後執行一系列基準測試和負載測試。這些測試的結果應與組建版本綁定，方便您追蹤長時間的效能變化。對於長期執行的測試，您可以讓管道的這個部分與組建版本的其餘部分不同步。或者，您可以使用 Amazon EC2 Spot 執行個體在夜間執行效能測試。
- **負載產生：**您應建立一系列的測試指令碼來複寫綜合性或預錄的使用者旅程。這些指令碼應該是冪等及非耦合的形式，而且您可能需要納入預熱型指令碼才能產生有效的結果。您的測試指令碼應盡可能

地複寫生產環境中的使用行為。您可以使用軟體或 software-as-a-service (SaaS) 解決方案來產生負載。可以考慮使用 [AWS Marketplace](#) 解決方案和 [Spot 執行個體](#) — 這些可能是經濟實惠的負載產生方式。

- 效能可見度：關鍵指標應對您的團隊可見，尤其是針對每個組建版本的指標。這可讓您查看隨時間變化出現的任何顯著的正面或負面趨勢。您也應顯示錯誤或例外狀況數量的指標，以確保您測試的是可運作的系統。
- 視覺化：使用視覺化技術可以清楚指出何處出現效能問題、熱點、等待狀態或較低的利用率。在架構圖上重疊效能指標 — 呼叫圖表或程式碼有助於快速識別問題。
- 定期審查程序：架構效能不佳通常是效能審查程序不存在或中斷的結果。如果您的架構效能不佳，則實作效能審查程序可讓您不斷反覆進行改善。
- 持續優化：培養文化以持續優化雲端工作負載效能達成效率。

下列問題著重於效能達成效率方面的這些考量。

PERF 5：您使用什麼程序來支援工作負載的更多效能效率？

在架構工作負載時，您可以採取一些原則和實務來協助您更有效率地執行高效能雲端工作負載。為了培養高效能雲端工作負載的文化，請考慮下列重要原則和實務。

資源

請參閱以下資源，進一步了解我們的效能達成效率最佳實務。

文件

- [Amazon S3 效能最佳化](#)
- [Amazon EBS磁碟區效能](#)

白皮書

- [效能達成效率支柱](#)

影片

- [AWS re : Invent 2019 : Amazon EC2基礎 \(CMP211-R2 \)](#)

- [AWS re : Invent 2019：領導工作階段：工會的儲存狀態 STG2 \(01-L \)](#)
- [AWS re : Invent 2019：領導工作階段：AWS 用途建置的資料庫 DAT2 \(09-L \)](#)
- [AWS re : Invent 2019：與 AWS 混合 AWS 網路架構的連線能力 \(NET317-R1 \)](#)
- [AWS re : Invent 2019：推動新一代 AmazonEC2：深入探索 Nitro 系統 \(CMP303-R2 \)](#)
- [AWS re : Invent 2019：擴展到前 1,000 萬使用者 \(ARC211-R \)](#)

成本最佳化

成本最佳化支柱包括以最低價格執行系統來產生商業價值的能力。

成本最佳化支柱概述了設計原則、最佳實務和相關問題。您可以在[成本最佳化支柱白皮書](#)中找到有關實作的規定性指引。

主題

- [設計原則](#)
- [定義](#)
- [最佳實務](#)
- [資源](#)

設計原則

雲端成本最佳化有五個設計原則：

- **實作雲端財務管理：**為了達成財務成功並加速在雲端實現商業價值，請投資雲端財務管理和成本最佳化。您的組織應投入時間和資源，在這個新的技術與使用管理領域中打造能力。與安全或卓越營運能力類似，您需要透過知識累積、計畫、資源和程序打造能力，以成為具成本效率的組織。
- **採用消費模式：**僅為您需要的運算資源付費，依照業務要求增減用量，不必倚賴複雜的預測。例如，開發與測試環境通常僅於一週工作日的一天八小時當中使用。您可在不使用這些資源時加以停止，有潛力可節省 75% 成本 (40 小時相對於 168 小時)。
- **衡量整體效率：**測量工作負載的商業輸出和遞送的相關成本。以此測量值可得知您從增加輸出與降低成本獲取的增益。
- **停止將金錢花在差異不大的繁重工作上：**AWS 會處理資料中心營運的繁重工作，例如架設伺服器。通過受管服務，它也免除了管理作業系統和應用程式這些營運負擔。這可讓您專注於客戶和業務專案，而非 IT 基礎設施。

- 分析和歸因支出：雲端可讓您輕鬆準確識別系統的用量和成本，繼而允許將 IT 成本透明化地歸因至個別工作負載擁有者。如此有助於測量投資報酬率 (ROI)，並且讓工作負載擁有者有機會優化資源和降低成本。

定義

雲端成本最佳化的最佳實務有五個方面：

- 實作雲端財務管理
- 支出和用量感知
- 具有經濟效益的資源
- 管理需求與供應資源
- 隨時間最佳化

如同 Well-Architected 架構內的其他支柱，有權衡事項需要考量，例如，該針對上市速度還是成本進行最佳化。在某些情況下，更有效的方式是針對速度來最佳化，例如快速上市、推出新功能，或滿足截止日期，而不是投資預付成本最佳化。設計決策有時會因倉促而不是資料來引導，因為總是會有「以防萬一」過度補償的趨向，而不是花時間為最經濟實惠的部署做基準化分析測試。這恐怕會導致過度佈建和最佳化不足的部署。不過，若必須將內部部署環境內的資源「平移」至雲端，然後再實施最佳化，這是理性的選擇。前期對成本最佳化策略進行適當投資，並達成一致奉行最佳實務，避免不必要的過度佈建，可讓您更穩當地體現雲端的經濟效益。以下各節提供初始和持續實作工作負載雲端財務管理和成本最佳化的技術和最佳實務。

最佳實務

主題

- [實作雲端財務管理](#)
- [了解支出和用量](#)
- [具有經濟效益的資源](#)
- [管理需求與供應資源](#)
- [隨時間優化](#)

實作雲端財務管理

採用雲端之後，技術團隊因核准、採購和基礎設施部署週期縮短而加快創新速度。實現商業價值和財務成功需要新的雲端財務管理方法。此方法為雲端財務管理，透過在整個組織實作知識建置、計畫、資源和程序，打造整個組織的能力。

許多組織是由許多不同的單位組成，每個單位都具有不同的優先事項。以下能力將協助建立更高效的組織：讓您的組織與一系列約定的財務目標保持一致，並為組織提供達成這些目標所需的機制。有能力的組織將更快速地創新和建立，且面對任何內部或外部因素時更靈活、適應性更強。

在 AWS 中，您可以使用 Cost Explorer、Amazon Athena (選用)、搭配成本和用量報告 (CUR) 的 Amazon QuickSight，在整個組織中提供成本和用量感知。AWS Budgets 可針對成本和用量提供主動通知。AWS 部落格提供新服務和功能的相關資訊，可確認您能夠隨時掌握最新的服務版本。

下列問題著重於成本最佳化方面的這些考量。(如需成本最佳化問題清單和最佳實務的清單，請參閱[附錄](#)。)

COST 1：如何實作雲端財務管理？

透過實作雲端財務管理，就可幫助組織最佳化成本和用量，並且在 AWS 上進行規模調整，進而實現商業價值和財務上的成功。

建立成本最佳化職能部門時，使用團隊成員，並在團隊中增加 CFM 和成本最佳化方面的專家。現有的團隊成員將會了解組織目前的運作方式，以及如何快速實作改善。同時也考慮納入具有輔助或專業技能集的人員，例如分析和專案管理方面的人員。

在組織中實作成本感知時，改善現有的計畫和程序或在此基礎上進行建置。在現有的程序和計畫中新增內容會比建立新的程序和計畫快得多。這會更快實現結果。

了解支出和用量

雲端提供的增強彈性和敏捷性，可促進創新和快節奏開發和部署。它減少了與佈建內部部署基礎設施相關的手動程序和時間，包括識別硬體規格、協商價格報價、管理採購訂單、安排裝運以及部署資源。然而，欲享有易用性和幾乎無限制的隨需容量，對於支柱需要換上新思維。

許多企業是以各種團隊執行多個系統之下所組成。能將資源成本歸因至個別組織或產品擁有者，能帶動高效使用的行為模式，有助於減少浪費。準確的成本歸因可讓您知道哪些產品具有真正的獲利能力，並就預算分配做出更明智的決策。

在 AWS 中，您可以使用 AWS Organizations 或 AWS Control Tower 來建立帳戶結構，如此可實現區隔並協助您分配成本和用量。您也可以對資源使用標記，利用商業和組織資訊確定用量和成本情況。使用 AWS Cost Explorer 查看您的成本和用量，或使用 Amazon Athena 和 Amazon QuickSight 建立自訂儀表板和分析。透過 AWS Budgets 的通知，以及使用 AWS Identity and Access Management (IAM) 和 Service Quotas 的控制，來控制成本和用量。

下列問題著重於成本最佳化方面的這些考量。

COST 2：您如何管控用量？

制訂政策和機制以驗證產生合理的成本，同時達成目標。您可以運用相互制衡的方法，在不超支的情況下進行創新。

COST 3：您如何監控用量和成本？

制訂政策和程序以監控並適當分配成本。這樣能夠讓您衡量並改善此工作負載的成本效益。

COST 4：如何進行資源除役？

從啟動到結束專案期間，控制變更並管理資源。這有助於關閉未使用的資源，以減少浪費。

您可以使用成本分配標籤為 AWS 用量和成本進行分類和追蹤。當您對 AWS 資源 (例如 EC2 執行個體或 S3 儲存貯體) 加上標籤時，AWS 就能以您的用量和標籤產生成本和使用報告。您可加上代表組織類別 (例如成本中心、工作負載名稱或擁有者) 的標籤，以便跨多項服務安排成本。

確認您在成本與用量報告和監控中使用正確的詳細資訊和精細度層級。如需高層級的洞察和趨勢，請透過 AWS Cost Explorer 使用每日精細度。如需更深入的分析 and 檢查，請使用 AWS Cost Explorer 中的每小時精細度，或 Amazon Athena 和搭配成本和用量報告 (CUR) 的 Amazon QuickSight 中的每小時精細度。

將加有標籤的資源結合實體生命週期追蹤 (員工、專案)，可識別不再為組織產生價值且應當除役的孤立資源或專案。您可以設定帳單提醒，通知您預測的超支。

具有經濟效益的資源

為您的工作負載使用適當的執行個體和資源，是節約成本的關鍵。例如，假設報告程序在較小的伺服器上執行時要花五小時，但在兩倍昂貴的較大伺服器上執行只需一小時。這兩種伺服器產出的結果相同，但較小的伺服器經過一段時間會形成較高成本。

架構完善的工作負載會用最具有成本效益的資源，帶來明顯正面的經濟影響。您並有機會可利用受管服務來降低成本。例如，與其維護伺服器以遞送電子郵件，可使用以訊息為單位收費的服務。

AWS 備有各種具有彈性且經濟的定價選項，讓您以更有效地符合需要的方式獲取 Amazon EC2 和其他服務的執行個體。隨需執行個體可讓您按時數為運算容量付費，無最低承諾的要求。Savings Plans 和預留執行個體可提供高達 75% 的隨需定價折扣。使用 Spot 執行個體，您可善用未用的 Amazon EC2 容量，與隨需定價相較可節省高達 90% 的成本。Spot 執行個體適合用在系統能耐受使用伺服器機群之處，其中個別伺服器能動態性地來去，例如無狀態 Web 伺服器、批次處理，或使用 HPC 和大數據時。

選擇適當的服務也能降低用量和成本；例如 CloudFront 能將資料傳輸降至最低或降低成本，例如在 Amazon RDS 上利用 Amazon Aurora 免於昂貴的資料庫授權成本。

下列問題著重於成本最佳化方面的這些考量。

COST 5：您選取服務時如何評估成本？

Amazon EC2、Amazon EBS 和 Amazon S3 是基礎的 AWS 服務。Amazon RDS 和 Amazon DynamoDB 等受管服務為更高等級或應用程式等級的 AWS 服務。選取適當的基礎和受管服務，就可最佳化此工作負載的成本。舉例來說，您可以使用受管服務減少或省下大部分管理和營運開銷，讓您從事應用程式或企業相關活動。

COST 6：您選取資源類型、大小和數量時，如何達成成本目標？

確認您為手邊的任務選取適當的資源大小和數量。選取最具成本效益的類型、大小和數量，就能盡量減少浪費。

COST 7：您如何使用定價模式降低成本？

使用最適合您資源的定價模式，就能盡量減少支出。

COST 8：您如何規劃資料傳輸費？

確實規劃和監控資料傳輸費，以便做出盡量減少成本的架構決策。小規模而有效的架構變更能夠隨時間大幅減少營運成本。

透過在選擇服務時考慮成本因素，並以 Cost Explorer 和 AWS Trusted Advisor 等工具定期審查 AWS 的使用情形，您可積極監測使用率，並隨之調整部署。

管理需求與供應資源

待您移至雲端後，即可僅為所需付費。您可以在需要時供應資源以符合工作負載需求，減少因過度佈建付出高昂成本和造成浪費。您也可以使用調節、緩衝區或佇列來修改需求，以讓需求變得平緩，並以較少的資源來滿足需求，從而降低成本，或稍後使用批次服務來處理。

在 AWS 中，您可自動佈建資源以符合工作負載需求。Auto Scaling 使用基於需求或時間的方法，可讓您視需要新增和移除資源。若您能預期需求變更，則可省下更多成本，並驗證資源符合工作負載需求。您可以使用 Amazon API Gateway 實作限流，或使用 Amazon SQS 在工作負載中實作佇列。這兩者都可讓您修改工作負載元件的需求。

下列問題著重於成本最佳化方面的這些考量。

COST 9：如何管理需求和供應資源？

針對支出和效能達到平衡的工作負載，確認您購買的每一個項目都用到，並避免極少使用執行個體。往任一端傾斜的使用指標，對您組織在營運成本 (因過度使用而降低效能) 或浪費的 AWS 花費 (因過度佈建) 方面會造成負面影響。

在設計修改需求與供給資源時，請主動思考用量模式、佈建新資源所需的時間，以及需求模式的可預測性。管理需求時，確認您的佇列或緩衝區大小正確，而且在所需的時間內回應工作負載需求。

隨時間優化

隨著 AWS 推出新服務和功能，最佳實務是檢視現有架構決策，以確認其持續發揮最大成本效益。隨著您的要求變更，請主動將不再需要的資源、整項服務和系統加以除役。

透過實作新功能或資源類型可逐步最佳化工作負載，同時盡量減少實作變更所需的工作量。這可隨著時間持續提高效率，並讓您持續使用最新的技術來降低營運成本。您也可以使用新的服務來取代工作負載

中的元件，或將新元件新增至工作負載中。這可以大幅提高效率，因此定期審核工作負載並實作新服務和功能至關重要。

下列問題著重於成本最佳化方面的這些考量。

COST 10：您如何評估新服務？

隨著 AWS 推出新服務和功能，最佳實務是檢視現有架構決策，以確認其持續發揮最大成本效益。

在定期審查您的部署時，請評估較新的服務能如何為您節省成本。例如，Amazon RDS 上的 Amazon Aurora 能降低關聯式資料庫的成本。使用 Lambda 等無伺服器函數時，無需操作和管理執行個體來執行程式碼。

COST 11：如何評估工作的成本？

評估雲端維運的工作成本，審核耗時的雲端維運，並透過採用相關 AWS 服務、第三方產品或自訂工具將其自動化，以減少人力和成本。

資源

請參閱以下資源，進一步了解我們成本最佳化的最佳實務。

文件

- [AWS 文件](#)

白皮書

- [成本最佳化支柱](#)

永續性

永續性支柱的重點在於環境影響，尤其是能源消耗和效率，因為這方面是架構師採取直接行動來減少資源使用的重要手段。可以在[永續性支柱白皮書](#)中找到實作指引。

主題

- [設計原則](#)
- [定義](#)
- [最佳實務](#)
- [資源](#)

設計原則

在雲端中實現永續性有六個設計原則：

- **了解您的影響：**衡量您雲端工作負載的影響，並建立工作負載未來影響的模型。包含所有影響來源，包括客戶使用您產品所產生的影響，以及產品最後除役和淘汰所產生的影響。審視每個工作單元所需的資源和排放量，將生產輸出與雲端工作負載的總體影響進行比較。使用此資料來建立關鍵績效指標 (KPIs)、評估提高生產力同時降低影響的方法，以及估算提議變更隨著時間的影響。
- **建立永續性目標：**對於每個雲端工作負載，建立長期永續性目標，例如減少每項交易所需的運算和儲存資源。針對改進現有工作負載永續性的投資回報建立模型，並為業主提供必須投資於永續性目標所需的資源。針對成長著手規劃，以及建構您的工作負載，讓成長導致的每個使用者或每項交易 (根據適當的單位測量) 影響強度降低。這些目標有助於支持貴企業或組織更廣泛的永續性發展目標、識別迴歸，以及排定潛在改進領域的優先順序。
- **最大化使用率：**調整合適的工作負載規模並實作高效率設計，以確認高使用率，並將底層硬體的能源效率發揮到最大。由於每部主機會有基準耗電量，因此兩部以 30% 使用率執行的主機效率，會低於一部以 60% 執行的主機效率。在此同時減少或最小化閒置資源、處理和儲存，以減少為工作負載供電所需的能源總量。
- **預測並採用新的、更有效率的軟硬體產品：**支援合作夥伴和供應商進行上游改進，以協助您減少雲端工作負載的影響。持續監控和評估更有效率的新軟硬體產品。針對靈活性進行設計，以允許快速採用高效率的新技術。
- **使用受管服務：**在廣泛的客戶群中共用服務，有助於最大化資源利用，進而減少支援雲端工作負載所需的基礎設施數量。例如，客戶可以將工作負載遷移至 AWS 雲端 並採用受管服務，例如 AWS Fargate for serverless Container，其中 會大規模 AWS 運作，並負責其高效操作，以共享電力和聯網等常見資料中心元件的影響。使用 受管服務，有助於將影響降至最低，例如使用 Amazon S3 生命週期組態或 Amazon EC2 Auto Scaling 自動將不常存取的資料移至冷儲存，以調整容量以符合需求。
- **減少雲端工作負載的下游影響：**減少使用您的服務所需的能源或資源量。減少客戶為了使用您的服務而升級裝置的需求。使用 Device Farm 進行測試以了解預期影響，並與客戶進行測試以了解使用您服務的實際影響。

定義

雲端永續性有六個最佳實務領域：

- 區域選擇
- 因應需求
- 軟體和架構
- 資料
- 硬體和服務
- 程序和文化

雲端中的永續性是一項近乎持續的工作，主要專注於透過從佈建的資源中實現最大效益並最大限度地減少所需的總資源，來減少工作負載所有元件的節能和效率。此工作的範圍包括最初選擇高效的程式設計語言、採用現代演算法、使用高效的資料儲存技術、部署至正確大小和高效的運算基礎設施，以及最大限度地減少對高效能最終使用者硬體的要求。

最佳實務

主題

- [區域選擇](#)
- [因應需求](#)
- [軟體和架構](#)
- [資料管理](#)
- [硬體和服務](#)
- [程序和文化](#)

區域選擇

工作負載的 區域選擇會大幅影響其 KPIs，包括效能、成本和碳足跡。若要改善這些 KPIs，您應該根據業務需求和永續性目標，為工作負載選擇區域。

下列問題著重於這些永續性方面的考量。(如需永續性問題清單和最佳實務，請參閱[附錄](#)。)

SUS 1：如何為工作負載選取區域？

工作負載的區域選擇會大幅影響其 KPIs，包括效能、成本和碳足跡。若要改善這些 KPIs，您應該根據業務需求和永續性目標，為工作負載選擇區域。

因應需求

使用者和應用程式使用工作負載和其他資源的方式，可協助您找到改善的機會，以達成永續性目標。擴展基礎架構以持續符合需求，並確認您僅使用支援使用者所需的最低資源。讓服務層級符合客戶需求。妥善放置資源，以限制使用者和應用程式使用資源所需的網路。移除未使用的資產。為團隊成員提供滿足其需求的裝置，同時將對永續性的影響降至最低。

下列問題著重於永續性方面的考量：

SUS 2：如何根據需求調整雲端資源？

使用者和應用程式使用工作負載和其他資源的方式，可協助您找到改善的機會，以達成永續性目標。擴展基礎架構以持續符合需求，並確認您僅使用支援使用者所需的最低資源。讓服務層級符合客戶需求。妥善放置資源，以限制使用者和應用程式使用資源所需的網路。移除未使用的資產。為團隊成員提供滿足其需求的裝置，同時將對永續性的影響降至最低。

隨使用者負載擴展基礎設施：識別使用率低或無使用率的時期，並調整資源規模以減少過剩容量、提高效率。

SLAs 與永續性目標保持一致：定義和更新服務層級協議（SLAs），例如可用性或資料保留期，以盡可能減少支援工作負載所需的資源數量，同時繼續滿足業務需求。

停止建立和維護未使用的資產：分析應用程式資產（例如預先編譯的報告、資料集和靜態影像）和資產存取模式，識別冗餘、未充分利用和可以除役的目標。合併具有冗餘內容的產生資產（例如，具有重疊或通用資料集與輸出的每月報告），以減少重複輸出時消耗資源。將未使用的資產除役（例如不再販售產品的影像）以釋放消耗的資源，並減少用於支援工作負載的資源數量。

針對使用者位置最佳化工作負載的地理位置：分析網路存取模式，識別客戶的地理連接位置。選取可減少網路流量傳輸距離的區域和服務，以減少支援工作負載所需的總網路資源。

為執行的活動最佳化團隊成員資源：最佳化提供給團隊成員的資源，以盡量減少對永續性的影響，同時支援他們的需求。例如，在使用率高的共用雲端桌面上執行複雜的操作（例如渲染和編譯），而不是在使用率低的高功率單一使用者系統上執行。

軟體和架構

實施可執行負載順暢並保持已部署資源一致高使用率的模式，將資源消耗降至最低。由於使用者行為隨時間改變，元件可能會因缺乏使用而閒置。修改模式和架構來整合未充分利用的元件，提高整體使用率。淘汰不再需要的元件。了解工作負載元件的效能，並最佳化消耗最多資源的元件。留意客戶用來存取服務的裝置，並實施盡量減少裝置升級需求的模式。

下列問題著重於這些永續性方面的考量：

SUS 3：如何利用軟體和架構模式來支援永續性目標？

實施可執行負載順暢並保持已部署資源一致高使用率的模式，將資源消耗降至最低。由於使用者行為隨時間改變，元件可能會因缺乏使用而閒置。修改模式和架構來整合未充分利用的元件，提高整體使用率。淘汰不再需要的元件。了解工作負載元件的效能，並最佳化消耗最多資源的元件。留意客戶用來存取服務的裝置，並實施盡量減少裝置升級需求的模式。

最佳化非同步與排程任務的軟體和架構：使用高效率的軟體設計和架構，將每個工作單元所需的平均資源降至最低。實作可平均利用元件的機制，減少任務之間的閒置資源，並將負載尖峰的影響降至最低。

移除或重構使用量低或完全未使用的工作負載元件：監控工作負載活動，識別各元件使用率隨時間的變化。移除未使用且不再需要的元件，並重構使用率低的元件，減少資源浪費。

最佳化程式碼中耗用最多時間或資源的區域：監控工作負載活動，識別消耗最多資源的應用程式元件。最佳化這些元件中執行的程式碼，將資源使用量降至最低，同時將效能發揮至最大。

最佳化對客戶裝置和設備的影響：了解客戶用來使用您服務的裝置和設備、其預期生命週期，以及更換這些元件對財務和永續性的影響。實作軟體模式和架構，將客戶更換裝置和升級設備的需求降至最低。例如，實作使用與較早硬體和作業系統版本向後相容的程式碼的新功能，或管理承載的大小，不讓其超過目標裝置的儲存容量。

使用最有效支援資料存取和儲存模式的軟體模式和架構：了解資料在工作負載中的使用方式、使用者的使用方式、傳輸方式以及儲存方式。選取可將資料處理和儲存要求降至最低的技術。

資料管理

下列問題著重於這些永續性方面的考量：

SUS 4：如何利用資料管理政策和模式來支援永續性目標？

實作資料管理實務來減少支援工作負載所需的佈建儲存，以及減少為了使用它所需的資源。了解您的資料，並使用最有效支援資料業務價值及其使用方式的儲存技術和組態。當需求減少時，將資料循環到效率較高、效能較低的儲存，並刪除不再需要的資料。

實作資料分類政策：將資料分類，以了解其對業務成果的重要性。使用此資訊來確定何時可將資料移動到更節能的儲存，或是可以安全刪除它。

使用支援資料存取和儲存模式的技術：使用最有效支援您的資料存取和儲存方式的儲存技術，以在支援工作負載的同時，也將佈建的資源降至最低。例如，固態裝置（SSDs）比磁性磁碟機更耗能，且應僅用於作用中的資料使用案例。針對不常存取的資料，使用節能的存檔類別儲存。

使用生命週期政策來刪除不需要的資料：管理所有資料的生命週期並自動執行刪除時間表，將工作負載的總儲存需求降至最低。

將區塊儲存中的過度佈建降至最低：若要最小化總佈建儲存，請建立具有適合工作負載之大小分配的區塊儲存。使用彈性磁碟區，隨著資料成長擴展儲存，無需調整連接到運算資源的儲存大小。定期審查彈性磁碟區，並縮減過度佈建的磁碟區以符合目前的資料大小。

移除不需要或多餘的資料：必要時才複製資料，將消耗的總儲存空間降至最低。使用在檔案和區塊層級刪除重複資料的備份技術。限制使用獨立磁碟機的備援陣列（RAID）組態，除非需要滿足 SLAs。

使用共用檔案系統或物件儲存體存取通用資料：採用共用儲存和單一真實來源，避免資料重複並降低工作負載的總儲存需求。僅在需要時從共用儲存體擷取資料。分離未使用的磁碟區以釋放資源。最小化跨網路的資料移動：使用共用儲存，並從區域資料存放區存取資料，將支援工作負載資料移動所需的總聯網資源降至最低。

僅在難以重新建立時才備份資料：為了將儲存消耗降至最低，僅備份具有業務價值或需要滿足合規要求的資料。檢查備份政策，並在復原案例中排除沒有價值的暫時性儲存。

硬體和服務

透過變更硬體管理實務，尋求降低工作負載永續性影響的機會。將佈建和部署所需的硬體量降至最低，並為個別工作負載選取最高效率的硬體和服務。

下列問題著重於這些永續性方面的考量：

SUS 5：如何在架構中選擇並使用雲端硬體和服務來支援永續性目標？

透過變更硬體管理實務，尋求降低工作負載永續性影響的機會。將佈建和部署所需的硬體量降至最低，並為個別工作負載選取最高效率的硬體和服務。

使用最低數量的硬體來滿足需求：使用雲端功能，您可以頻繁變更工作負載實作。隨著需求變更，更新已部署的元件。

使用影響最小的執行個體類型：持續關注新執行個體類型的發佈，並運用能源效率改進，包括旨在支援特定工作負載 (例如機器學習訓練和推論以及影片轉碼) 的執行個體類型。

使用 受管服務：受管服務會將維持部署硬體高平均使用率和永續性最佳化的責任轉移到 AWS。使用受管服務，將服務的永續性影響分散給服務的所有租用戶，降低您的個人佔比。

最佳化您對 的使用GPUs：圖形處理單元 (GPUs) 可以是高耗電的來源，而且許多GPU工作負載具有高度可變性，例如轉譯、轉碼和機器學習訓練和建模。只在所需的時間內執行GPUs執行個體，並在不需將耗用的資源降至最低時，透過自動化停用執行個體。

程序和文化

透過變更開發、測試和部署實務來尋找降低永續性影響的機會。

下列問題著重於這些永續性方面的考量：

SUS 6：您的組織程序如何支援您的永續性目標？

透過變更開發、測試和部署實務來尋找降低永續性影響的機會。

採用可快速導入永續性改進的操作：在將潛在改善部署到生產環境之前，先對其進行測試和驗證。在計算改善所帶來的未來潛在利益時，應考慮測試成本。開發低成本測試操作，以推動小改進的交付。

讓您的工作負載保持最新狀態：Up-to-date作業系統、程式庫和應用程式可以提高工作負載效率，並建立更有效率的技術採用。Up-to-date 軟體也可能包含功能，以更精確地測量工作負載的永續性影響，因為廠商提供功能來滿足自己的永續性目標。

提高建置環境的使用率：使用自動化和基礎設施即程式碼，在需要時啟動生產前環境，並在不使用時將其關閉。常見的模式是排程可用性時間，使之與開發團隊成員的工作時間一致。休眠是一種有用的工

具，可保留狀態，並在需要時快速讓執行個體上線。使用具有高載容量的執行個體類型、Spot 執行個體、彈性資料庫服務、容器和其他技術，以根據使用量調整開發和測試容量。

使用受管 Device Farm 進行測試：受管 Device Farm 可將硬體製造和資源使用的永續性影響分散給多個租用戶。受管 Device Farm 提供多種裝置類型，因此您可以支援較早且較不熱門的硬體，並避免不必要的裝置升級對客戶的永續性造成影響。

資源

請參閱下列資源，進一步了解我們的永續性最佳實務。

白皮書

- [永續性支柱](#)

影片

- [氣候承諾](#)

審查程序

架構審查的執行方式必須一致，採行鼓勵深入探索的無譴責作法。應為輕量程序 (數小時而非數日)，屬於一種對話而非稽核。就架構進行審查的目的是識別可能需要解決的重要問題，或是有改進空間之處。審查的結果是一套行動，應能提升客戶使用工作負載得到的體驗。

如同「論架構」一節所討論，建議由各團隊成員對其架構的品質負起責任。我們建議建置架構的團隊成員使用 Well-Architected 架構以持續審查其架構，而非舉行正式審查會議。採取近乎持續作法可讓您的團隊成員隨著架構演進更新答案，並隨著您遞送功能而提升架構。

AWS Well-Architected 架構與內部審核系統和服務的方式 AWS 一致。它以影響架構方法的一組設計原則為基礎，以及驗證人員不會忽略根本原因分析 (RCA) 中經常出現的區域的問題。每當內部系統、AWS 服務或客戶發生重大問題時，我們會查看 RCA，看看我們是否可以改善我們使用的審核程序。

審查應在產品生命週期的重要里程碑，並於設計階段早期實施，以免成為單向門戶難以變更，而且需趕在正式運作日期之前。(許多決定為可逆的雙向門戶。這些決定可採用輕量程序。單向門戶難以、甚至無法逆轉，實施之前需要更多檢查工作。) 進入生產環境之後，您的工作負載可隨著新增功能和變更技術實作而繼續演進。工作負載的架構會隨時間而變化。您必須遵守良好的衛生實務，以阻止您推動演進的同時，其架構上的特性隨之衰退。在您作出重要的架構變更時，應遵照一套衛生程序，包括 Well-Architected 審查。

如果您想以審查作為一次性的快照或獨立測量，建議確認在對話中包含所有適當人員。我們經常發現，到審查時團隊才初次真正了解實作了些什麼。審查另一個團隊的工作負載時，一種效果良好的方式就是就其架構進行一連串非正式對話，能探詢出大多數問題的答案。接著您即可透過一兩次會議進行追蹤，釐清或深入探索模稜兩可或看出有風險的領域。

開會時的一些建議項目如下：

- 有白板的會議室
- 任何圖或設計備註的列印紙本
- 需要 out-of-band 研究才能回答的問題動作清單 (例如，「我們是否啟動加密？」)

在您完成審查之後，應列有問題清單，可根據業務環境排列優先順序。您還需要考慮這些問題對您團隊 day-to-day 工作的影響。如果您及早解決這些問題，即可空出時間創造商業價值，不必忙於解決重複發生的問題。當您解決問題時，可以更新審查，了解架構改良的情形。

雖然審查完成後，其價值所在自然明朗，但您可能會發現新的團隊起初可能會有所抗拒。經由對團隊教育審查的益處，可解決下列幾項反對說法：

- 「我們太忙了！」（團隊預備進行盛大推出時，往往會這麼說。）
 - 既然預備進行盛大推出，一定希望過程能夠順利。審查可讓您了解可能漏掉的任何問題。
 - 建議您在產品生命週期之中及早實施審查，以發現風險並開發配合功能遞送藍圖的減緩計畫。
- 「我們沒有時間處理結果！」（往往在作為目標的活動無法挪動，例如超級盃時會這麼說。）
 - 這些活動無法挪動。您是否真的想在對於架構所具風險不知情的情況下迎接活動？就算無法解決所有的問題，仍然可在發生狀況時握有處理問題的程序手冊。
- 「我們不想讓解決方案實作的秘密外流！」
 - 如果您向團隊指出 Well-Architected 架構中的疑問，他們就能看出這些疑問完全不會顯露商業或技術上的專屬資訊。

在您與組織內的團隊實施多重審查之時，可能會識別主題上的問題。例如，可能會發現一群團隊的問題集中在特定支柱或主題上。建議以全面方式審視所有的審查，並識別有助於解決這些主題問題的任何機制、培訓或首席工程設計對談。

結論

AWS Well-Architected Framework 提供跨六個支柱的架構最佳實務，用於設計和操作雲端中可靠、安全、高效、經濟實惠且永續的系統。該架構提供一套問題，允許您審查現有或提議的架構。它也為每個支柱提供一組 AWS 最佳實務。在您的架構中使用該架構可協助您產生穩定且有效率的系統，讓您能夠專注於功能需求。

貢獻者

下列個人和組織為本文件作出了貢獻：

- Brian Carlson : Amazon Web Services Well-Architected 營運主管
- Ben Potter , Amazon Web Services Well-Architected 安全主管
- Seth Eliot , Amazon Web Services Well-Architected 可靠性主管
- Eric Pullen , Amazon Web Services 資深解決方案架構師
- Rodney Lester , Amazon Web Services 首席解決方案架構師
- Jon Steele , Amazon Web Services 資深技術客戶經理
- Max Ramsay , Amazon Web Services 首席安全解決方案架構師
- Callum Hughes , Amazon Web Services 解決方案架構師
- Ben Mergen , Amazon Web Services 資深成本主管與解決方案架構師
- Chris Kozlowski , Amazon Web Services 企業支援資深專家與技術客戶經理
- Alex Livingstone , Amazon Web Services 雲端操作首席專家與解決方案架構師
- Paul Moran , Amazon Web Services 企業支援首席技術專家
- Peter Mullen , Amazon Web Services 專業服務諮詢顧問
- Chris Pates , Amazon Web Services 企業支援資深專家與技術客戶經理
- Arvind Raghunathan , Amazon Web Services 企業支援首席專家與技術客戶經理
- Sam Mokhtari , Amazon Web Services 資深效率主管與解決方案架構師

深入閱讀

[AWS 架構中心](#)

[AWS 雲端合規](#)

[AWS Well-Architected Partner 計畫](#)

[AWS Well-Architected Tool](#)

[AWS Well-Architected 首頁](#)

[卓越營運支柱白皮書](#)

[安全支柱白皮書](#)

[可靠性支柱白皮書](#)

[效能達成效率支柱白皮書](#)

[成本最佳化支柱白皮書](#)

[永續性支柱白皮書](#)

[Amazon 建置者資料中心](#)

文件修訂

若要收到此白皮書更新的通知，請訂閱 RSS 摘要。

變更	描述	日期
主要更新	最佳實務已更新，在可靠性、安全性、卓越營運、永續性和效能效率方面加入了新指引。可靠性支柱經過大幅度的整理，並更新了許多最佳實務。安全性和卓越營運方面的指引已更新，內容經細部調整後加入了新的服務與生成式 AI 建議。永續性方面根據 AWS 服務和新的最佳實務進行了多項更新。	2024 年 11 月 6 日
主要更新	整個支柱進行了大規模最佳實務更新。安全性和成本都獲得了新的最佳實務。	2024 年 6 月 27 日
主要更新	主要支柱更新。	2023 年 10 月 3 日
主要更新	最佳實務已更新，納入了規範性指引，並增加了新的最佳實務。安全性和成本最佳化支柱加入了新問題。	2023 年 4 月 10 日
次要更新	已在附錄中新增工作量的定義和更新最佳實務。	2022 年 10 月 20 日
主要更新	已新增永續性支柱和更新了連結。	2021 年 12 月 2 日
主要更新	永續性支柱已新增到框架中。	2021 年 11 月 20 日
次要更新	已移除非包容性語言。	2021 年 4 月 22 日

次要更新	已修正數個連結。	2021 年 3 月 10 日
次要更新	整體的小幅度編輯變更。	2020 年 7 月 15 日
主要更新	檢閱和重寫大多數問題和答案。	2020 年 7 月 8 日
白皮書已更新	新增 AWS Well-Architected Tool，連結至 AWS Well-Architected 實驗室及 AWS Well-Architected 合作夥伴、小處修復以促成架構有多種語言版本。	2019 年 7 月 1 日
白皮書已更新	審查並重新撰寫大多數的問題和答案，以確保問題一次聚焦在一個主題之上。這使得部分先前的問題分為數個問題。新增定義的共同詞彙 (工作負載、元件等)。變更主要本文中的問題呈現，以含入描述性文字。	2018 年 11 月 1 日
白皮書已更新	更新以簡化問題文字，將答案標準化，並提升可讀性。	2018 年 6 月 1 日
白皮書已更新	卓越營運移至支柱前端並重新撰寫，使其成為其他支柱的框架。重新整理其他支柱，以反映 AWS 的演進。	2017 年 11 月 1 日
白皮書已更新	更新架構以含入卓越營運支柱，並修訂及更新其他支柱以減少重複，並納入與數千客戶一同執行審查之所學。	2016 年 11 月 1 日
次要更新	以目前的 Amazon CloudWatch Logs 資訊更新附錄。	2015 年 11 月 1 日

[初次出版](#)

已發布 AWS Well-Architected Framework。 2015 年 10 月 1 日

 Note

若要訂閱 RSS 更新，必須為所使用的瀏覽器啟用 RSS 外掛程式。

框架版本：

- [2024-06-27](#)
- [2023-10-03](#)
- [2023-04-10](#)
- [2022-03-31](#)

附錄：問題與最佳實務

本附錄總結了所有與 AWS Well-Architected Framework 相關的問題與最佳實務。

支柱

- [操作效能](#)
- [安全](#)
- [可靠性](#)
- [效能效率](#)
- [成本最佳化](#)
- [永續性](#)

操作效能

卓越營運 (OE) 是一項承諾，代表致力於妥善設計軟體，同時持續提供絕佳的客戶體驗。卓越營運支柱包括組織團隊、設計工作負載、大規模運作工作負載及促使其進化的最佳實務。您可以在[卓越營運支柱白皮書](#)中找到實作指引。

最佳實務領域

- [組織](#)
- [準備](#)
- [操作](#)
- [演進](#)

組織

問題

- [OPS 1. 如何決定您的優先事項？](#)
- [OPS 2. 如何建構組織以支援業務成果？](#)
- [OPS 3. 您的組織文化如何支援您的業務成果？](#)

OPS 1. 如何決定您的優先事項？

每個人都應了解自己在實現商業價值過程中的角色。擁有共同目標以設定資源優先順序。這會充分發揮您所做努力的優勢。

最佳實務

- [OPS01-BP01 評估外部客戶需求](#)
- [OPS01-BP02 評估內部客戶需求](#)
- [OPS01-BP03 評估治理要求](#)
- [OPS01-BP04 評估合規要求](#)
- [OPS01-BP05 評估威脅態勢](#)
- [OPS01-BP06 評估權衡的同時管理利益和風險](#)

OPS01-BP01 評估外部客戶需求

讓關鍵利害關係人 (包括業務、開發和營運團隊) 參與進來，以確定將工作重點放在哪些外部客戶需求上。這驗證了您對實現預期業務成果所需的營運支援有透徹的了解。

預期成果：

- 您可以透過客戶成果逆向工作。
- 您了解營運實務如何支援業務成果和目標。
- 您與所有相關各方接觸。
- 您擁有捕捉外部客戶需求的機制。

常見的反模式：

- 您已決定不在核心上班時間以外的時間提供客戶支援，但尚未檢閱歷史支援請求資料。您不知道這是否會對您的客戶產生影響。
- 您正在開發新功能，但尚未與客戶互動，以了解是否需要該功能，若需要又應以何種形式提供，而且未進行試驗以驗證交付的需求和方法。

建立此最佳實務的優勢：需求得到滿足的客戶更有可能繼續成為客戶。評估和了解外部客戶的需求，將讓您了解如何安排工作的優先順序來實現商業價值。

未建立此最佳實務時的曝險等級：高

實作指引

了解業務需求：只有業務、開發及營運團隊等利害關係人擁有共同的目標並達成共識，方能讓業務取得成功。

審查外部客戶的業務目標、需求和優先事項：與關鍵利害關係人 (包括業務、開發和營運團隊) 進行互動，以討論外部客戶的目標、需求和優先事項。這將確保您對實現業務和客戶成果所需的營運支援有透徹的了解。

建立共識：在以下方面建立共識：工作負載的業務功能、每個團隊在工作負載營運過程中的角色，以及這些因素如何支援內外部客戶的共同業務目標。

資源

相關的最佳實務：

- [OPS11-BP03 實作回饋迴圈](#)

OPS01-BP02 評估內部客戶需求

讓關鍵利益相關者 (包括業務、開發和營運團隊) 參與進來，以確定將工作重點放在哪些內部客戶需求上。這將確保您對實現業務成果所需的營運支援有透徹的了解。

預期成果：

- 利用您制定的優先事項，聚焦於改善作業，因為它們能發揮最大的影響力 (例如，發展團隊技能、改善工作負載效能、降低成本、自動化執行手冊或提升監控力)。
- 根據需求變更更新您的優先順序。

常見的反模式：

- 您已決定在不向產品團隊諮詢的情況下，變更他們的 IP 位址配置，以便更輕鬆地管理網路。您不知道這會對您的產品團隊產生什麼影響。
- 您正在實作新的開發工具，但尚未讓內部客戶了解是否需要它，或者它是否與他們現有的實務相容。
- 您正在實作新的監控系統，但尚未聯絡內部客戶，以了解他們是否有應該考慮的監控或報告需求。

建立此最佳實務的優勢：評估和了解內部客戶的需求，可讓您了解如何安排工作的優先順序來實現商業價值。

未建立此最佳實務時的曝險等級：高

實作指引

- 了解業務需求：只有業務、開發及營運團隊等利益相關者擁有共同的目標並達成共識，方能實現業務成功。
- 審查內部客戶的業務目標、需求和優先事項：與關鍵利益相關者 (包括業務、開發和營運團隊) 進行互動，以討論內部客戶的目標、需求和優先事項。這將確保您對實現業務和客戶成果所需的營運支援有透徹的了解。
- 建立共識：在以下方面建立共識，即工作負載的業務功能、每個團隊在工作負載營運過程中的角色，以及這些因素如何支援內外部客戶的共同業務目標。

資源

相關的最佳實務：

- [OPS11-BP03 實作意見回饋循環](#)

OPS01-BP03 評估治理要求

管控是政策、規則或架構的集合，供公司用來達成其業務目標。管控要求產生自您的組織內部。這些要求可能會影響到您所選擇的技術類型，或是您操作工作負載的方式。將組織管控要求納入您的工作負載中。合規是指展現您已實作管控要求的能力。

預期成果：

- 管控要求會併入工作負載的架構設計和操作中。
- 您可以提供您已遵循管控要求的證明。
- 定期審查並更新管控要求。

常見的反模式：

- 您的組織規定根帳戶需進行多重要素驗證。您未能實行此要求，根帳戶遭到損害。
- 在設計工作負載期間，您選擇了未經 IT 部門核准的執行個體類型。您無法啟動工作負載，而必須執行重新設計。
- 您必須有災難復原計畫。您未建立該計畫，且工作負載遭逢長時間的中斷。
- 您的團隊想要使用新的執行個體，但您的管理要求尚未更新予以允許。

建立此最佳實務的優勢：

- 遵循管控要求，可讓您的工作負載符合較大組織的政策。
- 管控要求會反映組織的產業標準和最佳實務。

未建立此最佳實務時的曝險等級：高

實作指引

與利益相關者和管控組織共同識別管控要求。將管控要求納入您的工作負載中。能夠證明您已遵循管控要求。

客戶範例

在 AnyCompany 零售，雲端營運團隊會與整個組織的利益相關者合作，以開發治理要求。例如，它們禁止SSH存取 Amazon EC2執行個體。如果團隊需進行系統存取，他們必須使用 AWS Systems Manager Session Manager。雲端營運團隊會在新服務推出時定期更新管控要求。

實作步驟

1. 識別工作負載的利益相關者，包括任何集中團隊。
2. 與利益相關者共同識別管控要求。
3. 產生清單後，請排定改善項目的優先順序，並開始在您的工作負載中加以實作。
 - a. 使用 等服務[AWS Config](#)來建立 governance-as-code和驗證遵循了治理要求。
 - b. 如果您使用 [AWS Organizations](#)，則可以利用服務控制政策來實作管控要求。
4. 提供驗證實作情形的文件。

實作計劃的工作量：中。實作遺漏的管控要求可能會導致工作負載重新作業。

資源

相關的最佳實務：

- [OPS01-BP04 評估合規要求](#) - 合規性類似於管控，但來自組織外部。

相關文件：

- [AWS 管理和治理雲端環境指南](#)

- [多帳戶環境中 AWS Organizations 服務控制政策的最佳實務](#)
- [中的治理 AWS 雲端：敏捷性與安全之間的正確平衡](#)
- [什麼是治理、風險與合規（GRC）？](#)

相關影片：

- [AWS 管理和治理：組態、合規和稽核 - AWS 線上技術講座](#)
- [AWS re：Inforce 2019：雲端時代的治理（DEM12-R1）](#)
- [AWS re：Invent 2020：使用實現合規作為程式碼 AWS Config](#)
- [AWS re：Invent 2020：靈活管理 AWS GovCloud \(US\)](#)

相關範例：

- [AWS Config Conformance Pack 範例](#)

相關服務：

- [AWS Config](#)
- [AWS Organizations - 服務控制政策](#)

OPS01-BP04 評估合規要求

法規、產業和內部合規要求是定義組織優先順序的重要因子。您的合規架構可能會禁止使用特定技術或地理位置。若未識別出外部合規架構，請運用盡職調查。產生驗證合規性的稽核或報告。

如果聲明您的產品符合特定的合規標準，您必須有內部程序來確保持續的合規性。合規標準的例子包括 PCI DSS、FedRAMP 和 HIPAA。適用的合規標準取決於各種因素，例如解決方案存放或傳輸的資料類型，以及解決方案支援的地理區域。

預期成果：

- 將法規、產業和內部合規要求併入架構選擇中。
- 您可以驗證合規性並產生稽核報告。

常見的反模式：

- 您的工作負載有部分屬於支付卡產業資料安全標準 (PCI-DSS) 架構下，但您的工作負載儲存信用卡資料時並未予以加密。
- 您的軟體開發人員和架構師不知道您的組織必須遵循的合規架構。
- 年度 Systems and Organizations Control (SOC2) Type II 稽核即將到來，但您無法驗證已設置控制。

建立此最佳實務的優勢：

- 評估和了解套用到工作負載的合規要求，可讓您了解如何安排工作的優先順序來實現商業價值。
- 您可以選擇與合規架構相符的適當位置和技术。
- 針對可稽核性設計工作負載，有助於證明您確實遵循合規架構。

未建立此最佳實務時的曝險等級：高

實作指引

若實作此最佳實務，即表示您會在架構設計程序中併入合規要求。您的團隊成員將得知必要的合規架構。您會驗證合規性符合架構。

客戶範例

AnyCompany Retail 儲存客戶的信用卡資訊。卡片儲存團隊的開發人員了解他們必須遵從 PCI-DSS 架構。他們執行了相關步驟，驗證信用卡資訊以安全方式儲存和存取，並遵從 PCI-DSS 架構。他們每年都會與安全團隊共同驗證合規性。

實作步驟

1. 與安全和管控團隊合作，確認您的工作負載必須遵循哪些產業、法規或內部合規架構。在您的工作負載中併入合規架構。
 - a. 使用諸如 [AWS Compute Optimizer](#) 和 [AWS Security Hub CSPM](#) 之類的服務驗證 AWS 資源的持續合規性。
2. 讓團隊成員了解合規要求，使其能據以操作及設計工作負載。合規要求應包含在架構和技术選擇中。
3. 根據合規架構，您可能必須產生稽核或合規報告。請與組織合作，盡可能將此程序自動化。
 - a. 使用諸如 [AWS Audit Manager](#) 等服務來驗證合規性並產生稽核報告。
 - b. 可以使用 [AWS Artifact](#) 下載 AWS 安全性和合規性文件。

實作計劃的工作量：中。實作合規架構可能並不容易。產生稽核報告或合規文件，會增添額外的複雜性。

資源

相關的最佳實務：

- [SEC01-BP03 識別並驗證控制目標](#) - 安全控制目標是整體合規性的重要組成部分。
- [SEC01-BP06 自動化管道中安全控制的測試和驗證](#) - 作為管道的一部分，驗證安全控制。您也可以產生新變更的合規文件。
- [SEC07-BP02 定義資料保護控制](#) - 許多合規性框架都基於資料處理和儲存政策。
- [SEC10-BP03 準備鑑識功能](#) - 鑑識功能有時可用於稽核合規性。

相關文件：

- [AWS 合規中心](#)
- [AWS 合規資源](#)
- [AWS 風險與合規白皮書](#)
- [AWS 共同責任模型](#)
- [合規計劃範圍內的 AWS 服務](#)

相關影片：

- [AWS re:Invent 2020：使用 AWS Compute Optimizer 實現合規即程式碼](#)
- [AWS re:Invent 2021 - 雲端合規性、保證和稽核](#)
- [AWS Summit ATL 2022 - 在 AWS 上實作合規性、保證和稽核 \(COP202\)](#)

相關範例：

- [AWS 上的 PCI DSS 和 AWS 基礎安全最佳實務](#)

相關服務：

- [AWS Artifact](#)
- [AWS Audit Manager](#)

- [AWS Compute Optimizer](#)
- [AWS Security Hub CSPM](#)

OPS01-BP05 評估威脅態勢

評估對業務的威脅 (例如，競爭、業務風險和負債、營運風險和資訊安全威脅)，並將最新的資訊保存在風險登記表內。決定工作重點的領域時，加入風險影響。

[Well-Architected Framework](#) 強調學習、衡量和改善。它提供一致的方法來評估架構，並實作會隨著時間擴展的設計。AWS 提供 [AWS Well-Architected Tool](#)，協助您在開發之前檢閱方法、生產前的工作負載狀態，以及生產中工作負載的狀態。您可以將它們與最新的 AWS 架構最佳實務進行比較，監控工作負載的整體狀態，並深入了解潛在風險。

AWS 客戶有資格接受其任務關鍵工作負載的引導式 Well-Architected Review，以根據 AWS 最佳實務[測量其架構](#)。企業支援客戶有資格獲得[營運審查](#)，該審查旨在助其識別在雲端營運的方法中的差距。

這些審查的跨團隊參與有助於建立對您的工作負載以及團隊角色可如何助力成功的共識。透過審查識別的需求可以助您確定優先順序。

[AWS Trusted Advisor](#) 是一款可存取核心檢查集的工具，這些檢查提出了優化建議，可能有助您確定優先事項。[商業和企業支援客戶](#)可存取針對安全性、可靠性、效能和成本優化的其他檢查，從而進一步協助確定他們的優先事項。

預期成果：

- 您可以定期檢閱 Well-Architected 和 Trusted Advisor 輸出並採取行動
- 您已知道服務的最新修補程式狀態
- 您了解已知威脅的風險和影響，並採取相應的行動
- 視需要實作緩和措施
- 對行動和背景進行溝通

常見的反模式：

- 您在產品中使用舊版的軟體程式庫。您不知道程式庫的安全性更新是否會對工作負載產生意外影響。
- 競爭對手剛剛發布了其產品的一個版本，它可以解決許多客戶對您產品的投訴。您並未優先處理這些已知問題。
- 監管機構一直在追查像你們這樣不符合法律監管合規要求的公司。您尚未排定處理任何未解決合規要求之事項的優先順序。

建立此最佳實務的優勢：識別並了解組織和工作負載所面臨的威脅，有助於您判斷要解決哪些威脅、它們的優先順序以及執行此作業所需的資源。

未建立此最佳實務時的曝險等級：中

實作指引

- 評估威脅態勢：評估對業務的威脅 (例如，競爭、業務風險和負債、營運風險和資訊安全威脅)，以便您可以在決定工作重點時考量其影響。
 - [AWS 最新安全公告](#)
 - [AWS Trusted Advisor](#)
- 維護威脅模型：建立和維護用於識別潛在威脅、已規劃和就地緩解措施及其優先順序的威脅模型。審查顯示為事件的威脅的機率、從這些事件中復原的成本、導致的預期傷害，以及防止這些事件的成本。當威脅模型的內容變更時，修改優先順序。

資源

相關的最佳實務：

- [SEC01-BP07 使用威脅模型識別威脅並排定緩解優先順序](#)

相關文件：

- [AWS 雲端 合規](#)
- [AWS 最新安全公告](#)
- [AWS Trusted Advisor](#)

相關影片：

- [AWS re:Inforce 2023 - 協助改善威脅模型的工具](#)

OPS01-BP06 評估權衡的同時管理利益和風險

來自多方的競爭利益可能會使安排工作的優先順序、建立能力、交付與業務策略一致的成果變得具有挑戰性。例如，您可能會被要求加快新功能的上市速度，而不是最佳化 IT 基礎架構成本。這會使利益雙方發生衝突。在這種情況下，需要將決策提交給更高的權利層來解決衝突。需要使用資料來消除決策過程中的情感依戀。

同樣的挑戰可能會發生在戰術層面。例如，選擇使用關聯式或非關聯式資料庫技術，可能會對應用程式的操作產生重大影響。了解各種選擇的可預測結果至關重要。

AWS 可以協助您教育您的團隊有關 AWS 及其服務的知識，從而增進他們對自己的選擇會如何影響工作負載的了解。使用 [支援](#) ([AWS 知識中心](#)、[AWS 論壇](#) 和 [支援中心](#)) 和 [AWS 文件](#) 提供的資源來教育您的團隊。如有其他問題，請聯絡支援。

AWS 還可以在 [Amazon 建置者資料中心](#) 中分享操作最佳實務和模式。您可透過 [AWS 部落格](#) 和 [官方 AWS 播客](#) 獲得其他各種實用資訊。

預期成果：您擁有明確定義的決策管控框架，可以促進雲端交付組織內每個層級的重要決策。該框架包括很多功能，例如風險登記表、有權做出決策的已定義角色以及可以做出的每個決策級別的定義模型。此框架預先定義了如何解決衝突，需要呈現哪些資料，以及如何確定選項的優先級，以便一旦做出決定，就可以立即提交。決策框架包括一個標準化方法，可審查和衡量每個決策的利益和風險，以了解權衡。這可能包括外部因素，例如遵守法規合規要求。

常見的反模式：

- 您的投資者要求您證明支付卡產業資料安全標準 (PCI DSS) 的合規性。您沒有考量滿足要求和繼續您目前開發工作之間的權衡取捨。相反地，您繼續開發工作，而不證明合規性。由於對平台安全性及其投資的擔憂，您的投資者會停止對公司的支援。
- 您已決定包含其中一位開發人員在網際網路上發現的程式庫。您尚未評估從未知來源採用此程式庫的風險，並且不知道它是否包含弱點或惡意程式碼。
- 遷移的最初業務理由是基於應用程式工作負載 60% 的現代化。然而，由於技術上的困難，決定只對 20% 進行現代化，導致長期計劃收益減少，基礎架構團隊手動支援舊式系統的操作員工作量增加，並且更依賴於在沒有規劃此變更的基礎架構團隊中開發新技能。

建立此最佳實務的優勢：充分協調和支援董事會層級的業務優先事項、了解取得成功的風險、制定明智決策，以及在風險阻礙成功機會時採取適當行動。了解決策的影響和後果有助於您優先考慮您的選擇，並更快地讓領導者達成一致，從而改善業務成果。確定您的選擇的可用好處並了解組織的風險，可以幫助您制定資料驅動型決策，而不是依賴於軼事。

未建立此最佳實務時的曝險等級：中

實作指引

管理利益和風險應由推動關鍵決策要求的管理機構來定義。您希望根據決策如何使組織受益來做出決策並確定優先級，並了解所涉及的風險。準確的資訊對於制定組織決策至關重要。這應該基於可靠的測量

結果，並由成本效益分析的常見行業實務進行定義。要做出這些類型的決策，請在集中式和分散式授權之間取得平衡。總有一種權衡，了解每個選擇如何影響定義的策略和期望的業務成果至關重要。

實作步驟

1. 在整體式雲端治理架構中正式確定效益衡量實務。
 - a. 在決策的中央控制與某些決策的分散權力之間取得平衡。
 - b. 了解強加給每項決策的繁瑣決策流程可能會減緩您的速度。
 - c. 將外部因素納入決策流程中 (例如合規要求)。
2. 為各級決策建立一個商定的決策框架，其中包括誰需要為受衝突利益影響的決策清除障礙。
 - a. 集中化可能不可逆轉的單向門決策。
 - b. 允許由較低級別的組織領導者做出雙向門決策。
3. 了解和管理利益和風險。在決策的收益與所涉及的風險之間取得平衡。
 - a. 確定收益：根據業務目標、需求和優先事項確定收益。範例包括企劃案影響、上市時間、安全性、可靠性、效能和成本。
 - b. 確定風險：根據業務目標、需求和優先事項確定風險。範例包括上市時間、安全性、可靠性、效能和成本。
 - c. 根據風險評估收益並做出明智決策：根據利益相關者 (包括業務、開發和營運團隊) 的目標、需求和優先事項，確定收益和風險的影響。評價收益的價值時要考慮發生風險的可能性及其代價。例如，強調上市速度優先於可靠性，可能提供競爭優勢。不過，如果發生可靠性問題，則可能會縮短正常執行時間。
4. 以程式設計方式強制執行關鍵決策，讓您自動遵守合規要求。
5. 利用已知的產業架構和功能，例如「價值串流分析」和 LEAN，對目前的狀態效能、業務指標進行基準分析，並定義這些指標的進度迭代。

實作計劃的工作量：中高

資源

相關的最佳實務：

- [OPS01-BP05 評估威脅環境](#)

相關文件：

- [Amazon 的 Day 1 文化要素 | 做出高品質、高速度的決策](#)

- [雲端治理](#)
- [管理與治理雲端環境](#)
- [雲端和數位時代的治理：第一部分和第二部分](#)

相關影片：

- [播客 | Jeff Bezos | 關於如何制定決策](#)

相關範例：

- [使用資料做出明智決策 \(DevOps Sagas\)](#)
- [使用開發價值串流映射來識別 DevOps 結果的限制](#)

OPS 2. 如何建構組織以支援業務成果？

您的團隊必須了解其在達成業務成果中所扮演的角色。團隊應了解本身在促成其他團隊成功的過程中所扮演的角色、其他團隊在獲致成功的過程中所扮演的角色，以及擁有共同目標。了解責任、擁有權、決策方式，以及誰有權制定決策，將有助於找到工作重點，並充分發揮團隊的優勢。

最佳實務

- [OPS02-BP01 資源已識別擁有者](#)
- [OPS02-BP02 流程和程序已識別擁有者](#)
- [OPS02-BP03 已為營運活動識別負責其效能的擁有者](#)
- [OPS02-BP04 存在管理責任和擁有權的機制](#)
- [OPS02-BP05 存在用於要求新增、變更和例外狀況的機制](#)
- [OPS02-BP06 團隊之間的責任是預先定義或經過協商的](#)

OPS02-BP01 資源已識別擁有者

工作負載的資源必須已識別變更控制、疑難排解和其他功能的擁有者。系統會為工作負載、帳戶、基礎設施、平台和應用程式指派擁有者。擁有權會使用集中註冊或連接至資源的中繼資料等工具來記錄。元件的商業價值會透露其適用的流程和程序。

預期成果：

- 資源已使用中繼資料或中央寄存器識別擁有者。

- 團隊成員可以識別誰擁有資源。
- 帳戶在可能的情況下擁有單一擁有者。

常見的反模式：

- 您的替代聯絡人 AWS 帳戶 不會填入。
- 資源缺少可識別其屬於哪個團隊的標籤。
- 您有一個沒有電子郵件映射的ITSM佇列。
- 兩個團隊對基礎設施的關鍵部分的擁有權重疊。

建立此最佳實務的優勢：

- 透過指派擁有權，資源的變更控制很簡單。
- 疑難排解問題時，可以讓合適的擁有者參與。

未建立此最佳實務時的曝險等級：高

實作指引

定義擁有權對環境中的資源使用案例的意義。擁有權可能意味著誰監督資源的變更、在疑難排解期間支援資源或者是負有財務責任的人員。指定並記錄資源的擁有者，包括名稱、聯絡資訊、組織和團隊。

客戶範例

AnyCompany 零售將所有權定義為擁有變更和資源支援之團隊或個人。他們利用 AWS Organizations 來管理其 AWS 帳戶。備選帳戶連絡人正在使用群組收件匣進行設定。每個ITSM佇列都會對應至電子郵件別名。標籤會識別誰擁有 AWS 資源。對於其他平台和基礎設施，他們有 Wiki 頁面會指出擁有權和聯絡資訊。

實作步驟

1. 首先為您的組織定義擁有權。擁有權可能表示誰擁有資源的風險、誰擁有資源的變更，或誰在疑難排解時支援資源。擁有權也可能意味著資源的財務或管理擁有權。
2. 使用 [AWS Organizations](#) 管理帳戶。您可以集中管理帳戶的替代聯絡人。
 - a. 只要使用公司擁有的電子郵件地址和電話號碼作為聯絡資訊，即使聯絡資訊所屬的個人已離職，您仍可存取這些資訊。例如，為帳單、營運和安全建立各別的電子郵件分發清單，在每個作用中

AWS 帳戶中將這些設定為帳戶、安全和營運聯絡人。即使有人正在休假、變更角色或離開公司，多人也會收到 AWS 通知並能夠回應。

- b. 如果帳戶不是由 [AWS Organizations](#) 管理，備選帳戶聯絡人便會協助 AWS 適時與相關人員取得聯繫。設定帳戶的備選聯絡人以將其指向團體而非個人。
3. 使用標籤來識別 AWS 資源的擁有者。您可以在不同的標籤中指定擁有者及其聯絡資訊。
 - a. 可以使用 [AWS Config](#) 規則來強制資源具有所需的擁有權標籤。
 - b. 如需有關如何為組織建立標記策略的深入指引，請參閱 [AWS Tagging Best Practices whitepaper](#)。
4. 使用 [Amazon Q Business](#)，這是一個使用生成式 AI 的對話式助理，可提高員工生產力、回答問題並根據企業系統中的資訊完成任務。
 - a. 將 Amazon Q Business 連接到您公司的資料來源。Amazon Q Business 為超過 40 個支援的資料來源提供預先建置的連接器，包括 Amazon Simple Storage Service (Amazon S3)、Microsoft SharePoint、Salesforce 和 Atlassian Confluence。如需詳細資訊，請參閱[企業版 Amazon Q 連接器](#)。
5. 對於其他資源、平台和基礎設施，請建立識別擁有權的文件。所有團隊成員都應可以存取。

實作計劃的工作量：低。利用帳戶聯絡資訊和標籤來指派 AWS 資源的所有權。對於其他資源，您可以使用 Wiki 中的資料表來記錄擁有權和聯絡資訊，或使用 ITSM 工具映射擁有權。

資源

相關的最佳實務：

- [OPS02-BP02 程序已識別擁有者](#)
- [OPS02-BP04 存在機制來管理責任和所有權](#)

相關文件：

- [AWS 帳戶管理 - 更新聯絡資訊](#)
- [AWS Organizations - 更新組織中的替代聯絡人](#)
- [《AWS 標記最佳實務》白皮書](#)
- [使用 Amazon Q Business and Identity Center 建置私有且 AWS IAM 安全的企業生成 AI 應用程式](#)
- [Amazon Q Business 現已正式推出，可透過生成式 AI 提升員工生產力](#)
- [AWS 雲端 Operations & Migrations 部落格 - 使用 和 實作自動化 AWS Config 和集中式標記控制項 AWS Organizations](#)

- [AWS 安全部落格 - 使用 擴展您的預先遞交掛鉤 AWS CloudFormation Guard](#)
- [AWS DevOps 部落格 - AWS CloudFormation Guard 整合至 CI/CD 管道](#)

相關研討會：

- [AWS 研討會 - 標記](#)

相關範例：

- [AWS Config 規則 - EC2具有必要標籤和有效值的 Amazon](#)

相關服務：

- [AWS Config 規則 - 必要標籤](#)
- [AWS Organizations](#)

OPS02-BP02 流程和程序已識別擁有者

了解誰具有個別流程和程序的擁有權、為何使用特定流程和程序，以及為何該擁有權存在。了解使用特定流程和程序的原因，能夠幫助發現改進機會。

預期成果：您的組織擁有一套明確定義且維護良好的操作任何流程和程序。流程和程序儲存在中心位置，並可供您的團隊成員使用。透過明確指定的擁有權經常更新流程和程序。如果可能，指令碼、範本和自動化文件會以程式碼的形式實作。

常見的反模式：

- 未記錄處理程序。隔離的操作員工作站上可能存在碎片化指令碼。
- 有關如何使用指令碼的知識由少數個人掌握，或非正式地作為團隊知識。
- 舊版流程由於更新而到期，但更新的擁有權不清楚，原始著作人不再是組織的一部分。
- 流程和指令碼不容易發現，因此在需要時 (例如，回應事故) 無法立即使用。

建立此最佳實務的優勢：

- 流程和程序可大幅提升您操作工作負載的效率。
- 新的團隊成員更快速地變得高效。

- 減少事故緩解的時間。
- 不同的團隊成員 (和團隊) 可以以一致的方式使用相同的流程和程序。
- 團隊可以透過可重複的流程擴展其流程。
- 標準化流程和程序有助於減輕團隊之間轉移工作負載責任的影響。

未建立此最佳實務時的曝險等級：高

實作指引

- 流程和程序已經確定負責其定義的擁有者。
 - 識別為支援工作負載所執行的營運活動。將這些活動記錄在可探索的位置中。
 - 唯一識別負責活動規格的個人或團隊。他們負責確認具備適當技能的團隊成員能夠成功執行該活動，且該團隊成員具備正確許可、存取權和工具。如果執行該活動時發生問題，執行該活動的團隊成員需負責提供改善活動所需的詳細回饋。
 - 透過 AWS Systems Manager 等服務，使用文件和 AWS Lambda 來擷取活動成品中繼資料中的擁有權。使用標籤或資源群組擷取資源擁有權，並指定擁有權和聯絡資訊。使用 AWS Organizations 建立標記政策，並擷取擁有權和聯絡資訊。
- 隨著時間的推移，這些程序應逐步發展為可以作為程式碼執行，從而減少人為干預的需求。
 - 例如，請考慮 AWS Lambda 函數、CloudFormation 範本或 AWS Systems Manager 自動化文件。
 - 在適當的儲存庫中執行版本控制。
 - 包括合適的資源標記，以便可以很容易地識別擁有者和文件。

客戶範例

AnyCompany Retail 將擁有權定義為擁有應用程式或應用程式群組 (共享共同的架構實務和技術) 流程的團隊或個人。最初，流程和程序作為分步指南記錄在文件管理系統中，可使用託管應用程式的 AWS 帳戶中的標籤以及帳戶內特定資源群組中的標籤進行探索。他們利用 AWS Organizations 來管理他們的 AWS 帳戶。隨著時間的推移，這些流程會轉換為程式碼，並使用基礎設施即程式碼 (例如 CloudFormation 或 AWS Cloud Development Kit (AWS CDK) 範本) 來定義資源。操作流程會變成 AWS Systems Manager 或 AWS Lambda 函數中的自動化文件，這些自動化文件可以作為排程任務啟動，以回應諸如 AWS CloudWatch 警示或 AWS EventBridge 事件之類的事件，或由 IT 服務管理 (ITSM) 平台內的請求啟動。所有流程都有標籤來識別擁有權。在流程的程式碼儲存庫所產生的 wiki 頁面中維護自動化和流程的文件。

實作步驟

1. 記錄現有的流程和程序。
 - a. 審核並將其維持在最新狀態。
 - b. 識別每個流程或程序的擁有者。
 - c. 將其置於版本控制之下。
 - d. 在可能的情況下，在共用架構設計的工作負載和環境之間共用流程和程序。
2. 建立回饋和改進機制。
 - a. 定義審核流程頻率的政策。
 - b. 定義審核者與核准者的流程。
 - c. 實作問題或票證隊列，以提供和追蹤意見回饋。
 - d. 在可能的情況下，流程和程序應得到變更核准委員會 (CAB) 的預先批准和風險分類。
3. 驗證流程和程序是否可供需要執行流程和程序的人員來存取和探索。
 - a. 使用標籤來指示可以針對工作負載存取流程和程序的位置。
 - b. 使用有意義的錯誤和事件訊息來指出可解決問題的適當流程或程序。
 - c. 使用 Wiki 和文件管理，讓整個組織可一致搜尋流程和程序。
4. 使用 [Amazon Q Business](#)，這是一個使用生成式 AI 的對話式助理，可提高員工生產力、回答問題並根據企業系統中的資訊完成任務。
 - a. 將 Amazon Q Business 連接到您公司的資料來源。Amazon Q Business 為 40 多個受支援的資料來源提供預先建置的連接器，其中包括 Amazon S3、Microsoft SharePoint、Salesforce 和 Atlassian Confluence。如需詳細資訊，請參閱 [Amazon Q 連接器](#)。
5. 適時進行自動化。
 - a. 當服務和技術提供 API 時，應該開發自動化。
 - b. 對流程進行充分的教育。制定使用者故事和要求以自動化這些流程。
 - c. 成功衡量流程和程序的使用情況，並建立問題或票證以支援反覆改進。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS02-BP01 已為資源識別擁有者](#)
- [OPS02-BP04 存在管理責任和擁有權的機制](#)

- [OPS11-BP04 執行知識管理](#)

相關文件：

- [AWS 白皮書 - DevOps on AWS 簡介](#)
- [AWS 白皮書 - 標記 AWS 資源的最佳實務](#)
- [AWS 白皮書 - 使用多個帳戶組織您的 AWS 環境](#)
- [AWS 雲端 操作和遷移部落格 - 使用 Amazon Q Business 簡化您的操作](#)
- [AWS 雲端 操作和遷移部落格 - 建立雲端自動化實務以實現卓越營運：AWS Managed Services 的最佳實務](#)
- [AWS 雲端 操作和遷移部落格 - 使用 AWS Config 和 AWS Organizations 實作自動化和集中式標記控制](#)
- [AWS 安全部落格 - 使用 AWS CloudFormation Guard 擴展您的預先提交勾點](#)
- [AWS DevOps 部落格 - 將 AWS CloudFormation Guard 整合到 CI/CD 管道中](#)

相關研討會：

- [AWS Well-Architected 卓越營運研討會](#)
- [AWS 研討會 - 標記](#)

相關影片：

- [如何在 AWS 上自動執行 IT 操作](#)
- [AWS re:Invent 2020 - 使用 AWS Systems Manager 自動執行任何操作](#)
- [AWS re:Inforce 2022 - 使用 AWS 自動化修補程式管理和合規性 \(NIS306\)](#)
- [支援 支援您 - 深入了解 AWS Systems Manager](#)

相關服務：

- [AWS Systems Manager - 自動化](#)
- [AWS Service Management Connector](#)

OPS02-BP03 已為營運活動識別負責其效能的擁有者

了解誰負責在已定義的工作負載上執行特定活動，以及為什麼該責任存在。透過了解誰負責執行活動，可得知誰將會進行活動、驗證結果，以及提供回饋給活動擁有者。

預期成果：

您的組織清楚地規定了對已定義的工作負載執行特定活動並回應工作負載所產生事件的責任。該組織記錄了流程和履行的擁有權，並使此資訊可被發現。您可以在組織變更發生時審核並更新責任，而團隊會追蹤並衡量缺陷和低效率識別活動的效能。可以實作回饋機制來追蹤缺陷和改進並支援迭代改進。

常見的反模式：

- 您不記錄責任。
- 隔離的操作員工作站上存在碎片化指令碼。只有少數人知道如何進行使用，或非正式地將其稱為團隊知識。
- 舊版流程由於更新而到期，但沒有人知道誰擁有該流程，並且原始著作人不再是組織的一部分。
- 流程和指令碼無法被發現，在需要時 (例如，回應事故) 無法立即使用。

建立此最佳實務的優勢：

- 了解誰負責執行活動，在需要採取動作時通知誰，以及誰會執行動作、驗證結果並為活動擁有者提供意見回饋。
- 流程和程序可大幅提升您操作工作負載的效率。
- 新的團隊成員更快速地變得高效。
- 可以減少緩解事故所需的時間。
- 不同的團隊可採取一致的方式來使用相同的流程和程序。
- 團隊可以透過可重複的流程擴展其流程。
- 標準化流程和程序有助於減輕團隊之間轉移工作負載責任的影響。

未建立此最佳實務時的曝險等級：高

實作指引

若要開始定義職責，請先從現有文件開始，例如責任矩陣、流程與程序、角色與責任，以及工具與自動化。審查並主持有關文件化流程責任的討論。與團隊一起審查，以識別文件責任和流程之間的不一致。與該團隊的內部客戶討論提供的服務，以確定團隊之間的期望差距。

分析並解決差異。找出改進的機會，並尋找經常請求的資源密集型活動，這些活動通常需要改進。探索最佳實務、模式和規範性指引，以簡化和標準化改進。記錄改善機會，並追蹤改進完成情況。

隨著時間的推移，這些程序應逐步發展為可以作為程式碼執行，從而減少人為干預的需求。例如，程序可以啟動為 AWS Lambda 函數、CloudFormation 範本或 AWS Systems Manager 自動化文件。驗證這些程序是否在適當的儲存庫中受版本控制，並包含適當的資源標記，以便團隊能夠輕鬆識別擁有者和文件。記錄執行活動的責任，然後監控用於成功啟動和操作的自動化，以及期望成果的績效。

客戶範例

AnyCompany Retail 將擁有權定義為擁有應用程式或應用程式群組 (共享共同的架構實務和技術) 流程的團隊或個人。最初，公司將流程和程序記錄為文件管理系統中的分步指南。其會使用託管應用程式的 AWS 帳戶中的標籤以及帳戶內特定資源群組中的標籤來探索程序，並使用 AWS Organizations 來管理其 AWS 帳戶。隨著時間的推移，AnyCompany Retail 會將這些流程轉換為程式碼，並使用基礎設施即程式碼 (例如 CloudFormation 或 AWS Cloud Development Kit (AWS CDK) 範本) 來定義資源。操作流程會變成 AWS Systems Manager 或 AWS Lambda 函數中的自動化文件，它們可以作為排程任務啟動，以回應諸如 Amazon CloudWatch 警示或 Amazon EventBridge 事件之類的事件，或由 IT 服務管理 (ITSM) 平台內的請求啟動。所有流程都有標籤，可用來識別其擁有者。團隊會在流程的程式碼儲存庫所產生的 wiki 頁面中管理自動化和流程的文件。

實作步驟

1. 記錄現有的流程和程序。
 - a. 審核並確認其是否為最新狀態。
 - b. 確認每個流程或程序都有擁有者。
 - c. 將程序置於版本控制之下。
 - d. 在可能的情況下，在共用架構設計的工作負載和環境之間共用流程和程序。
2. 建立回饋和改進機制。
 - a. 定義審核流程頻率的政策。
 - b. 定義審核者與核准者的流程。
 - c. 實作問題或票證隊列，以提供並追蹤意見回饋。
 - d. 在可能的情況下，為流程和程序提供變更核准委員會 (CAB) 的預先批准和風險分類。
3. 讓流程和程序可供需要執行它們的人員進行存取和探索。
 - a. 使用標籤來指示可以針對工作負載存取流程和程序的位置。
 - b. 使用有意義的錯誤和事件訊息來指出可解決問題的適當流程或程序。
 - c. 使用 Wiki 或文件管理，讓整個組織可一致搜尋流程和程序。

4. 在適當的時候實現自動化。
 - a. 當服務和技術提供 API 時，開發自動化。
 - b. 驗證是否充分理解流程，制定使用者故事和要求以自動化這些流程。
 - c. 衡量流程和程序的成功使用情況，並追蹤問題以支援反覆改進。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS02-BP01 已為資源識別擁有者](#)
- [OPS02-BP02 流程和程序已識別擁有者](#)
- [OPS02-BP04 存在管理責任和擁有權的機制](#)
- [OPS02-BP05 存在識別責任和擁有權的機制](#)
- [OPS11-BP04 執行知識管理](#)

相關文件：

- [AWS 白皮書 | DevOps on AWS 簡介](#)
- [AWS 白皮書 | 標記 AWS 資源的最佳實務](#)
- [AWS 白皮書 | 使用多個帳戶組織您的 AWS 環境](#)
- [AWS 雲端 操作和遷移部落格 | 建立雲端自動化實務以實現卓越營運：AWS Managed Services 的最佳實務](#)
- [AWS 研討會 - 標記](#)
- [AWS Service Management Connector](#)

相關影片：

- [AWS 知識中心直播 | 標記 AWS 資源](#)
- [AWS re:Invent 2020 | 使用 AWS Systems Manager 自動執行任何操作](#)
- [AWS re:Inforce 2022 | 使用 AWS 自動化修補程式管理和合規性 \(NIS306\)](#)
- [支援 支援您 | 深入了解 AWS Systems Manager](#)

OPS02-BP04 存在管理責任和擁有權的機制

了解您角色的責任以及您為業務成果做出貢獻的方式，因為這種了解可明確任務的優先順序以及您的角色為何重要。這有助於團隊成員辨識需求並適當地回應。當團隊成員知道的角色時，他們可以建立擁有、找出改進機會，以及了解如何影響或進行適當的變更。

有時候，責任可能沒有明確的擁有者。在這些情況下，設計一種機制來解決此差距。為有權指派擁有權或計劃解決需求的人員建立明確定義的上報路徑。

預期成果：組織內的團隊擁有明確定義的職責，其中包括他們與資源、要執行的動作、流程及程序的關係。這些職責符合團隊的責任和目標，以及其他團隊的責任。您可以用一致且可探索的方式記錄上報路徑，並將這些決策輸入到文件成品中，例如責任矩陣、團隊定義或 Wiki 頁面。

常見的反模式：

- 團隊的責任含糊不清或定義不佳。
- 團隊沒有將角色與責任保持一致。
- 該團隊沒有調整其總體目標和具體目標以及其責任，這使得它難以衡量成功。
- 團隊成員的責任與團隊和更廣泛的組織不一致。
- 您的團隊不會更新職責，這導致其與團隊執行的任務不一致。
- 確定職責的上報路徑尚未定義或不清楚。
- 上報路徑沒有單一執行緒擁有者，以確保及時回應。
- 角色、責任和上報路徑不容易發現，在需要時 (例如，回應事故) 無法立即使用。

建立此最佳實務的優勢：

- 了解誰擁有責任或擁有權時，可讓您聯絡適當的團隊或團隊成員，以提出請求或轉換任務。
- 為了降低不作為和未解決需求的風險，您已經確定了有權指派責任或擁有權的人員。
- 當您明確定義責任範圍時，團隊成員將獲得自主權和擁有權。
- 您的責任決定了您所做的決定、您採取的動作，以及如何將活動交給其適當的擁有者。
- 很容易識別被放棄的責任，因為您清楚地了解哪些責任不在團隊責任範圍內，這有助於您上報以進行澄清。
- 團隊可以避免混亂和緊張，並且可以更充分地管理工作負載和資源。

未建立此最佳實務時的曝險等級：高

實作指引

確定團隊成員的角色和責任，並確保他們了解其角色的期望。讓此資訊可供探索，如此組織的成員便能夠確定他們針對特定需求需要聯絡的人員 (團隊或個人)。隨著組織尋求利用機會在 AWS 上進行遷移和現代化，角色和職責也可能會改變。讓您的團隊及其成員了解他們的責任，並對其進行適當的培訓，以便在此次變更期間執行其任務。

確定應接收上報的角色或團隊，以確認責任和擁有權。此團隊可以與各種利益相關者互動以做出決定。但是，他們應該擁有決策制定流程的管理權。

為您的組織成員提供可存取的機制，以探索和識別擁有權和責任。這些機制教導他們應該聯絡誰以滿足特定需求。

客戶範例

AnyCompany Retail 最近使用平移方法，完成了在 AWS 中將工作負載從內部部署環境遷移到其登陸區域。他們執行了操作審查以反映他們如何完成一般操作任務，並驗證其現有的責任矩陣是否反映了新環境中的操作。當他們從內部部署遷移到 AWS 時，他們減少了與硬體和實體基礎結構相關的基礎架構團隊的責任。此舉還揭示了為其工作負載發展營運模式的新機會。

他們在確定、處理並記錄大部分職責的同時，還定義了任何遺漏或隨著營運實務演變而可能需要變更的任何責任的上報路徑。若要探索標準化和改善工作負載效率的新機會，請提供 AWS Systems Manager 等操作工具以及 AWS Security Hub CSPM 和 Amazon GuardDuty 等安全工具的存取權。AnyCompany Retail 根據他們希望首先解決的改進來審查責任和策略。由於公司採用新的工作方式和技術模式，他們會更新其責任矩陣以進行匹配。

實作步驟

1. 從現有文件開始。一些典型的來源文件可能包括：
 - a. 責任或負責者、當責者、事先諮詢者和事後告知者 (RACI) 矩陣
 - b. 團隊定義或 Wiki 頁面
 - c. 服務定義和產品
 - d. 角色或職位描述
2. 審查並主持有關文件化責任的討論：
 - a. 與團隊進行審查，以確定文件化責任與團隊通常執行的責任之間的不一致。
 - b. 討論內部客戶提供的潛在服務，以確定團隊之間的期望差距。
3. 分析並解決差異。
4. 找出改進機會。

- a. 找出頻繁發出的資源密集型請求，這類請求通常表示需要改進。
 - b. 尋找最佳實務、了解模式、遵循規範性指引，並簡化和標準化改進。
 - c. 記錄改進機會，並追蹤完成情況。
5. 如果團隊尚未負責管理和追蹤職責指派，請指定團隊中負責此職責的人員。
6. 為團隊定義一個流程以請求澄清責任。
- a. 審核流程，並確認其清晰且易用。
 - b. 確保有人擁有並追蹤他們結論的上報。
 - c. 建立運營指標以衡量有效性。
 - d. 建立意見回饋機制，以確認團隊是否能突顯改進機會。
 - e. 實作定期審查機制。
7. 文件存放在可發現且可存取的位置。
- a. Wiki 或文件入口網站是常見選擇。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS01-BP06 評估權衡](#)
- [OPS03-BP02 授權團隊成員在成果有風險時採取動作](#)
- [OPS03-BP03 鼓勵向上呈報](#)
- [OPS03-BP07 適當地為團隊提供資源](#)
- [OPS09-BP01 使用指標衡量營運目標與 KPI](#)
- [OPS09-BP03 審核營運指標並優先改進](#)
- [OPS11-BP01 建立持續改進程序](#)

相關文件：

- [AWS 白皮書 - DevOps on AWS 簡介](#)
- [AWS 白皮書 - AWS 雲端 採用架構：營運觀點](#)
- [AWS Well-Architected Framework 卓越營運 - 工作負載層級操作模型拓撲](#)
- [AWS 方案指引 - 建置雲端操作模式](#)

- [AWS 方案指引 - 為雲端操作模式建立 RACI 或 RASCI 矩陣](#)
- [AWS 雲端 操作和遷移部落格 - 透過雲端平台團隊提供商業價值](#)
- [AWS 雲端 操作和遷移部落格 - 為何選擇雲端操作模式？](#)
- [AWS DevOps 部落格 - 組織如何實現雲端操作的現代化](#)

相關影片：

- [AWS Summit Online - 加速轉型的雲端操作模式](#)
- [AWS re:Invent 2023 - 面向未來的雲端安全性：一種新的操作模式](#)

OPS02-BP05 存在用於要求新增、變更和例外狀況的機制

您可以向流程、程序和資源的擁有者提出請求。請求包含新增、變更和例外狀況。這些請求會經歷變更管理程序。評估收益和風險後，若可行並經判斷是合適的行為，則應制定明智的決策以核准請求。

預期成果：

- 可以根據分配的擁有權請求變更流程、程序和資源。
- 變更是經過深思熟慮的，權衡了利益和風險。

常見的反模式：

- 必須更新部署應用程式的方式，但無法向操作團隊請求變更部署程序。
- 必須更新災難復原計畫，但沒有確定的擁有者可以請求變更。

建立此最佳實務的優勢：

- 流程、程序和資源可隨需求的變化而發展。
- 擁有者可以在進行變更時做出明智決策。
- 變更是經過深思熟慮的。

未建立此最佳實務時的曝險等級：中

實作指引

若要實作此最佳實務，必須能夠要求變更流程、程序和資源。變更管理流程可以很簡單。記錄變更管理流程。

客戶範例

AnyCompany Retail 使用責任指派 (RACI) 矩陣來確定誰擁有流程、程序和資源的變更。他們擁有記錄在案的變更管理流程，簡單且易於遵循。使用 RACI 矩陣和流程，任何人都可以提交變更請求。

實作步驟

1. 確定工作負載的流程、程序和資源，以及各自的擁有者。將其記錄在您的知識管理系統中。
 - a. 如果尚未實作 [OPS02-BP01 資源已識別擁有者](#)、[OPS02-BP02 流程和程序已識別擁有者](#) 或 [OPS02-BP03 已為營運活動識別負責其效能的擁有者](#)，請先從這些項目開始。
2. 與組織中的利益相關者合作，制定變更管理流程。此流程應涵蓋資源、流程及程序的新增、變更及例外狀況。
 - a. 可以使用 [AWS Systems Manager Change Manager](#) 做為工作負載資源的變更管理平台。
3. 在您的知識管理系統中記錄變更管理流程。

實作計劃的工作量：中。制定變更管理流程需要與組織中的多個利益相關者保持一致。

資源

相關的最佳實務：

- [OPS02-BP01 資源已識別擁有者](#) - 在建立變更管理流程之前，資源需要確定的擁有者。
- [OPS02-BP02 流程和程序已識別擁有者](#) - 在建立變更管理流程之前，流程需要確定的擁有者。
- [OPS02-BP03 已為營運活動識別負責其效能的擁有者](#) - 在建立變更管理流程之前，操作活動需要確定的擁有者。

相關文件：

- [AWS 方案指引 - AWS 大型移轉的基礎說明手冊：建立 RACI 矩陣](#)
- [雲端白皮書中的變更管理](#)

相關服務：

- [AWS Systems Manager Change Manager](#)

OPS02-BP06 團隊之間的責任是預先定義或經過協商的

團隊間已定義或協商說明如何相互配合及支援的協議 (例如，回應時間、服務水準目標或服務水準協議)。團隊間的溝通管道記錄於文件中。透過了解團隊工作對於商業成果和其他團隊及組織成果的影響，可得知其任務的優先順序，並協助他們適當地回應。

如果責任和擁有權未定義或未知，則您會面臨風險，不僅無法及時處理必要的活動，在解決這些需求時還會出現冗餘和可能相互衝突的工作。

預期成果：

- 團隊間的工作或支援協議已達成一致並記錄在案。
- 彼此支援或合作的團隊已定義了溝通渠道和回應期望。

常見的反模式：

- 生產環境中發生問題，而且兩個不同的團隊開始彼此獨立地進行疑難排解。他們各自為政的工作延長了中斷時間。
- 運營團隊需要開發團隊的幫助，但沒有商定回應時間。請求卡在待辦事項中。

建立此最佳實務的優勢：

- 團隊知道如何互動和互相支援。
- 對回應的期望是眾所周知的。
- 溝通渠道已明確定義。

未建立此最佳實務時的曝險等級：低

實作指引

實作此最佳實務意味著團隊之間的合作方式沒有歧義。正式協議規定了團隊如何一起工作或相互支援。團隊間的溝通渠道記錄於文件中。

客戶範例

AnyCompany Retail 的 SRE 團隊與其開發團隊簽訂了服務水準協議。每當開發團隊在其票證系統中提出請求時，他們可以期望在十五分鐘內得到回應。如果發生站點中斷，SRE 團隊在開發團隊的支援下牽頭進行調查。

實作步驟

1. 與組織中的利益相關者合作，根據流程和程序來制定團隊之間的協議。
 - a. 如果在兩個團隊之間共享一個流程或程序，則制定有關團隊將如何一起工作的執行手冊。
 - b. 如果團隊之間存在依賴關係，請同意請求的回應 SLA。
2. 在知識管理系統中記錄責任。

實作計劃的工作量：中。如果團隊之間沒有現有協議，則可能需要努力與組織中的利益相關者達成協議。

資源

相關的最佳實務：

- [OPS02-BP02 流程和程序已識別擁有者](#) - 在團隊之間達成協議之前，必須確定流程所有權。
- [OPS02-BP03 已為營運活動識別負責其效能的擁有者](#) - 在團隊之間達成協議之前，必須確定運營活動所有權。

相關文件：

- [AWS Executive Insights - 透過雙披薩團隊賦予創新力量](#)
- [DevOps on AWS 簡介 - 雙披薩團隊](#)

OPS 3. 您的組織文化如何支援您的業務成果？

為您的團隊成員提供支援，讓他們能夠更有效地採取動作以及支援業務成果。

最佳實務

- [OPS03-BP01 提供高層支持](#)
- [OPS03-BP02 授權團隊成員在成果有風險時採取動作](#)
- [鼓勵 OPS03-BP03 升級](#)
- [OPS03-BP04 溝通需及時、清楚且可行](#)

- [OPS03-BP05 鼓勵進行試驗](#)
- [OPS03-BP06 管理團隊成員維持和發展自己的技能集](#)
- [OPS03-BP07 資源團隊適當](#)

OPS03-BP01 提供高層支持

在最高層面上，高級領導層充當執行倡議者，以清楚地設定組織成果的期望和方向，包括評估其成功。倡議者倡導並推動最佳實務的採用和組織的發展。

預期成果：努力採用、轉型和最佳化雲端操作的組織可針對預期成果設定明確的領導層和問責製。該組織了解組織實現新成果所需的每種能力，並將擁有權分配給職能團隊進行開發。領導層積極制定此方向、分配擁有權、承擔責任並定義工作。因此，整個組織的個人都可以調動起來、備受鼓舞並積極朝著預期目標努力。

常見的反模式：

- 在沒有明確的雲端營運倡議者和計劃的情況下，工作負載擁有者必須將工作負載遷移至 AWS。這導致團隊無法自覺地協作以改善和充分發展其營運能力。缺乏營運最佳實務標準使團隊不堪重負 (例如營運商辛苦工作、隨叫隨到以及技術債務)，這會限制創新。
- 一個新的全組織目標已設定，即在不提供領導倡議者和策略的情況下採用新興技術。團隊以不同的方式闡述目標，這會導致對於將精力集中在哪裡、為什麼重要以及如何衡量影響感到困惑。因此，組織在採用技術方面失去了動力。

建立此最佳實務的優勢：當高層的支援明確傳達並分享願景、方向和目標時，團隊成員知道他們的期望是什麼。當領導者積極參與時，個人和團隊開始將精力集中在相同的方向上，以實現定義的目標。因此，組織得以運用最強大的能力來獲得成功。當您評估成功時，您可以更好地識別成功障礙，以便透過高層支持的干預來解決這些障礙。

未建立此最佳實務時的曝險等級：高

實作指引

- 在雲端之旅的每個階段 (遷移、採用或最佳化)，成功都需要最高領導層與指定執行倡議者的積極參與。執行倡議者可根據定義的策略來調整團隊的心態、技能組合和工作方式。
 - 解釋原因：澄清並解釋願景和策略背後的原因。
 - 設定期望：為您的組織定義和發佈目標，包括如何衡量進度和成功。
 - 追蹤目標的達成情況：定期衡量目標逐步達成的情況 (不只是完成任務)。分享結果，以便在成果有風險時可以採取適當的行動。

- 提供實現目標所需的資源：將人員和團隊聚集在一起，共同合作並構建正確的解決方案，以實現定義的成果。這可減少或消除組織摩擦。
- 倡導您的團隊：與您的團隊保持互動，以便了解他們的表現以及是否有影響他們的外部因素。找出阻礙您團隊進度的障礙。代表您的團隊來協助解決障礙並消除不必要的負擔。當您的團隊受到外部因素影響時，請重新評估目標並適當地調整目標。
- 推動採用最佳實務：確認提供量化效益的最佳實務，並認可建立者和採用者。鼓勵進一步採用，以擴大已達成的效益。
- 鼓勵團隊發展：創造持續改進的文化，並主動從取得的進步和失敗中學習。鼓勵人員和組織的成長和發展。利用資料和軼事來發展願景和策略。

客戶範例

AnyCompany Retail 正在透過快速重新打造客戶體驗、提高生產力，以及透過生成式 AI 加速成長來實現業務轉型。

實作步驟

1. 建立單一執行緒領導，並指派主要執行倡議者來領導和推動轉型。
2. 定義轉型的明確業務成果，並指派擁有權和責任。讓主要執行人員具有領導和做出關鍵決策的權限。
3. 驗證您的轉型策略非常清晰，並且由執行倡議者廣泛傳達到組織的各個層級。
 - a. 為 IT 和雲端計畫建立明確定義的業務目標。
 - b. 記錄關鍵業務指標，以推動 IT 和雲端轉型。
 - c. 將願景一致地傳達給負責策略各部分的所有團隊和個人。
4. 制定溝通規劃矩陣，指定需要傳遞給特定領導人、經理和個別貢獻者的訊息。指定應傳遞此訊息的人員或團隊。
 - a. 一致且可靠地履行溝通計劃。
 - b. 定期透過面對面活動設定並管理期望。
 - c. 接受有關溝通有效性的反饋，並相應地調整溝通並進行計劃。
 - d. 安排溝通活動以主動了解團隊的挑戰，並建立一致的回饋迴圈，以便在必要時糾正過程。
5. 從領導角度積極參與每項舉措，以驗證所有受影響的團隊都了解他們應負責實現的成果。
6. 在每次狀態會議上，執行倡議者都應該查找阻礙因素，檢查既定的指標、軼事或團隊的反饋，並衡量實現目標的進展情況。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS03-BP04 溝通需及時、清楚且可行](#)
- [OP11-BP01 建立持續改進程序](#)
- [OPS11-BP07 執行營運指標審查](#)

相關文件：

- [解決組織困擾：高度一致](#)
- [正在實施轉型：務實地應對變化](#)
- [成為面向未來的企業](#)
- [建置 CCOE 時應避開的 7 大陷阱](#)
- [導覽雲端：成功的關鍵績效指標](#)

相關影片：

- [AWS re:Invent 2023：生成式 AI 的領導者指南：利用歷史塑造未來 \(SEG204\)](#)

相關範例：

- [Prosci：主要倡議者的角色和重要性](#)

OPS03-BP02 授權團隊成員在成果有風險時採取動作

由領導層灌輸的擁有權文化行為使所有員工都感到有權代表整個公司行事，超出了其定義的角色和責任範圍。員工可以採取行動，在風險出現時主動識別風險，並採取適當的措施。這樣的文化使員工能夠在了解情況的同時做出高價值決策。

例如，Amazon 使用[領導方針](#)作為指引，以推動員工在各種情況下前進、解決問題、處理衝突並採取行動的期望行為。

預期成果：領導層的影響力擴大，培養出新的文化，讓組織內較低層級的個人和團隊也能做出關鍵決策 (只要是透過可稽核的許可和安全機制來定義決策即可)。失敗並不氣餒，團隊反覆學習以改善他們的決

策和應對措施，從而解決未來的類似情況。如果某人的行動導致可以使其他團隊受益的改進，則他們會主動分享此類行動中的知識。領導層會衡量運營改進，並激勵個人和組織採用此類模式。

常見的反模式：

- 在識別風險時，組織中沒有明確的指引或機制來進行應對。例如，當員工發現網路釣魚攻擊時，他們無法向安全團隊報告，導致組織中很大一部分人受到攻擊。這會導致資料外洩。
- 客戶抱怨服務無法使用，這主要源於部署失敗。SRE 團隊負責部署工具，而部署的自動復原則在他們的長期藍圖中。在最近推出的應用程式中，其中一位工程師設計了一種解決方案，將其應用程式自動還原到舊版本。雖然他們的解決方案可以成為 SRE 團隊的模式，但其他團隊不會採用，因為沒有追蹤此類改進的流程。組織繼續受到部署失敗的困擾，影響了客戶並導致進一步的負面情緒。
- 為了保持一致，資訊安全團隊會監督一個歷史悠久的流程，以代表連線至其 Amazon EC2 Linux 執行個體的營運商定期輪換共用的 SSH 金鑰。資訊安全團隊需要幾天的時間才能完成輪換金鑰，而且會阻止您連線到這些執行個體。資訊安全團隊內外沒有人建議在 AWS 上使用其他選項來實現相同的結果。

建立此最佳實務的優勢：透過分散權限來制定決策，並授權團隊決定關鍵決策，您可以更快地解決問題，提高成功率。此外，團隊開始意識到主人翁精神，並且失敗是可以接受的。實驗成為文化中流砥柱。經理和董事並不覺得他們在工作的各個方面都受到了微觀管理。

未建立此最佳實務時的曝險等級：中

實作指引

1. 培養一種預期可能會發生失敗的文化。
2. 為組織內各個職能區域定義明確的擁有權和責任。
3. 向每個人傳達擁有權和責任感，以便個人知道誰可以幫助他們促進分散式決策。
4. 定義您的單向和雙向門決策，以幫助個人了解何時需要升級到更高的領導層級。
5. 建立組織意識，讓所有員工有能力在結果出現風險時，在不同層面採取行動。為您的團隊成員提供管控文件、權限級別、工具和機會，以練習有效回應所需的技能。
6. 讓您的團隊成員有機會練習應對各種決策所需的技能。定義決策等級後，請執行演練日，以驗證所有個別參閱者是否了解並能夠演示該過程。
 - a. 提供替代的安全環境，在其中可測試和培訓流程及程序。
 - b. 承認並意識到，當結果出現預先定義的風險等級時，團隊成員有權採取行動。
 - c. 透過指派權限和對其支援的工作負載和元件的存取權，定義團隊成員採取動作的權限。
7. 讓團隊能夠分享學習經驗 (運營成功和失敗)。

8. 使團隊能夠挑戰現狀，並提供機制來追蹤和衡量改進，以及這些改進對組織的影響。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS01-BP06 評估權衡，同時管理收益和風險](#)
- [OPS02-BP05 存在識別責任和擁有權的機制](#)

相關文件：

- [AWS 部落格文章 | 敏捷企業](#)
- [AWS 部落格文章 | 衡量成功：悖論和計畫](#)
- [AWS 部落格文章 | 放手：實現團隊的自主權](#)
- [集中化還是分散化？](#)

相關影片：

- [re:Invent 2023 | 如何不破壞轉型 \(SEG201\)](#)
- [re:Invent 2021 | Amazon 建置者資料中心：Amazon 的卓越營運](#)
- [集中化與分散化](#)

相關範例：

- [使用架構決策記錄來簡化軟體開發專案的技術決策](#)

鼓勵 OPS03-BP03 升級

如果團隊成員認為預期成果存在風險並且達不到預期標準，領導層鼓勵他們將問題和疑慮向上呈報給更高層級的決策者和利益相關者。這是組織文化的一個特徵，並且在各個層面推動。應該儘早且經常向上呈報，以便識別風險，並防止風險引發事件。領導層不會譴責向上呈報問題的個人。

預期成果：整個組織的個人都很樂意將問題呈報給他們的直接和更高級別的領導層。領導層刻意和有意識地建立了期望，即他們的團隊在呈報任何問題時感到非常安全。存在一種機制來呈報組織內每個層級

的問題。當員工呈報到經理時，他們共同決定影響的程度，以及是否應該呈報問題。為了啟動呈報，員工需要包含建議的工作計畫以解決問題。如果直接管理層沒有及時採取行動，則鼓勵員工在強烈認為組織面臨的風險需要呈報時，將問題呈報給最高領導層。

常見的反模式：

- 在雲端轉型計畫狀態會議期間，高層主管沒有提出足夠的探究性問題，以找出問題和阻礙發生的位置。只有好消息被呈現為狀態。CIO 已明確表示她只喜歡聽到好消息，因為提出的任何挑戰都會讓 CEO 認為程式失敗。
- 您是雲端操作工程師，您注意到應用程式團隊並未廣泛採用新的知識管理系統。該公司投入了一年時間和數百萬美元來實作這個新的知識管理系統，但人們仍然在本地編寫他們的執行手冊，並在組織的雲端共享中進行分享，這使得很難找到與所支援的工作負載相關的知識。您試圖引起領導層的注意，因為持續使用此系統可以提高運營效率。當您將其提交給領導該知識管理系統實作的主管時，她會譴責您，因為它讓投資受到質疑。
- 負責強化運算資源的資訊技術團隊已決定制定程序，該程序需要執行必要的掃描，以確保 EC2 執行個體在運算團隊釋出資源以供使用之前完全安全。這為要部署的資源建立了額外一週的時間延遲，這會中斷其 SLA。運算團隊害怕透過雲端將此問題呈報給 VP，因為這會讓資訊安全 VP 出醜。

建立此最佳實務的優勢：

複雜或重大問題在影響業務之前得到解決。浪費的時間更少。風險最小化。在解決問題時，團隊變得更積極主動並注重結果。

未建立此最佳實務時的曝險等級：高

實作指引

在組織的各個層面自由呈報的意願和能力是一個組織和文化基礎，應該透過強調培訓、領導溝通、期望設定以及在組織各個層面的機制部署來有意識地進行發展。

實作步驟

1. 定義組織的政策、標準和期望。
 - a. 確保廣泛採用和理解政策、期望和標準。
2. 在不符合標準時，鼓勵、培訓和授權工人儘早和頻繁呈報。
3. 組織認可儘早且頻繁呈報是最佳實務。接受向上呈報可能經證明是毫無根據的，然而有機會防止事件的發生好過於不向上呈報而錯過機會。

- a. 建立呈報機制 (例如 Andon Cord 系統)。
 - b. 制定書面程序，定義向上呈報的時機與方式。
 - c. 定義一系列具有越來越多權限可採取或批准行動的人員，以及每個利益相關者的聯絡資訊。
4. 當進行呈報時，它應該繼續下去，直到團隊成員確信透過領導層推動的行動緩解了風險。
- a. 呈報應包括：
 - i. 情況描述和風險性質
 - ii. 情況嚴重性
 - iii. 誰或什麼受到影響
 - iv. 影響有多大
 - v. 發生影響時的緊迫性
 - vi. 建議的補救措施和緩解計畫
 - b. 保護向上呈報的員工。如果團隊成員圍繞無回應決策制定者或利益相關者向上呈報，則制定保護團隊成員免受報復的政策。制定機制以識別是否發生此情況，並適當地做出回應。
5. 鼓勵在組織生產的一切產品中建立持續改進回饋迴圈的文化。回饋迴圈充當負責個人的小型呈報，他們可確定改進機會，即使不需要升級。持續改進的文化迫使每個人都更加積極主動。
6. 領導層應定期重新強調政策、標準、機制，以及對公開呈報和持續回饋迴圈而不受報復的願望。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS02-BP05 存在用於要求新增、變更和例外狀況的機制](#)

相關文件：

- [如何培養一種持續改進並向 Andon 和呈報系統學習的文化？](#)
- [Andon Cord \(IT 革命\)](#)
- [AWS DevOps 指引 | 建立明確的呈報路徑，並鼓勵建設性的分歧](#)

相關影片：

- [Jeff Bezos 如何制定決策 \(並提高速度\)](#)

- [Toyota 產品系統：停止生產，一個按鈕以及一個 Andon 電氣板](#)
- [LEAN製造業的 Andon Cord](#)

相關範例：

- [在 Incident Manager 中使用呈報計畫](#)

OPS03-BP04 溝通需及時、清楚且可行

領導層負責建立強大而有效的溝通，尤其是當組織採用新策略、技術或工作方式時。領導者應該為所有員工設定期望，以實現公司目標。設計溝通機制，在負責運行由領導資助和贊助計劃的團隊之間可建立和保持意識。利用跨組織的多樣性，並用心聆聽多個獨特觀點。使用此觀點來增加創新、挑戰假設，並降低確認偏差的風險。在團隊中培養包容性、多樣性和可及性，以獲得有益的觀點。

預期成果：您的組織設計溝通策略以解決變更對組織的影響。團隊保持知情並積極繼續彼此合作，而不是彼此對抗。個人了解他們的角色對於實現既定目標是多麼重要。電子郵件只是一種被動的溝通機制，並相應地使用。管理層花時間與他們的個人貢獻者溝通，以幫助他們了解其責任、要完成的任務，以及他們的工作如何為整體使命做出貢獻。必要時，領導者直接在較小的場所與人們互動，以傳達訊息並確認這些訊息是否有效地傳遞。由於良好的溝通策略，組織的表現達到或超出領導層的期望。領導層鼓勵和尋求團隊內部和團隊之間的不同意見。

常見的反模式：

- 您的組織有五年計劃，可將所有工作負載遷移至 AWS。雲端的業務案例包括所有工作負載的 25% 進行現代化，以充分利用無伺服器技術。CIO 將此策略傳達給直接下屬，並期望每位領導者將此演示文稿提交給經理、董事和個人貢獻者，而無須進行任何面對面溝通。CIO 退後一步，並期望其組織執行新的策略。
- 領導層不提供或使用回饋機制，並且預期差距不斷增長，導致專案停滯。
- 系統會要求您對安全群組進行變更，但不會為您提供任何詳細資訊，說明需要進行哪些變更、變更可能對所有工作負載造成什麼影響，以及何時發生變更。經理會轉發來自資訊安全網副總裁的電子郵件，並新增訊息「去實現 (Make this happen)」。
- 對您的遷移策略進行了變更，將規劃的現代化數量從 25% 降低到 10%。這對營運組織具有下游影響。他們沒有被告知這一策略變化，因此，他們還沒有準備好足夠的技術能力來支援更多的工作負載平移到 AWS 中。

建立此最佳實務的優勢：

- 您的組織充分了解新的或變更的策略，他們會以強烈的動力採取相應的行動，以幫助彼此實現領導層設定的總體目標和指標。
- 存在的機制可用來及時通知團隊成員已知的風險和計劃的事件。
- 組織會更有效地採用新的工作方式 (包括對人員或組織、程序或技術的變更) 以及所需技能，而且您的組織可更快速地實現企業利益。
- 團隊成員了解溝通事項，他們可以更有效地完成工作。

未建立此最佳實務時的曝險等級：高

實作指引

若要實作此最佳實務，您必須與組織中的利益相關者合作，以同意溝通標準。在組織內將這些標準公告週知。對於任何重大的 IT 轉型，與忽略此實務的組織相比，已確立的規劃團隊可以更成功地管理變更對其人員的影響。大型組織在管理變更時更具挑戰性，因為與所有個別貢獻者一起建立對新策略的強有力支援至關重要。如果沒有此類轉型規劃團隊，領導層 100% 負責有效溝通。建立轉型規劃團隊時，指派團隊成員與所有組織領導層合作，以定義和管理各個層面的有效溝通。

客戶範例

AnyCompany Retail 已註冊 AWS Enterprise Support，並依賴其他第三方提供商進行其雲端運營。該公司透過聊天和 chatops 作為其運營活動的主要溝通媒介。特定管道會填入提醒和其他資訊。人們必須展開行動時，他們會明確說明預期成果，且在許多情況下他們會接收執行手冊或程序手冊以供使用。他們可使用變更行事曆來排程生產系統的重大變更。

實作步驟

1. 在組織內建立一個核心團隊，負責為組織內多個層面發生的變更制定和啟動溝通計畫。
2. 建立單線程擁有權以實現監督。賦予各個團隊獨立創新的能力，並平衡使用一致的機制，從而實現適當的檢查水平和方向性願景。
3. 與組織中的利益相關者合作，以同意溝通標準、實務和計劃。
4. 確認核心溝通團隊是否與組織和計劃領導層協作，代表領導者為適當的員工製作訊息。
5. 建立策略性溝通機制，透過公告、共用行事曆、全體會議，以及面對面或一對一的方法來管理變更，讓團隊成員對應採取的行動有適當的期望。
6. 提供必要的背景知識、詳細資訊和時間 (如果可能的話)，以確定是否需要採取行動。當需要採取行動時，請提供所需的動作及其影響。
7. 實作可促進戰術溝通的工具，例如內部聊天、電子郵件和知識管理。

8. 實施機制以衡量和驗證所有溝通是否都能達到預期成果。
9. 建立一種回饋循環，用以衡量所有溝通的有效性，尤其是與整個組織中抗拒改變的力量相關的溝通。
10. 對於所有 AWS 帳戶，建立帳單、安全性和操作的[替代聯絡人](#)。理想情況下，每個聯絡人應該是電子郵件分發，而不是特定的個人聯絡人。
11. 建立呈報和逆向呈報溝通計畫，以與您的內部和外部團隊 (包括 AWS Support 和其他第三方供應商) 互動。
12. 在每個轉型計劃的生命週期中始終如一地啟動和執行溝通策略。
13. 排定可重複動作的優先順序，以便大規模安全地自動化。
14. 在具有自動化動作的情況下進行通訊時，通訊目的應該是通知團隊、進行稽核或作為變更管理流程的一部分。
15. 分析來自提醒系統的通訊，找出不斷建立的誤報或提醒。移除或變更這些提醒，使其僅在需要人為介入時啟動。如果啟動了提醒，請提供執行手冊或程序手冊。
 - a. 您可以使用 [AWS Systems Manager 文件](#)，為提醒建立程序手冊和執行手冊。
16. 已設立機制，以清楚且可行的方式提供風險或計劃事件的通知，並提供足夠的通知，以便適當的回應。使用電子郵件清單或聊天管道，在計劃性事件發之前傳送通知。
 - a. [AWS Chatbot](#) 可用於在組織訊息傳遞平台中傳送提醒並回應事件。
17. 提供可存取的資訊來源，您可以在其中發現計劃的事件。提供來自相同系統之計劃事件的通知。
 - a. [AWS Systems Manager 變更行事曆](#) 可用於在發生變更時建立變更時段。這可為團隊成員提供有關於何時可安全進行變更的通知。
18. 監控漏洞通知和修補程式資訊，了解外部漏洞以及與工作負載元件相關的潛在風險。提供通知給團隊成員，以便讓他們可以採取動作。
 - a. 您可以訂閱 [AWS 安全公告](#)，以接收 AWS 上的漏洞通知。
19. 尋求不同的意見和觀點：鼓勵每個人的貢獻。為代表人數不夠的群體提供溝通機會。在會議中輪換角色和職責。
 - a. 詳細闡述角色和職責：為團隊成員提供機會，讓他們承擔他們可能不會承擔的角色。他們會透過角色，以及與他們可能不會與之互動的新團隊成員互動，而獲得經驗和觀點。他們還將自己的經驗和觀點帶到新的角色，並帶給和他們互動的團隊成員。隨著觀點增加，確定新興的業務機會或者新的改進機會。在團隊成員之間輪流處理其他人通常執行的常見任務，以了解執行這些任務的需求和影響。
 - b. 提供安全且友善的環境：制定政策和控制措施，以保護組織內團隊成員的心理和身體安全。團隊成員應該能夠在不擔心報復行為的情況下進行互動。當團隊成員感到安全且受歡迎時，他們才更有可能參與進來並具備生產力。您的組織越多樣化，您就越能了解所支援的人員，包括您的客

戶。當您的團隊成員感到安心、可以自在的暢所欲言，而且有信心他們的聲音不會被淹沒，他們才更有可能分享寶貴的洞見 (例如，行銷機會、可及性的需求、尚未提供服務的市場區隔，以及環境中未確認的風險)。

- c. 鼓勵團隊成員充分參與：提供員工充分參與所有與工作相關的活動所需的資源。面對日常挑戰的團隊成員會發展出解決挑戰的技能。這些以獨特方式發展的技能可為組織提供顯著的效益。為團隊成員提供必要的便利性支援，將可從他們的貢獻中獲得更高的效益。

資源

相關的最佳實務：

- [OPS03-BP01 提供高層支持](#)
- [OPS07-BP03 使用執行手冊執执行程序](#)
- [OPS07-BP04 使用程序手冊調查問題](#)

相關文件：

- [AWS 部落格文章 | 問責制和授權是高效能敏捷組織的關鍵](#)
- [AWS Executive Insights | 學習擴展創新，而不是複雜性 | 單線程領導者](#)
- [AWS 安全公告](#)
- [開放式 CVE](#)
- [Slack 中用於管理支援案例的 支援 應用程式](#)
- [使用聊天應用程式中的 Amazon Q Developer 來管理 Slack 頻道中的 AWS 資源](#)

相關服務：

- [聊天應用程式中的 Amazon Q Developer](#)
- [AWS Systems Manager 變更行事曆](#)
- [AWS Systems Manager 文件](#)

OPS03-BP05 鼓勵進行試驗

試驗是將新構想轉化為產品和功能的觸媒。試驗可加速學習，讓團隊成員保持興趣和參與度。我們鼓勵團隊成員經常進行試驗以推動創新。即便結果不如預期仍有其價值，至少我們了解到什麼是不該做的。團隊成員不會因取得不理想結果的成功試驗而受懲罰。

預期成果：

- 您的組織鼓勵試驗以促進創新。
- 試驗被視為一種學習機會。

常見的反模式：

- 您想要執行 A/B 測試，但沒有相關機制可執行試驗。您在沒有測試能力的情況下部署了 UI 變更。其結果導致了負面客戶體驗。
- 您的公司只有模擬和生產環境。沒有沙盒環境可用來試驗新功能或產品，因此您必須在生產環境內試驗。

建立此最佳實務的優勢：

- 試驗可帶動創新。
- 透過試驗，您可以更快回應使用者的意見反映。
- 組織可培養學習文化。

未建立此最佳實務時的曝險等級：中

實作指引

試驗應以安全的方式執行。利用多種環境進行試驗，而不會損害生產資源。使用 A/B 測試和功能旗標來測試試驗。為團隊成員提供在沙盒環境中執行試驗的能力。

客戶範例

AnyCompany Retail 鼓勵試驗。團隊成員可將其 20% 的工時投入於試驗或學習新技術。他們有沙盒環境可供創新之用。他們可對新功能進行 A/B 測試，用實際使用者的意見反映加以驗證。

實作步驟

1. 與組織中的領導階層共同推行試驗風氣。應鼓勵團隊成員以安全的方式執行試驗。
2. 為團隊成員提供可安全進行試驗的環境。他們必須能夠存取類似生產環境的環境。
 - a. 可以使用單獨的 AWS 帳戶來建立用於實驗的沙盒環境。[AWS Control Tower](#) 可用於佈建這些帳戶。
3. 使用功能旗標和 A/B 測試安全地進行試驗，並收集使用者的意見反映。

- a. [AWS AppConfig Feature Flags](#) 能夠建立功能旗標。
- b. 可以使用 [AWS Lambda 版本](#) 部署新版本的函數以進行 beta 測試。

實作計劃的工作量：高。為團隊成員提供可安全執行試驗的環境，可能需要可觀的投資。為了使用功能旗標或支援 A/B 測試，您可能需要修改應用程式程式碼。

資源

相關的最佳實務：

- [OPS11-BP02 執行事後分析](#) - 從事件中學習是創新以及實驗的重要驅動力。
- [OPS11-BP03 實作意見回饋循環](#) - 回饋迴圈是實驗的重要組成部分。

相關文件：

- [Amazon 文化內貌：實驗、失敗和客戶至上](#)
- [在 AWS 中建立和管理沙盒帳戶的最佳實務](#)
- [建立由雲端啟用的實驗文化](#)
- [在 SulAmérica Seguros 實現雲端實驗和創新](#)
- [實驗越多，失敗越少](#)
- [使用多個帳戶來組織您的 AWS 環境 - 沙盒 OU](#)
- [使用 AWS AppConfig 功能旗標](#)

相關影片：

- [AWS On Air ft. Amazon CloudWatch Evidently | AWS 活動](#)
- [AWS On Air San Fran Summit 2022 ft. AWS AppConfig 功能旗標與 Jira 整合](#)
- [AWS re:Invent 2022 - 部署不是發行版本：透過功能旗標控制您的啟動 \(BOA305-R\)](#)
- [使用 AWS Control Tower 以程式設計方式建立 AWS 帳戶](#)
- [設定使用 AWS Organizations 最佳實務的多帳戶 AWS 環境](#)

相關範例：

- [AWS 創新沙盒](#)

- [電子商務端對端個人化 101](#)

相關服務：

- [Amazon CloudWatch Evidently](#)
- [AWS AppConfig](#)
- [AWS Control Tower](#)

OPS03-BP06 管理團隊成員維持和發展自己的技能集

團隊必須發展自己的技能集，以採用新技術，並支援需求和責任的變更，以支援您的工作負載。新技術的技能成長通常是團隊成員滿意度的來源，並可支援創新。支援團隊成員追求和維持產業認證，以驗證和認可他們不斷成長的技能。交叉培訓以促進知識轉移，並在失去熟練的、經驗豐富且具備機構知識的成員時，降低重大影響的風險。提供學習專用的結構化時間。

AWS 可提供資源，包括 [AWS 入門資源中心](#)、[AWS 部落格](#)、[AWS 線上技術講座](#)、[AWS 活動和網路研討會](#) 以及 [AWS Well-Architected 實驗室](#)，提供指引、範例和詳細的逐步解說來教育您的團隊。

諸如 [支援](#) ([AWS re:Post](#)、[支援 Center](#)) 等資源和 [AWS 文件](#) 有助於消除技術障礙並改善操作。透過 [支援中心](#) 聯絡支援，以獲取相關問題的幫助。

AWS 還分享了我們透過在 [Amazon 建置者資料中心](#) 操作 AWS 所學到的最佳實務和模式，以及透過 [AWS 部落格](#) 和 [官方 AWS 部落格](#) 提供的各種其他有用的教育材料。

[AWS 培訓和認證](#) 透過自定進度的數位課程提供免費培訓，以及按角色或領域制定的學習計劃。您還可以報名參加講師指導下的培訓，以進一步協助開發團隊的 AWS 技能。

預期成果：您的組織不斷評估技能差距，並透過結構化預算和投資進行彌補。團隊透過提高技能的活動來鼓勵和激勵其會員，例如獲得領先的行業認證。團隊利用專門的交叉分享知識計劃，例如午餐學習會、Immersion Day、黑客松和演練日。您的組織將其知識系統保持在最新狀態，並與交叉培訓團隊成員相關，包括新員工入職培訓。

常見的反模式：

- 在沒有結構化培訓計劃和預算的情況下，團隊會遇到不確定性，因為他們試圖跟上技術發展的步伐，從而增加損耗。
- 作為遷移到 AWS 的一部分，您的組織展現了團隊間的技能差距和不同的雲端流暢度。如果不努力提高技能，團隊會發現自己被雲端環境的傳統和效率低下的管理所累，這會導致操作員的工作量增加。這種倦怠會加劇員工的不滿。

建立此最佳實務的優勢：當您的組織有意識地投資於改善其團隊的技能時，它也有助於加速和擴展雲端採用和最佳化。針對性的學習計劃可推動創新並建立營運能力，讓團隊為處理各種事件做好準備。團隊有意識地投資於最佳實務的實作和發展。團隊士氣高漲，團隊成員重視自己對業務的貢獻。

未建立此最佳實務時的曝險等級：中

實作指引

為了採用新技術、推動創新、並跟上需求和責任變化的步伐來支援您的工作負載，請不斷投資於團隊的專業成長。

實作步驟

1. 使用結構化雲端宣傳計劃：[AWS Skills Guild](#) 提供諮詢培訓，以提高雲端技能信心並培養持續學習的文化。
2. 提供教育資源：提供專門的結構化時間，以及取用培訓教材和實驗室資源的權限，並支援參與會議和專業組織，從中為教育工作者和同儕提供學習機會。為資淺團隊成員提供接近資深團隊成員的機會，讓資深團隊成員成為導師，或允許資淺團隊成員參觀資深團隊成員的工作，並接觸他們的方法和技能。鼓勵學習與工作不直接相關的內容，以便取得更廣泛的視野。
3. 鼓勵使用專家技術資源：利用諸如 [AWS re:Post](#) 之類的資源來存取精選知識和充滿活力的社群。
4. 建置並維護最新的知識儲存庫：使用知識共享平台，例如 Wiki 和執行手冊。使用 [AWS re:Post Private](#) 建立您自己的可重複使用的專家知識來源，以簡化協同合作、提高生產力並加速員工入職。
5. 團隊教育和跨團隊參與：規劃團隊成員的持續教育需求。為團隊成員提供機會 (暫時或永久地) 加入其他團隊，以分享讓整個組織受益的技能和最佳實務。
6. 支援追求和維持產業認證：支援團隊成員取得與維持可驗證所學知識並認可其成就的產業認證。

實作計劃的工作量：高

資源

相關的最佳實務：

- [OPS03-BP01 提供高層支持](#)
- [OPS11-BP04 執行知識管理](#)

相關文件：

- [AWS 白皮書 | 雲端採用架構：個人視角](#)

- [投資持續學習以發展組織的未來](#)
- [AWS Skills Guild](#)
- [AWS 培訓 和認證](#)
- [支援](#)
- [AWS re:Post](#)
- [AWS 資源中心入門](#)
- [AWS 部落格](#)
- [AWS 雲端 合規](#)
- [AWS 文件](#)
- [官方 AWS 播客。](#)
- [AWS 線上技術講座](#)
- [AWS 活動和研討會](#)
- [AWS Well-Architected 實驗室](#)
- [Amazon 建置者資料中心](#)

相關影片：

- [AWS re:Invent 2023 | 以雲端的速度重塑技能：將員工變為企業家](#)
- [AWS re:Invent 2023 | 透過遊戲化建立好奇心文化](#)

OPS03-BP07 資源團隊適當

提供適量訓練有素的團隊成員，並提供工具和資源來支援工作負載需求。負擔過重的團隊成員會增加人為錯誤的風險。對工具和資源的投資 (例如自動化) 可以提升團隊的效率，並協助他們支援更多工作負載，而不需要額外的生產能力。

預期成果：

- 您已適當地配置團隊，以取得他們 AWS 根據您的遷移計劃操作工作負載所需的技能組合。由於您的團隊在遷移專案期間自我擴展，因此他們已精通業務計劃在遷移或現代化其應用程式時使用的核心 AWS 技術。
- 您已經仔細調整人員配置計畫，以利用自動化和工作流程來有效利用資源。較小的團隊現在可以代表應用程式開發團隊管理更多基礎設施。

- 隨著營運優先順序的轉變，會主動識別任何資源人員配置限制，以保護業務計畫的成功。
- 審核報告操作辛勞的操作指標 (例如待命疲勞或過度呼叫)，以驗證員工是否不堪重負。

常見的反模式：

- 當您接近多年雲端遷移計畫時，您的員工尚未提高 AWS 技能，這些風險支援工作負載並降低員工士氣。
- 您的整個 IT 組織正在轉向敏捷的工作方式。企業正優先考慮產品組合，並為需要首先開發的功能設定指標。敏捷流程不需要團隊將故事點分配給他們的工作計畫。因此，不可能了解下一工作量所需的能力水平，或者您是否擁有分配給該工作的正確技能。
- 您正在讓 AWS 合作夥伴遷移工作負載，而且一旦合作夥伴完成遷移專案，您就沒有團隊的支援轉換計畫。您的團隊難以高效地支援工作負載。

建立此最佳實務的優勢：您的組織中有適當技能的團隊成員來支援工作負載。資源配置可適應不斷變化的優先順序，而不會影響效能。結果是團隊精通工作負載的支援，同時最大限度地利用時間專注於為客戶創新，從而提高員工滿意度。

未建立此最佳實務時的曝險等級：中

實作指引

雲端遷移的資源規劃應該在組織層級進行，它符合遷移計畫以及為支援新雲端環境而實作的所需操作模型。這應該包括了解為業務和應用程式開發團隊部署哪些雲端技術。基礎設施和營運領導者應該為領導雲端採用的工程師規劃技能差距分析、培訓和角色定義。

實作步驟

1. 使用諸如員工生產力等相關營運指標 (例如，支援工作負載的成本或事件期間所花費的操作員時數)，定義團隊成功的標準。
2. 定義資源容量計畫與檢驗機制，以驗證合格容量的適當平衡在需要時可用，並且可隨時間進行調整。
3. 建立機制 (例如，向團隊傳送每月調查問卷)，以了解影響團隊的工作相關挑戰 (例如責任增加、技術變化、人員損失或支援的客戶增加)。
4. 使用這些機制與團隊互動，並發現可能導致員工生產力挑戰的趨勢。當您的團隊受到外部因素影響時，請重新評估目標並適當地調整目標。找出阻礙您團隊進度的障礙。
5. 定期檢閱目前佈建的資源是否仍然足夠，或是否需要額外資源，並做出適當的調整以支援團隊。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS我們鼓勵 03-BP06 團隊成員維護和提升其技能組合](#)
- [OPS09-BP03 檢閱操作指標並排定改善優先順序](#)
- [OPS10-BP01 使用事件、事件和問題管理的程序](#)
- [OPS10-BP07 自動化對事件的回應](#)

相關文件：

- [AWS 雲端 採用架構：人員觀點](#)
- [成為面向未來的企業](#)
- [優先考慮員工的技能以推動業務增長](#)
- [高效能組織 - Amazon 雙披薩團隊](#)
- [雲端成熟企業如何成功](#)

準備

問題

- [OPS 4. 如何在工作負載中實作可觀測性？](#)
- [OPS 5. 如何減少缺陷、幫助輕鬆修復，以及改善生產流程？](#)
- [OPS 6. 如何緩解部署風險？](#)
- [OPS 7. 如何知道自己準備好支援工作負載？](#)

OPS 4. 如何在工作負載中實作可觀測性？

在工作負載中實作可觀測性，以便了解其狀態，並根據業務需求做出資料驅動的決策。

最佳實務

- [OPS04-BP01 識別關鍵績效指標](#)
- [OPS04-BP02 實作應用程式遙測](#)

- [OPS04-BP03 實作使用者體驗遙測](#)
- [OPS04-BP04 實作相依性遙測](#)
- [OPS04-BP05 實作分散式追蹤](#)

OPS04-BP01 識別關鍵績效指標

想在工作負載中實作可觀測性，要先了解工作負載狀態，並根據業務需求做出資料驅動型決策。確保監控活動與業務目標保持一致的最有效方法之一，是透過定義和監控關鍵績效指標（KPIs）。

預期成果：有效率的、可觀測性實務會與業務目標密切保持一致，確保監控工作始終能夠帶來實際的業務成果。

常見的反模式：

- 未定義 KPIs：在沒有明確的情況下工作KPIs可能會導致監控太多或太少，遺失重要訊號。
- 靜態 KPIs：不會KPIs隨著工作負載或業務目標的發展而重新檢視或精簡。
- 未能保持一致：專注於與業務成果沒有直接關係的技術指標，或難與實際問題相關聯的技術指標。

建立此最佳實務的優勢：

- 問題識別的易用性：企業KPIs通常會比技術指標更清晰地呈現問題。企業的下降KPI可以比篩選多個技術指標更有效找出問題。
- 業務一致性：確保監控活動可直接支援業務目標。
- 效率：優先監控資源並關注重要指標。
- 主動積極：找出並解決問題，不讓問題擴大影響業務。

未建立此最佳實務時的曝險等級：高

實作指引

若要有效定義工作負載KPIs：

1. 從業務成果開始著手：在深入研究指標之前，請先了解預期業務成果。想要增加銷售量、提高使用者參與度，還是加快回應時間？
2. 將技術指標與業務目標相關聯：並非所有技術指標都會直接影響業務成果。識別這樣做的人，但使用業務識別問題通常更為簡單KPI。
3. 使用 [Amazon CloudWatch](#)：使用 CloudWatch 來定義和監控代表您的指標KPIs。

4. 定期檢閱和更新 KPIs：隨著工作負載和業務的發展，請保持KPIs關聯。
5. 讓利益相關者參與：讓技術和業務團隊參與定義和檢閱 KPIs。

實作計劃的工作量：中

資源

相關的最佳實務：

- [the section called “OPS04-BP02 實作應用程式遙測”](#)
- [the section called “OPS04-BP03 實作使用者體驗遙測”](#)
- [the section called “OPS04-BP04 實作相依性遙測”](#)
- [the section called “OPS04-BP05 實作分散式追蹤”](#)

相關文件：

- [AWS 可觀測性最佳實務](#)
- [CloudWatch 使用者指南](#)
- [AWS 可觀測性技能建置器課程](#)

相關影片：

- [研擬可觀測性策略](#)

相關範例：

- [一個可觀測性研討會](#)

OPS04-BP02 實作應用程式遙測

應用程式遙測是工作負載可觀測性的基礎。發出遙測至關重要，它為您的應用程式狀態以及技術和業務成果的實現提供了可行洞見。從疑難排解到測量新功能的影響，或確保與業務金鑰效能指標（KPIs）保持一致，應用程式遙測會告知您建置、操作和發展工作負載的方式。

指標、日誌和追蹤構成了可觀測性的三個主要支柱。其可作為描述應用程式狀態的診斷工具。隨著時間的推移，它們有助於建立基準並識別異常。不過，為了確保監控活動與業務目標之間的一致性，定義和監控至關重要KPIs。與技術指標相比，企業KPIs通常更容易識別問題。

其他遙測類型，例如真實使用者監控（RUM）和合成交易，可補充這些主要資料來源。RUM 提供即時使用者互動的洞見，而合成交易模擬潛在的使用者行為，有助於在實際使用者遇到瓶頸之前進行偵測。

預期成果：獲得工作負載效能且可付諸行動的洞見。這些洞見可讓您做出有關效能最佳化的主動決策、提高工作負載穩定性、使 CI/CD 程序更順暢，並且有效利用資源。

常見的反模式：

- 不完整的可觀測性：忽略在工作負載的每一層納入可觀測性，導致出現可能遮蔽重要系統效能和行為洞見的盲點。
- 分散的資料檢視：當資料分散在多個工具和系統中時，便難以提供涵蓋工作負載運作狀況和效能的全面概覽。
- 使用者報告問題：缺乏透過遙測和業務KPI監控主動偵測問題的跡象。

建立此最佳實務的優勢：

- 知情決策：透過遙測和業務的洞察KPIs，您可以做出資料驅動的決策。
- 改善運作效率：資料驅動的資源利用率可帶來成本效益。
- 提高工作負載穩定性：更快偵測並解決問題，進而改善正常運作。
- 更順暢的 CI/CD 程序：從遙測資料獲得的洞見，有助於改進程序並交付可靠的程式碼。

未建立此最佳實務時的曝險等級：高

實作指引

若要為工作負載實作應用程式遙測，請使用 AWS [Amazon CloudWatch](#) 和 等服務[AWS X-Ray](#)。Amazon CloudWatch 提供全方位的監控工具套件，可讓您在內部部署環境中觀察資源 AWS 和應用程式。還會收集、追蹤和分析指標、合併和監控日誌資料，並且回應資源的變更，以增進您對工作負載運作方式的了解。在串聯中，AWS X-Ray 可讓您追蹤、分析和偵錯應用程式，讓您深入了解工作負載的行為。透過服務地圖、延遲分佈和追蹤時間表等功能，AWS X-Ray 可讓您深入了解工作負載的效能和影響工作負載的瓶頸。

實作步驟

1. 確定要收集的資料：確定可提供工作負載運作狀況、效能和行為實質洞見的重要指標、日誌和追蹤。

2. 部署[CloudWatch代理程式](#)：代理 CloudWatch 程式對於從您的工作負載及其基礎基礎設施中取得系統和應用程式指標和日誌至關重要。CloudWatch 代理程式也可以用來收集 OpenTelemetry 或 X-Ray 追蹤，並將其傳送至 X-Ray。
3. 實作日誌和指標的異常偵測：使用[CloudWatch 日誌異常偵測](#)和[CloudWatch指標異常偵測](#)，自動識別應用程式操作中的異常活動。這些工具使用機器學習演算法來偵測異常並發出提醒，進而提升監控能力，並加快對潛在中斷或安全威脅的回應時間。設定這些功能以主動管理應用程式運作狀態和安全性。
4. 安全敏感日誌資料：使用 [Amazon CloudWatch Logs 資料保護](#)來遮罩日誌中的敏感資訊。此功能可在存取敏感資料前進行自動偵測和遮罩，從而有助於維護隱私權與合規性。實作資料遮罩，以安全地處理和保護敏感詳細資訊，例如個人識別資訊（PII）。
5. 定義和監控業務 KPIs：建立與您的[業務成果相符的自訂指標](#)。
6. 使用 來測試您的應用程式 AWS X-Ray：除了部署 CloudWatch代理程式之外，[測試應用程式](#)以發出追蹤資料也很重要。此程序可提供工作負載行為和效能的進一步洞見。
7. 將整個應用程式的資料收集標準化：將整個應用程式的資料收集實務標準化。採取一致的方式有助於找出資料關聯並進行分析，進而提供應用程式行為的全面概覽。
8. 實作跨帳戶可觀測性：AWS 帳戶 使用 [Amazon CloudWatch 跨帳戶可觀測性](#)增強跨多個 的監控效率。使用此功能，您可以將不同帳戶的指標、日誌和警示合併為單一檢視，可簡化管理並改善組織 AWS 環境中已識別問題的回應時間。
9. 分析資料並採取行動：資料收集和標準化完成後，請使用 [Amazon CloudWatch](#) 進行指標和日誌分析，以及[AWS X-Ray](#)追蹤分析。這類分析可產生有關工作負載運作狀況、效能和行為的洞見，進而引導您進行決策。

實作計劃的工作量：高

資源

相關的最佳實務：

- [OPS04-BP01 定義工作負載 KPIs](#)
- [OPS04-BP03 實作使用者活動遙測](#)
- [OPS04-BP04 實作相依性遙測](#)
- [OPS04-BP05 實作交易可追蹤性](#)

相關文件：

- [AWS 可觀測性最佳實務](#)
- [CloudWatch使用者指南](#)
- [AWS X-Ray 開發人員指南](#)
- [檢測分散式系統，以了解運作狀態](#)
- [AWS 可觀測性 Skill Builder 課程](#)
- [Amazon 的新功能 CloudWatch](#)
- [新功能 AWS X-Ray](#)

相關影片：

- [AWS re : Invent 2022 - Amazon 的可觀測性最佳實務](#)
- [AWS re : Invent 2022 - 開發可觀測性策略](#)

相關範例：

- [一個可觀測性研討會](#)
- [AWS 解決方案庫：使用 Amazon 進行應用程式監控 CloudWatch](#)

OPS04-BP03 實作使用者體驗遙測

深入了解客戶體驗以及與應用程式的互動情形非常重要。實際使用者監控（RUM）和合成交易可做為此目的強大工具。RUM 提供有關真實使用者互動的資料，授予使用者滿意度的未篩選觀點，而合成交易模擬使用者互動，即使在影響真實使用者之前，也有助於偵測潛在問題。

預期成果：提供使用者體驗、主動偵測問題及最佳化使用者互動的整體概觀，從而獲得順暢的數位體驗。

常見的反模式：

- 沒有實際使用者監控的應用程式（RUM）：
 - 延遲問題偵測：如果沒有 RUM，在使用者投訴之前，您可能不會發現效能瓶頸或問題。這種被動回應的方式可能導致客戶不滿意。
 - 缺乏使用者體驗洞察：不使用 RUM 表示您遺失了關鍵資料，這些資料顯示真實使用者如何與您的應用程式互動，限制了您最佳化使用者體驗的能力。
- 沒有綜合交易的應用程式：

- 缺少邊緣案例：綜合交易可協助您測試一般使用者可能不常使用，但對於某些業務功能來說相當關鍵的路徑和功能。缺少的話，這些路徑可能無法正常運作並遭到忽視。
- 在應用程式未使用的情況下檢查問題：定期綜合測試可模擬實際使用者未積極與您的應用程式互動的情況，進而確保系統隨時正常運作。

建立此最佳實務的優勢：

- 主動偵測問題：找出並解決潛在問題，避免進一步影響實際使用者。
- 最佳化使用者體驗：RUM協助精簡和增強整體使用者體驗的持續意見回饋。
- 裝置和瀏覽器效能的相關洞見：了解您的應用程式在不同裝置和瀏覽器上的效能表現，以便進一步最佳化。
- 經驗證的業務工作流程：定期綜合交易可確保核心功能和重要路徑維持正常且高效率的運作。
- 增強應用程式效能：利用收集自實際使用者資料的洞見來改善應用程式回應能力和可靠性。

未建立此最佳實務時的曝險等級：高

實作指引

為了利用 RUM和 合成交易進行使用者活動遙測，AWS 提供 [Amazon CloudWatch RUM](#) 和 [Amazon CloudWatch Synthetics](#) 等服務。指標、日誌和追蹤搭配使用者活動資料，可提供深入應用程式運作狀態和使用者體驗的全方位檢視。

實作步驟

1. 部署 Amazon CloudWatch RUM：將應用程式與 CloudWatch RUM 整合，以收集、分析和呈現實際使用者資料。
 - a. 使用 [CloudWatch RUM JavaScript 程式庫](#) RUM與您的應用程式整合。
 - b. 設定儀表板以視覺化和監控實際使用者資料。
2. 設定 CloudWatch 合成：建立 Canary 或指令碼常式，以模擬使用者與應用程式的互動。
 - a. 定義關鍵應用程式工作流程和路徑。
 - b. 使用 [CloudWatch Synthetics 指令碼](#) 設計 Canary，以模擬這些路徑的使用者互動。
 - c. 排定依指定間隔執行 Canary 並進行監控，確保一致的效能檢查。
3. 分析資料並對資料採取行動：利用 RUM和 合成交易中的資料來取得洞察，並在偵測到異常時採取修正措施。使用 CloudWatch 儀表板和警示來隨時掌握最新資訊。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS04-BP01 識別關鍵績效指標](#)
- [OPS04-BP02 實作應用程式遙測](#)
- [OPS04-BP04 實作相依性遙測](#)
- [OPS04-BP05 實作分散式追蹤](#)

相關文件：

- [Amazon CloudWatch RUM 指南](#)
- [Amazon CloudWatch Synthetics 指南](#)

相關影片：

- [使用 Amazon 透過最終使用者洞察最佳化應用程式 CloudWatch RUM](#)
- [AWS Air ft 上的 。 Amazon 的真實使用者監控 CloudWatch](#)

相關範例：

- [一個可觀測性研討會](#)
- [Amazon CloudWatch RUM Web Client 的 Git 儲存庫](#)
- [使用 Amazon CloudWatch Synthetics 來測量頁面載入時間](#)

OPS04-BP04 實作相依性遙測

對於監控工作負載所依賴的外部服務和元件運作狀況與效能，相依性遙測至關重要，可提供連線能力、逾時，以及像是 DNS、資料庫或第三方 API 等其他與相依性相關重要事件的寶貴洞見。檢測應用程式以產生有關這些相依性的指標、日誌和追蹤，便可更清楚了解可能影響工作負載的潛在瓶頸、效能問題或故障。

預期成果：確保工作負載所依賴的相依性如預期般正常運作，讓您能夠主動解決問題並確保最佳的工作負載效能。

常見的反模式：

- 忽略外部相依性：僅關注內部應用程式指標，而忽略與外部相依性相關的指標。
- 缺乏主動監控：等待問題出現，而非持續監控相依性的運作狀況與效能。
- 單獨運作的監控：使用多種分散的監控工具，如此可能導致僅片段掌握相依性運作狀況且獲得的資訊不一致。

建立此最佳實務的優勢：

- 改善工作負載可靠性：確保外部相依性穩定運作並保持最佳效能。
- 更快偵測並解決問題：主動找出並解決相依性相關問題，不讓問題影響工作負載。
- 全方位視角：獲得全方位視角，有效掌握影響工作負載運作狀況的內部和外部元件。
- 增強工作負載可擴展性：了解外部相依性的可擴展性限制與效能特性。

未建立此最佳實務時的曝險等級：高

實作指引

從識別您的工作負載所依賴的服務、基礎設施和程序開始，實作相依性遙測。將相依性正常運作時的良好條件量化，然後判斷衡量時所需的資料。有了這些資訊，您就可以打造儀表板並設定警示，以便為營運團隊提供這些相依性狀態的洞見。相依性無法按需求運作時，使用 AWS 工具探索並量化其影響。持續重新檢視您的策略，以考量優先順序、目標和獲得的洞見的變化。

實作步驟

若要有效實作相依性遙測：

1. 識別外部相依性：與利益相關者協作，共同找出工作負載所依賴的外部相依性。外部相依性可能包含各種服務，像是外部資料庫、第三方 API、前往其他環境的網路連線能力路由，以及 DNS 服務。實現有效相依性遙測的第一步，就是徹底了解這些相依性。
2. 擬訂監控策略：清楚了解外部相依性之後，就可以為其量身打造監控策略。這包括了解每一項相依性的重要性、預期行為，以及任何相關的服務層級協議或目標 (SLA 或 SLT)。設定主動警示，以便在發生狀態變更或效能偏差時通知您。
3. 使用[網路監控](#)：使用[網際網路監控](#)和[網路監控](#)，全面了解全球網際網路和網路狀況。這些工具可協助您了解並回應影響外部相依性的停機、中斷或效能降低。

4. 利用 [AWS Health](#) 隨時掌握新知：AWS Health 是 AWS 雲端 資源運作狀態的權威資訊來源。使用 AWS Health 視覺化並接收有關任何目前服務事件和近期變更的通知 (例如規劃的生命週期事件)，如此您就能採取行動來緩解衝擊。
 - a. 透過 [AWS 使用者通知](#) [建立符合用途的 AWS Health 事件通知](#)，以利用電子郵件和聊天管道傳送，並[透過 Amazon EventBridge 或 AWS Health API 以程式設計方式與您的監控和警示工具整合](#)。
 - b. 透過 Amazon EventBridge 或 AWS Health API 整合變更管理或您可能已在使用的 ITSM 工具 (如 [Jira](#) 或 [ServiceNow](#))，以規劃並追蹤需要採取行動的運作狀態事件進度。
 - c. 如果您使用 AWS Organizations，請啟用 [AWS Health 的組織檢視](#)，以彙總帳戶之間的 AWS Health 事件。
5. 使用 [AWS X-Ray](#) 檢測您的應用程式：AWS X-Ray 提供了深入了解應用程式及其基礎相依性運作效能的洞見。透過從頭到尾追蹤請求，您就可以找出應用程式所依賴的外部服務或元件的瓶頸或故障。
6. 使用 [Amazon DevOps Guru](#)：這項機器學習驅動的服務可識別操作問題，預測重大問題可能在什麼時候發生，並且建議可採取的特定行動。對於獲得相依性洞見並確保其不是造成操作問題的根源來說，這項服務非常寶貴。
7. 定期監控：持續監控與外部相依性相關的指標和日誌。針對非預期的行為或效能降低的情況設定警示。
8. 變更後驗證：每當有任何外部相依性更新或變更，便驗證其效能並檢查是否符合您的應用程式需求。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS04-BP01 定義工作負載 KPI](#)
- [OPS04-BP02 實作應用程式遙測](#)
- [OPS04-BP03 實作使用者活動遙測](#)
- [OPS04-BP05 實作交易可追溯性](#)
- [OP08-BP04 建立可執行的提醒](#)

相關文件：

- [Amazon Personal Health 儀板表 使用者指南](#)
- [AWS 網際網路監控使用者指南](#)
- 《AWS X-Ray 開發人員指南》<https://docs.aws.amazon.com/xray/latest/devguide/aws-xray.html>
- [AWS DevOps Guru 使用者指南](#)

相關影片：

- [深入了解影響應用程式效能的網際網路問題](#)
- [Amazon DevOps Guru 簡介](#)
- [使用 AWS Health 大規模管理資源生命週期事件](#)

相關範例：

- [AWS Health Aware](#)
- [使用以標籤為基礎的篩選來大規模管理 AWS Health 監控和提醒](#)

OPS04-BP05 實作分散式追蹤

分散式追蹤可讓您監控和以視覺化的方式了解，在分散式系統中各種來回移動元件的請求。透過從多個來源擷取追蹤資料並在統一的檢視中進行分析，團隊就能更了解請求的流程、瓶頸出現的位置，以及最佳化工作應著重的地方。

預期成果：提供分散式系統請求流程的全面概覽，實現精確偵錯、最佳化效能，並改善使用者體驗。

常見的反模式：

- 不一致的檢測：並非所有分散式系統中的服務都經過檢測可進行追蹤。
- 忽略延遲：僅專注於錯誤，而未考慮延遲或效能逐漸降低的現象。

建立此最佳實務的優勢：

- 全方位的系統概觀：從進入到退出，徹底視覺化整個請求路徑。
- 強化偵錯：快速識別失敗或效能問題發生的位置。
- 改善使用者體驗：根據實際使用者資料進行監控與最佳化，確保系統符合實際需求。

未建立此最佳實務時的曝險等級：高

實作指引

首先，識別工作負載中需要檢測的所有元素。計算所有元件後，請利用 AWS X-Ray 和 等工具 OpenTelemetry 來收集追蹤資料，以便使用 X-Ray 和 Amazon CloudWatch ServiceLens Map 等工具進行分析。與開發人員進行定期檢閱，並使用 Amazon DevOpsGuru、X-Ray Analytics 和 X-Ray Insights 等工具來補充這些討論，以協助探索更深入的調查結果。從追蹤資料建立警示，以便在工作負載監視計畫中定義的結果存在風險時發出通知。

實作步驟

若要有效實作分散式追蹤：

1. 採用 [AWS X-Ray](#)：將 X-Ray 整合到您的應用程式中，以獲得深入其行為的洞見、了解效能，並且找出瓶頸的確切位置。利用 X-Ray Insights 進行自動化追蹤分析。
2. 測試您的服務：確認從 [AWS Lambda](#) 函數到 [EC2 執行個體](#) 的每個服務都會傳送追蹤資料。您測試的服務越多，檢視越清晰 end-to-end。
3. 整合 [CloudWatch 實際使用者監控](#) 和 [合成監控](#)：將實際使用者監控（RUM）和合成監控與 X-Ray 整合。這樣就能擷取實際使用者體驗並模擬使用者互動，以從中找出潛在問題。
4. 使用 [CloudWatch 代理程式](#)：代理程式可以從 X-Ray 或 傳送追蹤 OpenTelemetry，增強所取得洞見的深度。
5. 使用 [Amazon DevOpsGuru](#)：DevOpsGuru 使用 X-Ray 的資料 CloudWatch AWS Config，AWS CloudTrail 並提供可行的建議。
6. 分析追蹤：定期檢閱追蹤資料，以找出可能影響應用程式效能的模式、異常或瓶頸。
7. 設定警示：在 中設定警示 [CloudWatch](#) 是否有異常模式或延長延遲，允許主動解決問題。
8. 持續改善：隨著服務增加或修改重新檢視您的追蹤策略，以擷取所有相關資料點。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS04-BP01 識別關鍵績效指標](#)
- [OPS04-BP02 實作應用程式遙測](#)
- [OPS04-BP03 實作使用者體驗遙測](#)
- [OPS04-BP04 實作相依性遙測](#)

相關文件：

- [AWS X-Ray 開發人員指南](#)
- [Amazon CloudWatch 代理程式使用者指南](#)
- [Amazon DevOpsGuru 使用者指南](#)

相關影片：

- [使用 AWS X-Ray Insights](#)
- [AWS Air ft 上的 。 可觀測性：Amazon CloudWatch 和 AWS X-Ray](#)

相關範例：

- [測試您的應用程式 AWS X-Ray](#)

OPS 5. 如何減少缺陷、幫助輕鬆修復，以及改善生產流程？

採用改善改變生產流程的方法，藉此推動重構、快速提供品質意見回饋及修復錯誤。這些方法會加快有助益的改變發揮作用的速度、限制部署問題，並快速識別和修復部署活動造成的問題。

最佳實務

- [OPS05-BP01 使用版本控制](#)
- [OPS05-BP02 測試並驗證變更](#)
- [OPS05-BP03 使用組態管理系統](#)
- [OPS05-BP04 使用建置和部署管理系統](#)
- [OPS05-BP05 執行修補程式管理](#)
- [OPS05-BP06 共用設計標準](#)
- [OPS05-BP07 實作用於提高程式碼品質的實務](#)
- [OPS05-BP08 使用多個環境](#)
- [OPS05-BP09 進行頻繁、小型、可逆的變更](#)
- [OPS05-BP10 完全自動化整合和部署](#)

OPS05-BP01 使用版本控制

使用版本控制來追蹤變更和發布。

許多 AWS 服務都提供版本控制功能。使用修訂版或[原始程式碼控制](#)系統 (例如 [Git](#)) 來管理程式碼和其他成品，例如基礎結構的版本控制 [AWS CloudFormation](#) 範本。

期望的結果：您的團隊在程式碼上進行協作。合併後，程式碼會是一致的，且變更不會遺失。透過正確的版本控制就能輕鬆復原錯誤。

常見的反模式：

- 您已在工作站上開發和儲存程式碼。您的工作站發生無法復原的儲存錯誤，造成程式碼遺失。
- 變更覆寫現有的程式碼之後，您重新啟動應用程式卻無法運作。您無法還原變更。
- 您對其他人要編輯的報告檔案加上了寫入鎖定。他們會與您聯絡，要求您停止處理該檔案，以便完成任務。
- 您的研究團隊一直在進行詳細的分析，以規劃您未來的工作。某人意外地將自己的購物清單儲存在最終報告中。您無法還原變更，且必須重新建立報告。

建立此最佳實務的優勢：透過使用版本控制功能，您可以輕鬆還原為已知的良好狀態和舊版本，並有效降低資產遺失的風險。

未建立此最佳實務時的曝險等級：高

實作指引

在版本控制的儲存庫中維護資產。此舉可實現變更追蹤、新版本部署、對現有版本的變更偵測以及還原到先前的版本 (例如，在發生故障時復原到已知的良好狀態)。將組態管理系統的版本控制功能整合到您的程序中。

資源

相關的最佳實務：

- [OPS05-BP04 使用建置和部署管理系統](#)

相關影片：

- [AWS re:Invent 2023 - Lockheed Martin 如何採用 DevSecOps 技術更快建置軟體](#)
- [AWS re:Invent 2023 - GitHub 如何將 AI 營運化以應用到團隊協作和生產力](#)

OPS05-BP02 測試並驗證變更

所部署的每項變更都必須經過測試，以避免在生產環境中發生錯誤。此一最佳實務著重於各種變更 (從版本控制到成品組建) 的測試。除了應用程式碼變更之外，測試應包括基礎設施、組態、安全控制和操作程序。測試採取多種形式，從單元測試到軟體元件分析 (SCA) 都包括在內。將測試進一步納入軟體整合和交付程序中，可進一步確保成品的品質。

您的組織必須針對所有軟體成品制定測試標準。自動化測試可減少辛勞並避免手動測試錯誤。在某些情況下可能需要手動測試。開發人員必須有權存取自動化測試結果，以建立可改善軟體品質的回饋迴圈。

預期成果：您的軟體變更在交付前都經過測試。開發人員有權存取測試結果和驗證。您的組織具有適用於所有軟體變更的測試標準。

常見的反模式：

- 您在部署新軟體變更時未進行任何測試。它無法在生產環境中執行，這會導致中斷。
- 新的安全群組透過 AWS CloudFormation 進行部署，而未在生產前環境中測試。安全群組會讓您的客戶無法存取您的應用程式。
- 方法被修改，但沒有單元測試。軟體部署至生產環境時失敗。

建立此最佳實務的優勢：降低了軟體部署的變更失敗率。軟體品質獲得改善。開發人員更能感知其程式碼的可行性。可以安心推出安全政策，以支援組織的合規性。基礎設施變更 (例如自動化擴展政策更新) 會事先經過測試，以符合流量需求。

未建立此最佳實務時的曝險等級：高

實作指引

作為持續整合實務的一部分，對從應用程式碼到基礎設施的所有變更進行測試。發布測試結果，以便開發人員快速獲得意見回饋。您的組織具有所有變更都必須通過的測試標準。

透過 Amazon Q Developer，利用生成式 AI 的強大功能，提升開發人員生產力和程式碼品質。Amazon Q Developer 包括程式碼建議的產生 (以大型語言模型為基礎)、單元測試的生產 (包括邊界條件)，以及透過偵測和修復安全漏洞增強程式碼安全性。

客戶範例

在其持續整合管道中，AnyCompany Retail 對所有軟體成品執行了數種類型的測試。他們實行測試驅動型開發，所以所有軟體都有單元測試。建置成品後，它們會執行端到端測試。在第一輪測試完成後，

他們會執行靜態應用程式安全性掃描，以尋找已知的漏洞。通過每個測試門時，開發人員會收到訊息。所有測試都完成後，軟體成品即儲存在成品儲存庫中。

實作步驟

1. 與組織中的利益相關者合作制定軟體成品的測試標準。所有成品都應該通過哪些標準測試？測試涵蓋範圍中是否必須包含合規性或管控要求？您是否需要進程式碼品質測試？測試完成時，誰需要得知？
 1. [AWS 部署管道參考架構](#) 包含可在整合管道中對軟體成品執行之測試類型的授權清單。
2. 根據您的軟體測試標準，以必要的測試檢測您的應用程式。每組測試應在十分鐘內完成。測試應執行為整合管道的一部分。
 - a. 使用 [Amazon Q Developer](#)，這是一種生成式 AI 工具，可協助建立單元測試案例 (包括邊界條件)、使用程式碼和註解產生函數，以及實作眾所周知的演算法。
 - b. 使用 [Amazon CodeGuru Reviewer](#) 來測試您的應用程式碼是否存在故障。
 - c. 可使用 [AWS CodeBuild](#) 對軟體成品執行測試。
 - d. [AWS CodePipeline](#) 可將您的軟體測試安排到管道中。

資源

相關的最佳實務：

- [OPS05-BP01 使用版本控制](#)
- [OPS05-BP06 共用設計標準](#)
- [OPS05-BP07 實作用於提高程式碼品質的實務](#)
- [OPS05-BP10 完全自動化整合和部署](#)

相關文件：

- [採用測試驅動的開發方法](#)
- [使用 Amazon Q 加速您的軟體開發生命週期](#)
- [Amazon Q Developer 現已正式推出，包含可重新構想開發人員體驗的新功能預覽](#)
- [在 IDE 中使用 Amazon Q Developer 的終極速查表](#)
- [左移工作負載，利用 AI 建立測試](#)
- [Amazon Q 開發人員中心](#)

- [使用 Amazon CodeWhisperer 快速建置應用程式的 10 種方式](#)
- [使用 Amazon CodeWhisperer 超越程式碼覆蓋範圍](#)
- [使用 Amazon CodeWhisperer 進行提示詞工程的最佳實務](#)
- [使用 TaskCat 和 CodePipeline 的自動化 AWS CloudFormation 測試管道](#)
- [使用開放原始碼 SCA、SAST 和 DAST 工具建置端對端 AWS DevSecOps CI/CD 管道](#)
- [開始測試無伺服器應用程式](#)
- [CI/CD 管道是我的發行主管](#)
- [《在 AWS 上實行持續整合和持續交付》白皮書](#)

相關影片：

- [使用適用於軟體開發的 Amazon Q Developer 代理程式來實作 API](#)
- [使用 JetBrains IDE 安裝、設定和使用 Amazon Q Developer \(操作說明\)](#)
- [掌握 Amazon CodeWhisperer 的藝術 - YouTube 播放清單](#)
- [AWS re:Invent 2020：可測試的基礎設施：在 AWS 上進行整合測試](#)
- [AWS Summit ANZ 2021 - 透過 CDK 和測試驅動的開發施行測試優先策略](#)
- [使用 AWS CDK 測試基礎設施即程式碼](#)

相關資源：

- [AWS 部署管道參考架構 - 應用程式](#)
- [AWS Kubernetes DevSecOps 管道](#)
- [使用 AWS CodeBuild 為 GitHub 中的 Node.js 應用程式執行單元測試](#)
- [使用 Serverspec 進行基礎設施程式碼的測試驅動開發](#)

相關服務：

- [Amazon Q Developer](#)
- [Amazon CodeGuru Reviewer](#)
- [AWS CodeBuild](#)
- [AWS CodePipeline](#)

OPS05-BP03 使用組態管理系統

使用組態管理系統進行和追蹤組態變更。這些系統可減少由手動程序引起的錯誤，並減少部署變更的工作量。

靜態組態管理會在初始化資源時設定值，這些值預期會在資源的整個生命週期內保持一致。動態組態管理會在初始化時設定值，這些值可能或是預期會在資源的整個生命週期內保持一致。例如，您可以設定功能切換，透過組態變更啟動程式碼中的功能，或在事件期間變更日誌詳細資訊等級。

組態應以已知且一致的狀態部署。應該使用自動化檢測來持續監控跨環境和區域的資源組態。這些控制項應定義為已自動化的程式碼和管理，以確保規則在各個環境中一致套用。組態變更應透過商定的變更控制程序進行更新，並一致地套用，以遵守版本控制。應用程式組態的管理應該獨立於應用程式和基礎設施程式碼。這允許在多個環境中進行一致的部署。組態變更不會導致重建或重新部署應用程式。

預期成果：您會在持續整合、持續交付 (CI/CD) 管道中進行設定、驗證和部署。您會進行監控，以確認組態正確無誤。這會將終端使用者和客戶受到的任何負面影響降到最低。

常見的反模式：

- 您手動更新整個機群的 Web 伺服器組態，但由於更新錯誤，導致多部伺服器無法回應。
- 您在數小時內手動更新應用程式伺服器機群。變更期間的組態不一致會導致未預期的行為。
- 某人已更新您的安全群組，無法再存取您的 Web 伺服器。若不知道進行了哪些變更，您就需要花大量時間來調查問題，復原時間也會跟著拉長。
- 您可以透過 CI/CD 將生產前組態推送到生產環境中，而不需進行驗證。您讓使用者和客戶面臨使用不正確的資料和服務。

建立此最佳實務的優勢：採用組態管理系統可減少進行和追蹤變更的工作量，以及手動程序造成的錯誤頻率。組態管理系統提供了管控、合規和法規需求方面的保證。

未建立此最佳實務時的曝險等級：中

實作指引

組態管理系統可用來追蹤和實作應用程式與環境組態的變更。組態管理系統也可用來減少手動程序所造成的錯誤、讓組態變更可重複且可稽核，以及減少工作量。

在 AWS 上，可以使用 [AWS Config](#) 持續監控 [跨帳號和區域](#) 的 AWS 資源組態。它可協助您追蹤其組態歷史記錄、了解組態變更如何影響其他資源、以及針對預期或所需的組態進行稽核，方法是使用 [AWS Config 規則](#) 和 [AWS Config Conformance Packs](#)。

對於在 Amazon EC2 執行個體、AWS Lambda、容器、行動應用程式或 IoT 裝置上執行的應用程式中的動態組態，可以使用 [AWS AppConfig](#) 在整個環境中進行設定、驗證、部署和監控。

實作步驟

1. 確定組態擁有者。
 - a. 讓組態擁有者得知任何合規、管控或法規需求。
2. 確定組態項目與交付成果。
 - a. 組態項目是指受到 CI/CD 管線內部署影響的所有應用程式和環境組態。
 - b. 交付成果包括成功條件、驗證及監控對象。
3. 請根據您的業務需求和交付管道選取工具來進行組態管理。
4. 請考慮針對重大組態變更進行加權部署 (例如 Canary 部署)，以盡量減少錯誤組態造成的影響。
5. 將組態管理整合到 CI/CD 管道中。
6. 驗證所有推送的變更。

資源

相關的最佳實務：

- [OPS06-BP01 計畫變更失敗](#)
- [OPS06-BP02 測試部署](#)
- [OPS06-BP03 採用安全的部署策略](#)
- [OPS06-BP04 自動化測試和復原](#)

相關文件：

- [AWS Control Tower](#)
- [AWS 登陸區域加速器](#)
- [AWS Config](#)
- [什麼是 AWS Config ?](#)
- [AWS AppConfig](#)
- [什麼是 AWS CloudFormation ?](#)
- [AWS 開發人員工具](#)
- [AWS CodeBuild](#)

- [AWS CodePipeline](#)
- [AWS CodeDeploy](#)

相關影片：

- [AWS re:Invent 2022 - AWS 工作負載的主動管控與合規](#)
- [AWS re:Invent 2020：使用 AWS Config 實現合規即程式碼](#)
- [使用 AWS AppConfig 管理和部署應用程式組態](#)

OPS05-BP04 使用建置和部署管理系統

使用建置和部署管理系統。這些系統可減少由手動程序引起的錯誤，並減少部署變更的工作量。

在 AWS 中，可以使用 [AWS 開發人員工具](#) (例如 [AWS CodeBuild](#)、[AWS CodePipeline](#) 和 [AWS CodeDeploy](#)) 等服務來建置持續整合/持續部署 (CI/CD) 管道。

預期成果：您的建置和部署管理系統可支援組織的持續整合持續交付 (CI/CD) 系統，提供了使用正確組態自動化安全推展的功能。

常見的反模式：

- 在開發系統中編譯程式碼之後，您將可執行檔複製到生產系統中，卻無法啟動。本機日誌檔案指出其因缺少相依性而失敗。
- 您在開發環境中使用新功能成功建置應用程式，並提供程式碼以進行品質保證 (QA)。它未通過 QA，因為缺少靜態資產。
- 週五，在經過一番努力之後，您成功在開發環境中手動建置應用程式，包括新編碼的功能。到了週一，您卻無法重複成功建置應用程式的步驟。
- 您執行為新版本建立的測試。然後，您會在下週設定測試環境，並執行所有現有的整合測試，接著執行效能測試。新的程式碼具有無法接受的效能影響，必須重新開發及測試。

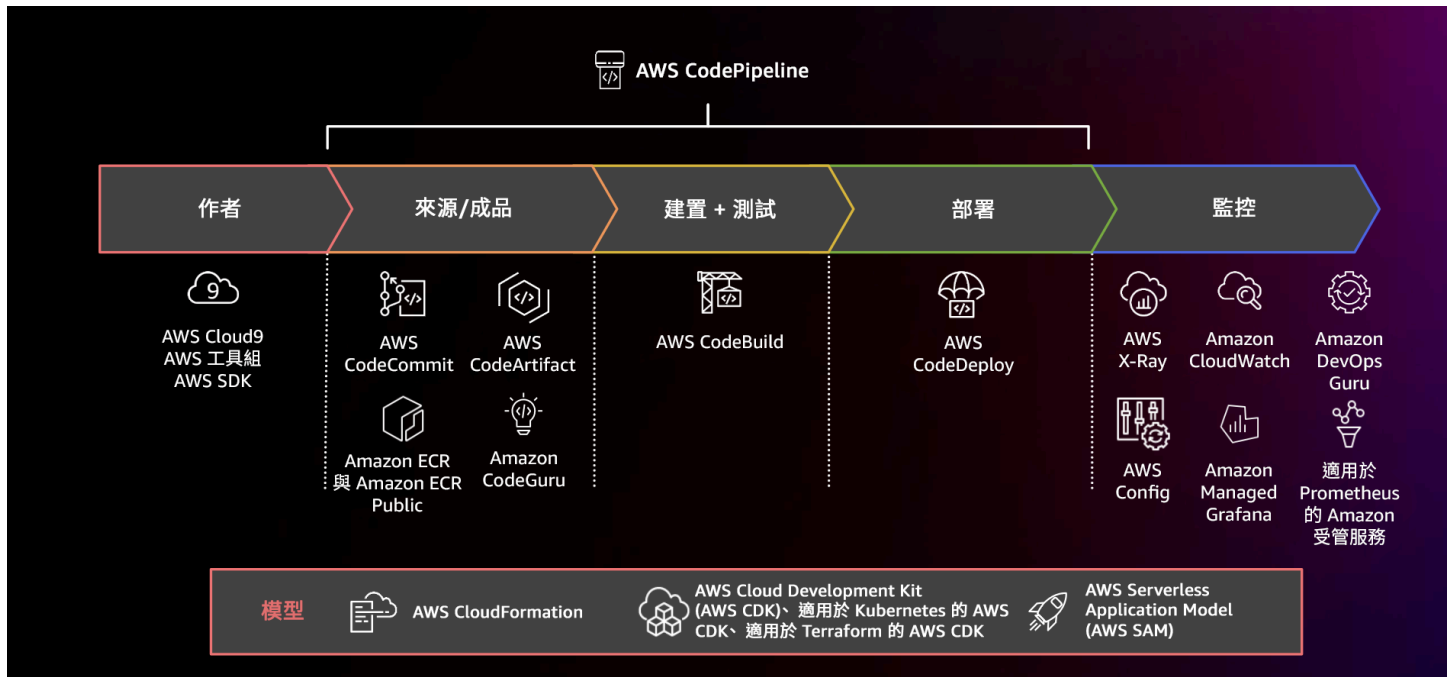
建立此最佳實務的優勢：透過提供用於管理建置和部署活動的機制，您可以減少執行重複性任務的工作量，讓團隊成員專注於高價值的創意任務，並減少手動程序導致的錯誤。

未建立此最佳實務時的曝險等級：中

實作指引

建置和部署管理系統可用來追蹤和實作變更、減少手動程序導致的錯誤，以及減少安全部署所需的工作量。從程式碼簽入到建置、測試、部署和驗證，完全自動化整合和部署管道。此舉可縮短前置時間、降低成本、促進增加變更頻率、減少工作量，並且增進協作。

實作步驟



圖中顯示使用 AWS CodePipeline 和相關服務的 CI/CD 管道

1. 使用版本控制系統來儲存和管理資產 (例如文件、原始程式碼和二進位檔案)。
2. 使用 CodeBuild 可編譯原始碼、執行單元測試，並產生可立即部署的成品。
3. 將 CodeDeploy 用作一項部署服務，可自動將應用程式部署至 [Amazon EC2](#) 執行個體、內部部署執行個體、[無伺服器 AWS Lambda 函數](#) 或 [Amazon ECS](#)。
4. 監控您的部署。

資源

相關的最佳實務：

- [OPS06-BP04 自動化測試和復原](#)

相關文件：

- [AWS 開發人員工具](#)
- [什麼是 AWS CodeBuild ?](#)
- [AWS CodeBuild](#)
- [什麼是 AWS CodeDeploy ?](#)

相關影片：

- [AWS re:Invent 2022 - 適用於 DevOps on AWS 的 AWS Well-Architected 最佳實務](#)

OPS05-BP05 執行修補程式管理

執行修補程式管理以取得功能、解決問題並保持遵循管控。自動化修補程式管理，以減少由手動程序引起的錯誤、進行擴展，並減少修補工作量。

修補程式和漏洞管理屬於您利益和風險管理活動的一部分。最好擁有不可變的基礎設施，並在已驗證的已知良好狀態下部署工作負載。如果這種方法不可行，剩下的方法就是進行修補。

[AWS Health](#) 是有關規劃的生命週期事件，以及其他影響 AWS 雲端 資源運作狀態且須採取行動的事件的權威資訊來源。您應得知即將進行的變更和更新。主要的規劃生命週期事件至少會提前六個月傳送。

[Amazon EC2 Image Builder](#) 提供管道來更新機器映像。作為修補程式管理的一部分，請考慮使用 [AMI 映像管道](#)的 [Amazon Machine Images](#) (AMI) 或具有 [Docker 映像管道](#)的容器映像，同時 AWS Lambda 會為[自訂執行時期和其他程式庫](#)提供模式以移除漏洞。

應使用 [Amazon EC2 Image Builder](#) 管理適用於 Linux 或 Windows Server 映像的 [Amazon Machine Images](#) 的更新。可以使用 [Amazon Elastic Container Registry \(Amazon ECR\)](#) 搭配現有管道來管理 Amazon ECS 映像並管理 Amazon EKS 映像。Lambda 包含[版本管理功能](#)。

若未先在安全環境中進行測試，就不應在生產系統上執行修補程式。只有在修補程式能夠支援營運或業務成果時，才應套用修補程式。在開啟 AWS 時，可以使用 [AWS Systems Manager Patch Manager](#) 來自動化修補受管系統的流程，並使用 [Systems Manager 維護時段](#)來排程活動。

預期成果：您的 AMI 和容器映像已完成修補、處於最新狀態，並準備好啟動。您可以追蹤所有已部署映像的狀態，並了解修補程式的合規狀況。您可以通報目前狀態，並設立程序來滿足合規需求。

常見的反模式：

- 您必須在兩小時內套用所有新的安全修補程式，結果導致應用程式與修補程式不相容而發生多次停機。

- 未修補的程式庫導致意外後果發生，因為有不明對象利用其中的漏洞來存取您的工作負載。
- 您自動修補開發人員環境，而未通知開發人員。您收到來自開發人員的多次投訴，表示其環境如預期停止運作。
- 您尚未在持續執行的執行個體上修補商用現成軟體。當軟體發生問題而您聯絡廠商時，他們會通知您不支援該版本，您必須修補至特定程度才能獲得協助。
- 您使用的加密軟體近期發佈了修補程式，使效能獲得大幅改善。未修補的系統因未修補仍存在效能問題。
- 收到發生零時差漏洞的通知時，需緊急修正並手動修補所有環境。
- 您不知道維護資源所需的重大行動，例如強制性版本更新，因為您未檢閱即將進行的規劃生命週期事件和其他資訊。您錯失規劃和執行的關鍵時間，導致團隊須緊急變更，且可能造成影響或非預期的停機時間。

建立此最佳實務的優勢：透過建立修補程式管理程序 (包括修補準則和在各環境中散佈的方法)，您就能擴展和報告修補程度。這樣可保證修補過程安全無虞，並確保能清楚看見已知修正的狀態。如此可促進採用所需的功能、迅速消除問題，並持續遵循管控要求。實作修補程式管理系統和自動化，以減少部署修補程式的工作量，並限制手動程序引起的錯誤。

未建立此最佳實務時的曝險等級：中

實作指引

修補系統以補救問題，獲得所需的功能，並保持符合管控政策和廠商支援需求。在不可變系統中，部署適當的修補程式集以實現所需的結果。自動化修補程式管理機制，以縮短修補時間、避免手動程序引起的錯誤，並減少修補工作量。

實作步驟

對於 Amazon EC2 Image Builder：

1. 使用 Amazon EC2 Image Builder 指定管道詳細資訊：
 - a. 建立映像管道並命名
 - b. 定義管道排程和時區
 - c. 設定任何相依性
2. 選擇配方：
 - a. 選取現有配方或建立新配方
 - b. 選取映像類型

- c. 提供配方的名稱和版本
 - d. 選取基礎映像
 - e. 新增組建元件並新增至目標登錄檔
3. 選用 - 定義您的基礎設施組態。
 4. 選用 - 定義組態設定。
 5. 審核設定。
 6. 定期維護配方乾淨度。

對於 Systems Manager Patch Manager：

1. 建立修補基準。
2. 選取修補操作方法。
3. 啟用合規報告和掃描。

資源

相關的最佳實務：

- [OPS06-BP04 自動化測試和復原](#)

相關文件：

- [什麼是 Amazon EC2 Image Builder](#)
- [使用 Amazon EC2 Image Builder 建立映像管道](#)
- [建立容器映像管道](#)
- [AWS Systems Manager Patch Manager](#)
- [使用 Patch Manager](#)
- [使用修補程式合規報告](#)
- [AWS 開發人員工具](#)

相關影片：

- [AWS 上適用於無伺服器應用程式的 CI/CD](#)
- [設計時考量 Ops](#)

相關範例：

- [AWS Systems Manager Patch Manager 教學課程](#)

OPS05-BP06 共用設計標準

在團隊之間共用最佳實務，以提高認識並最大化開發工作的效益。記載它們並且隨著您的架構演進讓它們保持在最新狀態。如果您的組織中強制執行共用標準，則必須存在用於請求標準新增、變更及例外狀況的機制。如果沒有此選項，標準就會限制創新。

預期成果：設計標準在貴組織的團隊之間共用。它們會隨著最佳實務的演變進行記錄和保存 up-to-date。

常見的反模式：

- 兩個開發團隊各自建立了使用者身分驗證服務。您的使用者必須針對要存取的系統的每一部分，維護一組單獨的憑證。
- 每個團隊管理他們自己的基礎設施。新的合規要求會強制變更您的基礎設施，每個團隊會以不同的方式實作。

建立此最佳實務的優勢：以共用的標準支援來實踐最佳實務，讓開發工作量發揮最大效益。記錄和更新設計標準可讓組織 up-to-date 符合最佳實務、安全和合規要求。

未建立此最佳實務時的曝險等級：中

實作指引

在團隊之間共用現有的最佳實務、設計標準、檢查清單、操作程序以及指引和管控要求。對於請求對設計標準進行變更、新增和例外設立程序，以支援改進和創新。讓團隊得知發布的內容。擁有機制，以在新最佳實務出現時保持設計標準 up-to-date。

客戶範例

AnyCompany Retail 有一個跨職能架構團隊，可建立軟體架構模式。這個團隊會建置具有內建合規和管控的架構。採用這些共用標準的團隊會獲得具有內建合規和管控的優點。他們可以快速地在設計標準的基礎上建置。架構團隊每季開會一次，評估架構模式並且視需要更新。

實作步驟

1. 識別擁有開發和更新設計標準的跨部門團隊。這個團隊應與整個組織的利益相關者合作，共同開發設計標準、操作程序、檢查清單、指引和管控需求。記錄設計標準並且在組織內共用。

- a. [AWS Service Catalog](#) 可以用來建立套裝服務，代表使用基礎設施即程式碼的設計標準。您可以與所有帳戶共用套裝服務。
2. 在識別新的最佳實務時，有適當的機制來保持設計標準 up-to-date。
3. 如果設計標準是集中強制執行，設立程序來請求變更、更新和豁免。

實作計劃的工作量：中。開發程序來建立和共用設計標準，即可與整個組織的利益相關者協調和合作。

資源

相關的最佳實務：

- [OPS01-BP03 評估治理要求](#) - 管控需求會影響設計標準。
- [OPS01-BP04 評估合規要求](#) - 合規是建立設計標準中的重要輸入。
- [OPS07-BP02 確保對營運準備度進行一致的審查](#) - 營運準備度檢查清單是在設計您的工作負載時實作設計標準的機制。
- [OPS11-BP01 建立持續改進程序](#) - 更新設計標準是持續改善的一部分。
- [OPS11-BP04 執行知識管理](#) - 在您的知識管理實務中，記錄和共用設計標準。

相關文件：

- [AWS Backup使用 自動化 AWS Service Catalog](#)
- [AWS Service Catalog 帳戶工廠增強型](#)
- [Expedia Group 如何使用 建置資料庫即服務 \(DBaaS \) 方案 AWS Service Catalog](#)
- [維護使用雲端架構模式的可見性](#)
- [簡化在 AWS Organizations 設定中共用 AWS Service Catalog 產品組合](#)

相關影片：

- [AWS Service Catalog – 入門](#)
- [AWS re : Invent 2020：像專家一樣管理您的 AWS Service Catalog 產品組合](#)

相關範例：

- [AWS Service Catalog 參考架構](#)
- [AWS Service Catalog 研討會](#)

相關服務：

- [AWS Service Catalog](#)

OPS05-BP07 實作用於提高程式碼品質的實務

實作相關實務以提高程式碼品質，並盡可能減少缺陷。部分範例包括測試驅動的開發、程式碼審查、標準採用和配對程式設計。將這些實務併入您的持續整合和交付程序。

預期成果：貴組織使用例如程式碼審核或配對程式設計的最佳實務來改善程式碼品質。開發人員和操作人員在軟體開發生命週期過程中採用程式碼品質最佳實務。

常見的反模式：

- 您將程式碼遞交至應用程式的主要分支，而未進程式碼審核。變更會自動部署至生產環境，並導致中斷。
- 會開發一個新的應用程式，沒有進行任何單元、端到端或整合測試。無法在部署之前測試應用程式。
- 您的團隊在生產中進行手動變更，以解決問題。變更不會經過測試或程式碼審核，而且不會在持續整合或交付程序中擷取或記錄。

建立此最佳實務的優勢：透過採用實務來提高程式碼品質，就能協助盡量減少生產環境中引發的問題。程式碼品質最佳實務包括配對程式設計、程式碼審核以及 AI 生產力工具的實作。

未建立此最佳實務時的曝險等級：中

實作指引

實作實務以提高程式碼品質，在部署之前將故障降至最低。使用像是測試驅動的開發、程式碼審核和配對程式設計等實務來提高開發的品質。

透過 Amazon Q Developer，利用生成式 AI 的強大功能，提升開發人員生產力和程式碼品質。Amazon Q Developer 包括程式碼建議的產生 (以大型語言模型為基礎)、單元測試的生產 (包括邊界條件)，以及透過偵測和修復安全漏洞增強程式碼安全性。

客戶範例

AnyCompany Retail 採用數個實務來改善程式碼品質。它們採用了測試驅動的開發做為撰寫應用程式的標準。對於某些新功能，它們會讓開發人員在衝刺期間一起進行配對程式設計。每個提取請求都會先經過資深開發人員的程式碼審核，然後再整合和部署。

實作步驟

1. 在您的持續整合和交付程序中，採用像是測試驅動開發、程式碼審核和配對程式設計的程式碼品質實務。使用這些技術來改善軟體品質。
 - a. 使用 [Amazon Q Developer](#)，這是一種生成式 AI 工具，可協助建立單元測試案例 (包括邊界條件)、使用程式碼和註解產生函數、實作眾所周知的演算法、偵測程式碼中的安全政策違規和漏洞、偵測機密、掃描基礎設施即程式碼 (IaC)、記錄程式碼以及更快速地學習第三方程式碼程式庫。
 - b. [Amazon CodeGuru Reviewer](#) 可以使用機器學習針對 Java 和 Python 程式碼提供程式設計建議。

實作計劃的工作量：中。有許多方式可以實作此最佳實務，但是要讓整個組織採用可能會是一項挑戰。

資源

相關的最佳實務：

- [OPS05-BP02 測試並驗證變更](#)
- [OPS05-BP06 共用設計標準](#)

相關文件：

- [採用測試驅動的開發方法](#)
- [使用 Amazon Q 加速您的軟體開發生命週期](#)
- [Amazon Q Developer 現已正式推出，包含可重新構想開發人員體驗的新功能預覽](#)
- [在 IDE 中使用 Amazon Q Developer 的終極速查表](#)
- [左移工作負載，利用 AI 建立測試](#)
- [Amazon Q 開發人員中心](#)
- [使用 Amazon CodeWhisperer 快速建置應用程式的 10 種方式](#)
- [使用 Amazon CodeWhisperer 超越程式碼覆蓋範圍](#)
- [使用 Amazon CodeWhisperer 進行提示詞工程的最佳實務](#)
- [敏捷式軟體指南](#)
- [CI/CD 管道是我的發行主管](#)
- [使用 Amazon CodeGuru Reviewer 自動進行程式碼審核](#)

- [採用測試驅動的開發方法](#)
- [DevFactory 如何使用 Amazon CodeGuru 建置更佳的應用程式](#)
- [關於配對程式設計](#)
- [RENGA Inc. 使用 Amazon CodeGuru 自動進程式碼審核](#)
- [敏捷開發的藝術：測試驅動的開發](#)
- [程式碼審核為何重要 \(而且確實可節省時間！\)](#)

相關影片：

- [使用適用於軟體開發的 Amazon Q Developer 代理程式來實作 API](#)
- [使用 JetBrains IDE 安裝、設定和使用 Amazon Q Developer \(操作說明\)](#)
- [掌握 Amazon CodeWhisperer 的藝術 - YouTube 播放清單](#)
- [AWS re:Invent 2020：使用 Amazon CodeGuru 持續改善程式碼品質](#)
- [AWS Summit ANZ 2021 - 透過 CDK 和測試驅動的開發施行測試優先策略](#)

相關服務：

- [Amazon Q Developer](#)
- [Amazon CodeGuru Reviewer](#)
- [Amazon CodeGuru Profiler](#)

OPS05-BP08 使用多個環境

使用多個環境來試驗、開發和測試您的工作負載。當環境接近生產環境時提高控制層級，以確保您的工作負載在部署後依預期執行。

預期成果：您有多個環境可反映您的合規和管控需求。在環境中測試並推廣程式碼，以逐步進行生產。

1. 您的組織可透過建立登陸區域來執行此操作，該區域提供治理、控制、帳戶自動化、聯網、安全和營運可觀測性。使用多個環境來管理這些登陸區域功能。常見的範例是沙盒組織，可用於開發和測試 [AWS Control Tower](#) 型登陸區域的變更，其中包括 [AWS IAM Identity Center](#) 和政策 (例如[服務控制政策 \(SCP\)](#))。這些元素全都可能對登陸區域內 AWS 帳戶 的存取和操作產生重大影響。
2. 除了這些服務之外，您的團隊還可利用 AWS 和 AWS 合作夥伴發布的解決方案，或您組織內開發的自訂解決方案來擴展登陸區域功能。AWS 所發布解決方案的範例包括 [Customizations for AWS Control Tower \(CfCT\)](#) 和 [AWS Control Tower Account Factory for Terraform \(AFT\)](#)。

3. 在付諸實際執行之前，您的組織會在過程中透過環境針對登陸區域套用相同的測試、推廣程式碼和政策變更原則。此策略為您的應用程式和工作負載團隊提供了穩定且安全的登陸區域環境。

常見的反模式：

- 您在共用開發環境中進行開發，而另一名開發人員覆寫了您的程式碼變更。
- 對共用開發環境的限制性安全控制可防止您試驗新服務和功能。
- 您對生產系統執行負載測試，並造成使用者停機。
- 在生產環境中發生導致資料遺失的嚴重錯誤。在您的生產環境中，您試圖重建導致資料遺失的條件，以便了解此情況如何發生，並防止它再次發生。為防止更多資料在測試期間遺失，必須讓使用者無法使用應用程式。
- 您正在操作多租用戶服務，且無法支援客戶對專用環境的要求。
- 您不一定會進行測試，但要測試時，您會在生產環境中進行。
- 您認為簡單的單一環境會覆寫環境內變更的影響範圍。
- 您升級一項重要的登陸區域功能，但變更卻有損您的團隊為新專案或現有工作負載供應帳戶的能力。
- 您將新的控制項套用至 AWS 帳戶，但變更卻影響工作負載團隊在其 AWS 帳戶 內部署變更的能力。

建立此最佳實務的優勢：當您部署多個環境時，您可以支援多個同時開發、測試和生產的環境，而不會在開發人員或使用者社群之間產生衝突。對於像是登陸區域等複雜功能，它可大幅降低變更的風險、簡化改善程序，並降低對環境進行重大更新的風險。使用登陸區域的組織自然受益於其 AWS 環境中的多個帳戶，因為具有帳戶結構、治理、網路和安全組態。經過一段時間後，隨著組織成長，登陸區域可能會進一步保護和組織您的工作負載和資源。

未建立此最佳實務時的曝險等級：中

實作指引

使用多個環境，並且對開發人員沙盒環境實施最低限度的控制，以協助實驗。提供多個單獨的開發環境，以協助實現並行工作，進而提高開發敏捷性。在環境逐漸達到生產環境的條件時，實施更嚴格的控制，以允許開發人員創新。使用基礎設施即程式碼和組態管理系統來部署所設定控制條件與生產環境一致的環境，以確保系統在部署後依預期執行。當不使用環境時，關閉環境以避免產生與閒置資源相關的成本 (例如，在夜間和週末關閉開發系統)。進行負載測試時，部署與生產環境同等的環境，以改善有效的結果。

平台工程、聯網和安全營運等團隊通常會在擁有不同需求的組織層級管理各種功能。僅是將帳戶分隔開來並不足以提供及維護實驗、開發和測試各自所需的環境。在這種情況下，可建立個別的 AWS Organizations 執行個體。

資源

相關文件：

- [Instance Scheduler on AWS](#)
- [什麼是 AWS CloudFormation？](#)
- [使用多個帳戶來組織您的 AWS 環境 - 多個組織 - 測試整體 AWS 環境的變更](#)
- [AWS Control Tower 指南](#)

OPS05-BP09 進行頻繁、小型、可逆的變更

頻繁、細微和可逆的變更會縮小變更的範圍和影響。與變更管理系統、組態管理系統以及建置與交付系統搭配使用時，頻繁、細微和可逆的變更可縮小變更的範圍和影響。透過回復變更，可以更有效地進行疑難排解並加快修復速度。

常見的反模式：

- 您每季部署應用程式的新版本，這表示在這段變更期間，核心服務為關閉狀態。
- 您經常對資料庫結構描述進行變更，但未在您的管理系統中追蹤變更。
- 您執行手動就地更新，並覆寫現有的安裝和組態，但沒有明確的回復計畫。

建立此最佳實務的優勢：頻繁部署小型變更，可加快開發工作的速度。若變更幅度很小，就更容易了解變更是否會產生意外的後果，也更容易回復。如果變更可逆，由於復原過程較單純，因此實作變更的風險也會降低。變更程序的風險降低，而且變更失敗的影響也會降低。

未建立此最佳實務時的曝險等級：低

實作指引

透過頻繁、細微和可逆的變更來縮小變更的範圍和影響。這樣可以簡化疑難排解，有助於加速修復，並提供回復變更的選項。另外還可以提高您為企業帶來價值的速度。

資源

相關的最佳實務：

- [OPS05-BP03 使用組態管理系統](#)
- [OPS05-BP04 使用建置和部署管理系統](#)
- [OPS06-BP04 自動化測試和復原](#)

相關文件：

- [在上實作 Microservices AWS](#)
- [微型服務 - 可觀測性](#)

OPS05-BP10 完全自動化整合和部署

自動化工作負載的建置、部署和測試。此舉可減少由手動程序引起的錯誤，以及部署變更的工作量。

使用[資源標籤](#)和 [AWS Resource Groups](#) 並遵循一致的[標記策略](#)，來套用中繼資料，以幫助識別您的資源。標記您的資源，以用於組織、成本會計、存取控制，以及將自動執行營運活動設為目標。

預期成果：開發人員使用工具交付程式碼並推廣至生產環境。開發人員不必登入 AWS 管理主控台 就可以交付更新。有完整的變更與組態稽核記錄，可滿足管控和合規的需求。程序可在各團隊重複執行並且標準化。開發人員可全心專注於開發和程式碼推送，從而提高生產力。

常見的反模式：

- 週五，您完成了為功能分支編寫新程式碼。週一，執程式碼品質測試指令碼和每個單位測試指令碼之後，請為下一排程版本檢查程式碼。
- 系統會指派您編寫修正程式碼，以解決影響生產環境中大量客戶的重大問題。測試修正後，您遞交程式碼和電子郵件變更管理內容，以請求核准將其部署到生產環境中。
- 您以開發人員身分登入 AWS 管理主控台，以使用非標準方法和系統來建立新的開發環境。

建立此最佳實務的優勢：透過實作自動化建置和部署管理系統，您可以減少手動程序引起的錯誤，以及部署變更的工作量，協助您的團隊成員專注於提供商業價值。您在推廣至生產環境的同時，加快了交付速度。

未建立此最佳實務時的曝險等級：低

實作指引

您使用建置和部署管理系統來追蹤和實作變更，以減少由手動程序引起的錯誤，並減少工作量。從程式碼簽入到建置、測試、部署和驗證，完全自動化整合和部署管道。此舉可縮短前置時間、促進增加變更頻率、減少工作量、加快上市速度、提高生產力，並且在您推廣到生產環境時提高程式碼的安全性。

資源

相關的最佳實務：

- [OPS05-BP03 使用組態管理系統](#)
- [OPS05-BP04 使用建置和部署管理系統](#)

相關文件：

- [什麼是 AWS CodeBuild？](#)
- [什麼是 AWS CodeDeploy？](#)

相關影片：

- [AWS re:Invent 2022 - 適用於 DevOps on AWS 的 AWS Well-Architected 最佳實務](#)

OPS 6. 如何緩解部署風險？

採用可快速提供品質意見回饋，並從成果不盡理想的改變中快速復原的方法。使用這些實務可緩解部署變更所帶來問題的影響。

最佳實務

- [OPS06-BP01 計畫變更失敗](#)
- [OPS06-BP02 測試部署](#)
- [OPS06-BP03 採用安全的部署策略](#)
- [OPS06-BP04 自動化測試和復原](#)

OPS06-BP01 計畫變更失敗

計劃在部署造成非預期成果時恢復到已知的良好狀態，或者在生產環境中進行修復。擁有制定這類計畫的政策可以協助所有團隊訂立政策，從失敗變更中恢復。一些範例策略包括部署和回復步驟、變更政

策、功能旗標、流量隔離和流量轉移。單一版本可能包含多個相關元件變更。策略要能提供您承受或從任何失敗元件變更中恢復的能力。

預期成果：您已經為失敗變更準備了詳細的恢復計畫。此外，您也縮減了發行版本的大小，如此一來，對其他工作負載元件的潛在影響將降到最低。因此，您可以縮短因變更失敗而造成的可能停機時間，並提高回復時間的彈性和效率，進而降低對業務的影響。

常見的反模式：

- 您執行了部署，而您的應用程式變得不穩定，但系統中似乎有作用中使用者。您必須決定是否要復原變更並影響作用中使用者，或在知道使用者無論如何都會受到影響的情況下，等待復原變更。
- 在進行路由變更後，您可以存取新的環境，但其中一個子網路變成無法連線。您必須決定是否要復原所有項目，或嘗試修正無法存取的子網路。當您做出該決定時，子網路仍無法連線。
- 您的系統架構並不允許以較小版本進行更新。因此，在部署失敗期間，您無法回復這些大量變更。
- 您未使用基礎設施即程式碼 (IaC)，並且以手動方式更新了基礎設施，從而造成不理想的組態。您無法有效追蹤和還原手動變更。
- 由於您尚未測量部署的增加頻率，您的團隊不會想降低變更規模和改善每次變更的回復計畫，進而造成更高的風險和失敗率。
- 請不要測量因變更失敗而導致中斷的總持續時間。您的團隊無法排定優先順序並改善其部署程序和回復計畫的效能。

建立此最佳實務的好處：擁有從失敗變更中復原的計畫，可最大限度地減少復原的平均時間（MTTR），並減少您的業務影響。

未建立此最佳實務時的曝險等級：高

實作指引

發行團隊採用的一致記錄政策和實務可讓組織規劃發生失敗變更時應採取的動作。該政策應允許在特定情況下向前修正。在任何情況下，向前修正或回復計畫都應該在部署到現場生產之前先經過詳細記錄和測試，以將回復變更所需的時間降到最低。

實作步驟

1. 記錄要求團隊有效計畫在特定期間內還原變更的政策。
 - a. 政策應指明允許向前修正的情況。
 - b. 要求所有參與者皆能存取記錄完善的回復計畫。

- c. 指定回復需求 (例如：發現部署未經授權的變更時)。
2. 分析工作負載每個元件相關所有變更的影響程度。
 - a. 如果可重複的變更遵循強制執行變更政策的一致工作流程，則允許這些變更進行標準化、範本化和預先授權。
 - b. 縮小變更規模以減少任何變更的潛在影響，進而降低回復時間和對業務的影響。
 - c. 確保回復程序會將程式碼回復到已知的良好狀態，避免可能發生的意外。
3. 整合工具和工作流程，以程式設計方式執行政策。
4. 讓其他工作負載擁有者可以看到變更相關資料，以改善任何無法回復失敗變更的診斷速度。
 - a. 使用可見的變更資料來衡量這項實務的成效，並識別出反覆改進方式。
5. 使用監控工具來驗證部署成敗，以加速回復的決策過程。
6. 測量失敗變更期間的中斷時間，以持續修正回復計畫。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS06-BP04 自動化測試和復原](#)

相關文件：

- [AWS Builders Library | 確保部署期間的復原安全](#)
- [AWS 白皮書 | 雲端的變更管理](#)

相關影片：

- [re:Invent 2019 | Amazon 的高可用性部署方法](#)

OPS06-BP02 測試部署

使用與生產環境相同的部署組態、安全控制、步驟和程序，在生產前測試發行程序。驗證所有部署的步驟均按照預期完成，例如檢查檔案、組態和服務。透過功能、整合和負載測試以及任何監控 (例如運作狀態檢查) 進一步測試所有變更。透過這些測試，您可以及早發現部署問題，有機會在生產前進行規劃和問題緩解。

您可以建立暫時的平行環境來測試每項變更。使用基礎設施即程式碼 (IaC) 來自動化測試環境的部署，協助減少涉及的工作量，並確保穩定性、一致性和更快的功能交付。

預期成果：您的組織採用測試驅動型開發文化，其中包含測試部署。如此一來，便能確保團隊專注於交付商業價值，而非管理發行版本。團隊會及早找出部署風險，並訂定適當的緩解方案。

常見的反模式：

- 使用生產版本期間，因為未經測試的部署經常會導致問題，而需要疑難排解或升級處理。
- 您的版本包含更新現有資源的基礎設施即程式碼 (IaC)。您不確定 IaC 是否會成功執行，或對資源造成影響。
- 您為應用程式部署一個新功能。該功能無法按照您的預期運作，且在受影響的使用者回報之前無法預見問題。
- 您更新憑證。您不小心將憑證安裝到錯誤的元件，這些元件未被偵測並因為無法建立與網站的安全連線，而影響了網站訪客。

建立此最佳實務的優勢：針對部署程序的生產前階段及其帶來的變更進行廣泛測試，將部署步驟對生產的潛在負面影響降到最低。這麼做能增加產品發行期間的信心，並盡可能減少操作支援，同時不影響交付變更的速度。

未建立此最佳實務時的曝險等級：高

實作指引

測試部署程序與測試部署所產生的變更同樣重要。您可以在生產前環境中測試部署步驟，盡可能準確反映生產環境。諸如不完整或錯誤部署步驟，或者配置錯誤等常見問題都能在生產環境之前偵測。此外，您也可以測試回復步驟。

客戶範例

作為持續整合和持續交付（CI/CD）管道的一部分，AnyCompany Retail 會執行在類似生產環境中為其客戶發佈基礎設施和軟體更新所需的定義步驟。流程包含許多預先檢查程序，可以在部署之前偵測到資源偏移（偵測 IaC 以外所執行的資源變更），以及驗證 IaC 啟動時所採取的動作。這個程序會驗證部署步驟，例如確認特定檔案和組態已準備就緒，或服務處於執行狀態，並在向負載平衡器重新註冊之前，正確回應本機上的運作狀態檢查。此外，所有變更都標記了許多自動化測試，例如功能、安全性、迴歸、整合和負載測試。

實作步驟

1. 執行安裝前檢查，將生產前環境反映到生產環境。

- a. 使用[偏離偵測](#)來偵測 資源是否已在 之外變更 CloudFormation。
- b. 使用[變更集](#)來驗證堆疊更新的意圖是否符合啟動變更集時 CloudFormation 採取的動作。
2. 這會觸發 [AWS CodePipeline](#) 中的手動核准步驟，以授權生產前環境的部署。
3. 使用[AWS CodeDeploy AppSpec](#)檔案等部署組態來定義部署和驗證步驟。
4. 在適用的情況下，[AWS CodeDeploy 與其他 AWS 服務整合](#)，或與[AWS CodeDeploy 合作夥伴產品和服務整合](#)。
5. 使用 Amazon CloudWatch AWS CloudTrail、和 Amazon SNS事件通知來[監控部署](#)。
6. 執行部署後自動化測試，包括功能、安全性、迴歸、整合和負載測試。
7. 對部署問題[進行故障診斷](#)。
8. 成功驗證前述步驟後，應該啟動手動核准工作流程，以授權部署到生產環境。

實作計劃的工作量：高

資源

相關的最佳實務：

- [OPS05-BP02 測試並驗證變更](#)

相關文件：

- [AWS Builders' Library | 自動化安全、移交部署 | 測試部署](#)
- [AWS 白皮書 | 在上練習持續整合和持續交付 AWS](#)
- [阿波羅的故事 - Amazon 的部署引擎](#)
- [如何在運送程式碼之前在 AWS CodeDeploy 本機進行測試和偵錯](#)
- [整合網路連線測試與基礎設施部署](#)

相關影片：

- [re:Invent 2020 | 在 Amazon 測試軟體和系統](#)

相關範例：

- [教學課程 | 透過驗證測試部署 和 Amazon ECS服務](#)

OPS06-BP03 採用安全的部署策略

安全的生產部署可控制有益變更的流程，目的是將這些變更對客戶造成的任何影響降到最低。安全控制提供檢查機制，以驗證所需的結果，並限制變更引入的任何缺陷或部署失敗造成的影響範圍。安全推出可能包括諸如功能旗標、一體式、滾動式 (Canary 版本)、不可變、流量拆分和藍/綠部署等策略。

預期成果：您的組織使用持續整合持續交付 (CI/CD) 系統，可提供自動化安全部署的功能。團隊必須使用適當的安全推出策略。

常見的反模式：

- 您一次性將失敗的變更部署至所有生產環境。因此，所有客戶都會同時受到影響。
- 同時部署到所有系統中引入的缺陷需要緊急釋放。為所有客戶修正它需要數天的時間。
- 管理生產發行需要多個團隊的規劃和參與。這會限制您為客戶經常更新功能的能力。
- 您透過修改現有系統來執行可變部署。發現變更失敗之後，您必須再次修改系統以還原舊版本，這會延長回復時間。

建立此最佳實務的優勢：自動化部署在推出速度和持續為客戶提供有益變更之間取得了平衡。限制影響範圍可以預防損失慘重的部署失敗，並盡可能提高團隊有效回應故障的能力。

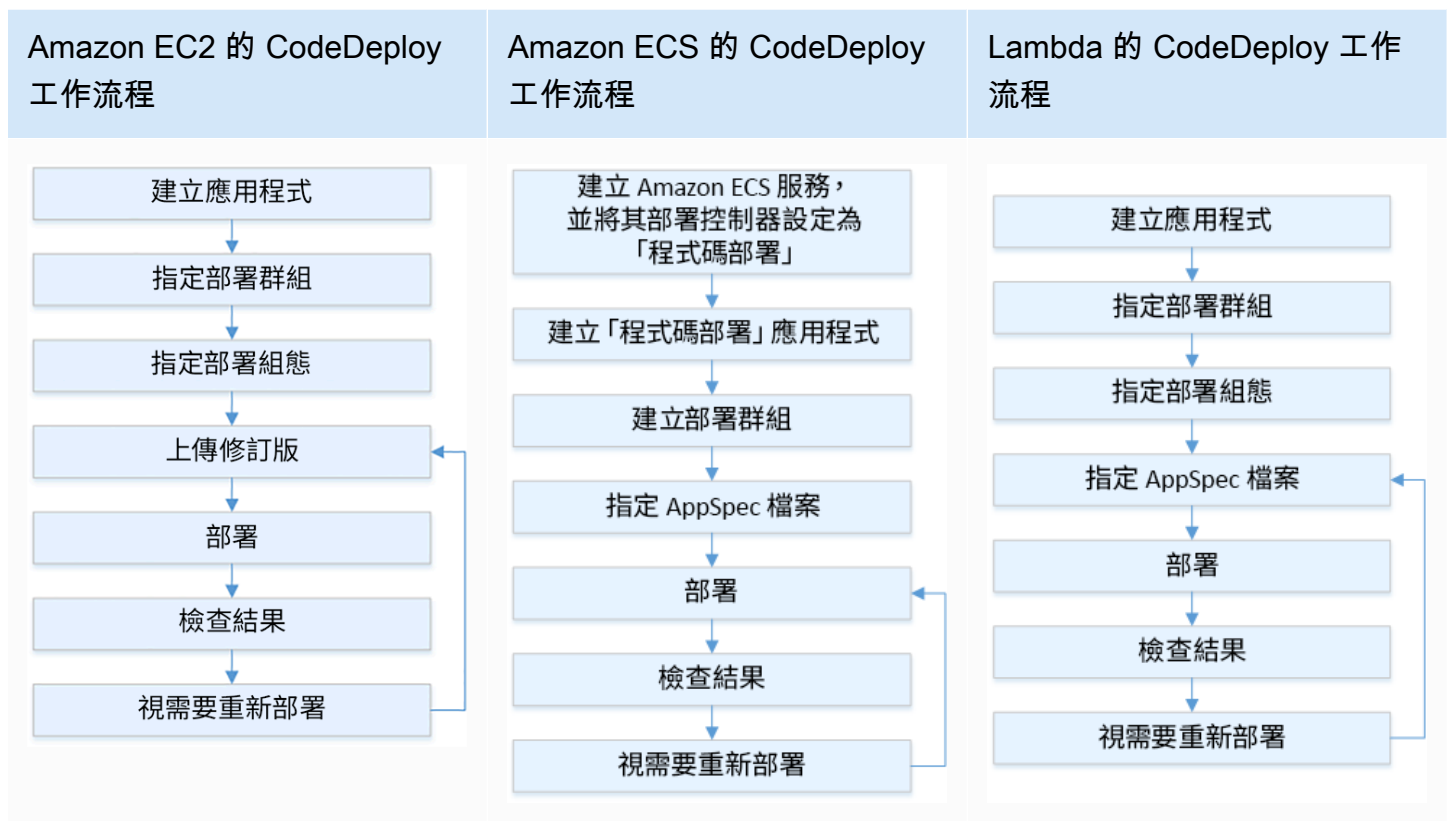
未建立此最佳實務時的曝險等級：中

實作指引

持續交付失敗可能導致服務可用性降低和糟糕的客戶體驗。為了盡可能提高部署成功率，請在端對端發行程序中實施安全控制，盡量減少部署錯誤，而目標則是實現零部署失敗。

客戶範例

AnyCompany Retail 的使命是實現最小至零停機時間的部署，這表示部署期間沒有對使用者產生任何可感知的負面影響。為此，公司已建立了部署模式 (請參閱下列工作流程圖表)，例如滾動部署和藍/綠部署。所有團隊都在其 CI/CD 管道中採用一個或多個模式。



實作步驟

- 推廣到生產之後，使用核准工作流程可啟動生產推出步驟的一系列動作。
- 使用自動化部署系統，例如 [AWS CodeDeploy](#)。AWS CodeDeploy [部署選項](#) 包括適用於 EC2/內部部署的就地部署，以及適用於 EC2/內部部署、AWS Lambda 和 Amazon ECS 的藍/綠部署 (請參閱上述工作流程圖)。
 - 在適用的情況下，[將 AWS CodeDeploy 與其他 AWS 服務整合](#)，或者[將 AWS CodeDeploy 與合作夥伴的產品和服務整合](#)。
- 對 [Amazon Aurora](#) 和 [Amazon RDS](#) 等資料庫使用藍/綠部署。
- 使用 Amazon CloudWatch、AWS CloudTrail 和 Amazon Simple Notification Service (Amazon SNS) 事件通知來[監控部署](#)。
- 執行部署後自動化測試，包括功能、安全性、迴歸、整合和任何負載測試。
- 對部署問題[進行故障診斷](#)。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS05-BP02 測試並驗證變更](#)
- [OPS05-BP09 進行頻繁、小型、可逆的變更](#)
- [OPS05-BP10 完全自動化整合和部署](#)

相關文件：

- [AWS 建置者資料中心 | 自動化安全、無人為介入的部署 | 生產部署](#)
- [AWS 建置者資料中心 | 我的 CI/CD 管道是我的發行主管 | 安全、自動的生產發行](#)
- [AWS 白皮書 | 在 AWS 上實行持續整合和持續交付 | 部署方法](#)
- [AWS CodeDeploy 使用者指南](#)
- [使用 AWS CodeDeploy 中的部署組態](#)
- [設定 API Gateway Canary 發行部署](#)
- [Amazon ECS 部署類型](#)
- [Amazon Aurora 和 Amazon RDS 中的全受管藍/綠部署](#)
- [使用 AWS Elastic Beanstalk 的藍/綠部署](#)

相關影片：

- [re:Invent 2020 | 無人為介入：Amazon 的自動化持續交付管道](#)
- [re:Invent 2019 | Amazon 的高可用性部署方法](#)

相關範例：

- [在 AWS CodeDeploy 中嘗試使用藍/綠部署範例](#)
- [研討會 | 使用 AWS CDK 建置 Lambda Canary 部署的 CI/CD 管道](#)
- [研討會 | 利用 Amazon ECS 建置您的第一個 DevOps 藍/綠管道](#)
- [研討會 | 利用 Amazon EKS 建置您的第一個 DevOps 藍/綠管道](#)
- [研討會 | EKS GitOps 搭配 ArgoCD](#)
- [研討會 | AWS 上的 CI/CD 研討會](#)

- [針對容器型 Lambda 函數使用 AWS SAM 實作跨帳戶 CI/CD](#)

OPS06-BP04 自動化測試和復原

為了提高部署程序的速度和可靠性，請在生產前和生產環境中制定自動化測試和回復功能的策略。在部署到生產環境時自動化測試，以模擬人類與系統的互動，驗證部署的變更。自動回復以快速回復到之前已知的良好狀態。回復應該在預先定義的條件下自動啟動，例如當未達到預期成果或自動化測試失敗時。將這兩項活動的自動化可以提高部署成功率，盡可能縮短回復時間，並減少對業務的潛在影響。

預期成果：您的自動化測試和回復策略將整合至持續整合與持續交付 (CI/CD) 管道。您的監控能夠根據您的成功條件進行驗證，並在失敗時啟動自動回復。這會將終端使用者和客戶受到的任何負面影響降到最低。例如，當所有測試結果都達到標準時，您可以利用相同的測試案例，將程式碼提升至啟動自動迴歸測試的生產環境。若迴歸測試結果不符預期，則會在管線工作流程中啟動自動回復。

常見的反模式：

- 您的系統架構並不允許以較小版本進行更新。因此，在部署失敗期間，您無法回復這些大量變更。
- 您的部署程序包含一系列手動步驟。將變更部署到工作負載之後，即可開始部署後測試。測試之後，您會發現工作負載無法運作，且客戶中斷連線。然後您開始回復到之前的版本。所有這些手動步驟都會延遲整體系統回復，並對客戶造成長期影響。
- 您花時間為應用程式中不常使用的功能開發自動化測試案例，因而大幅降低了自動化測試功能的投資報酬率。
- 您的版本包含彼此獨立的應用程式、基礎設施、修補程式和組態更新。但是，您有一個 CI/CD 管道可以一次交付所有變更。一個元件故障會強迫您還原所有變更，進而使回復過程變得複雜且效率低下。
- 您的團隊在第一個衝刺階段完成編碼，並開始衝刺兩項工作，但直到第三個衝刺階段，計畫中都不包括測試。最終，自動化測試找出第一個衝刺階段的缺漏，必須在測試第二個衝刺階段前解決，才能啟動交付項目，因此整個版本延遲，進而降低您的自動化測試效率。
- 生產版本的自動迴歸測試案例已經完成，但您並未監控工作負載的運作狀況。由於無法查看是否已重啟服務，您不確定是否需要回復或已啟動回復。

建立此最佳實務的優勢：自動化測試可以提高測試流程的透明度，以及您在更短時間內顧及更多功能的能力。在生產環境中測試和驗證變更，可以立即識別出問題。改善自動化測試工具的一致性可以更精確地偵測問題。透過自動回復至舊版本，將對客戶的影響降至最低。自動化回復最終可減少業務影響，讓您對部署功能更有信心。整體而言，這些功能會減少 time-to-delivery，同時確保品質。

未建立此最佳實務時的曝險等級：中

實作指引

自動測試已部署的環境，更快確認是否達到預期成果。當無法達成預先定義的結果時，自動還原到先前的良好狀態，以盡量縮短還原時間，並減少由手動程序引起的錯誤。將測試工具與管道工作流程整合，持續進行測試並減少手動輸入。優先處理自動化測試案例，例如減緩最高風險且需要在每次變更時經常測試的案例。此外，還可以根據測試計畫中預先定義的特定條件進行自動回復。

實作步驟

1. 為您的開發生命週期建立測試生命週期，定義需求規劃到測試案例開發、工具配置、自動化測試和測試案例結案等每個測試程序階段。
 - a. 根據您的整體測試策略建立針對特定工作負載的測試方式。
 - b. 在整個開發生命週期中，考慮適當的連續測試策略。
2. 根據您的業務需求和管道投資，選擇用於測試和回復的自動化工具。
3. 決定您應該分別自動化和手動執行哪些測試案例。這些內容皆可以根據受測功能的業務價值優先順序來決定。使所有團隊成員隨時接收計畫最新資訊，並確認執行手動測試的權責分配。
 - a. 將自動化測試功能應用於對自動化有意義的特定測試案例，例如可重複或經常執行的案例、需要重複作業的案例，或跨多個組態所需的案例。
 - b. 在自動化工具中定義測試自動化指令碼和成功條件，如此一來，當特定案例失敗時，可以啟動持續的工作流程自動化。
 - c. 定義自動回復的特定失敗條件。
4. 測試案例其中複雜度和人工互動具較高的失敗風險，因此必須排定測試自動化的優先順序，透過詳盡的測試案例開發來產生一致的結果。
5. 將您的自動化測試和回復工具整合到 CI/CD 管道。
 - a. 為變更制定明確的成功條件。
 - b. 監控觀察以偵測這些條件，並在符合特定回復條件時自動回復變更。
6. 執行不同類型的自動化生產測試，例如：
 - a. A/B 測試，以顯示結果比較兩個使用者測試組之間的當前版本。
 - b. Canary 測試讓您能在將變更發佈給所有使用者之前，先將其發佈給一部分使用者。
 - c. 功能旗標測試允許您從應用程式外部標記新版本的功能 (每次僅限一個)，進而使每個新功能皆能逐一進行驗證。
 - d. 迴歸測試，以現有的關聯元件驗證新功能。
7. 透過其他應用程式和元件，監控應用程式、交易和互動的操作面向。開發報告，以便按照工作負載顯示變更成功率，讓您得以識別出能夠進一步最佳化的自動化和工作流程部分。

- a. 開發測試結果報告，協助您快速決定是否應該調用回復程序。
 - b. 實施策略，允許根據一個或多個測試方法導出的預定義失敗條件進行自動回復。
8. 開發自動化測試用例，以便在未來可重複的變更中重複使用。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS06-BP01 計畫變更失敗](#)
- [OPS06-BP02 測試部署](#)

相關文件：

- [AWS Builders Library | 確保部署期間的復原安全](#)
- [使用 重新部署和復原部署 AWS CodeDeploy](#)
- [使用 自動化部署時的 8 個最佳實務 AWS CloudFormation](#)

相關範例：

- [使用 Selenium AWS Lambda AWS Fargate、和 AWS 開發人員工具進行無伺服器 UI 測試](#)

相關影片：

- [re:Invent 2020 | 無人為介入：Amazon 的自動化持續交付管道](#)
- [re:Invent 2019 | Amazon 的高可用性部署方法](#)

OPS 7. 如何知道自己準備好支援工作負載？

評估工作負載、流程和程序及人員的營運準備度，了解工作負載相關營運風險。

最佳實務

- [OPS07-BP01 確保人員能力](#)
- [OPS07-BP02 確保對營運準備度進行一致的審查](#)
- [OPS07-BP03 使用執行手冊執行程序](#)

- [OPS07-BP04 使用程序手冊來調查問題](#)
- [OPS07-BP05 做出部署系統和變更的明智決策](#)
- [OPS07-BP06 建立生產工作負載的支援計劃](#)

OPS07-BP01 確保人員能力

建立一種機制，用於驗證您有適當數量受過培訓的人員來支援工作負載。他們必須受過組成您的工作負載的平台和服務的培訓。為他們提供操作工作負載所需的知識。您必須擁有足夠受過培訓的人員，才能支援工作負載的一般操作，並且針對會發生的任何事件進行疑難排解。擁有足夠的人員，以便您可以輪替待命和休假的人員，避免倦怠。

預期成果：

- 有足夠受過培訓的人員可以在工作負載可用時支援工作負載。
- 您為人員提供組成您的工作負載的軟體和服務的培訓。

常見的反模式：

- 在沒有受過培訓操作使用中平台和服務的團隊成員之情況下，部署工作負載。
- 沒有足夠的人員可以支援待命輪替或人員休假。

建立此最佳實務的優勢：

- 擁有熟練的團隊成員有助於有效支援您的工作負載。
- 具有足夠的團隊成員，您可以支援工作負載和待命輪替，同時降低倦怠風險。

未建立此最佳實務時的曝險等級：高

實作指引

驗證人員是否已經過充分培訓，可支援工作負載。確認擁有足夠且訓練有素的團隊成員，以妥善應對一般營運活動，包括待命輪替。

客戶範例

AnyCompany Retail 確保支援工作負載的團隊有適當的配備人員且訓練有素。他們有足夠的工程師可以支援待命輪替。人員會獲得工作負載建置基礎的軟體和平台的培訓，並且鼓勵他們考取認證。有足夠的人員讓員工可以休假，同時仍然支援工作負載和待命輪替。

實作步驟

1. 指派足夠的人員來操作和支援工作負載，包括待命職責、安全問題和生命週期事件，例如終止支援和憑證輪換任務。
2. 為您的人員提供組成您的工作負載的軟體和平台的培訓。
 - a. [AWS 培訓和認證](#) 有一個關於 AWS 的課程庫。它們提供免費和付費的線上或面授課程。
 - b. [AWS 舉辦活動和網路研討會](#)，您可向 AWS 專家學習。
3. 定期執行下列工作：
 - 隨著操作條件和工作負載變更，評估團隊規模和技能。
 - 調整團隊大小和技能以符合操作要求。
 - 透過 AWS Health 驗證 [處理規劃的生命週期事件](#)、意外的安全性和操作通知的能力和容量。

實作計劃的工作量：高。招聘和培訓團隊來支援工作負載需要大量的努力，但是會有重大的長期優點。

資源

相關的最佳實務：

- [OPS11-BP04 執行知識管理](#) - 團隊成員必須擁有操作和支援工作負載所需的資訊。知識管理是提供這項能力的關鍵。

相關文件：

- [AWS 活動和研討會](#)
- [AWS 培訓和認證](#)

OPS07-BP02 確保對營運準備度進行一致的審查

使用營運準備度審查 (ORR) 來確認您可以運作工作負載。ORR 是在 Amazon 開發的機制，可確認團隊是否可放心地運行工作負載。ORR 是使用需求檢查清單的審查和檢查程序。ORR 是一種自助服務體驗，團隊會透過此體驗來進行工作負載的認證。ORR 包含的最佳實務皆汲取我們多年來建置軟體所獲得的經驗。

ORR 檢查清單包含架構建議、營運程序、事件管理和發行品質。錯誤糾正 (CoE) 程序是這些項目的主要驅動要素。您專屬的事件後分析應有助於專屬 ORR 的發展。ORR 不只是遵循最佳實務，還能防止先前發生過的事件再發。最後，ORR 中也能夠包含安全性、管控和合規需求。

在工作負載啟動以全面供應前，並在整個軟體開發生命週期執行 ORR。在啟動前執行 ORR 可改善安全運行工作負載的能力。定期針對工作負載重新執行 ORR 可捕捉最佳實務中的任何偏移。您可以為新服務的推出制定 ORR 檢查清單，並為定期審查制定 ORR。此可協助您掌握新出現的最佳實務最新狀態，並採納從事件後分析獲得的經驗。隨著您可以更熟練地使用雲端後，您就可以在架構中建置 ORR 需求作為預設值。

預期成果：制定 ORR 檢查清單，內含組織的最佳實務。ORR 會在工作負載啟動前執行。ORR 會在工作負載生命週期的過程中定期執行。

常見的反模式：

- 您啟動工作負載，但不知道自己是否能夠運行工作負載。
- 啟動工作負載的認證中未納入管控和安全性需求。
- 不會定期重新評估工作負載。
- 工作負載啟動，但不需設置必要的程序。
- 您可以在多個工作負載中看到重複出現的相同根本原因失敗。

建立此最佳實務的優勢：

- 工作負載包含架構、程序和管理最佳實務。
- 經驗已納入 ORR 程序中。
- 工作負載啟動時，已設置必要的程序。
- ORR 會在工作負載的整個軟體生命週期執行。

若未建立此最佳實務的風險等級：高

實作指引

ORR 有兩個部分：程序和檢查清單。貴組織應採用 ORR 程序，並由執行主辦人支援此程序。至少，必須在工作負載啟動以全面供應前執行 ORR。在整個軟體開發生命週期執行 ORR，使其與最佳實務或新需求保持同步。ORR 檢查清單應包含組態項目、安全性和管控需求，以及來自貴組織的最佳實務。隨著時間的推移，可以使用諸如 [AWS Config](#)、[AWS Security Hub CSPM](#) 和 [AWS Control Tower Guardrails](#) 之類的服務，將 ORR 的最佳實務建置為防護機制，以自動偵測最佳實務。

客戶範例

在發生數個生產事件後，AnyCompany Retail 決定實作 ORR 程序。他們建立了一份檢查清單，其中由最佳實務、管控和合規需求，以及從中斷中汲取的經驗教訓所組成。在工作負載啟動前，新的工作負載

會執行 ORR。每個工作負載每年都會使用一部分的最佳實務來執行 ORR，以便納入在 ORR 檢查清單中新增的最佳實務和需求。隨著時間的推移，AnyCompany Retail 使用 [AWS Config](#) 來偵測一些最佳實務，從而加快 ORR 流程。

實作步驟

要了解有關 ORR 的更多資訊，請閱讀 [《營運準備度審查 \(ORR\)》白皮書](#)。其中提供詳細的資訊，說明 ORR 程序的歷史、如何建立您專屬的 ORR 實務，以及如何制定 ORR 檢查清單。以下步驟是該文件的精簡版本。如需深入了解 ORR 是什麼，以及如何建立您專屬的 ORR，我們建議閱讀該白皮書。

1. 召集關鍵利益相關者，包含安全性、營運和開發等團隊的代表人員。
2. 請每位利益相關者提供至少一個需求。對於第一次的反覆測試，請嘗試將項目數限制在三十個以下。
 - 《營運準備度審查 (ORR)》白皮書的 [附錄 B：ORR 問題範例](#) 包含您可以開始使用的範例問題。
3. 將需求集中放在試算表中。
 - 您可以在 [AWS Well-Architected Tool](#) 中使用 [自訂聚焦](#) 來開發 ORR，並在您的帳戶和 AWS 組織中進行共用。
4. 找出要在其中執行 ORR 的一個工作負載。啟動前的工作負載或內部工作負載是理想的選擇。
5. 演練 ORR 檢查清單，並記下任何所探索的項目。如果採取緩解措施，那就可能接受探索。對於缺少緩解措施的任何探索，請將那些探索新增至項目的待辦清單中，然後在啟動前加以實作。
6. 隨著時間持續在 ORR 檢查清單中新增最佳實務和需求。

擁有 Enterprise Support 的支援客戶可向其技術客戶經理請求 [營運準備度審查研討會](#)。該研討會是一個交互式回溯會議，用於開發您自己的 ORR 檢查清單。

實作計劃的工作量：高。在組織中採用 ORR 實務需要高層和利益相關者的支持。使用貴組織提供的各方意見，來建立和更新檢查清單。

資源

相關的最佳實務：

- [OPS01-BP03 評估治理要求](#) – ORR 檢查清單原本就很適合用來管控需求。
- [OPS01-BP04 評估合規要求](#) – ORR 檢查清單中有時會包含合規需求。有些時候，它們會是獨立的程序。
- [OPS03-BP07 資源團隊適當](#) – 團隊能力是 ORR 需求的絕佳候選項。
- [OPS06-BP01 計畫變更失敗](#) – 啟動工作負載前，必須先建立回復或向前回復計劃。

- [OPS07-BP01 確保人員能力](#) – 若要支援工作負載，您必須具備所需的人員。
- [SEC01-BP03 識別和驗證控制目標](#) — 安全性控制目標是絕佳的 ORR 需求。
- [REL13-BP01 定義停機時間和資料遺失的復原目標](#) — 災難復原計劃是絕佳的 ORR 需求。
- [COST02-BP01 根據您的組織需求制定政策](#) — 成本管理政策適合包含在 ORR 檢查清單中。

相關文件：

- [AWS Control Tower - AWS Control Tower 中的防護機制](#)
- [AWS Well-Architected Tool - 自訂聚焦](#)
- [Adrian Hornsby 提供的營運準備度審查範本](#)
- [《營運準備度審查 \(ORR\)》白皮書](#)

相關影片：

- [AWS 支援 為您提供支援 | 建立有效的營運準備度審查 \(ORR\)](#)

相關範例：

- [營運準備度審查 \(ORR\) 聚焦範例](#)

相關服務：

- [AWS Config](#)
- [AWS Control Tower](#)
- [AWS Security Hub CSPM](#)
- [AWS Well-Architected Tool](#)

OPS07-BP03 使用執行手冊執执行程序

執行手冊是為了達成特定成果而記錄的程序。執行手冊由一系列可供遵循以完成某項工作的步驟組成。早在航空業早期，就已使用執行手冊。在雲端操作中，我們使用執行手冊來降低風險及達到預期成果。簡言之，執行手冊就是完成一項工作的檢查清單。

執行手冊是操作工作負載的重要組成部分。從新團隊成員入職到部署重大版本，執行手冊是經過編纂的流程，無論誰使用這些執行手冊，都能提供一致的成果。應該在中央位置發佈執行手冊，並隨著流程的

發展進行更新，因為更新執行手冊是變更管理流程的關鍵組成部分。它們還應該包括當發生問題時有關錯誤處理、工具、權限、異常以及向上呈報的指引。

隨著組織的成熟，開始自動化執行手冊。從簡短且經常使用的執行手冊開始。使用指令碼語言來自動化步驟或讓步驟更容易執行。當您自動化前幾個執行手冊時，將花費時間來自動化更複雜的執行手冊。隨著時間的推移，大多數執行手冊都應該以某種方式自動化。

預期成果：您的團隊擁有一系列用於執行工作負載任務的分步指南。執行手冊中包含預期成果、必要的工具和許可，以及錯誤處理指示。它們會集中存放 (版本控制系統)，並且經常更新。例如，執行手冊可讓您的團隊在應用程式警示、操作問題和計劃的生命週期事件期間監控、溝通和回應關鍵帳戶的 AWS Health 事件。

常見的反模式：

- 依靠記憶體來完成流程的每個步驟。
- 手動部署變更，無須檢查清單。
- 不同的團隊成員執行相同的過程，但具有不同的步驟或成果。
- 讓執行手冊脫離系統變更和自動化。

建立此最佳實務的優勢：

- 降低手動任務的錯誤率。
- 以一致的方式執行操作。
- 新的團隊成員可以更快地開始執行任務。
- 可以自動化執行手冊以減少辛勞。

未建立此最佳實務時的曝險等級：中

實作指引

執行手冊可以根據組織的成熟度等級採用多種形式。至少，其中應包含一個循序漸進的文字文件。應明確指出預期成果。清楚記錄必要的特殊權限或工具。如果發生問題，提供有關錯誤處理和呈報的詳細指引。列出執行手冊擁有者並將其發佈在中央位置。執行手冊被記錄下來之後，透過讓團隊中的其他人執行它來進行驗證。隨著程序的發展，請根據您的變更管理流程來更新執行手冊。

隨著組織的成熟，應自動化文字執行手冊。使用 [AWS Systems Manager Automation](#) 等服務，可以將純文字轉換為可針對工作負載執行的自動化功能。可以執行這些自動化功能以回應事件，減少維護工作

負載的操作負擔。AWS Systems Manager Automation 也提供低程式碼的[視覺化設計體驗](#)，以便更輕鬆地建立自動化執行手冊。

客戶範例

AnyCompany Retail 必須在軟體部署期間執行資料庫結構描述更新。雲端操作團隊與資料庫管理團隊合作，建立用於手動部署這些變更的執行手冊。執行手冊以檢查清單的形式列出流程中的每個步驟。它包括發生問題時關於錯誤處理的部分。他們在其內部 wiki 中發布該執行手冊以及其他執行手冊。雲端操作團隊計劃在未來的衝刺中自動化該執行手冊。

實作步驟

如果您沒有現有的文件儲存庫，版本控制儲存庫是開始建置執行手冊庫的好地方。可以使用 Markdown 來構建執行手冊。我們提供了一個執行手冊範本範例，您可以使用它來開始構建執行手冊。

```
# Runbook Title
## Runbook Info
| Runbook ID | Description | Tools Used | Special Permissions | Runbook Author | Last
Updated | Escalation POC |
|-----|-----|-----|-----|-----|-----|-----|
| RUN001 | What is this runbook for? What is the desired outcome? | Tools | Permissions
| Your Name | 2022-09-21 | Escalation Name |
## Steps
1. Step one
2. Step two
```

1. 如果您沒有現有的文件儲存庫或 wiki，請在版本控制系統中建立新的版本控制儲存庫。
2. 識別沒有執行手冊的流程。理想的流程是半定期執行，步驟數量少，並且具有低影響故障。
3. 在文件儲存庫中，使用範本建立新的 Markdown 草稿文件。填寫執行手冊標題和執行手冊資訊下的必填欄位。
4. 從第一個步驟開始，填寫執行手冊的「步驟」部分。
5. 將執行手冊交給團隊成員。讓他們使用執行手冊來驗證步驟。如果缺少某些內容或需要澄清，請更新執行手冊。
6. 將執行手冊發佈到您的內部文件存放區。發佈後，告知您的團隊和其他利益相關者。
7. 隨著時間的推移，您將建置執行手冊的程式庫。隨著程式庫的增長，開始努力自動化執行手冊。

實作計劃的工作量：低。執行手冊的最低標準是循序漸進的文字指南。自動化執行手冊可以增加實作工作量。

資源

相關的最佳實務：

- [OPS02-BP02 流程和程序已識別擁有者](#)
- [OPS07-BP04 使用程序手冊調查問題](#)
- [OPS10-BP01 使用程序進行事件、事故和問題管理](#)
- [OPS10-BP02 每個提醒建立一個程序](#)
- [OPS11-BP04 執行知識管理](#)

相關文件：

- [使用自動化程序手冊和執行手冊實現卓越的營運](#)
- [AWS Systems Manager：使用執行手冊](#)
- [適用於 AWS 大型遷移的遷移程序手冊 - 任務 4：改善您的遷移執行手冊](#)
- [使用 AWS Systems Manager Automation 執行手冊來解決操作任務](#)

相關影片：

- [AWS re:Invent 2019：執行手冊、事件報告和事件回應的 DIY 指南](#)
- [如何在 AWS 上自動執行 IT 操作 | Amazon Web Services](#)
- [將指令碼整合到 AWS Systems Manager](#)

相關範例：

- [Well-Architected 實驗室：使用程序手冊和執行手冊將操作自動化](#)
- [AWS 部落格文章：建立雲端自動化實務以實現卓越營運：AWS Managed Services 的最佳實務](#)
- [AWS Systems Manager：自動化演練](#)
- [AWS Systems Manager：從最新的快照執行手冊中還原根磁碟區](#)
- [使用 Jupyter 筆記本和 CloudTrail Lake 建置 AWS 事件回應執行手冊](#)
- [Gitlab - 執行手冊](#)
- [Rubix - 用於在 Jupyter 筆記本中構建執行手冊的 Python 庫](#)
- [使用文件建置器建立自訂執行手冊](#)

相關服務：

- [AWS Systems Manager Automation](#)

OPS07-BP04 使用程序手冊來調查問題

程序手冊是用來調查事件的逐步指南。事件發生時，我們會使用程序手冊來調查、確認影響範圍和找出根本原因。程序手冊適用於各種情況，從失敗的部署到安全性事故。在許多案例中，程序手冊可釐清根本原因，而執行手冊則用來緩解該根本原因。程序手冊是組織事件應變計劃的關鍵要素。

一個好的程序手冊有幾個關鍵功能。它循序漸進地引導使用者完成探索過程。從外到內思考，應該遵循哪些步驟來診斷事件？在程序手冊中明確定義程序手冊中是否需要特殊工具或更高權限。制定溝通計劃，向利益相關者通報調查進展情況，這非常關鍵。在無法確定根本原因的情況下，程序手冊應具有升級計劃。如果確定了根本原因，程序手冊應該指向說明如何解決問題的執行手冊。程序手冊應集中存放並定期維護。如果程序手冊用於特定提醒，請在提醒中為您的團隊提供指向程序手冊的指引。

隨著組織的成熟，會將您的程序手冊自動化。從涵蓋低風險事件的程序手冊開始。使用指令碼自動執行探索步驟。確認您有配套的執行手冊來減輕常見根本原因。

預期成果：您的組織擁有常見事件的程序手冊。程序手冊存放在中心位置，並可供您的團隊成員使用。程序手冊會經常更新。對於任何已知的根本原因，都會構建配套的執行手冊。

常見的反模式：

- 調查事件沒有標準方法。
- 團隊成員依賴肌肉記憶或機構知識來疑難排解失敗的部署。
- 新團隊成員學習如何透過試驗和錯誤來調查問題。
- 調查問題的最佳實務不會跨團隊共用。

建立此最佳實務的優勢：

- 程序手冊可加強您減輕事故的努力。
- 不同的團隊成員可以使用相同的程序手冊，以一致的方式識別根本原因。
- 您可以為已知的根本原因制定執行手冊，進而縮短復原時間。
- 程序手冊有助於團隊成員更快地開始做出貢獻。
- 團隊可以透過可重複的程序手冊擴展其程序。

未建立此最佳實務時的曝險等級：中

實作指引

建置和使用程序手冊的方式取決於組織的成熟度。如果您是雲端新手，請在中央文件儲存庫中以文字形式建立程序手冊。隨著組織的成熟，程序手冊可以使用 Python 之類的指令碼語言進行半自動化。這些指令碼可以在 Jupyter 筆記本內部運行，以加快發現速度。進階組織具有完全自動化的程序手冊，可解決使用執行手冊自動修復的常見問題。

列出工作負載發生的常見事件，開始建置程序手冊。為低風險並且根本原因已縮小到幾個問題的事件選擇程序手冊以開始。在您擁有更簡單案例的程序手冊之後，請轉到風險較高的案例或根本原因尚不明確的案例。

隨著組織的成熟，應自動化文字程序手冊。使用 [AWS Systems Manager Automation](#) 等服務，可以將純文字轉換為自動化功能。可以針對您的工作負載執行這些自動化，以加快調查速度。可以啟動這些自動化以回應事件，減少發現和解決事故的平均時間。

客戶可以使用 [AWS Systems Manager Incident Manager](#) 來回應事故。此服務提供單一介面來分類事故、在發現和緩解期間通知利益相關者，並在整個事故中進行協同合作。它使用 AWS Systems Manager Automation 來加速偵測和復原。

客戶範例

生產事故影響了 AnyCompany Retail。隨時待命的工程師使用程序手冊來調查問題。隨著他們逐步完成這些步驟，他們會讓程序手冊中確定的關鍵利益相關者了解最新狀況。工程師將根本原因確定為後端服務中的競爭條件。使用執行手冊，工程師重新啟動了服務，使 AnyCompany Retail 重新上線。

實作步驟

如果您沒有現有的文件儲存庫，建議您為程序手冊庫建立版本控制儲存庫。您可以使用 Markdown 構建程序手冊，它可與大多數程序手冊自動化系統相容。如果您是從頭開始，請使用下列範例程序手冊範本。

```
# Playbook Title
## Playbook Info
| Playbook ID | Description | Tools Used | Special Permissions | Playbook Author | Last
Updated | Escalation POC | Stakeholders | Communication Plan |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| RUN001 | What is this playbook for? What incident is it used for? | Tools |
Permissions | Your Name | 2022-09-21 | Escalation Name | Stakeholder Name | How will
updates be communicated during the investigation? |
```

Steps

1. Step one
2. Step two

1. 如果您沒有現有的文件儲存庫或 wiki，請在版本控制系統中為您的程序手冊建立新的版本控制儲存庫。
2. 找出需要調查的常見問題。這種情況應該是根本原因僅限於少數問題且解決方案風險較低。
3. 使用 Markdown 範本，填寫 [程序手冊名稱] 部分和 [程序手冊資訊] 下方的欄位。
4. 填寫疑難排解步驟。盡可能明確要執行哪些操作或應該調查哪些領域。
5. 將程序手冊交給團隊成員，讓他們仔細閱讀以驗證。如果有任何遺漏或不清楚的內容，請更新程序手冊。
6. 在文件儲存庫中發佈程序手冊，並通知您的團隊和任何利益相關者。
7. 此程序手冊庫會隨著您新增更多程序手冊而增加。當您擁有多本程序手冊之後，請使用 AWS Systems Manager Automation 等工具開始進行自動化，讓自動化和程序手冊保持同步。

實作計劃的工作量：低。您的程序手冊應該是存放在中央位置的文字文件。更成熟的組織將推進程序手冊自動化。

資源

相關的最佳實務：

- [OPS02-BP02 流程和程序已識別擁有者](#)
- [OPS07-BP03 使用執行手冊執执行程序](#)
- [OPS10-BP01 使用程序進行事件、事故和問題管理](#)
- [OPS10-BP02 每個提醒建立一個程序](#)
- [OPS11-BP04 執行知識管理](#)

相關文件：

- [使用自動化程序手冊和執行手冊實現卓越的營運](#)
- [AWS Systems Manager：使用執行手冊](#)
- [使用 AWS Systems Manager Automation 執行手冊來解決操作任務](#)

相關影片：

- [AWS re:Invent 2019：執行手冊、事件報告和事件回應的 DIY 指南 \(SEC318-R1\)](#)
- [AWS Systems Manager Incident Manager - AWS 虛擬研討會](#)
- [將指令碼整合到 AWS Systems Manager](#)

相關範例：

- [AWS 客戶程序手冊架構](#)
- [AWS Systems Manager：自動化演練](#)
- [使用 Jupyter 筆記本和 CloudTrail Lake 建置 AWS 事件回應執行手冊](#)
- [Rubix - 用於在 Jupyter 筆記本中構建執行手冊的 Python 庫](#)
- [使用文件建置器建立自訂執行手冊](#)

相關服務：

- [AWS Systems Manager Automation](#)
- [AWS Systems Manager Incident Manager](#)

OPS07-BP05 做出部署系統和變更的明智決策

為成功和失敗的工作負載變更建立程序。事前剖析是一種演練，團隊可藉此模擬失敗，制定緩解策略。使用事前剖析可預測失敗並適時建立程序。評估將變更部署到您的工作負載的優點和風險。確認所有變更都符合管控。

預期成果：

- 您在將變更部署到您的工作負載時做出明智決策。
- 變更符合管控。

常見的反模式：

- 將變更部署到我們的工作負載，而沒有處理失敗部署的程序。
- 對不符合管控要求的生產環境進行變更。
- 部署新版本的工作負載，而未建立資源使用率的基準。

建立此最佳實務的優勢：

- 您對工作負載的失敗變更已做好準備。
- 變更您的工作負載符合管控政策。

未建立此最佳實務時的曝險等級：低

實作指引

使用事前剖析來開發失敗變更的程序。記載失敗變更的程序。確定所有變更都符合管控。評估將變更部署到您的工作負載的優點和風險。

客戶範例

AnyCompany Retail 定期執行事前剖析來驗證他們失敗變更的程序。他們在共用 Wiki 中記載程序並且頻繁更新。所有變更都符合管控要求。

實作步驟

1. 在將變更部署到您的工作負載時做出明智決策。建立及審核成功部署的準則。開發會啟動變更回復的情境或準則。權衡部署變更的優點與失敗變更的風險。
2. 確認所有變更都符合管控政策。
3. 使用事前剖析為失敗變更進行規劃並且記載緩解策略。執行桌上模擬演練來建立失敗變更的模型，並且驗證回復程序。

實作計劃的工作量：中。實作事前剖析的實務需要貴組織利益相關者的協調和努力

資源

相關的最佳實務：

- [OPS01-BP03 評估治理要求](#) - 管控要求是判斷是否部署變更的關鍵因素。
- [OPS06-BP01 計畫變更失敗](#) - 建立計畫來緩解失敗的部署並且使用事前剖析來進行驗證。
- [OPS06-BP02 測試部署](#) - 每個軟體變更都應該在部署之前先適當的進行測試，以便在生產中減少缺陷。
- [OPS07-BP01 確保人員能力](#) - 擁有支援工作負載的足夠受過培訓的人員，對於為部署系統變更做出明智決策相當重要。

相關文件：

- [Amazon Web Services：風險與合規](#)

- [AWS 共同責任模型](#)
- [AWS 雲端 中的管控：敏捷性和安全性之間的正確平衡](#)

OPS07-BP06 建立生產工作負載的支援計劃

針對您的生產工作負載所依賴的任何軟體和服務啟用支援。根據您的生產服務層級需求選取適當的支援等級。這些相依性的支援計劃的存在有其必要性，以便應對服務中斷或軟體問題。記錄支援計劃，以及如何要求所有服務和軟體供應商的支援。實作相關機制以確認支援的聯絡窗口是最新的。

預期成果：

- 為生產工作負載所依賴的軟體和服務實作支援計劃。
- 根據服務層級需求選擇適當的支援計劃。
- 記錄支援計劃、支援等級，以及如何要求支援。

常見的反模式：

- 您沒有主要軟體供應商的支援計劃。您的工作負載因此受到影響，且您無法加速進行修正，或及時獲得供應商提供的更新。
- 擔任軟體供應商主要聯絡窗口的開發人員已離開公司。您無法直接聯繫供應商支援人員。您必須花時間研究及瀏覽通用聯絡系統，因此必要時的回應將更為耗時。
- 軟體供應商發生生產中斷。目前沒有文件說明如何提出支援案例。

建立此最佳實務的優勢：

- 透過適當的支援等級，您將可在必要的時間範圍內獲得回應以滿足服務層級需求。
- 受支援的客戶可在遇到生產問題時加以呈報。
- 軟體和服務供應商可在事件發生期間協助進行疑難排解。

未建立此最佳實務時的曝險等級：低

實作指引

針對您的生產工作負載所依賴的任何軟體和服務供應商啟用支援計劃。設定適當的支援計劃以滿足服務層級需求。對於 AWS 客戶而言，這意味著在任何有生產工作負載的帳戶上啟動 AWS Business Support 或更高等級。定期與支援供應商聯繫，取得關於支援優惠、程序和聯絡人的更新。記錄如何要

求軟體和服務供應商的支援，包括如何在中斷發生時加以呈報。實作相關機制以保有最新的支援聯絡資料。

客戶範例

在 AnyCompany Retail，所有商業軟體和服務相依性都有支援計劃。例如，他們在所有具有生產工作負載的帳戶上啟動了 AWS Enterprise Support。任何開發人員都可在問題發生時提出支援案例。有 Wiki 頁面提供了相關資訊說明如何要求支援、應通知誰，以及加速處理案例的最佳實務為何。

實作步驟

1. 與組織中的利益相關者合作，識別您的工作負載所依賴的軟體和服務供應商。記錄這些相依性。
2. 確認工作負載的服務層級需求。選取相對應的支援計劃。
3. 針對商業軟體和服務，與供應商共同建立支援計劃。
 - a. 為所有生產帳戶訂閱 AWS Business Support 或更高等級可獲得 AWS 支援 更快的回應時間，極力建議這麼做。如果您沒有付費支援，則必須有處理問題的行動計劃，而這需要 AWS 支援 的協助。AWS 支援 提供了多種工具和技術、人員及方案，旨在主動協助您最佳化效能、降低成本和加快創新速度。此外，AWS Business Support 提供額外的權益，包括透過 API 存取 AWS Trusted Advisor 和 AWS Health，以便與您的系統進程式設計整合，以及其他存取方法，例如 AWS 管理主控台 和 Amazon EventBridge 管道。
4. 在您的知識管理工具中記錄支援計劃。納入如何要求支援、在提出支援案例時應通知誰，以及在事件發生時如何加以呈報等資訊。任何人在得知支援程序或聯絡資料有所變更時，都可以利用 Wiki 這項機制對文件進行必要的更新。

實作計劃的工作量：低。大部分的軟體和服務供應商都提供選擇加入支援計劃。在您的知識管理系統上記錄並分享支援最佳實務，可確保您的團隊知道在生產問題發生時應如何因應。

資源

相關的最佳實務：

- [OPS02-BP02 流程和程序已識別擁有者](#)

相關文件：

- [AWS 支援 計劃](#)

相關服務：

- [AWS 商業支援](#)
- [AWS 企業支援](#)

操作

問題

- [OPS 8. 如何在組織中利用工作負載可觀測性？](#)
- [OPS 9. 如何了解營運狀況？](#)
- [OPS 10. 如何管理工作負載和營運事件？](#)

OPS 8. 如何在組織中利用工作負載可觀測性？

利用可觀測性確保最佳的工作負載運作狀況。利用相關指標、日誌和追蹤，全面掌握工作負載效能並有效解決問題。

最佳實務

- [OPS08-BP01 分析工作負載指標](#)
- [OPS08-BP02 分析工作負載日誌](#)
- [OPS08-BP03 分析工作負載追蹤](#)
- [OPS08-BP04 建立可執行的提醒](#)
- [OPS08-BP05 建立儀表板](#)

OPS08-BP01 分析工作負載指標

實作應用程式遙測之後，請定期分析收集到的指標。雖然延遲、請求、錯誤和容量 (或配額) 可提供深入了解系統效能的洞見，但務必將檢閱業務成果指標視為優先事項。這樣做可確保您所做的資料驅動決策符合您的業務目標。

預期成果：獲得深入工作負載效能的精確洞見，有助於做出資料驅動的決策，確保與業務目標保持一致。

常見的反模式：

- 單獨分析指標，未能考慮到其對業務目標的影響。
- 過度依賴技術指標，而輕忽業務指標。

- 未能時常檢閱指標，而錯失即時決策的機會。

建立此最佳實務的優勢：

- 增進對於技術表現與業務成果之間相互關聯的了解。
- 透過即時資料改善了決策過程。
- 主動識別並緩解問題，不讓問題影響業務成果。

未建立此最佳實務時的曝險等級：中

實作指引

利用 Amazon 等工具 CloudWatch 執行指標分析。CloudWatch 異常偵測和 Amazon DevOpsGuru 等 AWS 服務可用來偵測異常，特別是靜態閾值未知或行為模式更適合異常偵測時。

實作步驟

1. 分析與檢閱：定期檢閱和解讀您的工作負載指標。
 - a. 將業務成果指標視為優先於純粹技術指標的事項。
 - b. 了解資料中峰值、下降或模式的重要性。
2. 使用 Amazon CloudWatch：使用 Amazon CloudWatch 進行集中式檢視和深入分析。
 - a. 設定 CloudWatch 儀表板以視覺化您的指標，並隨時間進行比較。
 - b. 使用 [中的百分位數 CloudWatch](#) 來取得指標分佈的清晰檢視，這有助於定義SLAs和了解異常值。
 - c. 設定 [CloudWatch 異常偵測](#) 以識別異常模式，而不必依賴靜態閾值。
 - d. 實作 [CloudWatch 跨帳戶可觀測性](#)，以監控和疑難排解跨區域內多個帳戶的應用程式。
 - e. 使用 [CloudWatch Metric Insights](#) 查詢和分析帳戶和區域的指標資料，識別趨勢和異常。
 - f. 套用 [CloudWatch 指標數學](#) 來轉換、彙總或執行指標的計算，以取得更深入的洞見。
3. 使用 Amazon DevOpsGuru：將 [Amazon DevOpsGuru](#) 納入其機器學習增強型異常偵測，以識別無伺服器應用程式的早期操作問題跡象，並在影響客戶之前對其進行修復。
4. 根據洞見最佳化：根據您的指標分析做出明智的決策，以調整和改善您的工作負載。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS04-BP01 識別關鍵績效指標](#)
- [OPS04-BP02 實作應用程式遙測](#)

相關文件：

- [The Wheel 部落格 - 強調持續檢閱指標的重要性](#)
- [百分位數很重要](#)
- [使用 AWS Cost Anomaly Detection](#)
- [CloudWatch 跨帳戶可觀測性](#)
- [使用 CloudWatch Metrics Insights 查詢您的指標](#)

相關影片：

- [在 Amazon 中啟用跨帳戶可觀測性 CloudWatch](#)
- [Amazon DevOpsGuru 簡介](#)
- [使用 持續分析指標 AWS Cost Anomaly Detection](#)

相關範例：

- [一個可觀測性研討會](#)
- [AIOps使用 Amazon DevOpsGuru 取得操作洞見](#)

OPS08-BP02 分析工作負載日誌

定期分析工作負載日誌對於深入了解應用程式的操作層面至關重要。藉由有效率地篩選、視覺化和解讀日誌資料，可持續最佳化應用程式效能和安全。

預期成果：從徹底的日誌分析中獲得深入應用程式行為和操作的豐富洞見，以確保主動偵測和緩解問題。

常見的反模式：

- 忽略日誌分析，直到出現嚴重問題。
- 沒有使用可用於日誌分析的完整工具套件，錯過了關鍵洞見。
- 只倚賴手動檢閱日誌，而未利用自動化和查詢功能。

建立此最佳實務的優勢：

- 主動找出操作瓶頸、安全威脅及其他潛在問題。
- 有效利用日誌資料，以實現持續的應用程式最佳化。
- 加強對應用程式行為的理解，幫助偵錯和疑難排解。

未建立此最佳實務時的曝險等級：中

實作指引

[Amazon CloudWatch Logs](#) 是日誌分析的強大工具。CloudWatch Logs Insights 和 Contributor Insights 等整合功能，讓從日誌中擷取有意義的資訊的過程變得直覺且有效。

實作步驟

1. 設定 CloudWatch 日誌：設定應用程式和服務將日誌傳送至 CloudWatch 日誌。
2. 使用日誌異常偵測：利用 [Amazon CloudWatch Logs 異常偵測](#) 自動識別並提醒異常日誌模式。此工具可協助您主動管理日誌中的異常，並儘早偵測潛在問題。
3. 設定 CloudWatch Logs Insights：使用 [CloudWatch Logs Insights](#) 以互動方式搜尋和分析您的日誌資料。
 - a. 製作查詢以找出模式、視覺化日誌資料，並產生可付諸行動的洞見。
 - b. 使用 [CloudWatch Logs Insights 模式分析](#) 來分析和視覺化常用日誌模式。此功能可協助您了解日誌資料中常見的操作趨勢和潛在的異常值。
 - c. 使用 [CloudWatch Logs compare \(diff \)](#) 在不同時段或不同日誌群組之間執行差異分析。使用此功能可精確找出變更，並評估其對系統效能或行為的影響。
4. 使用 Live Tail 即時監控日誌：使用 [Amazon CloudWatch Logs Live Tail](#) 即時檢視日誌資料。您可以在應用程式的操作活動發生時進行主動監控，以便立即掌握系統效能和潛在問題。
5. 利用 Contributor Insights：使用 [CloudWatch Contributor Insights](#) 來識別 IP 地址或使用者代理等高基度維度的熱門發言者。
6. 實作 CloudWatch 日誌指標篩選條件：設定 [CloudWatch 日誌指標篩選條件](#)，將日誌資料轉換為可操作的指標。如此您就能設定警報或進一步分析模式。
7. 實作 [CloudWatch 跨帳戶可觀測性](#)：監控和疑難排解跨區域內多個帳戶的應用程式。
8. 定期檢閱和改進：定期檢閱您的日誌分析策略，以擷取所有相關資訊並持續最佳化應用程式效能。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS04-BP01 識別關鍵績效指標](#)
- [OPS04-BP02 實作應用程式遙測](#)
- [OPS08-BP01 分析工作負載指標](#)

相關文件：

- [使用 Logs Insights 分析 CloudWatch 日誌資料](#)
- [使用 CloudWatch 貢獻者洞察](#)
- [建立和管理 CloudWatch 日誌指標篩選條件](#)

相關影片：

- [使用 Logs Insights 分析 CloudWatch 日誌資料](#)
- [使用 CloudWatch 貢獻者洞察分析高基數資料](#)

相關範例：

- [CloudWatch 記錄範例查詢](#)
- [一個可觀測性研討會](#)

OPS08-BP03 分析工作負載追蹤

分析追蹤資料對於實現應用程式營運歷程的全面檢視至關重要。透過視覺化和了解各種不同元件之間的互動，就能微調效能、找出瓶頸，並且增強使用者體驗。

預期成果：清楚掌握應用程式的分散式操作，就能更快解決問題並增強使用者體驗。

常見的反模式：

- 忽略追蹤資料，只依賴日誌和指標。
- 不會將追蹤資料與相關日誌建立關聯。
- 忽略從追蹤產生的指標，如延遲和故障率。

建立此最佳實務的優勢：

- 改善故障診斷並減少解決的平均時間（MTTR）。
- 深入了解依賴性及其影響。
- 快速識別和糾正效能問題。
- 利用追蹤衍生的指標制定明智的決策。
- 透過最佳化元件互動改善使用者體驗。

未建立此最佳實務時的曝險等級：中

實作指引

[AWS X-Ray](#) 提供了全方位的追蹤資料分析套件，能讓您深入了解服務互動的各個層面、監控使用者活動，以及偵測效能問題。ServiceLens、X-Ray Insights、X-Ray Analytics 和 Amazon DevOpsGuru 等功能可增強從追蹤資料衍生的可操作洞察深度。

實作步驟

下列步驟提供結構化方法，可有效使用 AWS 服務實作追蹤資料分析：

1. 整合 AWS X-Ray：確保 X-Ray 與您的應用程式整合，以擷取追蹤資料。
2. 分析 X-Ray 指標：深入研究 X-Ray 追蹤衍生的指標，例如延遲、請求率、錯誤率和回應時間分佈，使用[服務地圖](#)來監控應用程式運作狀態。
3. 使用 ServiceLens：利用[ServiceLens地圖](#)增強服務和應用程式的可觀測性。如此就能將追蹤、指標、日誌、警報和其他運作狀況資訊整合在一起檢視。
4. 啟用 X-Ray Insights：
 - a. 開啟 [X-Ray Insights](#)，以自動偵測追蹤中的異常。
 - b. 檢查洞見以找出明確的模式並確定根本原因，例如故障率或延遲增加。
 - c. 請參考 Insights 時間軸，依時間順序查看所偵測到問題的分析。
5. 使用 X-Ray Analytics：[X-Ray Analytics](#) 可讓您徹底探索追蹤資料、精確定位模式並擷取洞見。
6. 使用 X-Ray 中的群組：在 X-Ray 中建立群組，即可根據如高延遲等條件篩選追蹤，以進行更針對性的分析。
7. 合併 Amazon DevOpsGuru：讓 [Amazon DevOpsGuru](#) 受益於機器學習模型，以找出追蹤中的操作異常。

8. 使用 CloudWatch Synthetics：使用 [CloudWatch Synthetics](#) 建立 Canary，以持續監控您的端點和工作流程。這些 Canary 可與 X-Ray 整合，以提供追蹤資料，用來對要測試的應用程式進行深入分析。
9. 使用實際使用者監控（RUM）：使用 [AWS X-Ray](#) 和 [CloudWatch RUM](#)，您可以從應用程式的最終使用者開始透過下游 AWS 受管服務分析和偵錯請求路徑。這樣做有助於找出影響最終使用者的延遲趨勢和錯誤。
- 10 與日誌建立關聯：將[追蹤資料與 X-Ray 追蹤檢視中的相關日誌](#)建立關聯，以深入了解應用程式行為。如此可讓您檢視與追蹤的交易直接相關的日誌事件。
- 11 實作[CloudWatch 跨帳戶可觀測性](#)：監控和疑難排解跨區域內多個帳戶的應用程式。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS08-BP01 分析工作負載指標](#)
- [OPS08-BP02 分析工作負載日誌](#)

相關文件：

- [使用 ServiceLens 監控應用程式運作狀態](#)
- [使用 X-Ray Analytics 探索追蹤資料](#)
- [使用 X-Ray Insights 偵測追蹤中的異常狀況](#)
- [使用 CloudWatch Synthetics 持續監控](#)

相關影片：

- [使用 Amazon CloudWatch Synthetics & 分析和偵錯應用程式 AWS X-Ray](#)
- [使用 AWS X-Ray Insights](#)

相關範例：

- [一個可觀測性研討會](#)
- [使用 實作 X-Ray AWS Lambda](#)

- [CloudWatch Synthetics Canary 範本](#)

OPS08-BP04 建立可執行的提醒

及時偵測並回應您的應用程式行為中的偏差至關重要。尤其重要的是要了解基於關鍵績效指標 (KPI) 的結果何時處於危險之中，或者何時出現意外異常。以 KPI 為基礎的提醒可確保您收到的訊號直接與業務或營運影響產生關係。這種可採取動作的提醒方法可促進主動回應，並有助於維持系統效能與可靠性。

預期成果：接收及時、相關且可行的提醒，以便快速識別和緩解潛在問題，尤其是在 KPI 結果面臨風險時。

常見的反模式：

- 設定太多非嚴重性提醒會導致提醒疲勞。
- 不會根據 KPI 來排定提醒的優先順序，因此難以了解問題的業務影響。
- 忽視解決根本原因導致同一問題的重複提醒。

建立此最佳實務的優勢：

- 透過專注於可操作且相關的提醒來減少提醒疲勞。
- 透過主動偵測和緩解問題，改善系統運作時間和可靠性。
- 透過與熱門的提醒和通訊工具整合，強化團隊協同作業並加快解決問題的速度。

未建立此最佳實務時的曝險等級：高

實作指引

若要建立有效的提醒機制，使用指標、日誌和追蹤資料至關重要，其會在基於 KPI 的結果出現風險或偵測到異常時進行標記。

實作步驟

1. 確定關鍵績效指標 (KPI)：確定應用程式的 KPI。提醒應與這些關鍵績效指標相關聯，以準確反映業務影響。
2. 實作異常偵測：
 - 使用 Amazon CloudWatch 異常偵測：設定 [Amazon CloudWatch 異常偵測](#) 以自動偵測異常模式，這可協助您僅針對真正的異常產生提醒。

- 使用 AWS X-Ray Insights :
 - a. 設定 [X-Ray Insights](#) 以偵測追蹤資料中的異常。
 - b. 設定 [X-Ray Insights 的通知](#)，以便在偵測到問題時收到提醒。
- 與 Amazon DevOps Guru 整合 :
 - a. 利用 [Amazon DevOps Guru](#) 的機器學習功能，偵測現有資料的操作異常情況。
 - b. 導覽至 DevOps Guru 中的[通知設定](#)，以設定異常提醒。
- 3. 實作可執行的提醒：設計提醒，為立即採取行動提供足夠資訊。
 - 1. [使用 Amazon EventBridge 規則監控 AWS Health 事件](#)，或以程式設計方式與 AWS Health API 整合，以便在您收到 AWS Health 事件時自動執行動作。這些動作可以是一般動作 (例如將所有規劃的生命週期事件訊息傳送至聊天介面) 或是特定動作 (例如在 IT 服務管理工具中啟動工作流程)。
 - 4. 減少提醒疲勞：將非嚴重性提醒降至最低。當團隊對眾多微不足道的提醒感到不知所措時，他們可能會失去對重大問題的監督，從而降低提醒機制的整體有效性。
 - 5. 設定複合警示：使用 [Amazon CloudWatch 複合警示](#)來合併多個警示。
 - 6. 與提醒工具整合：整合諸如 [Ops Genie](#) 和 [PagerDuty](#) 等工具。
 - 7. 採用聊天應用程式中的 Amazon Q Developer：整合[聊天應用程式中的 Amazon Q Developer](#)，以便將警示轉送至 Amazon Chime、Microsoft Teams 和 Slack。
 - 8. 基於日誌的提醒：使用 CloudWatch 中的[日誌指標篩選器](#)，根據特定的日誌事件建立警示。
 - 9. 審查並反覆：定期重新檢視並調整提醒組態。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS04-BP01 識別關鍵績效指標](#)
- [OPS04-BP02 實作應用程式遙測](#)
- [OPS04-BP03 實作使用者體驗遙測](#)
- [OPS04-BP04 實作相依性遙測](#)
- [OPS04-BP05 實作分散式追蹤](#)
- [OPS08-BP01 分析工作負載指標](#)
- [OPS08-BP02 分析工作負載日誌](#)

- [OPS08-BP03 分析工作負載追蹤](#)

相關文件：

- [使用 Amazon CloudWatch 警示](#)
- [建立複合警示](#)
- [根據異常偵測建立 CloudWatch 警示](#)
- [DevOps Guru 通知](#)
- [X-ray Insights 通知](#)
- [透過互動式 ChatOps 監控和操作您的 AWS 資源並進行疑難排解](#)
- [Amazon CloudWatch 整合指南 | PagerDuty](#)
- [整合 Opsgenie 與 Amazon CloudWatch](#)

相關影片：

- [在 Amazon CloudWatch 中建立複合警示](#)
- [聊天應用程式中的 Amazon Q Developer 概觀](#)
- [聊天應用程式中的 Amazon Q Developer 中的 AWS On Air ft. 可變命令](#)

相關範例：

- [使用 Amazon CloudWatch 在雲端進行警示、事件管理和修復](#)
- [教學課程：建立將通知傳送至聊天應用程式中的 Amazon Q Developer 的 Amazon EventBridge 規則](#)
- [一個可觀測性研討會](#)

OPS08-BP05 建立儀表板

儀表板是以人為本的工作負載遙測資料檢視。雖然它們提供了重要的視覺介面，但它們不應該取代警報機制，而是補充它們。經過精心打造的儀表板不僅能提供快速了解系統運作狀況和效能的洞見，還能對利益相關者呈現有關業務成果和問題影響層面的即時資訊。

預期成果：

使用視覺呈現的方式，提供清楚、深入系統與業務運作狀況且可付諸行動的洞見。

常見的反模式：

- 包含太多指標、過於複雜的儀表板。
- 仰賴沒有異常偵測提醒的儀表板。
- 儀表板未隨著工作負載發展而更新。

建立此最佳實務的優勢：

- 立即掌握關鍵系統指標和 KPI。
- 增強利益相關者的溝通和理解。
- 快速深入洞察操作問題的影響層面。

未建立此最佳實務時的風險等級：中

實作指引

以業務為中心的儀表板

專為業務 KPI 量身打造的儀表板，可與更廣泛的利益相關者進行互動。儘管這些人可能對系統指標不感興趣，但他們熱衷於了解這些數字的業務含義。以業務為中心的儀表板可確保所有受監控和分析的技術和營運指標都與總體業務目標保持同步。這種一致性提供了清晰度，確保每個人在什麼重要以及什麼不重要的問題上意見一致。此外，突出顯示業務 KPI 的儀表板往往更具可操作性。利益相關者可以快速了解營運的運作狀態、需要注意的領域以及對業務成果的潛在影響。

考慮到這一點，在建立儀表板時，請確保技術指標和業務 KPI 之間保持平衡。兩者都至關重要，但兩者迎合不同的受眾。在理想情況下，您應有能夠提供全方位視角儀表板，以便深入掌握系統運作狀況與效能，同時也要強調關鍵業務成果及其影響。

Amazon CloudWatch 儀表板是 CloudWatch 主控台中可自訂的首頁，可讓您在單一檢視中監控資源，甚至是分散在不同的 AWS 區域和帳戶中的那些資源。

實作步驟

1. 建立基本儀表板：[在 CloudWatch 中建立新儀表板](#)，並為其提供描述性名稱。
2. 使用 Markdown 小工具：在深入研究指標之前，請[使用 Markdown 小工具](#)在儀表板頂端新增文字內容。此內容應說明儀表板涵蓋的內容、所呈現指標的重要性，還可以包含其他儀表板和疑難排解工具的連結。

3. 建立儀表板變數：在適當位置[合併儀表板變數](#)，以允許動態且靈活的儀表板檢視。
4. 建立儀表板小工具：[新增儀表板小工具](#)以便將應用程式產生的各種不同指標視覺化，並調整這些小工具以便有效呈現系統運作狀況和業務成果。
5. Log Insights 查詢：利用 [CloudWatch Log Insights](#) 從日誌中導出可操作的指標，並在儀表板上顯示這些洞見。
6. 設定警示：將 [CloudWatch 警示](#) 整合到儀表板中，以便快速查看違反其閾值的任何指標。
7. 使用 Contributor Insights：整合 [CloudWatch Contributor Insights](#) 以分析高基數欄位，並更清楚地了解資源的主要貢獻者。
8. 設計自訂小工具：對於標準小工具未滿足的特定需求，請考慮建立[自訂小工具](#)。這些小工具可從各種資料來源中提取資料，或以獨特的方式呈現資料。
9. 使用 AWS Health：AWS Health 是 AWS 雲端 資源運作狀態的權威資訊來源。使用現成的 [AWS Health 儀表板](#)，或使用您自己的儀表板和工具中的 AWS Health 資料，以便擁有正確的資訊來做出明智的決策。
10. 反覆執行並改進：隨著應用程式發展，請定期重新檢視您的儀表板，以確保其相關性。

資源

相關的最佳實務：

- [OPS04-BP01 識別關鍵績效指標](#)
- [OPS08-BP01 分析工作負載指標](#)
- [OPS08-BP02 分析工作負載日誌](#)
- [OPS08-BP03 分析工作負載追蹤](#)
- [OPS08-BP04 建立可執行的提醒](#)

相關文件：

- [建置用於檢視營運狀況的儀表板](#)
- [使用 Amazon CloudWatch 儀表板](#)

相關影片：

- [建立跨帳戶和跨區域 CloudWatch 儀表板](#)
- [AWS re:Invent 2021 - 透過 AWS 雲端 了解企業 \(營運儀表板\)](#)

相關範例：

- [一個可觀測性研討會](#)
- [使用 Amazon CloudWatch 監控應用程式](#)
- [AWS Health 事件智慧儀表板和洞見](#)
- [使用 Amazon Managed Grafana 視覺化 AWS Health 事件](#)

OPS 9. 如何了解營運狀況？

定義、擷取和分析營運指標，掌握營運事件，以便採取適當行動。

最佳實務

- [OPS09-BP01 使用指標衡量營運目標與 KPI](#)
- [OPS09-BP02 傳達狀態和趨勢以確實掌控營運狀況](#)
- [OPS09-BP03 審核營運指標並優先改進](#)

OPS09-BP01 使用指標衡量營運目標與 KPI

從您的組織取得定義營運成功的目標和 KPI，並決定反映這些目標的指標。設定基準做為參考點，並定期重新評估。制定機制以便從團隊收集這些指標以進行評估。[DevOps Research and Assessment \(DORA\)](#) 指標提供了衡量軟體交付的 DevOps 實務進度的常用方法。

預期成果：

- 組織會發布並分享營運團隊的目標和 KPI。
- 您負責建立反映這些 KPI 的指標。範例可能包括：
 - 票證佇列深度或票證平均存留時間
 - 依問題類型分組的票證計數
 - 處理問題所花的時間，無論是否有標準作業程序 (SOP)
 - 從失敗的程式碼推送復原所花的時間長度
 - 通話音量

常見的反模式：

- 錯過部署期限，因為開發人員須分心處理疑難排解工作。開發團隊要求更多人力，但無法提出確切需要的人力數量，因為無法衡量被佔用的時間。

- 設立了 1 級服務台來處理使用者通話。經過一段時間後，加入了更多工作負載，但並沒有分配更多人員給 1 級服務台。客戶滿意度受到通話時間增加及問題未解決的時間拉長影響而下降，但管理層看不到這些現象的指標，未能採取任何行動。
- 有問題的工作負載已交由另一個營運團隊進行維護。與其他工作負載不同的是，並未針對這個新工作負載提供適當的文件和執行手冊。因此，團隊花費更長的時間進行疑難排解和解決失敗情況。然而，沒有任何指標記載此情況，因此無法明確究責。

建立此最佳實務的優勢：只要工作負載監控顯示我們應用程式和服務的狀態，監控營運團隊就可讓擁有者深入了解這些工作負載取用者之間的變化，例如業務需求轉變。藉由建立能夠反映營運狀態的指標來衡量這些團隊的效用，並依據業務目標進行評估。指標可突顯支援問題，或識別何時發生偏離服務層級目標的情形。

未建立此最佳實務時的曝險等級：中

實作指引

安排時間與企業領導者和利害關係人一起確定服務未來的整體目標。確定各個不同營運團隊應負責的任務，以及能夠應對哪些挑戰。使用這些來集思廣益，找出能夠反映這些營運目標的關鍵績效指標 (KPI)。這些目標可能包括客戶滿意度、從形成功能概念到部署的時間、平均問題解決時間，以及成本效率。

從 KPI 中找出最能反映這些目標的資料指標和來源。客戶滿意度可能由各種不同的指標組合而成，例如通話等待或回應時間、滿意度分數，以及提出的問題類型。部署時間可能是測試和部署，加上任何需要新增的部署後修正所需時間的總和。顯示不同類型的問題所花費時間 (或是這些問題的計數) 的統計資料，可提供一個切入視角，以了解需要針對性處理的地方。

資源

相關文件：

- [Quick - 使用 KPI](#)
- [Amazon CloudWatch - 使用指標](#)
- [建置儀表板](#)
- [如何使用 KPI 儀表板追蹤成本最佳化 KPI](#)
- [AWS DevOps 指引](#)

相關範例：

- [使用原生 AWS 監控和可觀測性工具來監控軟體交付的效能](#)
- [利用 DORA 指標在部署速度與穩定性之間取得平衡](#)
- [金融服務產業中的 MLOps 營運指標範例](#)
- [如何使用 KPI 儀表板追蹤成本最佳化 KPI](#)

OPS09-BP02 傳達狀態和趨勢以確實掌控營運狀況

您須了解營運狀態及趨勢方向，以確定成果何時可能存在風險、是否可支援新增的工作，或是變更對您的團隊造成的影響。營運事件發生時，有提供使用者和營運團隊參考資訊的狀態頁面，就能減輕溝通管道的壓力，並有效傳播資訊。

預期成果：

- 主管對團隊處理的通話量類型和正在進行的工作 (例如部署) 可以一目瞭然。
- 發生影響擴及正常營運的情況時，利益相關者和使用者社群就會收到警示。
- 組織領導階層和利益相關者可查看狀態頁面以便回應警示或影響，並且獲得有關營運事件的資訊，例如聯絡窗口、票證資訊及預估的復原時間。
- 領導階層和其他利益相關者會收到報告，報告中會顯示營運統計資料，例如某一段時間內的通話量、使用者滿意度分數、待處理票證數量及其存留時間。

常見的反模式：

- 工作負載停擺，造成服務無法使用。通話量暴增，因為使用者要求得知發生什麼情況。主管也要求得知誰在處理問題，因而增加了通話量。不同的營運團隊重複投入嘗試調查的工作。
- 由於需要新功能，因而轉派數名人員進行工程工作。但未回補空缺，使得解決問題的時間大幅拉長。領導階層並未獲得這些資訊，而是在經過數週且收到使用者不滿意的意見回饋後才察覺此問題。

建立此最佳實務的優勢：在業務受到影響的營運事件中，各個團隊可能會浪費大量時間和精力來查詢資訊，以試圖了解情況。透過建立廣泛傳播的狀態頁面和儀表板，利益相關者就能迅速獲得資訊，例如是否偵測到問題、誰負責處理問題，或是預計何時恢復正常營運。如此一來，團隊成員就不需花太多時間與其他人溝通狀態，因而有更多時間解決問題。

此外，儀表板和報告可以為決策者和利益相關者提供洞見，以了解營運團隊如何能夠回應業務需求以及如何分配其資源。這對於確定是否有足夠的資源來支援業務至關重要。

未建立此最佳實務時的曝險等級：中

實作指引

建置儀表板，以顯示營運團隊目前的關鍵指標，並且讓營運主管和管理層隨時可存取這些資訊。

建置可快速更新的狀態頁面，以顯示事故或事件何時發生、負責人是誰，以及誰負責協調回應。在此頁面上分享使用者應考量的任何步驟或因應措施，並廣泛傳播位置。鼓勵使用者遇到未知的問題時，先查看此位置。

收集並提供報告，以顯示長時間的營運狀況，並將此資訊傳達給主管和決策者，以說明運營工作及挑戰和需求。

在團隊之間共用這些最能反映目標和 KPI 的指標和報告，以及這些資訊在推動變革方面的影響力。花時間進行這些活動，以在團隊內部和團隊之間提高營運的重要性。

使用 [AWS Health](#) 搭配您擁有的儀表板，或將 AWS Health 事件整合到其中，讓您的團隊能夠找出應用程式問題與 AWS 服務狀態之間的相互關聯。

資源

相關的最佳實務：

- [OPS09-BP01 使用指標衡量營運目標與 KPI](#)

相關文件：

- [測量進度](#)
- [建置用於檢視營運狀況的儀表板](#)

相關範例：

- [資料操作](#)
- [如何使用 KPI 儀表板追蹤成本最佳化 KPI](#)
- [大規模雲端遷移關鍵績效指標 \(KPI\) 的重要性](#)

OPS09-BP03 審核營運指標並優先改進

預留審核營運狀態的專屬時間和資源，以確保依舊優先處理日常業務線所需的服務。召集營運主管和利益相關者定期審核指標、重新確認或修改各項目標，並優先改進。

預期成果：

- 營運主管和員工定期開會，以審核一段特定報告期間的指標。說明挑戰、一同慶祝成就，並分享學到的經驗。
- 利益相關者與企業領導者會定期收到營運狀態的簡報，並徵求有關目標、KPI 和未來計畫的意見。討論服務交付、營運和維護之間的權衡，並納入相關環境中。

常見的反模式：

- 新產品已推出，但 1 級和 2 級營運團隊未接受足夠的培訓來提供支援，或未配置額外的人員。領導者未看見指出支援單解決次數減少且事故量增加的指標。訂閱數量隨著不滿的使用者離開平台而開始減少，但數週後才採取行動。
- 長久以來一直採用手動程序來執行工作負載維護工作。雖然渴望自動化，但由於系統的重要性較低，因此優先順序較低。然而經過一段時間後，系統的重要性已提高，而現在這些手動程序佔用了大多數營運時間。未安排資源來提供更多營運工具，導致員工隨著工作負載增加而倦怠。等到有員工離職並加入其他競爭對手，領導階層才察覺到此情況。

建立此最佳實務的優勢：在某些組織中，將相同的時間和注意力分配給服務交付和新產品或方案可能會是一項挑戰。發生這種情況時，業務線可能會因為預期的服務層級逐漸惡化而受到影響。這是因為營運未隨著業務成長而改變和發展，並且可能很快就會落後。假如未定期審核營運收集的洞見，那麼察覺到業務風險時，便可能為時已晚。透過分配時間與營運員工和領導階層一起審核指標和程序，就能持續掌握營運所扮演的重要角色，並且能夠在風險達到嚴重等級之前發現。營運團隊能夠更深入洞察即將發生的業務變化與計畫，進而採取積極的行動。領導階層對於營運指標的掌握程度，展現了這些團隊在內部和外部的客戶滿意度方面所扮演的角色，並讓他們在各種選擇當中權衡出更適當的優先順序，或確保營運團隊有時間和資源能夠隨著新的業務和工作負載計畫做出改變與發展。

未建立此最佳實務時的曝險等級：中

實作指引

花時間與利益相關者和營運團隊一起審核營運指標，並審核報告資料。將這些報告與組織的目標相互比對，以確定是否符合這些目標。找出目標不明確，或者要求與付出之間存在衝突的模糊地帶來源。

找出時間、人員和工具能夠協助實現營運成果的地方。確定哪些 KPI 會受到影響，以及哪些應是成功的目標。定期重新檢視，以確保營運資源充足，可支援業務線。

資源

相關文件：

- [Amazon Athena](#)

- [Amazon CloudWatch 指標和維度參考](#)
- [Amazon Quick](#)
- [AWS Glue](#)
- [AWS Glue Data Catalog](#)
- [使用 Amazon CloudWatch Agent 從 Amazon EC2 執行個體和內部部署伺服器收集指標和日誌](#)
- [使用 Amazon CloudWatch 指標](#)

OPS 10. 如何管理工作負載和營運事件？

準備和驗證回應事件的程序，大幅降低工作負載中斷情形。

最佳實務

- [OPS10-BP01 使用程序進行事件、事故和問題管理](#)
- [OPS10-BP02 每個提醒建立一個程序](#)
- [OPS10-BP03 根據業務影響確定營運事件的優先順序](#)
- [OPS10-BP04 定義呈報路徑](#)
- [OPS10-BP05 為影響服務的事件定義客戶溝通計劃](#)
- [OPS10-BP06 透過儀表板傳達狀態](#)
- [OPS10-BP07 自動化對事件的回應](#)

OPS10-BP01 使用程序進行事件、事故和問題管理

有效管理事件、事故和問題的能力是維持工作負載運作狀態和效能的關鍵。識別和理解這些元素之間的差異，以制定有效的回應和解決策略至關重要。為每個方面建立並遵循明確定義的流程，有助於您的團隊迅速且有效地處理出現的任何運營挑戰。

預期成果：您的組織透過詳細記錄且集中儲存的流程，有效地管理營運事件、事故和問題。這些流程會持續更新以反映變更，簡化處理並維持高服務可靠性和工作負載效能。

常見的反模式：

- 您會反應性地 (而非主動) 回應事件。
- 對不同類型的事件或事故採取不一致的方法。
- 您的組織不會分析事件並從中學習，以防止未來再次發生。

建立此最佳實務的優勢：

- 簡化且標準化的回應流程。
- 減少事件對服務和客戶的影響。
- 加速解決問題。
- 持續改善營運流程。

未建立此最佳實務時的曝險等級：高

實作指引

實作此最佳實務表示您正在追蹤工作負載事件。您有處理事件和問題的程序。會經常記錄、共用和更新這些程序。問題經識別後會定出優先順序，然後獲得修正。

了解事件、事故和問題

- 事件：事件是對動作、狀況或狀態變化的觀察。事件可以經過計劃或未計劃，並且事情可以在工作負載內部或外部產生。
- 事故：事故是指需要回應的事件，例如意外中斷或服務品質下降。它們表示需要立即注意以恢復正常工作負載操作的中斷。
- 問題：問題是一個或多個事故的根本原因。識別和解決問題涉及更深入地研究事故，以防止將未來再次發生。

實作步驟

事件

1. 監控事件：

- [實作可觀測性](#)並[利用工作負載可觀測性](#)。
- 使用者、角色或 AWS 服務所採取的監控動作會記錄為 [AWS CloudTrail](#) 中的事件。
- 使用 [Amazon EventBridge](#) 即時對應用程式中的操作變更進行回應。
- 使用 [AWS Config](#) 持續評估、監控和記錄資源組態變更。

2. 建立程序：

- 制定一個程序來評估哪些事件重要並需要監控。這涉及設定正常和異常活動的閾值和參數。
- 確定將事件升級為事故的條件。這可以基於嚴重性、對使用者的影響或與預期行為的偏差。
- 定期審核事件監控和回應程序。這包括分析過去的事件、調整閾值以及完善警示機制。

事故

1. 回應事故：

- 使用可觀測性工具的洞察力，快速識別並回應事故。
- 實作 [AWS Systems Manager Ops Center](#) 以彙總、組織營運項目和事故，並排定優先順序。
- 使用 [Amazon CloudWatch](#) 和 [AWS X-Ray](#) 等服務，進行更深入的分析 and 疑難排解。
- 考慮使用 [AWS Managed Services \(AMS\)](#) 來加強事件管理，利用其主動、預防性和偵測功能。AMS 透過監控、事故偵測和回應以及安全管理等服務擴展了營運支援。
- Enterprise Support 客戶可利用 [AWS 事件偵測與回應](#) 功能，為生產工作負載提供持續的主動監控和事件管理。

2. 建立事件管理程序：

- 建立結構化的事件管理流程，包括清晰的角色、通訊協定和解決步驟。
- 將事件管理與聊天應用程式中的 [Amazon Q Developer](#) 這類工具整合，以實現有效率的回應和協調。
- 依嚴重性將事件分類，並針對每個類別預先定義[事件回應計劃](#)。

3. 學習和改進：

- 進行[事件後分析](#)以了解根本原因和解決方案有效性。
- 根據審查和不斷發展的實務，持續更新和改進回應計劃。
- 記錄並分享跨團隊所學到的經驗教訓，以增強營運彈性。
- Enterprise Support 客戶可向其技術客戶經理請求參加[事件管理研討會](#)。這個指導性研討會可測試您現有的事件回應計畫，並協助您找出需要改進的領域。

問題

1. 識別問題：

- 使用先前事件的資料來識別可能指出更深層次系統性問題的週期性模式。
- 運用諸如 [AWS CloudTrail](#) 和 [Amazon CloudWatch](#) 等工具來分析趨勢並發現潛在問題。
- 與包括營運、開發和業務單位在內的跨職能團隊合作，以獲得有關根本原因的不同觀點。

2. 建立問題管理程序：

- 制定問題管理的結構化程序，專注於長期解決方案，而不是快速修復。
- 整合根本原因分析 (RCA) 技術，以調查並了解事件的根本原因。

• 根據調查結果更新營運政策、程序和基礎設施，以防止重複發生。

3. 持續改善：

- 培養不斷學習和改進的文化，鼓勵團隊積極識別和解決潛在問題。
- 定期審查和修訂問題管理程序和工具，以配合不斷發展的業務和技術環境。
- 在整個組織中分享見解和最佳實務，以建立更具彈性且更有效率的營運環境。

4. 聯絡 AWS 支援：

- 使用 AWS 支援資源，例如 [AWS Trusted Advisor](#)，取得主動式指引和最佳化建議。
- Enterprise Support 客戶可以在重大事件期間存取 [AWS Countdown](#) 等專業計畫以取得支援。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS04-BP01 識別關鍵績效指標](#)
- [OPS04-BP02 實作應用程式遙測](#)
- [OPS07-BP03 使用執行手冊執执行程序](#)
- [OPS07-BP04 使用程序手冊來調查問題](#)
- [OPS08-BP01 分析工作負載指標](#)
- [OPS11-BP02 執行事後分析](#)

相關文件：

- [AWS 安全事件回應指南](#)
- [AWS 事件偵測與回應](#)
- [AWS Cloud Adoption Framework：營運角度 - 事件與問題管理](#)
- [DevOps 和 SRE 時代的事件管理](#)
- [PagerDuty - 什麼是事件管理？](#)

相關影片：

- [來自 AWS 的重要事件回應秘訣](#)
- [AWS re:Invent 2022 - Amazon 建置者資料中心：25 年 Amazon 卓越營運](#)

- [AWS re:Invent 2022 - AWS 事故偵測與回應 \(SUP201\)](#)
- [介紹 AWS Systems Manager 中的 Incident Manager](#)

相關範例：

- [AWS 主動服務 – 事件管理研討會](#)
- [如何使用 PagerDuty 和 AWS Systems Manager Incident Manager 來自動化事故回應](#)
- [使用 AWS Systems Manager Incident Manager 中的隨時待命的時間表與事故回應人員互動](#)
- [改善 AWS Systems Manager Incident Manager 中事故處理期間的可見性和協作](#)
- [AMS 中的事故報告和服務請求](#)

相關服務：

- [Amazon EventBridge](#)

OPS10-BP02 每個提醒建立一個程序

為系統中的每個提醒建立清晰明確的程序，對於有效且高效的事件管理至關重要。此做法可確保每個提醒都能產生特定且可行的回應，從而改善操作的可靠性和回應能力。

預期成果：每個提醒都會啟動特定且明確定義的回應計劃。在可能的情況下，回應會自動化，具有明確的擁有權和定義的呈報路徑。提醒會連結至最新的知識庫，以便任何操作員都能一致且有效地回應。回應迅速且全面一致，可提升營運效率和可靠性。

常見的反模式：

- 提醒沒有預定義的回應流程，導致臨時和延遲的解決方案。
- 提醒過載會導致重要提醒被忽略。
- 由於缺乏明確的擁有權和責任，提醒的處理不一致。

建立此最佳實務的優勢：

- 透過僅提高可操作的提醒來減少提醒疲勞。
- 減少操作問題的平均解決時間 (MTTR)。
- 減少平均調查時間 (MTTI)，有助於降低 MTTR。

- 增強擴展操作回應的能力。
- 提高了處理操作事件中的一致性和可靠性。

例如，您已有既定流程來處理重要帳戶的 AWS Health 事件，包括應用程式警示、營運問題及規劃的生命週期事件 (例如，在叢集自動更新之前更新 Amazon EKS 版本)，而且您為團隊提供主動監控、溝通和回應這些事件的能力。這些動作有助於防止 AWS 端變更造成的服務中斷，或是在發生非預期的問題時更快緩解。

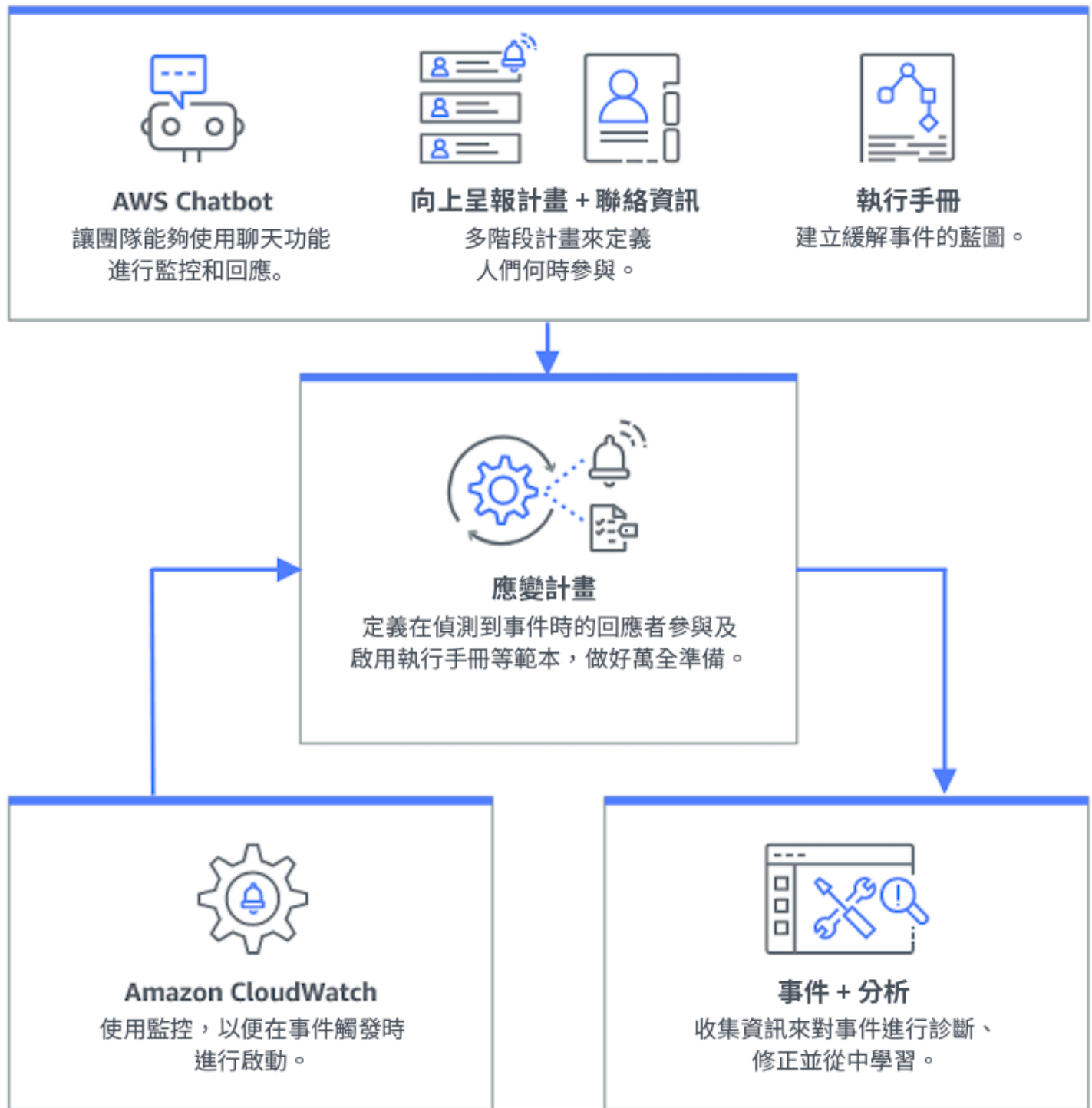
未建立此最佳實務時的曝險等級：高

實作指引

為每個提醒制定一個流程，包括：為每個提醒建立清晰的回應計劃；在可能的情況下自動化回應；並根據營運意見回饋和不斷發展的需求持續完善這些流程。

實作步驟

下圖說明 [AWS Systems Manager Incident Manager](#) 中的事件管理工作流程。它的設計目的是透過自動建立事件來回應 [Amazon CloudWatch](#) 或 [Amazon EventBridge](#) 中的特定事件，迅速回應營運問題。自動或手動建立事件時，Incident Manager 會集中管理事件，組織相關的 AWS 資源資訊，並啟動預先定義的回應計劃。這包括執行 Systems Manager Automation 執行手冊以立即採取行動，以及在 OpsCenter 中建立父作業工作項目以追蹤相關任務和分析。此簡化的流程可加速並協調整個 AWS 環境中的事件回應。



1. 使用複合警示：在 CloudWatch 中建立[複合警示](#)，將相關警示分組，從而降低噪音並允許更有意義的回應。

2. 利用 [AWS Health](#) 隨時掌握新知：AWS Health 是 AWS 雲端 資源運作狀態的權威資訊來源。使用 AWS Health 視覺化並接收有關任何目前服務事件和近期變更的通知 (例如規劃的生命週期事件)，如此您就能採取行動來緩解衝擊。
 - a. 透過 [AWS 使用者通知](#) 建立符合用途的 [AWS Health 事件通知](#)，以利用電子郵件和聊天管道傳送，並透過 [Amazon EventBridge](#) 或 [AWS Health API](#) 以程式設計方式與您的監控和警示工具整合。
 - b. 透過 [Amazon EventBridge](#) 或 [AWS Health API](#) 整合變更管理或您可能已在使用的 ITSM 工具 (如 [Jira](#) 或 [ServiceNow](#))，以規劃並追蹤需要採取行動的運作狀態事件進度。
 - c. 如果您使用 AWS Organizations，請啟用 [AWS Health 的組織檢視](#)，以彙總帳戶之間的 AWS Health 事件。
3. 整合 Amazon CloudWatch 警示與 Incident Manager 設定 CloudWatch 警示，以便在 [AWS Systems Manager Incident Manager](#) 中自動建立事件。
4. 整合 Amazon EventBridge 與 Incident Manager：建立 [EventBridge 規則](#)以回應事件並使用定義的回應計劃建立事件。
5. 為 Incident Manager 中的事件做好準備：
 - 在 Incident Manager 中針對每種提醒類型建立詳細的[回應計劃](#)。
 - 透過與 Incident Manager 中回應計劃相連的[聊天應用程式中的 Amazon Q Developer](#) 來建立聊天頻道，以便在 Slack、Microsoft Teams 和 Amazon Chime 等平台的事件期間進行即時通訊。
 - 將 [Systems Manager Automation 執行手冊](#)納入 Incident Manager 中，以推動對事件的自動回應。

資源

相關的最佳實務：

- [OPS04-BP01 識別關鍵績效指標](#)
- [OPS08-BP04 建立可執行的提醒](#)

相關文件：

- [AWS Cloud Adoption Framework：營運角度 - 事件與問題管理](#)
- [使用 Amazon CloudWatch 警示](#)
- [設定 AWS Systems Manager Incident Manager](#)
- [為 Incident Manager 中的事件做好準備](#)

相關影片：

- [來自 AWS 的重要事件回應秘訣](#)
- [re:Invent 2023 | 使用 AWS Health 大規模管理資源生命週期事件](#)

相關範例：

- [AWS 研討會 - AWS Systems Manager Incident Manager - 自動化對安全事件的事件回應](#)

OPS10-BP03 根據業務影響確定營運事件的優先順序

及時回應操作事件至關重要，但並非所有事件都是平等的。當您根據業務影響排定優先順序時，也可以優先處理可能產生重大後果的事件，例如安全、財務損失、違反法規或聲譽損害。

預期成果：根據對業務營運和目標的潛在影響，對營運事件的回應進行優先級排序。這使得回應高效且有效。

常見的反模式：

- 每個事件都按相同的緊急程度處理，會導致解決關鍵問題時出現混亂和延遲。
- 您無法區分高影響和低影響事件，導致資源分配錯誤。
- 您的組織缺乏明確的優先順序排定架構，導致對營運事件的回應不一致。
- 事件的優先級基於其報告的順序，而不是其對業務成果的影響。

建立此最佳實務的優勢：

- 確保關鍵業務功能首先獲得關注，將潛在損害降至最低。
- 改善多個並行事件期間的資源配置。
- 增強組織維持信任並遵守法規要求的能力。

未建立此最佳實務時的曝險等級：高

實作指引

當面對多個營運事件時，根據影響和緊迫性來確定優先順序的結構化方法至關重要。這種方法可幫助您做出明智的決策，直接在最需要的地方做出努力，並降低業務持續性的風險。

實作步驟

1. 評估影響：制定分類系統，根據事件對業務營運和目標的潛在影響來評估事件的嚴重性。下列範例顯示了影響類別：

影響層級	Description
高	影響許多員工或客戶，財務影響大，聲譽受損或受傷嚴重。
中	影響一組員工或客戶，中度財務影響，或中度聲譽受損。
低	影響個別員工或客戶，財務影響小，聲譽受損不嚴重。

2. 評估緊迫性：定義事件需要回應的速度之緊急程度，考慮安全、財務影響和服務水準協議 (SLA) 等因素。下列範例示範緊急類別：

緊急程度	Description
高	損害呈指數級增長，時間敏感性工作受到影響，迫在眉睫的升級，VIP 使用者或群體受到影響。
中	損害隨時間增長，或單個 VIP 使用者或群體受到影響。
低	邊際損害隨時間增長，或非時間敏感性工作受到影響。

3. 建立一個優先級矩陣：
 - 使用矩陣來交叉參考影響和緊迫性，將優先級別分配給不同的組合。
 - 讓負責營運事件回應的所有團隊成員都能存取和理解矩陣。
 - 下列範例矩陣會根據緊急性 and 影響來顯示事件嚴重性：

緊迫性和影響	高	中	低
高	嚴重	緊急	高
中	緊急	高	正常
低	高	正常	低

4. 培訓與溝通：對回應團隊進行優先級矩陣的培訓，並強調在事件期間遵循該矩陣的重要性。向所有利益相關者傳達優先級排序過程，以設定明確的期望。
5. 與事件回應整合：
 - 將優先級矩陣整合到您的事件回應計畫和工具中。
 - 盡可能自動化事件的分類和優先順序，以加快回應時間。
 - 企業支援客戶可利用 [AWS 事件偵測與回應](#) 功能，為生產工作負載提供全年無休的主動監控和事件管理。
6. 審查和調整：定期審查優先順序排定程序的有效性，並根據業務環境中的意見回饋和變化進行調整。

資源

相關的最佳實務：

- [鼓勵 OPS03-BP03 升級](#)
- [OPS08-BP04 建立可執行的提醒](#)
- [OPS09-BP01 使用指標衡量營運目標與 KPI](#)

相關文件：

- [Atlassian - 了解事件嚴重性層級](#)
- [IT 流程圖 - 檢查清單事件優先順序](#)

OPS10-BP04 定義呈報路徑

在您的事件回應協定中建立明確的呈報路徑，以促進及時且有效的活動。這包括指定呈報提示、詳細說明呈報流程，以及預先核准動作，以加速決策並縮短平均解決時間 (MTTR)。

預期成果：結構化且有效率的流程，可將事件呈報給適當的人員，將回應時間和影響降到最低。

常見的反模式：

- 復原程序不明確會導致在關鍵事件期間採取臨時應對措施。
- 當需要緊急行動時，缺少已定義的權限和擁有權會導致延遲。
- 利益相關者和客戶沒有按照預期得到通知。
- 重要決策被推遲。

建立此最佳實務的優勢：

- 透過預先定義的呈報程序來簡化事件回應。
- 透過預先核准的動作和明確的擁有權，減少停機時間。
- 根據事件嚴重性來改善資源配置和支援層級調整。
- 改善與利益相關者和客戶的溝通。

未建立此最佳實務時的曝險等級：中

實作指引

正確定義的呈報路徑對於快速事件回應至關重要。AWS Systems Manager Incident Manager 支援設定結構化呈報計畫和隨時待命的排程，它們會提醒合適人員，以便他們在事件發生時做好行動準備。

實作步驟

1. 設定呈報提示：設定 [CloudWatch 警示](#) 以在 [AWS Systems Manager Incident Manager](#) 中建立事件。
2. 設定隨時待命的排程：在 Incident Manager 中建立與您的呈報路徑保持一致的[隨時待命的排程](#)。為隨時待命的人員提供必要的權限和工具，以迅速採取行動。
3. 詳細說明呈報程序：
 - 確定應在哪些特定條件下呈報事件。
 - 在 Incident Manager 中建立[呈報計畫](#)。
 - 呈報渠道應包括聯絡人或隨時待命的時間表。
 - 定義團隊在每個呈報級別的角色和職責。
4. 預先核准的緩解措施：與決策者協同合作，針對預期情況預先核准動作。使用與 Incident Manager 整合的 [Systems Manager Automation 執行手冊](#)，加快事件解決速度。

5. 指定擁有權：針對呈報路徑的每個步驟，清楚識別內部擁有者。
6. 詳細說明第三方呈報：
 - 記錄第三方服務水準協議 (SLA)，並使其與內部目標保持一致。
 - 為事件期間的供應商溝通制定明確的協定。
 - 將供應商聯絡資訊整合至事件管理工具，以便直接存取。
 - 定期進行演練，包括第三方回應方案。
 - 保持供應商呈報資訊有據可查且易於存取。
7. 培訓和演練呈報計劃：對您的團隊進行呈報流程培訓，並定期進行事件回應演習或練習。企業支援客戶可申請[事件管理研討會](#)。
8. 持續改善：定期審核呈報路徑的有效性。根據事件發生後的經驗教訓和持續回饋來更新您的流程。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS08-BP04 建立可執行的提醒](#)
- [OPS10-BP02 每個提醒建立一個程序](#)
- [OPS11-BP02 執行事後分析](#)

相關文件：

- [AWS Systems Manager Incident Manager 呈報計劃](#)
- [在 Incident Manager 中使用隨時待命的時間表](#)
- [建立和管理執行手冊](#)
- [使用 AWS IAM Identity Center 進行臨時提升的存取管理](#)
- [Atlassian - 有效事件管理的呈報政策](#)

OPS10-BP05 為影響服務的事件定義客戶溝通計劃

在影響服務的事件中，有效的溝通對於維護與客戶的信任和透明度至關重要。明確定義的溝通計劃可協助您的組織在事件發生期間快速且清楚地分享資訊，包括內部和外部。

預期成果：

- 強大的溝通計劃可在影響服務的事件中有效地通知客戶和利益相關者。
- 溝通中的透明度可建立信任並減少客戶焦慮。
- 盡量減少服務影響事件對客戶體驗和業務運營的影響。

常見的反模式：

- 溝通不充分或延遲會導致客戶困惑和不滿。
- 過於技術化或模糊的消息傳遞無法傳達對使用者的實際影響。
- 沒有預先定義的溝通策略，導致不一致且被動的消息傳遞。

建立此最佳實務的優勢：

- 透過主動和清晰的溝通，增強客戶的信任和滿意度。
- 透過搶先解決客戶問題，減輕支援團隊的負擔。
- 改善有效管理事件並從中復原的能力。

未建立此最佳實務時的曝險等級：中

實作指引

為影響服務的事件制定全面的溝通計劃涉及多個方面，從選擇正確的渠道到精心製作消息和基調。該計劃應具有適應性、可擴展性，並適應不同的停機情況。

實作步驟

1. 定義角色和責任：

- 指派一名主要事件管理者來監督事件回應活動。
- 指定一名溝通管理者，其負責協調所有外部與內部溝通。
- 包括支援管理者，以透過支援票證提供一致的溝通。

2. 確定溝通渠道：選取諸如工作場所聊天、電子郵件、簡訊、社交媒體、應用程式內通知和狀態頁面等渠道。這些渠道應具有彈性，並且能夠在影響服務的事件期間獨立運作。

3. 快速、清晰、定期地與客戶溝通：

- 為各種服務損害場景開發模板，強調簡單性和基本細節。包括有關服務損害、預期解決時間和影響等資訊。

- 透過推播通知、應用程式內通知、電子郵件、文字訊息、語音訊息和自訂渠道上的訊息，使用 Amazon Pinpoint 來提醒客戶。
 - 使用 Amazon Simple Notification Service (Amazon SNS) 以程式設計方式或透過電子郵件、行動推送通知和文字訊息來提醒訂閱用戶。
 - 透過公開共用 Amazon CloudWatch 儀表板，使用儀表板溝通狀態。
 - 鼓勵社交媒體參與：
 - 積極監控社交媒體以了解客戶情緒。
 - 在社交媒體平台上發布公共更新和社區參與情況。
 - 準備範本以進行一致且清晰的社交媒體溝通。
4. 協調內部溝通：使用如聊天應用程式中的 Amazon Q Developer 等工具實作內部通訊協定，以進行團隊協調和溝通。使用 CloudWatch 儀表板來溝通狀態。
5. 使用專用工具和服務協調溝通：
- 搭配使用 AWS Systems Manager Incident Manager 和聊天應用程式中的 Amazon Q Developer 即可設定專屬的聊天頻道，以便在事件發生時進行即時內部溝通和協調。
 - 在事件發生期間，透過 Amazon Pinpoint、Amazon SNS 或第三方工具 (例如社交媒體平台)，使用 AWS Systems Manager Incident Manager 執行手冊將客戶通知自動化。
 - 在執行手冊中合併核准工作流程，以便在傳送前選擇性地審核和授權所有外部通訊。
6. 實踐和改善：
- 針對溝通工具和策略的使用進行培訓。讓團隊能夠在事件發生時及時做出決策。
 - 透過定期演習或演練日測試溝通計劃。使用這些測試來精簡消息傳遞並評估渠道的有效性。
 - 實作意見回饋機制，以評估事件期間的溝通效率。根據意見回饋和不斷變化的需求不斷發展溝通計劃。

實作計劃的工作量：高

資源

相關的最佳實務：

- [OPS07-BP03 使用執行手冊執执行程序](#)
- [OPS10-BP06 透過儀表板傳達狀態](#)
- [OPS11-BP02 執行事後分析](#)

相關文件：

- [Atlassian - 事件溝通最佳實務](#)
- [Atlassian - 如何編寫良好的狀態更新](#)
- [PagerDuty - 事件溝通指南](#)

相關影片：

- [Atlassian - 建立您自己的事件溝通計劃：事件範本](#)

相關範例：

- [AWS Health 儀表板](#)

OPS10-BP06 透過儀表板傳達狀態

使用儀表板作為戰略工具，將即時營運狀態和關鍵指標傳達給不同的受眾，包括內部技術團隊、領導層和客戶。這些儀表板提供系統運作狀態和業務績效的集中式視覺化呈現，從而提高透明度和決策效率。

預期成果：

- 儀表板提供與不同利益相關者相關的系統和業務指標的全面檢視。
- 利益相關者可以主動存取營運資訊，減少頻繁的狀態請求。
- 在正常操作和事件期間增強實時決策。

常見的反模式：

- 加入事件管理通話的工程師要求更新狀態以加快速度。
- 依靠手動報告進行管理，這會導致延遲和潛在的不準確性。
- 事件發生期間，營運團隊經常因為狀態更新而受到干擾。

建立此最佳實務的優勢：

- 使利益相關者能夠立即存取關鍵資訊，有助於制定明智決策。
- 透過最大限度地減少手動報告和頻繁狀態查詢，減少操作效率低下問題。
- 透過即時掌握系統效能和業務指標，提高透明度和信任度。

未建立此最佳實務時的曝險等級：中

實作指引

儀表板可有效地傳達系統和業務指標的狀態，並可根據不同受眾群體的需求進行量身打造。Amazon CloudWatch 儀表板和 Amazon Quick 等工具可協助您建立用於系統監控和商業智慧的互動式即時儀表板。

實作步驟

1. 確定利益相關者的需求：確定不同受眾群體的特定資訊需求，例如技術團隊、領導層和客戶。
2. 選擇正確的工具：選取適當的工具，例如，使用 [Amazon CloudWatch 儀表板](#) 進行系統監控，以及使用 [Amazon Quick](#) 獲得互動式商業情報。[AWS Health](#) 在 [AWS Health 儀表板](#) 中提供了立即可用的體驗，您也可以使用 Amazon EventBridge 中的運作狀態事件或透過 AWS Health API 來擴增您的儀表板。
3. 設計高效儀表板：
 - 設計儀表板以清楚呈現相關指標和 KPI，確保其易於理解和操作。
 - 視需要整合系統層級與企業層級檢視。
 - 包括高階 (用於廣泛概述) 和低階 (用於詳細分析) 儀表板。
 - 在儀表板中整合自動警示，以突顯重大問題。
 - 使用重要指標閾值和目標為儀表板加上註釋，以實現即時可見性。
4. 整合資料來源：
 - 使用 [Amazon CloudWatch](#) 彙總和顯示來自各種 AWS 服務的指標，並[查詢其他資料來源的指標](#)，建立系統運作狀態和商業指標的統一檢視。
 - 使用 [CloudWatch Logs Insights](#) 等功能，查詢並視覺化來自不同應用程式和服務的日誌資料。
 - 使用 AWS Health 事件透過 [AWS Health API](#) 或 [Amazon EventBridge 上的 AWS Health 事件](#)，隨時掌握 AWS 服務的操作狀態和已確認的操作問題。
5. 提供自助服務存取：
 - 與相關利益相關者共用 CloudWatch 儀表板，以使用[儀表板共用功能](#)進行自助式資訊存取。
 - 確保儀表板易於存取，並提供即時、最新的資訊。
6. 定期更新和完善：
 - 持續更新和完善儀表板，以滿足不斷變化的業務需求和利益相關者的意見回饋。
 - 定期審核儀表板，使其保持相關性並可有效傳達必要資訊。

資源

相關的最佳實務：

- [OPS08-BP05 建立儀表板](#)

相關文件：

- [建置用於檢視營運狀況的儀表板](#)
- [使用 Amazon CloudWatch 儀表板](#)
- [使用儀表板變數建立彈性儀表板](#)
- [共享 CloudWatch 儀表板](#)
- [從其他資料來源中查詢指標](#)
- [將自訂小工具新增至 CloudWatch 儀表板](#)

相關範例：

- [一個可觀測性研討會 - 儀表板](#)

OPS10-BP07 自動化對事件的回應

自動化事件回應是快速、一致且無誤操作處理的關鍵。建立簡化的流程，並使用工具自動管理和回應事件，將手動干預降至最低，並提高營運效率。

預期成果：

- 透過自動化減少人為錯誤並縮短解決時間。
- 一致且可靠的操作事件處理。
- 提高運營效率和系統可靠性。

常見的反模式：

- 手動事件處理會導致延遲和錯誤。
- 在重複的關鍵任務中，自動化被忽略。
- 重複的手動任務會導致警示疲勞，並遺漏重大問題。

建立此最佳實務的優勢：

- 加速事件回應，減少系統停機時間。
- 可靠的操作，自動化且一致的事件處理。

未建立此最佳實務時的曝險等級：中

實作指引

整合自動化以建立有效的操作工作流程，並將手動干預降至最低。

實作步驟

1. 識別自動化機會：確定自動化的重複性任務，例如問題修復、工單擴充、容量管理、擴展、部署和測試。
2. 識別自動化提示：
 - 使用 [Amazon CloudWatch 警示動作](#) 來評估和定義啟動自動回應的特定條件或指標。
 - 使用 [Amazon EventBridge](#) 來回應 AWS 服務、自訂工作負載和 SaaS 應用程式中的事件。
 - 考慮啟動事件，例如[特定日誌項目](#)、[效能指標閾值](#) 或 AWS 資源的[狀態變更](#)。
3. 實作事件驅動型自動化：
 - 使用 AWS Systems Manager Automation Runbook 簡化維護、部署和修復任務。
 - 在 [Incident Manager 中建立事件](#) 會自動收集有關事件所涉及 AWS 資源的詳細資訊，並將其新增至事件。
 - 使用 [Quota Monitor for AWS](#) 主動監控配額。
 - 使用 [AWS Auto Scaling](#) 自動調整容量，以維持可用性和效能。
 - 使用 [Amazon CodeCatalyst](#) 將開發管道自動化。
 - 煙霧測試或持續監控端點，APIs [並使用合成監控](#)。
4. 透過自動化執行風險緩解：
 - 實作 [自動化安全回應](#)，迅速解決風險。
 - 使用 [AWS Systems Manager 狀態管理器](#) 來減少組態偏差。
 - [使用修復不合規資源 AWS Config 規則](#)。

實作計劃的工作量：高

資源

相關的最佳實務：

- [OPS08-BP04 建立可執行的提醒](#)
- [OPS10-BP02 每個提醒建立一個程序](#)

相關文件：

- [搭配 Incident Manager 使用 Systems Manager Automation 執行手冊](#)
- [在 Incident Manager 中建立事件](#)
- [AWS 服務配額](#)
- [監控資源使用情況並在接近配額時傳送通知](#)
- [AWS Auto Scaling](#)
- [什麼是 Amazon CodeCatalyst？](#)
- [使用 Amazon CloudWatch 警示](#)
- [使用 Amazon CloudWatch 警示動作](#)
- [使用 修復不合規資源 AWS Config 規則](#)
- [使用篩選條件從日誌事件建立指標](#)
- [AWS Systems Manager State Manager](#)

相關影片：

- [使用 建立 Automation Runbook AWS Systems Manager](#)
- [如何在 上自動化 IT 操作 AWS](#)
- [AWS Security Hub CSPM 自動化規則](#)
- [使用 Amazon CodeCatalyst 藍圖快速啟動軟體專案](#)

相關範例：

- [Amazon CodeCatalyst 教學課程：使用現代三層式 Web 應用程式藍圖建立專案](#)
- [一個可觀測性研討會](#)
- [使用 Incident Manager 回應事件](#)

演進

問題

- [OPS 11. 如何改善營運？](#)

OPS 11. 如何改善營運？

投入時間和資源，盡量持續逐漸改善，以加強營運的效果和效率。

最佳實務

- [OPS11-BP01 建立持續改進程序](#)
- [OPS11-BP02 執行事後分析](#)
- [OPS11-BP03 實作意見回饋循環](#)
- [OPS11-BP04 執行知識管理](#)
- [OPS11-BP05 定義改進驅動因素](#)
- [OPS11-BP06 驗證洞察](#)
- [OPS11-BP07 執行操作指標檢閱](#)
- [OPS11-BP08 記錄和分享獲得的經驗](#)
- [OPS11-BP09 分配時間進行改善](#)

OPS11-BP01 建立持續改進程序

根據內部和外部架構最佳實務評估您的工作負載。進行頻繁、有目的的工作負載審查。根據您的軟體開發步調制定改進機會的優先順序。

預期成果：

- 根據架構最佳實務頻繁分析工作負載。
- 在軟體開發過程中，改進機會與功能獲得同等優先級。

常見的反模式：

- 您在數年前部署工作負載後，即未對其執行過架構審查。
- 您給予改進機會較低的優先級。與新功能相比，這些機會仍在待辦事項中。
- 沒有針對組織的最佳實務實作修改標準。

建立此最佳實務的優勢：

- 您的工作負載在架構最佳實務中保持最新狀態。
- 您以故意的方式發展工作負載。
- 您可以利用組織最佳實務來改進所有工作負載。
- 您可以獲得具有累積影響的邊際收益，從而提高效率。

未建立此最佳實務時的曝險等級：高

實作指引

經常對工作負載進行架構審查。使用內部和外部最佳實務，評估您的工作負載並識別改進機會。根據您的軟體開發步調制定改進機會的優先順序。

實作步驟

1. 以商定的頻率對生產工作負載進行定期架構審查。使用包含 AWS 特定最佳實務的已記載架構標準。
 - a. 使用內部定義的標準進行這些審查。如果沒有內部標準，可使用 AWS Well-Architected Framework。
 - b. 使用 AWS Well-Architected Tool 來建立內部最佳實務的自訂聚焦，並執行架構審查。
 - c. 請聯絡您的 AWS 解決方案架構師或技術客戶經理，在引導下執行工作負載的 Well-Architected Framework 審查。
2. 在您的軟體開發程序中，為在審查期間找出的改進機會制定優先順序。

實作計劃的工作量：低。可以使用 AWS Well-Architected Framework 進行每年的架構審查。

資源

相關的最佳實務：

- [OPS11-BP02 執行事件後分析](#)
- [OPS11-BP08 記錄和分享獲得的經驗](#)
- [OPS04 實作可觀測性](#)

相關文件：

- [AWS Well-Architected Tool - 自訂聚焦](#)
- [AWS Well-Architected 白皮書 - 審查程序](#)
- [使用自訂聚焦和 AWS Well-Architected Tool 來自訂 Well-Architected 審查](#)
- [在您的組織中實作 AWS Well-Architected 自訂聚焦生命週期](#)

相關影片：

- [AWS re:Invent 2023 - 擴展組織的 AWS Well-Architected 最佳實務](#)

相關範例：

- [AWS Well-Architected Tool](#)

OPS11-BP02 執行事後分析

審查影響客戶的事件，並識別造成問題的因素和預防性動作。使用此資訊來開發緩解措施，以限制或防止事件再次發生。制定可快速有效回應的程序。適當地傳達成因和為目標受眾量身打造的糾正措施。

預期成果：

- 您已建立包含事件後分析的事件管理程序。
- 您已制定可觀測性計畫，可收集有關事件的資料。
- 透過這些資料，您就可以了解並收集支援事件後分析程序的指標。
- 您可以從事件中學習，以改善未來的成果。

常見的反模式：

- 管理應用程式伺服器。大約每 23 小時 55 分鐘，您的所有活動工作階段都會終止。您試圖確定應用程式伺服器出了什麼問題。您懷疑這可能是網路問題，但由於網路團隊太忙而無法支援您，因此無法與他們合作。您缺乏可遵循的預定義流程來獲得支援並收集所需資訊以確定發生了什麼。
- 您的工作負載中有資料遺失。這是第一次發生，原因尚不清楚。您認為這並不重要，因為您可以重新建立資料。資料遺失開始更頻繁地出現，從而影響客戶。當您還原遺失的資料時，這也會為您帶來額外的操作負擔。

建立此最佳實務的優勢：

- 您有一個預先定義的程序來判斷造成事故的元件、條件、動作和事件，這有助於您找出改進機會。
- 可以使用事件後分析的資料進行改善。

未建立此最佳實務時的曝險等級：高

實作指引

使用程序判斷成因。審查所有影響客戶的事件。建立程序來識別和記錄事件的成因，以便您可以制定緩解措施來限制或防止事件再次發生。另外，您還可以制定快速有效地做出回應的程序。酌情溝通事件根本原因，並根據目標受眾量身定制溝通方式。在組織內公開分享學習成果。

實作步驟

1. 收集諸如部署變更、組態變更、事件開始時間、警示時間、參與時間、緩解開始時間和事件解決時間等指標。
2. 描述時間軸上的關鍵時間點，以了解事故的事件。
3. 請提出以下問題：
 - a. 您可以縮短偵測時間嗎？
 - b. 是否有指標和警示的更新，可以更快地檢測到事件？
 - c. 可以改善診斷時間嗎？
 - d. 回應計畫或呈報計畫是否有更新，可以更快地吸引合適的回應方？
 - e. 可以改善緩解時間嗎？
 - f. 是否有可以新增或改善的執行手冊或說明手冊步驟？
 - g. 可以防止未來的事件發生嗎？
4. 建立檢查清單和動作。追蹤並傳遞所有動作。

實作計劃的工作量：中

資源

相關的最佳實務：

- [OPS11-BP01 建立持續改進程序](#)
- [OPS 4 - 實作可觀測性](#)

相關文件：

- [在 Incident Manager 中執行事件後分析](#)
- [營運準備情況審核](#)

OPS11-BP03 實作意見回饋循環

回饋迴圈提供可推動決策的可行洞見。在程序和工作負載中建立回饋迴圈。此可協助您找出問題和需要改善的地方。回饋迴圈也會驗證在改善中所做的投資。這些回饋迴圈是持續改善工作負載的基礎。

回饋迴圈分為兩類：即時回饋和追溯性分析。透過審查營運活動的績效和成果來收集即時的回饋。此回饋來自團隊成員、客戶或活動的自動化輸出。接收 A/B 測試和交付新功能等方面的即時回饋，對於快速檢錯非常重要。

定期進行追溯性分析，以從對營運成果和指標的審查中獲取回饋。這些追溯性分析會在衝刺結束，按規律或在主要版本或事件後發生。這類回饋迴圈會驗證對營運或工作負載所做的投資。其可協助您衡量成功並驗證策略。

預期成果：您使用即時回饋和追溯性分析來推動改善。存在可擷取使用者和團隊成員回饋的機制。追溯性分析會用來找出可推動改善的趨勢。

常見的反模式：

- 您推出新功能，但沒有辦法收到客戶對該功能的回饋。
- 針對營運改善投入資源和時間後，您無法執行追溯性分析來進行驗證。
- 您收集客戶的回饋，但未能定期審查回饋。
- 回饋迴圈讓我們得以提議行動項目，但軟體開發程序中未納入這些項目。
- 客戶沒有收到他們提議之改善的回饋。

建立此最佳實務的優勢：

- 您可以反過來與客戶合作來推動新功能。
- 您的組織文化可以更快速地應對變化。
- 趨勢會用來找出改善的機會。
- 追溯性分析可驗證對工作負載和營運所做的投資。

未建立此最佳實務時的曝險等級：高

實作指引

實作此最佳實務表示您同時使用即時回饋和追溯性分析。這些回饋迴圈可推動改善。有許多機制可用來處理即時回饋，包含調查、客戶投票和回饋表單。組織也會使用追溯性分析來找出改善的機會並驗證計畫。

客戶範例

AnyCompany 零售建立了一個 Web 表單，客戶可以提供意見回饋或報告問題。在每週 Scrum 期間，軟體開發團隊會評估使用者回饋。該團隊會定期使用回饋來為其平台的發展釐清方向。他們會在每次衝刺結束時執行追溯性分析，來找出他們想要改善的項目。

實作步驟

1. 即時回饋

- 您需要制定機制來接收來自客戶和團隊成員的回饋。您也可以設定營運活動來提供自動化的回饋。
- 組織需要制定程序來審查此回饋、判斷需要改善的項目，並安排改善項目。
- 您必須將回饋新增至軟體開發程序。
- 在您著手改善後，請與回饋提交者追蹤後續進展。
 - 您可以使用 [AWS Systems Manager OpsCenter](#) 來建立和追蹤這些改進，如 [OpsItems](#)。

2. 追溯性分析

- 在開發週期結束時，以固定的規律或在主要版本之後，執行追溯性分析。
- 召集工作負載中參與的利益相關者，進行回顧會議。
- 在白板或試算表建立三個欄位：停止、開始和持續。
 - 停止是指您希望團隊停止做的任何事。
 - 開始是指您希望開始執行的想法。
 - 保持是指您希望持續執行的項目。
- 詢問在場人士的想法，收集利益相關者的回饋。
- 排列回饋的優先順序。將動作和利益相關者指派至任何「開始」或「持續」項目。
- 將動作新增至軟體開發程序中，並在您執行改善項目時向利益相關者告知最新的狀態。

實作計劃的工作量：中。若要實作此最佳實務，您需要找到方法來擷取即時回饋並進行分析。此外，您需要建立追溯性分析程序。

資源

相關的最佳實務：

- [OPS01-BP01 評估外部客戶需求](#)：回饋迴圈是一種機制，可收集外部客戶的需求。
- [OPS01-BP02 評估內部客戶需求](#)：內部利益相關者可以使用回饋迴圈來表達需要和需求。
- [OPS11-BP02 執行事後分析](#)：事件後分析是在事件後執行的追溯性分析的一種重要形式。
- [OPS11-BP07 執行操作指標檢閱](#)：營運指標審查會找出趨勢和待改善的地方。

相關文件：

- [建置 時要避免的 7 個陷阱 CCOE](#)
- [Atlassian 團隊程序手冊 - 追溯性](#)
- [電子郵件定義：回饋迴圈](#)
- [根據 AWS Well-Architected Framework Review 建立意見回饋循環](#)
- [IBM 車庫方法 - 保留回溯性](#)
- [投資 – PDICS週期](#)
- [最大化開發人員的效能 \(作者：Tim Cochran\)](#)
- [操作就緒審核 \(ORR \) 白皮書 - 迭代](#)
- [ITIL CSI - 持續服務改進](#)
- [當 Toyota 遇見電子商務：Amazon 的精實原則](#)

相關影片：

- [建立有效的客戶回饋迴圈](#)

相關範例：

- [Astuto - 開放原始碼客戶回饋工具](#)
- [AWS 解決方案 - Q nABot on AWS](#)
- [Fider - 整理客戶回饋的平台](#)

相關服務：

- [AWS Systems Manager OpsCenter](#)

OPS11-BP04 執行知識管理

知識管理可協助團隊成員尋找資訊以執行其作業。在學習組織中，資訊是任意共用的，助個人一臂之力。資訊可以探索和搜尋。資訊是準確且最新的。存在機制以建立新資訊、更新現有資訊，以及封存過時資訊。最常見的知識管理平台範例是內容管理系統，例如 Wiki。

預期成果：

- 團隊成員可以存取及時、準確的資訊。
- 資訊是可搜尋的。
- 存在機制以新增、更新和封存資訊。

常見的反模式：

- 沒有集中式知識儲存。團隊成員會在他們的本機電腦上管理他們自己的備註。
- 您有自我託管的 Wiki，但是沒有管理資訊的機制，導致資訊過時。
- 某人識別遺漏的資訊，但是沒有要求在團隊 Wiki 中新增它的程序。他們自行新增，但是遺漏關鍵步驟，導致中斷。

建立此最佳實務的優勢：

- 因為資訊任意共用，所以團隊成員握有能力。
- 因為文件是最新的且可搜尋，所以新的團隊成員可以更快上線。
- 資訊是及時、準確且可行的。

未建立此最佳實務時的曝險等級：高

實作指引

知識管理是學習組織的重要面向。若要開始，您需要集中儲存庫來存放您的知識 (常見的範例是自我託管的 Wiki)。您必須開發新增、更新和封存知識的程序。開發應該記載哪些項目的標準，並且讓所有人做出貢獻。

客戶範例

AnyCompany 零售託管儲存所有知識的內部 Wiki。團隊成員受到鼓勵在他們執行每日職責時新增至知識庫。跨功能團隊每季會評估哪些頁面最少更新，並且判斷它們是否應該封存或更新。

實作步驟

1. 從識別存放知識所在的內容管理系統開始。跨組織取得利益相關者的同意。
 - a. 如果您沒有現有內容管理系統，請考慮執行自我託管 Wiki 或使用版本控制儲存庫做為起點。
2. 開發新增、更新和封存資訊的執行手冊。向您的團隊教育這些程序。
3. 識別哪些知識應該存放在內容管理系統中。從團隊成員執行的每日活動 (執行手冊和程序手冊) 開始。與利益相關者合作來排列新增知識的優先順序。
4. 定期與利益相關者合作，以識別 out-of-date 資訊並將其封存或更新。

實作計劃的工作量：中。如果您沒有現有內容管理系統，您可以設定自我託管 Wiki 或版本控制文件儲存庫。

資源

相關的最佳實務：

- [OPS11-BP08 記錄和分享獲得的經驗](#) - 知識管理可促進所學習課程的資訊共用。

相關文件：

- [Atlassian - 知識管理](#)

相關範例：

- [DokuWiki](#)
- [Gollum](#)
- [MediaWiki](#)
- [Wiki.js](#)

OPS11-BP05 定義改進驅動因素

確定改進驅動因素，以幫助您依據資料和回饋迴圈來評估改進機會並排定其優先順序。探索系統和程序中的改進機會，並在適當的情況下自動化。

預期成果：

- 可以追蹤您的整個環境的資料。
- 可以將事件和活動與業務成果相關聯。
- 可以在環境和系統之間進行比較和對比。
- 可以維護部署和成果的詳細活動歷史記錄。
- 可收集資料以支援您的安全狀態。

常見的反模式：

- 可以從整個環境中收集資料，但不會關聯事件和活動。
- 可以從整個資產中收集詳細資料，這會導致較高的 Amazon CloudWatch 和 AWS CloudTrail 活動與成本。但是，您不會有目的地使用此資料。
- 在定義改進驅動因素時，您不會考慮業務成果。
- 您不會測量新功能的效果。

建立此最佳實務的優勢：

- 透過確定改進標準，可以將事件型動機或情緒投資的影響降到最低。
- 您可以回應商業活動，而不僅僅是技術事件。
- 測量您的環境，以確定需要改進的領域。

未建立此最佳實務時的曝險等級：中

實作指引

- 了解改進驅動因素：僅在支援理想結果時才對系統進行變更。
 - 所需能力：在評估改進機會時，評估所需的功能和能力。
 - [AWS 最新消息](#)
 - 不可接受的問題：在評估改進機會時，評估不可接受的問題、錯誤和漏洞。追蹤合適的選項，並尋求優化機會。
 - [AWS 最新安全公告](#)
 - [AWS Trusted Advisor](#)
 - [雲端智慧儀表板](#)

- 合規要求：在審查改進機會時，評估保持法規、政策的遵從性或保持受到第三方支援所需的更新和變更。
 - [AWS 合規](#)
 - [AWS 合規計劃](#)
 - [AWS 合規性最新資訊](#)

資源

相關的最佳實務：

- [OPS01 組織優先事項](#)
- [OPS02 關係和擁有權](#)
- [OPS04-BP01 識別關鍵績效指標](#)
- [OPS08 利用工作負載可觀測性](#)
- [OPS09 了解運作狀態](#)
- [OPS11-BP03 實作回饋迴圈](#)

相關文件：

- [Amazon Athena](#)
- [快速](#)：
- [AWS 合規](#)
- [AWS 合規性最新資訊](#)
- [AWS 合規計劃](#)
- [AWS Glue](#)
- [AWS 最新安全公告](#)
- [AWS Trusted Advisor](#)
- [將日誌資料匯出至 Amazon S3](#)
- [AWS 最新消息](#)
- [以客戶為中心的創新的必要條件](#)
- [數位轉型：炒作還是戰略需要？](#)

相關影片

- [AWS re:Invent 2023 - 透過 支援 \(SUP310\) 提高營運效率和彈性](#)

OPS11-BP06 驗證洞察

與跨職能團隊和企業擁有者一起審查您的分析結果和回應。透過這些審查建立共識，確定其他影響並確定行動方案。適當調整回應。

預期成果：

- 定期與企業擁有者審查洞見。企業擁有者為新獲得的洞察提供額外的內容。
- 可以檢閱洞見並請求技術同儕們的意見回饋，並在各個團隊之間分享您的學習成果。
- 可以發布資料和洞見，供其他技術和業務團隊審核。將所學知識納入其他部門的新實務中。
- 與資深主管一起總結和審查新洞見。資深主管使用新洞見來定義策略。

常見的反模式：

- 您發佈了一個新功能。此功能會變更部分客戶行為。您的可觀測性不會考慮這些變更。您不會量化這些變更的好處。
- 您推送新的更新並忽略重新整理您的 CDN。CDN 快取不再與最新版本相容。測量發生錯誤的請求百分比。您的所有使用者在與後端伺服器通訊時，都會報告 HTTP 400 個錯誤。調查用戶端錯誤，並發現由於您測量了錯誤的維度，所以時間被浪費了。
- 您的服務水準協議規定正常執行時間為 99.9%，而您的復原點目標為四小時。服務擁有者維護系統為零停機時間。實作昂貴且複雜的複寫解決方案，這會浪費時間和金錢。

建立此最佳實務的優勢：

- 與企業擁有者和領域專家驗證洞見時，建立共識並更有效地引導改進。
- 您會發現隱藏的問題，並將其納入未來決策中。
- 重點從技術成果轉移到業務成果。

未建立此最佳實務時的曝險等級：中

實作指引

- 驗證洞見：與企業擁有者和領域專家互動，確保您收集資料的意義得到眾人理解和同意。識別其他疑慮、潛在影響，並確定行動方案。

資源

相關的最佳實務：

- [OPS01-BP06 在管理效益和風險時評估權衡](#)
- [OPS團隊之間的 02-BP06 責任已預先定義或協商](#)
- [OPS11-BP03 實作意見回饋循環](#)

相關文件：

- [設計 Cloud Center of Excellence \(CCOE \)](#)

相關影片：

- [建立可觀測性以提高復原能力](#)

OPS11-BP07 執行操作指標檢閱

與來自不同業務領域的跨團隊參與者定期進行營運指標的追溯性分析。透過這些審查確定改進機會、可能的行動方案並分享獲得的經驗。尋找所有環境 (例如開發、測試和生產) 中的改善機會。

預期成果：

- 經常審查影響業務的指標
- 可以透過可觀測性功能來偵測和審查異常
- 可以使用資料來支援業務成果和目標

常見的反模式：

- 維護時段會中斷重要的零售促銷活動。如果還有其他影響企業的事件，企業仍然不知道是否有可能會延遲的標準維護時段。
- 您經歷了長時間的中斷，因為您經常使用組織中過時的程式庫。之後您已遷移到支援的程式庫。組織中的其他團隊不知道他們正面臨風險。
- 您不會定期檢閱客戶的達成情況SLAs。您即將不符合您的客戶SLAs。未滿足客戶會受到財務處罰SLAs。

建立此最佳實務的優勢：

- 當您定期開會以審查營運指標、事件和事故時，可以在團隊之間保持共識。
- 您的團隊會定期會面，以檢閱指標和事件，這些指標和事件可讓您針對風險採取行動並識別客戶 SLAs。
- 您分享學到的經驗教訓，為業務成果的優先順序和有針對性的改進提供資料。

未建立此最佳實務時的曝險等級：中

實作指引

- 與來自不同業務領域的跨團隊參與者定期進行營運指標的追溯性分析。
- 與包括業務、開發和營運團隊在內的利益相關者進行互動，以驗證您從即時回饋和追溯性分析獲得的發現，並分享經驗教訓。
- 利用這些洞見確定改進機會和可能的行動方案。

資源

相關的最佳實務：

- [OPS08-BP05 建立儀表板](#)
- [OPS09-BP03 檢閱操作指標並排定改善優先順序](#)
- [OPS10-BP01 使用事件、事件和問題管理的程序](#)

相關文件：

- [Amazon CloudWatch](#)
- [Amazon CloudWatch 指標和維度參考](#)
- [發佈自訂指標](#)
- [使用 Amazon CloudWatch 指標](#)
- [使用的儀表板和視覺化 CloudWatch](#)

OPS11-BP08 記錄和分享獲得的經驗

記錄並分享在操作活動中獲得的經驗，以便您可以在內部以及跨團隊使用它們。應分享您的團隊獲得的經驗，以提高整個組織的效益。共用資訊和資源，以防止可避免的錯誤及簡化開發工作，並專注於交付所需的功能。

使用 AWS Identity and Access Management (IAM) 來定義權限，允許對您希望在帳戶內及帳戶間分享的資源的受控存取。

預期成果：

- 使用版本控制的儲存器來分享應用程式程式庫、執行指令碼的程序、程序文件及其他系統文件。
- 可以將基礎設施標準共用為版本控制的 AWS CloudFormation 範本。
- 審核團隊學到的經驗教訓。

常見的反模式：

- 您經歷了長時間的中斷，因為您的組織經常使用錯誤的程式庫。之後您已遷移到可靠的程式庫。組織中的其他團隊不知道他們正面臨風險。沒有人記錄和分享有關此程式庫的經驗，他們沒有意識到風險。
- 您已在內部共用的微型服務中找出導致工作階段終止的邊緣案例。您已更新對服務的呼叫，以避免此邊緣案例。組織中的其他團隊不知道他們正面臨風險。
- 您已找到一個方法，可大幅降低其中一個微型服務所需的 CPU 使用率。您不知道是否有任何其他團隊可以利用此技術。

建立此最佳實務的優勢：分享經驗教訓以支援改進並最大限度地發揮經驗的優勢。

未建立此最佳實務時的曝險等級：低

實作指引

- 記錄和分享獲得的經驗：制定程序來記錄從執行營運活動和追溯性分析中學到的經驗教訓，以便其他團隊可以使用。
- 分享經驗：制定程序在團隊之間分享經驗教訓和相關成品。例如，透過可存取的 Wiki 分享更新的程序、指南、管控和最佳實務；透過公共儲存庫分享指令碼、程式碼和程式庫。
 - 利用 [AWS re:Post Private](#) 做為知識服務，讓組織內的協作和知識共享更順暢。

資源

相關的最佳實務：

- [OPS02-BP06 團隊之間的責任是預先定義或經過協商的](#)
- [OPS05-BP01 使用版本控制](#)

- [OPS05-BP06 共用設計標準](#)
- [OPS11-BP03 實作回饋迴圈](#)
- [OPS11-BP07 執行營運指標審查](#)

相關文件：

- [利用 AWS re:Post Private 增進協作並安全地共享雲端知識](#)
- [利用文件即程式碼解決方案減少專案延遲](#)

相關影片：

- [AWS re:Invent 2023 - 利用 AWS re:Post Private 在您的公司內協作並與 AWS 合作](#)
- [支援 為您提供支援 | 探索事件管理桌上模擬演練](#)

OPS11-BP09 分配時間進行改善

在流程中投入時間和資源，以持續逐漸改善。

預期成果：

- 您可以建立臨時環境複本，從而降低試驗和測試的風險、工作量及成本。
- 這些重複的環境可用於測試從您的分析、試驗和開發得出的結論，以及測試計劃的改善。
- 您會執行遊戲日，並使用 Fault Injection Service (FIS) 提供團隊在類似生產環境中執行實驗所需的控制項和防護機制。

常見的反模式：

- 您的應用程式伺服器存在已知的效能問題。它會新增到每個計劃功能實作的待辦項目中。如果計劃功能的新增速率保持不變，則效能問題永遠不會解決。
- 為協助持續改進，您核准管理員和開發人員使用他們額外的時間來選取和實作改進項目。改進永遠不會有完成的一天。
- 操作驗收已完成，您不會再測試操作實務。

建立此最佳實務的優勢：透過在程序中投入時間和資源，您可以實現持續逐漸改善。

未建立此最佳實務時的曝險等級：低

實作指引

- 分配時間進行改進：在流程中投入時間和資源，以持續逐漸改善。
- 實作變更以改進和評估結果，從而確定成功與否。
- 如果結果未能達到目標，並且改進仍然是優先事項，則應採取替代行動方案。
- 在演練日模擬生產工作負載，並利用這些模擬中的知識進行改進。

資源

相關的最佳實務：

- [OPS05-BP08 使用多個環境](#)

相關影片：

- [AWS re : Invent 2023 - 使用 AWS Fault Injection Service 改善應用程式彈性](#)

安全

安全支柱包含能夠保護資料、系統和資產，以利用雲端技術來改善安全性。可以在[安全支柱白皮書](#)中找到實作指引。

最佳實務領域

- [安全基礎](#)
- [身分與存取管理](#)
- [偵測](#)
- [基礎設施保護](#)
- [資料保護](#)
- [事件反應](#)
- [應用程式安全](#)

安全基礎

問題

- [SEC 1. 如何安全地操作工作負載？](#)

SEC 1. 如何安全地操作工作負載？

若要安全地操作工作負載，您必須將整體的最佳實務套用到每個安全區域。採用您在組織和工作負載層級於卓越營運中定義的要求和程序，並將其應用於所有區域。掌握最新的 AWS 和產業建議以及威脅情報，可協助您發展威脅模型並控制目標。自動化安全程序、測試和驗證，讓您能夠擴展安全操作。

最佳實務

- [SEC01-BP01 使用帳戶區隔工作負載](#)
- [SEC01-BP02 保護帳戶根使用者和屬性](#)
- [SEC01-BP03 識別和驗證控制目標](#)
- [SEC01-BP04 隨時掌握安全威脅和建議](#)
- [SEC01-BP05 減少安全管理範圍](#)
- [SEC01-BP06 自動化標準安全控制的部署](#)
- [SEC01-BP07 使用威脅模型識別威脅並優先考慮緩解措施](#)
- [SEC01-BP08 定期評估和實作新的安全服務和功能](#)

SEC01-BP01 使用帳戶區隔工作負載

透過多帳戶策略在環境 (例如生產、開發和測試) 與工作負載之間建立共通的防護機制和隔離。強烈建議帳戶層級的區隔，因為這在安全性、帳單和存取方面提供了有力的隔離界限。

預期成果：一種帳戶結構，可將雲端作業、不相關的工作負載和環境隔離為單獨的帳戶，從而提高整個雲端基礎設施的安全性。

常見的反模式：

- 將多個具有不同資料敏感度等級且不相關的工作負載置於相同的帳戶中。
- 定義不良的組織單位 (OU) 結構。

建立此最佳實務的優勢：

- 若工作負載遭到意外存取，縮小影響範圍。
- 集中管控對 AWS 服務、資源和區域的存取。
- 利用政策以及集中管理安全服務，維護雲端基礎設施的安全性。

- 自動化帳戶建立和維護流程。
- 集中稽核您的基礎設施以滿足合規性和法規需求。

未建立此最佳實務時的風險暴露等級：高

實作指引

AWS 帳戶 在以不同的敏感度等級操作的工作負載或資源之間提供安全隔離界限。AWS 提供工具透過多帳戶策略大規模管理您的雲端工作負載，以利用此隔離界限。如需有關 AWS 中的多帳戶策略的概念、模式和實作指引，請參閱 [Organizing Your AWS Environment Using Multiple Accounts](#)。

當您集中管理多個 AWS 帳戶 時，應該將您的帳戶組織成由組織單位 (OU) 層定義的階層。接著可以組織安全控制並套用至 OU 和成員帳戶，在組織內的成員帳戶上建立一致的預防性控制。安全控制是繼承的，讓您能夠篩選位於 OU 階層較低層級的成員帳戶可用的許可。良好的設計可利用此繼承關係來降低必要的安全政策數目和複雜度，達成每個成員帳戶預期的安全控制。

[AWS Organizations](#) 和 [AWS Control Tower](#) 是可以用來在 AWS 環境中實作和管理此多帳戶結構的兩個服務。AWS Organizations 可讓您將帳戶組織到由一個或多個 OU 層定義的階層中，每個 OU 都包含許多成員帳戶。[服務控制政策](#) (SCP) 可讓組織管理員在成員帳戶上建立精細的預防性控制，並且 [AWS Config](#) 可用於在成員帳戶上建立主動性和偵測性控制。許多 AWS 服務與 [AWS Organizations](#) [整合](#)，以提供委派的管理控制，並在組織的所有成員帳戶中執行服務特定任務。

[AWS Control Tower](#) 位於 AWS Organizations 之上，為具有[登陸區域](#)的多帳戶 AWS 環境提供一鍵式最佳實務設定。該登陸區域是通往由 Control Tower 所建立之多帳戶環境的進入點。與 AWS Organizations 相比，Control Tower 具有數個[好處](#)。提供改進的帳戶管控的三個優點是：

- 整合式強制性安全控制，會自動套用至獲准加入組織的帳戶。
- 選擇性控制，可針對指定 OU 集合開啟或關閉。
- [AWS Control Tower Account Factory](#) 提供自動化帳戶部署，其中包含組織內部預先核准的基準和組態選項。

實作步驟

1. 設計組織單位結構：設計合理的組織單位結構可減少建立及維護服務控制政策及其他安全性控制所需的管理負擔。您的組織單位結構應[與業務需求、資料敏感度和工作負載結構保持一致](#)。
2. 為多帳戶環境建立登陸區域：登陸區域可提供一致的安全性和基礎設施基礎，您的組織可以從中快速開發、啟動和部署工作負載。可以使用[自訂建置的登陸區域或 AWS Control Tower](#) 來協調您的環境。

3. 建立防護機制：透過登陸區域為您的環境實作一致的安全防護機制。AWS Control Tower 提供可部署的[強制性](#)控制與[選擇性](#)控制清單。實作 Control Tower 時會自動部署強制性控制。審核強烈建議的控制和選擇性控制清單，並實作符合您需求的控制。
4. 限制對新增區域的存取：對於新 AWS 區域，IAM 資源 (例如使用者和角色) 只會傳播到您指定的區域。可以在[使用 Control Tower 時透過主控台](#)執行此動作，或[在 AWS Organizations 中調整 IAM 權限政策](#)。
5. 考慮 AWS [CloudFormation StackSets](#)：StackSets 可協助您將資源 (包括 IAM 政策、角色和群組) 從核准的範本中部署到不同的 AWS 帳戶 帳戶和區域。

資源

相關的最佳實務：

- [SEC02-BP04 仰賴集中式身分提供者](#)

相關文件：

- [AWS Control Tower](#)
- [《AWS 安全性稽核指南》](#)
- [IAM 最佳實務](#)
- [使用 CloudFormation StackSets 在多個 AWS 帳戶 和區域中佈建資源](#)
- [Organizations 常見問答集](#)
- [AWS Organizations 術語與概念](#)
- [AWS Organizations 多帳戶環境中服務控制政策的最佳實務](#)
- [AWS 帳戶管理參考指南](#)
- [使用多個帳戶組織您的 AWS 環境](#)

相關影片：

- [透過自動化和管控大規模採用 AWS](#)
- [以 Well-Architected 方式提供安全最佳實務](#)
- [使用 AWS Control Tower 建置和管控多個帳戶](#)
- [為現有組織啟用 Control Tower](#)

SEC01-BP02 保護帳戶根使用者和屬性

根使用者是 AWS 帳戶 中最具特權的使用者，對帳戶內的所有資源具備完整的管理存取權，並且在某些情況下，不受安全政策的限制。停用對根使用者的程式設計存取，為根使用者建立適當的控制，以及避免例行使用根使用者，可降低意外暴露根憑證及後續危及雲端環境的風險。

預期成果：保護根使用者有助於減少因濫用根使用者憑證而造成意外或蓄意損壞的機會。建立偵測控制也能在當使用根使用者採取動作時警告適當的人員。

常見的反模式：

- 將根使用者用於需要根使用者憑證以外的工作。
- 疏於定期測試緊急應變計劃以確認重大基礎設施、程序和人員在緊急情況下的運作情形。
- 僅考慮一般帳戶登入流程而疏於考慮或測試替代帳戶復原方法。
- 未將 DNS、電子郵件伺服器 and 電話提供者作為重要安全周邊的一部分來處理，因為其會用於帳戶復原流程。

建立此最佳實務的優勢：保護對根使用者的存取，可建立對帳戶中的動作加以控制和稽核的信心。

未建立此最佳實務時的風險暴露等級：高

實作指引

AWS 提供眾多工具來協助保護您的帳戶。然而，由於預設情況下不會開啟其中一些措施，因此您必須採取直接行動加以實作。考慮將這些建議作為保護 AWS 帳戶 的基本步驟。實作這些步驟時，務必建立程序以持續評估和監視安全控制。

首次建立 AWS 帳戶 時，您是從一個對帳戶中所有 AWS 服務和資源具有完全存取權的身分開始。此身分稱為 AWS 帳戶 根使用者。您可以使用您用來建立帳戶的電子郵件地址和密碼，以根使用者的身分登入。由於授予給 AWS 根使用者更高的存取權限，您必須限制 AWS 根使用者的使用，才能執行[特別需要它](#)的任務。根使用者登入憑證必須受嚴密防護，並且多重要素驗證 (MFA) 應始終用於 AWS 帳戶 根使用者。

除了一般驗證流程 (使用使用者名稱、密碼和多重要素驗證 (MFA) 裝置登入根使用者) 之外，還有帳戶復原流程會登入 AWS 帳戶 根使用者，而其能夠存取與您的帳戶相關聯的電子郵件地址和電話號碼。因此，保護傳送復原電子郵件的根使用者電子郵件帳戶以及與帳戶相關聯的電話號碼同樣也很重要。另外，對於與根使用者相關聯的電子郵件地址託管在相同 AWS 帳戶 的電子郵件伺服器或網域名稱服務 (DNS) 資源上的情況，也要考慮可能的循環相依性。

使用 AWS Organizations 時，會有多個 AWS 帳戶，各自都有根使用者。將一個帳戶指定為管理帳戶，接著可以在該管理帳戶之下新增數層成員帳戶。優先保護您的管理帳戶根使用者後，再來處理成員帳戶根使用者。保護管理帳戶根使用者的策略可不同於成員帳戶根使用者，而且您可以對成員帳戶根使用者設立預防性安全控制。

實作步驟

以下是為根使用者建立控制的建議實作步驟。如果適用，建議將交叉引用到 [CIS AWS Foundations 基準版本 1.4.0](#)。除了這些步驟之外，請參閱[保護您的 AWS 帳戶 及其資源的 AWS 最佳實務指引](#)。

預防性控制

1. 為帳戶設定準確的[聯絡資訊](#)。
 - a. 此資訊會用於遺失密碼復原流程、遺失 MFA 裝置帳戶復原流程，以及與您的團隊進行重大安全相關通訊。
 - b. 使用由您的企業網域所託管的電子郵件地址 (最好是使用分發清單) 作為根使用者的電子郵件地址。使用分發清單而不是個人的電子郵件帳戶可對長期存取根帳戶提供額外的備援和持續性。
 - c. 聯絡資訊上所列的電話號碼應該是針對此用途的專用安全電話。不應公布或與他人共用電話號碼。
2. 請勿為根使用者建立存取金鑰。若存在存取金鑰，請將其移除 (CIS 1.4)。
 - a. 去除根使用者任何長期存留的程式設計憑證 (存取和秘密金鑰)。
 - b. 如果根使用者存取金鑰已存在，您應該使用這些金鑰來轉換程序，以使用來自 AWS Identity and Access Management (IAM) 角色的臨時存取金鑰，然後[刪除根使用者存取金鑰](#)。
3. 確定您是否需要儲存根使用者的憑證。
 - a. 如果您使用 AWS Organizations 建立新成員帳戶，則成員帳戶上的根使用者的初始密碼會設為隨機值，並且不會向您公開。如有需要，請考慮使用來自 AWS 組織管理帳戶的密碼重設流程，[以取得成員帳戶的存取權](#)。
 - b. 對於獨立 AWS 帳戶 或管理 AWS 組織帳戶，請考慮建立根使用者的憑證並安全存放。為根使用者使用 MFA。
4. 在 AWS 多帳戶環境中為成員帳戶根使用者使用預防性控制。
 - a. 考慮針對成員帳戶使用[不允許為根使用者建立根存取金鑰](#)預防性防護機制。
 - b. 考慮針對成員帳戶使用[不允許以根使用者身分執行動作](#)預防性防護機制。
5. 如果您需要根使用者的憑證：
 - a. 使用複雜密碼。

- b. 為根使用者開啟多重要素驗證 (MFA)，尤其是 AWS Organizations 管理 (付款人) 帳戶 (CIS 1.5)。
 - c. 考慮硬體 MFA 裝置以獲得彈性和安全性，因為一次性裝置可減少包含 MFA 代碼的裝置重複用於其他用途的可能性。確認定期更換使用電池的硬體 MFA 裝置。(CIS 1.6)
 - 若要為根使用者設定 MFA，請遵循建立[虛擬 MFA](#) 或[硬體 MFA 裝置](#)的指示。
 - d. 考慮註冊多個 MFA 裝置進行備份。[每個帳戶最多允許 8 個 MFA 裝置](#)。
 - 請注意，[如果 MFA 裝置遺失](#)，為根使用者註冊多個 MFA 裝置會自動關閉復原帳戶的流程。
 - e. 請將密碼妥善保管，如果以電子方式儲存密碼，請考慮循環相依性。儲存密碼時，請勿以需要存取相同的 AWS 帳戶 來取得密碼的方式儲存。
6. 選擇性：考慮為根使用者建立定期密碼輪流排程。
- 憑證管理最佳實務取決於您法規和政策需求。受 MFA 保護的根使用者不依賴把密碼當作單一驗證要素。
 - 定期[變更根使用者密碼](#)可降低不慎公開密碼遭到濫用的風險。

偵測性控制

- 建立警示以偵測根憑證的使用 (CIS 1.7)。[Amazon GuardDuty](#) 可以透過 [RootCredentialUsage](#) 調查結果來監控根使用者 API 憑證使用情況，並發出提醒。
- 評估並實作[適用於 AWS Config 的《AWS Well-Architected 安全支柱》一致性套件](#)中包含的偵測控制，或者如果使用 AWS Control Tower，則評估並實作 Control Tower 內可用的[強烈建議的控制](#)。

操作指引

- 確定組織內誰應該存取根使用者憑證。
 - 使用雙人規則，如此沒有單獨一人可以存取所有必要的憑證和 MFA 來取得根使用者存取權。
 - 確認組織而不是單一個人持有對與帳戶相關聯的電話號碼和電子郵件別名 (用於密碼重設和 MFA 重設程序) 的控制權。
- 只在特殊情況下使用根使用者 (CIS 1.7)。
 - AWS 根使用者不可用於日常任務，即使管理任務也一樣。只有以根使用者身分登入，才能執行[需要根使用者的 AWS 任務](#)。所有其他動作都應該由其他擔任適當角色的使用者執行。
- 定期檢查根使用者的存取權操作正常，以便在發生需要使用根使用者憑證的緊急情況之前，測試相關程序。

- 定期檢查與帳戶相關聯的電子郵件地址，以及[替代連絡人](#)下列出的電子郵件地址。監控這些電子郵件收件匣，查看您可能接收的來自 <abuse@amazon.com> 的安全通知。另外確保與帳戶相關聯的任何電話號碼都有效。
- 準備事件回應程序以回應根帳戶誤用的情況。請參閱 [AWS Security Incident Response Guide](#) 和 [安全支柱白皮書中「事件回應」一節](#)中的最佳實務，了解如何為 AWS 帳戶建立事件回應策略的詳細資訊。

資源

相關的最佳實務：

- [SEC01-BP01 使用帳戶區隔工作負載](#)
- [SEC02-BP01 使用強式登入機制](#)
- [SEC03-BP02 授予最低權限存取權](#)
- [SEC03-BP03 建立緊急存取程序](#)
- [SEC10-BP05 預先佈建存取權](#)

相關文件：

- [AWS Control Tower](#)
- [《AWS 安全性稽核指南》](#)
- [IAM 最佳實務](#)
- [Amazon GuardDuty – 根憑證使用情況警示](#)
- [透過 CloudTrail 監控根憑證使用情況的逐步指引](#)
- [已核准可與 AWS 搭配使用的 MFA 權杖](#)
- [在 AWS 上實作緊急存取](#)
- [在 AWS 帳戶 中要改善的十大安全項目](#)
- [如果發現我的 AWS 帳戶 中有未經授權的活動，該怎麼辦？](#)

相關影片：

- [透過自動化和管控大規模採用 AWS](#)
- [以 Well-Architected 方式提供安全最佳實務](#)

- [限制使用 AWS 根憑證](#)，來自 AWS re:inforce 2022 – AWS IAM 安全最佳實務

SEC01-BP03 識別和驗證控制目標

根據合規需求以及從威脅模型識別的風險，衍生並驗證您需要套用到工作負載的控制目標和控制。對控制目標與控制持續進行驗證，可協助您測量風險降低的有效性。

預期成果：企業的安全控制目標是明確定義的，並符合您的合規要求。控制項是透過自動化和政策來實作和強制執行的，並持續評估其在達成目標方面的有效性。在一個時間點和一段時間內的有效性證據可以很容易地向稽核人員報告。

常見的反模式：

- 您的企業對可保證安全的法規要求、市場預期和業界標準並未充分了解
- 網路安全架構和控制目標不符合業務需求
- 控制措施的實作並未以可衡量的方式與您的控制目標保持一致
- 您不使用自動化來報告控制措施的有效性

未建立此最佳實務時的曝險等級：高

實作指引

有許多常見的網路安全框架可以構成安全控制目標的基礎。考慮企業的法規要求、市場期望和業界標準，以決定哪些架構最能支援您的需求。範例包括 [AICPA SOC 2](#)、[HITRUST](#)、[PCI-DSS](#)、[ISO 27001](#) 和 [NIST SP 800-53](#)。

針對您確定的控制目標，了解您使用的 AWS 服務如何協助您達成這些目標。使用 [AWS Artifact](#) 尋找符合目標架構的文件和報告，這些文件和報告可說明 AWS 涵蓋的責任範圍，以及您負責的剩餘範圍的指引。如需符合各種架構控制聲明的詳細服務特定指引，請參閱 [AWS Customer Compliance Guides](#)。

當您定義實現目標的控制措施時，使用預防性控制措施對執行進行整理，並使用偵測性控制來自動化緩解措施。使用[服務控制政策 \(SCP\)](#)，協助防止您的 AWS Organizations 中的不合規資源組態和動作。在 [AWS Config](#) 中實作規則以監控和報告不合規的資源，然後在對其行為有信心後，將規則切換為強制執行模型。若要部署符合網路安全架構的預定義和受控規則集，請首先評估 [AWS Security Hub CSPM 標準](#) 的使用。AWS Foundational Service Best Practices (FSBP) 標準和 CIS AWS Foundations Benchmark 是良好的起點，具有與多個標準框架共享的許多目標保持一致的控制措施。如果 Security Hub CSPM 本質上沒有所需的控制偵測，則可以使用 [AWS Config 一致性套件](#) 來補充它。

使用 AWS 全球安全與合規加速 (GSCA) 團隊建議的 [APN 合作夥伴套件](#)，在需要時從安全顧問、諮詢機構、證據收集和報告系統、稽核人員和其他輔助服務那裡取得協助。

實作步驟

1. 評估常見的網路安全架構，並使您的控制目標與選擇的目標保持一致。
2. 使用 AWS Artifact 取得有關架構指引和責任的相關文件。了解哪些法規遵循部分屬於共同責任模式的 AWS 方面，以及哪些部分是您的責任。
3. 使用 SCP、資源政策、角色信任政策及其他防護機制來防止不合規的資源組態和動作。
4. 評估部署符合您控制目標的 Security Hub CSPM 標準和 AWS Config 一致性套件。

資源

相關的最佳實務：

- [SEC03-BP01 定義存取需求](#)
- [SEC04-BP01 設定服務和應用程式日誌記錄](#)
- [SEC07-BP01 了解您的資料分類方案](#)
- [OPS01-BP03 評估管控要求](#)
- [OPS01-BP04 評估合規要求](#)
- [PERF01-BP05 使用政策和參考架構](#)
- [COST02-BP01 根據貴組織的需求制定政策](#)

相關文件：

- [AWS 客戶合規指南](#)

相關工具：

- [AWS Artifact](#)

SEC01-BP04 隨時掌握安全威脅和建議

透過監控業界威脅情報刊物和資料摘要以取得更新，以便隨時掌握最新的威脅和緩解措施。評估根據最新威脅資料自動更新的受管服務產品。

預期成果：隨著產業刊物更新，隨時掌握最新威脅和建議的資訊。在識別新威脅時，使用自動化技術偵測潛在的漏洞和暴露。您可以採取緩解措施對抗這些威脅。您採用 AWS 的服務會自動更新為最新的威脅情報。

常見的反模式：

- 沒有可靠且可反覆執行的機制，因而無法隨時掌握新威脅情報。
- 手動維護技術產品組合、工作負載和相依項的庫存清單，而這些都需要人員審查才能得知是否有潛在的漏洞和暴露。
- 沒有既定的機制能夠將工作負載和相依項更新到可用的最新版本，因而無法取得已知的威脅緩解措施。

建立此最佳實務的優勢：使用威脅情報來源隨時掌握新資訊，即可盡量避免錯過可能影響業務的威脅態勢中發生的重大變化。採用自動化的方式取代手動，以掃描、偵測和修復工作負載及其相依項中存在的潛在漏洞或暴露，如此就能協助您事先預測並快速降低風險。這有助於控制與漏洞緩解相關的時間和成本。

未建立此最佳實務時的曝險等級：高

實作指引

檢閱值得信賴的威脅情報刊物，以掌握威脅態勢。如需已知對手戰術、技術和程序的文件（ ），請參閱 [MITRE ATT&CK](#) 知識庫TTPs。檢閱 MITRE的[常見漏洞與暴露](#)（ CVE ）清單，以隨時掌握您所依賴產品中已知漏洞的資訊。使用 Open Worldwide Application Security Project（ OWASP ）的熱門[OWASP前 10 大](#)專案，了解 Web 應用程式的關鍵風險。

使用 AWS 適用於 AWS 的安全[公告](#)，隨時掌握安全事件和建議的修復步驟CVEs。

為了減少您保持最新狀態的整體工作量和額外負荷，請考慮使用 AWS 服務，隨著時間自動整合新的威脅情報。例如，[Amazon GuardDuty](#) 會隨時掌握最新的產業威脅情報，以偵測帳戶中的異常行為和威脅簽章。[Amazon Inspector](#) 會自動將CVEs用於其持續掃描功能的 資料庫保持在最新狀態。[AWS WAF](#) 和 [AWS Shield Advanced](#) 兩者都提供受管規則群組，這些群組會隨著新的威脅出現時自動更新。

檢閱 [Well-Architected 卓越營運支柱](#)，了解自動機群管理和修補的資訊。

實作步驟

- 訂閱與您的業務和產業相關的威脅情報刊物更新。訂閱 AWS 安全公告。

- 考慮採用自動整合新威脅情報的服務，例如 Amazon GuardDuty 和 Amazon Inspector。
- 部署符合 Well-Architected 卓越營運支柱最佳實務的機群管理和修補策略。

資源

相關的最佳實務：

- [SEC01-BP07 使用威脅模型識別威脅並排定緩解的優先順序](#)
- [OPS01-BP05 評估威脅態勢](#)
- [OPS11-BP01 擁有持續改善的程序](#)

SEC01-BP05 減少安全管理範圍

判斷您是否可以使用將某些控制項的管理轉移到 AWS（受管服務）的服務 AWS 來減少安全範圍。這些服務有助於減少安全維護任務，例如基礎設施佈建、軟體設定、修補或備份。

預期結果：您在為工作負載選取 AWS 服務時，會考慮安全管理的範圍。除了其他 Well-Architected 考量之外，管理額外負荷和維護任務的成本（總擁有成本，或 TCO）會權衡您選擇的服務成本。您可以將 AWS 控制和合規文件納入控制評估和驗證程序。

常見的反模式：

- 部署工作負載時，並未徹底了解您所選取服務的共用責任模式。
- 在虛擬機器上託管資料庫和其他技術時，未先行評估對等的受管服務。
- 比較受管服務選項時，未將安全管理任務納入虛擬機器上託管技術的總體擁有成本中。

建立此最佳實務的優勢：使用受管服務可以減輕您管理營運安全控制的整體負擔，進而降低安全風險和總體擁有成本。否則，時間可能會花費在某些安全任務上，而無法轉投入其他為企業創造更多價值的任務。受管服務也可以透過將某些控制要求轉移到 AWS，以縮小合規要求的範圍。

未建立此最佳實務時的曝險等級：中

實作指引

您可以透過多種方式在 AWS 上整合工作負載的元件。在 Amazon EC2 執行個體上安裝和執行技術通常需要您承擔整體安全責任的最大份額。為了協助減少操作特定控制項的負擔，請識別可降低您共同責任模型範圍的 AWS 受管服務，並了解如何在現有架構中使用它們。範例包括使用 [Amazon](#)

[Relational Database Service \(Amazon RDS \)](#) 部署資料庫、使用 [Amazon Elastic Kubernetes Service \(Amazon EKS \)](#) 或 [Amazon Elastic Container Service \(Amazon ECS \)](#) 協調容器，或使用[無伺服器選項](#)。建置新的應用程式時，請仔細考量哪些服務有助於減少實作和管理安全控制方面的時間和成本。

合規要求也是選取服務時的考量因素。受管服務可以將某些要求的合規轉移到 AWS。與您的合規團隊討論其稽核相關稽核 AWS 報告中所操作之服務的各個層面，以及管理和接受控制陳述式的自在程度。您可以[AWS Artifact](#)向稽核人員或監管機構提供 中的稽核成品，作為 AWS 安全控制的證據。您也可以使用某些 AWS 稽核成品提供的責任指南，以及[AWS 客戶合規指南](#) 來設計您的架構。本指引可協助您確定應採取哪些額外的安全控制，以便支援系統的特定使用案例。

使用 受管服務時，請熟悉將資源更新至較新版本的程序（例如，更新 Amazon 管理的資料庫版本 RDS，或 AWS Lambda 函數的程式設計語言執行時間）。雖然受管服務可能會自動為您執行此操作，但設定更新的時間並了解對操作的影響仍然是您的責任。[AWS Health](#) 等工具可以協助您在整個環境中追蹤和管理這些更新。

實作步驟

1. 評估可取代為受管服務的工作負載元件。
 - a. 如果您要將工作負載遷移至 AWS，請考慮減少管理（時間和費用）並降低風險，同時評估您是否應該重新託管、重構、修改、重建或取代工作負載。有時候，在一開始遷移時的額外投資，長遠來看可能帶來大幅的節省。
2. 請考慮實作 受管服務，例如 Amazon RDS，而不是安裝和管理您自己的技術部署。
3. 使用 中的責任指引 AWS Artifact，協助判斷您應該為工作負載設定的安全控制。
4. 保留使用中的資源清單，並持續 up-to-date 使用新的服務和方法，以識別減少範圍的新機會。

資源

相關的最佳實務：

- [PERF02-BP01 為您的工作負載選取最佳運算選項](#)
- [PERF03-BP01 使用最能夠支援資料存取和儲存需求的專用資料存放區](#)
- [SUS05-BP03 使用受管服務](#)

相關文件：

- [的計劃生命週期事件 AWS Health](#)

相關工具：

- [AWS Health](#)
- [AWS Artifact](#)
- [AWS 客戶合規指南](#)

相關影片：

- [如何使用 遷移至 Amazon RDS 或 Aurora MySQL 資料庫執行個體 AWS DMS ?](#)
- [AWS re : Invent 2023 - 使用 大規模管理資源生命週期事件 AWS Health](#)

SEC01-BP06 自動化標準安全控制的部署

在開發和部署整個 AWS 環境的標準安全控制時，套用現代 DevOps 實務。使用基礎架構作為程式碼 (IaC) 範本定義標準安全控制和組態、擷取版本控制系統中的變更、測試變更作為 CI/CD 管道的一部分，以及自動部署 AWS 環境的變更。

預期成果：IaC 範本會擷取標準化的安全控制，並將其遞交至版本控制系統。CI/CD 管道位於可偵測變更、自動測試和部署 AWS 環境的位置。防護機制準備好在進行部署之前，先偵測範本中的組態錯誤並發出提醒。工作負載會部署到已採取標準控制的環境中。團隊有權透過自助服務機制部署經核准的服務組態。已制定安全的備份和復原策略來控制組態、指令碼和相關資料。

常見的反模式：

- 透過 Web 主控台或命令列介面，手動變更標準安全控制。
- 仰賴個別工作負載團隊手動實作中央團隊定義的控制。
- 仰賴中央安全團隊應工作負載團隊的要求部署工作負載層級的控制。
- 允許相同的個人或團隊開發、測試和部署安全控制自動化指令碼，而未能妥善區分職責，或適當地對其加以制衡。

建立此最佳實務的優勢：使用範本定義標準安全控制，可讓您使用版本控制系統追蹤和比較一段時間的變更。使用自動化方式測試和部署變更可建立標準化和可預測性，從而提高成功部署的機會，並減少手動重複任務。為工作負載團隊提供自助服務機制來部署核准的服務和組態，可降低組態錯誤和濫用的風險。這也有助於讓團隊及早在開發過程中納入控制。

未建立此最佳實務時的風險暴露等級：中

實作指引

遵循 [SEC01-BP01 中所述的實務 使用帳戶 分隔工作負載](#)時，您會針對使用 管理的不同環境，最終獲得多個 AWS 帳戶 AWS Organizations。雖然其中每個環境和工作負載可能需要不同的安全控制，但您可以將整個組織的某些安全控制標準化。範例包括整合集中式身分提供者、定義網路和防火牆，以及設定用於儲存和分析日誌的標準位置。同樣地，您可以使用基礎設施即程式碼 (IaC) 將同樣嚴謹的應用程式程式碼開發程序套用至基礎設施佈建，也可以使用 IaC 來定義和部署標準安全控制。

請盡可能以宣告的方式定義安全控制 (例如在 [AWS CloudFormation](#) 中)，並將其儲存在來源控制系統中。使用 DevOps 實務來自動化部署控制項以取得更可預測的版本、使用 等工具進行自動測試 [AWS CloudFormation Guard](#)，以及偵測部署控制項與所需組態之間的偏離。您可以使用 [AWS CodePipeline](#)、[AWS CodeBuild](#) 和 [AWS CodeDeploy](#) 等服務來建構 CI/CD 管道。請考慮 [使用多個帳戶組織您的 AWS 環境](#)，在與其他部署管道分開的自己的帳戶中設定這些服務的指南。

您也可以定義範本，以標準化定義和部署、AWS 帳戶服務和組態。此技術可讓中央安全團隊管理這些定義，並透過自助服務方法將其提供給工作負載團隊。實現這一點的方法之一是使用 [Service Catalog](#)，您可以在其中將範本發佈為產品，讓工作負載團隊可以將這些產品納入自己的管道部署中。如果您使用的是 [AWS Control Tower](#)，某些範本和控制可以作為起點。Control Tower 還提供 [Account Factory](#) 功能，可讓工作負載團隊使用您定義的標準建立新的 AWS 帳戶。此功能有助於消除對中央團隊的依賴，以在工作負載團隊視需要識別新帳戶時核准和建立新帳戶。您可能需要這些帳戶，以便根據諸如提供的功能、所處理資料的敏感性或其行為等原因，區隔不同的工作負載元件。

實作步驟

1. 確定如何在版本控制系統中儲存和維護範本。
2. 建立 CI/CD 管道以測試和部署您的範本。定義測試以檢查是否有組態錯誤，以及範本是否符合您公司的標準。
3. 為工作負載團隊建立標準化範本的目錄，以根據您的需求部署 AWS 帳戶 和服務。
4. 針對控制組態、指令碼和相關資料實作安全的備份和復原策略。

資源

相關的最佳實務：

- [OPS05-BP01 使用版本控制](#)
- [OPS05-BP04 使用建置和部署管理系統](#)
- [REL08-BP05 透過自動化部署變更](#)
- [SUS06-BP01 採用可快速引入永續發展改進的方法](#)

相關文件：

- [使用多個帳戶組織您的 AWS 環境](#)

相關範例：

- [使用 Service Catalog 自動化帳戶建立和資源佈建 AWS Organizations，以及 AWS Lambda](#)
- [使用 AWS Secrets Manager、和 增強 DevOps 管道 AWS KMS並保護資料 AWS Certificate Manager](#)

相關工具：

- [AWS CloudFormation Guard](#)
- [上的登陸區域加速器 AWS](#)

SEC01-BP07 使用威脅模型識別威脅並優先考慮緩解措施

執行威脅建模，為您的工作負載識別並保有潛在威脅及關聯緩解措施的最新記錄。排定威脅的優先順序並調整安全控制緩解措施，以防止、偵測和回應威脅。就您的工作負載的情況，以及不斷演變的安全態勢，重新檢視和維護此工作。

未建立此最佳實務時的曝險等級：高

實作指引

什麼是威脅建模？

「威脅建模以保護有價值物為目標，識別、溝通和了解威脅及緩解措施。」 – [開放式 Web 應用程式安全專案 \(OWASP\) 應用卡程式威脅建模](#)

為何使用威脅模型？

系統本身錯綜複雜，並且隨時間變得更複雜且更具能力，從而實現更大的商業價值及更高的客戶滿意度和參與度。這意味著 IT 設計決策需要考慮不斷增加的使用案例數量。這種複雜性和使用案例數量的排列通常使得非結構化方法無法有效尋找和緩解威脅。反之，您需要一套系統化方法來列舉對系統的潛在威脅，以及制定緩解措施，並以這些緩解措施為優先來確保組織的有限資源能在改善系統整體安全狀態上發揮最大的影響力。

威脅建模旨在提供這套系統化方法，目的是要在設計過程中及早尋找和解決問題，此時進行緩解的成本和精力與生命週期稍後相比要來得低。此方法與[往前移安全性](#)的業界原則相一致。威脅建模最終會與組織的風險管理程序整合，透過使用威脅驅動的方法，協助推動要實作哪些控制決策。

何時應執行威脅建模？

在工作負載的生命週期中及早開始威脅建模，可給予您更大的彈性來決定要如何處理所識別的威脅。就跟軟體錯誤一樣，越早識別威脅，就能以越具成本效益的方式加以解決。威脅模型是不斷更新的文件，並且應該持續隨著工作負載的變更而演進。隨時間重新檢視您的威脅模型，包括當有重大變更、威脅形勢有變化，或是採用新功能或服務時。

實作步驟

我們如何執行威脅建模？

有許多不同的方式來執行威脅建模。就跟程式設計語言一樣，各有優缺點，而您應選擇最適合您的方式。其中一個方法是從 [Shostack 針對威脅建模的 4 個問題框架](#) 開始著手，當中提出自由回答的問題會為您的威脅建模練習提供結構：

1. 我們目前在做什麼？

此問題的目的是協助您了解正在建置的系統並對之取得一致的意見，以及該系統與安全相關的詳細資訊。建立模型或圖表是回答此問題最受歡迎的方法，因為這可協助您將正在建置的項目視覺化，例如使用[資料流程圖](#)。寫下關於您的系統的假設和重要詳細資訊也有助於您定義涵蓋的範圍。這可讓對參與威脅模型的每個人專注於相同的事情，並避免耗時地繞入不在範圍內的主題 (包括系統的過時版本)。舉例來說，如果您正在建置 Web 應用程式，可能不值得花時間為瀏覽器用戶端建立作業系統信任開機順序的模型，因為您無法透過您的設計對此產生影響。

2. 什麼可能出錯？

這是您識別對系統的威脅之處。威脅是意外或故意的動作或事件，會帶來不必要的影響，並且可能會影響系統安全。如果對可能出錯之處沒有清楚的了解，則無法對症下藥。

對於什麼可能出錯，您並沒有標準的清單可循。建立此清單需要團隊內的每個人與[涉及的相關角色](#)在威脅建模練習中集思廣益和共同協作。您可以使用識別威脅的模型來協助集思廣益，例如[STRIDE](#)，這會建議不同的類別以進行評估：詐騙、竄改、否認、資訊洩露、拒絕服務和提升權限。此外，您可能會想要審核現有的清單並研究以獲得靈感來協助集思廣益，包括[OWASP 前十大](#)、[HiTrust 威脅型錄](#)，以及您組織本身的威脅型錄。

3. 我們要做何處理？

就跟上一個問題一樣，對於所有可能的緩解措施並沒有標準的清單可循。此步驟的輸入是上一步識別的威脅、動作和改進之處。

安全與合規是[您和 AWS 之間共同責任](#)。了解當您提出「我們要做何處理？」時，也是在問「誰要對其負責？」，這一點很重要。了解您與 AWS 之間的責任制衡有助於您將威脅建模練習的範圍定在您控制之下的緩解措施，這通常是 AWS 服務組態選項與您自身的系統特定緩解措施的組合。

對於共同責任的 AWS 部分，您將發現 [AWS 服務在許多合規計畫的範圍之內](#)。這些計畫可協助您了解 AWS 在維護雲端安全和合規方面設立的強大控制。來自這些計畫的稽核報告可供 AWS 客戶從 [AWS Artifact](#) 下載。

無論您使用何種 AWS 服務，其始終涉及客戶責任，而您的威脅模型中應包含與這些責任一致的緩解措施。對於 AWS 服務本身的安全控制緩解措施，您要考慮跨領域實作安全控制，包括身分和存取管理 (驗證和授權)、資料保護 (靜態和傳輸中)、基礎設施安全、日誌記錄和監控等領域。每個 AWS 服務的文件都有[專屬的安全章節](#)，提供將安全控制視為緩解措施的指引。重要的是，考慮您正在編寫的程式碼及其程式碼相依性，並思考您可以設立以解決這些威脅的控制。這些控制可以是[輸入驗證](#)、[工作階段處理](#)和[界限處理](#)等事項。大多數漏洞通常是在自訂程式碼中引入，因此請專注於此區域。

4. 我們處理得當嗎？

目標是讓您的團隊和組織改進威脅模型的品質以及隨時間執行威脅建模的速度。這些改進出自練習、學習、教導和審查的組合。若要更加深入並實作，建議您與您的團隊完成[建置人員建立威脅模型的正確方式訓練課程](#)或[研討會](#)。此外，如果您正在尋找有關如何將威脅建模整合至您組織的應用程式開發生命週期，請參閱 AWS 安全部落格上的[如何進行威脅建模](#)文章。

威脅編寫器

為了協助並指導您執行威脅建模，請考慮使用[威脅編寫器](#)工具，該工具旨在縮短威脅建模實現價值的時間。此工具可協助您執行下列操作：

- 撰寫與[威脅文法](#)相符、可在自然非線性工作流程中使用的有用威脅陳述式
- 產生人類可讀的威脅模型
- 產生機器可讀的威脅模型，以讓您將威脅模型視為程式碼
- 使用洞察儀表板協助您快速識別品質和涵蓋範圍有所改進的領域

如需進一步的參考，請造訪「威脅編寫器」，並切換到系統定義的範例工作區。

資源

相關的最佳實務：

- [SEC01-BP03 識別和驗證控制目標](#)
- [SEC01-BP04 隨時掌握安全威脅和建議](#)
- [SEC01-BP05 減少安全管理範圍](#)
- [SEC01-BP08 定期評估和實作新的安全服務和功能](#)

相關文件：

- [如何進行威脅建模 \(AWS 安全部落格\)](#)
- [NIST：以資料為中心的系統威脅建模指南](#)

相關影片：

- [AWS Summit ANZ 2021 - 如何進行威脅建模](#)
- [AWS Summit ANZ 2022 - 擴展安全性 – 針對快速和安全交付進行最佳化](#)

相關訓練：

- [建置人員建立威脅模型的正確方式 – AWS Skill Builder 虛擬自定進度培訓](#)
- [建置人員建立威脅模型的正確方式 – AWS 研討會](#)

相關工具：

- [威脅編寫器](#)

SEC01-BP08 定期評估和實作新的安全服務和功能

評估和實作 AWS 和 AWS 合作夥伴提供的安全服務和功能，讓您發展工作負載的安全狀態。

預期成果：您擁有一套標準實務，可通知您 AWS 和 AWS 合作夥伴發行的新功能和服務。您會評估這些新功能對於環境和工作負載目前和新控制的設計有何影響。

常見的反模式：

- 您未訂閱 AWS 部落格和 RSS 摘要，因此未能迅速得知相關的新功能和服務
- 您仰賴第二手來源得知安全服務和功能的最新消息和更新
- 您沒有鼓勵組織中的 AWS 使用者隨時掌握最新更新

建立此最佳實務的優勢：如果您能隨時掌握新的安全服務和功能，就可以在雲端環境和工作負載中實作控制方面做出明智的決策。這些來源有助於提高對於不斷變化的安全態勢的意識，以及說明如何使用 AWS 服務來防範新出現的威脅。

未建立此最佳實務時的曝險等級：低

實作指引

AWS 會透過幾種通道通知客戶新的安全服務和功能：

- [AWS 最新消息](#)
- [AWS 新聞部落格](#)
- [AWS 安全部落格](#)
- [AWS 安全公告](#)
- [AWS 文件概觀](#)

您可以使用 Amazon Simple Notification Service (Amazon SNS) 訂閱 [AWS 每日功能更新](#) 主題，以獲得完整的每日更新摘要。某些安全服務 (例如 [Amazon GuardDuty](#) 和 [AWS Security Hub CSPM](#)) 會提供自己的 SNS 主題，以持續掌握與這些特定服務有關的新標準、調查結果及其他更新。

每年全球各地也會舉行多場[會議、活動和網路研討會](#)，於會中宣佈並詳細描述新服務和功能。其中特別值得關注的是年度 [AWS re:Inforce](#) 安全會議，以及較為常態的 [AWS re:Invent](#) 會議。先前所述的 AWS 新聞頻道會分享這些有關安全和其他服務的會議公告，您可以在 YouTube 上的 [AWS 活動頻道](#) 線上觀看更深入的教育性分組會議。

您也可以向您的 [AWS 帳戶 團隊](#) 詢問有關最新安全服務更新和建議的資訊。如果沒有直接聯絡資訊，您可以透過[銷售人員支援表單](#)聯繫您的團隊。同樣地，如果您訂閱 [AWS 企業支援](#)，將會收到來自技術客戶經理 (TAM) 的每週更新，您可以與他們安排定期審查會議。

實作步驟

1. 使用您偏好的 RSS 閱讀器訂閱各種不同的部落格和公告，以及訂閱每日功能更新 SNS 主題。
2. 評估要參加哪些 AWS 活動，以獲得有關新功能和服務的第一手資訊。

3. 與您的 AWS 帳戶 團隊排定會議，以討論有關更新安全服務和功能的任何問題。
4. 考慮訂閱企業支援，即可定期向技術客戶經理 (TAM) 諮詢。

資源

相關的最佳實務：

- [PERF01-BP01 了解並熟悉可用的雲端服務和功能](#)
- [COST01-BP07 及時了解新的服務版本](#)

身分與存取管理

問題

- [SEC 2. 如何管理人員和機器的身分驗證？](#)
- [SEC 3. 如何管理人員和機器的許可？](#)

SEC 2. 如何管理人員和機器的身分驗證？

處理安全的 AWS 工作負載時，您需要管理兩種身分類型。

- 真人身分：需要存取您的 AWS 環境和應用程式的真人身分可分為三個群組：人力資源、第三方和使用者。

人力資源群組包括管理員、開發人員，以及身為您組織成員的操作人員。他們需要存取權來管理、建置和操作您的 AWS 資源。

第三方是外部協作者，例如承包商、廠商或合作夥伴。他們與您組織互動的過程中，會需要與 AWS 資源互動。

使用者是您應用程式的取用者。他們透過 Web 瀏覽器、用戶端應用程式、行動應用程式或互動式命令列工具存取您的 AWS 資源。

- 機器身分：您的工作負載應用程式、操作工具和元件需要身分才能向 AWS 服務發出請求，例如讀取資料。這些身分也包括您 AWS 環境 (如 Amazon EC2 執行個體或 AWS Lambda 函數) 內執行的機器。您也可以管理需要存取您 AWS 環境之外部對象的機器身分，或 AWS 外部的機器。

最佳實務

- [SEC02-BP01 使用強式登入機制](#)
- [SEC02-BP02 使用暫時憑證](#)
- [SEC02-BP03 安全地儲存和使用機密](#)
- [SEC02-BP04 仰賴集中式身分提供者](#)
- [SEC02-BP05 定期稽核和輪換憑證](#)
- [SEC02-BP06 採用使用者群組和屬性](#)

SEC02-BP01 使用強式登入機制

登入 (使用登入憑證進行驗證) 可預防當沒有使用多重要素驗證 (MFA) 等機制時的風險，尤其是在登入憑證遭到意外暴露或輕易被猜出的情況下。使用強式登入機制透過要求 MFA 和強式密碼政策來降低這些風險。

預期成果：透過對 [AWS Identity and Access Management \(IAM\)](#) 使用者、[AWS 帳戶 根使用者](#)、[AWS IAM Identity Center](#) 和協力廠商身分提供者使用強式登入機制，來降低 AWS 中意外存取憑證的風險。這意味著要求使用 MFA，強制強式密碼政策，以及偵測異常的登入行為。

常見的反模式：

- 沒有為您的身分強制強式密碼政策，包括複雜密碼和 MFA。
- 在不同使用者之間共用相同的憑證。
- 沒有針對可疑的登入使用偵測控制。

未建立此最佳實務時的曝險等級：高

實作指引

真人身分可採用數種方法來登入 AWS。AWS 最佳實務是依賴使用聯合 (在 AWS IAM 與集中式 IdP 之間的直接 SAML 2.0 聯合或使用 AWS IAM Identity Center) 的集中式身分提供者，向 AWS 進行身分驗證。在這種情況下，以您的身分提供者或 Microsoft Active Directory 建立安全的登入程序。

當您第一次開啟 AWS 帳戶 時，是從 AWS 帳戶 根使用者開始。您應僅使用帳戶根使用者來設定使用者的存取權 (以及[需要根使用者的任務](#))。務必在開啟 AWS 帳戶 之後，立即為帳戶根使用者開啟多重要素驗證 (MFA)，並使用 [AWS 最佳實務指南](#) 來保護根使用者。

AWS IAM Identity Center 是專為員工使用者所設計，您可以在服務內建立和管理使用者身分，並使用 MFA 保護登入程序。AWS 另一方面，Cognito 是專為客戶身分和存取管理 (CIAM) 所設計，它為應用程式中的外部使用者身分提供使用者集區和身分提供者。

如果您在 AWS IAM Identity Center 中建立使用者，請保護該服務中的登入程序，並[開啟 MFA](#)。對於您應用程式中的外部使用者身分，您可以使用 [Amazon Cognito 使用者集區](#)並保護該服務中的登入程序，或透過 Amazon Cognito 使用者集區支援的其中一個身分提供者。

此外，對於 AWS IAM Identity Center 中的使用者，您可以在將存取權授予 AWS 資源之前，使用 [AWS Verified Access](#) 來確認使用者的身分和裝置狀態，藉此提供一層額外的保護。

如果您使用 [AWS Identity and Access Management \(IAM\)](#) 使用者，請使用 IAM 保護登入程序。

您可以同時使用 AWS IAM Identity Center 和直接 IAM 聯合來管理 AWS 的存取權。您可以使用 IAM 聯合來管理 AWS 管理主控台 和服務的存取權，以及使用 IAM Identity Center 來管理 Quick 或 Amazon Q Business 等業務應用程式的存取權。

無論登入方法為何，強制執行強式登入政策至關重要。

實作步驟

以下是一般的強式登入建議。您設定的實際設定應由貴公司政策來規定，或使用如 [NIST 800-63](#) 的標準。

- Require MFA (需要 MFA)。 [IAM 最佳實務是對真人身分和工作負載要求 MFA](#)。開啟 MFA 可提供多一層安全防護，要求使用者提供登入憑證和一次性密碼 (OTP)，或是從硬體裝置以密碼編譯方式驗證和產生的字串。
- 強制密碼長度下限，此為密碼強度的要素。
- 強制密碼複雜性，使密碼更難猜測。
- 允許使用者變更自己的密碼。
- 建立個別身分，而不是共用憑證。透過建立個別身分，您可以為每個使用者提供一組獨一無二的安全憑證。個別使用者可讓您稽核每個使用者的活動。

IAM Identity Center 建議：

- 當使用預設目錄來建立密碼長度、複雜性和重複使用需求時，IAM Identity Center 提供預先定義的[密碼政策](#)。
- 當身分來源為預設目錄、AWS Managed Microsoft AD 或 AD Connector 時，[開啟 MFA](#) 並為 MFA 設定內容感知或永遠開啟設定。
- 允許使用者[註冊自己的 MFA 裝置](#)。

Amazon Cognito 使用者集區目錄建議：

- 設定[密碼強度](#)設定。
- 使用者[需要 MFA](#)。
- 針對[適應性身分驗證](#) (這可封鎖可疑登入) 等功能使用 Amazon Cognito 使用者集區[進階安全設定](#)。

IAM 使用者建議：

- 在理想情況下，您使用 IAM Identity Center 或直接聯合。然而，您可能需要 IAM 使用者。在這種情況下，為 IAM 使用者[設定密碼政策](#)。您可以使用密碼政策來定義需求，例如最短長度或是密碼是否需要非字母字元等。
- 建立 IAM 政策以[強制 MFA 登入](#)，以便允許使用者管理自己的密碼和 MFA 裝置。

資源

相關的最佳實務：

- [SEC02-BP03 安全地儲存和使用機密](#)
- [SEC02-BP04 仰賴集中式身分提供者](#)
- [SEC03-BP08 在組織內安全地共用資源](#)

相關文件：

- [AWS IAM Identity Center 密碼政策](#)
- [IAM 使用者密碼政策](#)
- [設定 AWS 帳戶 根使用者密碼](#)
- [Amazon Cognito 密碼政策](#)
- [AWS 憑證](#)
- [IAM 安全最佳實務](#)

相關影片：

- [使用 AWS IAM Identity Center 大規模管理使用者許可](#)
- [在每一層都能掌握身分](#)

SEC02-BP02 使用暫時憑證

當進行任何類型的驗證時，最好是使用臨時憑證，而不是長期憑證，以降低或消除風險，例如憑證遭到意外洩露、共用或遭竊。

預期成果：為了降低長期憑證的風險，對於人員和機器身分，請盡可能使用臨時憑證。長期憑證會產生許多風險，例如，經由上傳到公有儲存庫而暴露。透過使用臨時憑證，您可大幅降低憑證遭到入侵的可能性。

常見的反模式：

- 開發人員使用取自 IAM 使用者的長期存取金鑰，而不是使用聯合從 CLI 取得暫時憑證。
- 開發人員將長期存取金鑰內嵌在程式碼中，並將該程式碼上傳到公有 Git 儲存庫。
- 開發人員將長期存取金鑰內嵌在行動應用程式中，之後在應用程式商店中提供該行動應用程式。
- 使用者與其他使用者共用長期存取金鑰，或是擁有長期存取金鑰的離職員工仍持有金鑰。
- 對機器身分可以使用暫時憑證時，卻使用長期存取金鑰。

未建立此最佳實務時的曝險等級：高

實作指引

對所有 AWS API 和 CLI 請求使用暫時安全憑證，而不是長期憑證。幾乎在任何情況下，對 AWS 服務的 API 和 CLI 請求都必須使用 [AWS 存取金鑰](#) 簽署。您可以使用暫時或長期憑證簽署這些請求。您唯一應使用長期憑證 (又稱為長期存取金鑰) 的時候是在使用 [IAM 使用者](#) 或 [AWS 帳戶 根使用者](#) 時。當您聯合至 AWS 或透過其他方法擔任 [IAM 角色](#) 時，系統會產生暫時憑證。每當您使用登入憑證存取 AWS 管理主控台 時，系統會為您產生暫時憑證以呼叫 AWS 服務。在幾種情況下，您將需要長期憑證，並能夠使用暫時憑證完成幾乎所有任務。

避免使用長期憑證而改用暫時憑證，同時實行減少使用 IAM 使用者並支持聯合和 IAM 角色的策略。雖然對人類和機器身分過去以來都是使用 IAM 使用者，我們現在建議不要使用它們以避免使用長期存取金鑰的風險。

實作步驟

真人身分

對於人力資源身分，例如員工、管理員、開發人員、操作員和客戶：

- 您應 [仰賴集中式身分提供者](#) 並 [要求人類使用者以聯合搭配身分提供者](#)，使用臨時憑證存取 AWS。您可以 [直接聯合至各個 AWS 帳戶](#) 或使用 [AWS IAM Identity Center](#) 和您選擇的身分提供者，為您的使

用戶進行聯合。與使用 IAM 使用者相比，聯合除了可消除長期憑證外，還提供一些優勢。您的使用者也可以從命令列進行[直接聯合](#)，或使用 [IAM Identity Center](#) 請求暫時憑證。這表示有少數使用案例會需要 IAM 使用者，或使用者需要長期憑證。

對於第三方身分：

- 當授權讓第三方 (例如軟體即服務 (SaaS) 提供者) 存取 AWS 帳戶 中的資源時，您可以使用[跨帳戶角色](#)和[以資源為基礎的政策](#)。此外，您可以針對 B2B SaaS 客戶或合作夥伴，使用 [Amazon Cognito OAuth 2.0 授權](#)用戶端憑證流程。

透過 Web 瀏覽器、用戶端應用程式、行動應用程式或互動式命令列工具存取您的 AWS 資源的使用者身分。

- 如果需要授權應用程式供消費者或客戶存取您的 AWS 資源，您可以使用 [Amazon Cognito 身分集區](#)或 [Amazon Cognito 使用者集區](#)來提供臨時憑證。憑證的許可是透過 IAM 角色來設定。您也可以為未驗證的訪客使用者，另外定義有限許可的 IAM 角色。

機器身分

對於機器身分，您可能需要使用長期憑證。在這些情況下，您應[要求工作負載使用暫時憑證，並以 IAM 角色存取 AWS](#)。

- 對於 [Amazon Elastic Compute Cloud](#) (Amazon EC2)，您可以使用[適用於 Amazon EC2 的角色](#)。
- [AWS Lambda](#) 可讓您設定 [Lambda 執行角色，授予服務許可](#)以使用暫時憑證執行 AWS 動作。有許多其他類似的模型可供 AWS 服務使用 IAM 角色授予暫時憑證。
- 對於 IoT 裝置，您可以使用 [AWS IoT Core 憑證提供者](#)來請求暫時憑證。
- 對於內部部署系統或是在 AWS 之外執行並需要存取 AWS 資源的系統，您可以使用 [IAM Roles Anywhere](#)。

在某些情況下不支援臨時憑證，因此需要使用長期憑證。在這些情況下，[定期稽核和輪換這些憑證並定期輪換存取金鑰](#)。對於有嚴格限制的 IAM 使用者存取金鑰，請考慮下列其他安全措施：

- 授予有嚴格限制的許可：
 - 遵守最低權限原則 (具體指定動作、資源和條件)。
 - 考慮僅授予 IAM 使用者一個特定角色的 AssumeRole 操作。根據內部部署架構而定，此方法有助於隔離和保護長期 IAM 憑證。

- 在 IAM 角色信任政策中，限制允許的網路來源和 IP 位址。
- 監控用量並設定未使用許可或濫用的警示 (使用 AWS CloudWatch Logs 指標篩選器和警報)。
- 強制執行[許可界限](#) (服務控制政策 (SCP) 和許可界限來彼此互補 - SCP 較為粗略，而許可界限較為精細)。
- 實作程序來佈建並安全地儲存 (在內部部署保存庫中) 憑證。

其他適用於需要長期憑證的案例的選項包括：

- 建置自己的權杖供應 API (使用 Amazon API Gateway)。
- 在您必須使用長期憑證，或是使用 AWS 存取金鑰以外的憑證 (例如資料庫登入) 的情況下，您可以使用專為管理機密而設計的服務，例如 [AWS Secrets Manager](#)。Secrets Manager 可簡化加密、輪換及安全儲存加密機密的工作。許多 AWS 服務都可使用 Secrets Manager 支援[直接整合](#)。
- 對於多重雲端整合，您可以根據來源憑證服務提供者 (CSP) 憑證使用聯合身分 (請參閱 [AWS STS AssumeRoleWithWebIdentity](#))。

如需有關輪換長期憑證的詳細資訊，請參閱[輪換存取金鑰](#)。

資源

相關的最佳實務：

- [SEC02-BP03 安全地儲存和使用機密](#)
- [SEC02-BP04 仰賴集中式身分提供者](#)
- [SEC03-BP08 在組織內安全地共用資源](#)

相關文件：

- [暫時安全憑證](#)
- [AWS 憑證](#)
- [IAM 安全最佳實務](#)
- [IAM 角色](#)
- [IAM Identity Center](#)
- [身分提供者與聯合](#)
- [輪換存取金鑰](#)
- [安全合作夥伴解決方案：存取與存取控制](#)

- [AWS 帳戶根使用者](#)
- [使用 Google Cloud Platform 原生工作負載身分存取 AWS](#)
- [如何使用 AWS Security Token Service 存取 Microsoft Entra ID 租用戶中的 AWS 資源](#)

相關影片：

- [使用 AWS IAM Identity Center 大規模管理使用者許可](#)
- [在每一層都能掌握身分](#)

SEC02-BP03 安全地儲存和使用機密

工作負載需要能夠自動向資料庫、資源和第三方資源證明其身分。這需使用私密存取憑證來完成，例如 API 存取金鑰、密碼和 OAuth 字符。使用專用服務來儲存、管理和輪換這些憑證有助於降低這些憑證遭到入侵的可能性。

預期成果：實作安全管理應用程式憑證的機制，以達成下列目標：

- 識別工作負載需要何種機密。
- 盡可能以短期憑證取代長期憑證，來減少所需的長期憑證數目。
- 建立安全的存放區並自動輪換其餘的長期憑證。
- 稽核對存在於工作負載中的機密的存取。
- 持續監控以確認原始程式碼在開發過程中沒有內嵌機密。
- 降低憑證遭意外洩露的可能性。

常見的反模式：

- 沒有輪換憑證。
- 將長期憑證存放在原始程式碼或設定檔中。
- 未加密儲存靜態憑證。

建立此最佳實務的優勢：

- 已加密儲存靜態和傳輸中的機密。
- 透過 API 限制憑證的存取 (將其視為憑證自動販賣機)。
- 稽核並記錄對憑證的存取 (包括讀寫)。

- 區隔顧慮：由不同的元件執行憑證輪換，而該元件可與其餘的架構分離。
- 自動將機密隨需散發到軟體元件並集中進行輪換。
- 可以精細的方式控制對憑證的存取。

未建立此最佳實務時的風險暴露等級：高

實作指引

以往，憑證用於向資料庫進行驗證，而第三方 API、字符和其他機密可能內嵌在原始程式碼或環境檔案中。AWS 提供數種機制以安全儲存這些憑證，自動輪換並稽核它們的使用情況。

著手機密管理的最佳方法是遵循移除、取代和輪換的指引。最安全的憑證是您不用儲存、管理或處理的憑證。有些憑證對於工作負載的運作不再是必要的，故能夠安全移除。

對於工作負載適當運作仍舊是必要的憑證，可能有機會以暫時或短期憑證取代長期憑證。例如，與其對 AWS 私密存取金鑰進行硬式編碼，考慮使用 IAM 角色以臨時憑證取代長期憑證。

部分長期存留的機密可能無法移除或取代。您可以將這些機密儲存在 [AWS Secrets Manager](#) 之類的服務中，進行集中儲存、管理和定期輪換。

對工作負載的原始程式碼和設定檔的稽核，可能顯現多種類型的憑證。下表概述處理常見憑證類型的策略：

憑證類型	描述	建議策略
IAM 存取金鑰	用於在工作負載內擔任 IAM 角色的 AWS IAM 存取和私有金鑰	取代：改用指派給運算執行個體的 IAM 角色 (例如 Amazon EC2 或 AWS Lambda)。為了與需要存取您的 AWS 帳戶中的資源的第三方進行互通，請詢問他們是否支援 AWS 跨帳戶存取權 。對於行動應用程式，請考慮透過 Amazon Cognito 身分集區 (聯合身分) 使用臨時憑證。對於 AWS 外部執行的工作負載，請考慮 IAM Roles Anywhere 或 AWS Systems Manager 混合啟用 。關於容

憑證類型	描述	建議策略
		器，請參閱 Amazon ECS 任務 IAM 角色 或 Amazon EKS 節點 IAM 角色 。
SSH 金鑰	用於手動或做為自動化程序的一部分登入 Linux EC2 執行個體的 Secure Shell 私有金鑰	取代：使用 AWS Systems Manager 或 EC2 Instance Connect 透過 IAM 角色提供對 EC2 執行個體的程式設計和手動存取。
應用程式和資料庫憑證	密碼 – 純文字字串	輪換：將憑證儲存在 AWS Secrets Manager 中並建立自動輪換 (如果可能)。
Amazon RDS 和 Aurora Admin 資料庫憑證	密碼 – 純文字字串	取代：使用 Secrets Manager 與 Amazon RDS 的整合 或 Amazon Aurora 。此外，某些 RDS 資料庫類型可以在某些使用案例中使用 IAM 角色而非密碼 (如需詳細資訊，請參閱 IAM 資料庫身分驗證)。
OAuth 字符	私密字符 – 純文字字串	輪換：將字符儲存在 AWS Secrets Manager 中並設定自動輪換。
API 字符和金鑰	私密字符 – 純文字字串	輪換：儲存 AWS Secrets Manager 中，並在可能的情況下建立自動輪換。

常見的反模式是將 IAM 存取金鑰內嵌在原始程式碼、組態檔案或行動應用程式內。當需要 IAM 存取金鑰與 AWS 服務通訊時，請使用 [暫時 \(短期\) 安全憑證](#)。您可以透過 [適用於 EC2 的 IAM 角色](#) 執行個體、[執行角色](#) (用於 Lambda 函數)、[Cognito IAM 角色](#) (用於行動使用者存取)，以及 [IoT Core 政策](#) (用於 IoT 裝置) 提供這些短期憑證。與第三方互動時，偏好 [委派 IAM 角色的存取權](#) (包含對帳戶資源的必要存取權)，而不是設定 IAM 使用者並將該使用者的私密存取金鑰傳送給第三方。

在很多情況下，工作負載需要儲存機密才能與其他服務和資源相互操作。[AWS Secrets Manager](#) 是專為安全管理這些憑證所打造的，可儲存、使用和輪換 API 字符、密碼和其他憑證。

AWS Secrets Manager 提供五項重要功能以確保敏感憑證的安全儲存和處理：[靜態加密](#)、[傳輸中加密](#)、[全面性稽核](#)、[精細存取控制](#)，以及[可擴充的憑證輪換](#)。來自 AWS 合作夥伴的其他機密管理服務，或本機開發並提供類似功能和保證的解決方案也可接受。

擷取機密時，您可以使用 Secrets Manager 用戶端快取元件進行快取以供未來使用。擷取快取的秘密比從 Secrets Manager 中擷取要快。此外，由於呼叫 Secrets Manager API 會產生費用，因此使用快取可以降低成本。如需您可以擷取機密的所有方法，請參閱[取得機密](#)。

Note

某些語言可能會要求您針對用戶端快取實作自己的記憶體內加密。

實作步驟

1. 使用 [Amazon CodeGuru](#) 等自動工具識別包含硬式編碼憑證的程式碼路徑。
 - a. 使用 Amazon CodeGuru 掃描您的程式碼儲存庫。檢閱完成後，在 CodeGuru 中篩選 Type=Secrets 以尋找有問題的程式碼行。
2. 識別可移除或取代的憑證。
 - a. 識別不再需要的憑證並標示以進行移除。
 - b. 對於內嵌在原始程式碼中的 AWS 機密金鑰，請將其取代為與必要資源關聯的 IAM 角色。如果您部分的工作負載位於 AWS 之外，但需要 IAM 憑證來存取 AWS 資源，請考慮 [IAM Roles Anywhere](#) 或 [AWS Systems Manager 混合啟用](#)。
3. 對於其他第三方長期存留且需要使用輪換策略的機密，將 Secrets Manager 整合至程式碼中以在執行時期擷取第三方機密。
 - a. CodeGuru 主控台可以使用已探索的憑證自動[在 Secrets Manager 中建立機密](#)。
 - b. 將 Secrets Manager 的機密擷取整合至您的應用程式程式碼中。
 - i. 無伺服器 Lambda 函數可以使用與語言無關的 [Lambda 延伸](#)。
 - ii. 對於 EC2 執行個體或容器，AWS 提供範例[用戶端程式碼，可以數種熱門的程式設計語言從 Secrets Manager 擷取機密](#)。
4. 定期審查您的程式碼庫並重新掃描，以確認程式碼中未加入新的機密。
 - a. 考慮使用 [git-secrets](#) 之類的工具以防將新機密認可到您的原始程式碼儲存庫。
5. [監控 Secrets Manager 活動](#) 以尋找非預期使用、不當私密存取或嘗試刪除機密的跡象。

6. 減少對憑證的人員接觸。將讀寫和修改憑證的存取權限於專門用於此用途的 IAM 角色，並且只將擔任該角色的存取權提供給一小組可操作的使用者子集。

資源

相關的最佳實務：

- [SEC02-BP02 使用臨時憑證](#)
- [SEC02-BP05 定期稽核和輪換憑證](#)

相關文件：

- [開始使用 AWS Secrets Manager](#)
- [身分提供者與聯合](#)
- [Amazon CodeGuru 推出機密偵測器](#)
- [AWS Secrets Manager 使用 AWS Key Management Service 的方式](#)
- [Secrets Manager 中的機密加密和解密](#)
- [Secrets Manager 部落格文章](#)
- [Amazon RDS 宣佈與 AWS Secrets Manager 整合](#)

相關影片：

- [大規模管理、擷取和輪換機密的最佳實務](#)
- [使用 Amazon CodeGuru 機密偵測器尋找硬式編碼的機密](#)
- [使用 AWS Secrets Manager 保護混合式工作負載的機密](#)

相關研討會：

- [在 AWS Secrets Manager 中儲存、擷取和管理敏感憑證](#)
- [AWS Systems Manager 混合啟用](#)

SEC02-BP04 仰賴集中式身分提供者

人力身分 (員工和承包商) 可仰賴身分供應商來集中管理身分。由於您是從單一位置建立、指派、管理、撤銷和稽核存取權，因此這樣一來可以更好管理多個應用程式和系統中的存取權。

預期成果：擁有集中式身分提供者，可集中管理員工使用者、身分驗證政策 (例如，要求多重要素驗證 (MFA))，以及對系統和應用程式進行授權 (例如，根據使用者的群組成員資格或屬性指派存取權)。您的員工使用者登入集中身分提供者並聯合 (單一登入) 至內部和外部應用程式，如此一來，使用者就不需記住多個憑證。您的身分提供者與您的人力資源 (HR) 系統整合，因此人事變更會自動同步至您的身分提供者。例如，若有人離開您的組織，您可以自動撤銷聯合應用程式和系統 (包括 AWS) 的存取權。您已在身分提供者中啟用詳細稽核日誌記錄，並監控這些日誌以找出不尋常的使用者行為。

常見的反模式：

- 您未使用聯合和單一登入。您的員工使用者在多個應用程式和系統中建立了不同的使用者帳戶和憑證。
- 您尚未將員工使用者的身分生命週期自動化，例如透過整合身分提供者與您的 HR 系統。使用者離開您的組織或變更職務時，您採取手動程序在多個應用程式和系統中刪除或更新記錄。

建立此最佳實務的優勢：透過使用集中式身分提供者，您就可以從單一位置管理員工使用者身分和政策，而且能夠將應用程式存取權指派給使用者和群組，並監控使用者登入活動。透過與您的人力資源 (HR) 系統整合，使用者變更職務時，這些變更就會同步至身分提供者，並自動更新指派的應用程式和許可。使用者離開您的組織時，系統會自動停用他們在身分提供者中的身分，並撤銷他們對聯合應用程式和系統的存取權。

未建立此最佳實務時的風險暴露等級：高

實作指引

員工使用者存取 AWS 的指引：員工使用者 (例如組織中的員工和承包商) 可能需要使用 AWS 管理主控台 或 AWS Command Line Interface (AWS CLI) 存取 AWS 來執行其工作職能。您可以透過從集中式身分提供者，在兩個層級聯合至 AWS 的方式，將 AWS 存取權授予員工使用者：直接聯合至各個 AWS 帳戶，或聯合至您的 [AWS 組織](#) 中的多個帳戶。

若要將您的員工使用者直接與各個 AWS 帳戶 聯合，您可以使用集中式身分提供者來聯合至該帳戶中的 [AWS Identity and Access Management](#)。IAM 的彈性可讓您啟用單獨的 [SAML 2.0](#) 或 [Open ID Connect \(OIDC\)](#) 身分提供者用於各個 AWS 帳戶，並使用聯合身分使用者屬性進行存取控制。您的員工使用者將透過提供憑證 (例如密碼和 MFA 字符代碼) 的方式，使用自己的 Web 瀏覽器登入身分提供者。身分提供者會向其瀏覽器發出 SAML 聲明，並提交至 AWS 管理主控台 登入 URL，以允許使用者透過 [擔任 IAM 角色對 AWS 管理主控台](#) 進行單一登入。您的使用者也可以取得暫時 AWS API 憑證，以便在 [AWS CLI](#) 或 [AWS SDK](#) (從 [AWS STS](#)) 中使用，方法是 [使用來自身分提供者的 SAML 聲明擔任 IAM 角色](#)。

若要將您的員工使用者與 AWS 組織中的多個帳戶聯合，您可以使用 [AWS IAM Identity Center](#) 集中管理員工使用者對 AWS 帳戶 和應用程式的存取權。您可以為組織啟用 Identity Center，並設定您的身分來源。IAM Identity Center 提供了預設身分來源目錄，您可以使用此目錄來管理您的使用者和群組。或者，您可以選擇外部身分來源，方法是使用 SAML 2.0 [連線至您的外部身分提供者](#)，並使用 SCIM [自動佈建](#)使用者和群組，或是使用 [Directory Service](#) [連線至您的 Microsoft AD 目錄](#)。身分來源設定後，您可以透過在您的[許可集](#)中定義最低許可政策的方式，指派使用者和群組對 AWS 帳戶 的存取權。您的員工使用者可以進行身分驗證的方式包括：透過您的集中身分提供者登入 [AWS 存取入口網站](#)以及對 AWS 帳戶 和指派給他們的雲端應用程式進行單一登入。您的使用者可以設定 [AWS CLI v2](#) 以透過 Identity Center 進行身分驗證，並取得憑證來執行 AWS CLI 命令。Identity Center 也允許透過單一登入方式存取 AWS 應用程式，例如 [Amazon SageMaker AI Studio](#) 和 [AWS IoT IoT Sitewise Monitor 入口網站](#)。

依照上述指引進行後，您的員工使用者在 AWS 上管理工作負載時，將不再需要使用 IAM 使用者 和群組，可直接正常操作。反之，您的使用者和群組將在 AWS 外部進行管理，而且使用者能夠以聯合身分存取 AWS 資源。聯合身分會使用您的集中式身分提供者所定義的群組。您應該識別並移除您的 AWS 帳戶 中不再需要的 IAM 群組、IAM 使用者和長期存在的使用者憑證 (密碼和存取金鑰)。您可以使用 [IAM 憑證報告找到未使用的憑證](#)，[刪除相應的 IAM 使用者](#)和[刪除 IAM 群組](#)。您可以對組織套用[服務控制政策 \(SCP\)](#) 以協助防止建立新的 IAM 使用者和群組，並強制透過聯合身分存取 AWS。

Note

您負責處理 SCIM 存取權杖的輪換，如[自動佈建](#)文件中所述。此外，您也負責輪換支援您的聯合身分的憑證。

應用程式使用者的指引：您可以使用 [Amazon Cognito](#) 做為您的集中式身分提供者，來管理應用程式 (例如行動應用程式) 使用者的身分。Amazon Cognito 可為您的 Web 和行動應用程式啟用身分驗證、授權和使用者管理功能。Amazon Cognito 提供了可擴展至數百萬使用者的身分存放區、可支援社交與企業聯合身分，並且提供進階安全功能來協助保護您的使用者和業務。您可以將自訂 Web 或行動應用程式與 Amazon Cognito 整合，在幾分鐘內就能在應用程式中新增使用者身分驗證和存取控制。Amazon Cognito 是以 SAML 和 Open ID Connect (OIDC) 等開放身分標準為基礎所建置，可支援各種不同的合規法規，並與前端和後端開發資源整合。

實作步驟

員工使用者存取 AWS 的步驟

- 使用下列其中一種方法，透過集中式身分提供者將您的員工使用者聯合至 AWS：

- 使用 IAM Identity Center 透過與您的身分提供者聯合，對您的 AWS 組織中的多個 AWS 帳戶 啟用單一登入。
- 使用 IAM 將您的身分提供者直接連接到各個 AWS 帳戶，以實現聯合的精細存取。
- 識別並移除已由聯合身分取代的 IAM 使用者和群組。

應用程式使用者的步驟

- 使用 Amazon Cognito 作為應用程式的集中式身分提供者。
- 使用 OpenID Connect 和 OAuth 將您的自訂應用程式與 Amazon Cognito 整合。您可以使用 Amplify 程式庫來開發自訂應用程式，這些程式庫提供了簡單的介面，可與各種不同的 AWS 服務進行整合，例如用於身分驗證的 Amazon Cognito。

資源

相關的最佳實務：

- [SEC02-BP06 採用使用者群組和屬性](#)
- [SEC03-BP02 授予最低權限存取權](#)
- [SEC03-BP06 根據生命週期管理存取](#)

相關文件：

- [AWS 中的聯合身分](#)
- [IAM 中的安全最佳實務](#)
- [AWS Identity and Access Management 最佳實務](#)
- [IAM Identity Center 委派管理入門](#)
- [如何在 IAM Identity Center 中針對進階使用案例使用客戶受管政策](#)
- [AWS CLI v2：IAM Identity Center 憑證提供者](#)

相關影片：

- [AWS re:Inforce 2022 - AWS Identity and Access Management \(IAM\) 深入剖析](#)
- [AWS re:Invent 2022 - 使用 IAM Identity Center 簡化現有的員工存取權](#)
- [AWS re:Invent 2018：在每一層都能掌握身分](#)

相關範例：

- [研討會：使用 AWS IAM Identity Center 實現強大的身分管理](#)

相關工具：

- [AWS 安全能力合作夥伴：身分和存取管理](#)
- [saml2aws](#)

SEC02-BP05 定期稽核和輪換憑證

定期稽核和輪換憑證以限制憑證可用來存取資源的時限。長期憑證會產生許多風險，而透過定期輪換長期憑證可以降低這些風險。

預期成果：實作憑證輪換以協助降低與使用長期憑證關聯的風險。定期稽核和修復不符合憑證輪換政策的情況。

常見的反模式：

- 沒有稽核憑證的使用。
- 不必要地使用長期憑證。
- 使用長期憑證並且未定期輪換。

未建立此最佳實務時的曝險等級：高

實作指引

當您無法倚賴臨時憑證且需要長期憑證時，請稽核憑證以確保強制執行既定的控制 (例如[多重要素驗證 \(MFA\)](#))、定期輪換，並且具備適當的存取層級。

定期驗證 (最好是透過自動化工具) 是確認強制執行正確的控制項的必要項目。對於人類身分，您應要求使用者定期變更密碼，並淘汰存取金鑰而改用暫時憑證。隨著您從 AWS Identity and Access Management (IAM) 使用者移向集中式身分，您可以[產生憑證報告](#)以稽核您的使用者。

我們也建議您在身分提供者中強制和監控 MFA。您可以設定[AWS Config 規則](#)或使用 [AWS Security Hub CSPM 安全標準](#)來監控使用者是否已啟用 MFA。請考慮使用 [IAM Roles Anywhere](#) 為機器身分提供臨時憑證。在無法使用 IAM 角色和暫時憑證的情況下，必須經常稽核和輪換存取金鑰。

實作步驟

- 定期稽核憑證：稽核身分提供者和 IAM 中設定的身分有助於確認只有已授權的身分才能存取您的工作負載。此類身分可能包括但不限於 IAM 使用者、AWS IAM Identity Center 使用者、Active Directory 使用者，或不同上游身分提供者中的使用者。例如，移除離職的人員和移除不再需要的跨帳戶角色。對於 IAM 實體存取的服務，備妥程序來定期稽核許可。這有助您識別需要修改的政策，以移除任何不使用的許可。使用憑證報告和 [AWS Identity and Access Management Access Analyzer](#) 來稽核 IAM 憑證和許可。您可以使用 [Amazon CloudWatch](#) 來設定對特定 API 呼叫 (在 AWS 環境中呼叫) 的警示。[Amazon GuardDuty](#) 也可以向您提醒未預期的活動，這可指出對 IAM 憑證過於寬鬆的存取或意外存取。
- 定期輪換憑證：當您無法使用暫時憑證時，定期輪換長期 IAM 存取金鑰 (最長每 90 天)。如果在您不知情的情況下意外洩漏了存取金鑰，這可限制憑證可用來存取資源的時限。如需有關輪換 IAM 使用者的存取金鑰的詳細資訊，請參閱[輪換存取金鑰](#)。
- 審查 IAM 許可：為提升 AWS 帳戶的安全，請定期審查和監控每個 IAM 政策。確認政策遵守最低權限的原則。
- 考慮自動化 IAM 資源建立和更新：[IAM Identity Center](#) 會自動執行許多 IAM 任務，例如角色和政策管理。或者，可以使用 AWS CloudFormation 自動化 IAM 資源 (包括角色和政策) 的部署，以減少人為錯誤，因為可針對範本進行驗證和版本控制。
- 對於機器身分，使用 IAM Roles Anywhere 取代 IAM 使用者：[IAM Roles Anywhere](#) 可讓您在傳統上無法使用的區域 (例如內部部署伺服器) 使用角色。IAM Roles Anywhere 使用可信的 [X.509 憑證](#) 向 AWS 進行身分驗證及接收臨時憑證。使用 IAM Roles Anywhere 讓您無需輪換這些憑證，因為長期憑證不再儲存於內部部署環境中。請注意，您將需要監視 X.509 憑證，並在快到期時輪換。

資源

相關的最佳實務：

- [SEC02-BP02 使用臨時憑證](#)
- [SEC02-BP03 安全地儲存和使用機密](#)

相關文件：

- [開始使用 AWS Secrets Manager](#)
- [IAM 最佳實務](#)
- [身分提供者與聯合](#)
- [安全合作夥伴解決方案：存取與存取控制](#)

- [暫時安全憑證](#)
- [取得 AWS 帳戶 的憑證報告](#)

相關影片：

- [大規模管理、擷取和輪換機密的最佳實務](#)
- [使用 AWS IAM Identity Center 大規模管理使用者許可](#)
- [在每一層都能掌握身分](#)

SEC02-BP06 採用使用者群組和屬性

根據使用者群組和屬性定義許可，有助於減少政策的數量並降低複雜性，因而更容易實現最低權限原則。您可以利用使用者群組，根據人員在組織中的職務，從單一位置管理多人的許可。部門、專案或位置這類屬性可在人員職務相似的情況下，針對不同的資源子集提供另一層許可範圍。

預期成果：您可以根據職務，針對執行該職務的所有使用者套用許可變更。群組成員資格和屬性會控管使用者許可，進而減少管理個別使用者層級許可的需求。您在身分提供者 (IdP) 中定義的群組和屬性會自動傳播到您的 AWS 環境。

常見的反模式：

- 管理個別使用者的許可，並且針對許多使用者重複此操作。
- 為群組定義的層級過高，授予的許可範圍過廣。
- 為群組定義的層級過於精細，致使成員資格發生重複和混淆的情形。
- 在可以使用屬性替代的情況下，仍使用在資源子集中具有重複許可的群組。
- 未透過整合在您 AWS 環境中的標準化身分提供者管理群組、屬性和成員資格。
- 使用 AWS IAM Identity Center 工作階段時使用角色鏈結

未建立此最佳實務時的曝險等級：中

實作指引

AWS 許可是在與主體 (例如使用者、群組、角色或資源) 相關聯且稱為政策的文件中定義。您可以根據工作職能、工作負載及 SDLC 環境來組織許可指派 (群組、許可、帳戶)，藉此調整許可管理的規模。針對您的員工，這可讓您根據使用者為組織執行的職務，而不是所存取的資源來定義群組。

例如，WebAppDeveloper 群組可能已附加政策，用於設定開發帳戶內的 Amazon CloudFront 等服務。AutomationDeveloper 群組可能有部分許可與 WebAppDeveloper 群組重疊。您可在個別政策中擷取這些常用的許可，並讓這些許可同時與這兩個群組相關聯，而不要讓這兩種職務的使用者屬於 CloudFrontAccess 群組。

除了群組之外，您還可以使用屬性來進一步設定存取範圍。例如，您的 WebAppDeveloper 群組中的使用者可能有「專案」屬性，可用來將其專案特定的資源設定至存取範圍內。如果使用此技術，則不需要在所擁有許可相同的情況下，針對處理不同專案的應用程式開發人員建立不同的群組。您在許可政策中參照屬性的方式是根據其來源，無論其定義為聯合通訊協定的一部分 (例如 SAML、OIDC 或 SCIM)、為自訂 SAML 聲明，還是在 IAM Identity Center 內設定。

實作步驟

1. 確定您將定義群組和屬性的位置：

- a. 依照 [SEC02-BP04 仰賴集中式身分提供者](#) 中的指引，您可以確定是否需要在身分提供者內、IAM Identity Center 內或在特定帳戶中使用 IAM 使用者群組定義群組和屬性。

2. 定義群組：

- a. 根據職務和所需的存取範圍確定您的群組。考慮使用階層式結構或命名慣例來有效組織群組。
- b. 如果在 IAM Identity Center 內定義，請建立群組，並使用許可集關聯所需的存取層級。
- c. 如果在外部身分提供者內定義，請確定提供者是否支援 SCIM 通訊協定，並考慮在 IAM Identity Center 內啟用自動佈建。此功能可在您的提供者與 IAM Identity Center 之間同步群組的建立、成員資格和刪除。

3. 定義屬性：

- a. 如果您使用外部身分提供者，SCIM 和 SAML 2.0 通訊協定預設都會提供特定屬性。您可以使用 SAML 聲明與 `https://aws.amazon.com/SAML/Attributes/PrincipalTag` 屬性名稱來定義和傳遞其他屬性。如需定義和設定自訂屬性的指引，請參閱身分提供者的文件。
- b. 如果您在 IAM Identity Center 內定義角色，請啟用屬性型存取控制 (ABAC) 功能並視需要定義屬性。考慮符合組織結構或資源標記策略的屬性。

如果您需要透過 IAM Identity Center 擔任的 IAM 角色進行 IAM 角色鏈結，則 `source-identity` 和 `principal-tags` 這類值就不會傳播。如需詳細資訊，請參閱 [啟用和設定存取控制的屬性](#)。

1. 根據群組和屬性設定許可的範圍：

- a. 考慮在許可政策中加入條件，用來比較您主體的屬性與所存取資源的屬性。例如，您可以定義一項條件，規定僅在 `PrincipalTag` 條件索引鍵的值與相同名稱的 `ResourceTag` 索引鍵的值相符時，才允許對相關資源的存取。

- b. 定義 ABAC 政策時，請遵循 [ABAC 授權](#) 最佳實務和範例中的指引。
- c. 隨著組織的需求發展，定期審查並更新群組和屬性結構，以確保最佳的許可管理。

資源

相關的最佳實務：

- [SEC02-BP04 仰賴集中式身分提供者](#)
- [SEC03-BP02 授予最低權限存取權](#)
- [COST02-BP04 實作群組和角色](#)

相關文件：

- [IAM 最佳實務](#)
- [在 IAM Identity Center 中管理身分](#)
- [AWS 的 ABAC 是什麼？](#)
- [IAM Identity Center 中的 ABAC](#)
- [ABAC 政策範例](#)

相關影片：

- [使用 AWS IAM Identity Center 大規模管理使用者許可](#)
- [在每一層都能掌握身分](#)

SEC 3. 如何管理人員和機器的許可？

管理許可，以控制對需要存取 AWS 和工作負載的人員和機器身分的存取。您可利用許可來控制誰可以在何種條件下存取哪些內容。將許可設定為特定的真人和機器身分，即可授予對特定資源上特定服務動作的存取權。此外，您可以指定必須為 true 才能授予存取的條件。

最佳實務

- [SEC03-BP01 定義存取需求](#)
- [SEC03-BP02 授予最低權限存取權](#)
- [SEC03-BP03 建立緊急存取程序](#)

- [SEC03-BP04 持續減少許可](#)
- [SEC03-BP05 為您的組織定義許可防護機制](#)
- [SEC03-BP06 根據生命週期管理存取](#)
- [SEC03-BP07 分析公有和跨帳戶存取權](#)
- [SEC03-BP08 在組織內安全地共用資源](#)
- [SEC03-BP09 安全地與第三方共用資源](#)

SEC03-BP01 定義存取需求

管理員、終端使用者或其他元件都需要存取工作負載的每個元件或資源。請明確定義應該有權存取每個元件的人員和機器，選擇適當的身分類型及驗證和授權方法。

常見的反模式：

- 將機密硬式編碼或儲存在應用程式中。
- 為每名使用者授予自訂許可。
- 使用長期憑證。

未建立此最佳實務時的曝險等級：高

實作指引

管理員、終端使用者或其他元件都需要存取工作負載的每個元件或資源。請明確定義應該有權存取每個元件的人員和機器，選擇適當的身分類型及驗證和授權方法。

應提供組織內對 AWS 帳戶 的定期存取 (使用[聯合存取](#)或集中式身分提供者)。您應集中進行身管理，並確保有既定的實務，可將 AWS 存取整合至員工的存取生命週期。例如，當員工改為擔任具有不同存取層級的任務角色時，其群組成員資格也應變更，以反映新的存取需求。

為非人類身分定義存取需求時，請判斷哪些應用程式和組成部分需要存取權，以及如何授予許可。使用透過最低權限存取模型建置的 IAM 角色是建議的方法。[AWS 受管政策](#)提供預先定義的 IAM 政策，其中涵蓋最常見的使用案例。

AWS 服務 (例如 [AWS Secrets Manager](#) 和 [AWS Systems Manager Parameter Store](#)) 可協助將機密從應用程式或工作負載中安全地分離。在 Secrets Manager 中，您可以為憑證建立自動輪換。您可以使用 Systems Manager 來參考指令碼、命令、SSM 文件、組態和自動化工作流程中的參數，方法是使用您在建立參數時指定的唯一名稱。

您可以使用 [AWS IAM Roles Anywhere](#) 取得 [IAM 中的臨時安全憑證](#)，以用於在 AWS 外部執行的工作負載。工作負載可以使用您在 AWS 應用程式中使用的相同 [IAM 政策](#) 和 [IAM 角色](#) 來存取 AWS 資源。

可能的話，請選擇短期暫時憑證，而不是長期靜態憑證。對於您希望使用者具備程式設計存取權和長期憑證的情況，請使用 [存取金鑰上次使用的資訊](#) 來輪換和移除存取金鑰。

若使用者想要與 AWS 管理主控台 之外的 AWS 互動，則需要程式化存取權。授與程式化存取權的方式取決於存取 AWS 的使用者類型。

若要授予使用者程式設計存取權，請選擇下列其中一個選項。

哪個使用者需要程式設計存取權？	到	根據
IAM	(建議事項) 將主控台憑證用作臨時憑證來簽署對 AWS CLI、AWS SDK 或 AWS API 的程式設計請求。	請依照您要使用的介面所提供的指示操作。 - 關於 AWS CLI，請參閱《AWS Command Line Interface 使用者指南》中的 登入以進行 AWS 本機開發。 - 如需 AWS SDK，請參閱《AWS SDK 和工具參考指南》中的 登入以進行 AWS 本機開發。
人力資源身分 (IAM Identity Center 中管理的使用者)	使用臨時憑證簽署對 AWS CLI、AWS SDK 或 AWS API 的程式設計請求。	請依照您要使用的介面所提供的指示操作。 - 關於 AWS CLI，請參閱《AWS Command Line Interface 使用者指南》中的 設定 AWS CLI 來使用 AWS IAM Identity Center。 - 關於 AWS SDK、工具和 AWS API，請參閱《AWSSDK 和工具參考

哪個使用者需要程式設計存取權？	到	根據
		指南》中的 IAM Identity Center 驗證 。
IAM	使用臨時憑證簽署對 AWS CLI、AWS SDK 或 AWS API 的程式設計請求。	請遵循《IAM 使用者指南》中 使用臨時憑證搭配 AWS 資源 中的指示。
IAM	(不建議使用) 使用長期憑證簽署 AWS CLI、AWS SDK 或 AWS API 的程式設計要求。	請依照您要使用的介面所提供的指示操作。 - 關於 AWS CLI，請參閱《AWS Command Line Interface 使用者指南》中的 使用 IAM 使用者憑證進行驗證。 - 關於 AWS SDK 和工具，請參閱《AWSSDK 和工具參考指南》中的 使用長期憑證進行驗證。 - 關於 AWS API，請參閱《IAM 使用者指南》中的 管理 IAM 使用者的存取金鑰。

資源

相關文件：

- [屬性型存取控制 \(ABAC\)](#)
- [AWS IAM Identity Center](#)
- [IAM Roles Anywhere](#)
- [IAM Identity Center 的 AWS 受管政策](#)
- [AWS IAM 政策條件](#)
- [IAM 使用案例](#)

- [移除不必要的憑證](#)
- [使用 政策](#)
- [如何根據 AWS 帳戶、OU 或組織控制對 AWS 資源的存取權](#)
- [使用 AWS Secrets Manager 中增強的搜尋功能來輕鬆識別、安排和管理機密](#)

相關影片：

- [在 60 分鐘內精通 IAM 政策](#)
- [責任區隔、最低權限、委派和 CI/CD](#)
- [簡化身分和存取管理，以促進創新](#)

SEC03-BP02 授予最低權限存取權

僅授予使用者在特定情況下對特定資源執行特定動作所需的存取權。使用群組和身分屬性大規模動態設定許可，而不是定義個別使用者的許可。例如，您可以允許一組開發人員的存取權，以只管理其專案的資源。如此一來，當開發人員退出專案時，其存取權將自動被撤銷，而無需變更基礎存取政策。

預期成果：使用者只擁有其特定工作職能所需的最低許可。您可以分開 AWS 帳戶 來隔離開發人員與正式作業環境。當開發人員需要存取特定任務的正式作業環境時，他們會獲得授予有限且受控的存取權，且僅限於任務期間有效。他們的正式作業存取權會在完成必要的工作後立即撤銷。您定期審查許可，並在不再需要時立即撤銷許可，例如，使用者變更角色或離開組織時。您可以將管理員權限侷限在一小群可信任的群組，以減少暴險。您僅給予機器或系統帳戶執行其預期任務所需的最低許可。

常見的反模式：

- 根據預設，您會對使用者授予管理員許可。
- 您使用根使用者帳戶從事日常活動。
- 您建立過度寬鬆的政策，而未適當界定範圍。
- 您不常檢閱許可，這會導致許可滲透。
- 您完全依賴屬性型存取控制來隔離環境或管理許可。

未建立此最佳實務時的曝險等級：高

實作指引

[最低權限](#)原則指出，僅應允許身分執行完成特定任務所需的最小動作集。這平衡了可用性、效率 and 安全性。根據此原則運作有助於限制意外存取，也有助於追蹤誰有權存取哪些資源。IAM 使用者和角色在

預設情況下沒有任何許可。根使用者預設擁有完整存取權，應受到嚴格監控，並僅用於[需要根存取權的任務](#)。

IAM 政策用於明確授予許可給 IAM 角色或特定資源。例如，以身分為基礎的政策可以連接到 IAM 群組，而 S3 儲存貯體可由以資源為基礎的政策控制。

建立 IAM 政策時，您就可以指定必須為 true 的服務動作、資源和條件，以便 AWS 允許或拒絕存取。AWS 支援各種條件，以協助您縮減存取權範圍。例如，透過使用 PrincipalOrgID [條件金鑰](#)，就可以在請求者不屬於您的 AWS 組織時拒絕動作。

此外，您還可以使用 CalledVia 條件金鑰控制 AWS 服務代您發出的請求，例如建立 AWS Lambda 函數的 AWS CloudFormation。您可以將不同的政策類型分層，以便建立深度防禦並限制使用者的整體許可。您還可以限制在什麼條件下，可以授予哪些許可。例如，您可以允許工作負載團隊為他們建置的系統建立自己的 IAM 政策，但前提是須套用[許可界限](#)來限制他們可授予的最大許可。

實作步驟

- 實作最低權限政策：將具有最低權限的存取政策指派給 IAM 群組和角色，以反映您已定義的使用者角色或職能。
- 透過分隔 AWS 帳戶 來隔離開發和正式作業環境：針對開發和正式作業環境使用不同的 AWS 帳戶，並使用[服務控制政策](#)、資源政策和身分政策來控制兩者之間的存取權。
- API 使用方式的基本政策：確定所需許可的一種方法是審核 AWS CloudTrail 日誌。您可以利用這種檢閱根據使用者在 AWS 內實際執行的動作，建立適合的許可。[IAM Access Analyzer](#) 可根據存取活動[自動產生](#) IAM 政策。您可以在組織或帳戶層級使用 IAM Access Advisor 來[追蹤特定政策的最後存取資訊](#)。
- 考慮使用 [AWS 受管政策來執行工作職能](#)：當您開始制定精細的許可政策時，針對帳單、資料庫管理員和資料科學家等常見的任務角色使用 AWS 受管政策會很實用。這些政策可在您決定如何實施最低權限政策的同時，協助縮小使用者的存取權。
- 移除不必要的許可：偵測並移除未使用的 IAM 實體、憑證和許可，以實踐最低權限原則。您可以使用 [IAM Access Analyzer](#) 來識別外部和未使用的存取權，而 [IAM Access Analyzer 政策產生](#) 功能可協助微調許可政策。
- 確保使用者對生產環境具有有限的存取權：使用者應只能存取具有有效使用案例的生產環境。在使用者執行完需要生產存取權的特定任務後，就應撤銷存取權。限制對生產環境的存取，有助於防止發生會影響生產的意外事件，也能降低意外存取的影響範圍。
- 考慮使用許可界限：[許可界限](#)是一種使用受管政策功能，其主要設定以身分為基礎的政策可授予 IAM 實體的最大許可權限。實體的許可界限可讓其只執行由身分類型政策和其許可界限同時允許的動作。

- 使用屬性型存取控制和資源標籤讓存取權更精簡：(若支援) 使用資源標籤的[屬性型存取控制 \(ABAC\)](#) 可用來精簡許可。您可以使用 ABAC 模型，此模型會比較主體標籤與資源標籤，以根據您定義的自訂維度來精簡存取權。這種方法可以簡化並減少組織中的許可政策數量。
- 建議僅在主體和資源皆為您的 AWS Organization 所擁有時，才使用 ABAC 進行存取控制。外部對象可以將與您的組織相同的標籤名稱和值，用於其自有主體和資源。如果您完全依賴這些名稱-值對來對外部對象的主體或資源授予存取權，則可能會提供意外的許可。
- 使用 AWS Organizations 的服務控制政策：[服務控制政策](#)可集中控制組織中成員帳戶的最大可用許可數。重要的是，您可以使用服務控制政策來限制成員帳戶中的根使用者許可。此外，還可以考慮使用 AWS Control Tower，它提供規範性受管控制，可以豐富 AWS Organizations。您也可以可以在 Control Tower 中定義自己的控制。
- 為您的組織制定使用者生命週期政策：使用者生命週期政策定義了當使用者上線至 AWS、變更職務或範圍或不再需要存取 AWS 時要執行的任務。在使用者生命週期的每個步驟進行許可審查，以確定許可受到適當限制並避免許可滲透的問題。
- 建立定期審查許可的排程，並移除任何不需要的許可：您應該定期審查使用者存取權，確認使用者沒有過度寬鬆的存取權。[AWS Config](#) 和 IAM Access Analyzer 可在使用者許可稽核過程中提供協助。
- 建立職務矩陣：職務矩陣會以視覺化的方式顯示您 AWS 足跡內所需的各種角色和存取層級。利用職務對照表，即可根據組織內的使用者職責定義和區分許可。使用群組，而不是將許可直接套用至個別使用者或角色。

資源

相關文件：

- [授予最低權限](#)
- [IAM 實體的許可界限](#)
- [用於寫入最低權限 IAM 政策的技巧](#)
- [IAM Access Analyzer 會根據存取活動產生 IAM 政策，因此更容易實作最低權限許可](#)
- [使用 IAM 許可界限將許可管理委派給開發人員](#)
- [使用上次存取的資訊精簡許可](#)
- [IAM 政策類型以及何時使用這些政策](#)
- [使用 IAM 政策模擬器測試 IAM 政策](#)
- [AWS Control Tower 中的防護機制](#)
- [零信任架構：AWS 觀點](#)
- [如何使用 CloudFormation StackSets 實作最低權限原則](#)

- [屬性型存取控制 \(ABAC\)](#)
- [檢視使用者活動以縮小政策範圍](#)
- [檢視角色存取](#)
- [使用標記來組織您的環境並推動問責制](#)
- [AWS 標記策略](#)
- [標記 AWS 資源](#)

相關影片：

- [下一代許可管理](#)
- [零信任：AWS 觀點](#)

SEC03-BP03 建立緊急存取程序

建立一項程序，在集中式身分提供者發生問題時，緊急存取您的工作負載。

您必須針對可能導致緊急事件發生的不同故障模式設計不同的程序。例如，正常情況下，您的員工使用者會使用集中式身分提供者 ([SEC02-BP04](#)) 聯合至雲端，以管理其工作負載。不過，如果您的集中式身分提供者發生錯誤，或是雲端中聯合的組態經過修改，則您的員工使用者可能無法連至雲端。緊急存取程序可讓授權的管理員透過替代方式 (例如聯合或直接使用者存取的替代形式) 存取您的雲端資源，以修正聯合組態或工作負載的問題。緊急存取程序會持續使用，直到恢復正常聯合機制為止。

預期成果：

- 您已定義並記載可視為緊急情況的故障模式：請考慮正常情況以及使用者用來管理工作負載的系統。考慮這些相依性如何發生錯誤並導致緊急情況。您可以在[可靠性支柱](#)中找到問題與最佳實務，有助於識別故障模式並架構更具彈性的系統，以盡量降低故障的可能性。
- 您已記載確認故障為緊急情況須遵循的步驟。例如，您可以要求身分管理員檢查主要和待命身分提供者的狀態，如果兩者都無法使用，則發佈身分提供者發生錯誤緊急事件。
- 您已針對每一種緊急或故障模式類型定義了緊急存取程序。明確定義可減少部分使用者過度使用一般程序，來處理所有類型的緊急情況。您的緊急存取程序描述了各個程序應在何種情況下使用，以及不應在哪些情況下使用，並指出可能適用的替代程序。
- 您的程序完整記載了詳細指示和程序手冊，可快速有效地遵循。請記住，緊急事件對使用者來說可能會非常緊張，他們可能面對極大的時間壓力，因此程序的設計應盡可能簡單。

常見的反模式：

- 您沒有詳細記載且經過充分測試的緊急存取程序。您的使用者未準備好面對緊急情況，而在緊急事件發生時只能隨機應變。
- 您的緊急存取程序與正常存取機制依賴相同的系統 (例如集中式身分提供者)。這表示，一旦這類系統發生故障，就可能同時影響您的正常和緊急存取機制，並損及您從故障復原的能力。
- 您的緊急存取程序用在非緊急情況。例如，您的使用者經常濫用緊急存取程序，因為他們發現直接進行變更透過管道提交變更更容易。
- 您的緊急存取程序未產生足夠的日誌來稽核程序，或是日誌未受監控，無法在發生可能濫用程序的情況時發出提醒。

建立此最佳實務的優勢：

- 只要有詳細記載且經充分測試的緊急存取程序，就能縮短使用者回應和解決緊急事件所花的時間。這樣就能進一步減少停機時間，並為客戶帶來更高的服務可用性。
- 您可以追蹤每一項緊急存取請求，以及偵測未經授權的人士試圖濫用程序來處理非緊急事件的情況，並發出提醒。

未建立此最佳實務時的風險暴露等級：中

實作指引

本節提供建立緊急存取程序的指引，用於處理與 AWS 上部署的工作負載相關的數種故障模式，一開始先介紹適用於所有故障模式的通用指引，接著再根據故障模式類型說明特定指引。

適用所有故障模式的通用指引

為故障模式設計緊急存取程序時，請考慮下列事項：

- 記載程序的前提和假設：應該和不應該使用程序的時機。這樣做有助於詳細說明故障模式並記載假設，例如其他相關系統的狀態。舉例來說，故障模式 2 的程序假設身分提供者可以使用，但 AWS 上的組態已經過修改或已過期。
- 預先建立緊急存取程序所需的資源 ([SEC10-BP05](#))。例如，在所有工作負載帳戶中預先建立具有 IAM 使用者和角色的緊急存取 AWS 帳戶，以及跨帳戶 IAM 使用者角色。這樣就可確定這些資源在緊急事件發生時立即可用。透過預先建立資源，您就不必依賴 AWS [控制平面](#) API (用來建立和修改 AWS 資源)，因為它們在緊急情況下可能無法使用。此外，預先建立 IAM 資源就不需考慮[因最終一致性而可能發生的延遲](#)。
- 請將緊急存取程序納入您的事件管理計畫中 ([SEC10-BP02](#))。記載緊急事件的追蹤方式，並傳達給組織中的其他人，例如同儕團隊、您的領導階層，以及適時向外傳達給您的客戶和業務合作夥伴。

- 在您現有的服務請求工作流程系統 (若有的話) 中定義緊急存取請求程序。一般來說，這類工作流程系統可讓您建立接收表單來收集有關請求的資訊、在工作流程的每個階段追蹤請求，以及新增自動和手動核准步驟。將每一個請求與事件管理系統中追蹤的對應緊急事件建立關聯。採用統一的緊急存取系統，可讓您在單一系統中追蹤這些請求、分析使用趨勢並改善程序。
- 確認您的緊急存取程序只能由經授權的使用者啟動，並且視情況要求使用者同儕或管理層的核准。核准程序在營業時間內外都要能夠有效運作。定義在主要核准者沒有空的情況下，如何由次要核准者核准請求，以及如何在您的管理鏈中向上呈報，直到請求獲得核准。
- 實作可靠的記錄、監控和警示機制，以因應緊急存取程序和機制。同時針對成功和失敗的嘗試產生詳細的稽核日誌，以便取得緊急存取權。在事件管理系統中將活動與正在發生的事件相互關聯，並在動作發生於預期時段之外，或在正常操作期間使用緊急存取帳戶時發出警示。緊急存取帳戶只能在緊急情況下存取，因為破窗程序也可視為後門。與您的安全資訊和事件管理 (SIEM) 工具或 [AWS Security Hub CSPM](#) 整合，以在緊急存取期間回報和稽核所有活動。恢復正常操作時，自動輪換緊急存取憑證，並通知相關團隊。
- 定期測試緊急存取程序，以確認步驟是否清楚，並且快速有效地授予正確的存取層級。您的緊急存取程序應在事件回應模擬的過程中 ([SEC10-BP07](#)) 和災難復原測試中 ([REL13-BP03](#)) 進行測試。

故障模式 1：用於聯合至 AWS 的身分提供者無法使用

如 [SEC02-BP04 仰賴集中化的身分提供者](#) 中所述，我們建議您仰賴集中式身分提供者來聯合您的員工使用者，以授予 AWS 帳戶的存取權。您可以使用 IAM Identity Center 聯合至您 AWS 組織中的多個 AWS 帳戶，或是使用 IAM 聯合至個別 AWS 帳戶。在這兩種情況下，員工使用者都會先透過集中式身分提供者進行身分驗證，然後才重新導向至 AWS 登入端點進行單一登入。

萬一您的集中式身分提供者無法使用，您的員工使用者就無法聯合至 AWS 帳戶 或管理其工作負載。在此緊急事件中，您可以提供緊急存取程序讓一小群管理員存取 AWS 帳戶，以便執行無法等到集中式身分提供者恢復連線後才處理的重要工作。例如，您的身分提供者停擺了 4 小時，而在此期間，您需要修改生產帳戶中 Amazon EC2 Auto Scaling 群組的上限，以處理客戶流量意外暴增的情況。您的緊急管理員應遵循緊急存取程序，才能獲得特定生產 AWS 帳戶的存取權並進行必要的變更。

緊急存取程序依賴預先建立的緊急存取 AWS 帳戶，該帳戶單純用於緊急存取，並擁有 AWS 資源 (例如 IAM 角色和 IAM 使用者) 可支援緊急存取程序。在正常操作期間，任何人都不應該存取緊急存取帳戶，而且您必須監控濫用此帳戶的情況並發出提醒 (如需詳細資訊，請參閱前一節「通用指引」)。

緊急存取帳戶具有緊急存取 IAM 角色，有權在需要緊急存取的 AWS 帳戶中擔任跨帳戶角色。這些 IAM 角色會預先建立並設定信任政策，以便信任緊急帳戶的 IAM 角色。

緊急存取程序可以使用下列其中一種方法：

- 您可以預先建立一組 [IAM 使用者](#) 並包含關聯的強式密碼和 MFA 字符，以供緊急存取帳戶中的緊急管理員使用。這些 IAM 使用者有權擔任 IAM 角色，且後續可在需要緊急存取時跨帳戶存取 AWS 帳戶。我們建議這類使用者的數量越少越好，並且將每一位使用者指派給單一緊急管理員。在緊急情況下，緊急管理員使用者會使用其密碼和 MFA 字符代碼登入緊急存取帳戶，切換到緊急帳戶中的緊急存取 IAM 角色，最後再切換到工作負載帳戶中的緊急存取 IAM 角色，以執行緊急變更動作。這種方法的優點是，每個 IAM 使用者都會指派給一名緊急管理員，而且您可以透過審核 CloudTrail 事件得知登入的使用者。缺點是，您必須維護多個 IAM 使用者及其關聯的長期存在密碼和 MFA 字符。
- 您可以使用緊急存取 [AWS 帳戶 根使用者](#) 來登入緊急存取帳戶、擔任緊急存取的 IAM 角色，並且在工作負載帳戶中擔任跨帳戶角色。我們建議您為根使用者設定強式密碼和多個 MFA 字符。同時也建議您，將密碼和 MFA 字符儲存在強制執行強式身分驗證和授權的安全企業憑證保存庫中。您應確保密碼和 MFA 字符重設要素的安全性：將帳戶的電子郵件地址設定為受到您的雲端安全管理員監控的電子郵件分發清單，並將帳戶的電話號碼設定為同樣受到安全管理員監控的共用電話號碼。這種方法的優點是，只需管理一組根使用者憑證。缺點是，由於這是共用使用者，因此有多個管理員能夠以根使用者的身分登入。您必須稽核企業保存庫日誌事件，以確定哪個管理員簽出了根使用者密碼。

故障模式 2：AWS 上的身分提供者組態已經過修改或已過期

若要讓您的員工使用者聯合至 AWS 帳戶，您可以使用外部身分提供者設定 IAM Identity Center，或建立 IAM 身分提供者 ([SEC02-BP04](#))。一般來說，您可以匯入身分提供者提供的 SAML 中繼資料 XML 文件來進行這些設定。中繼資料 XML 文件包含一個 X.509 憑證，對應於身分提供者用來簽署其 SAML 聲明的私有金鑰。

AWS 端的這些組態可能遭到管理員誤改或誤刪。另一種情況是，匯入 AWS 中的 X.509 憑證可能過期，而具有新憑證的新中繼資料 XML 尚未匯入 AWS 中。這兩種情況都可能使員工使用者的 AWS 聯合中斷，導致緊急情況發生。

在這類緊急事件中，您可以提供 AWS 的存取權給身分管理員，以修正聯合問題。例如，您的身分管理員使用緊急存取程序登入緊急存取 AWS 帳戶、切換為 Identity Center 管理員帳戶中的角色，並透過從您的身分提供者匯入最新的 SAML 中繼資料 XML 文件來更新外部身分提供者組態，以重新啟用聯合。聯合修復後，您的員工使用者繼續使用正常操作程序來聯合至其工作負載帳戶。

您可以依照先前「故障模式 1」中詳述的方法來建立緊急存取程序。您可以將最低權限許可授予身分管理員，以限制他們只能存取 Identity Center 管理員帳戶以及在該帳戶中對 Identity Center 執行動作。

故障模式 3：Identity Center 中斷

萬一發生 IAM Identity Center 或 AWS 區域中斷的情況，建議您設定一個可用來暫時存取 AWS 管理主控台的組態。

緊急存取程序會在緊急帳戶中，使用您的身分提供者對 IAM 的直接聯合。如需有關程序和設計考量的詳細資訊，請參閱[設定 AWS 管理主控台的緊急存取](#)。

實作步驟

適用所有故障模式的通用步驟

- 建立緊急存取程序專用的 AWS 帳戶。在帳戶中預先建立所需的 IAM 資源，例如 IAM 角色或 IAM 使用者，也可以選擇建立 IAM 身分提供者。此外，在工作負載 AWS 帳戶中預先建立跨帳戶 IAM 角色，並與緊急存取帳戶中對應的 IAM 角色建立信任關係。您可以使用 [CloudFormation StackSets 搭配 AWS Organizations](#) 在組織的成員帳戶中建立此類資源。
- 建立 AWS Organizations [服務控制政策](#) (SCP) 以拒絕刪除和修改成員 AWS 帳戶中的跨帳戶 IAM 角色。
- 為緊急存取 AWS 帳戶啟用 CloudTrail，並將追蹤事件傳送到日誌收集 AWS 帳戶中的中央 S3 儲存貯體。如果您使用 AWS Control Tower 來設定和管控您的 AWS 多帳戶環境，則您使用 AWS Control Tower 建立或在 AWS Control Tower 中註冊的每個帳戶都會預設啟用 CloudTrail，並傳送至專用日誌封存 AWS 帳戶中的 S3 儲存貯體。
- 透過建立在主控台登入時比對的 EventBridge 規則來監控活動的緊急存取帳戶，以及透過緊急 IAM 角色監控 API 活動。當活動於事件管理系統中追蹤的持續緊急事件之外發生時，傳送通知給您的安全營運中心。

適用「故障模式 1：用於聯合至 AWS 的身分提供者無法使用」及「故障模式 2：AWS 上的身分提供者組態已經過修改或已過期」的其他步驟

- 根據您選擇的緊急存取機制預先建立資源：
 - 使用 IAM 使用者：預先建立 IAM 使用者並設定強式密碼和關聯的 MFA 裝置。
 - 使用緊急帳戶根使用者：設定根使用者使用強式密碼，並將密碼儲存在您的企業憑證保存庫中。將多個實體 MFA 裝置與根使用者建立關聯，並將裝置儲存在您的緊急管理員小組成員可快速存取的位置。

適用「故障模式 3：Identity Center 中斷」的其他步驟

- 如[設定 AWS 管理主控台的緊急存取](#)中所述，在緊急存取 AWS 帳戶中，建立 IAM Identity Provider 身分提供者，以從您的身分提供者啟用直接 SAML 聯合。
- 在 IdP 中建立緊急操作群組，但不新增任何成員。

- 在緊急存取帳戶中建立對應於緊急操作群組的 IAM 角色。

資源

相關 Well-Architected 的最佳實務：

- [SEC02-BP04 仰賴集中式身分提供者](#)
- [SEC03-BP02 授予最低權限存取權](#)
- [SEC10-BP02 制定事件管理計畫](#)
- [SEC10-BP07 執行演練日](#)

相關文件：

- [設定 AWS 管理主控台 的緊急存取](#)
- [使 SAML 2.0 聯合身分使用者能夠存取 AWS 管理主控台](#)
- [緊急存取](#)

相關影片：

- [AWS re:Invent 2022 - 使用 IAM Identity Center 簡化現有的員工存取權](#)
- [AWS re:Inforce 2022 - AWS Identity and Access Management \(IAM\) 深入剖析](#)

相關範例：

- [AWS 緊急存取角色](#)
- [AWS 客戶程序手冊架構](#)
- [AWS 事件回應程序手冊範例](#)

SEC03-BP04 持續減少許可

在團隊確定所需的存取權時，請移除不需要的許可，並建立審查程序以達到最低權限的許可。持續監視人類和機器存取權，並移除不使用的身分和許可。

預期成果：許可政策應遵循最低權限原則。隨著工作職責和角色的定義變得更具體，您需要審查許可政策以移除不必要的許可。若憑證遭到意外洩露或以其他方式在未經授權下遭存取，此方法可縮小影響範圍。

常見的反模式：

- 預設授予使用者管理員許可。
- 建立過於寬鬆的政策，但不具完整的管理員權限。
- 保留不再需要的許可政策。

未建立此最佳實務時的曝險等級：中

實作指引

在團隊和專案剛開始時，可使用寬鬆的許可政策來激發創新和敏捷性。例如，在開發或測試環境中，可以讓開發人員存取廣泛的 AWS 服務。我們建議您持續評估存取權，並將存取權限於完成目前工作所需的這些服務和服務動作。我們建議對人類和機器身分進行此項評估。機器身分有時候稱為系統或服務帳戶，是提供 AWS 存取權給應用程式或伺服器的身分。此存取權在生產環境中尤為重要，因為過於寬鬆的許可可能影響廣大而且可能暴露客戶資料。

AWS 提供多種方法可協助識別未使用的使用者、角色、許可和憑證。AWS 也有助於分析 IAM 使用者和角色的存取活動，包括關聯的存取金鑰，以及對 AWS 資源的存取，例如 Amazon S3 儲存貯體中的物件。AWS Identity and Access Management Access Analyzer 政策產生可協助您根據某主體進行互動的實際服務和動作來建立限制性許可。[屬性型存取控制 \(ABAC\)](#) 可以協助簡化許可管理，因為您可以使用使用者的屬性提供許可給使用者，而不是將許可政策直接連接至每個使用者。

實作步驟

- 使用 [AWS Identity and Access Management Access Analyzer](#)：IAM Access Analyzer 可協助您識別組織和帳戶中[與外部實體共用](#)的資源，例如 Amazon Simple Storage Service (Amazon S3) 儲存貯體或 IAM 角色。
- 使用 [IAM Access Analyzer 政策產生](#)：IAM Access Analyzer 政策產生可協助您[根據 IAM 使用者或角色的存取活動建立精細的許可政策](#)。
- 在正式作業前測試較低環境的許可：一開始先使用[較不重要的沙盒和開發環境](#)，利用 IAM Access Analyzer 來測試各種不同工作職能所需的許可。然後在測試、品質保證和預備環境逐步加強並驗證這些許可，再將許可套用至正式作業環境。較低的環境一開始可以擁有較寬鬆的許可，因為服務控制政策 (SCP) 會透過限制授予的許可數上限來強制執行防護機制。
- 為 IAM 使用者和角色確定可接受的時間範圍和使用政策：使用[上次存取的時間戳記](#)以[識別未使用的使用者和角色](#)並將其移除。審核服務和動作上次存取的資訊，以識別和[設定特定使用者和角色的許可範圍](#)。例如，您可以使用上次存取的資訊來識別您的應用程式角色所需的特定 Amazon S3 動作，並將該角色的存取權僅限於這些動作。AWS 管理主控台 中有提供「上次存取的資訊」功能，並且可讓您透過程式設計的方式將這些功能併入基礎設施工作流程和自動化工具中。

- 考慮在 [AWS CloudTrail 中日誌記錄資料事件](#)：在預設情況下，CloudTrail 不會記錄資料事件，例如 Amazon S3 物件層級活動 (如 GetObject 和 DeleteObject) 或 Amazon DynamoDB 資料表活動 (如 PutItem 和 DeleteItem)。考慮對這些事件使用日誌記錄功能，以確定哪些使用者和角色需要存取特定 Amazon S3 物件或 DynamoDB 資料表項目。

資源

相關文件：

- [授予最低權限](#)
- [移除不必要的憑證](#)
- [什麼是 AWS CloudTrail？](#)
- [使用 政策](#)
- [DynamoDB 中的日誌記錄和監控](#)
- [對您的 Amazon S3 儲存貯體和物件使用 CloudTrail 事件日誌記錄](#)
- [取得 AWS 帳戶 的憑證報告](#)

相關影片：

- [在 60 分鐘內精通 IAM 政策](#)
- [責任區隔、最低權限、委派和 CI/CD](#)
- [AWS re:Inforce 2022 - AWS Identity and Access Management \(IAM\) 深入剖析](#)

SEC03-BP05 為您的組織定義許可防護機制

使用許可防護機制縮小可授予主體的可用許可範圍。許可政策評估鏈包括您的防護機制，可在進行授權決策時確定主體的有效許可。您可以採用分層方式定義防護機制。對整個組織廣泛套用一些防護機制，另外對暫時存取工作階段套用一些精細的防護機制。

預期成果：您可以使用個別 AWS 帳戶 對環境進行明確的隔離。服務控制政策 (SCP) 用於定義整個組織的許可防護機制。較寬鬆的防護機制設定於最靠近組織根目錄的階層層級，較嚴謹的防護機制則設定於較靠近個別帳戶的層級。

在受支援的情況下，資源政策會定義主體必須滿足才能取得資源存取權的條件。資源政策也會適時縮小允許的動作範圍。許可界限會設置在管理工作負載許可的主體上，以將許可管理委派給個別工作負載擁有者。

常見的反模式：

- 在 [AWS 組織](#) 內建立成員 AWS 帳戶，但未使用 SCP 來限制其根憑證適用的用途和許可。
- 根據最低權限指派許可，但未對可授予的許可集上限設置防護機制。
- 仰賴 AWS IAM 的隱含拒絕基礎來限制許可，相信政策不會授予不需要的明確允許許可。
- 在相同 AWS 帳戶 中執行多個工作負載環境，然後仰賴 VPC、標籤或資源政策等機制來強制執行許可界限。

建立此最佳實務的優勢：許可防護機制有助於建立信心，確保不會有不需要的許可授予情況，即使許可政策嘗試這樣做也不必擔心。此最佳實務可透過縮小需考量的許可範圍上限來簡化定義和管理許可。

未建立此最佳實務時的曝險等級：中

實作指引

建議您採用分層方式為您的組織定義許可防護機制。此方法能夠隨著套用額外的分層，有系統地減少可能的許可集上限。這可協助您根據最低權限原則授予存取權，降低了因政策組態錯誤導致意外存取的風險。

建立許可防護機制的第一步，是將您的工作負載和環境隔離到個別 AWS 帳戶 中。在沒有明確許可的情況下，某一帳戶中的主體無法存取另一帳戶中的資源，即使兩個帳戶在相同 AWS 組織中或在相同 [組織單位 \(OU\)](#) 下亦是如此。您可以使用 OU 將您要管理的帳戶分組為一個單位。

下一步是減少您可授予組織的成員帳戶內主體的許可集上限。您可以使用 [服務控制政策 \(SCP\)](#) 達到此目的，這些政策可套用至 OU 或帳戶。SCP 可強制執行通用的存取控制，例如限制對特定 AWS 區域的存取、協助防止資源遭到刪除，或停用有潛在風險的服務動作。您套用至組織根目錄的 SCP 只會影響其成員帳戶，而不會影響管理帳戶。SCP 只會控管組織內的主體。您的 SCP 不會控管組織外部存取您資源的主體。

如果您使用 [AWS Control Tower](#)，則可以利用其 [控制項](#) 和 [登陸區域](#) 做為許可防護機制和多帳戶環境的基礎。登陸區域提供預先設定的安全基準環境，當中採用不同的帳戶來處理不同的工作負載和應用程式。防護機制透過服務控制政策 (SCP)、AWS Config 規則及其他組態的組合，強制執行有關安全、操作和合規的強制性控制。不過，同時使用 Control Tower 防護機制和登陸區域與自訂 Organization SCP 時，務必遵循 AWS 文件中所述的最佳實務，以避免衝突並確保適當的控管。如需在 Control Tower 環境中管理 SCP、帳戶和組織單位 (OU) 的詳細建議，請參閱 [AWS Organizations 的 AWS Control Tower 指引](#)。

只要遵循這些指引，您就能有效利用 Control Tower 的防護機制、登陸區域和自訂 SCP，同時減少前在衝突，並確保適當控管與控制多帳戶 AWS 環境。

再下一步是使用 [IAM 資源政策](#) 來設定您可對其控管的資源執行的可用動作範圍，以及設定執行動作的主體必須符合的任何條件。這個範圍可以很廣泛，例如只要主體屬於組織的一部分就允許所有動作 (使用 PrincipalOrgId [條件金鑰](#))，也可以很精細，例如只允許特定 IAM 角色執行特定動作。您可以在 IAM 角色信任政策中採取類似方法，並附帶條件。如果資源或角色信任政策明確指名相同帳戶中的某個主體作為其控管的角色或資源，則該主體不需要有授予相同許可的連接 IAM 政策。如果主體位於與資源不同的帳戶中，則該主體確實需要有授予這些許可的連接 IAM 政策。

通常，工作負載團隊會希望管理其工作負載需要的許可。這樣一來，他們便需要建立新的 IAM 角色和許可政策。您可以擷取允許團隊在 [IAM 許可界限](#) 中授予的許可範圍上限，並將此文件與 IAM 角色關聯，之後團隊就可使用該角色來管理其 IAM 角色和許可。這種方法可讓他們彈性地完成其工作，同時降低擁有 IAM 管理存取權的風險。

更詳細的步驟是實作特殊權限存取管理 (PAM) 和暫時提升存取管理 (TEAM) 技術。PAM 的範例是，要求主體在採取特殊權限動作之前執行多重要素驗證。如需詳細資訊，請參閱 [設定受 MFA 保護的 API 存取](#)。TEAM 需要使用解決方案來管理允許主體擁有已提升存取權的核准和時間範圍。其中一種方法是暫時將主體新增至具有已提升存取權之 IAM 角色的角色信任政策中。另一種方法是在正常操作情況下，使用 [工作階段政策](#) 縮小 IAM 角色授予主體的許可範圍，然後在核准時段暫時解除此限制。若要進一步了解已經過 AWS 和精選合作夥伴驗證的解決方案，請參閱 [暫時提升的存取權](#)。

實作步驟

1. 將您的工作負載和環境隔離到個別 AWS 帳戶中。
2. 使用 SCP 減少可授予組織的成員帳戶內主體的許可集上限。
 - a. 定義 SCP 以減少可授予組織成員帳戶中主體的許可集上限時，您可以選擇允許清單或拒絕清單方法。允許清單策略會明確指定允許的存取權，並隱含地封鎖所有其他存取權。拒絕清單策略會明確指定不允許的存取權，並預設允許所有其他存取權。這兩種策略各有其優點和權衡，而適當的選擇取決於組織的特定要求和風險模型。如需詳細資訊，請參閱 [使用 SCP 的策略](#)。
 - b. 此外，請查看 [服務控制政策範例](#)，了解如何有效地建構 SCP。
3. 使用 IAM 資源政策可縮小資源上許可動作的範圍並指定條件。在 IAM 角色信任政策中使用條件來建立擔任角色的限制。
4. 將 IAM 許可界限指派至 IAM 角色，之後工作負載團隊可以使用這些角色來管理自己的工作負載 IAM 角色和許可。
5. 根據您的需求評估 PAM 和 TEAM 解決方案。

資源

相關文件：

- [上的資料周邊AWS](#)
- [使用資料周邊建立許可防護機制](#)
- [政策評估邏輯](#)

相關範例：

- [服務控制政策範例](#)

相關工具：

- [AWS 解決方案：暫時提升存取管理](#)
- [適用 TEAM 的已驗證安全合作夥伴解決方案](#)

SEC03-BP06 根據生命週期管理存取

監控並調整授予您的主體 (使用者、角色和群組) 在組織內其整個生命週期的許可。隨著使用者變更角色調整群組成員資格，並在使用者離開組織時移除存取權。

預期成果您可在組織內監控並調整主體整個生命週期的許可，進而降低不必要權限帶來的風險。您可在建立使用者時授予適當的存取權。您可以隨著使用者的職責變更修改存取權，並且在使用者不再為作用中狀態或離開組織時移除存取權。您可以集中管理使用者、角色和群組的變更。您使用自動化的方式將變更傳播到您的 AWS 環境。

常見的反模式：

- 事先授予身分過多或過廣的存取權限，超過最初所需的範圍。
- 當身分的角色和職責隨時間發生變更後，未能審查並調整存取權限。
- 未移除非作用中或已終止身分的作用中存取權限。此舉會增加未經授權存取的風險。
- 未利用自動化來管理身分的生命週期。

未建立此最佳實務時的風險暴露等級：中

實作指引

在身分的整個生命週期中，仔細管理和調整您授予身分 (例如使用者、角色、群組) 的存取權限。此生命週期涵蓋初始入職階段、後續角色和職責變更，以及最終離職或終止。根據生命週期階段主動管理存取權，以維護適當的存取層級。遵守最低權限原則，以降低過度或不必要存取權限帶來的風險。

您可以直接在 AWS 帳戶 內管理 IAM 使用者的生命週期，或透過從人力資源身分提供者至 [AWS IAM Identity Center](#) 的聯合來進行管理。針對 IAM 使用者，您可以在 AWS 帳戶 內建立、修改和刪除使用者及其關聯的許可。針對聯合身分使用者，您可以使用 IAM Identity Center 管理生命週期，方法是透過 [跨網域身分管理系統](#) (SCIM) 協定，同步源自您組織身分提供者的使用者和群組資訊。

SCIM 是開放標準協定，可在不同系統中自動佈建和解除佈建使用者身分。透過將身分提供者與使用 SCIM 的 IAM Identity Center 整合，您就可以自動同步使用者和群組資訊，協助驗證組織是否根據其授權身分來源的變更授予、修改或撤銷存取權限。

隨著組織內員工的角色和職責改變，調整他們的存取權限。您可以使用 IAM Identity Center 的許可集來定義不同的工作角色或職責，並將其與適當的 IAM 政策和許可關聯。當員工的角色變更時，您可以更新其指派的許可集，以反映他們的新職責。確認他們具有必要的存取權，同時遵守最低權限原則。

實作步驟

1. 定義並記錄存取管理生命週期流程，包括授予初始存取權、定期審查和離職的程序。
2. 實作 [IAM 角色、群組和許可界限](#)，以共同管理存取權，並強制執行受允許的最高存取層級。
3. 使用 IAM Identity Center 與 [聯合身分提供者](#) (例如 Microsoft Active Directory、Okta、Ping Identity) 整合，使其做為使用者和群組資訊的授權來源。
4. 使用 [SCIM](#) 協定可將來自身分提供者的使用者和群組資訊同步到 IAM Identity Center 的身分存放區中。
5. 在 IAM Identity Center 中建立 [許可集](#)，以代表組織內不同的工作角色或職責。為每個許可集定義適當的 IAM 政策和許可。
6. 實作定期存取權審查、提示撤銷存取權，以及持續改進存取權管理生命週期流程。
7. 為員工提供有關存取權管理最佳實務的培訓和認知。

資源

相關的最佳實務：

- [SEC02-BP04 仰賴集中式身分提供者](#)

相關文件：

- [管理您的身分來源](#)
- [管理 IAM Identity Center 中的身分](#)
- [使用 AWS Identity and Access Management Access Analyzer](#)

- [IAM Access Analyzer 政策產生](#)

相關影片：

- [AWS re:Inforce 2023 - 使用 AWS IAM Identity Center 管理暫時提升的存取權](#)
- [AWS re:Invent 2022 - 使用 IAM Identity Center 簡化現有的員工存取權](#)
- [AWS re:Invent 2022 - 使用 Access Analyzer 掌握 IAM 政策之力並駕馭許可](#)

SEC03-BP07 分析公有和跨帳戶存取權

持續監控強調公有和跨帳戶存取權的調查結果。減少僅對需要此存取之特定資源的公有存取權和跨帳戶存取權。

預期成果：知道您共用了哪些 AWS 資源以及共用的對象。持續監控和稽核您共用的資源以確認這些資源僅與已授權主體共用。

常見的反模式：

- 沒有維持共用資源的詳細目錄。
- 未遵循程序來核准跨帳戶或資源的公有存取權。

未建立此最佳實務時的曝險等級：低

實作指引

如果您的帳戶位於 AWS Organizations 中，您可以將資源的存取權授予整個組織、特定組織單位或個別帳戶。如果您的帳戶不是組織的成員，您可以與個別帳戶共用資源。您可以使用以資源為基礎的政策 (例如 [Amazon Simple Storage Service \(Amazon S3\) 儲存貯體政策](#)) 來授予直接跨帳戶存取權，或允許另一個帳戶中的主體擔任您帳戶中的 IAM 角色。當使用資源政策時，確認僅將該存取權授予已授權的主體。定義程序，來核准所有需要公開提供的資源。

[AWS Identity and Access Management Access Analyzer](#) 採用 [可證明的安全性](#) 來識別從其帳戶外部存取資源的所有路徑。它會持續審查資源政策，並報告公有和跨帳戶存取權的調查結果，讓您輕鬆分析潛在的各種存取。考慮使用 AWS Organizations 設定 IAM Access Analyzer，以確認您對所有帳戶具有可見性。IAM Access Analyzer 也允許您在部署資源許可之前 [預覽調查結果](#)。這可讓您驗證政策變更是否僅授予對您資源的預期公有和跨帳戶存取權。設計多帳戶存取權時，您可以使用 [信任政策](#) 來控制可以擔任角色的情況。例如，您可以使用 [PrincipalOrgId](#) 條件金鑰來拒絕嘗試從 [AWS Organizations](#) 外部擔任角色的動作。

[AWS Config](#) 可以報告設定不當的資源，並可透過 AWS Config 政策檢查來偵測已設定公開存取的資源。諸如 [AWS Control Tower](#) 和 [AWS Security Hub CSPM](#) 的服務可簡化在 AWS Organizations 間部署控制和防護機制的作業，以識別和修正公開暴露的資源。例如，AWS Control Tower 具備受管的防護機制，可偵測是否有任何[可由 AWS 帳戶 還原的 Amazon EBS 快照](#)。

實作步驟

- 考慮對 [AWS Organizations](#) 使用 [AWS Config](#)：AWS Config 可讓您將 AWS Organizations 內來自多個帳戶的調查結果彙總到一個委派的管理員帳戶。這提供了全面性檢視，並讓您[在帳戶間部署 AWS Config 規則 以偵測公開可存取的資源](#)。
- 設定 AWS Identity and Access Management Access Analyzer：IAM Access Analyzer 可協助您識別組織及帳戶中[與外部實體共用](#)的資源，例如 Amazon S3 儲存貯體或 IAM 角色。
- 使用 AWS Config 中的自動修復以回應 Amazon S3 儲存貯體的公開存取設定中的變更：[您可以自動開啟 Amazon S3 儲存貯體的封鎖公開存取設定](#)。
- 實作監控和提醒以識別 Amazon S3 儲存貯體是否已變為公有：您必須設立[監控和提醒](#)以識別何時關閉 Amazon S3 封鎖公開存取，以及 Amazon S3 儲存貯體是否變為公有。此外，如果您正在使用 AWS Organizations，可以建立[服務控制政策](#)來防止對 Amazon S3 公開存取政策進行變更。[AWS Trusted Advisor](#) 會檢查具有公開存取許可的 Amazon S3 儲存貯體。將上傳或刪除存取權授予每個人的儲存貯體許可，可讓任何人在儲存貯體中新增、修改或移除項目，進而產生潛在的安全問題。Trusted Advisor 檢查會分析明確的儲存貯體許可，以及可能覆寫儲存貯體許可的關聯儲存貯體政策。您也可以使用 AWS Config 來監控 Amazon S3 儲存貯體的公開存取。如需詳細資訊，請參閱[如何使用 AWS Config 來監控及回應允許公開存取的 Amazon S3 儲存貯體](#)。

檢閱 Amazon S3 儲存貯體的存取控制時，請務必考量儲存於其中的資料性質。[Amazon Macie](#) 是一項服務，其設計在於協助您探索和保護敏感資料，例如個人身分識別資訊 (PII)、受保護健康資訊 (PHI)，以及私有金鑰或 AWS 存取金鑰等憑證。

資源

相關文件：

- [使用 AWS Identity and Access Management Access Analyzer](#)
- [AWS Control Tower 控制程式庫](#)
- [AWS 基礎安全最佳實務標準](#)
- [AWS Config 受管規則](#)
- [AWS Trusted Advisor 檢查參考](#)
- [使用 Amazon EventBridge 監控 AWS Trusted Advisor 檢查結果](#)

- [管理組織內所有帳戶間的 AWS Config 規則](#)
- [AWS Config 和 AWS Organizations](#)
- [在 Amazon EC2 中公開提供您的 AMI](#)

相關影片：

- [保護多帳戶環境的最佳實務](#)
- [深入了解 IAM Access Analyzer](#)

SEC03-BP08 在組織內安全地共用資源

隨著工作負載數量增加，您可能需要在這些工作負載內共用對資源的存取，或在多個帳戶間多次佈建資源。您可能具備劃分環境 (例如擁有開發、測試和生產環境) 的建構模組。然而，擁有分隔建構模組並不會限制您安全共用的能力。透過共用重疊的元件，您可以降低營運負擔並允許一致的體驗，而不用猜測在多次建立相同的資源時可能錯過了什麼。

預期成果：使用安全方法在組織內共用資源，藉此充分減少意外存取，並協助您的資料外洩防護計畫。減輕與管理個別元件相較下的營運負擔，減少多次手動建立相同元件的錯誤，以及增加工作負載的可擴展性。您可以從多點失敗案例中更短的解決時間獲益，並更有信心確定何時不再需要某元件。如需有關分析外部共用的資源的方案指引，請參閱 [SEC03-BP07 分析公有和跨帳戶存取權](#)。

常見的反模式：

- 缺乏可持續監控和自動發出意外外部共用通知的程序。
- 對於應該和不應該共用的內容缺乏基準。
- 預設採用廣泛的開放政策而不是在必要時明確共用。
- 必要時手動建立重疊的基礎資源。

未建立此最佳實務時的曝險等級：中

實作指引

建構您的存取控制和模式來管控安全地取用共用資源並只與信任的實體共用。監控共用資源並持續審查共用資源存取，在不當或意外共用時獲得提醒。審核[分析公開和跨帳戶存取權](#)協助您確立管控能力以減少外部存取，而僅限於需要存取的資源，以及建立程序持續監控並自動提供提醒。

在 AWS Organizations 內跨帳戶共用受到[數個 AWS 服務](#)的支援，例如 [AWS Security Hub CSPM](#)、[Amazon GuardDuty](#) 和 [AWS Backup](#)。這些服務允許將資料共用到中央帳戶，從中央帳戶存

取，或從中央帳戶管理資源和資料。例如，AWS Security Hub CSPM 可以將調查結果從個別帳戶轉移到中央帳戶，讓您能夠檢視所有調查結果。AWS Backup 可以對資源進行備份並在帳戶之間共用。您可以使用 [AWS Resource Access Manager \(AWS RAM\)](#) 來共用其他常見的資源，例如 [VPC 子網路和 Transit Gateway 附件](#)、[AWS Network Firewall](#) 或 [Amazon SageMaker AI 管道](#)。

若要將您的帳戶限制為僅共用組織內的資源，請使用[服務控制政策 \(SCP\)](#) 防止存取外部主體。當共用資源時，結合身分型控制和網路控制為您的組織建立資料周邊，以協助預防意外存取。資料周邊是一組預防性防護機制，可協助確認只有可信的身分從預期網路存取可信的資源。這些控制應適當限制可以共用哪些資源，並防止共用或公開不應該允許的資源。例如，作為資料周邊的一部分，您可以使用 VPC 端點政策和 `AWS:PrincipalOrgId` 條件來確保存取 Amazon S3 儲存貯體的身分屬於您的組織。請務必注意，[SCP 不適用於連結服務的角色或 AWS 服務主體](#)。

當使用 Amazon S3 時，請[關閉 Amazon S3 儲存貯體的 ACL](#) 並使用 IAM 政策來定義存取控制。若要限制從 [Amazon CloudFront](#) 對 Amazon S3 原點的存取，請從原始存取身分 (OAI) 遷移至原始存取控制 (OAC)，後者支援額外的功能，包括使用 [AWS Key Management Service](#) 的伺服器端加密。

在某些情況下，您可能會想要允許在組織外部共用資源或將資源的存取權授予第三方。如需有關管理許可以在外部共用資源的方案指引，請參閱[許可管理](#)。

實作步驟

1. 使用 AWS Organizations：AWS Organizations 是一項帳戶管理服務，可讓您將多個 AWS 帳戶整合到您所建立和集中管理的組織中。您可以將帳戶編組成組織單位 (OU) 並將不同的政策連接至各個 OU，以協助滿足您的預算、安全和合規需求。您也可以控制 AWS 人工智慧 (AI) 和機器學習 (ML) 服務收集和儲存資料的方式，並使用與 Organizations 整合的 AWS 服務的多帳戶管理功能。
2. 將 AWS Organizations 與 AWS 服務整合：當您啟用 AWS 服務代表您在組織的成員帳戶中執行任務時，AWS Organizations 會在每個成員帳戶中為該服務建立 IAM 服務連結角色 (SLR)。您應該使用 AWS 管理主控台、AWS API 或 AWS CLI 來管理可信存取。如需有關開啟可信存取的方案指引，請參閱[使用 AWS Organizations 與其他 AWS 服務](#)以及[您可以搭配 Organizations 使用的 AWS 服務](#)。
3. 建立資料周邊：資料周邊提供明確的信任和所有權界限。在 AWS 上，它通常表示為 AWS Organizations 所管理的組織 AWS，以及存取您的 AWS 資源的任何內部部署網路或系統。資料周邊的目標是要確認若身分可信、資源可信且是預期的網路，則允許存取。不過，建立資料周邊並非一體適用的方法。請根據您的特定安全風險模型和須求，評估和採用在[AWS 上建立周邊白皮書](#)中所述的控制目標。您應仔細考慮您獨特的風險狀態，並實作符合您安全需求的周邊控制。
4. 使用 AWS 服務中的資源共用並適當限制：許多 AWS 服務都可讓您與另一個帳戶共用資源，或鎖定另一個帳戶中的資源，例如 [Amazon Machine Image \(AMI\)](#) 和 [AWS Resource Access Manager \(AWS RAM\)](#)。限制 `ModifyImageAttribute` API 以指定可信帳戶來共用 AMI。當使用 AWS

RAM 來限制僅共用至您的組織時，指定 `ram:RequestedAllowsExternalPrincipals` 條件來協助防止不受信任的身分的存取。如需方案指引和考量，請參閱[資源共用和外部目標](#)。

5. 使用 AWS RAM 在帳戶中或與其他 AWS 帳戶 安全地共用：[AWS RAM](#) 可協助您安全地將您使用帳戶中的角色和使用者所建立的資源與其他 AWS 帳戶 共用。在多帳戶環境中，AWS RAM 可讓您建立資源一次並與其他帳戶共用。這個方法有助於降低營運負擔，同時透過與 Amazon CloudWatch 和 AWS CloudTrail 的整合提供一致性、可見性和可稽核性，這是在使用跨帳戶存取權時所沒有的。

如果您擁有之前使用以資源為基礎的政策共用的資源，可以使用 [PromoteResourceShareCreatedFromPolicy API](#) 或同等項目將資源共用升級到完整 AWS RAM 資源共用。

在某些情況下，您可能需要採取額外步驟來共用資源。例如，若要共用加密快照，您需要[共用 AWS KMS 金鑰](#)。

資源

相關的最佳實務：

- [SEC03-BP07 分析公有和跨帳戶存取權](#)
- [SEC03-BP09 安全地與第三方共用資源](#)
- [SEC05-BP01 建立網路層](#)

相關文件：

- [儲存貯體擁有者將跨帳戶許可授予非其擁有的物件](#)
- [如何使用信任政策搭配 IAM](#)
- [在 AWS 上建置資料周邊](#)
- [向第三方授予對 AWS 資源的存取權限時如何使用外部 ID](#)
- [可與 AWS Organizations 搭配使用的 AWS 服務](#)
- [在 AWS 上建立資料周邊：僅允許可信身分存取公司資料](#)

相關影片：

- [使用 AWS Resource Access Manager 進行精密的存取](#)
- [使用 VPC 端點確保資料周邊的安全](#)
- [在 AWS 上建立資料周邊](#)

相關工具：

- [資料周邊政策範例](#)

SEC03-BP09 安全地與第三方共用資源

您雲端環境的安全並不止於您的組織。您的組織可能仰賴第三方來管理您的部分資料。針對第三方管理的系統的許可管理應該遵循即時存取的做法，採用最低權限的原則搭配暫時憑證。透過與第三方密切合作，您可以同時減少影響範圍以及意外存取的風險。

預期成果：您會避免使用長期 AWS Identity and Access Management (IAM) 憑證，例如存取金鑰和私密金鑰，因為這些金鑰遭到濫用的情況下，會帶來安全風險。您改用 IAM 角色和臨時憑證來改善您的安全狀態，並將管理長期憑證的營運負擔降至最低。授予第三方存取權時，您使用通用唯一識別碼 (UUID) 做為 IAM 信任政策中的外部 ID，並持續掌控與角色連接的 IAM 政策，以確保最低權限存取。如需有關分析外部共用資源的方案指引，請參閱 [SEC03-BP07 分析公開和跨帳戶存取權](#)。

常見的反模式：

- 無條件地使用預設 IAM 信任政策。
- 使用長期 IAM 憑證和存取金鑰。
- 重複使用外部 ID。

未建立此最佳實務時的曝險等級：中

實作指引

您可能會想要允許在 AWS Organizations 之外共用資源或將帳戶存取權授予第三方。例如，第三方可能提供監控解決方案，而該解決方案需要存取您帳戶中的資源。在這些情況下，僅以第三方需要的權限來建立 IAM 跨帳戶角色。此外，請使用[外部 ID 條件](#)定義信任政策。當使用外部 ID 時，您或第三方可以為每個客戶、第三方或租用戶產生唯一 ID。在建立唯一 ID 後，其不應該受除了您之外的任何人控制。第三方必須實作程序，以安全、可稽核且可重新產生的方式將外部 ID 與客戶關聯。

您還可以使用 [IAM Roles Anywhere](#) 為 AWS 之外使用 AWS API 的應用程式管理 IAM 角色。

如果第三方不再需要存取您的環境，請移除該角色。避免為第三方提供長期憑證。持續關注其他支援共用的 AWS 服務，例如，允許與其他 AWS 帳戶 [共用工作負載](#)的 AWS Well-Architected Tool，以及可協助您與其他帳戶安全共用您所擁有 AWS 資源的 [AWS Resource Access Manager](#)。

實作步驟

1. 使用跨帳戶角色提供存取權給外部帳戶。[跨帳戶角色](#)可減少外部帳戶和第三方為了服務客戶所儲存的敏感資訊量。跨帳戶角色允許您在帳戶中將 AWS 資源的存取權安全地授予第三方，例如 AWS 合作夥伴或組織內的其他帳戶，同時維持管理和稽核該存取權的能力。第三方可能從混合式基礎設施為您提供服務，或將資料提取至異地。[IAM Roles Anywhere](#) 可協助您讓第三方工作負載安全地與您的 AWS 工作負載進行互動，並進一步減少使用長期憑證的需要。

您不應該使用與使用者關聯的長期憑證或存取金鑰來提供外部帳戶存取權。反而應該使用跨帳戶角色來提供跨帳戶存取權。

2. 執行盡職調查並確保第三方 SaaS 提供者的安全存取。與第三方 SaaS 提供者共用資源時，請執行徹底的盡職調查，以確保他們採取安全且負責任的方法來存取您的 AWS 資源。評估他們的共同責任模型，以了解他們提供哪些安全措施，以及您的責任範圍。確保 SaaS 提供者採取安全且可稽核的程序來存取您的資源，包括使用[外部 ID](#) 和最低權限存取原則。使用外部 ID 有助於解決[權利義務混淆問題](#)。

實作安全控制，以確保在授予存取權給第三方 SaaS 提供者時，能夠安全存取並遵守最低權限原則。這可能包括使用外部 ID、通用唯一識別碼 (UUID) 和 IAM 信任政策，這些會將存取權限於絕對必要的身分。與 SaaS 提供者密切合作，以建立安全存取機制、定期審查他們對您 AWS 資源的存取，並進行稽核以確保遵循您的安全需求。

3. 棄用客戶提供的長期憑證。棄用長期憑證並使用跨帳戶角色或 IAM Roles Anywhere。如果您必須使用長期憑證，請制定計畫以遷移至角色型存取。如需有關管理金鑰的詳細資訊，請參閱[身分管理](#)。另外也與您的 AWS 帳戶團隊和第三方合作建立風險緩解執行手冊。如需有關回應和緩解潛在安全事件的衝擊的方案指引，請參閱[事件回應](#)。
4. 確認設定具有方案指引且已自動化。外部 ID 不會被視為機密，但外部 ID 不能是容易猜測的值，例如電話號碼、名稱或帳戶 ID。將外部 ID 設為唯讀欄位，而使外部 ID 不能為了冒充設定的目的而遭到變更。

您或第三方可以產生外部 ID。定義程序以決定由誰負責產生 ID。無論建立外部 ID 的實體為何，第三方都要在客戶間一致地強制唯一性和格式。

您的帳戶中為跨帳戶存取權建立的政策必須遵循[最低權限原則](#)。第三方必須提供角色政策文件，或使用 AWS CloudFormation 範本或對您來說同等的自動設定機制。這可減少發生與手動政策建立相關聯之錯誤的機率，並提供可稽核的記錄。如需有關使用 AWS CloudFormation 範本來建立跨帳戶角色的詳細資訊，請參閱[跨帳戶角色](#)。

第三方應該提供自動化、可稽核的設定機制。然而，透過使用概述所需存取權的角色政策文件，您應該可自動設定角色。使用 AWS CloudFormation 範本或同等項目，您應該透過偏移偵測來監控變更，以做為稽核實務的一部份。

5. 將變更列入考量。您的帳戶結構、對第三方的需求或他們提供的服務方案可能發生變更。您應該預期變更和失敗，並透過合適的人員、程序和技术相應進行規劃。定期稽核您提供的存取層級，並實作偵測方法以在發生意外變更時通知您。監控和稽核角色和外部 ID 資料儲存的使用。您應該準備好在發生意外變更或存取模式時撤銷第三方存取權，無論是暫時或永久撤銷。另外，衡量撤銷作業的衝擊，包括執行所花的時間、牽涉的人員、成本，以及對其他資源的衝擊。

如需有關偵測方法的方案指引，請參閱[偵測最佳實務](#)。

資源

相關的最佳實務：

- [SEC02-BP02 使用臨時憑證](#)
- [SEC03-BP05 為您的組織定義許可防護機制](#)
- [SEC03-BP06 根據生命週期管理存取](#)
- [SEC03-BP07 分析公有和跨帳戶存取權](#)
- [SEC04 偵測](#)

相關文件：

- [儲存貯體擁有者將跨帳戶許可授予非其擁有的物件](#)
- [如何搭配 IAM 角色使用信任政策](#)
- [使用 IAM 角色在 AWS 帳戶 間委派存取權](#)
- [如何使用 IAM 存取另一個 AWS 帳戶 中的資源？](#)
- [IAM 中的安全最佳實務](#)
- [跨帳戶政策評估邏輯](#)
- [將 AWS 資源的存取權授予第三方時如何使用外部 ID](#)
- [從在外部帳戶中使用自訂資源建立的 AWS CloudFormation 資源收集資訊](#)
- [安全地使用外部 ID 存取其他人擁有的 AWS 帳戶](#)
- [使用 IAM Roles Anywhere 將 IAM 角色擴展到 IAM 之外的工作負載](#)

相關影片：

- [如何允許不同 AWS 帳戶 中的使用者或角色存取我的 AWS 帳戶？](#)
- [AWS re:Invent 2018：在 60 分鐘內精通 IAM 政策](#)
- [AWS 知識中心直播：IAM 最佳實務和設計決策](#)

相關範例：

- [設定對 Amazon DynamoDB 的跨帳戶存取權](#)
- [AWS STS 網路查詢工具](#)

偵測

問題

- [SEC 4. 如何偵測和調查安全事件？](#)

SEC 4. 如何偵測和調查安全事件？

從日誌和指標中擷取並分析事件，以獲得可見性。請針對安全事件和潛在威脅採取行動，以協助保護工作負載。

最佳實務

- [SEC04-BP01 設定服務和應用程式日誌記錄](#)
- [SEC04-BP02 在標準化的位置擷取日誌、調查結果和指標](#)
- [SEC04-BP03 建立安全提醒的相互關聯並增添其豐富性](#)
- [SEC04-BP04 針對不合規資源實施補救措施](#)

SEC04-BP01 設定服務和應用程式日誌記錄

保留服務和應用程式的安全事件日誌。這是稽核、調查和操作使用案例的基礎原則，以及由管控、風險和合規 (GRC) 標準、政策和程序所推動的常見安全需求。

預期成果：組織應該能夠在需要執行內部程序或義務時 (例如安全事件回應)，以可靠且一致的方式及時從 AWS 服務和應用程式擷取安全事件日誌。考慮集中日誌以達到最佳的營運成果。

常見的反模式：

- 日誌存放太久或太早刪除。
- 每個人都能存取日誌。
- 日誌的管控和使用完全仰賴手動程序。
- 儲存每一種日誌以備不時之需。
- 只在必要時檢查日誌完整性。

建立此最佳實務的優勢：對安全事件和證據來源實作根本原因分析 (RCA) 機制，以履行管控、風險和合規義務。

未建立此最佳實務時的曝險等級：高

實作指引

根據您的需求進行安全調查或其他使用案例期間，您需要能夠審查相關日誌以記錄和了解該事件的全部範圍和時間表。產生提醒也需要日誌，以指出特定關注的動作已發生。選取、開啟、儲存和設定查詢與擷取機制和提醒至關重要。

實作步驟

- 選取並使用日誌來源。在安全調查之前，您需要擷取相關日誌以追溯的方式重新建構 AWS 帳戶 中的活動。選取與您的工作負載相關的日誌來源。

日誌來源選擇條件應該根據您的業務所需的使用案例。使用 AWS CloudTrail 或 AWS Organizations 線索為每個 AWS 帳戶 建立線索，以及為其設定 Amazon S3 儲存貯體。

AWS CloudTrail 是一種日誌記錄服務，會追蹤對 AWS 帳戶 的 API 呼叫以擷取 AWS 服務活動。預設開啟時，此服務會保留 90 天的管理活動，而其能使用 AWS 管理主控台、AWS CLI 或 AWS SDK [透過 CloudTrail 事件歷史記錄擷取](#)。如需較長的保留期間和資料事件的能見度，可[建立 CloudTrail 線索](#)並將其與 Amazon S3 儲存貯體關聯，也可以選擇與 Amazon CloudWatch 日誌群組關聯。或者，您可以建立 [CloudTrail Lake](#)，這會將 CloudTrail 日誌保留長達七年，並提供以 SQL 為基礎的查詢設施

AWS 建議使用 VPC 的客戶分別使用 [VPC 流程日誌](#)和 [Amazon Route 53 解析器查詢日誌](#)來開啟網路流量和 DNS 日誌，並將它們串流處理到 Amazon S3 儲存貯體或 CloudWatch 日誌群組。您可以為 VPC、子網路或網路介面建立 VPC 流程日誌。對於 VPC 流程日誌，您可以選擇何時何地使用流程日誌來降低成本。

AWS CloudTrail 日誌、VPC 流程日誌和 Route 53 解析器查詢日誌是在 AWS 中支援安全調查的基本日誌記錄來源。您還可以使用 [Amazon Security Lake](#) 以 Apache Parquet 格式和 Open

Cybersecurity Schema Framework (OCSF) 收集、正規化並儲存此日誌資料，此種格式隨時可供查詢。Security Lake 還支援其他 AWS 日誌以及來自第三方來源的日誌。

AWS 服務可產生基本日誌來源未擷取的日誌，例如 Elastic Load Balancing 日誌、AWS WAF 日誌、AWS Config 記錄器日誌、Amazon GuardDuty 調查結果、Amazon Elastic Kubernetes Service (Amazon EKS) 稽核日誌和 Amazon EC2 執行個體作業系統及應用程式日誌。如需記錄和監控選項的完整清單，請參閱 [AWS Security Incident Response Guide](#) 的 [附錄 A：雲端功能定義 – 日誌記錄和事件](#)。

- 每個 AWS 服務和應用程式的研究日誌記錄功能：每個 AWS 服務和應用程式都為您提供日誌儲存的選項，而其各自有自己的保留和生命週期功能。兩個最常見的日誌儲存服務是 Amazon Simple Storage Service (Amazon S3) 和 Amazon CloudWatch。如需長期保留，建議使用具成本效益和彈性生命週期功能的 Amazon S3。如果主要日誌記錄選項是 Amazon CloudWatch Logs，則您應該考慮將較不常存取的日誌封存到 Amazon S3，作為一種選項。
- 選取日誌儲存：日誌儲存的選擇通常與您使用的查詢工具、保留功能、熟悉度和成本相關。日誌儲存的主要選項是 Amazon S3 儲存貯體或 CloudWatch 日誌群組。

Amazon S3 儲存貯體提供符合成本效益、耐用的儲存方式，並且具備可選擇的生命週期政策。儲存在 Amazon S3 儲存貯體的日誌可使用 Amazon Athena 之類的服務進行查詢。

CloudWatch 日誌群組透過 CloudWatch Logs Insights 提供耐用的儲存方式和內建查詢設施。

- 識別適當的日誌保留時間：當您使用 Amazon S3 儲存貯體或 CloudWatch 日誌群組來儲存日誌時，您必須為每個日誌來源建立充分的生命週期，以最佳化儲存和擷取成本。客戶一般擁有三個月到一年的時間使日誌隨時可供查詢，並且最長可保留七年。可用性和保留時間的選擇應該配合您的安全需求與各種法令、法規和業務規定。
- 依照適當的保留和生命週期政策對每個 AWS 服務和應用程式使用日誌記錄：對於組織內的每個 AWS 服務或應用程式，尋找特定的日誌記錄組態指引：
 - [設定 AWS CloudTrail 線索](#)
 - [設定 VPC 流程日誌](#)
 - [設定 Amazon GuardDuty 調查結果匯出](#)
 - [設定 AWS Config 記錄](#)
 - [設定 AWS WAF Web ACL 流量](#)
 - [設定 AWS Network Firewall 網路流量日誌](#)
 - [設定 Elastic Load Balancing 存取日誌](#)
 - [設定 Amazon Route 53 解析器查詢日誌](#)
 - [設定 Amazon RDS 日誌](#)

- [設定 Amazon EKS 控制平面日誌](#)
- [為 Amazon EC2 執行個體和內部部署伺服器設定 Amazon CloudWatch 代理程式](#)
- 為日誌選取並實作查詢機制：對於日誌查詢，您可以使用 [CloudWatch Logs Insights](#) (適用於儲存在 CloudWatch 日誌群組中的資料) 以及 [Amazon Athena](#) 和 [Amazon OpenSearch Service](#) (適用於儲存在 Amazon S3 中的資料)。您還可以使用第三方查詢工具，例如安全資訊和事件管理 (SIEM) 服務。

選取日誌查詢工具的過程中應該考慮安全營運的人員、程序和技术層面。選取符合營運、業務和安全需求的工具，並且可供存取和長期維護。請記住，將要掃描的日誌數目維持在日誌查詢工具限制之內，以便以最佳狀態運作。因為成本或技術限制的關係，擁有多個查詢工具十分常見。

例如，您可能使用第三方安全資訊和事件管理 (SIEM) 工具對過去 90 天的資料執行查詢，但由於 SIEM 的日誌擷取成本，而使用 Athena 來執行超過 90 天的查詢。無論實作方式為何，請確認您的方法將所需的工具數量最小化以最大化營運效率，尤其是在安全事件調查期間。

- 使用日誌提供提醒：AWS 透過數種安全服務提供提醒功能：
 - [AWS Config](#) 可監控和記錄 AWS 資源組態，並讓您根據所需的組態自動評估和修復。
 - [Amazon GuardDuty](#) 是威脅偵測服務，可持續監控惡意活動和未經授權的行為，以保護 AWS 帳戶 和工作負載。GuardDuty 會擷取、彙總和分析來自如 AWS CloudTrail 管理和資料事件、DNS 日誌、VPC 流程日誌和 Amazon EKS 稽核日誌等來源的資訊。GuardDuty 會直接從 CloudTrail、VPC 流程日誌、DNS 查詢日誌和 Amazon EKS 提取獨立資料串流。您不需要管理 Amazon S3 儲存貯體或修改您收集和儲存日誌的方式。仍舊建議您保留這些日誌，供自身調查和合規用途。
 - [AWS Security Hub CSPM](#) 提供以單一位置從多個 AWS 服務和選用的第三方產品將安全提醒或調查結果加以彙總、組織和排列優先順序，為您提供安全提醒和合規狀態的全面檢視。

您還可以使用自訂提醒產生引擎，取得這些服務未涵蓋的安全提醒或與您的環境相關的特定提醒。如需有關建置這些提醒和偵測的資訊，請參閱 [《AWS 安全事件回應指南》中的偵測](#)。

資源

相關的最佳實務：

- [SEC04-BP02 在標準化的位置擷取日誌、調查結果和指標](#)
- [SEC07-BP04 定義可擴展的資料生命週期管理](#)
- [SEC10-BP06 預先部署工具](#)

相關文件：

- [AWS 安全事件回應指南](#)
- [Amazon Security Lake 入門](#)
- [入門：Amazon CloudWatch Logs](#)

相關影片：

- [AWS re:Invent 2022 - Amazon Security Lake 簡介](#)

相關範例：

- [Assisted Log Enabler for AWS](#)
- [AWS Security Hub CSPM 調查結果歷史匯出](#)

SEC04-BP02 在標準化的位置擷取日誌、調查結果和指標

安全團隊仰賴日誌和調查結果來分析可能代表未經授權活動或意外變更的事件。為了簡化此分析，您可在標準化的位置擷取安全日誌和調查結果。這樣就能提供關注的資料點來建立相互關聯，並且可簡化工具整合。

預期成果：您採用標準化方法來收集、分析和視覺化日誌資料、調查結果和指標。安全團隊能夠有效率地跨分散的系統建立相互關聯、分析和視覺化安全資料，藉此探索可能發生的安全事件並識別異常。安全資訊和事件管理 (SIEM) 系統或其他機制經整合後，即可查詢和分析日誌資料，以便及時回應、追蹤和向上呈報安全事件。

常見的反模式：

- 多個團隊各自擁有並管理日誌記錄和指標收集工作，而他們的工作方式卻與組織的日誌記錄策略不一致。
- 團隊沒有適當的存取控制可用來限制所收集資料的可見性和更改。
- 團隊未將控管安全日誌、調查結果和指標納入其資料分類政策中。
- 團隊在設定資料收集時，忽略了資料主權和本地化需求。

建立此最佳實務的優勢：擁有標準化的日誌記錄解決方案可用來收集和查詢日誌資料和事件，如此就能改善從內含的資訊中產生的洞察。為收集的日誌資料設定自動化生命週期，可降低日誌儲存所伴隨的成

本。您可以根據團隊所需的資料敏感度和存取模式，為收集的日誌資訊建置精細的存取控制。您可以整合工具來建立資料的相互關聯、視覺化資料，以及從資料中產生洞察。

未建立此最佳實務時的風險暴露等級：中

實作指引

隨著組織內 AWS 的使用增加，分散的工作負載和環境數量也會增加。由於這些工作負載和環境會各自產生其內部活動的相關資料，因此在本地擷取和儲存這些資料會為安全營運方面帶來挑戰。安全團隊會使用安全資訊和事件管理 (SIEM) 系統等工具從分散的來源收集資料，並進行相互關聯、分析和回應工作流程。這需要管理一組複雜的許可來存取各種資料來源，還會在操作擷取、轉換和載入 (ETL) 程序上帶來額外的負擔。

為了克服這些挑戰，請考慮依照[使用多個帳戶整理您的 AWS 環境](#)所述，將所有相關的安全日誌資料來源彙總到日誌封存帳戶中。這包括來自您的工作負載和 AWS 服務所產生日誌的所有安全相關資料，例如 [AWS CloudTrail](#)、[AWS WAF](#)、[Elastic Load Balancing](#) 和 [Amazon Route 53](#)。在標準化的位置透過具有適當跨帳戶許可的個別 AWS 帳戶擷取此資料有幾個好處。此實務有助於防止受害的工作負載和環境內的日誌遭到竄改、可為其他工具提供單一整合點，並提供更簡化的模式來設定資料保留和生命週期。評估資料主權、合規範圍和其他法規的影響，以判斷是否需要多個安全資料儲存位置和保留期。

為了輕鬆擷取和標準化日誌和調查結果，請在您的日誌封存帳戶中評估 [Amazon Security Lake](#)。您可以將 Security Lake 設定為自動從常見的來源擷取資料，例如 CloudTrail、Route 53、[Amazon EKS](#) 和 [VPC 流程日誌](#)。您也可以將 AWS Security Hub CSPM 設定為 Security Lake 中的資料來源，如此就能將其他 AWS 服務 (例如 [Amazon GuardDuty](#) 和 [Amazon Inspector](#)) 的調查結果與您的日誌資料相互關聯。您也可以使用第三方資料來源整合，或設定自訂資料來源。所有整合都會將您的資料標準化為 [Open Cybersecurity Schema Framework \(OCSF\)](#) 格式，並以 Parquet 檔案形式儲存在 [Amazon S3](#) 儲存貯體中，因此不需要進行 ETL 處理。

將安全資料儲存在標準化的位置可提供進階分析功能。AWS 建議您將在 AWS 環境中操作的安全分析工具部署到[安全工具](#)帳戶中，與您的日誌封存帳戶加以區隔。此方法可讓您深入實作控制，以保護日誌和日誌管理程序的完整性和可用性，有別於用於存取日誌的工具。考慮使用 [Amazon Athena](#) 等服務來執行與多個資料來源相互關聯的隨需查詢。您也可以整合視覺化工具，例如 [Quick](#)。採用 AI 技術的解決方案越來越普遍可得，並且能夠執行許多功能，例如將調查結果轉譯成人類可讀的摘要以及自然語言互動等。擁有標準化的資料儲存位置可供查詢，通常會更容易整合這些解決方案。

實作步驟

1. 建立日誌封存和安全工具帳戶

- a. 使用 AWS Organizations，在安全組織單位下[建立日誌封存和安全工具帳戶](#)。如果您使用 AWS Control Tower 來管理組織，則會自動為您建立日誌封存和安全工具帳戶。視需要設定存取和管理這些帳戶的角色與許可。
2. 設定標準化的安全資料位置
 - a. 確定您用來建立標準化安全資料位置的策略。您可以透過通用資料湖架構方法、第三方資料產品或 [Amazon Security Lake](#) 等選項達成此目的。AWS 建議您從為帳戶設為[選擇加入](#)的 AWS 區域擷取安全資料 (即使沒有在作用中)。
3. 設定將資料來源發佈到標準化位置
 - a. 識別安全資料的來源，並將它們設定為發佈到您的標準化位置。評估以所需格式自動匯出資料的選項，而不是需要開發 ETL 程序的選項。透過 Amazon Security Lake，您可以從受支援的 AWS 來源和經過整合的第三方系統[收集資料](#)。
4. 設定工具以存取標準化位置
 - a. 設定 Amazon Athena、Quick 或第三方解決方案等工具，使其具備存取您的標準化位置所需的權限。設定這些工具，以適時透過對日誌封存帳戶的跨帳戶讀取存取權在安全工具帳戶之外操作。[在 Amazon Security Lake 中建立訂閱者](#)，以便為這些工具提供對您資料的存取權。

資源

相關的最佳實務：

- [SEC01-BP01 使用帳戶區隔工作負載](#)
- [SEC07-BP04 定義資料生命週期管理](#)
- [SEC08-BP04 強制存取控制](#)
- [OPS08-BP02 分析工作負載日誌](#)

相關文件：

- [AWS 白皮書：使用多個帳戶整理您的 AWS 環境](#)
- [AWS 方案指引：AWS 安全性參考架構 \(AWS SRA\)](#)
- [AWS 方案指引：應用程式擁有者的日誌記錄和監控指南](#)

相關範例：

- [使用 Amazon Athena 和 Quick 彙總、搜尋和視覺化來自分散來源的日誌資料](#)

- [如何使用 Quick 視覺化 Amazon Security Lake 調查結果](#)
- [使用 Amazon SageMaker AI Studio 和 Amazon Bedrock 為 Amazon Security Lake 產生 AI 技術支援的洞察](#)
- [使用 Amazon SageMaker AI 識別 Amazon Security Lake 資料中的網路安全異常](#)
- [擷取和轉換 Amazon Security Lake 發布的事件並傳遞至 Amazon OpenSearch Service](#)
- [使用 CloudTrail Lake 中的自然語言查詢生成功能來簡化 AWS CloudTrail 日誌分析](#)

相關工具：

- [Amazon Security Lake](#)
- [Amazon Security Lake 合作夥伴整合](#)
- [Open Cybersecurity Schema Framework \(OCSF\)](#)
- [Amazon Athena](#)
- [快速](#)：
- [Amazon Bedrock](#)。

SEC04-BP03 建立安全提醒的相互關聯並增添其豐富性

非預期的活動可能會導致不同來源產生多個安全提醒，因此需要進一步在建立這些來源之間的相互關聯並增添豐富性，才能了解完整的內容。實作安全提醒的自動化相互關聯並增添其豐富性，有助於更準確地識別和回應事件。

預期成果：當活動在您的工作負載和環境內產生不同的提醒時，自動化機制會建立資料的相互關聯，並使用其他資訊增添該資料的豐富性。此預先處理程序呈現了對事件更詳細的了解，進而協助調查人員判斷事件的關鍵性，以及它是否構成需要正式回應的事件。此程序可減輕監控和調查團隊的負擔。

常見的反模式：

- 不同組合的人員對不同系統產生的調查結果和提醒進行調查 (除非是在因職責區分需求而另有規定的情況下)。
- 您的組織將所有安全調查結果和提醒資料收集到標準位置，但要求調查人員手動建立相互關聯和添加資訊。
- 您只仰賴威脅偵測系統的情報來回報調查結果和確定關鍵性。

建立此最佳實務的優勢：自動建立提醒的相互關聯和增添其豐富性，有助於減輕調查人員的整體認知負擔和手動準備資料的負荷。這種做法可縮短判斷事件是否為「事件」及正式回應所需的時間。額外的內容還可協助您準確評估事件的嚴重性，因為實際的嚴重性可能高於或低於任何提醒表明的嚴重性。

未建立此最佳實務時的曝險等級：低

實作指引

安全提醒可能來自 AWS 內多個不同的來源，包括：

- [Amazon GuardDuty](#)、[AWS Security Hub CSPM](#)、[Amazon Macie](#)、[Amazon Inspector](#)、[AWS Config](#)、[AWS Identity and Access Management Access Analyzer](#) 和 [網路存取分析器](#) 等服務
- 來自 AWS 服務、基礎設施和應用程式日誌的自動化分析的提醒，例如來自 [Security Analytics for Amazon OpenSearch Service](#)。
- 回應帳單活動變更的警示，來自下列來源：[Amazon CloudWatch](#)、[Amazon EventBridge](#) 或 [AWS Budgets](#)。
- 第三方來源，例如來自 AWS Partner Network 的威脅情報摘要和 [安全合作夥伴解決方案](#)
- [來自 AWS 信任與安全或其他來源的聯絡資訊](#)，例如客戶或內部員工。
- 使用 [Threat Technique Catalog by AWS\(TTC\)](#)，透過入侵指標 (IoC) 識別法幫助識別和關聯威脅行為者的行為。TTC 是 MITRE ATT&CK 架構的延伸，可針對 AWS 資源的所有已知和觀察到的威脅行為和技術進行分類。

提醒基本上包含有關誰 (主體或身分)、做了什麼 (採取的動作)，以及對象是誰 (受影響的資源) 的資訊。對於其中每個來源，請確定是否有能夠在這些身分、動作和資源的識別符之間建立映射的方式，以作為建立相互關聯的基礎。可能的形式包括整合提醒來源與安全資訊和事件管理 (SIEM) 工具，以便為您自動建立相互關聯、建置您自己的資料管道和處理流程，或結合上述兩者。

可為您建立相互關聯的服務範例為 [Amazon Detective](#)。偵測會持續接收來自各種 AWS 和第三方來源的提醒，並使用不同形式的情報來構成視覺化圖形以呈現其關係，進而協助調查。

雖然提醒的初始關鍵性可協助判斷優先順序，但提醒發生的環境則決定了其真實的關鍵性。舉例來說，[Amazon GuardDuty](#) 可以針對您工作負載內的 Amazon EC2 執行個體正在查詢非預期的網域名稱而發出警示。GuardDuty 可能會自行對此提醒指派「低關鍵性」標籤。然而，如果在提醒發出之時與其他活動自動建立相互關聯，您就可能會發現有數百個 EC2 執行個體是由相同身分部署的，這種情況會增加整體營運成本。在此事件中，此相互關聯事件內容會發布新的安全提醒，且關鍵性可能會調整為高，進而加速採取進一步行動。

實作步驟

1. 識別安全提醒資訊的來源。了解來自這些系統的提醒如何表示身分、動作和資源，以確定可能的相互關聯。
2. 建立一個機制來擷取不同來源的提醒。考慮用於這類用途的服務，例如 Security Hub CSPM、EventBridge 和 CloudWatch。
3. 識別資料相互關聯和增添豐富性的來源。範例來源包括 [AWS CloudTrail](#)、[VPC 流程日誌](#)、[Route 53 Resolver 日誌](#)，以及基礎設施和應用程式日誌。這些日誌的任一種或全部都可能透過與 [Amazon Security Lake](#) 的單一整合來取用。
4. 將提醒與資料相互關聯和增添豐富性的來源整合在一起，以建立更詳細的安全事件內容並構成關鍵性。
 - a. Amazon Detective、SIEM 工具或其他第三方解決方案可以自動執行特定層級的擷取、相互關聯和豐富性。
 - b. 您也可以使用 AWS 服務來自行建置。例如，您可以調用 AWS Lambda 函數來對 AWS CloudTrail 或 Amazon Security Lake 執行 Amazon Athena 查詢，並將結果發佈至 EventBridge。

資源

相關的最佳實務：

- [SEC10-BP03 準備鑑識功能](#)
- [OPS08-BP04 建立可付諸行動的提醒](#)
- [REL06-BP03 傳送通知 \(即時處理和警示\)](#)

相關文件：

- [AWS 安全事件回應指南](#)

相關範例：

- [如何使用帳戶中繼資料增添 AWS Security Hub CSPM 調查結果的豐富性](#)

相關工具：

- [Amazon Detective](#)

- [Amazon EventBridge](#) ()
- [AWS Lambda](#)
- [Amazon Athena](#)

SEC04-BP04 針對不合規資源實施補救措施

您的偵測控制可能針對不符合您組態需求的資源發出提醒。您可以手動或自動實施以程式設計方式定義的補救措施，以修正這些資源並協助盡可能將影響降到最低。以程式設計方式定義補救措施時，您可以採取快速且一致的行動。

雖然自動化可以增強安全操作，但您應謹慎實作和管理自動化程序。設置適當的監督和控制機制，以確認自動化回應是否有效、準確，且合乎組織政策和風險偏好。

預期成果：您會定義資源組態標準，以及在偵測到資源不合規時的修復步驟。在可能的情況下，您已透過程式設計方式定義補救措施，以供人員手動或透過自動化方式啟動。已設置偵測系統，其可識別不合規的資源，並在由安全人員監控的集中式工具中發佈提醒。這些工具支援手動或自動執行程式設計的補救措施。自動化補救措施設有適當的監督和控制機制來管理其使用。

常見的反模式：

- 您實作自動化，但未徹底測試和驗證補救動作。這可能會導致意外的後果，例如中斷正當的業務營運或導致系統不穩定。
- 您透過自動化改善回應時間和程序，但未設置適當的監控與機制，無法在需要時允許人為介入和判斷。
- 您只仰賴補救措施，而不是將補救措施視為更廣泛的事件回應和復原計畫的一部分。

建立此最佳實務的優勢：自動化補救措施能夠比手動流程更快回應組態錯誤，進而有助於將可能對業務造成的影響降至最低，並且減少意外使用的機會。當您以程式設計方式定義補救措施時，就能一致套用這些措施，進而降低人為錯誤的風險。自動化還可同時處理更大量的提醒，這點對於大規模操作的環境來說尤其重要。

未建立此最佳實務時的曝險等級：中

實作指引

如 [SEC01-BP03 識別和驗證控制目標](#) 中所述，[AWS Config](#) 和 [AWS Security Hub CSPM](#) 等服務可協助您監控帳戶中資源的組態，以確保符合您的需求。偵測到不合規的資源時，AWS Security Hub CSPM 這類服務可協助適當路由警示和修復。這些解決方案為您的安全調查人員提供了一個集中的位置，方便監控問題並採取修正動作。

除了 AWS Security Hub CSPM 之外，AWS 還推出了 [Security Hub Advanced](#)。此服務在 re:Invent 2025 上宣布，轉變組織最關鍵安全問題的優先順序，並大規模回應以保護其雲端環境。增強型 Security Hub 現在使用進階分析技術，在您的雲端環境中自動關聯、增強和排定安全訊號的優先順序。Security Hub 與 [Amazon GuardDuty](#)、[Amazon Inspector](#)、[Amazon Macie](#) 和 [AWS Security Hub CSPM](#) 完全整合。Security Hub 中的相關調查結果可能會導致新的調查結果，稱為暴露調查結果，其中包括根據每個資源中找到的漏洞所假設的攻擊路徑。

有些不合規資源的情況獨特，需要人為判斷來進行修復，有些情況則有標準回應，您可透過程式設計方式定義這類回應。例如，對於設定錯誤的 VPC 安全群組，其標準回應可能是移除不允許的規則並通知擁有者。您可以在 [AWS Lambda](#) 函數中、[AWS Systems Manager Automation](#) 文件中，或透過您慣用的其他程式碼環境來定義回應。確定環境能夠使用具有採取修正動作所需之最低許可權的 IAM 角色來對 AWS 進行身分驗證。

定義所需的補救措施後，您就可以決定您偏好的補救措施實施方法。AWS Config 可以為您[實施補救措施](#)。如果您使用 Security Hub CSPM，則可透過[自訂動作](#)來執行此操作，該動作會將調查結果資訊發佈至 [Amazon EventBridge](#)。然後，EventBridge 規則就能實施補救措施。您可以透過 Security Hub CSPM 設定自動或手動執行補救措施。

對於程式化的補救措施，建議您留存所執行動作的完整日誌和稽核，以及其結果。審核並分析這些日誌，以評估自動化流程的有效性，並識別改進之處。擷取 [Amazon CloudWatch Logs](#) 中的日誌，以及 Security Hub CSPM 中作為[調查結果備註](#)的修復結果。

一開始，請考慮 [AWS 上的自動化安全性回應](#)，其中包含預先建置的補救措施，可用來解決常見的安全錯誤組態。

實作步驟

1. 分析提醒並排定優先順序。
 - a. 將來自各種不同 AWS 服務的安全提醒合併到 Security Hub CSPM 中，以提供集中查看、排定優先順序和修復的方式。
2. 制定補救措施。
 - a. 使用 Systems Manager 和 AWS Lambda 等服務來執行程式化的補救措施。
3. 設定實施補救措施的方式。
 - a. 使用 Systems Manager 定義將調查結果發佈到 EventBridge 的自訂動作。設定手動或自動啟動這些動作。
 - b. 您也可以使用 [Amazon Simple Notification Service \(SNS\)](#) 傳送通知和提醒給相關的利害關係人 (例如，安全團隊或事件回應團隊)，以便在必要時進行人為介入或向上呈報。

4. 審核並分析補救措施日誌，以了解有效性和改進之處。
 - a. 將日誌輸出傳送至 CloudWatch Logs。擷取 Security Hub CSPM 中作為調查結果備註的結果。

資源

相關的最佳實務：

- [SEC06-BP03 減少手動管理和互動式存取](#)

相關文件：

- [AWS 安全事件回應指南 - 偵測](#)

相關範例：

- [上的自動化安全回應AWS](#)
- [使用 AWS Config 監控 EC2 執行個體金鑰對](#)
- [使用 AWS CloudFormation Guard 政策建立 AWS Config 自訂規則](#)
- [自動修復未加密的 Amazon RDS 資料庫執行個體和叢集](#)

相關工具：

- [AWS Systems Manager Automation](#)
- [上的自動化安全回應AWS](#)

基礎設施保護

問題

- [SEC 5. 如何保護網路資源？](#)
- [SEC 6. 如何保護運算資源？](#)

SEC 5. 如何保護網路資源？

任何具有某種網路連線能力的工作負載，無論是網際網路或私有網路，都需要多層防禦來協助保護其不受外部和內部網路威脅的影響。

最佳實務

- [SEC05-BP01 建立網路層](#)
- [SEC05-BP02 控制網路層內的流量流程](#)
- [SEC05-BP03 實作以檢查為基礎的保護](#)
- [SEC05-BP04 自動化網路保護](#)

SEC05-BP01 建立網路層

以工作負載元件的邏輯分組為基礎，根據其資料敏感性和存取需求將您的網路拓模區分成不同層。將需要從網際網路進行傳入存取的元件 (例如公用 Web 端點) 和只需要內部存取的元件 (例如資料庫) 加以區分。

預期成果：您的網路層是整體安全深度防禦方法的一部分，可與工作負載的身分驗證和授權策略相輔相成。根據資料敏感性和存取需求妥善區分的網路層，且具有適當的流量流程和控制機制。

常見的反模式：

- 您在單一 VPC 或子網路中建立所有資源。
- 您建構網路層時未考量資料敏感性需求、元件行為或功能。
- 您將 VPC 和子網路作為所有網路層考量的預設值，而且未考慮 AWS 受管服務對拓模的影響。

建立此最佳實務的優勢：建立網路層是透過網路限制非必要路徑的第一步，尤其是前往關鍵系統和資料的路徑。這可讓未經授權的行為者更難存取您的網路並瀏覽至網路內的其他資源。分散的網路層有利於縮小檢測系統的分析範圍，例如針對入侵偵測或防範惡意軟體。這樣可減少誤報和不必要的處理負擔。

未建立此最佳實務時的曝險等級：高

實作指引

在設計工作負載架構時，根據元件的責任將其劃分至不同層是常見的做法。例如，Web 應用程式可擁有呈現層、應用程式層和資料層。您可以在設計網路拓模時採用類似的方法。基礎網路控制可協助強制執行工作負載的資料存取需求。例如，在擁有三層的 Web 應用程式架構中，您可以將靜態呈現層檔案儲存在 [Amazon S3](#) 上，並從內容交付網路 (CDN) (例如 [Amazon CloudFront](#)) 提供它們。應用程式層可以包含 [Application Load Balancer \(ALB\)](#) 在 [Amazon VPC](#) 公有子網路 (類似非軍事區即 DMZ) 中提供的公有端點，且加上部署到私有子網路中的後端服務。託管資料庫和共用檔案系統等資源的資料層，可位於與應用程式層的資源所在位置不同的私有子網路。在這些層的邊界 (CDN、公有子網路、私有子網路)，您可以部署控制，藉此僅允許授權的流量跨越這些邊界。

類似於根據工作負載元件的功能用途建立網路層模型，請一併考量要處理的資料敏感性。以 Web 應用程式為例，雖然您所有的工作負載服務可能都位於應用程式層內，但不同的服務可能會處理不同敏感程度的資料。在這種情況下，根據您的控制需求，針對每一種程度的資料敏感性使用多個私有子網路、相同 AWS 帳戶 中的不同 VPC，甚至是不同 AWS 帳戶 中的不同 VPC 來區分應用程式層較為適當。

對網路層的進一步考量是工作負載元件的行為一致性。繼續以上述範例說明，在應用程式層中，您可能接受來自最終使用者或外部系統整合輸入的服務，而導向這些服務的輸入在本質上比對其他服務的輸入帶有更高風險。範例包括檔案上傳、要執行的程式碼指令碼、電子郵件掃描等。將這些服務放置在其自己的網路層中，有助於在其周圍建立更強大的隔離邊界，並可防止其獨特行為造成檢測系統中的誤報提醒。

在設計過程中，請考慮使用 AWS 受管服務對網路拓樸的影響。探索像是 [Amazon VPC Lattice](#) 等服務如何讓您更輕鬆地跨網路層實現工作負載元件的互通性。使用 [AWS Lambda](#) 時，請在您的 VPC 子網路中部署，除非受到特定原因限制而無法這樣做。確定 VPC 端點和 [AWS PrivateLink](#) 可在哪些地方簡化遵循限制網際網路閘道存取的安全政策。

實作步驟

1. 審核您的工作負載架構。根據元件和服務提供的功能、處理的資料敏感性以及其行為，將其邏輯分組。
2. 對於回應網際網路請求的元件，請考慮使用負載平衡器或其他代理來提供公有端點。探索如何使用 CloudFront、[Amazon API Gateway](#)、Elastic Load Balancing 和 [AWS Amplify](#) 等受管服務轉移安全控制，以託管公有端點。
3. 對於在運算環境中執行的元件 (例如 Amazon EC2 執行個體、[AWS Fargate](#) 容器或 Lambda 函數)，一開始即根據您的群組將這些元件部署到私有子網路中。
4. 對於全受管 AWS 服務，例如 [Amazon DynamoDB](#)、[Amazon Kinesis](#) 或 [Amazon SQS](#)，請考慮使用 VPC 端點作為透過私有 IP 位址存取的預設方式。

資源

相關的最佳實務：

- [REL02 規劃您的網路拓撲](#)
- [PERF04-BP01 了解聯網如何影響效能](#)

相關影片：

- [AWS re:Invent 2023 - AWS 聯網基礎](#)

相關範例：

- [VPC 範例](#)
- [透過使用 AWS Fargate、AWS PrivateLink 和 Network Load Balancer，在 Amazon ECS 上私有存取容器應用程式](#)
- [使用 Amazon CloudFront 在 Amazon S3 儲存貯體中透過 VPC 提供靜態內容](#)

SEC05-BP02 控制網路層內的流量流程

在網路層內，使用進一步的分隔方式，將流量限於每個工作負載所需的流程。首先，專注於控制網際網路或其他外部系統到工作負載與您的環境之間的流量 (南北流量)。接著查看不同元件和系統之間的流量 (東西流量)。

預期成果：您只允許工作負載的元件所需的網路流程互相通訊，以及與其用戶端和其相依的任何其他服務進行通訊。您的設計考量到公有與私有輸入和輸出之間的比較、資料分類、區域法規以及協定需求等因素。在設計最低權限原則的過程中，盡可能採用點對點流程，而非網路對等。

常見的反模式：

- 您採用以周邊為基礎的網路安全方法，僅控制網路層邊界處的流量流程。
- 您假設網路層內的所有流量都經過驗證和授權。
- 您只對輸入流量或輸出流量實施控制，而不是對兩者都實施。
- 您只仰賴工作負載元件和網路控制項來驗證和授權流量。

建立此最佳實務的優勢：此實務有助於降低網路內發生未經授權行動的風險，並且為您的工作負載增加一層額外的授權。透過執行流量流程控制，您就可以限制安全事件的影響範圍，並加快偵測和回應速度。

未建立此最佳實務時的曝險等級：高

實作指引

雖然網路層有助於在具有類似功能、資料敏感性層級和行為的工作負載元件周圍建立邊界，但您可以利用一些技術進一步區隔這些層內的元件，藉此建立遵循最低權限原則且更精細的流量控制層級。在 AWS 內，網路層主要是根據 Amazon VPC 內的 IP 位址範圍使用子網路定義。您也可以使用不同的 VPC 定義網路層，其用途包括根據業務領域將微型服務環境分組等。使用多個 VPC 時，使用 [AWS Transit Gateway](#) 調解路由。雖然這可在第 4 層 (IP 位址和連接埠範圍) 使用安全群組和路由表提供流

量控制，但您可以使用 [AWS PrivateLink](#)、[Amazon Route 53 Resolver DNS 防火牆](#)、[AWS Network Firewall](#) 和 [AWS WAF](#) 等其他服務進一步控制流量。

針對連線啟動方、連接埠、協定和網路層等方面，了解並清查工作負載的資料流程和通訊需求。評估可用於建立連線和傳輸資料的協定，以選取符合您的防護需求的協定 (例如 HTTPS 而非 HTTP)。在網路邊界和每一層內擷取這些需求。確定這些需求後，探索僅允許必要的流量流經每個連線點的選項。一開始在 VPC 內使用安全群組會是合適的選擇，因為安全群組可以連接至使用彈性網路介面 (ENI) 的資源，例如 Amazon EC2 執行個體、Amazon ECS 任務、Amazon EKS Pod 或 Amazon RDS 資料庫。與第 4 層防火牆不同的是，安全群組可以設置一項規則來依識別碼允許來自另一個安全群組的流量，藉此盡量減少群組內的資源隨時間改變所需的更新。您也可以使用安全群組的傳入和傳出規則來篩選流量。

當流量在 VPC 之間移動時，針對簡便路由使用 VPC 對等或針對複雜路由使用 AWS Transit Gateway 的情況很常見。使用這些方法可讓來源和目的地網路的 IP 位址範圍之間的流量更順暢。然而，如果您的工作負載只需要讓流量在不同 VPC 中的特定元件之間流動，則可考慮使用 [AWS PrivateLink](#) 的點對點連線。若要這樣做，請確定哪些服務應作為生產者，哪些服務應作為取用者。為生產者部署相容的負載平衡器，對應地開啟 PrivateLink，然後接受取用者的連線請求。接著從取用者的 VPC 指派私有 IP 位址給生產者服務，取用者可使用該位址提出後續請求。這種方法減少了對等網路的需求。在評估 PrivateLink 的過程中，請納入資料處理和負載平衡的成本。

安全群組和 PrivateLink 有助於控制工作負載的元件之間的流程，而另一項重要的考量則是如何控制允許您的資源存取哪些 DNS 網域 (如有的話)。根據您 VPC 的 DHCP 組態而定，您可以考慮兩種不同的 AWS 服務來達成此目的。大多數客戶會使用在 CIDR 範圍的 +2 位址供 VPC 使用的預設 Route 53 Resolver DNS 服務 (也稱為 Amazon DNS 伺服器或 AmazonProvidedDNS)。使用此方法，您可以建立 DNS 防火牆規則，並將它們與 VPC 關聯，以確定要對您提供的網域清單執行哪些動作。

如果您不使用 Route 53 Resolver，或是想要用網域篩選以外更深入的檢測和流程控制功能來輔助 Resolver，請考慮部署 AWS Network Firewall。此服務會使用無狀態或有狀態規則來檢測個別封包，以確定是否拒絕或允許流量。您可以採用類似的方法，使用 AWS WAF 篩選前往公有端點的傳入 Web 流量。如需有關這些服務的進一步指引，請參閱 [SEC05-BP03 實作檢測型防護措施](#)。

實作步驟

1. 識別工作負載元件之間的必要資料流程。
2. 藉由對傳入和傳出流量採用深度防禦方法，套用多項控制，包括使用安全群組和路由表。
3. 使用防火牆對 VPC 上輸入、輸出和 VPC 之間的網路流量定義精細的控制，例如 Route 53 Resolver DNS 防火牆、AWS Network Firewall 和 AWS WAF。考慮使用 [AWS Firewall Manager](#) 集中設定和管理整個組織的防火牆規則。

資源

相關的最佳實務：

- [REL03-BP01 選擇如何劃分工作負載](#)
- [SEC09-BP02 強制傳輸中加密](#)

相關文件：

- [VPC 的安全最佳實務](#)
- [AWS 網路最佳化秘訣](#)
- [AWS 上的網路安全指引](#)
- [在 AWS 雲端 中保護您 VPC 的傳出網路流量](#)

相關工具：

- [AWS Firewall Manager](#)

相關影片：

- [適用於許多 VPC 的 AWS Transit Gateway 參考架構](#)
- [使用 Amazon CloudFront、AWS WAF 和 AWS Shield 的應用程式加速和保護](#)
- [AWS re:Inforce 2023：防火牆和設置位置](#)

SEC05-BP03 實作以檢查為基礎的保護

在網路層之間設定流量檢測點，以確保傳輸中的資料符合預期的類別和模式。分析流量流程、中繼資料和模式，以協助更有效地識別、偵測及回應事件。

預期成果：在網路層之間穿梭的流量會經過檢測和授權。允許和拒絕決策取決於明確的規則、威脅情報以及偏離基準行為的程度。流量越接近敏感資料，防護就會越嚴格。

常見的反模式：

- 僅仰賴以連接埠和通訊協定為準的防火牆規則。未善加利用情報系統。
- 根據可能隨時變更的特定目前威脅模式制訂防火牆規則。
- 僅檢測從私有子網路傳輸到公有子網路的流量，或從公有子網路傳輸到網際網路的流量。

- 沒有網路流量基準點可比較，以識別行為異常。

建立此最佳實務的優勢：檢測系統可讓您制訂智慧型規則，例如，只有在流量資料內存在特定條件時，才允許或拒絕流量。根據最新的威脅情報，隨著威脅態勢隨著時間的變化，從 AWS 和 合作夥伴的受管規則集中受益。這可減輕維護規則和研究入侵指標的工作負擔，進而降低誤報的可能性。

未建立此最佳實務時的曝險等級：中

實作指引

使用 AWS Network Firewall或其他[防火牆](#)和[入侵預防系統](#)（IPS）精細控制具狀態和無狀態網路流量 AWS Marketplace，您可以在 [Gateway Load Balancer（GWLB）](#) 後方部署。AWS Network Firewall 支援 [Suricata 相容的](#)開放原始碼IPS規格，以協助保護您的工作負載。

AWS Network Firewall 和廠商解決方案都支援GWLB不同的內嵌檢查部署模型。例如，您可以根據基準VPC執行檢查、集中在檢查中VPC，或在混合模型中部署，其中東西流量會流經檢查，VPC而網際網路傳入會按檢查VPC。另一個考量是解決方案是否支援取消包裝 Transport Layer Security（TLS），可針對任一方向啟動的流量進行深度封包檢查。如需這些組態的相關資訊和深入資訊，請參閱 [AWS Network Firewall 最佳實務指南](#)。

如果您使用的是執行 out-of-band檢查的解決方案，例如從以半透明模式操作的網路介面進行封包資料的 pcap 分析，則可以設定[VPC流量鏡像](#)。鏡像流量會計入介面的可用頻寬，並且您需支付與非鏡像流量相同的資料傳輸費用。您可以查看這些設備的虛擬版本是否可在上使用[AWS Marketplace](#)，這可能支援在後方的內嵌部署GWLB。

對於透過 HTTP型通訊協定進行交易的元件，請使用 Web 應用程式防火牆（WAF）保護您的應用程式免受常見威脅。[AWS WAF](#) 是一種 Web 應用程式防火牆，可讓您在傳送至 Amazon API Gateway、Amazon CloudFront AWS AppSync 或 Application Load Balancer 之前，監控並封鎖符合您可設定規則的 HTTP（S）請求。當您評估 Web 應用程式防火牆的部署時，請考慮進行深度封包檢查，因為有些會要求您在流量檢查TLS之前終止。若要開始使用 AWS WAF，您可以[AWS 受管規則](#)搭配自己的使用，或使用現有的[合作夥伴整合](#)。

您可以使用 集中管理組織中的 AWS Network Firewall AWS WAF AWS Shield Advanced、和 Amazon VPC安全群組[AWS Firewall Manager](#)。

實作步驟

1. 判斷您是否可以廣泛地範圍檢查規則，例如透過檢查 VPC，或者是否需要每個VPC方法更精細。
2. 對於內嵌檢測解決方案：

- a. 如果使用 AWS Network Firewall，請建立規則、防火牆政策和防火牆本身。上述這些設定完成後，您可以[將流量路由至防火牆端點](#)以啟用檢測。
 - b. 如果搭配 Gateway Load Balancer（GWLB）使用第三方設備，請在一或多個可用區域中部署和設定您的設備。然後，建立您的 GWLB、端點服務、端點，並設定流量的路由。
3. 對於 out-of-band 檢查解決方案：
1. 在應鏡像傳入和傳出流量的介面上開啟 VPC 流量鏡像。您可以使用 Amazon EventBridge 規則來叫用 AWS Lambda 函數，以在建立新資源時開啟介面上的流量鏡像。將流量鏡像工作階段指向處理流量的設備前方的 Network Load Balancer。
4. 對於傳入 Web 流量解決方案：
- a. 若要設定 AWS WAF，請先設定 Web 存取控制清單（web ACL）。Web ACL 是具有序列處理預設動作（ALLOW 或 DENY）的規則集合，可定義您的 WAF 如何處理流量。您可以在 Web 中建立自己的規則和群組，或使用 AWS 受管規則群組 ACL。
 - b. 設定 Web ACL 後，請將 Web ACL 與 AWS 資源（例如 Application Load Balancer、API Gateway REST API 或 CloudFront 分佈）建立關聯，以開始保護 Web 流量。

資源

相關文件：

- [什麼是流量鏡像？](#)
- [使用第三方安全設備實作內嵌流量檢測](#)
- [AWS Network Firewall 具有路由的範例架構](#)
- [使用 AWS Gateway Load Balancer 和 的集中式檢查架構 AWS Transit Gateway](#)

相關範例：

- [部署 Gateway Load Balancer 的最佳實務](#)
- [TLS 加密輸出流量和 的檢查組態 AWS Network Firewall](#)

相關工具：

- [AWS Marketplace IDS/IPS](#)

SEC05-BP04 自動化網路保護

使用 DevOps 實務自動化網路保護的部署，例如基礎設施，例如程式碼（IaC 和 CI/CD 管道。這些實務可協助您透過版本控制系統追蹤網路防護措施中的變更、縮短部署變更所需的時間，並協助您偵測網路防護措施是否偏離所需的組態。

預期成果：您會使用範本定義網路防護措施，並將其遞交至版本控制系統中。當進行新的變更時，自動化管道會因此而啟動，以協調其測試和部署。設置了政策檢查和其他靜態測試，可在部署前驗證變更。您可以將變更部署到模擬環境中，以驗證控制是否如預期運作。控制得到核准後，也會自動將其部署到實際執行環境中。

常見的反模式：

- 仰賴個別工作負載團隊各自定義自己的整套網路堆疊、防護措施和自動化程序。未集中發佈網路堆疊和防護措施的標準層面，供工作負載團隊取用。
- 仰賴中央網路團隊來定義網路、防護措施和自動化的所有層面。未將網路堆疊和防護措施的工作負載特定層面委派給該工作負載的團隊。
- 集中和委派的情況在網路團隊與工作負載團隊之間達到適當的平衡，但未對 IaC 範本和 CI/CD 管道整體實施一致的測試和部署標準。未在工具中擷取所需的組態，以致無法檢查範本是否遵循規範。

建立此最佳實務的優勢：使用範本定義網路防護措施，可讓您使用版本控制系統追蹤和比較一段時間的變化。使用自動化方式測試和部署變更可建立標準化和可預測性，提高成功部署的機會，並減少重複的手動組態。

未建立此最佳實務時的曝險等級：中

實作指引

[SEC05-BP02 控制網路層內的流量流](#)和 [SEC05-BP03 實作檢查型保護](#) 中所述的一些網路保護控制項，隨附可依據最新威脅情報自動更新的受管規則系統。保護 Web 端點的範例包括 [AWS WAF 受管規則](#) 和 [AWS Shield Advanced 自動應用程式層 DDoS 緩解](#)。使用 [AWS Network Firewall 受管規則群組](#) 隨時掌握有關信譽不良網域清單和威脅特徵的最新資訊。

除了受管規則之外，我們建議您使用 DevOps 實務來自動部署網路資源、保護和您指定的規則。您可以在 [AWS CloudFormation](#) 或您選擇的其他基礎設施即程式碼 (IaC) 工具中擷取這些定義，將它們遞交至版本控制系統，並使用 CI/CD 管道部署它們。使用此方法取得 DevOps 管理網路控制項的傳統優點，例如更可預測的版本、使用等工具進行自動測試 [AWS CloudFormation Guard](#)，以及偵測部署環境與所需組態之間的偏離。

根據您在 [SEC05-BP01 建立網路層](#) 中所做的決策，您可能需要建立專用VPCs於輸入、輸出和檢查流程的中央管理方法。如[AWS 安全參考架構 \(AWS SRA\)](#) 中所述，您可以在VPCs專用[網路基礎設施帳戶](#)中定義這些屬性。您可以使用類似的技術，集中定義其他帳戶中工作負載VPCs使用的、其安全群組、AWS Network Firewall 部署、Route 53 Resolver 規則和DNS防火牆組態，以及其他網路資源。您可以透過 [AWS Resource Access Manager](#) 將這些資源與其他帳戶共用。使用此方法，您可以簡化對網路控制的自動測試並將該控制部署到網路帳戶的程序，同時只需管理一個目的地。您可以在混合模式中，透過集中部署和共用特定控制，並將其他控制委派給個別工作負載團隊及其各自的帳戶，來執行上述操作。

實作步驟

1. 建立擁有權來規範要集中定義網路和防護措施的哪些方面，以及您的工作負載團隊可以維護哪些方面。
2. 建立環境來測試變更，並將變更部署至您的網路及其防護措施。例如，使用網路測試帳戶和網路實際執行帳戶。
3. 確定如何在版本控制系統中儲存和維護範本。將儲存中央範本的儲存庫與工作負載儲存庫分開，而工作負載範本可以儲存在專屬於該工作負載的儲存庫中。
4. 建立 CI/CD 管道以測試和部署範本。定義測試以檢查是否有組態錯誤，以及範本是否符合您公司的標準。

資源

相關的最佳實務：

- [SEC01-BP06 自動化標準安全控制的部署](#)

相關文件：

- [AWS Security Reference Architecture - 網路帳戶](#)

相關範例：

- [AWS 部署管道參考架構](#)
- [NetDevSecOps 將 AWS 網路部署現代化](#)
- [將 AWS CloudFormation 安全測試與 AWS Security Hub CSPM 和 AWS CodeBuild 報告整合](#)

相關工具：

- [AWS CloudFormation](#)
- [AWS CloudFormation Guard](#)
- [cfn_nag](#)

SEC 6. 如何保護運算資源？

工作負載中的運算資源需要多層防禦，以協助防範外部和內部威脅。運算資源包括 EC2 執行個體、容器、AWS Lambda 函數、資料庫服務、IoT 裝置等。

最佳實務

- [SEC06-BP01 執行漏洞管理](#)
- [SEC06-BP02 從強化的影像佈建運算](#)
- [SEC06-BP03 減少手動管理和互動式存取](#)
- [SEC06-BP04 驗證軟體完整性](#)
- [SEC06-BP05 自動化運算保護](#)

SEC06-BP01 執行漏洞管理

經常掃描和修補程式碼、相依性和基礎設施中的漏洞，以協助防禦新的威脅。

預期成果：您擁有解決方案，可持續掃描工作負載以找出軟體漏洞、潛在缺陷和意料之外的網路暴露情形。您已建立流程和程序，可根據風險評估標準來識別、排定優先順序和修復這些漏洞。此外，您已為運算執行個體實作自動修補管理。您的漏洞管理計畫已整合至您的軟體開發生命週期，並且備有解決方案可在 CI/CD 管道中掃描原始程式碼。

常見的反模式：

- 沒有漏洞管理計畫。
- 執行系統修補而不考慮嚴重性或避免風險。
- 使用已過廠商提供的結束生命週期日期的軟體。
- 在分析程式碼的安全問題之前將其部署至生產環境。

未建立此最佳實務時的曝險等級：高

實作指引

漏洞管理是維護安全可靠的雲端環境的關鍵層面。它是一整套全方位的程序，包括安全性掃描、識別問題並排定優先順序，以及修補程式操作，以解決找到的漏洞。自動化在此過程中扮演關鍵角色，因為它有助於持續掃描工作負載，以找出潛在問題和意料之外的網路暴露情形，以及提供補救措施。

[AWS 共同責任模型](#)是支援漏洞管理的基本概念。根據此模型，AWS 負責保護基本的基礎設施，包括執行 AWS 服務的硬體、軟體、網路和設備。另一方面，您負責保護與 Amazon EC2 執行個體和 Amazon S3 物件等服務相關聯的資料、安全組態和管理任務。

AWS 提供一系列服務來支援漏洞管理計畫。[Amazon Inspector](#) 會持續掃描 AWS 工作負載，以找出軟體漏洞和意料之外的網路存取，而 [AWS Systems Manager Patch Manager](#) 會協助管理所有 Amazon EC2 執行個體的修補工作。這些服務可與 [AWS Security Hub CSPM](#) 整合，這個雲端安全狀態管理服務可自動化 AWS 安全檢查、將安全提醒集中在一起，並提供全方位的視角來了解組織的安全狀態。此外，[Amazon CodeGuru 安全工具](#) 會在開發階段使用靜態程式碼分析，以識別 Java 和 Python 應用程式中的潛在問題。

透過將漏洞管理實務納入軟體開發生命週期，您就可以在漏洞進入實際執行環境之前主動解決漏洞，從而降低安全事件的風險，並將漏洞可能造成的影響降至最低。

實作步驟

1. 了解共同責任模型：檢閱 AWS 共同責任模型，以了解您在雲端保護工作負載和資料的責任。AWS 負責保護基礎雲端基礎設施，而您負責保護您使用的應用程式、資料和服務。
2. 實作漏洞掃描：設定漏洞掃描服務 (例如 Amazon Inspector) 以自動掃描您的運算執行個體 (例如，虛擬機器、容器或無伺服器函數)，藉此找出軟體漏洞、潛在缺陷和意料之外的網路暴露情形。
3. 建立漏洞管理程序：定義流程和程序來識別漏洞、排定優先順序和修復漏洞。這可能包括設定定期漏洞掃描排程、建立風險評估條件，以及根據漏洞嚴重性定義修復時間表。
4. 設定修補程式管理：使用修補程式管理服務來自動化修補運算執行個體的程序，包括作業系統和應用程式。您可以設定服務來掃描執行個體以找出缺少的修補程式，並依照排程自動進行安裝。考慮使用 AWS Systems Manager Patch Manager 來提供此功能。
5. 設定惡意軟體防護：實作各種機制來偵測您環境中的惡意軟體。例如，您可以使用 [Amazon GuardDuty](#) 等工具來分析和偵測 EC2 和 EBS 磁碟區中的惡意軟體並發出警示。GuardDuty 也可以掃描新上傳到 Amazon S3 的物件，以找出潛藏的惡意軟體或病毒，並採取行動來隔離它們，以防它們進入下游程序。
6. 在 CI/CD 管道中整合漏洞掃描：如果您使用 CI/CD 管道進行應用程式部署，請將漏洞掃描工具整合到您的管道中。Amazon CodeGuru 安全工具和開放原始碼選項等工具可掃描您的原始程式碼、相依性和成品，以找出潛在的安全問題。

7. 設定安全監控服務：設定安全監控服務 (例如 AWS Security Hub CSPM)，以全面檢視多種雲端服務的安全狀態。監控服務應從各種不同的來源收集安全調查結果，並以標準化格式呈現，以更容易排定優先順序和修復。
8. 實作 Web 應用程式滲透測試：如果您的應用程式是 Web 應用程式，且您的組織具備必要的技能或能夠取得外部協助，請考慮實作 Web 應用程式滲透測試，以找出應用程式中的潛在漏洞。
9. 利用基礎設施即程式碼實現自動化：使用基礎設施即程式碼 (IaC) 工具 (例如 [AWS CloudFormation](#)) 將資源的部署和組態自動化，包括上述安全服務在內。此做法可協助您在多個帳戶和環境中建立更一致且標準化的資源架構。
- 10 監控並持續改進：持續監控漏洞管理計畫的有效性，並視需要改進。檢閱安全調查結果、評估補救措施的有效性，並據以調整您的程序和工具。

資源

相關文件：

- [AWS Systems Manager](#)
- [AWS Lambda 的安全概觀](#)
- [Amazon CodeGuru](#)
- [透過全新的 Amazon Inspector 改進、自動化雲端工作負載的漏洞管理](#)
- [使用 Amazon Inspector 和 AWS Systems Manager 自動化 AWS 中的漏洞管理和修復 – 第 1 部分](#)

相關影片：

- [保護無伺服器容器服務的安全](#)
- [Amazon EC2 執行個體中繼資料服務的安全最佳實務](#)

SEC06-BP02 從強化的映像佈建運算

透過從強化的映像部署，就可減少意外存取執行時期環境的機會。僅從受信任的登錄檔取得執行時期相依項 (例如容器映像和應用程式庫)，並驗證其簽章。建立自己的私有登錄檔來儲存受信任的映像和程式庫，以供您的建置和部署程序使用。

預期成果：您的運算資源是從強化的基準映像佈建。您只會從受信任的登錄檔擷取外部相依項 (例如容器映像和應用程式庫)，並驗證其簽章。這些都會儲存在私有登錄檔中，以供您的建置和部署程序參考。您會定期掃描和更新映像與相依項，以協助防禦任何新發現的漏洞。

常見的反模式：

- 從受信任的登錄檔取得映像和程式庫，但未先驗證其簽章或執行漏洞掃描，即逕行使用。
- 強化映像，但未定期測試映像以確認是否有新的漏洞或更新到最新版本。
- 安裝或未移除在預期的映像生命週期內不需要的軟體套件。
- 僅仰賴修補來讓實際執行運算資源保持最新狀態。單單是修補就仍有可能導致運算資源在經過一段時間後，偏離強化的標準。修補也可能無法移除威脅行為者在安全事件期間安裝的惡意軟體。

建立此最佳實務的優勢：強化映像有助於減少您的執行時期環境中可能成為未經授權使用者或服務意外存取路徑的數目。此外還能在發生任何意外存取的情況時，縮小影響的範圍。

未建立此最佳實務時的風險暴露等級：高

實作指引

若要強化您的系統，請從作業系統、容器映像和應用程式庫的最新版本開始。套用已知問題的修補程式。移除任何不需要的應用程式、服務、裝置驅動程式、預設使用者和其他憑證，藉此盡可能縮減系統規模。採取任何其他必要的行動，例如，停用連接埠以建立只有工作負載所需資源和功能的環境。以此為基準，您就可以安裝用於監控工作負載或管理漏洞等操作所需的軟體、代理程式或其他程序。

您可以使用受信任來源提供的指南來減輕強化系統的負擔，例如[網際網路安全中心](#)（ CIS ）和國防資訊系統局（ DISA ）[安全技術實作指南（ STIGs ）](#)。我們建議您從 AWS 或 APN 合作夥伴發佈的 [Amazon Machine Image](#)（ AMI ）開始，並使用 AWS [EC2 Image Builder](#) 根據適當的 CIS 和 STIG 控制項組合來自動化組態。

雖然有可用的強化映像和 EC2 映像建置器配方會套用 CIS 或 DISA STIG 建議，但您可能會發現其組態阻止軟體順利執行。在這種情況下，您可以從非強化基本映像開始，安裝軟體，然後逐步套用 CIS 控制項來測試其影響。對於防止軟體執行的任何 CIS 控制項，請測試您是否可以在 DISA 中實作更精細的強化建議。追蹤您能夠成功套用的不同 CIS 控制項和 DISA STIG 組態。使用這些選項，相應地在 Image Builder 中定義您的 EC2 映像強化配方。

對於容器化工作負載，來自 Docker 的強化映像可在 [Amazon Elastic Container Registry（ ECR ）公有儲存庫](#) 上取得。您可以使用 EC2 Image Builder 搭配 [Amazon Elastic Container Registry](#) 來強化容器映像 AMIs。

與作業系統和容器映像類似，您可以透過 pip、npm、Maven 和 等工具，從公有儲存庫取得程式碼套件（或程式庫）NuGet。我們建議您藉由整合私有儲存庫（例如在 [AWS CodeArtifact](#) 內）與受信任的公有儲存庫來管理程式碼套件。此整合可以 up-to-date 為您處理擷取、儲存和保留套件。然後，您的應用程式建置程序可以使用 Software Composition Analysis（ SCA ）、Static Application Security

Testing (SAST) 和 Dynamic Application Security Testing () 等技術，取得並測試這些套件的最新版本DAST。

對於使用 的無伺服器工作負載 AWS Lambda，可簡化使用 [Lambda 層管理套件相依性](#)。使用 Lambda 層設定一組跨不同函數共用的標準相依項，並放入獨立的封存中。您可以透過自己的建置程序來建立和維護層，為您的函數提供保留 的中央方式 up-to-date。

實作步驟

- 強化作業系統。使用信任來源的基本映像作為建置強化 的基礎AMIs。使用 [EC2 Image Builder](#) 協助自訂安裝在映像上的軟體。
- 強化容器化資源。設定容器化資源以符合安全最佳實務。使用容器時，請在建置管道中實作[ECR映像掃描](#)，並定期針對映像儲存庫實作映像掃描，以便在CVEs容器中尋找。
- 搭配 使用無伺服器實作時 AWS Lambda，請使用 [Lambda 層](#)來分隔應用程式函數程式碼和共用相依程式庫。為 Lambda 設定[程式碼簽署](#)，確保只有受信任的程式碼能夠在您的 Lambda 函數中執行。

資源

相關的最佳實務：

- [OPS05-BP05 執行修補程式管理](#)

相關影片：

- [深入探索 AWS Lambda 安全性](#)

相關範例：

- [AMI使用 EC2 Image Builder 快速建置 STIG合規](#)
- [建置更好的容器映像](#)
- [使用 Lambda 層簡化開發流程](#)
- [使用無伺服器架構開發和部署 AWS Lambda 層](#)
- [使用開放原始碼 SCA、SAST和 DAST 工具建置 end-to-end AWS DevSecOps CI/CD 管道](#)

SEC06-BP03 減少手動管理和互動式存取

盡可能使用自動化方式來執行部署、組態、維護和調查任務。在發生緊急程序的情況下或在安全 (沙盒) 環境中無法啟用自動化時，請考慮手動存取運算資源。

預期成果：程式化的指令碼和自動化文件 (執行手冊) 會擷取運算資源上獲得授權的動作。這些執行手冊會自動啟動、透過變更偵測系統啟動，或是在需要人為判斷時手動啟動。只有在無法啟用自動化的緊急情況下，才能直接存取運算資源。所有手動活動都會加以記錄並納入審查程序中，以持續改善您的自動化功能。

常見的反模式：

- 透過 SSH 或 RDP 等協定互動式存取 Amazon EC2 執行個體。
- 維護個別使用者登入，例如 `/etc/passwd` 或 Windows 本機使用者。
- 在多個使用者之間共用密碼或私有金鑰以存取執行個體。
- 手動安裝軟體和建立或更新組態檔案。
- 手動更新或修補軟體。
- 登入執行個體以對問題進行疑難排解。

建立此最佳實務的優勢：透過自動化方式執行步驟，有助於降低意外變更和組態錯誤伴隨的操作風險。不再使用 Secure Shell (SSH) 和遠端桌面協定 (RDP) 進行互動式存取，因此縮小了運算資源的存取範圍。這樣也消除了常見的未經授權動作路徑。在自動化文件和程式化指令碼中寫入運算資源管理任務，提供了以更精細的細節程度定義和稽核完整的授權活動範圍的機制。

未建立此最佳實務時的曝險等級：中

實作指引

登入執行個體是系統管理的傳統方法。安裝伺服器作業系統後，使用者通常會手動登入以設定系統並安裝所需的軟體。在伺服器的生命週期內，使用者可能會登入以執行軟體更新、套用修補程式、變更組態及對問題進行疑難排解。

但是，手動存取伴隨著許多風險。它需要伺服器監聽請求，例如 SSH 或 RDP 服務，而這些服務可能成為未經授權的存取路徑。此外，它也會增加執行手動步驟時發生人為錯誤的風險。這些都可能導致工作負載事件、資料損壞或銷毀，或其他安全問題。人為存取也需要設置防護措施來防止憑證共用行為，因而產生額外的管理負擔。

為了降低這些風險，您可以實作代理程式型遠端存取解決方案，例如 [AWS Systems Manager](#)。AWS Systems Manager Agent (SSM Agent) 會啟動加密通道，因此不需仰賴偵聽外部發出的請求。請考慮設定 SSM Agent 以 [透過 VPC 端點建立此通道](#)。

Systems Manager 可讓您精細控制與受管執行個體互動的方式。您可以定義要執行的自動化程序、誰可以執行它們，以及何時可以執行。Systems Manager 不需互動式存取執行個體，即可套用修補程式、安裝軟體及進行組態變更。Systems Manager 還可提供對遠端 Shell 的存取權，並將工作階段期間調用的每個命令及其輸出記錄到日誌和 [Amazon S3](#)。[AWS CloudTrail](#) 會記錄 Systems Manager API 的調用以供檢測。

實作步驟

1. 在 Amazon EC2 執行個體上 [安裝 AWS Systems Manager Agent](#) (SSM Agent)。查看是否包含 SSM Agent，且它會作為基本 AMI 組態的一部分自動啟動。
2. 確認與您的 EC2 執行個體設定檔關聯的 IAM 角色是否包含 AmazonSSMManagedInstanceCore [受管 IAM 政策](#)。
3. 停用執行個體上執行的 SSH、RDP 及其他遠端存取服務。您可以藉由執行啟動範本的使用者資料區段中設定的指令碼，或使用 EC2 Image Builder 等工具建置自訂 AMI，以執行此操作。
4. 確認適用於 EC2 執行個體的安全群組輸入規則不允許在連接埠 22/tcp (SSH) 或連接埠 3389/tcp (RDP) 上的存取。使用 AWS Config 等服務實作偵測，並對設定錯誤的安全群組發出提醒。
5. 定義適當的自動化、執行手冊，並在 Systems Manager 中執行命令。使用 IAM 政策定義誰可以執行這些動作，以及允許執行這些動作的條件。在非實際執行環境中完整測試這些自動化程序。在必要時調用這些自動化程序，而非以互動方式存取執行個體。
6. 在必要時，使用 [AWS Systems Manager Session Manager](#) 提供對執行個體的互動式存取。開啟工作階段活動日誌記錄，以在 [Amazon CloudWatch Logs](#) 或 [Amazon S3](#) 中維護稽核記錄。

資源

相關的最佳實務：

- [REL08-BP04 使用不可變基礎設施進行部署](#)

相關範例：

- [使用 AWS Systems Manager 取代 SSH 存取以減輕管理和安全負擔](#)

相關工具：

- [AWS Systems Manager](#)

相關影片：

- [在 AWS Systems Manager Session Manager 中控制使用者工作階段對執行個體的存取權](#)

SEC06-BP04 驗證軟體完整性

使用加密驗證來驗證工作負載所使用之軟體成品 (包括映像) 的完整性。以加密方式簽署您的軟體，以防範未經授權的變更在您的運算環境內執行。

預期成果：所有成品都是從受信任的來源取得。廠商網站憑證經過驗證。下載的成品已藉由簽章以加密方式驗證。自有軟體會由您的運算環境以加密方式簽署和驗證。

常見的反模式：

- 信任信譽良好的廠商網站來取得軟體成品，但忽略憑證到期通知。未先確認憑證是否有效，即逕行下載。
- 驗證廠商網站憑證，但未以加密方式驗證從這些網站下載的成品。
- 僅仰賴摘要或雜湊值來驗證軟體完整性。雜湊值確定成品的原版未經修改，但未驗證其來源。
- 即使僅在您自己的部署中使用，也未簽署自有軟體、程式碼或程式庫。

建立此最佳實務的優勢：驗證您的工作負載所依賴之成品的完整性，有助於防止惡意軟體進入您的運算環境。簽署您的軟體有助於防範運算環境中發生未經授權執行的情況。藉由簽署和驗證程式碼來保護您的軟體供應鏈。

未建立此最佳實務時的風險暴露等級：中

實作指引

作業系統映像、容器映像和程式碼成品通常在散佈時會提供完整性檢查，例如透過摘要或雜湊值。這些檢查可讓用戶端運算自有的承載雜湊值並確認其與發佈的雜湊值相同，藉此驗證完整性。雖然這些檢查有助於驗證承載未遭到竄改，但不會驗證承載來自原始出處 (其來源)。驗證來源需要使用受信任的授權機構發出的憑證來數位簽署成品。

如果您在工作負載中使用下載的軟體或成品，請檢查提供者是否提供了用於驗證數位簽章的公有金鑰。以下一些範例說明 AWS 如何提供公有金鑰，以及如何驗證我們發佈的軟體：

- [EC2 Image Builder：驗證安裝下載的 AWS TOE 簽章](#)

- [AWS Systems Manager：驗證SSM客服人員的簽章](#)
- [Amazon CloudWatch：驗證 CloudWatch 客服人員套件的簽章](#)

將數位簽章驗證納入您用於取得和強化影像的程序，如 [SEC06-BP02 佈建來自強化影像的運算](#) 中所述。

您可以使用 [AWS Signer](#) 協助您管理簽章驗證，以及您自有軟體和成品的程式碼簽署生命週期。[AWS Lambda](#) 和 [Amazon Elastic Container Registry](#) 兩者都提供與 Signer 的整合，可用來驗證程式碼和映像的簽章。您可以使用「資源」區段中的範例，將 Signer 納入您的持續整合和交付 (CI/CD) 管道中，以便自動驗證簽章及自動簽署自有程式碼和映像。

資源

相關文件：

- [容器的加密簽署](#)
- [使用 協助保護容器映像建置管道的最佳實務 AWS Signer](#)
- [宣佈使用 AWS Signer 和 Amazon 進行容器映像簽署 EKS](#)
- [設定 的程式碼簽署 AWS Lambda](#)
- [Lambda 程式碼簽署的最佳實務和進階模式](#)
- [使用 AWS Certificate Manager 私有 CA AWS Key Management Service 和非對稱金鑰進程式碼簽署](#)

相關範例：

- [使用 Amazon CodeCatalyst 和 自動化 Lambda 程式碼簽署 AWS Signer](#)
- [使用 簽署和驗證OCI偽影 AWS Signer](#)

相關工具：

- [AWS Lambda](#)
- [AWS Signer](#)
- [AWS Certificate Manager](#)
- [AWS Key Management Service](#)
- [AWS CodeArtifact](#)

SEC06-BP05 自動化運算保護

自動化運算保護操作以減少人工介入的需求。使用自動化掃描偵測運算資源內的潛在問題，並透過自動化的程式化回應或機群管理操作進行修復。將自動化納入 CI/CD 程序，以部署具有最新相依項且值得信賴的工作負載。

預期成果：自動化系統會執行運算資源的所有掃描和修補工作。您會使用自動化的驗證方式檢查軟體映像和相依項來自受信任的來源，且未遭到竄改。透過自動化方式檢查工作負載是否有最新的相依項，並且簽署工作負載以在 AWS 運算環境中建立可靠性。偵測到不合規資源時，系統會啟動自動補救措施。

常見的反模式：

- 遵循不可變的基礎設施實務，但未備妥解決方案來因應緊急修補或取代實際執行系統。
- 使用自動化方式修復設定錯誤的資源，但未設置手動覆寫機制。可能會發生需要調整需求的情況，且您可能需要暫停自動化程序，直到完成這些變更為止。

建立此最佳實務的優勢：自動化可降低未經授權存取和使用您的運算資源的風險。它有助於防止錯誤的組態進入實際執行環境，並且在發生組態錯誤時偵測到該錯誤並加以修復。自動化還可協助偵測未經授權存取和使用運算資源的情況，進而縮短您回應的時間。如此還能進一步縮小問題的整體影響範圍。

未建立此最佳實務時的曝險等級：中

實作指引

您可以套用「安全支柱」實務中所述的自動化方式，以保護您的運算資源。[SEC06-BP01 執行漏洞管理](#)說明如何在您的 CI/CD 管道中使用 [Amazon Inspector](#)，以及如何運用它持續掃描您的執行時期環境，以找出已知的通用漏洞披露 (CVE)。您可以透過自動化執行手冊，使用 [AWS Systems Manager](#) 套用修補程式或從全新映像重新部署，讓您的運算機群隨時擁有最新的軟體和程式庫。使用這些技術可減少對手動程序和互動式存取運算資源的需求。請參閱 [SEC06-BP03 減少手動管理和互動式存取](#) 以了解更多資訊。

自動化在部署值得信賴的工作負載方面，也發揮了舉足輕重的作用，如 [SEC06-BP02 從強化的映像佈建運算](#) 和 [SEC06-BP04 驗證軟體完整性](#) 中所述。您可以使用 [EC2 Image Builder](#)、[AWS Signer](#)、[AWS CodeArtifact](#) 和 [Amazon Elastic Container Registry \(ECR\)](#) 等服務來下載、驗證、建構和儲存強化且經核准的映像和程式碼相依項。除了 Inspector 之外，這些都可在 CI/CD 程序中發揮作用。因此，只有在確認工作負載的相依項為最新狀態且來自受信任的來源時，工作負載才會進入實際執行環境。您的工作負載也會經過簽署，如此一來，[AWS Lambda](#) 和 [Amazon Elastic Kubernetes Service \(EKS\)](#) 等 AWS 運算環境就能在確認其未遭到竄改後，再允許其執行。

除了這些預防性控制之外，您還可以在偵測控制中針對運算資源使用自動化。舉例來說，[AWS Security Hub CSPM](#) 提供 [NIST 800-53 Rev. 5](#) 標準，其中包括如下述的檢查：[\[EC2.8\] EC2 執行個體應使用執行個體中繼資料服務第 2 版 \(IMDSv2\)](#)。IMDSv2 會使用工作階段驗證技術，封鎖包含 X-Forwarded-For HTTP 標頭以及網路 TTL 1 的請求，藉此阻止來自外部來源的流量擷取有關 EC2 執行個體的資訊。Security Hub CSPM 中的這項檢查可偵測 EC2 執行個體何時使用 IMDSv1，並實施自動補救措施。請參閱 [SEC04-BP04 針對不合規資源實施補救措施](#)，進一步了解自動化偵測和補救措施。

實作步驟

1. 使用 [EC2 Image Builder](#) 自動建立安全、合規且強化的 AMI。您可以產生映像，並於當中納入來自 Center for Internet Security (CIS) Benchmarks 的控制，或來自基本 AWS 和 APN 合作夥伴映像的安全技術實作指南 (STIG) 標準。
2. 自動化組態管理。藉由使用組態管理服務或工具，在您的運算資源中自動強制執行和驗證安全組態。
 - a. 使用 [AWS Config](#) 自動化組態管理
 - b. 使用 [AWS Security Hub CSPM](#) 自動化安全和合規狀態管理
3. 自動修補或取代 Amazon Elastic Compute Cloud (Amazon EC2) 執行個體。AWS Systems Manager Patch Manager 可自動執行透過安全相關及其他更新來修補受管執行個體的程序。您可以使用修補程式管理員以套用適用於作業系統和應用程式的修補程式。
 - a. [AWS Systems Manager Patch Manager](#)
4. 自動掃描運算資源以找出通用漏洞披露 (CVE)，並將安全掃描解決方案內嵌於您的建置管道中。
 - a. [Amazon Inspector](#)
 - b. [ECR 映像掃描](#)
5. 考慮使用 Amazon GuardDuty 執行自動化惡意軟體和威脅偵測，以保護運算資源。GuardDuty 還可在 [AWS Lambda](#) 函數於您的 AWS 環境中調用時，識別潛在問題。
 - a. [Amazon GuardDuty](#)
6. 考慮 AWS 合作夥伴解決方案。AWS 合作夥伴提供領先業界的產品，這些產品與您內部部署環境中的現有控制相當、相同或互相整合。這些產品可補充現有的 AWS 服務，讓您在雲端和內部部署環境中部署全方位的安全架構，以及擁有更流暢的體驗。
 - a. [基礎設施安全性](#)

資源

相關的最佳實務：

- [SEC01-BP06 自動部署標準安全控制](#)

相關文件：

- [在您的 AWS 基礎設施內充分利用 IMDSv2 的優勢並停用 IMDSv1](#)

相關影片：

- [Amazon EC2 執行個體中繼資料服務的安全最佳實務](#)

資料保護

問題

- [SEC 7. 如何分類資料？](#)
- [SEC 8. 如何保護靜態資料？](#)
- [SEC 9. 如何保護傳輸中的資料？](#)

SEC 7. 如何分類資料？

資料分類可讓您根據關鍵性和敏感度將資料分類，協助判定適當的保護和保留控制。

最佳實務

- [SEC07-BP01 了解您的資料分類機制](#)
- [SEC07-BP02 根據資料敏感性實施資料保護控制](#)
- [SEC07-BP03 自動識別和分類](#)
- [SEC07-BP04 定義可擴展的資料生命週期管理](#)

SEC07-BP01 了解您的資料分類機制

了解您的工作負載要處理的資料分類、其處理需求、關聯的業務流程、資料儲存在何處，以及誰是資料擁有者。您的資料分類和處理機制應考慮工作負載的適用法律和合規需求，以及需要何種資料控制。了解資料是資料分類之旅的第一步。

預期成果：您的工作負載中存在的資料類型已得到充分了解並加以記錄。設置了適當的控制，可根據資料分類來保護敏感資料。這些控制左右著下列考量：允許誰存取資料及存取的目的為何、資料儲存

在何處、該資料的加密政策及如何管理加密金鑰、資料的生命週期及其保留需求、適當的銷毀程序、設置了哪些備份和復原程序，以及存取權稽核。

常見的反模式：

- 未制定正式的資料分類政策來定義資料敏感程度及其處理需求
- 未充分了解工作負載內資料的敏感程度，也未在架構和營運文件中擷取這些資訊
- 未能根據您的資料分類和處理政策中所述的資料敏感程度和需求，對資料實施適當的控制
- 未能向政策負責人提供有關資料分類和處理需求的意見回饋。

建立此最佳實務的優勢：此實務可消除有關適當處理工作負載內資料的不確定性。實施正式政策來定義組織中資料的敏感程度及其所需防護措施，有助於符合法律規範和其他網路安全鑑定與認證。工作負載負責人清楚知道敏感資料儲存在何處以及設置了哪些保護控制，因而能夠放心。將這些資訊納入文件中，可協助新的團隊成員更充分了解並在任職期間及早採取控制。這些實務還可針對每一種資料類型實施適當的控制，進而有助於降低成本。

未建立此最佳實務時的曝險等級：高

實作指引

在設計工作負載時，您可能會考量採取直接了當的方式保護敏感資料。例如，在多租用戶應用程式中，直接將每一個租用戶的資料視為敏感資料並採取防護措施，讓租用戶無法存取其他租用戶的資料。同樣地，您可能會直接設計存取控制，只讓管理員修改資料，而其他使用者只擁有讀取層級存取權，或完全無存取權。

藉由在政策中定義並擷取這些資料敏感程度，以及其資料保護需求，您就能正式確定哪些資料要放在您的工作負載中。然後，您可以確定是否設置了正確的控制、是否可稽核控制，以及在發現資料遭不當處理的情況時，要採取何種適當的回應。

為了有助於識別敏感資料存放在工作負載中的位置，請考慮使用資料目錄。資料目錄是對應您組織中資料的資料庫，其位置、敏感度層級和既有控制機制都會保護該資料。此外，請考慮使用可用的[資源標籤](#)。例如，您可以對受保護醫療資訊 (PHI) 套用標籤索引鍵為 Classification 且標籤值為 PHI 的標籤，以及另一個標籤索引鍵為 Sensitivity 且標籤值為 High 的標籤。然後，您可以使用 [AWS Config](#) 等服務來監控這些資源是否發生變更，並且在發生修改後導致資源不符合保護需求 (例如變更加密設定) 的情況時發出提醒。您可以使用[標籤政策](#) (AWS Organizations 的功能) 擷取標籤索引鍵的標準定義和可接受的值。不建議在標籤索引鍵或值中包含私人或敏感資料。

實作步驟

1. 了解組織的資料分類機制和保護需求。
2. 識別工作負載處理的敏感資料類型。
3. 在資料目錄中擷取資料，該目錄提供單一檢視讓您得知資料在組織中的位置，以及該資料的敏感度層級。
4. 考慮使用資源和資料層級標記 (如可用) 來標記資料的敏感程度和其他操作中繼資料，以協助監控和回應事件。
 - a. AWS Organizations 標籤政策可用來強制實施標記標準。

資源

相關的最佳實務：

- [SUS04-BP01 實作資料分類政策](#)

相關文件：

- [資料分類白皮書](#)
- [標記 AWS 資源的最佳實務](#)

相關範例：

- [AWS Organizations 標籤政策語法和範例](#)

相關工具

- [AWS Tag Editor](#)

SEC07-BP02 根據資料敏感性實施資料保護控制

實施資料保護控制，為您分類政策中定義的每一個資料類別提供適當的控制層級。此實務可讓您保護敏感資料防止遭到未經授權的存取和使用，同時讓資料保持可用且實用。

預期成果：您設置了分類政策，在組織中定義不同程度的資料敏感性。您針對每一種敏感程度發布了清楚的指引，以界定核准的儲存和處理服務與位置，以及其所需的組態。您根據所需的保護層級及其關聯成本，針對每一種敏感程度實施控制。您設置了監控和提醒，以偵測資料是否出現在未經授權的

位置、在未經授權的環境中經過處理、遭到未經授權的人員存取，或是相關服務的組態是否變得不合規。

常見的反模式：

- 對所有資料實施相同層級的保護控制。這可能會導致對低敏感性資料過度佈建安全控制，或對高敏感性資料的保護不足。
- 在定義資料保護控制時，未邀集安全、合規和業務團隊的利害關係人參與此過程。
- 忽略實施和維護資料保護控制伴隨的營運支出和成本。
- 未定期審查資料保護控制，而未能持續遵循分類政策。
- 沒有靜態資料和傳輸中資料位置的完整庫存。

建立此最佳實務的優勢：藉由依照資料分類層級實施您的控制，您的組織就能在需要時投入更高層級的控制。這可能包括增加保障安全、監控、衡量、修復和報告方面的資源。在適度採取較少控制的情況下，您可以改善員工、客戶或成員使用的資料存取性和完整性。此方法為您的組織帶來了最大的資料使用彈性，同時遵守資料保護要求。

未建立此最佳實務時的曝險等級：高

實作指引

根據資料敏感程度實施資料保護控制的方式包含幾個重要的步驟。首先，識別工作負載架構內不同的資料敏感程度（例如公開、內部、機密和受限），並評估您儲存和處理這些資料的位置。接著，根據資料敏感程度定義其隔離界限。建議您使用[服務控制政策](#) (SCP) 將資料分隔到不同的 AWS 帳戶中，以限制每一種資料敏感程度允許的服務和動作。這樣一來，您就可以建立強大的隔離界限，並強制執行最低權限原則。

定義隔離界限之後，根據資料敏感程度實施適當的保護控制。請參閱[保護靜態資料](#)和[保護傳輸中的資料](#)的最佳實務，以實作加密、存取控制和稽核等相關控制。考慮採用記號化或匿名化等技術來降低資料的敏感程度。採用集中式系統進行記號化和去記號化，以簡化對整個企業套用一致的資料政策的程序。

持續監控和測試所實作控制的有效性。隨著組織的資料態勢和威脅發展，定期審查和更新資料分類機制、風險評估和保護控制。實作的資料保護控制務必遵循相關產業法規、標準和法律要求。此外，提供安全意識和培訓，協助員工了解資料分類機制及他們在處理和保護敏感資料方面的責任。

實作步驟

1. 識別工作負載內資料的分類和敏感程度。

2. 為每一種敏感程度定義隔離界限，並確定執行策略。
3. 評估您定義的控制是否確實有效控管您的資料分類政策規定的存取、加密、稽核、保留和其他方面。
4. 評估能適時降低資料敏感程度的選項，例如使用記號化或匿名化。
5. 使用自動測試和監控所設定資源的方式來驗證您的控制。

資源

相關的最佳實務：

- [PERF03-BP01 使用最能滿足資料存取和儲存需求的專用資料存放區](#)
- [COST04-BP05 強制執行資料保留政策](#)

相關文件：

- [資料分類白皮書](#)
- [安全、身分及合規最佳實務](#)
- [AWS KMS 最佳實務](#)
- [AWS 服務的加密最佳實務和功能](#)

相關範例：

- [建置無伺服器記號化解決方案為敏感資料提供遮罩](#)
- [如何使用記號化的方式來提高資料安全並縮小稽核範圍](#)

相關工具：

- [AWS Key Management Service \(AWS KMS\)](#)
- [AWS CloudHSM](#)
- [AWS Organizations](#)

SEC07-BP03 自動識別和分類

將資料的識別和分類自動化，可協助您實作正確的控制。使用自動化增強手動判斷，可降低人為錯誤和暴露的風險。

預期成果：您可根據您的分類和處理政策來確認是否有適當的控制。自動化工具和服務可協助您識別和分類資料的敏感程度。自動化還可協助您持續監控環境，以偵測並提醒未經授權的資料儲存或處理行為，以便快速採取更正動作。

常見的反模式：

- 僅仰賴手動程序來識別和分類資料，這個過程可能容易出錯且相當耗時。這樣做可能導致資料分類效率不彰且不一致，尤其隨著資料量增加會每況愈下。
- 未設置追蹤和管理整個組織中資料資產的機制。
- 即使資料在組織內移動和發展，組織仍然忽略持續監控和分類資料的需要。

建立此最佳實務的優勢：採取自動識別和分類資料的方式，能夠更一致且準確地實施資料保護控制，進而降低人為錯誤的風險。自動化還可讓您深入洞悉敏感資料存取和移動的情形，進而協助您偵測未經授權的處理，並採取更正動作。

未建立此最佳實務時的曝險等級：中

實作指引

在工作負載的初始設計階段常會採用人為判斷來分類資料，儘管如此，仍請考慮設置系統來自動識別和分類測試資料，以此做為預防性控制。例如，您可提供工具或服務讓開發人員用來掃描代表性的資料，以確定其敏感性。在 AWS 內，您可以將資料集上傳到 [Amazon S3](#)，並使用 [Amazon Macie](#)、[Amazon Comprehend](#) 或 [Amazon Comprehend Medical](#) 掃描這些資料集。同樣地，請考慮在單元和整合測試的過程中掃描資料，以偵測不該出現敏感資料的位置。在此階段發出有關敏感資料的提醒，就能在部署到實際執行環境之前，讓防護措施的落差浮現。其他諸如 [AWS Glue](#)、[Amazon SNS](#) 和 [Amazon CloudWatch](#) 中的敏感資料偵測等功能，也可用來偵測 PII 和採取緩解措施。對於任何自動化工具或服務，務必了解其如何定義敏感資料，並利用其他人為或自動化解決方案加強它，以視需要消除任何落差。

持續監控您的環境，以其做為偵測控制，藉以偵測敏感資料是否以不合規的方式儲存。這樣做有助於偵測出在未適當去識別化或修訂的情況下，將敏感資料發送到日誌檔案或複製到資料分析環境中的情形。您可以使用 Amazon Macie 持續監控儲存在 Amazon S3 中的資料，以偵測其中是否存在敏感資料。

實作步驟

1. 請檢閱 [SEC07-BP01](#) 中所述，了解您組織內的資料分類結構。
 - a. 只要了解您組織的資料分類結構，您就可以建立準確的流程用於自動化適合您公司政策的識別和分類。

2. 對您的環境執行初始掃描，以進行自動識別和分類。
 - a. 初始完整掃描資料有助於全面了解敏感資料在您環境中的位置。若一開始不需要或因成本考量而無法事先完成完整掃描，請評估資料取樣技術是否適合用來實現您的成果。例如，您可以設定 Amazon Macie 跨 S3 儲存貯體執行廣泛的自動化敏感資料探索操作。此功能使用的取樣技術會以符合成本效益的方式初步分析敏感資料的所在位置。然後，您可以使用敏感資料探索工作來深入分析 S3 儲存貯體。也可以將其他資料存放區匯出至 S3，以便讓 Macie 進行掃描。
 - b. 針對掃描中識別的資料儲存資源，依照 [SEC07-BP02](#) 中所定義來建立存取控制。
3. 設定環境的持續掃描。
 - a. Macie 的自動化敏感資料探索功能可用來持續掃描您的環境。若有任何經授權儲存敏感資料的已知 S3 儲存貯體，則可使用 Macie 中的允許清單將其排除在外。
4. 將識別和分類納入您的建置和測試程序中。
 - a. 識別開發人員可在工作負載開發過程中用來掃描資料以判斷敏感性的工具。在整合測試的過程中使用這些工具，以便在敏感資料意外出現時發出提醒，並防止進一步部署。
5. 在未經授權的位置發現敏感資料時，實作系統或執行手冊來採取行動。
 - a. 使用自動修復來限制對資料的存取。例如，若您使用屬性型存取控制 (ABAC)，就可以將此資料移至限制存取的 S3 儲存貯體或標記物件。此外，請考慮在偵測到資料時，為其加上遮罩。
 - b. 提醒您的資料防護和事件回應團隊調查事件的根本原因。他們在調查過程中學到的任何經驗，都有助於防止未來的事件發生。

資源

相關文件：

- [AWS Glue：偵測和處理敏感資料](#)
- [在 Amazon SNS 中使用受管資料識別碼](#)
- [Amazon CloudWatch Logs：使用遮罩協助保護敏感日誌資料](#)

相關範例：

- [使用 Macie 為 Amazon RDS 資料庫啟用資料分類](#)
- [使用 Macie 偵測 DynamoDB 中的敏感資料](#)

相關工具：

- [Amazon Macie](#)

- [Amazon Comprehend](#)
- [Amazon Comprehend Medical](#)
- [AWS Glue](#)

SEC07-BP04 定義可擴展的資料生命週期管理

了解您的資料生命週期需求，因為這些需求與您不同層級的資料分類和處理相關。這可能包括資料一開始進入環境時的處理方式、資料轉換的方式，以及銷毀資料的規則。請將保留期、存取、稽核和追蹤來源等因素納入考量。

預期成果：您的資料分類會盡可能接近擷取點和時間。當資料分類需要遮罩、記號化或其他降低敏感程度的處理時，您會在盡可能最接近擷取點和時間的條件下執行這些動作。

當資料不再適合保存時，您會遵循政策根據資料的分類將其刪除。

常見的反模式：

- 實作一體適用的方法來管理資料生命週期，而未考量不同的敏感程度和存取需求。
- 僅從資料為可用資料或備份資料的角度來考量生命週期管理，而非兩者均考量。
- 假設已輸入工作負載的資料有效，但未確定其價值或來源。
- 仰賴資料耐久性來替代資料備份和保護。
- 保留資料的時間超過其實用性和所需的保留期。

建立此最佳實務的優勢：定義明確且可擴展的資料生命週期管理策略有助於保持合規、提高資料安全性、最佳化儲存成本，以及在維持適當控制之下實現有效率的資料存取和共用。

未建立此最佳實務時的風險暴露等級：高

實作指引

工作負載內的資料通常是動態的。資料進入工作負載環境時採取的形式，可能與資料儲存或使用在商業邏輯、報告、分析或機器學習上的形式有所不同。此外，資料的價值可能隨時間而改變。有些資料本質上是暫時性的，會隨著時間失去其價值。請考量在您的資料分類機制與相關控制下，這些資料變更對評估的影響。可能的話，盡量使用自動化生命週期機制（如 [Amazon S3 生命週期政策](#) 和 [Amazon Data Lifecycle Manager](#)）來設定資料保留、封存和到期程序。對於儲存在 DynamoDB 中的資料，您可以使用 [存留時間 \(TTL\)](#) 功能來定義每個項目的過期時間戳記。

區分可供使用的資料與儲存為備份的資料。考慮使用 [AWS Backup](#) 自動備份 AWS 服務中的資料。[Amazon EBS 快照](#) 提供了複製 EBS 磁碟區並使用 S3 功能來儲存它的方式，包括生命週期、資料

保護和存取保護機制。其中兩種機制為 [S3 Object Lock](#) 和 [AWS Backup Vault Lock](#)，皆可提高您備份的安全性，並且讓您更有效地掌控備份。進行分明的職責和備份存取權劃分管理。在帳戶層級隔離備份，以便在事件發生期間與受影響的環境保持分離。

生命週期管理的另一方面，是記錄資料在工作負載中進度的歷史記錄，稱為資料來源追蹤。如此您就能確信自己知道資料來自何處、執行的任何轉換、哪些擁有者或處理程序做出這些變更，以及時間點。這份歷史記錄有助於在可能發生安全事件的期間進行問題的疑難排解和調查。例如，您可以在 [Amazon DynamoDB](#) 資料表中記錄有關轉換的中繼資料。在資料湖內，您可以針對每一個資料管道階段，將轉換後資料的副本保留在不同的 S3 儲存貯體中。將結構描述和時間戳記資訊儲存在 [AWS Glue Data Catalog](#) 中。無論您採用何種解決方案，請務必考量最終使用者的需求，以確定報告資料來源所需的適當工具。這樣做將協助您確定追蹤來源的最佳方式。

實作步驟

1. 分析工作負載的資料類型、敏感程度及存取需求，以分類資料並定義適當的生命週期管理策略。
2. 設計並實施符合法律、法規和組織要求的資料保留政策及自動銷毀程序。
3. 建立流程和自動化功能，以隨著工作負載需求和法規發展，持續監控、稽核和調整資料生命週期管理策略、控制及政策。
 - a. 使用 [AWS Config](#) 偵測尚未開啟自動生命週期管理的資源

資源

相關的最佳實務：

- [COST04-BP05 強制執行資料保留政策](#)
- [SUS04-BP03 使用政策來管理資料集的生命週期](#)

相關文件：

- [資料分類白皮書](#)
- [AWS 勒索軟體防禦藍圖](#)
- [DevOps 指引：利用資料來源追蹤改善可追溯性](#)

相關範例：

- [如何在 AWS 中於整個生命週期保護敏感資料](#)
- [使用 AWS Glue、Amazon Neptune 和 Spline 為資料湖建置資料歷程](#)

相關工具：

- [AWS Backup](#)
- [Amazon Data Lifecycle Manager](#)
- [AWS Identity and Access Management Access Analyzer](#)

SEC 8. 如何保護靜態資料？

實作多重控制來保護靜態資料，以降低未經授權存取或處理不當的風險。

最佳實務

- [SEC08-BP01 實作安全金鑰管理](#)
- [SEC08-BP02 強制靜態加密](#)
- [SEC08-BP03 自動化靜態資料保護](#)
- [SEC08-BP04 強制存取控制](#)

SEC08-BP01 實作安全金鑰管理

安全金鑰管理包括儲存、輪換、存取控制及監控保護工作負載的靜態資料所需的金鑰資料。

預期成果：您擁有可擴展、可重複且自動化的金鑰管理機制。此機制會對金鑰資料強制執行最低權限存取，並且在金鑰可用性、機密性和完整性之間取得正確的平衡。您會監控金鑰的存取情況，若需要輪換金鑰資料，您可以使用自動化程序進行輪換。您不允許真人操作員存取金鑰資料。

常見的反模式：

- 真人存取未加密的金鑰資料。
- 建立自訂加密演算法。
- 存取金鑰資料的許可過於廣泛。

建立此最佳實務的優勢：透過為工作負載建立安全的金鑰管理機制，就可以協助保護您的內容，防止未經授權的存取。此外，您可能需要依法加密您的資料。有效的金鑰管理解決方案能夠提供符合這些法規的技術機制，以保護金鑰資料。

未建立此最佳實務時的曝險等級：高

實作指引

靜態資料加密是基本的安全控制。為了實作此控制，您的工作負載須採取某種機制，以安全儲存和管理用於加密靜態資料的金鑰資料。

AWS 提供 AWS Key Management Service (AWS KMS) 來為 AWS KMS 金鑰提供耐用、安全和備援儲存。[許多 AWS 服務會與 AWS KMS 整合](#)，以支援資料加密。AWS KMS 使用 FIPS 140-3 3 級驗證的硬體安全模組來保護您的金鑰。沒有任何機制可將 AWS KMS 金鑰匯出為純文字。

使用多帳戶策略部署工作負載時，您應將 AWS KMS 金鑰與使用金鑰的工作負載保存在相同的帳戶中。[這個分散式模型](#)會將管理 AWS KMS 金鑰的責任歸在您的團隊身上。在其他使用案例中，您的組織可能會選擇將 AWS KMS 金鑰儲存到集中式帳戶中。此集中式結構須實施其他政策來實現跨帳戶存取權，才能讓工作負載帳戶存取儲存在集中式帳戶中的金鑰，但此結構可能較適合跨多個 AWS 帳戶共用單一金鑰的使用案例。

無論金鑰資料儲存在何處，您都應透過使用[金鑰政策](#)和 IAM 政策嚴格控管金鑰存取權。金鑰政策是控制 AWS KMS 金鑰存取權的主要方式。此外，AWS KMS 金鑰授權可提供 AWS 服務的存取權，以代表您加密和解密資料。檢閱 [AWS KMS 金鑰存取控制指引](#)。

您應監控加密金鑰的使用情況，以偵測不尋常的存取模式。使用儲存在 AWS KMS 中 AWS 管理的金鑰和客戶管理的金鑰執行的操作可記錄在 AWS CloudTrail 中，並且應定期審查。務必特別注意監控金鑰銷毀事件。為了減少意外或惡意銷毀金鑰資料的情況，金鑰銷毀事件並不會立即刪除金鑰資料。嘗試刪除 AWS KMS 中的金鑰會受到[等待期](#)的約束 (預設為 30 天，最短 7 天)，讓管理員有時間檢閱這些動作，並在必要時撤回請求。

大多數 AWS 服務會以顯而易見的方式使用 AWS KMS，您唯一要做的就是決定要使用 AWS 管理或客戶管理的金鑰。如果您的工作負載要求直接使用 AWS KMS 來加密或解密資料，則您應使用[封套加密](#)來保護您的資料。此 [AWS Encryption SDK](#) 可為您的應用程式提供用戶端加密基本類型，以實作封套加密並與 AWS KMS 整合。

實作步驟

1. 確定金鑰的適當[金鑰管理選項](#) (AWS 受管或客戶自管)。
 - a. 為了方便使用，AWS 為大多數服務提供了 AWS 擁有和 AWS 管理的金鑰，其提供靜態加密功能，而不需要管理金鑰資料或金鑰政策。
 - b. 使用客戶管理的金鑰時，請考慮使用預設金鑰存放區，以便在敏捷性、安全性、資料主權與可用性之間達到最佳平衡。其他使用案例可能會要求使用自訂金鑰存放區搭配 [AWS CloudHSM](#) 或[外部金鑰存放區](#)。

2. 審核您用於工作負載的服務清單，以了解 AWS KMS 與服務整合的方式。例如，EC2 執行個體可以使用加密的 EBS 磁碟區，因此要確認從這些磁碟區建立的 Amazon EBS 快照同樣是使用客戶管理的金鑰加密，並減少意外洩漏未加密的快照資料。
 - a. [AWS 服務使用 AWS KMS 的方式](#)
 - b. 如需 AWS 服務提供之加密選項的詳細資訊，請參閱該服務之使用者指南或開發人員指南中的靜態加密主題。
3. 實作 AWS KMS：AWS KMS 可讓您輕鬆建立和管理金鑰，並控制多種 AWS 服務和應用程式中的加密使用方式。
 - a. [入門：AWS Key Management Service \(AWS KMS\)](#)
 - b. 審核 [AWS KMS 金鑰存取控制的最佳實務](#)。
4. 考慮使用 AWS Encryption SDK：當您的應用程式需要在用戶端對資料進行加密時，可使用整合 AWS KMS 的 AWS Encryption SDK。
 - a. [AWS Encryption SDK](#)
5. 啟用 [IAM Access Analyzer](#) 以自動審核並在發現有過度廣泛的 AWS KMS 金鑰政策時發出通知。
 - a. 考慮使用 [自訂政策檢查](#)，以確認資源政策更新未將公開存取權授予 KMS 金鑰。
6. 啟用 [Security Hub CSPM](#) 以在金鑰政策設定錯誤、有排定要刪除的金鑰，或有未啟用自動輪替的金鑰時收到通知。
7. 確定適合 AWS KMS 金鑰的日誌記錄層級。由於 AWS KMS 的呼叫 (包括唯讀事件) 會加以記錄，因此與 AWS KMS 關聯的 CloudTrail 日誌可能會變得很龐大。
 - a. 有些組織偏好將 AWS KMS 日誌記錄活動分隔為單獨的追蹤。如需詳細資訊，請參閱《AWS KMS 開發人員指南》的 [Logging AWS KMS API calls with CloudTrail](#)。

資源

相關文件：

- [AWS Key Management Service](#)
- [AWS 加密服務和工具](#)
- [使用加密保護 Amazon S3 資料](#)
- [封套加密](#)
- [數位主權承諾](#)
- [揭密 AWS KMS 金鑰操作、攜帶自有金鑰、自訂金鑰存放區，以及密文可攜性](#)
- [AWS Key Management Service 密碼編譯詳細資訊](#)

相關影片：

- [在 AWS 中加密如何運作](#)
- [保護 AWS 上的區塊儲存安全](#)
- [AWS 資料保護：使用鎖定、金鑰、簽章和憑證](#)

相關範例：

- [使用 AWS KMS 實作進階存取控制機制](#)

SEC08-BP02 強制靜態加密

加密私有靜態資料以維持機密性，並提供一層額外的保護，以防止意外揭露或洩露資料。加密可保護資料，防止資料在未先解密的情況下遭到讀取或存取。清查和控制未加密的資料，以降低資料暴露伴隨的風險。

預期成果：您建立了預設加密靜態私有資料的機制。這些機制有助於維持資料的機密性，並提供一層額外的保護，以防有意或不慎揭露或洩露資料。您維護未加密資料的庫存，並了解為了保護資料而採取的控制措施。

常見的反模式：

- 未使用預設加密組態。
- 對解密金鑰提供過於寬鬆的存取權。
- 未監控加密和解密金鑰的使用。
- 在未加密的情況下儲存資料。
- 對所有資料使用相同的加密金鑰，無論資料使用方式、類型和分類。

未建立此最佳實務時的曝險等級：高

實作指引

在工作負載中將加密金鑰對應到資料分類。當對資料使用單一或極少數的加密金鑰時，此方法有助於防止過於寬鬆的存取權 (請參閱 [SEC07-BP01 了解您的資料分類機制](#))。

AWS Key Management Service (AWS KMS) 與許多 AWS 服務整合，更方便您加密靜態資料。例如，在 Amazon Elastic Compute Cloud (Amazon EC2) 中，您可以在帳戶上設定[預設加密](#)，以便自動加密

新的 EBS 磁碟區。當使用 AWS KMS 時，考慮需要嚴格限制資料的程度。AWS 會代表您管理及使用預設和服務控制的 AWS KMS 金鑰。對於需要對基礎加密金鑰的精細存取權之敏感資料，可考慮客戶自管金鑰 (CMK)。您可全權控制 CMK，包括透過使用金鑰政策進行輪換和存取管理。

此外，像是 Amazon Simple Storage Service ([Amazon S3](#)) 等服務現在預設會加密所有新物件。此實作提供了更強大的安全性，而且不會影響效能。

其他服務像是 [Amazon Elastic Compute Cloud](#) (Amazon EC2) 或 [Amazon Elastic File System](#) (Amazon EFS)，都可支援預設加密的設定。您也可以使用 [AWS Config 規則](#) 自動檢查您是否對 [Amazon Elastic Block Store \(Amazon EBS\) 磁碟區](#)、[Amazon Relational Database Service \(Amazon RDS\) 執行個體](#) 和 [Amazon S3 儲存貯體](#)，以及組織內的其他服務使用加密。

AWS 也提供用戶端加密選項，允許您在上傳到雲端之前加密資料。AWS Encryption SDK 提供使用[封套加密](#)來加密資料的方法。您提供包裝金鑰，而 AWS Encryption SDK 會為它加密的每個資料物件產生唯一的資料金鑰。如果您需要受管的單一租用戶硬體安全模組 (HSM)，可考慮 AWS CloudHSM。AWS CloudHSM 可讓您在 FIPS 140-2 3 級驗證的 HSM 上產生、匯入和管理加密金鑰。AWS CloudHSM 的一些使用案例包括保護用於核發憑證認證機構 (CA) 的私有金鑰，以及為 Oracle 資料庫開啟透明資料加密 (TDE)。AWS CloudHSM 用戶端 SDK 提供軟體，可讓您在將資料上傳到 AWS 之前，使用儲存在 AWS CloudHSM 內的金鑰加密資料用戶端。Amazon DynamoDB Encryption Client 還允許您在上傳到 DynamoDB 資料表之前，加密和簽署項目。

實作步驟

- 為新的 [Amazon EBS 磁碟區設定預設加密](#)：使用 AWS 提供的預設金鑰或您自行建立的金鑰，指定您希望以加密形式建立所有新的 Amazon EBS 磁碟區。
- 設定加密的 Amazon Machine Images (AMI)：複製已設定加密的現有 AMI 會自動加密根磁碟區和快照。
- 設定 [Amazon RDS 加密](#)：透過使用加密選項，為您的 Amazon RDS 資料庫叢集和靜態快照設定加密。
- 使用政策限制對適當主體的存取，為每個資料分類建立和設定 AWS KMS 金鑰：例如，建立一個 AWS KMS 金鑰用於加密生產資料，另一個金鑰用於加密開發或測試資料。您還可以提供金鑰來存取其他 AWS 帳戶。考慮針對開發和生產環境擁有不同的帳戶。如果您的生產環境需要解密開發帳戶中的成品，您可以編輯用來加密開發成品的 CMK 金鑰，使生產帳戶能夠解密這些成品。生產環境接著可以擷取解密的資料以用於生產。
- 在其他 AWS 服務中設定加密：對於您使用的其他 AWS 服務，請檢閱該服務的[安全文件](#)，以確定該服務的加密選項。

資源

相關文件：

- [AWS 加密工具](#)
- [AWS Encryption SDK](#)
- [AWS KMS 加密詳細資訊白皮書](#)
- [AWS Key Management Service](#)
- [AWS 加密服務和工具](#)
- [Amazon EBS 加密](#)
- [Amazon EBS 磁碟區的預設加密](#)
- [加密 Amazon RDS 資源](#)
- [如何啟用 Amazon S3 儲存貯體的預設加密？](#)
- [使用加密保護 Amazon S3 資料](#)

相關影片：

- [在 AWS 中加密如何運作](#)
- [保護 AWS 上的區塊儲存安全](#)

SEC08-BP03 自動化靜態資料保護

使用自動化來驗證和強制執行靜態資料控制。使用自動掃描來偵測資料儲存解決方案的錯誤組態，並盡可能透過自動化的程式化回應執行補救措施。將自動化納入 CI/CD 程序中，以在部署到實際執行環境之前先偵測是否有資料儲存組態錯誤的情形。

預期成果：自動化系統會掃描和監控資料儲存位置，找出是否有控制組態錯誤、未經授權存取及意外使用的情況。偵測到設定錯誤的儲存位置就會啟動自動化補救措施。自動化程序會建立資料備份，並將不可變的副本儲存在原始環境之外。

常見的反模式：

- 未考慮在受支援的情況下，啟用預設加密設定的選項。
- 制定自動備份和復原策略時，未考慮安全事件還有操作事件。
- 未強制執行儲存服務的公開存取設定。

- 未監控和稽核保護靜態資料的控制。

建立此最佳實務的優勢：自動化有助於防止發生資料儲存位置設定錯誤的風險。這有助於防止錯誤組態進入您的實際執行環境。此最佳實務也有助於偵測並修正錯誤組態 (如發生)。

未建立此最佳實務時的曝險等級：中

實作指引

自動化是貫穿保護靜態資料的實務的主題。[SEC01-BP06 自動部署標準安全控制](#)說明了如何使用基礎設施即程式碼 (IaC) 範本 (例如使用 [AWS CloudFormation](#)) 擷取資源的組態。這些範本已遞交至版本控制系統中，且用於在 AWS 上透過 CI/CD 管道部署資源。這些技術同樣適用於自動化資料儲存解決方案的組態，例如 Amazon S3 儲存貯體上的加密設定。

您可以在 CI/CD 管道中使用 [AWS CloudFormation Guard](#) 內的規則檢查您在 IaC 範本中定義的設定是否有組態錯誤。您可以透過 [AWS Config](#) 監控尚未在 CloudFormation 或其他 IaC 工具中提供的設定，以檢查是否有組態錯誤。Config 針對錯誤組態所產生的提醒可以自動修復，如 [SEC04-BP04 針對不合規資源實施補救措施](#)中所述。

將自動化納入您的許可管理策略中，也是整體自動化資料防護措施的一環。[SEC03-BP02 授予最低權限存取權](#)和 [SEC03-BP04 持續減少許可](#)中說明了如何設定最低權限存取權政策，這些政策會受到 [AWS Identity and Access Management Access Analyzer](#) 的持續監控，以在能夠減少許可時產生調查結果。除了自動化監控許可之外，您還可以設定 [Amazon GuardDuty](#) 來監看 [EBS 磁碟區](#) (藉由 EC2 執行個體)、[S3 儲存貯體](#)和支援的 [Amazon Relational Database Service 資料庫](#)中是否存在異常資料存取行為。

自動化也會在偵測到敏感資料儲存於未經授權的位置時，發揮重要的作用。[SEC07-BP03 自動識別和分類](#)說明了 [Amazon Macie](#) 如何監控您的 S3 儲存貯體是否有非預期的敏感資料，並產生可啟動自動化回應的提醒。

遵循 [REL09 備份資料](#)中的實務，制定自動化資料備份和復原策略。資料備份和復原對於操作事件，以及從安全事件中復原來說都相當重要。

實作步驟

1. 在 IaC 範本中擷取資料儲存組態。使用自動化檢查在 CI/CD 管道中偵測組態錯誤。
 - a. 您可以將 [CloudFormation](#) 用於您的 IaC 範本，並且將 [CloudFormation Guard](#) 用於檢查範本是否有組態錯誤。
 - b. 使用 [AWS Config](#) 在主動評估模式下執行規則。在建立資源之前，使用此設定做為 CI/CD 管道中的步驟來檢查資源是否合規。

2. 監控資源是否有資料儲存組態錯誤。
 - a. 設定 [AWS Config](#) 來監控資料儲存資源是否有控制組態方面的變更，並在偵測到組態錯誤時產生提醒，以調用修復動作。
 - b. 如需有關自動化補救措施的詳細指引，請參閱 [SEC04-BP04 針對不合規資源實施補救措施](#)。
3. 透過自動化持續監控並減少資料存取許可。
 - a. [IAM Access Analyzer](#) 可持續執行，並在有可能減少許可時產生提醒。
4. 監控並提醒異常資料存取行為。
 - a. [GuardDuty](#) 會監看 EBS 磁碟區、S3 儲存貯體及 RDS 資料庫等資料儲存資源中，是否存在已知的威脅特徵和偏離基準存取行為。
5. 監控並在敏感資料儲存於非預期位置時發出提醒。
 - a. 使用 [Amazon Macie](#) 持續掃描您的 S3 儲存貯體是否有敏感資料。
6. 自動保護和加密資料備份。
 - a. [AWS Backup](#) 是受管服務，會在 AWS 上建立各種資料來源的加密和安全備份。[彈性災難復原](#) 可讓您透過以秒為單位測量的復原點目標 (RPO)，複製完整的伺服器工作負載並維持持續的資料保護。您可以設定讓兩種服務搭配運作，以自動建立資料備份並將其複製到容錯移轉位置。這有助於在受到操作或安全事件影響時，保持資料的可用性。

資源

相關的最佳實務：

- [SEC01-BP06 自動部署標準安全控制](#)
- [SEC03-BP02 授予最低權限存取權](#)
- [SEC03-BP04 持續減少許可](#)
- [SEC04-BP04 針對不合規資源實施補救措施](#)
- [SEC07-BP03 自動識別和分類](#)
- [REL09-BP02 保護和加密備份](#)
- [REL09-BP03 自動執行資料備份](#)

相關文件：

- [AWS 方案指引：自動加密現有和新的 Amazon EBS 磁碟區](#)
- [在 AWS 上使用 NIST 網路安全架構 \(CSF\) 進行勒索軟體風險管理](#)

相關範例：

- [如何使用 AWS Config 主動式規則和 AWS CloudFormation 勾點來防止建立不合規的雲端資源](#)
- [使用 AWS Backup 自動化並集中管理 Amazon S3 的資料保護](#)
- [AWS re:Invent 2023 - 使用 Amazon EBS 快照實作主動式資料保護](#)
- [AWS re:Invent 2022 - 利用現代化資料保護建置並自動化以強化恢復能力](#)

相關工具：

- [AWS CloudFormation Guard](#)
- [AWS CloudFormation Guard 規則登錄檔](#)
- [IAM Access Analyzer](#)
- [Amazon Macie](#)
- [AWS Backup](#)
- [彈性災難復原](#)

SEC08-BP04 強制存取控制

要協助保護您的靜態資料，使用隔離和版本控制等機制來強制存取控制。套用最低權限和條件存取控制。避免授予您資料的公開存取權。

預期成果：確認只有授權使用者才能在必要時存取資料。您透過定期備份和版本控制來保護您的資料，以防有意或不慎修改或刪除資料。您將重要資料與其他資料分離，以保護其機密性和資料完整性。

常見的反模式：

- 將具有不同敏感度需求或分類的資料儲存在一起。
- 對解密金鑰使用過於寬鬆的許可。
- 資料分類不當。
- 未保留重要資料的詳細備份。
- 對生產資料提供持續存取權。
- 未稽核資料存取或定期審查許可。

未建立此最佳實務時的曝險等級：高

實作指引

保護靜態資料對於維護資料完整性、機密性和符合法規要求至關重要。您可以實作多個控制項來協助達成此目標，包括存取控制、隔離、條件式存取和版本控制。

您可以使用最低權限原則強制執行存取控制，該原則僅為使用者和服務提供執行其任務所需的許可權。這包括加密金鑰的存取權。檢閱您的 [AWS Key Management Service \(AWS KMS\) 政策](#)，確認您授予的存取權層級適當，並且適用相關條件。

您可以針對每個層級使用各自不同的 AWS 帳戶，根據不同的分類層級分隔資料，並使用 [AWS Organizations](#) 管理這些帳戶。此隔離有助於預防未經授權的存取，並將資料揭露風險降到最低。

定期審查 S3 儲存貯體政策中授予的存取權層級。除非絕對必要，否則避免使用可公開讀取或可寫入的儲存貯體。考慮使用 [AWS Config](#) 偵測公開可用的儲存貯體，以及使用 Amazon CloudFront 從 Amazon S3 提供內容。確認不允許公開存取的儲存貯體，已正確設定為防止公開存取。

針對存放在 Amazon S3 中的關鍵資料實作版本控制和物件鎖定機制。[Amazon S3 版本控制](#) 會保留舊版物件，以從意外刪除或覆寫中復原資料。[Amazon S3 物件鎖定](#) 為物件提供強制性存取控制功能，此功能可預防遭到刪除或覆寫，即使根使用者也一樣，直到鎖定到期為止。此外，[Amazon Glacier Vault Lock](#) 可為儲存在 Amazon Glacier 中的封存提供類似的功能。

實作步驟

1. 使用最低權限原則強制執行存取控制：

- 檢閱授予使用者和服務的存取權限，並確認他們只擁有執行任務所需的許可權。
- 檢查 [AWS Key Management Service \(AWS KMS\) 政策](#) 以檢閱加密金鑰的存取權。

2. 根據不同的分類層級分隔資料：

- 針對每個資料分類層級使用不同的 AWS 帳戶。
- 使用 [AWS Organizations](#) 管理這些帳戶。

3. 檢閱 Amazon S3 儲存貯體和物件許可權：

- 定期審查 S3 儲存貯體政策中授予的存取權層級。
- 除非絕對必要，否則避免使用可公開讀取或可寫入的儲存貯體。
- 考慮使用 [AWS Config](#) 偵測公開可用的儲存貯體。
- 使用 Amazon CloudFront 提供來自 Amazon S3 的內容。
- 確認不允許公開存取的儲存貯體，已正確設定為防止公開存取。
- 您可以對資料庫和使用 IAM 身分驗證的其他任何資料來源，套用相同的審查程序，例如 SQS 或第三方資料存放區。

4. 使用 AWS IAM Access Analyzer :

- 您可以設定讓 [AWS IAM Access Analyzer](#) 分析 Amazon S3 儲存貯體，並在 S3 政策將存取權授予外部實體時產生調查結果。

5. 實作版本控制和物件鎖定機制：

- 使用 [Amazon S3 版本控制](#) 保留物件的舊版本，如此即可將意外刪除或覆寫的資料復原。
- 使用 [Amazon S3 物件鎖定](#) 為物件提供強制存取控制，如此可避免遭到刪除或覆寫，即使根使用者也一樣，直到鎖定功能到期為止。
- 針對儲存在 Amazon Glacier 中的封存使用 [Amazon Glacier Vault Lock](#)。

6. 使用 Amazon S3 庫存清單

- 使用 [Amazon S3 庫存清單](#) 稽核及報告 S3 物件的複寫和加密狀態。

7. 檢閱 Amazon EBS 和 AMI 共用權限：

- 檢閱 [Amazon EBS](#) 和 [AMI 共用](#) 權限，確認您的映像和磁碟區未與在工作負載外部的 AWS 帳戶共用。

8. 定期審查 AWS Resource Access Manager Shares：

- 您可以使用 [AWS Resource Access Manager](#) 共用 Amazon VPC 內的資源，例如 AWS Network Firewall 政策、Amazon Route 53 Resolver 規則和子網路。
- 定期稽核共用的資源，並停止共用不再需要共用的資源。

資源

相關的最佳實務：

- [SEC03-BP01 定義存取需求](#)
- [SEC03-BP02 授予最低權限存取權](#)

相關文件：

- [AWS KMS 加密詳細資訊白皮書](#)
- [管理對 Amazon S3 資源的存取權限的簡介](#)
- [管理 AWS KMS 資源存取的概觀](#)
- [AWS Config 規則](#)
- [Amazon S3 + Amazon CloudFront：雲端中的最佳拍檔](#)
- [使用版本控制](#)

- [使用 Amazon S3 Object Lock 鎖定物件](#)
- [共用 Amazon EBS 快照](#)
- [共用 AMI](#)
- [在 Amazon S3 上託管單頁應用程式](#)
- [AWS 全域條件索引鍵](#)
- [在 AWS 上建置資料周邊](#)

相關影片：

- [保護 AWS 上的區塊儲存安全](#)

SEC 9. 如何保護傳輸中的資料？

實作多重控制來保護傳輸中的資料，以降低未經授權存取或遺失的風險。

最佳實務

- [SEC09-BP01 實作安全金鑰和憑證管理](#)
- [SEC09-BP02 強制傳輸中加密](#)
- [SEC09-BP03 驗證網路通訊](#)

SEC09-BP01 實作安全金鑰和憑證管理

Transport Layer Security (TLS) 憑證可用來保護網路通訊，和建立網際網路跟私有網路中的網站、資源和工作負載的身分。

預期成果：可以在公開金鑰基礎設施 (PKI) 中佈建、部署、儲存和更新憑證的安全憑證管理系統。安全金鑰與憑證管理機制可以防止憑證私有金鑰資料外洩，也能定期自動更新憑證。它也能與其他服務整合，為工作負載內的機器資源提供安全的網路通訊和身分識別。金鑰資料絕不可供真人身分存取。

常見的反模式：

- 在憑證部署或更新程序期間執行手動步驟。
- 設計私有 CA 時，忽略憑證授權單位 (CA) 階層。
- 針對公有資源使用自我簽署憑證。

建立此最佳實務的優勢：

- 透過自動化部署和更新來簡化憑證管理
- 鼓勵使用 TLS 憑證加密傳輸中的資料
- 增加憑證授權機構所採取憑證動作的安全性和可稽核性
- 在 CA 階層中不同層次的管理責任組織

未建立此最佳實務時的曝險等級：高

實作指引

現代工作負載可透過使用 TLS 等 PKI 協定，來廣泛使用加密網路通訊。PKI 憑證管理可能很複雜，但自動化憑證佈建、部署和更新可以減少憑證管理相關障礙。

AWS 提供兩種管理一般用途 PKI 憑證的服務：[AWS Certificate Manager](#) 和 [AWS 私有憑證授權單位 \(AWS 私有 CA\)](#)。ACM 是客戶在公有和私有 AWS 工作負載中，用來佈建、管理和部署憑證的主要服務。ACM 則使用 AWS 私有 CA 發行私有憑證，並與許多其他 AWS 受管服務[整合](#)，以提供安全的工作負載 TLS 憑證。ACM 也可以從 [Amazon Trust Services](#) 發出公有信任憑證。ACM 的公有憑證可用於面向公眾的工作負載，因為新型瀏覽器和作業系統預設信任這些憑證。

AWS 私有 CA 可讓您建立自己的根憑證授權單位或下層憑證授權單位，並透過 API 發行 TLS 憑證。在您控制和管理 TLS 連線用戶端信任鏈時，您可以使用這些憑證類型。除了 TLS 使用案例之外，AWS 私有 CA 還可以用來發行憑證給 Kubernetes Pod、重要裝置產品證明、程式碼簽署，以及其他具有[自訂範本](#)的使用案例。您也可以使用 [IAM Roles Anywhere](#) 提供暫時 IAM 憑證給具有您私有 CA 發行 X.509 憑證的內部部署工作負載。

除了 ACM 和 AWS 私有 CA 之外，[AWS IoT Core](#) 也為 IoT 裝置提供特別支援，用來佈建、管理和部署 PKI 憑證。AWS IoT Core 提供了專門的機制，可[將 IoT 裝置大規模加入](#)您的公有金鑰基礎設施。

有些 AWS 服務如 [Amazon API Gateway](#) 和 [Elastic Load Balancing](#)，會提供使用憑證來保護應用程式連線的自有功能。例如，API Gateway 和 Application Load Balancer (ALB) 都支援交互 TLS (mTLS)，其便是使用您使用 AWS 管理主控台、CLI 或 API 建立和匯出的用戶端憑證。

建立私有 CA 階層時的考量

在需要建立私有 CA 時，請務必特別留意，預先正確設計 CA 階層。建立私有 CA 階層時，最佳實務是將 CA 階層的每個層級部署到個別 AWS 帳戶。這個刻意的步驟會減少 CA 階層中每個層級的界面積，讓您更容易察覺 CloudTrail 日誌資料的異常狀況，並在出現未經授權的帳戶存取動作時降低存取或影響範圍。根 CA 應位於自己的個別帳戶中，且只能用來發行一個或多個中繼 CA 憑證。

接著，請在不同於根 CA 帳戶的其他帳戶中建立一個或多個中繼 CA，為終端使用者、裝置或其他工作負載發行憑證。最後，請將憑證從根 CA 發行至中繼 CA，這個動作會將憑證發行給您的終端使用者或

裝置。如需深入了解 CA 部署規劃和 CA 階層設計，包括恢復能力、跨區域複寫、在組織中共用 CA 等規劃，請參閱[規劃您的 AWS 私有 CA 部署](#)。

實作步驟

1. 確定使用案例所需的相關 AWS 服務：

- 許多使用案例都可以使用 [AWS Certificate Manager](#) 來利用現有的 AWS 公有金鑰基礎設施。ACM 可以用於為 Web 伺服器、負載平衡器或公開信任憑證的其他用途部署 TLS 憑證。
- 當您需要建立自己的私有憑證授權機構階層或需要存取可匯出的憑證時，考慮 [AWS 私有 CA](#)。然後，ACM 可用於使用 AWS 私有 CA 發行[許多類型的終端實體憑證](#)。
- 針對必須為嵌入式物聯網 (IoT) 裝置大規模佈建憑證的使用案例，請考慮 [AWS IoT Core](#)。
- 請考慮在 [Amazon API Gateway](#) 或 [Application Load Balancer](#) 等服務中使用原生 mTLS 功能。

2. 盡可能實作自動憑證續約：

- 使用 [ACM 受管續約](#) ACM 發行的憑證以及整合式 AWS 受管服務。

3. 建立日誌記錄和稽核記錄：

- 啟用 [CloudTrail 日誌](#)以追蹤對持有憑證授權機構之帳戶的存取權。請考慮在 CloudTrail 中設定日誌檔案完整性驗證，以驗證日誌資料的真實性。
- 定期產出並審核[稽核報告](#)，其中列出您的私有 CA 發行或撤銷的憑證。這些報告可以匯出到 S3 儲存貯體。
- 部署私有 CA 時，您也需要建立 S3 儲存貯體來儲存憑證撤銷清單 (CRL)。如需有關根據工作負載需求設定此 S3 儲存貯體的指引，請參閱[規劃憑證撤銷清單 \(CRL\)](#)。

資源

相關的最佳實務：

- [SEC02-BP02 使用暫時憑證](#)
- [SEC08-BP01 實作安全金鑰管理](#)
- [SEC09-BP03 驗證網路通訊](#)

相關文件：

- [如何在 AWS 中託管和管理整個私有憑證基礎設施](#)
- [如何保護用於汽車和製造領域的企業規模 ACM 私有 CA 階層](#)
- [私有 CA 最佳實務](#)

- [如何使用 AWS RAM 共用您的 ACM 私有 CA 跨帳戶](#)

相關影片：

- [啟動 AWS Certificate Manager 私有 CA \(研討會\)](#)

相關範例：

- [私有 CA 研討會](#)
- [IOT Device Management 研討會](#) (包括裝置佈建)

相關工具：

- [要使用 AWS 私有 CA 的 Kubernetes cert-manager 外掛程式](#)

SEC09-BP02 強制傳輸中加密

根據您組織的政策、法規義務和標準強制已定義的加密需求，以符合組織、法律和合規上的要求。只有在虛擬私有雲端 (VPC) 以外傳輸敏感資料時才使用加密協定。加密有助於保持資料完整性，甚至當資料傳輸於不受信任的網路。

預期成果：您將資源與網際網路之間的網路流量加密，以減少對資料的未經授權存取。您可以根據您的安全要求，加密內部 AWS 環境中的網路流量。您使用安全的 TLS 協定和密碼套件，為所有傳輸中的資料加密。

常見的反模式：

- 使用 SSL、TLS 和其他套件元件已棄用的版本 (例如，SSL v3.0、1024 位元 RSA 金鑰和 RC4 密碼)。
- 允許未加密的 (HTTP) 流量來往面向公眾的資源。
- 未監控 X.509 憑證並在到期前更換。
- 對 TLS 使用自我簽署的 X.509 憑證。

未建立此最佳實務時的曝險等級：高

實作指引

AWS 服務提供使用 TLS 的 HTTPS 端點以進行通訊，在與 AWS API 通訊時提供傳輸中加密。不安全的 HTTP 通訊協定可以在虛擬私有雲端 (VPC) 中，透過使用安全群組來進行稽核和封鎖。在 Amazon CloudFront 或 [Application Load Balancer](#) 中，HTTP 請求也可以[自動重新導向至 HTTPS](#)。您可以使用 [Amazon Simple Storage Service \(Amazon S3\) 儲存貯體政策](#)來限制透過 HTTP 上傳物件的能力，並有效地強制執行使用 HTTPS，將物件上傳至儲存貯體。您可以全權控制您的運算資源，以在各個服務中實作傳輸中加密。此外，還可以從外部網路或 [AWS Direct Connect](#) 使用 VPN 連線功能進入 VPC，加速流量加密。確認您的用戶端使用至少 TLS 1.2 對 AWS APIs 進行呼叫，因為 [AWS TLS 的早期版本已於 2024 年 2 月棄用](#)。建議您使用 TLS 1.3。如果您有傳輸中加密的特殊要求，您可以在 AWS Marketplace 中找到可用的第三方解決方案。

實作步驟

- 強制傳輸中加密：您定義的加密要求應符合最新標準和最佳實務，並僅允許採用安全協定。例如，設定安全群組，僅允許 HTTPS 協定連至 Application Load Balancer 或 Amazon EC2 執行個體。
- 在邊緣服務中設定安全協定：[使用 Amazon CloudFront 設定 HTTPS](#)，並使用[適用於您的安全狀態和使用案例的安全設定檔](#)。
- 使用 [VPN 進行外部連線](#)：考慮使用 IPsec VPN，保護點對點或網路對網路連線，以協助提供資料隱私和完整性。
- 在負載平衡器中設定安全協定：選取安全政策，以提供要連接到接聽程式的用戶端所支援的最強固的密碼套件。[為 Application Load Balancer 建立 HTTPS 接聽程式](#)。
- 在 Amazon Redshift 中設定安全協定：將您的叢集設定為要求 [Secure Socket Layer \(SSL\) 或 Transport Layer Security \(TLS\) 連線](#)。
- 設定安全協定：檢閱 AWS 服務文件以確定傳輸中加密功能。
- 設定上傳至 Amazon S3 儲存貯體時的安全存取：使用 Amazon S3 儲存貯體政策控制對資料[強制安全存取](#)。
- 考慮使用 [AWS Certificate Manager](#)：ACM 可讓您佈建、管理和部署 TLS 憑證與 AWS 服務搭配使用。
- 考慮使用 [AWS 私有憑證授權單位](#) 滿足私有 PKI 需求：AWS 私有 CA 可讓您建立私有憑證授權機構 (CA) 階層來發行可用於建立已加密 TLS 通道的終端實體 X.509 憑證。

資源

相關文件：

- [將 HTTPS 與 CloudFront 搭配使用](#)

- [使用 AWS Virtual Private Network 將 VPC 連線至遠端網路](#)
- [為 Application Load Balancer 建立 HTTPS 接聽程式](#)
- [教學課程：在 Amazon Linux 2 上設定 SSL/TLS](#)
- [使用 SSL/TLS 加密與資料庫執行個體的連線](#)
- [設定連線的安全選項](#)

SEC09-BP03 驗證網路通訊

使用支援身分驗證的協定 [如 Transport Layer Security (TLS) 或 IPsec] 來驗證通訊的身分。

設計工作負載，以在每當服務、應用程式或使用者之間進行通訊時，使用安全、經驗證的網路協定。使用支援驗證和授權的網路協定可提供更強大的網路流量控制能力，並減少未經授權存取所造成的影響。

預期成果：設計出工作負載，讓其有明確定義的服務間資料平面和控制平面流量流程。在技術允許的情況下，流量流程要使用經過驗證和加密的網路協定。

常見的反模式：

- 工作負載內有未經加密或驗證的流量流程。
- 在多個使用者或實體之間重複使用驗證憑證。
- 僅仰賴網路控制做為存取控制機制。
- 建立自訂驗證機制，而非仰賴產業標準的驗證機制。
- 服務元件或 VPC 中的其他資源之間有過於寬鬆的流量。

建立此最佳實務的優勢：

- 將未經授權存取所造成的影響範圍限制在工作負載的某個部分。
- 提供只會由已驗證實體執行動作的更高層級保證。
- 透過清楚地定義並強制執行預期的資料傳輸介面來改善服務的去耦。
- 透過請求歸因和明確定義的通訊介面，增強監控、日誌記錄和事件回應。
- 結合網路控制與驗證和授權控制，為您的工作負載提供深度防禦。

未建立此最佳實務時的曝險等級：低

實作指引

您工作負載的網路流量模式可分為兩個類別：

- 東西流量代表構成工作負載的服務之間的流量流程。
- 南北流量代表工作負載和取用者之間的流量流程。

雖然加密南北流量是常見的做法，但是使用經過驗證的協定來保護東西流量則較不常見。現代安全實務的建議是，單靠網路設計並無法讓兩個實體之間建立信任的關係。當兩個服務可能位於一個共通的網路邊界內時，最佳實務仍是對這些服務之間的通訊進行加密、驗證和授權。

舉例來說，無論請求來自哪個網路，AWS 服務 API 都會使用 [AWS 第 4 版簽署程序 \(SigV4\)](#) 簽署協定來驗證呼叫者。此驗證可確保 AWS API 可以驗證發出動作請求的身分，該身分接著可與政策結合來作出授權決策，確定是否應允許該動作。

[Amazon VPC Lattice](#) 和 [Amazon API Gateway](#) 等服務可讓您使用相同的 SigV4 簽署協定，為自己的工作負載中的東西流量新增驗證和授權功能。如果 AWS 環境以外的資源需要與要求進行 SigV4 型驗證和授權的服務進行通訊，您可以在非 AWS 資源上使用 [AWS Identity and Access Management \(IAM\) Roles Anywhere](#) 來取得暫時 AWS 憑證。使用這些憑證，便可透過 SigV4 簽署服務請求以授權存取。

用於驗證東西流量的另一種常見機制是 TLS 相互驗證 (mTLS)。許多物聯網 (IoT)、企業對企業應用程式和微型服務都使用 mTLS，透過使用用戶端和伺服器端 X.509 憑證來驗證 TLS 通訊兩端的身分。這些憑證可由 AWS 私有憑證授權單位 (AWS 私有 CA) 發行。您可以使用 [Amazon API Gateway](#) 等服務，為工作負載之間或工作負載內部的通訊提供 mTLS 驗證。[Application Load Balancer](#) 也支援內部或外部工作負載的 mTLS。雖然 mTLS 會為 TLS 通訊的兩端提供驗證資訊，但不提供授權機制。

最後，OAuth 2.0 和 OpenID Connect (OIDC) 是兩種常用於控制使用者對服務存取行為的協定，但現在也變成服務對服務流量的熱門協定。API Gateway 會提供 [JSON Web Token \(JWT\) 授權器](#)，可讓工作負載使用 OIDC 或 OAuth 2.0 身分提供者所發行的 JWT 來限制 API 路由的存取。OAuth2 的範圍可做為基本授權決策的來源，但仍需要在應用程式層實作授權檢查，而且單靠 OAuth2 範圍並無法支援更複雜的授權需求。

實作步驟

- 定義並記錄您的工作負載網路流程：實作深度防禦策略的第一步是定義工作負載的流量流程。
 - 建立可清楚定義構成工作負載的不同服務間資料傳輸方式的資料流程圖。此圖是透過已驗證的網路通道強制執行這些流程的第一步。
 - 在開發和測試階段檢測您的工作負載，以驗證資料流程圖是否準確反映工作負載在執行時期的行為。
 - 資料流程圖在執行威脅建模練習時也很有用，如 [SEC01-BP07 使用威脅模型識別威脅並優先考慮緩解措施](#) 中所述。

- 建立網路控制：考慮用來建立與資料流程一致的網路控制的 AWS 功能。雖然網路邊界不應成為唯一的安全控制，但其可在深度防禦策略中提供一個保護層，以保護您的工作負載。
- 使用[安全群組](#)建立定義和限制資源之間的資料流程。
- 考慮使用 [AWS PrivateLink](#) 與支援 AWS PrivateLink 的 AWS 和第三方服務進行通訊。透過 AWS PrivateLink 介面端點傳送的資料會保留在 AWS 網路骨幹內，不會周遊公用網際網路。
- 在工作負載中跨服務實作驗證和授權：選擇最適合用來在您工作負載中提供經驗證加密流量的 AWS 服務集。
- 考慮用來保護服務對服務通訊的 [Amazon VPC Lattice](#)。VPC Lattice 可以使用 [SigV4 驗證結合驗證政策](#)來控制服務對服務的存取。
- 對於使用 mTLS 的服務對服務通訊，請考慮 [API Gateway](#)、[Application Load Balancer](#)。 [AWS 私有 CA](#) 可用於建立能夠發行憑證以與 mTLS 搭配使用的私有 CA 階層。
- 與使用 OAuth 2.0 或 OIDC 的服務進行整合時，請考慮[使用 JWT 授權器的 API Gateway](#)。
- 對於工作負載和 IoT 裝置之間的通訊，請考慮 [AWS IoT Core](#)，它提供了幾種網路流量加密和驗證選項。
- 監控未經授權的存取：持續監控是否有意外的通訊管道、嘗試存取受保護資源的未經授權主體，以及其他不當的存取模式。
- 如果使用 VPC Lattice 來管理服務的存取，請考慮啟用和監控 [VPC Lattice 存取日誌](#)。這些存取日誌包括請求方實體的資訊、包括來源和目的地 VPC 在內的網路資訊，以及請求中繼資料。
- 考慮啟用 [VPC 流程日誌](#)來擷取網路流程上的中繼資料，並定期審查是否有異常狀況。
- 如需更多有關規劃、模擬和應對安全事件的指引，請參閱 [AWS 安全事件回應指南](#)和 AWS Well Architected Framework 安全支柱的[事件回應](#)章節。

資源

相關的最佳實務：

- [SEC03-BP07 分析公有和跨帳戶存取權](#)
- [SEC02-BP02 使用臨時憑證](#)
- [SEC01-BP07 使用威脅模型識別威脅並優先考慮緩解措施](#)

相關文件：

- [評估用來保護 Amazon API Gateway API 的存取控制方法](#)
- [為 REST API 設定相互 TLS 驗證](#)

- [如何使用 JWT 授權器保護 API Gateway HTTP 端點](#)
- [使用 AWS IoT Core 憑證提供者來授權 AWS 服務的直接呼叫](#)
- [AWS 安全事件回應指南](#)

相關影片：

- [AWS re:invent 2022：VPC Lattice 簡介](#)
- [AWS re:invent 2020：針對 AWS 上 HTTP API 的無伺服器 API 驗證](#)

相關範例：

- [Amazon VPC Lattice 研討會](#)
- [零信任第 1 集 - Phantom Service Perimeter 研討會](#)

事件反應

問題

- [SEC 10. 如何預測、回應事故以及從事故中復原？](#)

SEC 10. 如何預測、回應事故以及從事故中復原？

即使採用了成熟的預防和偵測控制，您的組織仍應實作機制，以回應並緩和安全事故的潛在影響。您的準備工作會大大地影響團隊在事故發生時能否有效運作，以隔離、遏制問題並進行鑑識，以及將營運恢復到已知的良好狀態。在安全事故發生前先備妥工具和存取權，然後在演練日期間定期練習事故回應，有助於確保您能夠復原，同時盡量減少業務中斷。

最佳實務

- [SEC10-BP01 識別關鍵人員和外部資源](#)
- [SEC10-BP02 制定事件管理計畫](#)
- [SEC10-BP03 準備鑑識功能](#)
- [SEC10-BP04 開發和測試安全事件回應程序手冊](#)
- [SEC10-BP05 預先佈建存取權](#)
- [SEC10-BP06 預先部署工具](#)

- [SEC10-BP07 執行模擬](#)
- [SEC10-BP08 建立從事件中學習的架構](#)

SEC10-BP01 識別關鍵人員和外部資源

識別可以協助您的組織回應事件的內部和外部人員、資源及法律義務。

預期成果：您備有一份關鍵人員名單，包含其聯絡資訊，以及他們在回應安全事件時扮演的角色。您可定期審查並更新此資訊，以在內部和外部工具中反映人員變更。您在記錄此資訊時考量所有第三方服務供應商和廠商，包括安全合作夥伴、雲端供應商，以及軟體即服務 (SaaS) 應用程式。在安全事件期間，人員會負起適當層級的責任、得知適當的關聯內容，並擁有適當的存取權能夠做出回應和進行復原。

常見的反模式：

- 回應安全事件時，未隨時備妥包含聯絡資訊、角色和職責的最新關鍵人員名單。
- 假設每個人都了解回應事件和從事件復原時的人員、相依關係、基礎設施和解決方案。
- 沒有能夠呈現主要基礎設施或應用程式設計的文件或知識儲存庫。
- 未制定適當的新員工上任流程，導致他們無法有效地參與安全事件回應工作，例如進行事件模擬。
- 在安全事件期間，當關鍵人員暫時聯絡不到或無法回應時，沒有向上呈報的管道。

建立此最佳實務的優勢：此實務可減少發生事件時，花在識別正確人員和其角色的權責劃分和回應時間。隨時備妥一份最新的關鍵人員及其角色的名單，您就能找到正確的人員進行權責劃分並從事件中復原，藉此在發生事件時盡量減少時間浪費。

未建立此最佳實務時的風險暴露等級：高

實作指引

識別組織中的關鍵人員：隨時備妥組織內應對特定事件所需的人員聯絡名單。發生人員變動 (例如組織變動、晉升和團隊變動) 時，定期審查並更新此資訊。這對於事件管理者、事件回應者和通訊負責人等關鍵角色尤其重要。

- **事件管理者：**事件管理者在事件回應期間具有整體授權。
- **事件回應者：**事件回應者負責調查和補救活動。這些人員可能根據事件類型而有所不同，但通常是負責受影響應用程式的開發人員和營運團隊。
- **通訊負責人：**通訊負責人負責內部和外部溝通，特別是與公家機關、監管機構和客戶之間的溝通。

- 上線程序：定期培訓新進員工，並讓他們上線操作，讓他們具備必要的技能和知識，以有效地為事件回應工作做出貢獻。將模擬和實作練習納入以當成上線程序的一部分，以催促他們做好準備
- 主題專家 (SME)：如果是分散式和自主團隊，我們建議您組成 SME 團隊來負責關鍵任務工作負載。他們負責提供對事件所涉及關鍵工作負載的操作和資料分類的深入洞悉。

資料表格式範例：

```
| Role | Name | Contact Information | Responsibilities |
1 | --- | --- | --- | --- |
2 | Incident Manager | Jane Doe | jane.doe@example.com | Overall authority during
  response |
3 | Incident Responder | John Smith | john.smith@example.com | Investigation and
  remediation |
4 | Communications Lead | Emily Johnson | emily.johnson@example.com | Internal and
  external communications |
5 | Communications Lead | Michael Brown | michael.brown@example.com | Insights on
  critical workloads |
```

請考慮使用 [AWS Systems Manager Incident Manager](#) 功能來擷取主要聯絡人、定義回應計畫、自動排定待命時間表，以及制定向上呈報計畫。自動排定待命時間表並輪替所有人員，藉此將工作負載的責任分散給其負責人。這樣有助於建立良好的實務，例如，發出相關的指標和日誌，以及定義對工作負載至關重要的警示閾值。

識別外部合作夥伴：企業使用由獨立軟體開發廠商 (ISV)、合作夥伴和子承包商建置的工具，為其客戶打造出獨特的解決方案。邀集上述多方的關鍵人員參與，他們可以協助回應事件並從中復原。我們建議您註冊適當層級的支援，以便透過支援案例迅速與 AWS 主題專家取得聯繫。請考慮針對工作負載的所有關鍵解決方案提供者做出類似的安排。有些安全事件會促使公開上市公司通知相關的公家機關和監管機構有關事件的情形和影響。維護並更新相關部門和負責人的聯絡資訊。

實作步驟

1. 設定事件管理解決方案。
 - a. 考慮在您的安全工具帳戶中部署 Incident Manager。
2. 在事件管理解決方案中定義聯絡人。
 - a. 針對每個聯絡人至少定義兩種類型的聯絡管道 (例如 SMS、電話或電子郵件)，以確保在發生事件時能夠與他們取得聯繫。
3. 定義回應計畫。

- a. 識別發生事件時最合適參與的聯絡人。定義符合要參與之人員角色 (而非個別聯絡人) 的向上呈報計畫。考慮納入可負責通知外部實體的聯絡人，即使他們並未直接參與事件解決工作。

資源

相關的最佳實務：

- [OPS02-BP03 已為營運活動識別負責其效能的擁有者](#)

相關文件：

- [AWS 安全事件回應指南](#)

相關範例：

- [AWS 客戶程序手冊架構](#)
- [準備和回應 AWS 環境中的安全事件](#)

相關工具：

- [AWS Systems Manager Incident Manager](#)

相關影片：

- [Amazon 在開發期間採取的安全方法](#)

SEC10-BP02 制定事件管理計畫

為事件回應制定的第一份文件是事件回應計畫。事件回應計畫應是您事件回應計畫和策略的基礎。

建立此最佳實務的優勢：開發全面且明確定義的事件回應程序，是成功且可擴展的事件回應計畫的關鍵。當安全事件發生時，明確的步驟和工作流程可協助您及時因應。您可能已具備現有的事件回應程序。無論您目前的狀態為何，都必須定期更新、重複執行和測試事件回應程序。

未建立此最佳實務時的曝險等級：高

實作指引

事件管理計畫對於回應、減輕安全事件所造成潛在影響並從中復原而言至關重要。事件管理計畫是結構清晰的程序，可及時識別、修復和回應安全事件。

雲端有許多在內部部署環境中所見的相同營運角色和需求。您在建立事件管理計畫時，必須將與業務成果及合規要求最相符的回應及復原策略納入考量。例如，若您在 AWS 中操作的工作負載符合美國的 FedRAMP，那麼請遵循《[NIST SP 800-61 電腦安全處理指南](#)》中的建議。同樣的，當您操作儲存個人身分識別資訊 (PII) 的工作負載時，請考慮如何保護和回應與資料落地和使用相關的問題。

為 AWS 中的工作負載建置事件管理計畫時，請從 [AWS 共同責任模式](#) 開始，以便建置事件回應的深度防禦方法。在此模型中，AWS 會管理雲端的安全，但維護雲端的安全是您的責任。此表示您保有控制權，並對您選擇實作的安全控制項負責。[AWS 安全事件回應指南](#) 詳細說明在建置以雲端為中心的事件管理計畫時的重要概念和基礎指引。

有效的事件管理計畫必須經過持續的反覆測試，以與您的雲端維運目標保持同步。在您建立和制定事件管理計畫時，請考慮使用以下詳述的實作計畫。

實作步驟

1. 定義您組織內處理安全事件的角色和責任。這應涉及各部門的代表，包括：
 - 人力資源 (HR)
 - 執行團隊
 - 法務部門
 - 應用程式擁有者和開發人員 (主題專家或 SME)
2. 明確概述事件進行期間誰應負責、當責、受諮詢和被告知 (RACI)。建立 RACI 圖表以促進快速直接的溝通，並明確概述事件不同階段的領導。
3. 在事件期間讓應用程式擁有者和開發人員 (SME) 參與，因為他們可以提供有價值的資訊和內容，有助於衡量影響範圍。與這些 SME 建立關係，並在實際事件發生之前與他們演練事件回應情境。
4. 讓受信任的合作夥伴或外部專家參與調查或回應程序，因為他們可以提供額外的專業知識和觀點。
5. 讓您的事件管理計畫和角色符合任何管理組織的當地法規或合規要求。
6. 定期練習和測試您的事件回應計畫，並納入所有定義的角色和責任。這有助於簡化程序，並確認您對於安全事件能有協調且有效的回應。
7. 定期審查和更新角色、責任和 RACI 圖表，或隨著您的組織結構或需求進行變更。

了解 AWS 回應團隊和支援

- AWS 支援
 - [支援](#) 提供各種方案，可支援您運用各種工具與專業知識，協助您的 AWS 解決方案獲得成功並正常運作。如果您需要技術支援和更多資源以協助規劃、部署和最佳化 AWS 環境，您可以選取最符合您 AWS 使用案例的支援計畫。
 - 考慮以 AWS 管理主控台 中的[支援中心](#) (需要登入) 作為中心聯絡窗口，以取得影響 AWS 資源的問題所需的支援。對 支援 的存取由 AWS Identity and Access Management 控制。如需有關取得 支援 功能存取權的詳細資訊，請參閱 [支援 入門](#)。
- AWS 客戶事件回應團隊 (CIRT)
 - AWS 客戶事件回應團隊 (CIRT) 是一個專門的全天候全球 AWS 團隊，在客戶端的有效安全事件期間為客戶提供支援 - [AWS 共同責任模式](#)。
 - 當 AWS CIRT 支援您時，他們會為 AWS 上的有效安全事件提供分類和復原方面的協助。他們可透過使用 AWS 服務日誌協助進行根本原因分析，並為您提供復原的建議。他們也可提供安全建議和最佳實務，以協助您避免事後發生安全事件。
 - AWS 客戶可以透過 [支援 案例](#) 洽詢 AWS CIRT。
- [AWS 安全事件應變](#)
 - AWS 安全事件應變 在 re:Invent 2024 發布，是一項託管安全事件回應服務，結合了現代分類技術和人機迴圈。此服務會擷取所有 GuardDuty 調查結果，以及傳送至 AWS Security Hub CSPM 進行分類的任何第三方調查結果，僅在需要調查的調查結果時提醒客戶。此服務也提供入口網站，在客戶注意到安全事件時提交反應性案例，並從 AWS 的進階事件回應團隊取得支援。
- DDoS 回應支援
 - AWS 提供 [AWS Shield](#)，其中包含受管的分散式拒絕服務 (DDoS) 保護服務，可為執行於 AWS 的 Web 應用程式提供保護。Shield 提供一律開啟的偵測和自動內嵌緩解措施，可將應用程式停機時間和延遲降至最低，讓您無需聯絡 支援 即可享有 DDoS 保護。Shield 有兩個層級：AWS Shield Standard 和 AWS Shield Advanced。若要了解這兩個層級的差異，請參閱 [Shield 功能文件](#)。
- AWS Managed Services (AMS)
 - [AWS Managed Services \(AMS\)](#) 可讓您持續管理 AWS 基礎設施，讓您專注在自己的應用程式上。透過實作維護基礎設施的最佳實務，AMS 有助於降低營運開銷和風險。AMS 可自動化常見的活動，例如，變更請求、監控、修補程式管理、安全性和備份服務，而且提供佈建、執行和支援基礎設施的完整生命週期服務。
 - AMS 負責部署安全偵測控制套件，並提供全年無休的第一線提醒應變措施。提醒啟動時，AMS 會依照一組標準的自動化和手動程序手冊來驗證回應的一致性。這些程序手冊會在上線期間與 AMS 客戶共享，讓他們能夠透過 AMS 來制定和協調應變措施。

制定事件回應計畫

事件回應計畫應是您事件回應計畫和策略的基礎。事件回應計畫應納入正式文件中。事件回應計畫通常包含下列章節：

- 事件回應團隊概觀：概述事件回應團隊的目標和職能。
- 角色和責任：列出事件回應利害關係人，並詳細說明他們在事件發生時的角色。
- 通訊計畫：詳細說明聯絡資訊，以及您在事件期間要如何進行通訊。
- 備份通訊方式：最佳實務是將頻外通訊作為事件通訊的備用方法。舉例來說，AWS Wickr 就是提供安全頻外通訊通道的應用程式。
- 事件回應的階段和應採取的行動：列舉事件回應的階段 (例如偵測、分析、消除、抑制及復原)，包括要在這些階段中採取的高階動作。
- 事件嚴重性和優先順序定義：詳細說明如何分類事件的嚴重性、如何排定事件的優先順序，以及嚴重性定義對於呈報程序有何影響。

儘管不同規模和產業的公司都會有這些章節，但每個組織的事件回應計畫都是獨一無二的。您必須建立最適合貴組織的事件回應計畫。

資源

相關的最佳實務：

- [SEC04 偵測](#)

相關文件：

- [AWS 安全事件回應指南](#)
- [NIST：電腦安全事件處理指南](#)

SEC10-BP03 準備鑑識功能

在安全事件發生之前，將開發鑑識功能納入考量，以協助安全事件調查。

未建立此最佳實務時的曝險等級：中

傳統內部部署鑑識的概念適用於 AWS。如需開始在 AWS 雲端 中建置鑑識功能的重要資訊，請參閱 [AWS 雲端 中的鑑識調查環境策略](#)。

設定鑑識的環境和 AWS 帳戶 結構後，請定義在四個階段有效執行合理鑑識方法所需的技術：

- 收集：收集相關 AWS 日誌，例如 AWS CloudTrail、AWS Config、VPC 流程日誌和主機層級日誌。收集受影響 AWS 資源的快照、備份和記憶體傾印 (如果有的話)。
- 檢查：檢查透過擷取和評估相關資訊所收集的資料。
- 分析：分析收集的資料，以了解事件並從中得出結論。
- 報告：呈現分析階段所產生的資訊。

實作步驟

準備鑑識環境

[AWS Organizations](#) 可協助您隨著 AWS 資源的成長和擴展，集中管理和控管 AWS 環境。AWS 組織會合併 AWS 帳戶，以便您可以將其當作一個單位進行管理。您可以使用組織單位 (OU) 將帳戶群組在一起，以單一單位的形式進行管理。

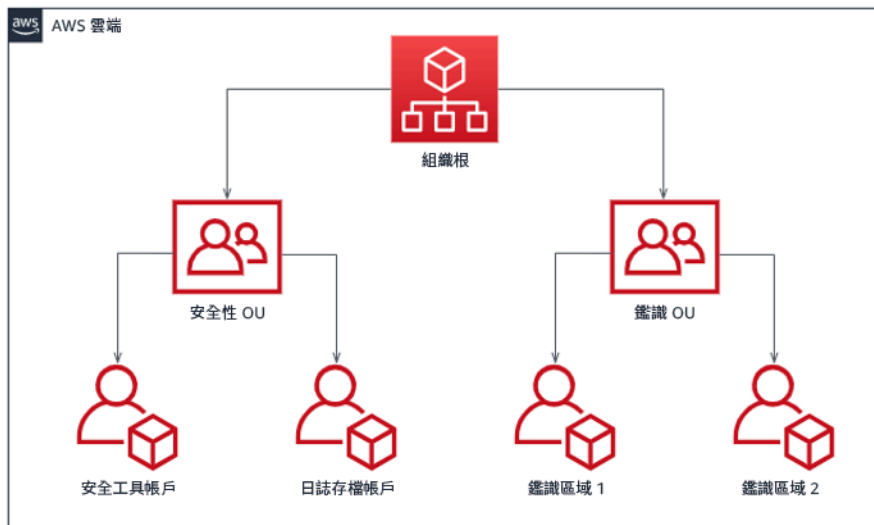
如需事件回應，建議您建立可支援事件回應功能的 AWS 帳戶結構，其中包括安全性 OU 和鑑識 OU。在安全性 OU 中，您應該擁有下列項目的帳戶：

- 日誌存檔：使用有限的許可彙總日誌存檔 AWS 帳戶 中的日誌。
- 安全性工具：將安全性服務集中在安全工具 AWS 帳戶 中。此帳戶會以安全性服務的委派系統管理員身分運作。

在鑑識 OU 中，您可以選擇為營運所在的每個區域實作一或多個鑑識帳戶，具體視哪個區域最適合您業務和營運模式而定。如果您為每個區域建立鑑識帳戶，則可以防止該區域以外的 AWS 資源建立，並降低將資源複製到非預期區域的風險。例如，如果您只在美國東部 (維吉尼亞北部) 區域 (us-east-1) 和美國西部 (奧勒岡) (us-west-2) 進行營運，則鑑識 OU 中會有兩個帳戶：一個用於 us-east-1，另一個用於 us-west-2。

您可以為多個區域建立鑑識 AWS 帳戶。您在將 AWS 資源複製到該帳戶時應小心，以確認是否符合資料主權要求。佈建新帳戶需要一些時間，因此必須在事件之前建立和檢測鑑識帳戶，以便回應者能夠有效地使用這些帳戶進行回應。

下圖顯示範例帳戶結構，包括具有每個區域鑑識帳戶的鑑識 OU：



事件回應的每個區域帳戶結構

擷取備份和快照

設定重要系統和資料庫的備份，對於從安全事件中復原和鑑識用途非常重要。備份就緒後，您可以將系統還原到先前的安全狀態。您可以在 AWS 上拍攝各種資源的快照。快照可為您提供那些資源的時間點備份。有許多 AWS 服務，可以在備份和復原方面為您提供支援。如需有關這些備份和復原之服務和方法的詳細資訊，請參閱[備份和復原方案指引](#)和[使用備份從安全事件中復原](#)。

尤其是當涉及勒索軟體等情況時，務必確保備份是否有充足的保護。如需有關保護備份的指引，請參閱在[AWS 中保護備份的 10 大安全最佳實務](#)。除了確保備份的安全之外，您還應該定期測試備份和還原程序，以確認您現有的技術和程序是否如預期般運作。

自動化鑑識

在安全事件期間，事件回應團隊必須能夠快速收集和分析證據，同時維持事件周圍期間的準確性 (例如擷取與特定事件或資源相關的日誌，或收集 Amazon EC2 執行個體的記憶體傾印)。事件回應團隊手動收集相關證據既具挑戰性又耗時，尤其是範圍遍及大量執行個體和帳戶時。此外，手動收集可能容易出現人為錯誤。基於這些原因，您應盡可能開發和實作鑑識的自動化。

AWS 為鑑識提供許多自動化資源，內容列於以下的資源區段。這些資源是我們已開發和客戶已實作的鑑識模式範例。雖然這些範例在一開始可能是有用的參考架構，但請根據環境、需求、工具和鑑識程序，考慮是否加以修改或建立新的鑑識自動化模式。

資源

相關文件：

- [AWS 安全事件回應指南 - 開發鑑識功能](#)
- [AWS 安全事件回應指南 - 鑑識資源](#)
- [AWS 雲端 中的鑑識調查環境策略](#)
- [如何在 AWS 中自動化鑑識磁碟收集](#)
- [AWS 方案指引 - 自動化事件回應和鑑識](#)

相關影片：

- [自動化事件回應和鑑識](#)

相關範例：

- [自動化事件回應和鑑識架構](#)
- [Amazon EC2 的自動化鑑識協調器](#)

SEC10-BP04 開發和測試安全事件回應程序手冊

準備事件回應流程的關鍵部分是制定程序手冊。事件回應程序手冊提供方案指引，以及安全事件發生時應遵循的步驟。提供清晰的結構和步驟簡化了回應的複雜度並減少人為錯誤的可能性。

未建立此最佳實務時的曝險等級：中

實作指引

應針對事件案例建立程序手冊，例如：

- 預期事件：應針對您預期的事件建立程序手冊。這包括拒絕服務 (DoS)、勒索軟體和憑證入侵等威脅。
- 已知安全調查結果或提醒：應針對已知安全調查結果和提醒 (例如 Amazon GuardDuty 調查結果) 建立執行手冊。當您收到 GuardDuty 調查結果時，手冊應提供明確的步驟，以避免處理不當或忽略提醒。如需更多修復相關資訊和指引，請參閱[修復 GuardDuty 發現的安全問題](#)。

程序手冊應包含安全分析師應完成的技術步驟，以便充分調查和應對潛在的安全事件。

AWS 的客戶事件回應團隊 (CIRT) 已發布[包含事件回應操作手冊的 GitHub 儲存庫](#)，依威脅情境、類型和資源整理。這些操作手冊可以調整到符合您現有的事件回應程序，或可當成開發新程序的基礎。

實作步驟

要納入程序手冊的項目包括：

- 程序手冊概觀：這份程序手冊可處理哪些風險或事件？程序手冊的目標是什麼？
- 先決條件：此事件案例需要哪些日誌、偵測機制和自動化工具？預期的通知是什麼？
- 溝通和向上呈報資訊：誰參與其中，其聯絡資訊為何？每個利害關係人的責任是什麼？
- 回應步驟：在事件回應的各個階段，應採取哪些戰術步驟？分析師應該執行哪些查詢？應該執行哪些程式碼以達到預期的成果？
 - 偵測：事件的偵測方式為何？
 - 分析：判斷影響範圍的方式為何？
 - 包含：隔離事件以限制範圍的方式為何？
 - 根除：將威脅從環境中移除的方式為何？
 - 復原：受影響的系統或資源重新投入生產環境的方式為何？
- 預期成果：執行查詢和程式碼後，程序手冊的預期結果是什麼？

資源

相關 Well-Architected 的最佳實務：

- [SEC10-BP02 - 制定事件管理計畫](#)

相關文件：

- [事件回應程序手冊的架構](#)
- [制定您自己的事件回應程序手冊](#)
- [事件回應程序手冊範例](#)
- [使用 Jupyter 程序手冊和 CloudTrail Lake 建置 AWS 事件回應執行手冊](#)

SEC10-BP05 預先佈建存取權

確認事件回應者具有在 AWS 中預先佈建的正确存取權限，以縮短調查直至復原所需的時間。

常見的反模式：

- 使用事件回應的根帳戶。
- 更改現有的帳戶。
- 當提供即時權限提升時直接操控 IAM 許可。

未建立此最佳實務時的曝險等級：中

實作指引

AWS 建議盡可能降低或避免對長期憑證的依賴，而是採用暫時憑證和即時權限提升機制。長期憑證容易發生安全性風險並會增加營運負擔。對於大多數管理任務，以及事件回應任務，我們建議您實作[聯合身分](#)以及[適用於管理存取權的暫時權限提升](#)。在此模型中，使用者會請求提升至較高層級的權限 (例如事件回應角色)；如果使用者符合提升的資格，則會將請求傳送至核准者。如果請求獲得核准，使用者就會收到一組暫時 [AWS 憑證](#)，使用者可使用此憑證來完成其任務。在這些憑證過期後，使用者就必須提交新的提升權限請求。

我們建議在大多數事件回應情境中，使用暫時權限提升。正確的做法是使用 [AWS Security Token Service](#) 和[工作階段政策](#)來界定存取權的範圍。

當發生聯合身分不可用的情況，例如：

- 與遭盜用身分提供者 (IdP) 相關的中斷。
- 設定錯誤或人為錯誤會導致聯合存取管理系統遭到破壞。
- 分散式拒絕服務 (DDoS) 事件或使系統無法使用之類的惡意活動。

在上述案例中，應會有已設定的緊急存取權，可協助調查和及時修復事件。我們建議您使用[具有適當許可的使用者](#)，來執行任務和存取 AWS 資源。僅將根憑證用於[需要根使用者存取權的任務](#)。若要確認事件回應者是否具有 AWS 和其他相關系統的正确存取權，我們建議預先佈建專屬帳戶。此類帳戶需要提升的存取權，且必須受到嚴格的控制和監控。必須以執行必要任務所需的最低權限來建置這些帳戶，而存取權層級應以事件管理計畫中建立的程序手冊為基礎。

使用專用和專屬的使用者及角色作為最佳實務。透過新增 IAM 政策而暫時提升權限的使用者和角色存取權，會同時使得使用者在事件發生期間的存取權不明確，又有無法將提升的權限撤銷的風險。

您必須盡可能移除相依性，來確認可在各種可能的失敗情境下獲得存取權。為了做到這一點，建立程序手冊來確認事件回應使用者的建立身分是專屬安全帳戶中的使用者，且不會透過任何現有的聯合或單一登入 (SSO) 解決方案來管理事件回應使用者。每個個別回應者必須具備其專屬的指定帳戶。帳戶組態必須強制執行[強式密碼政策](#)和多重要素驗證 (MFA)。如果事件回應程序手冊僅需要 AWS 管理主控台的存取權，使用者就不應設定存取金鑰，且應明確禁止使用者建立存取金鑰。您可以使用 IAM 政策或

服務控制政策 (SCP) 進行設定，如 [AWS Organizations SCP](#) 的 AWS 安全最佳實務中所述。除了在其他帳戶中擔任事件回應角色的能力外，使用者不應具備任何權限。

在事件期間，必須將存取權授予其他內部或外部人員，來協助調查、修復和復原活動。在此案例中，使用先前提到的程序手冊機制，而且必須制定程序，以確認在事件完成後，立即將任何其他存取權撤回。

若要確認事件回應角色的使用是否受到適當的監控和稽核，則必須確保未在人員之間共用為此目的建立的 IAM 帳戶，且除非[特定任務所需](#)，否則不得使用 AWS 帳戶根使用者。如果需要根使用者 (例如，特定帳戶的 IAM 存取權不可用時)，請使用獨立的程序，其中有可用的程序手冊，來確認根使用者登入憑證和 MFA 字符是否可用。

若要為事件回應角色設定 IAM 政策，請考慮使用 [IAM Access Analyzer](#) 來根據 AWS CloudTrail 日誌產生政策。若要這麼做，請向管理員授予在非生產帳戶上事件回應角色的存取權，並透過程序手冊加以執行。完成後，您就可以建立政策來僅允許所採取的動作。接著就可以將此政策套用至所有帳戶中的所有事件回應角色。您可能希望為每個程序手冊建立個別 IAM 政策，來讓管理和稽核作業更輕鬆。範例程序手冊可能包含勒索軟體、資料洩漏、生產存取權遺失和其他情境的回應計畫。

使用事件回應帳戶來擔任[在其他 AWS 帳戶中專屬事件回應 IAM 角色](#)。必須將這些角色設定為僅供安全帳戶中的使用者擔任，而信任關係必須要求呼叫主體使用 MFA 進行驗證。這些角色必須使用嚴格控制範圍的 IAM 政策來控制存取權。確保所有對這些角色的 AssumeRole 請求都記錄在 CloudTrail 中並據以發出提醒，而使用這些角色採取的任何動作都會記錄下來。

強烈建議必須清楚地命名 IAM 帳戶和 IAM 角色，因此您可以輕鬆地在 CloudTrail 日誌中找到這些帳戶和角色。這類範例便是將 IAM 帳戶命名為 `<USER_ID>-BREAK-GLASS` 以及將 IAM 角色命名為 `BREAK-GLASS-ROLE`。

[CloudTrail](#) 會用來在 AWS 帳戶中記錄 API 活動，且應用來[設定對事件回應角色使用情形的提醒](#)。請參閱部落格貼文，其中會說明使用根金鑰如何設定提醒。您可以修改說明，以設定與事件回應 IAM 角色相關的 AssumeRole 事件的 [Amazon CloudWatch](#) 指標篩選條件至篩選條件：

```
{ $.eventName = "AssumeRole" && $.requestParameters.roleArn =  
  "<INCIDENT_RESPONSE_ROLE_ARN>" && $.userIdentity.invokedBy NOT EXISTS && $.eventType !=  
  "AwsServiceEvent" }
```

由於事件回應角色可能具備很高的存取權限，因此必須將這些提醒傳送給多個群組，並據此快速採取行動。

在事件期間，回應者可能需要存取未受 IAM 直接保護的系統。其中可能包含 Amazon Elastic Compute Cloud 執行個體、Amazon Relational Database Service 資料庫或軟體即服務 (SaaS) 平台。強烈建議

使用此方法，而不是使用 SSH 或 RDP 等原生協定，[AWS Systems Manager Session Manager](#) 會用於對 Amazon EC2 執行個體的所有管理存取權。您可以使用安全且受稽核的 IAM 來控制此存取權。您也可以使用 [AWS Systems Manager Run Command 文件](#) 來自動化部分程序手冊，如此可減少使用者錯誤並縮短復原時間。如需資料庫和第三方工具的存取權，我們建議將存取憑證儲存在 AWS Secrets Manager 中，並將存取權授予事件回應者角色。

最後，應將事件回應 IAM 使用者帳戶的管理作業新增至[加入者、異動者和離職者程序](#)中，並定期審查和測試此管理作業，以確認僅允許預期的存取。

資源

相關文件：

- [管理對 AWS 環境的暫時提升存取權](#)
- [AWS 安全事件回應指南](#)
- [AWS Elastic Disaster Recovery](#)
- [AWS Systems Manager Incident Manager](#)
- [設定 IAM 使用者的帳戶密碼政策](#)
- [使用 AWS 中的多重要素驗證 \(MFA\)](#)
- [使用 MFA 設定跨帳戶存取權](#)
- [使用 IAM Access Analyzer 來產生 IAM 政策](#)
- [多帳戶環境中 AWS Organizations 服務控制政策的最佳實務](#)
- [如何在使用 AWS 帳戶的根存取金鑰時接收通知](#)
- [使用 IAM 受管政策來建立精細的工作階段許可](#)
- [緊急存取](#)

相關影片：

- [將 AWS 中的事件回應和鑑識自動化](#)
- [執行手冊、事件報告和事件回應的 DIY 指南](#)
- [準備和回應 AWS 環境中的安全事件](#)

SEC10-BP06 預先部署工具

確認安全人員具有預先部署的適當工具，以縮短調查直至復原的時間。

未建立此最佳實務時的曝險等級：中

實作指引

若要自動化安全回應和操作功能，您可以使用 AWS 提供的完整 API 和工具集。您可以將身分管理、網路安全、資料保護和監控功能完全自動化，並使用現有的熱門軟體開發方法遞送這些功能。建置安全自動化時，您的系統可以監控、審核和啟動回應，而不是讓人員監控您的安全地位並手動回應事件。

如果您的事件回應團隊持續以相同方式回應提醒，可能會形成提醒疲勞的風險。隨著時間的推移，團隊可能會變得對收到提醒不敏感，而且在處理一般情況時可能會犯錯，或是錯過不尋常的提醒。自動化使用能夠處理重複和一般提醒的功能，讓人員處理敏感和獨特的事件，有助於避免發生提醒疲倦的情形。整合異常偵測系統 (例如 Amazon GuardDuty、AWS CloudTrail Insights 和 Amazon CloudWatch 異常檢測) 可以減輕常見閾值型提醒的負擔。

您可以透過程式設計方式將程序中的步驟自動化，以改善手動程序。定義事件的補救模式之後，您可以將該模式分解為可行的邏輯，並撰寫程式碼來執行該邏輯。回應人員接著可以執行該程式碼來修復問題。隨著時間的推移，您可以將越來越多的步驟自動化，最終自動處理整個類別的常見事件。

在安全調查期間，您需要能夠審核相關日誌以記錄和了解該事件的完整範圍和時間表。產生提醒也需要日誌，以指出特定關注的動作已發生。選擇、啟用、儲存和設定查詢與擷取機制和設定提醒至關重要。此外，提供搜尋日誌資料之工具的有效方法是 [Amazon Detective](#)。

AWS 擁有 200 多種雲端服務和數千種功能。我們建議您審核可支援並簡化事件回應策略的服務。

除了日誌記錄之外，您還應該開發和實作[標記策略](#)。標記可以協助提供與 AWS 資源用途有關的上下文。標記也可用於自動化。

實作步驟

選取並設定日誌以進行分析和提醒

請參閱下列有關設定事件回應日誌記錄的文件：

- [安全事件回應的日誌記錄策略](#)
- [SEC04-BP01 設定服務和應用程式日誌記錄](#)

啟用安全服務以支援偵測和回應

AWS 提供偵測、預防性和回應式功能，而其他服務可用於建立自訂安全性解決方案的架構。如需安全事件回應最相關的服務清單，請參閱[雲端功能定義](#)和[安全事件回應首頁](#)。

制定和實作標記策略

取得有關業務使用案例和圍繞 AWS 資源的相關內部利害關係人的上下文資訊可能很困難。執行此操作的一種方法是使用標籤的形式，此形式會將中繼資料指派給 AWS 資源，並包含使用者定義的鍵值組。您可以建立標籤，依目的、擁有者、環境、處理的資料類型以及您選擇的其他條件來分類資源。

擁有一致的標記策略可讓您快速識別和辨別與 AWS 資源有關的情境資訊，從而加快回應時間並盡可能減少用在組織情境的時間。標籤也可以作為啟動回應自動化的機制。如需有關標記內容的詳細資訊，請參閱[標記您的 AWS 資源](#)。您需要先定義要在整個組織中實作的標籤。之後，您將實作並強制執行此標記策略。如需有關實作和強制執行的詳細資訊，請參閱[使用 AWS 標籤政策和服務控制政策 \(SCP\) 實作 AWS 資源標記策略](#)。

資源

相關 Well-Architected 的最佳實務：

- [SEC04-BP01 設定服務和應用程式日誌記錄](#)
- [SEC04-BP02 在標準化的位置擷取日誌、調查結果和指標](#)

相關文件：

- [安全事件回應的日誌記錄策略](#)
- [事件回應雲端功能定義](#)

相關範例：

- [使用 Amazon GuardDuty 和 Amazon Detective 進行威脅偵測與回應](#)
- [Security Hub 研討會](#)
- [使用 Amazon Inspector 管理漏洞](#)

SEC10-BP07 執行模擬

在組織隨著時間成長和發展時，威脅態勢也會跟著演變，因此持續審查事件回應能力是很重要的。執行模擬 (也稱為比賽日) 是可用於執行此評估的一種方法。模擬會使用真實世界的安全事件案例，這些案例旨在模擬威脅參與者的策略、技術和程序 (TTP)，並且讓組織可藉由回應這些可能發生在現實中的模擬網路事件，來運用和評估其事件回應能力。

建立此最佳實務的優勢：模擬具有多種優勢：

- 驗證網路整備程度和培養事件回應人員的信心。
- 測試工具和工作流程的正確性及效率。
- 根據您的事件回應計畫，精進溝通和呈報方法。
- 提供回應罕見媒介的機會。

未建立此最佳實務時的曝險等級：中

實作指引

主要的模擬類型有三種：

- **桌上模擬演練：**桌上模擬方法是基於討論的會議，涉及各種事件回應利害關係人的角色和責任練習，並使用已建立的溝通工具和程序手冊。模擬演練促進通常可在虛擬場地、實體場地或兩者的組合於一整天內完成。桌上模擬演練以討論為主軸，因此側重於程序、人員和協作。技術在討論中是不可或缺的一部分，但事件回應工具或指令碼的實際使用通常不是桌上模擬演練的一部分。
- **紫隊模擬演練：**紫隊模擬演練提高了事件回應人員 (藍隊) 和模擬威脅參與者 (紅隊) 之間的協作層級。藍隊由安全營運中心 (SOC) 的成員組成，但也可以包含在實際網路事件期間涉入的其他利害關係人。紅隊由滲透測試團隊或受過攻擊性安全培訓的主要利害關係人組成。紅隊在設計場景時會與模擬演練協調員合作，使場景精確且可行。在紫隊模擬演練期間，主要重點是偵測機制、工具和支援事件回應工作的標準操作程序 (SOP)。
- **紅隊模擬演練：**在紅隊模擬演練期間，攻方 (紅隊) 會進行模擬，以在預定範圍內達到某個目標或一組目標。守方 (藍隊) 不一定知道模擬演練的範圍和持續時間，這對他們應對實際事件的能力可呈現出更真實的評估。由於紅隊模擬演練可能是侵入性測試，請謹慎行事並施加控制，以確認該模擬演練不會對您的環境造成實際傷害。

考慮定期推行網路模擬。每種模擬演練類型都可以為參與者和整個組織提供特有的好處，因此您可以選擇從較不複雜的模擬類型 (例如桌上模擬演練) 開始著手，然後再進入更複雜的模擬類型 (紅隊模擬演練)。您應根據自身的安全成熟度、資源和所需的結果來選取模擬類型。由於複雜性和成本較高，有些客戶可能不會選擇執行紅隊模擬演練。

實作步驟

無論您選擇的模擬類型為何，模擬通常會執行下列實作步驟：

1. **定義核心演練元素：**定義模擬情境與模擬的目標。這兩者都應獲得領導階層的允許。
2. **識別關鍵利害關係人：**模擬演練至少需要模擬演練協調員和參與者。根據情境，可能會涉及法律、通訊或主管領導階層等其他利害關係人。

3. 建置和測試情境：如果特定元素不可行，則可能需要在情境建置期間加以重新定義。預計最終的情境會成為此階段的輸出。
4. 促進模擬：模擬的類型將決定使用的促進形式 (編撰的場景對比於高度技術性的模擬場景)。協調員應使其促進策略與模擬演練目標相對應，他們應盡可能吸引所有模擬演練參與者，以提供最大的效益。
5. 撰寫事後報告 (AAR)：識別進展順利的領域、可以改進的領域，以及潛在的差距。AAR 應衡量模擬的有效性以及團隊對於模擬事件的應變能力，以便在未來的模擬追蹤進展幅度。

資源

相關文件：

- [AWS 事件回應指南](#)

相關影片：

- [AWS GameDay - Security Edition](#)
- [執行有效的安全事件回應模擬](#)

SEC10-BP08 建立從事件中學習的架構

實作經驗教訓的架構和根本原因分析能力，不僅有助改善事件回應能力，還有助防止事件重複發生。透過學習每個事件，您可以協助避免重複相同的錯誤、披露或錯誤設定，不僅能夠改善安全狀態，還可以盡可能縮短因可預防情況而損失的時間。

未建立此最佳實務時的曝險等級：中

實作指引

實作經驗教訓是非常重要的，其可在高層級實現以下幾點：

- 什麼時候開設經驗教訓課程？
- 經驗教訓課程中包含哪些內容？
- 經驗教訓課程的進行方式？
- 這個課程的參與者以及參與方式？
- 如何識別待改善之處？

- 您將如何確保有效地追蹤和實作待改善之處？

此架構不應該針對或責怪個人，而應該專注於改善工具和流程。

實作步驟

除了前述所列的高層級結果之外，確保您提出正確問題以從流程中獲得最大價值 (即協助您找到可行改善之處的資訊) 非常重要。考慮這些問題，有助您發起經驗教訓的討論：

- 事件是什麼？
- 第一次識別事件的時間？
- 事件的識別方式？
- 哪些系統對活動發出提醒？
- 涉及哪些系統、服務和資料？
- 具體發生的事件？
- 哪些方面做得很好？
- 哪些方面做得不好？
- 哪個流程或程序失敗或未能擴展以回應事件？
- 在以下幾個領域有哪些可以改善之處：
 - 人物
 - 需要聯絡的對象實際上是否有空，並且聯絡人清單是最新的嗎？
 - 人們是否缺少有效回應和調查事件所需的培訓或能力？
 - 適當的資源是否已準備就緒且可供使用？
 - 流程
 - 是否遵循流程和程序？
 - 是否已記錄並提供這類事件的流程和程序？
 - 是否缺少必要的流程和程序？
 - 回應人員是否能夠即時存取所需的資訊以回應問題？
 - 技術
 - 現有的提醒系統是否能有效地識別活動，並據以發出提醒？
 - 我們如何將偵測時間縮短 50%？
 - 是否需要改善現有提醒，或是需要針對此類事件建立新的提醒？
 - 現有的工具是否允許對事件進行有效的調查 (搜尋/分析)？

- 可以做什麼來協助加快這類事件的識別速度？
- 可以做什麼來協助避免這類事件再次發生？
- 負責改善計畫的人是誰，您將如何測試是否已實作此計畫？
- 實作和測試其他監控或預防性控制和流程的時間表為何？

這份清單並不詳盡，但可作為起點，幫助您識別組織和企業的需求，以及如何分析這些需求，以便最有效地從事件中學習並持續改善安全狀態。最重要的是透過將經驗教訓納入事件回應流程，文件和利害關係人期望的標準部分。

資源

相關文件：

- [AWS 安全事件回應指南 - 建立從事件中學習的架構](#)
- [NCSC CAF 指引 - 經驗教訓](#)

應用程式安全

問題

- [SEC 11. 如何在橫跨應用程式設計、開發和部署的整個生命週期內融入安全屬性並進行驗證？](#)

SEC 11. 如何在橫跨應用程式設計、開發和部署的整個生命週期內融入安全屬性並進行驗證？

人員培訓、使用自動化測試、了解相依性，以及驗證工具和應用程式的安全屬性，有助於減少生產工作負載中發生安全問題的機率。

最佳實務

- [SEC11-BP01 應用程式安全訓練](#)
- [SEC11-BP02 自動化在整個開發和發佈生命週期的測試](#)
- [SEC11-BP03 定期進行滲透測試](#)
- [SEC11-BP04 執行程式碼檢閱](#)
- [SEC11-BP05 集中化套件和相依性的服務](#)
- [SEC11-BP06 以程式設計方式部署軟體](#)

- [SEC11-BP07 定期評估管道的安全屬性](#)
- [SEC11-BP08 打造在工作負載團隊中納入安全所有權的計畫](#)

SEC11-BP01 應用程式安全訓練

為您的團隊提供安全開發和操作實務的培訓，協助他們建置安全的優質軟體。此實務可協助您的團隊在開發生命週期中提早預防、偵測和修復安全問題。請考慮提供涵蓋建立威脅模型、安全編碼實務，以及使用服務進行安全組態設定和操作的培訓。透過自助服務資源，提供培訓內容存取權給您的團隊，並定期收集他們的意見回饋以確保持續進步。

預期成果：您可以從一開始就將安全性納入考量，為團隊提供設計和建置軟體所需的知識和技能。經過有關建立威脅模型和安全開發實務的培訓後，您的團隊將對潛在安全風險，以及如何在軟體開發生命週期 (SDLC) 期間減輕這些風險有更深入的了解。這種主動式安全措施是團隊文化的一部分，能夠讓您及早識別和修復潛在的安全問題。因此，您的團隊可以更有效率地提供高品質、安全的軟體和功能，進而加速整體交付時間表。您在組織中擁有協作且具包容性的安全文化，並將安全擁有權與所有建置者共享。

常見的反模式：

- 您等到安全審查階段，才開始考慮系統的安全屬性。
- 您將所有的安全性決定工作全部留給中央安全團隊。
- 您未在 SDLC 溝通如何做出與整體安全期待或組織政策相關的決定。
- 您太晚執行安全審查程序。

建立此最佳實務的優勢：

- 可在開發生命週期初期更清楚了解組織對於安全的要求。
- 可以更快識別、修復安全問題，進而加快功能交付速度。
- 改善軟體和系統的品質。

未建立此最佳實務時的曝險等級：中

實作指引

若要建置安全且高品質的軟體，請為您的團隊提供有關安全開發和操作應用程式的常見實務培訓。此實務可協助您的團隊在開發生命週期中提早預防、偵測和修復安全問題，以加速交付時間表。

若要完成此實務，請考慮使用[建立威脅模型研討會](#)等的 AWS 資源，訓練您的團隊建立威脅模型。建立威脅模型可以讓您的團隊從一開始就了解潛在安全風險，並從安全考量來設計系統。此外，您可以提供有關安全開發實務的[AWS 培訓 and Certification](#)、產業或AWS合作夥伴培訓。如需更多設計、開發、保護和有效率地大規模操作的完整方法相關資訊，請參閱[AWS DevOps 指引](#)。

明確定義和傳達組織的安全審查程序，並概述團隊、安全團隊和其他利害關係人的責任。發布展示如何達到您的安全要求的自助式指引、程式碼範例和範本。您可以使用[AWS CloudFormation](#)、[AWS Cloud Development Kit \(AWS CDK\) \(AWS CDK\) 建構模組](#)和[Service Catalog](#)等 AWS 服務，提供預先核准的安全組態，並減少自訂設定的需求。

定期收集團隊的安全審查程序與培訓體驗方面的意見回饋，並藉由該回饋持續改進。執行演練日或錯誤修復日活動，以識別和解決安全問題，同時增強團隊的技能。

實作步驟

1. 識別培訓需求：透過調查、程式碼檢閱或與團隊成員討論，評估團隊中有關安全開發實務的目前技能水平和知識差距。
2. 規劃培訓：根據確定的需求建立培訓計畫，培訓範圍涵蓋建立威脅模型、安全編碼實務、安全測試和安全部署實務等相關主題。使用[建立威脅模型研討會](#)、[AWS 培訓 and Certification](#)，以及產業或AWS合作夥伴培訓計畫等資源。
3. 排程和交付培訓：為您的團隊安排定期培訓課程或研討會。這些可以由講師指導或自行安排進度，端視團隊的偏好設定和可用狀況而定。鼓勵實作練習和實際範例，以強化學習。
4. 定義安全審查程序：與您的安全團隊和其他利害關係人合作，明確定義應用程式的安全審查程序。記錄參與程序的每個團隊或個人的責任，包括您的開發團隊、安全團隊和其他利害關係人。
5. 建立自助式資源：製作自助式指引、程式碼範例和範本，示範如何做才能符合組織的安全要求。考慮採用[CloudFormation](#)、[AWS CDK 建構模組](#)和[Service Catalog](#)等 AWS 服務，以提供預先核准的安全組態，並減少自訂設定的需求。
6. 溝通和社交：有效地將安全審查程序和可用的自助式資源傳達給您的團隊。進行培訓課程或研討會，讓團隊成員熟悉這些資源，並確認他們了解如何使用這些資源。
7. 收集意見回饋並改進：定期收集團隊對於安全審查程序與培訓體驗的意見回饋。利用此回饋識別需要改進的領域，並持續改良培訓教材、自助式資源和安全審查程序。
8. 執行安全防護演練：安排演練日或錯誤修復日活動，以識別和解決應用程式內的安全問題。這些練習不僅有助於發現潛在的漏洞，同時也為您的團隊提供實際的學習機會，增強他們在安全開發和操作方面的技能。
9. 持續學習和改進：鼓勵您的團隊掌握最新的安全開發實務、工具和技術。定期審查和更新您的培訓教材和資源，以反映不斷變化的資安情勢和最佳實務。

資源

相關的最佳實務：

- [SEC11-BP08 打造在工作負載團隊中納入安全所有權的計畫](#)

相關文件：

- [AWS 培訓 和認證](#)
- [如何思考雲端安全管控](#)
- [如何建立威脅模型](#)
- [加速訓練 – AWS 技能培養](#)
- [AWS DevOps Sagas](#)

相關影片：

- [預防性安全：考量與方法](#)

相關範例：

- [威脅建模相關的研討會](#)
- [開發人員的產業認知](#)

相關服務：

- [AWS CloudFormation](#)
- [AWS Cloud Development Kit \(AWS CDK\) \(AWS CDK\) 建構模組](#)
- [Service Catalog](#)

SEC11-BP02 自動化在整個開發和發佈生命週期的測試

自動化在整個開發和發佈生命週期的安全署性測試。自動化可以讓軟體在發佈之前更容易一致且重複地識別潛在問題，因此能降低將供應軟體的安全問題風險。

預期成果：自動化測試的目標是提供以程式設計方式，及早偵測潛在問題，而且常常是遍及整個開發生命週期。啟用自動化迴歸測試時，您可以對變更之後的軟體重新執行功能性與非功能性的測試，確認

先前測試過的軟體運作仍如預期。如果定義安全單元測試來檢查常見的錯誤組態，例如損壞或缺失的驗證，您就能夠在開發過程中及早發現和修正這些問題。

測試自動化會使用專用測試個案來進行應用程式驗證，測試期間以應用程式需求和所需功能為基礎。自動化測試會將產生的測試輸出與其個別預期輸出進行比較，得到最後結果，進而加快整體的測試生命週期。包括像是迴歸測試與單位測試套組的測試方法最適合自動化應用。自動化安全屬性測試，建置人員就能接收自動化意見回饋，而不需要等待舉行安全審核。採用靜態或靜態程式碼分析的自動化測試，可以提高程式碼品質，並協助及早在開發生命週期中偵測出潛在的軟體問題。

常見的反模式：

- 未傳達測試個案與自動化測試的測試結果。
- 僅在即將發佈前執行自動化測試。
- 自動化有經常改變需求的測試個案。
- 無法提供如何解決安全測試結果的指引。

建立此最佳實務的優勢：

- 降低人員評估系統安全署性的依賴性。
- 可在跨多個工作串流之間找到一致結果，進而提高一致性。
- 降低造成安全問題被導入產品線上軟體的可能性。
- 因提早捕捉到軟體問題，而縮短偵測與矯正之間的範圍時段。
- 提高跨多個工作串流之系統或重複行為的能見度，其可用來促進整體組織改進。

未建立此最佳實務時的風險暴露等級：中

實作指引

隨著軟體逐漸建置，採用各種不同機制來測試軟體，確保您正根據應用程式的業務邏輯為主的功能性需求，以及著重應用程式可靠性、效能和安全性的非功能性需求，進行應用程式的測試作業。

靜態應用程式安全測試 (SAST) 分析原始程式碼是否有異常的安全模式，並提供可能存在缺陷程式碼的提示。SAST 依賴靜態輸入來測試某個範圍的已知安全問題，這些輸入包括文件 (需求規格、設計文件，以及設計規格) 和應用程式原始程式碼。靜態程式碼分析器可協助加快大量程式碼的分析作業。[NIST 品質群組](#)則提供[原始程式碼安全分析器](#)的比較，其中包括能用於[位元組程式碼掃描器](#)和[二進位程式碼掃描器](#)的開放原始碼工具。

應用動態分析安全測試 (DAST) 方法以補充您的靜態測試，這個方法會在應用程式執行期間進行測試，識別潛在的意外行為。動態測試可用來偵測出靜態分析無法偵測出的潛在問題。在程式碼儲存、建置和管道等階段進行測試，您就可以檢查進入程式碼當中的各種不同潛在問題類型。[Amazon Q Developer](#) 提供程式碼建議，包括在建置者的 IDE 中執行安全掃描。[Amazon CodeGuru 安全工具](#) 可以識別關鍵問題、安全問題，以及在應用程式開發期間難以發現的錯誤，並提供改善程式碼品質的建議。擷取軟體材料清單 (SBOM) 也可讓您擷取正式紀錄，其中包含建置軟體時所使用的各種元件的詳細資訊和關係。如此您便能知會漏洞管理，並快速識別軟體或元件相依項及供應鏈風險。

[開發人員安全工作坊](#) 使用 AWS 開發人員工具 (如 [AWS CodeBuild](#)、[AWS CodeCommit](#) 和 [AWS CodePipeline](#)) 執行包括 SAST 和 DAST 測試方法的發布管道自動化。

隨著 SDLC 繼續進行，建立包括安全團隊定期進行應用程式審核的反覆程序。收集自這些安全審核的意見回饋應加以解決，並在發佈準備度審核時加以驗證。這些審核作業會建立堅實強大的應用程式安全狀態，並提供建置人員可解決潛在問題的可行動意見回饋。

實作步驟

- 實作一致的 IDE、程式碼審核，以及包括安全測試的 CI/CD 工具。
- 考慮 SDLC 中的哪個位置適合封鎖管道，而不只是通知建置者出現需要矯正的問題。
- [自動化安全協助程式 \(ASH\)](#) 是開放原始碼程式碼安全掃描工具的範例。
- 使用自動化工具 (例如，已與開發人員 IDE 完成整合的 [Amazon Q Developer](#)，以及可在遞交認可時掃描程式碼的 [Amazon CodeGuru 安全工具](#)) 執行測試或程式碼分析，協助建置者在適當時間取得意見回饋。
- 使用 AWS Lambda 進行建置時，您可以使用 [Amazon Inspector](#) 來掃描多個功能的應用程式程式碼。
- 如果將自動化測試納入 CI/CD 管道，您應該使用票證系統來追蹤通知，以及軟體問題的矯正。
- 如果是可能會產生調查結果的安全測試，連結矯正的指引將有助於建置人員改善程式碼品質。
- 定期分析自動化工具所找到的調查結果，以便找出下次自動化、建置者訓練或認知行銷活動的優先順序。
- 若要擷取 SBOM 做為 CI/CD 管道的一部分，請使用 [Amazon Inspector SBOM 產生器](#)，以 CycloneDX SBOM 格式為封存、容器映像、目錄、本機系統和編譯的 Go and Rust 二進位檔製作 SBOM。

資源

相關的最佳實務：

- [DevOps 指引：DL.CR.3 為程式碼任務建立明確的完成條件](#)

相關文件：

- [持續交付與持續部署](#)
- [AWS DevOps 能力合作夥伴](#)
- [AWS 安全能力合作夥伴 \(應用程式安全\)](#)
- [選擇 Well-Architected CI/CD 方法](#)
- [Amazon CodeGuru 安全工具中的機密偵測](#)
- [Amazon CodeGuru 安全工具偵測程式庫](#)
- [配合有效管控，加速在 AWS 的部署](#)
- [AWS 如何達到自動化安全、無人為介入的部署](#)
- [Amazon CodeGuru 安全工具如何協助您有效平衡安全性和速度](#)

相關影片：

- [無人為介入：Amazon 的自動化持續交付管道](#)
- [自動化跨帳戶 CI/CD 管道](#)
- [Amazon 的軟體開發程序](#)
- [Amazon 的測試軟體和系統](#)

相關範例：

- [開發人員的產業認知](#)
- [自動化安全協助程式 \(ASH\)](#)
- [AWS CodePipeline 管控 - Github](#)

SEC11-BP03 定期進行滲透測試

定期對您的軟體進行滲透測試。這項機制有助於識別自動化測試或手動程式碼審查時，未能偵測到的潛在軟體問題。此外還有助於了解偵測控制的效用。滲透測試應嘗試判斷軟體是否會透過非預期的方式執行，例如暴露原本應受保護的資料，或是授予超乎預期的較廣泛權限。

預期成果：滲透測試可用來為您的應用程式安全屬性進行偵測、修復和驗證。軟體開發生命週期 (SDLC) 期間應該進行定期與排程型滲透測試。從滲透測試找到的調查結果應事先解決，才能安排軟體發行。您應該分析從滲透測試得到的調查結果，並識別是否有任何問題可使用自動化找出。實施包括主動意見回饋機制的定期和可重複滲透測試程序，可協助建置人員得知指引，並改善軟體品質。

常見的反模式：

- 只對已知或普遍存在的安全問題進行滲透測試。
- 滲透測試應用程式 (不含相依第三方工具和程式庫)。
- 只對套件安全問題進行滲透測試，且不評估已實作的商業邏輯。

建立此最佳實務的優勢：

- 提高軟體在發行前的安全屬性信心。
- 可識別偏好應用程式模式，並藉以提高軟體品質的機會。
- 在開發生命週期初期進行的意見回饋循環流程，當中的自動化或額外訓練可以改善軟體的安全屬性。

未建立此最佳實務時的風險暴露等級：高

實作指引

滲透測試是一種結構化的安全測試練習，過程當中，您會執行計畫的安全性缺口情境，對安全控制進行偵測、修復與驗證。滲透測試從偵察活動開始，過程中會根據目前的應用程式設計與其相依性收集資料。已經建置並執行精選的安全特定測試情境清單。這些測試的主要目的在於找出您的應用程式中的安全問題，這些問題可能會被利用來非預期地存取環境，或未經授權存取資料。當您推出新功能，或是每當應用程式遭遇重大的功能變更或進行技術實作，您就應該進行滲透測試。

您應該識別開發生命週期中最適合進行滲透測試的階段。這項測試的執行時間應該盡量延到系統功能接近預定發行階段之時，而且要保留足夠修復任何問題的時間。

實作步驟

- 建立處理滲透測試範圍限制方式的結構化程序，前提是這個關於[威脅模型](#)的程序是維持內容的好方法。
- 識別開發週期中最適合進行滲透測試的時機。進行測試時應該是預期應用程式進行最少變更，而且有足夠時間進行修復。
- 訓練建置人員學會從滲透測試調查結果預期哪些內容，以及如何取得關於修復的資訊。
- 使用工具，透過自動化共通或可重複測試，加速滲透測試程序。

- 分析滲透測試調查結果來找出系統性安全問題，並使用這份資料，得知其他的自動化測試與持續進行的建置人員教育。

資源

相關的最佳實務：

- [SEC11-BP01 應用程式安全訓練](#)
- [SEC11-BP02 自動化在整個開發和發佈生命週期的測試](#)

相關文件：

- [AWS 滲透測試](#)提供在 AWS 上進行滲透測試的詳細指引
- [配合有效管控，加速在 AWS 的部署](#)
- [AWS 安全能力合作夥伴](#)
- [現代化您在 AWS Fargate 上的滲透測試架構](#)
- [AWS Fault Injection Simulator](#)

相關範例：

- [自動化配合 AWS CodePipeline 的 API 測試 \(GitHub\)](#)
- [自動化安全協助程式 \(GitHub\)](#)

SEC11-BP04 執程式碼檢閱

實作程式碼檢閱，協助驗證正在開發之軟體的品質和安全性。程式碼檢閱旨在讓原始程式碼作者以外的團隊成員檢閱程式碼是否潛藏任何問題、漏洞，以及是否遵守編碼標準和最佳實務。此處理程序有助於擷取原始開發人員可能忽略的錯誤、不一致和安全漏洞。使用自動化工具協助檢閱程式碼。

預期結果：您在開發期間納入程式碼檢閱流程，以提高撰寫的軟體品質。您可以利用在程式碼檢閱期間的學習，提升經驗不足之團隊成員的技能。您可以使用自動化工具和測試來識別自動化的機會，並支援程式碼檢閱程序。

常見的反模式：

- 您在部署之前不會檢閱程式碼。
- 撰寫程式碼和檢閱程式碼的是同一個人。

- 您未使用自動化程序和工具來協助或協調程式碼檢閱。
- 您未在建置者開始檢閱程式碼之前，先讓他們接受應用程式安全的訓練。

建立此最佳實務的優勢：

- 程式碼品質更高。
- 經由重複使用常用方法而使程式碼開發更具一致性。
- 減少在滲透測試與後期階段找出問題的數量。
- 團隊內部的知識轉移效能更高。

未建立此最佳實務時的曝險等級：中

實作指引

程式碼檢閱有助於在開發期間驗證軟體的品質和安全性。手動檢閱旨在讓原始程式碼作者以外的團隊成員檢閱程式碼是否潛藏任何問題、漏洞，以及是否遵守編碼標準和最佳實務。此處理程序有助於擷取原始開發人員可能忽略的錯誤、不一致和安全漏洞。

請考慮使用 [Amazon CodeGuru 安全工具](#) 協助自動執行程式碼檢閱。CodeGuru 安全工具使用機器學習和自動推理來分析您的程式碼，並識別潛在的安全漏洞和編碼問題。將自動化程式碼檢閱與現有程式碼儲存庫整合，並持續整合/持續部署 (CI/CD) 管道。

實作步驟

1. 建立程式碼檢閱程序：

- 定義進行程式碼檢閱的時間點，例如在將程式碼合併到主要分支之前，或在部署到生產環境之前。
- 確定誰應該參與程式碼檢閱程序，例如團隊成員、資深開發人員和安全專家。
- 決定程式碼檢閱方法，包括使用的程序和工具。

2. 設定程式碼檢閱工具：

- 評估並選取符合您團隊需求的程式碼檢閱工具，例如 GitHub 提取請求或 CodeGuru 安全工具。
- 將所選工具與您現有的程式碼儲存庫和 CI/CD 管道整合。
- 設定工具以強制執行程式碼檢閱要求，例如檢閱者人數下限與核准規則。

3. 定義程式碼檢閱檢查清單和指引：

- 建立程式碼檢閱檢查清單，或概述應檢閱之內容的指引。考慮程式碼品質、安全漏洞、遵守編碼標準和效能等因素。

- 與開發團隊共用檢查清單或指引，並確認每個人都了解預期情況。
4. 訓練開發人員程式碼檢閱最佳實務：
 - 為您的團隊提供有關如何執行有效程式碼檢閱的訓練。
 - 教育您的團隊有關在檢閱過程中，需注意的應用程式安全原則和常見漏洞。
 - 鼓勵知識分享和配對程式設計工作階段，以提升經驗不足之團隊成員的技能。
 5. 實作程式碼檢閱程序：
 - 將程式碼檢閱步驟整合到您的開發工作流程中，例如建立提取請求和指派檢閱者。
 - 要求在合併或部署之前，針對程式碼變更內容進程式碼檢閱。
 - 在檢閱過程中，鼓勵公開溝通和建設性的意見回饋。
 6. 監控並改善：
 - 定期審查程式碼檢閱程序的有效性，並收集團隊的意見回饋。
 - 識別自動化或工具改進的機會，以簡化程式碼檢閱程序。
 - 根據學習和產業最佳實務，持續更新和精簡程式碼檢閱檢查清單或指引。
 7. 培養程式碼檢閱文化：
 - 強調程式碼檢閱的重要性，以維持程式碼品質和安全性。
 - 慶祝程式碼檢閱程序的成功和學習成果。
 - 營造一個可以讓開發人員放心地提供和接收意見回饋的協作和支援環境。

資源

相關的最佳實務：

- [SEC11-BP02 自動化在整個開發和發佈生命週期的測試](#)

相關文件：

- [DevOps 指引：DL.CR.2 針對程式碼變更執行對等檢閱](#)
- [關於 GitHub 中的提取請求](#)

相關範例：

- [使用 Amazon CodeGuru 安全工具自動執行程式碼檢閱](#)
- [使用 Amazon CodeGuru 安全工具 CLI，自動偵測 CI/CD 管道中的安全漏洞與錯誤](#)

相關影片：

- [使用 Amazon CodeGuru 安全工具持續改善程式碼品質](#)

SEC11-BP05 集中化套件和相依性的服務

提供可讓您的團隊取得軟體套件和其他相依性的集中化服務。這樣套件就能先接受驗證，再納入撰寫的軟體，並提供您組織中使用的軟體分析資料來源。

預期結果：除了您撰寫的程式碼之外，也可以從外部軟體套件建置您的工作負載。重複使用的實作功能變得更加簡單，例如 JSON 剖析器或加密程式庫。您可以集中這些套件和相依項的來源，如此一來您的安全團隊就能在使用這些項目前，先進行驗證。使用這個方法再加上手動與自動測試流程，即能提升對於您開發之軟體品質的信心。

常見的反模式：

- 您從網際網路的任意儲存庫中取出套件。
- 您未測試新套件，便將其提供給建置者。

建立此最佳實務的優勢：

- 更清楚了解哪些套件將用於建置中的軟體。
- 可以在了解過實際使用情況而需要更新套件時通知工作負載團隊。
- 降低在軟體中納入有問題套件的風險。

未建立此最佳實務時的風險暴露等級：中

實作指引

提供可讓建置人員輕鬆取得的套件和其他相依性集中化服務。集中化服務可依照邏輯方式進行集中，而非實作成單一龐大的系統。這個方法可讓您用符合建置人員需求的方式提供服務。您應該實作一種能在發生更新或新需求萌生時，快速在儲存庫新增套件的方法。[AWS CodeArtifact](#) 或類似的 AWS 合作夥伴解決方案等 AWS 服務可提供發揮這種能力的方法。

實作步驟

- 實作依照邏輯方式集中，而且各種軟體開發所在環境均可使用的儲存庫服務。
- 將儲存庫的存取作業納入 AWS 帳戶 銷售程序。

- 建置自動化測試流程，在將套件發行至儲存庫之前先進行測試。
- 維護最常使用的套件、語言，以及變更程度最高團隊的規格表。
- 提供可讓建置人員團隊自動要求新套件與提供意見回饋的機制。
- 定期掃描儲存庫中的套件，識別最近所找到問題的潛在影響。

資源

相關的最佳實務：

- [SEC11-BP02 自動化在整個開發和發佈生命週期的測試](#)

相關文件：

- [DevOps 指引：DL.CS.2 在每次建置後，會在程式碼成品上簽名](#)
- [軟體成品的供應鏈層級 \(SLSA\)](#)

相關範例：

- [配合有效管控，加速在 AWS 的部署](#)
- [使用 CodeArtifact 套件來源控制工具組加強您的套件安全](#)
- [多重區域套件發行管道 \(GitHub\)](#)
- [使用 AWS CodePipeline 在 AWS CodeArtifact 上發行 Node.js 模組 \(GitHub\)](#)
- [AWS CDK CDK Java CodeArtifact 管道範例 \(GitHub\)](#)
- [使用 AWS CodeArtifact 分發私有 .NET NuGet 套件 \(GitHub\)](#)

相關影片：

- [預防性安全：考量與方法](#)
- [安全的 AWS 原則 \(re:Invent 2017\)](#)
- [當安全性、安全和緊迫性都很重要時：處理 Log4Shell](#)

SEC11-BP06 以程式設計方式部署軟體

盡可能以程式設計方式進行軟體部署。此方法可減少部署失敗或因人為疏失而發生非預期問題的機率。

預期結果：您測試的工作負載版本是您部署的版本，且每次都會一致地執行部署。您可以將工作負載的組態外部化，如此有助於在無變動的情況下部署到不同的環境。您可以使用軟體套件的加密簽章來驗證環境之間沒有任何變化。

常見的反模式：

- 手動部署軟體至生產環境。
- 手動執行因應不同環境需求的軟體變更。

建立此最佳實務的優勢：

- 提高軟體發佈程序的可信度。
- 降低變更失敗影響到業務功能的風險。
- 因變更風險降低而增加發佈規律。
- 部署其間意外事件的自動回復能力。
- 可以密碼編譯方式證明所測試的軟體就是實際部署的軟體。

未建立此最佳實務時的曝險等級：高

實作指引

要維護強大且可靠的應用程式基礎設施，請實作安全且自動化的部署實務。此實務與從生產環境移除持續的人工存取、使用 CI/CD 工具進行部署，以及將環境特定的組態資料外化相關。遵循此方法可以增強安全性、降低人為錯誤的風險，並可簡化部署程序。

您可以建置您的 AWS 帳戶 結構，以從生產環境中移除持續的人力存取。此做法可將未經授權的變更或意外修改的風險降至最低，進而改善生產系統的完整性。您可以使用 [AWS CodeBuild](#) 和 [AWS CodePipeline](#) 等 CI/CD 工具來執行部署，而不是直接以人力存取。您可以使用這些服務來自動化建置、測試和部署處理程序，進而減少手動介入並提高一致性。

若要進一步增強安全性和可追蹤性，您可以在測試後簽署應用程式套件，並在部署期間驗證這些簽署。若要這麼做，請使用密碼編譯工具，例如 [AWS Signer](#) 或 [AWS Key Management Service \(AWS KMS\)](#)。您可以透過簽署和驗證套件，確保僅將已授權和已驗證的程式碼部署到您的環境。

此外，您的團隊可以建立工作負載架構，以從外部來源取得環境特定的組態資料，例如 [AWS Systems Manager Parameter Store](#)。此做法會將應用程式的程式碼與組態資料分開，這樣就能協助您獨立管理和更新組態，而無需修改應用程式的程式碼本身。

若要簡化基礎設施佈建和管理，請考慮使用基礎設施即程式碼 (IaC) 工具，例如 [AWS CloudFormation](#) 或 [AWS CDK](#)。您可以使用這些工具將基礎設施定義為程式碼，以提高跨不同環境部署的一致性和可重複性。

考慮採用 Canary 部署，驗證軟體是否已成功部署。Canary 部署需要在部署至整個生產環境之前，對執行個體或使用子集進行變更。然後，您可以監控變更造成的影響，並視需要復原，將常發生的問題風險降至最低。

請遵循[使用多個帳戶組織您的 AWS 環境](#)白皮書中概述的建議。此白皮書提供有關將環境 (例如開發、預備和生產) 分隔為不同 AWS 帳戶的指引，進一步增強安全性和隔離性。

實作步驟

1. 設定 AWS 帳戶結構：

- 遵循[使用多個帳戶組織您的 AWS 環境](#)白皮書中的指引，針對不同的環境 (例如開發、預備和生產) 建立不同的 AWS 帳戶。
- 為每個帳戶設定適當的存取控制和許可權限，以限制人力對生產環境的直接存取。

2. 實作 CI/CD 管道：

- 使用 [AWS CodeBuild](#) 和 [AWS CodePipeline](#) 等服務來設定 CI/CD 管道。
- 設定管道以自動建置、測試和部署應用程式的程式碼至個別環境中。
- 將程式碼儲存庫與 CI/CD 管道整合，以便於進行版本控制和程式碼管理。

3. 簽署和驗證應用程式套件：

- 測試和驗證應用程式套件後，可使用 [AWS Signer](#) 或 [AWS Key Management Service \(AWS KMS\)](#) 簽署應用程式套件。
- 設定部署程序，以在將應用程式套件部署到目標環境之前驗證其簽章。

4. 外部化組態資料：

- 在 [AWS Systems Manager Parameter Store](#) 中儲存環境特定的組態資料。
- 修改應用程式的程式碼，以便在部署或執行時期從參數存放區擷取組態資料。

5. 基礎設施即程式碼 (IaC)：

- 使用 [AWS CloudFormation](#) 或 [AWS CDK](#) 等 IaC 工具，以定義基礎設施即程式碼，並加以管理。
- 建立 CloudFormation 範本或 CDK 指令碼，以佈建和設定應用程式的必要 AWS 資源。
- 將 IaC 與您的 CI/CD 管道整合，以自動部署基礎設施變更，以及應用程式的程式碼變更。

6. 實作 Canary 部署：

- 將部署程序設定為支援 Canary 部署，其會在您將變更部署到整個生產環境之前，先推廣到執行個體或使用子集。

- 使用 [AWS CodeDeploy](#) 或 [AWS ECS](#) 等服務管理 Canary 部署，並監控變更造成的影響。
- 如果在 Canary 部署期間偵測到問題，請實作復原機制以還原至先前的穩定版本。

7. 監控和稽核：

- 設定監控和記錄機制，以追蹤部署、應用程式效能和基礎設施變更。
- 使用 [Amazon CloudWatch](#) 和 [AWS CloudTrail](#) 等服務來收集和分析日誌和指標。
- 實作稽核與合規檢查，確認是否遵循安全性最佳實務及法規要求。

8. 持續改善：

- 定期審查和更新您的部署實務，並納入對於先前部署的意見回饋，和從中學到的經驗。
- 盡可能自動化部署程序，以減少手動介入和潛在的人為錯誤。
- 與跨部門團隊合作 (例如，維運或資安)，以協調並持續改善部署實務。

遵循這些步驟，您可以在AWS您的環境中實作安全且自動化的部署實務，藉此增強安全性、降低人為錯誤的風險，並簡化部署程序。

資源

相關的最佳實務：

- [SEC11-BP02 自動化在整個開發和發佈生命週期的測試](#)
- [DL.CI.2 觸發程序會在原始碼修改後自動建置](#)

相關文件：

- [配合有效管控，加速在 AWS 的部署](#)
- [自動化安全、無人為介入的部署](#)
- [使用 AWS Certificate Manager Private CA 和 AWS Key Management Service 非對稱金鑰的程式碼簽署](#)
- [程式碼簽署，AWS Lambda 的信任和完整性控制](#)

相關影片：

- [無人為介入：Amazon 的自動化持續交付管道](#)

相關範例：

- [使用 AWS Fargate 的藍/綠部署](#)

SEC11-BP07 定期評估管道的安全屬性

採用 Well-Architected 安全原則保護您的流程，特別注意權限的區隔。定期評估管道基礎設施的安全屬性。有效管理管道的安全，就能讓您的軟體通過管道中重重的安全性考驗。

預期成果：用於建置與部署軟體的管道，皆遵循針對環境中其他任何工作負載所建議的相同做法。您在管道中實作的測試，無法由使用管道的團隊編輯。您只提供使用臨時憑證進行部署時所需的許可權給管道。您可以實作防護措施，以避免將管道部署到錯誤的環境。您可以將管道設定為發出狀態，如此才能驗證建置環境的完整性。

常見的反模式：

- 安全測試可能遭建置人員避開。
- 部署管道的權限過於廣泛。
- 管道未設定進行輸入驗證。
- 未定期審查與 CI/CD 基礎設施關聯的許可。
- 使用長期有效或硬式編碼的登入資料。

建立此最佳實務的優勢：

- 經由此類管道完成建置與部署的軟體完整性具備更高的可信度。
- 可在發現可疑活動時停止部署作業。

未建立此最佳實務時的風險暴露等級：高

實作指引

您的部署管道是軟體開發生命週期的重要元件，應遵循與環境中其他任何工作負載相同的安全原則和做法。這包括實作適當的存取控制、驗證輸入，以及定期審查與稽核與您的 CI/CD 基礎設施相關聯的許可權。

確認負責建置和部署應用程式的團隊，無法編輯或略過管道中實作的安全測試和檢查。這種將有疑慮的分隔的做法，有助於維護建置和部署程序的完整性。

剛開始，請考慮採用 [AWS 部署管道參考架構](#)。要在 AWS 上建置 CI/CD 管道，此參考架構提供了安全且可擴展的基礎。

此外，您可以使用 [AWS Identity and Access Management Access Analyzer](#) 等服務，為兩種管道許可產生最低權限 IAM 政策，並做為管道中用於驗證工作負載許可權的步驟。這有助於確認您的管道和工作負載僅擁有其特定功能所需的必要許可權，以降低未經授權的存取或動作風險。

實作步驟

- 從 [AWS 部署管道參考架構](#) 開始行動。
- 考慮使用 [AWS IAM Access Analyzer](#)，以程式設計方式產生管道的最低權限 IAM 政策。
- 將您的管道與監控和警示功能整合，以便在發生意外或異常活動時收到通知；而對於 AWS 受管服務，[Amazon EventBridge](#) 可讓您將資料路由至 [AWS Lambda](#) 或 [Amazon Simple Notification Service](#) (Amazon SNS) 等目標。

資源

相關文件：

- [AWS 部署管道參考架構](#)
- [監控 AWS CodePipeline](#)
- [的安全最佳實務AWS CodePipeline](#)

相關範例：

- [DevOps 監控儀表板](#) (GitHub)

SEC11-BP08 打造在工作負載團隊中納入安全所有權的計畫

打造一項計畫或一種機制，賦予建置人員團隊能對其本身所建軟體做出安全決策的能力。您的安全團隊仍需在審查過程中確認這些決策，但是讓建置人員團隊與生俱來擁有安全決策權，就能建置更快速、更安全的工作負載。這項機制也能推動所有權文化，積極影響您所建置系統的運作。

預期成果：您在團隊中已內嵌安全性擁有權和決策。您已針對如何考慮安全性培訓您的團隊，或已透過內嵌或關聯的資安人員來強化這些人的團隊。如此您的團隊才能在開發週期的早期，做出更優質的資安決策。

常見的反模式：

- 將所有的安全決策全部留給安全團隊。

- 未及早在開發程序初期解決安全需求。
- 未諮詢建置人員與安全部門人員在計畫運作方面的意見回饋。

建立此最佳實務的優勢：

- 縮短完成安全審核的時間。
- 減少必須在安全審核階段中偵測出的安全問題。
- 改善所編寫軟體的整體品質。
- 有機會識別並了解具備高度改善價值的系統性問題或區域。
- 減少因安全審核調查結果而必須進行的重新作業量。
- 改善對於安全功能的感覺。

未建立此最佳實務時的曝險等級：低

實作指引

從 [SEC11-BP01 應用程式安全訓練](#) 的指引開始。接著識別您認為最適合組織的計畫操作模式。其中兩種主要模式分別是訓練建置人員，以及在建置人員團隊當中納入安全部門人員。在您決定初步方法之後，您應該透過單一或小組型工作負載團隊進行先行試驗，證明該模式適合您的組織。建置人員與組織的安全部門所提供的領導支援，有助於計畫達成與成功實施。隨著這項計畫不斷建置，您一定要選擇可以用來顯示計畫價值的矩陣。了解 AWS 如何解決這個問題可以學到相當多知識。這項最佳實務非常強調組織層面變更與文化。您所使用的工具應該能支援建置人員與安全社群之間的協作。

實作步驟

- 從訓練建置人員處理應用程式安全開始。
- 建立專為教育建置人員的社群和上線計畫。
- 挑選計畫名稱。守門人、擁護者或倡導者是常見手法。
- 識別要應用的模式：訓練建置人員、納入安全部門工程師，或是安排親和性安全角色。
- 從安全部門、建置人員和可能的其他相關小組當中，識別專案贊助者。
- 計畫當中所涉多人的追蹤矩陣，審核所花時間，以及建置人員與安全團隊人員的意見回饋。使用這些矩陣來達成改善。

資源

相關的最佳實務：

- [SEC11-BP01 應用程式安全訓練](#)
- [SEC11-BP02 自動化在整個開發和發佈生命週期的測試](#)

相關文件：

- [如何建立威脅模型](#)
- [如何思考雲端安全管控](#)
- [AWS 如何建置 Security Guardians 程式的這種分散安全所有權的機制](#)
- [如何建置 Security Guardians 程式以分散安全所有權](#)

相關影片：

- [預防性安全：考量與方法](#)
- [來自 AWS 和 Toyota Motor North America 的 AppSec 工具和文化秘訣](#)

可靠性

可靠性支柱包括工作負載如預期般正確、一致地執行其預期功能的能力。可以在[可靠性支柱白皮書](#)中找到實作指引。

最佳實務領域

- [基礎](#)
- [工作負載架構](#)
- [變更管理](#)
- [故障管理](#)

基礎

問題

- [REL 1. 如何管理 Service Quotas 和限制？](#)
- [REL 2. 如何規劃您的網路拓撲？](#)

REL 1. 如何管理 Service Quotas 和限制？

雲端型工作負載架構具有 Service Quotas (也稱為服務限制)。這些配額的存在是為了防止意外佈建超出所需數量的資源，並限制 API 操作的請求率，以保護服務免於遭到濫用。另外還有一些資源限制，例如可將位元下推到光纖纜線的速率，或實體磁碟的儲存量。

最佳實務

- [REL01-BP01 了解服務配額和限制](#)
- [REL01-BP02 管理跨帳戶和區域的服務配額](#)
- [REL01-BP03 透過架構適應固定服務配額和限制](#)
- [REL01-BP04 監控和管理配額](#)
- [REL01-BP05 自動化配額管理](#)
- [REL01-BP06 確保目前配額與最大用量之間存在足夠差距以適應容錯移轉](#)

REL01-BP01 了解服務配額和限制

了解工作負載架構的預設配額和管理配額增加要求。知道哪些雲端資源限制 (例如，磁碟或網路) 具有潛在影響。

預期成果：客戶可以實作適當的準則來監控關鍵指標、基礎設施審查和自動化修復步驟，以確認是否未達到可能導致服務降級或中斷的服務配額和限制，以防止 AWS 帳戶中的服務降級或中斷。

常見的反模式：

- 在部署工作負載時，未了解軟、硬體配額及其對所用服務的限制。
- 部署替換工作負載時，未事先分析及重新設定必要的配額或聯絡支援人員。
- 假設雲端服務沒有限制，且在使用服務時無須考量費率、限制、計數和數量。
- 假設配額會自動增加。
- 不知道配額要求的程序和時間表。
- 假設每個服務在不同區域間的預設雲端服務配額都是相同的。
- 假設可以違反服務限制，且系統會將限制自動擴展或增加到資源限制以上
- 未以尖峰流量測試應用程式，以製造資源使用率的壓力。
- 佈建資源時未分析必要的資源大小。
- 選擇遠超出實際需求或預期尖峰的資源類型，而過度佈建容量。

- 未在新的客戶事件或部署新技術之前事先評估新流量層級的容量要求。

建立此最佳實務的優勢：監控和自動化服務配額和資源限制的管理可以主動減少故障。若未遵循最佳實務，客戶服務的流量模式變更即可能導致中斷或降級。藉由在所有區域和所有帳戶間監控並管理這些值，應用程式在遇到不良或非計劃性事件時將會有更高的彈性。

未建立此最佳實務時的曝險等級：高

實作指引

Service Quotas 是一項 AWS 服務，有助於您從單一位置管理超過 250 種 AWS 服務的配額。除了查閱配額值外，您也可以從 Service Quotas 主控台或使用 AWS SDK 來請求和追蹤配額增長。AWS Trusted Advisor 提供服務配額檢查功能，會顯示部分服務某些層面的用量和配額。每項服務的預設服務配額也根據各自服務列於 AWS 文件中 (例如，請參閱 [Amazon VPC Quotas](#))。

某些服務限制 (例如用於調節 API 的速率限制) 可藉由設定用量計畫在 Amazon API Gateway 中設定。在其各自服務上設為組態的一些限制包括佈建 IOPS、分配的 Amazon RDS 儲存體，以及 Amazon EBS 磁碟區分配。Amazon Elastic Compute Cloud 擁有自己的服務限制儀表板，有助於您管理執行個體、Amazon Elastic Block Store 和彈性 IP 位址限制。如果在您的使用案例中，服務配額會影響您的應用程式效能且無法根據您的需求調整，請聯絡 支援 以查看是否有緩解措施。

服務配額可能隨著區域而不同，或本質上是通用的。使用達到配額的 AWS 服務，將無法作為預期的正常用法，且可能導致服務中斷或降級。例如，服務配額會限制區域中使用的 DL Amazon EC2 執行個體數量。在使用 Auto Scaling 群組 (ASG) 的流量擴展事件期間，可能會達到該限制。

每個帳戶的服務配額均應定期受到用量評估，以確認該帳戶的適當服務限制為何。這些服務配額可作為操作上的防護機制，以防止不慎佈建超過您所需的資源。此外也可用來限制 API 操作的要求率，以保護服務免於濫用。

服務限制與服務配額不同。服務限制代表特定資源的類型為該資源定義的限制。這有可能是儲存容量 (例如，gp2 的大小限制為 1 GB - 16 TB) 或磁碟輸送量。制定資源類型的限制，並持續評估用量是否可能超出限制，是很重要的。若意外超出限制，帳戶的應用程式或服務可能會降級或中斷。

如果在某個使用案例中，服務配額會影響到應用程式的效能，且無法根據需求進行調整，請聯絡 支援 以查看是否有緩解措施。如需有關調整固定配額的詳細資訊，請參閱 [REL01-BP03 透過架構適應固定服務配額和限制](#)。

有許多 AWS 服務和工具可協助您監控及管理 Service Quotas。您應利用這些服務和工具，以提供配額層級的自動或手動檢查。

- AWS Trusted Advisor 提供服務配額檢查功能，會顯示部分服務某些層面的用量和配額。這有助於識別接近配額的服務。
- AWS 管理主控台 提供了相關方法，用以顯示服務配額值、管理及要求新配額、監控配額要求的狀態，以及顯示配額的歷史。
- AWS CLI 和 CDK 提供了程式化方法，可自動管理及監控服務配額層級與用量。

實作步驟

對於 Service Quotas：

- [檢視 AWS Service Quotas。](#)
- 若要了解現有的服務配額，請確認使用的服務 (例如 IAM Access Analyzer)。約有 250 個 AWS 服務受到服務配額控制。然後，確認每個帳戶和區域內可能使用的特定服務配額名稱。每個區域約有 3000 個服務配額名稱。
- 使用 AWS Config 擴充此配額分析，以查找 AWS 帳戶 中使用的所有 [AWS 資源](#)。
- 使用 [AWS CloudFormation 資料](#) 來確定您使用的 AWS 資源。查看在 AWS 管理主控台 中或使用 [list-stack-resources](#) AWS CLI 命令建立的資源。您也可以查看設定為自行在範本中部署的資源。
- 透過查看部署程式碼來確定工作負載所需的所有服務。
- 確認適用的服務配額。在 Trusted Advisor 和 Service Quotas 中使用可以程式設計方式存取的資訊。
- 建立自動監控方法 (請參閱 [REL01-BP02 管理跨帳戶和區域的服務配額](#) 和 [REL01-BP04 監控和管理配額](#))，以警示並通知服務配額是否接近或已達到上限。
- 建立自動化和程式化方法，以檢查某個區域中的服務配額是否已變更，但在同一帳戶的其他區域中未已變更 (請參閱 [REL01-BP02 管理跨帳戶和區域的服務配額](#) 和 [REL01-BP04 監控和管理配額](#))。
- 自動執行掃描應用程式日誌和指標，以確認是否有任何配額或服務限制錯誤。若有這類錯誤存在，請傳送提醒到監控系統。
- 一旦確定特定服務需要更大的配額，請建立工程程序，以計算配額中所需的變更 (請參閱 [REL01-BP05 自動化配額管理](#))。
- 建立佈建和核准工作流程，以要求變更服務配額。其中應包含要求遭拒絕或部分核准時的例外狀況工作流程。
- 建立工程方法以在佈建之前審查服務配額，並在推行至生產環境之前使用新的 AWS 服務。(例如，載入測試帳戶)。

針對服務限制：

- 建立監控和指標方法，針對接近資源限制的資源讀數發出提醒。適當利用 CloudWatch 進行指標或日誌監控。
- 為每個具有有效應用程式或系統限制的資源建立提醒閾值。
- 建立工作流程和基礎設施管理程序，以在限制接近使用率時變更資源類型。此工作流程應將負載測試納入作為最佳實務，以確認新類型是具有新限制的正確資源類型。
- 使用現有的程序和流程，將已識別的資源遷移至建議的新資源類型。

資源

相關的最佳實務：

- [REL01-BP02 管理跨帳戶和區域的服務配額](#)
- [REL01-BP03 透過架構適應固定服務配額和限制](#)
- [REL01-BP04 監控和管理配額](#)
- [REL01-BP05 自動化配額管理](#)
- [REL01-BP06 確保目前配額與最大用量之間存在足夠差距以適應容錯移轉](#)
- [REL03-BP01 選擇如何分割工作負載](#)
- [REL10-BP01 將工作負載部署至多個位置](#)
- [REL11-BP01 監控工作負載的所有元件以偵測故障](#)
- [REL11-BP03 將所有分層的修復自動化](#)
- [REL12-BP04 使用混沌工程測試彈性](#)

相關文件：

- [AWS Well-Architected Framework 的可靠性支柱：可用性](#)
- [AWS Service Quotas \(先前稱為 Service Limits\)](#)
- [AWS Trusted Advisor 最佳實務檢查 \(請參閱 Service Limits 章節\)](#)
- [AWS Answers 上的 AWS Limit Monitor](#)
- [Amazon EC2 Service Limits](#)
- [什麼是 Service Quotas？](#)
- [如何請求提高配額](#)
- [服務端點和配額](#)

- [Service Quotas 使用者指南](#)
- [Quota Monitor for AWS](#)
- [AWS 故障隔離界限](#)
- [備援的可用性](#)
- [AWS for Data](#)
- [什麼是持續整合？](#)
- [什麼是持續交付？](#)
- [APN 合作夥伴：可以幫助進行組態管理的合作夥伴](#)
- [管理 AWS 上每租用戶帳戶 SaaS 環境中的帳戶生命週期](#)
- [管理和監控工作負載中的 API 限流](#)
- [使用 AWS Organizations 大規模檢視 AWS Trusted Advisor 建議](#)
- [使用 AWS Control Tower 自動提高服務限制和企業支援](#)

相關影片：

- [AWS Live re:Inforce 2019 - Service Quotas](#)
- [使用 Service Quotas 檢視和管理 AWS 服務的配額](#)
- [AWS IAM 配額示範](#)

相關工具：

- [Amazon CodeGuru Reviewer](#)
- [AWS CodeDeploy](#)
- [AWS CloudTrail](#)
- [Amazon CloudWatch](#)
- [Amazon EventBridge](#)
- [Amazon DevOps Guru](#)
- [AWS Config](#)
- [AWS Trusted Advisor](#)
- [AWS CDK](#)

- [AWS Systems Manager](#)
- [AWS Marketplace](#)

REL01-BP02 管理跨帳戶和區域的服務配額

如果您使用多個帳戶或區域，請在生產工作負載執行的所有環境中都要求合適的配額。

預期成果：對於跨帳戶或區域的組態，或使用地區、區域或帳戶容錯移轉具有彈性設計的組態，服務和應用程式不應受到服務配額耗盡的影響。

常見的反模式：

- 允許一個隔離區域內的資源用量增長，但無維持其他隔離區域中容量的機制。
- 在隔離區域中單獨手動設定所有配額。
- 未考量彈性架構 (例如主動或被動) 日後在非主要區域降級期間對配額需求產生的影響。
- 未定期評估配額，並在工作負載執行所在的每個區域和帳戶中進行必要的變更。
- 不要利用[配額請求範本](#)在多個區域和帳戶之間請求增加。
- 因誤認為增加配額會產生成本上的影響 (例如運算保留要求) 而未更新服務配額。

建立此最佳實務的優勢：確認如果區域服務無法使用時，您可以處理次要區域或帳戶中目前的負載。這有助於降低區域中斷期間發生的錯誤數量或降級程度。

未建立此最佳實務時的曝險等級：高

實作指引

系統會針對每個帳戶追蹤服務配額。除非另有說明，否則每個配額都是 AWS 區域 特有的。除生產環境之外，也會在所有適用的非生產環境中管理配額，因此不會阻礙測試和開發。要維持高水準的彈性，必須持續評估服務配額 (無論自動還是手動)。

由於採用主動/主動、主動/被動 - 熱、主動/被動 - 冷，以及主動/被動- 指示燈方法實作設計，跨區域的工作負載越來越多，因此了解所有區域和帳戶配額級別至關重要。過去的流量模式不一定可明確指出服務配額是否正確設定。

同樣重要的是，每個區域的服務配額名稱限制不一定相同。在某個區域中，該值可能是五，而另一個區域中的值可能是十。這些配額的管理必須跨所有的相同服務、帳戶和區域，以在負載下提供一致的彈性。

在不同區域 (主動區域或被動區域) 間協調所有服務配額差異，並建立持續協調這類差異的程序。被動區域容錯移轉的測試計劃鮮少擴展至尖峰主動容量，意即演練日或桌面演練可能找不到區域之間的服務配額差異，因而無法維持正確的限制。

服務配額漂移是指特定指定配額的服務配額限制在一個區域而非所有區域中發生變更的情況，這對於追蹤和評估非常重要。您應考慮在具有流量甚或雲端承載流量的區域中變更配額。

- 根據您的服務要求、延遲、法規和災難復原 (DR) 要求，選取相關的帳戶和區域。
- 確定所有相關帳戶、區域和可用區域中的服務配額。限制範圍受限於帳戶和區域。您應比較這些值的差異。

實作步驟

- 審查可能超出使用風險等級的 Service Quotas 值。超出 80% 和 90% 閾值時，AWS Trusted Advisor 會提供提醒。
- 審查任何被動區域 (主動/被動設計中) 的服務配額值。確認在主要區域失敗時，負載將可在次要區域中成功執行。
- 自動評估相同帳戶中的區域之間是否發生了任何服務配額漂移，並採取因應措施以變更限制。
- 如果客戶的組織單位 (OU) 是以支援的方式建構的，則應更新服務配額範本，以反映應套用至多個區域和帳戶的任何配額中的變更。
 - 建立範本，並將區域關聯至配額變更。
 - 審查所有現有的服務配額範本，確認是否有任何必要的變更 (區域、限制和帳戶)。

資源

相關的最佳實務：

- [REL01-BP01 了解服務配額和限制](#)
- [REL01-BP03 透過架構適應固定服務配額和限制](#)
- [REL01-BP04 監控和管理配額](#)
- [REL01-BP05 自動化配額管理](#)
- [REL01-BP06 確保目前配額與最大用量之間存在足夠差距以適應容錯移轉](#)
- [REL03-BP01 選擇如何分割工作負載](#)
- [REL10-BP01 將工作負載部署至多個位置](#)

- [REL11-BP01 監控工作負載的所有元件以偵測故障](#)
- [REL11-BP03 將所有分層的修復自動化](#)
- [REL12-BP04 使用混沌工程測試彈性](#)

相關文件：

- [AWS Well-Architected Framework 的可靠性支柱：可用性](#)
- [AWS Service Quotas \(先前稱為 Service Limits\)](#)
- [AWS Trusted Advisor 最佳實務檢查 \(請參閱 Service Limits 章節\)](#)
- [AWS Answers 上的 AWS Limit Monitor](#)
- [Amazon EC2 Service Limits](#)
- [什麼是 Service Quotas？](#)
- [如何請求提高配額](#)
- [服務端點和配額](#)
- [Service Quotas 使用者指南](#)
- [Quota Monitor for AWS](#)
- [AWS 故障隔離界限](#)
- [備援的可用性](#)
- [AWS for Data](#)
- [什麼是持續整合？](#)
- [什麼是持續交付？](#)
- [APN 合作夥伴：可以幫助進行組態管理的合作夥伴](#)
- [管理 AWS 上每租用戶帳戶 SaaS 環境中的帳戶生命週期](#)
- [管理和監控工作負載中的 API 限流](#)
- [使用 AWS Organizations 大規模檢視 AWS Trusted Advisor 建議](#)
- [使用 AWS Control Tower 自動提高服務限制和企業支援](#)

相關影片：

- [AWS Live re:Inforce 2019 - Service Quotas](#)
- [使用 Service Quotas 檢視和管理 AWS 服務的配額](#)

- [AWS IAM 配額示範](#)

相關服務：

- [Amazon CodeGuru Reviewer](#)
- [AWS CodeDeploy](#)
- [AWS CloudTrail](#)
- [Amazon CloudWatch](#)
- [Amazon EventBridge](#)
- [Amazon DevOps Guru](#)
- [AWS Config](#)
- [AWS Trusted Advisor](#)
- [AWS CDK](#)
- [AWS Systems Manager](#)
- [AWS Marketplace](#)

REL01-BP03 透過架構適應固定服務配額和限制

請注意不可變更的服務配額、服務限制和實際資源限制。設計應用程式和服務的架構以防止這些限制影響可靠性。

範例包括 API 閘道的網路頻寬、無伺服器函數叫用承載大小、調節爆量速率，以及資料庫的使用者同時連線數目。

預期成果：在正常和高流量條件下，應用程式或服務會如預期般執行。它們已設計為在該資源的固定限制或服務配額內運作。

常見的反模式：

- 選擇使用一項服務的一項資源的設計，但未注意到擴展時會導致此項設計失效的設計限制。
- 執行不切實際的基準並且在測試期間達到服務固定配額。例如，以爆量限制執行測試，但是進行擴充的時間量。
- 選擇若超過固定服務配額時無法擴展或修改的設計。例如，SQS 承載大小為 256KB。
- 未設計可觀測性並且實作以監控和提醒在高流量活動期間可能有風險之服務配額的閾值

建立此最佳實務的優勢：確認應用程式是否會在所有預計的服務負載層級下執行，而不會中斷或降級。

未建立此最佳實務時的曝險等級：中

實作指引

不同於以較高容量單位取代的軟性服務配額，AWS 服務的固定配額無法變更。這表示這裡所有類型的 AWS 服務都必須在使用於應用程式設計中時評估其潛在硬性容量限制。

硬性限制會顯示在 Service Quotas 主控台中。如果資料欄顯示 ADJUSTABLE = No，則服務有硬性限制。硬性限制也會顯示在一些資源組態頁面中。例如，Lambda 有無法調整的特定硬性限制。

例如，設計 Python 應用程式在 Lambda 函數中執行時，應用程式應該評估以判斷 Lambda 是否有機會執行超過 15 分鐘。如果程式碼可能執行超過此服務配額限制，則必須考慮替代技術或設計。如果在生產部署後達到此限制，應用程式會遭受降級和中斷直到可以矯正為止。與軟性配額不同，沒有任何方法可以變更這些限制，即使是在緊急嚴重性 1 活動下。

一旦應用程式部署到測試環境，應該使用策略來尋找是否達到任何硬性限制。壓力測試、負載測試和混亂測試應該是引入測試計劃的一部分。

實作步驟

- 審核可用於應用程式設計階段的 AWS 服務完整清單。
- 審核這些服務的軟性配額限制和硬性配額限制。並非所有限制都會顯示在 Service Quotas 主控台中。有些服務[會在替代位置描述這些限制](#)。
- 隨著您設計您的應用程式，審核您的工作負載的業務和技術驅動來源，例如業務成果、使用案例、相依系統、可用性目標和災難復原物件。讓您的業務和技術驅動來源引導程序以識別適合您的工作負載的分散式系統。
- 分析區域和帳戶之間的服務負載。許多硬性限制對於服務是區域型的。不過，某些限制是帳戶型。
- 分析區域 (Zonal) 失敗和區域 (Regional) 失敗期間資源用量的彈性架構。在使用主動/主動、主動/被動 – 熱、主動/被動 – 冷和主動/被動 – 指示燈方法的多區預設定進度中，這些失敗案例會導致較高的用量。這會建立達到硬性限制的潛在使用案例。

資源

相關的最佳實務：

- [REL01-BP01 了解服務配額和限制](#)

- [REL01-BP02 管理跨帳戶和區域的服務配額](#)
- [REL01-BP04 監控和管理配額](#)
- [REL01-BP05 自動化配額管理](#)
- [REL01-BP06 確保目前配額與最大用量之間存在足夠差距以適應容錯移轉](#)
- [REL03-BP01 選擇如何分割工作負載](#)
- [REL10-BP01 將工作負載部署至多個位置](#)
- [REL11-BP01 監控工作負載的所有元件以偵測故障](#)
- [REL11-BP03 將所有分層的修復自動化](#)
- [REL12-BP04 使用混沌工程測試彈性](#)

相關文件：

- [AWS Well-Architected Framework 的可靠性支柱：可用性](#)
- [AWS Service Quotas \(先前稱為 Service Limits\)](#)
- [AWS Trusted Advisor 最佳實務檢查 \(請參閱 Service Limits 章節\)](#)
- [AWS Answers 上的 AWS Limit Monitor](#)
- [Amazon EC2 Service Limits](#)
- [什麼是 Service Quotas？](#)
- [如何請求提高配額](#)
- [服務端點和配額](#)
- [Service Quotas 使用者指南](#)
- [Quota Monitor for AWS](#)
- [AWS 故障隔離界限](#)
- [備援的可用性](#)
- [AWS for Data](#)
- [什麼是持續整合？](#)
- [什麼是持續交付？](#)
- [APN 合作夥伴：可以幫助進行組態管理的合作夥伴](#)
- [管理 AWS 上每租用戶帳戶 SaaS 環境中的帳戶生命週期](#)
- [管理和監控工作負載中的 API 限流](#)

- [使用 AWS Organizations 大規模檢視 AWS Trusted Advisor 建議](#)
- [使用 AWS Control Tower 自動提高服務限制和企業支援](#)
- [Service Quotas 的動作、資源和條件金鑰](#)

相關影片：

- [AWS Live re:Inforce 2019 - Service Quotas](#)
- [使用 Service Quotas 檢視和管理 AWS 服務的配額](#)
- [AWS IAM 配額示範](#)
- [AWS re:Invent 2018：閉環與開放思維：如何取得大小型系統的控制權](#)

相關工具：

- [AWS CodeDeploy](#)
- [AWS CloudTrail](#)
- [Amazon CloudWatch](#)
- [Amazon EventBridge](#)
- [Amazon DevOps Guru](#)
- [AWS Config](#)
- [AWS Trusted Advisor](#)
- [AWS CDK](#)
- [AWS Systems Manager](#)
- [AWS Marketplace](#)

REL01-BP04 監控和管理配額

評估潛在用量並適當地增加配額，以允許使用量按計劃增長。

預期成果：已經部署管理和監控的主動和自動化系統。這些操作解決方案可確保幾乎達到配額用量閾值。這些會由請求配額變更主動矯正。

常見的反模式：

- 未設定監控來檢查服務配額閾值

- 未設定監控硬性限制，即使這些值無法變更。
- 假設請求和保護軟性配額變更所需的時間量是立即或短期間。
- 設定了正在接近服務配額的警示，但無如何回應提醒的程序。
- 僅設定 AWS Service Quotas 所支援之服務的警示，而不監控其他服務 AWS。
- 未考慮多個區域彈性設計的配額管理，例如主動/主動、主動/被動 – 熱、主動/被動 - 冷和主動/被動 - 指示燈方法。
- 未評估區域之間的配額差異。
- 未評估每個區域特定配額增加請求的需求。
- 不利用[範本進行多區域配額管理](#)。

建立此最佳實務的優點：自動追蹤 AWS Service Quotas，並根據這些配額監控用量，可讓您在接近配額限制時查看。您也可以使用此監控資料來協助限制由於配額耗盡造成的任何降級。

未建立此最佳實務時的曝險等級：中

實作指引

針對支援的服務，您可以藉由設定可評估然後傳送提醒或警示的各種不同服務，來監控您的配額。這可協助監控用量並且可以在您接近配額時提醒您。這些警示可以從 AWS Config、Lambda 函數 CloudWatch、Amazon 或從叫用 AWS Trusted Advisor。您也可以在 CloudWatch 日誌上使用指標篩選條件來搜尋和擷取日誌中的模式，以判斷用量是否接近配額閾值。

實作步驟

針對監控：

- 擷取當前資源消耗 (例如，儲存貯體或執行個體)。使用 服務API操作，例如 Amazon EC2 DescribeInstances API來收集目前的資源消耗。
- 使用以下項目，擷取您目前基本且適用於服務的配額：
 - AWS Service Quotas
 - AWS Trusted Advisor
 - AWS 文件
 - AWS 服務特定頁面
 - AWS Command Line Interface (AWS CLI)
 - AWS Cloud Development Kit (AWS CDK)

- 使用 AWS Service Quotas，這項 AWS 服務可協助您從單一位置管理超過 250 個 AWS 服務的配額。
- 使用 Trusted Advisor 服務限制來監控您目前在各種閾值的服務限制。
- 使用服務配額歷史記錄（主控台或 AWS CLI）來檢查區域增加。
- 比較每個區域和每個帳戶中的服務配額變更，視需要建立等值。

針對管理：

- 自動化：設定 AWS Config 自訂規則，跨區域掃描服務配額，並比較差異。
- 自動化：設定排定的 Lambda 函數來掃描區域之間的服務配額，並且比較是否有差異。
- 手動：透過 AWS CLI、API 或 AWS 主控台掃描服務配額，以跨區域掃描服務配額並比較差異。報告差異。
- 如果區域之間識別出配額的差異，請視需要請求配額變更。
- 檢閱所有請求的結果。

資源

相關的最佳實務：

- [REL01-BP01 了解服務配額和限制](#)
- [REL01-BP02 管理跨帳戶和區域的服務配額](#)
- [REL01-BP03 透過架構適應固定服務配額和限制](#)
- [REL01-BP05 自動化配額管理](#)
- [REL01-BP06 確保目前配額與最大用量之間存在足夠差距以適應容錯移轉](#)
- [REL03-BP01 選擇如何分割工作負載](#)
- [REL10-BP01 將工作負載部署至多個位置](#)
- [REL11-BP01 監控工作負載的所有元件以偵測故障](#)
- [REL11-BP03 將所有分層的修復自動化](#)
- [REL12-BP04 使用混沌工程測試彈性](#)

相關文件：

- [AWS Well-Architected Framework 的可靠性支柱：可用性](#)

- [AWS Service Quotas \(先前稱為服務限制 \)](#)
- [AWS Trusted Advisor 最佳實務檢查 \(請參閱服務限制區段 \)](#)
- [AWS 限制對 AWS 答案的監控](#)
- [Amazon EC2 Service 限制](#)
- [什麼是 Service Quotas ?](#)
- [如何請求提高配額](#)
- [服務端點和配額](#)
- [Service Quotas 使用者指南](#)
- [的配額監控 AWS](#)
- [AWS 故障隔離界限](#)
- [備援的可用性](#)
- [AWS 適用於資料](#)
- [什麼是持續整合 ?](#)
- [什麼是持續交付 ?](#)
- [APN 合作夥伴：可協助進行組態管理的合作夥伴](#)
- [在 上管理 account-per-tenant SaaS 環境中的帳戶生命週期 AWS](#)
- [管理和監控工作負載中的API限流](#)
- [使用 大規模檢視 AWS Trusted Advisor 建議 AWS Organizations](#)
- [使用 自動化服務限制增加和企業支援 AWS Control Tower](#)
- [Service Quotas 的動作、資源和條件金鑰](#)

相關影片：

- [AWS Live re：Inforce 2019 - Service Quotas](#)
- [使用 AWS 服務配額檢視和管理Service Quotas](#)
- [AWS IAM Quotas 示範](#)
- [AWS re：Invent 2018：閉環和開場思維：如何控制大大小小的系統](#)

相關工具：

- [AWS CodeDeploy](#)

- [AWS CloudTrail](#)
- [Amazon CloudWatch](#)
- [Amazon EventBridge](#)
- [Amazon DevOpsGuru](#)
- [AWS Config](#)
- [AWS Trusted Advisor](#)
- [AWS CDK](#)
- [AWS Systems Manager](#)
- [AWS Marketplace](#)

REL01-BP05 自動化配額管理

服務配額也稱為 AWS 服務中的限制，是您的 AWS 帳戶 中資源的最大值。每一項 AWS 服務都會定義一組配額及其預設值。若要讓您的工作負載存取所需的所有資源，您可能需要提高服務配額值。

若工作負載耗用的 AWS 資源增加，則可能在超出配額時威脅到工作負載的穩定性，並且影響使用者體驗。實作工具以在工作負載接近限制時提醒您，並考慮自動建立配額增加請求。

預期成果：每個 AWS 帳戶 和區域中執行的工作負載都設定了適當的配額。

常見的反模式：

- 您未能考慮並適當地調整配額以符合工作負載需求。
- 您用來追蹤配額和用量的方法可能過舊，例如使用試算表。
- 您僅依照定期排程更新服務限制。
- 您的組織缺少審查現有配額以及在必要時請求增加服務配額的操作流程。

建立此最佳實務的優勢：

- 更富彈性的工作負載：您可避免超過 AWS 資源配額所造成的錯誤。
- 更簡單的災難復原：您在另一個 AWS 區域 中設定 DR 時，可重複使用在主要區域中建置的自動配額管理機制。

未建立此最佳實務時的曝險等級：中

實作指引

透過 AWS Service Quotas 主控台、AWS Command Line Interface (AWS CLI) 和 AWS SDK 等機制，檢視目前的配額並追蹤持續的配額使用量。您還可以將組態管理資料庫 (CMDB) 和 IT 服務管理 (ITSM) 系統與 AWS Service Quota API 整合。

如果配額用量達到定義的閾值，則產生自動提醒，並定義在收到提醒時提交配額增加請求的流程。如果基礎工作負載對您的業務至關重要，您可以自動化配額增加請求，但請仔細測試自動化程序，以避免增長回饋迴圈這類失控動作的風險。

通常會自動核准增加較少量的配額。較大的配額請求可能需要經過 AWS 支援人員手動處理，而且可能需要額外的時間來審查和處理。請預留額外的時間來處理多個請求或大量增加請求。

實作步驟

- 實作自動監控服務配額，並且在工作負載的資源使用率增長接近配額限制時發出警示。例如，適用於 AWS 的[配額監控](#)可自動監控服務配額。此工具會使用 Cloudformation StackSets 與 AWS Organizations 整合並進行部署，如此新帳戶就會在建立時自動受到監控。
- 使用 [Service Quotas 請求範本](#) 或 [AWS Control Tower](#) 等功能簡化新帳戶的 Service Quotas 設定。
- 在所有 AWS 帳戶和區域建置目前服務配額使用情形的儀表板，並視需要參考它們以避免超出配額。[Trusted Advisor Organizational \(TAO\) 儀表板](#) 是 [Cloud Intelligence 儀表板](#) 的一部分，可讓您快速開始使用此類儀表板。
- 追蹤服務限制增加請求。[Consolidated Insights from Multiple Accounts \(CIMA\)](#) 可提供所有請求的組織層級檢視內容。
- 透過在非實際執行帳戶中設定較低的配額閾值，以測試警示產生和任何配額增加請求的自動化程序。請勿在實際執行帳戶中進行這些測試。

資源

相關的最佳實務：

- [OPS10-BP07 自動回應事件](#)

相關文件：

- [APN 合作夥伴：可以幫助進行組態管理的合作夥伴](#)
- [AWS Marketplace：可追蹤限制的 CMDB 產品](#)
- [AWS Service Quotas \(先前稱為 Service Limits\)](#)

- [AWS Trusted Advisor 最佳實務檢查 \(請參閱 Service Limits 章節\)](#)
- [AWS 上的配額監控解決方案 - AWS 解決方案](#)
- [什麼是 Service Quotas ?](#)
- [什麼是 Service Quotas 請求範本 ?](#)

相關影片：

- [AWS Live re:Inforce 2019 - Service Quotas](#)
- [使用 AWS Control Tower 自動提高服務限制和提供企業支援](#)

相關工具：

- [適用於 AWS 的配額監控](#)

REL01-BP06 確保目前配額與最大用量之間存在足夠差距以適應容錯移轉

本文說明如何維護資源配額與用量之間的空間，以及如何讓您的組織受益。在完成使用資源之後，用量配額可能會繼續佔用該資源。這可能會導致資源失敗或無法存取。透過確認您的配額是否涵蓋無法存取資源及其替換項目的重疊，來防止資源失敗。計算此差距時，應考慮諸如網路失敗、可用區域失敗或區域失敗等案例。

預期成果：資源或資源可存取性中的小型或大型故障可涵蓋在目前的服務閾值內。已在資源規劃中考慮區域 (Zone) 失敗、網路失敗或甚至是區域 (Regional) 失敗。

常見的反模式：

- 根據目前的需求設定服務配額，而不考慮容錯移轉案例。
- 計算服務的尖峰配額時，未考慮靜態穩定性的主體。
- 計算每個區域所需的配額總計時，未考慮可能有無法存取的資源。
- 未針對某些服務及其潛在異常用量模式考慮 AWS 服務故障隔離界限。

建立此最佳實務的優勢：當服務中斷事件影響應用程式可用性時，請使用雲端來實作策略，以便從這些事件中復原。一個範例策略是建立額外的資源來取代無法存取的資源，以適應容錯移轉條件，而不會耗盡您的服務限制。

未建立此最佳實務時的曝險等級：中

實作指引

評估配額限制時，請考慮由於某些降級而可能發生的容錯移轉案例。請考慮下列容錯移轉情況。

- 中斷或無法存取的 VPC。
- 無法存取的子網路。
- 影響資源可存取性的降級可用區域。
- 聯網路由或輸入和輸出點遭到封鎖或變更。
- 影響資源可存取性的降級區域。
- 受區域或可用區域中的失敗所影響的資源子集。

容錯移轉的決策對於每個情況都是獨一無二的，因為業務影響有所不同。在決定容錯移轉應用程式或服務之前，先處理容錯移轉位置中的資源容量規劃和資源的配額。

審核每個服務的配額時，請考慮高於正常的活動峰值。這些峰值可能與由於聯網或權限而無法存取但仍處於活動狀態的資源相關。未終止的作用中資源會計入服務配額限制。

實作步驟

- 維持服務配額和最大用量之間的空間，以適應容錯移轉或可存取性的喪失。
- 確定服務配額。說明典型的部署模式、可用性需求和使用量增長。
- 視需要請求增加配額。預計配額增加請求的等待時間。
- 確定您的可靠性需求 (也稱為「幾個 9」)。
- 了解可能的故障案例，例如元件遺失、可用區域或區域。
- 建立您的部署方法 (範例包括 Canary、藍/綠、紅/黑或滾動)。
- 為當前配額限制新增適當的緩衝。範例緩衝為 15%。
- 適當時包含靜態穩定性的計算 (區域 (Zonal) 和區域 (Regional))。
- 規劃使用量增長並監控使用量趨勢。
- 考慮最關鍵工作負載的靜態穩定性影響。評估符合所有區域和可用區域中靜態穩定系統的資源。
- 考慮使用隨需容量保留，在任何容錯移轉之前排程容量。這是針對關鍵業務排程而實作的有用策略，可以降低在容錯移轉期間取得正確數量和資源類型的潛在風險。

資源

相關的最佳實務：

- [REL01-BP01 了解服務配額和限制](#)
- [REL01-BP02 管理跨帳戶和區域的服務配額](#)
- [REL01-BP03 透過架構適應固定服務配額和限制](#)
- [REL01-BP04 監控和管理配額](#)
- [REL01-BP05 自動化配額管理](#)
- [REL03-BP01 選擇如何分割工作負載](#)
- [REL10-BP01 將工作負載部署至多個位置](#)
- [REL11-BP01 監控工作負載的所有元件以偵測故障](#)
- [REL11-BP03 將所有分層的修復自動化](#)
- [REL12-BP04 使用混沌工程測試彈性](#)

相關文件：

- [AWS Well-Architected Framework 的可靠性支柱：可用性](#)
- [AWS Service Quotas \(先前稱為 Service Limits\)](#)
- [AWS Trusted Advisor 最佳實務檢查 \(請參閱 Service Limits 章節\)](#)
- [AWS Answers 上的 AWS Limit Monitor](#)
- [Amazon EC2 Service Limits](#)
- [什麼是 Service Quotas？](#)
- [如何請求提高配額](#)
- [服務端點和配額](#)
- [Service Quotas 使用者指南](#)
- [Quota Monitor for AWS](#)
- [AWS 故障隔離界限](#)
- [備援的可用性](#)
- [AWS for Data](#)
- [什麼是持續整合？](#)
- [什麼是持續交付？](#)
- [APN 合作夥伴：可以幫助進行組態管理的合作夥伴](#)
- [管理 AWS 上每租用戶帳戶 SaaS 環境中的帳戶生命週期](#)

- [管理和監控工作負載中的 API 限流](#)
- [使用 AWS Organizations 大規模檢視 AWS Trusted Advisor 建議](#)
- [使用 AWS Control Tower 自動提高服務限制和企業支援](#)
- [Service Quotas 的動作、資源和條件金鑰](#)

相關影片：

- [AWS Live re:Inforce 2019 - Service Quotas](#)
- [使用 Service Quotas 檢視和管理 AWS 服務的配額](#)
- [AWS IAM 配額示範](#)
- [AWS re:Invent 2018：閉環與開放思維：如何取得大小型系統的控制權](#)

相關工具：

- [AWS CodeDeploy](#)
- [AWS CloudTrail](#)
- [Amazon CloudWatch](#)
- [Amazon EventBridge](#)
- [Amazon DevOps Guru](#)
- [AWS Config](#)
- [AWS Trusted Advisor](#)
- [AWS CDK](#)
- [AWS Systems Manager](#)
- [AWS Marketplace](#)

REL 2. 如何規劃您的網路拓撲？

工作負載通常存在於多個環境中。其中包括多個雲端環境 (可公開存取與私有)，也可能包含您現有的資料中心基礎設施。計畫必須包括系統內和系統間連線、公有 IP 位址管理、私有 IP 位址管理和網域名稱解析等網路考量因素。

最佳實務

- [REL02-BP01 針對工作負載公有端點使用高可用性網路連線](#)

- [REL02-BP02 在雲端中的私有網路與內部部署環境之間佈建冗餘連線能力](#)
- [REL02-BP03 確保 IP 子網路分配帳戶具有擴展性和可用性](#)
- [REL02-BP04 偏好軸輻式拓撲而非多對多網狀拓撲](#)
- [REL02-BP05 在連線的所有私有地址空間中強制使用不重疊的私有 IP 位址範圍](#)

REL02-BP01 針對工作負載公有端點使用高可用性網路連線

建置與您的工作負載公有端點的高度可用網路連線，可協助您減少由於遺失連線的停機時間，並且改善您的工作負載的可用性和 SLA。為達成此目的，請使用高度可用的 DNS、內容交付網路 (CDN)、API 閘道、負載平衡或反向代理。

預期成果：為您的公用端點規劃、建置和操作高可用性網路連線至關重要。如果您的工作負載由於遺失連線而無法連線，即使您的工作負載正在執行且可用，您的客戶還是會看到您的系統是停機。藉由結合工作負載公有端點高度可用和具彈性的網路連線與工作負載本身的彈性架構，為您的客戶提供可行的最佳可用性和服務水準。

AWS Global Accelerator、Amazon CloudFront、Amazon API Gateway、AWS Lambda Function URL、AWS AppSync API 以及 Elastic Load Balancing (ELB) 全都可提供高可用性公用端點。Amazon Route 53 針對網域名稱解析提供高可用性的 DNS 服務，以確認您的公有端點地址是否可以解析。

您也可以評估用於負載平衡和代理的 AWS Marketplace 軟體設備。

常見的反模式：

- 設計高度可用的工作負載，而未規劃 DNS 和網路連線以取得高可用性。
- 在個別執行個體或容器上使用公有網際網路地址，並透過 DNS 管理其連線。
- 使用 IP 位址，而非網域名稱來定位服務。
- 未測試您的公有端點已遺失連線的情境。
- 未分析網路輸送量需求和分發模式。
- 未測試和規劃您的工作負載公有端點的網際網路網路連線可能遭到中斷的情境。
- 提供內容 (例如網頁、靜態資產或媒體檔案) 到大型地理區域，而不使用內容交付網路。
- 未針對分散式阻斷服務 (DDoS) 攻擊加以規劃。DDoS 攻擊所存在的風險會將合法流量阻擋在外，並減少使用者的可用性。

建立此最佳實務的優勢：針對高可用性和高彈性的網路連線進行設計，確保您的工作負載可供使用者存取且可用。

未建立此最佳實務時的曝險等級：高

實作指引

建置與您的公有端點的高度可用網路連線的核心是流量的路由。若要確認您的流量可以連線到端點，DNS 必須能夠將網域名稱解析為它們的對應 IP 位址。使用 Amazon Route 53 等高可用且可擴展的[網域名稱系統 \(DNS\)](#) 來管理網域的 DNS 記錄。也可以使用 Amazon Route 53 提供的運作狀態檢查。運作狀態檢查會確認您的應用程式可連線、可用並且可運作，它們可以透過模仿您的使用者行為的方式進行設定，例如請求網頁或特定 URL。發生失敗時，Amazon Route 53 會回應 DNS 解析請求，並且僅將流量導向到健康的端點。您也可以考慮使用 Amazon Route 53 提供的 Geo DNS 和以延遲為基礎的路由功能。

若要確認您的工作負載本身具有高可用性，請使用 Elastic Load Balancing (ELB)。Amazon Route 53 可用於將流量鎖定到 ELB，這會將流量分配到目標運算執行個體。也可以將 Amazon API Gateway 以及 AWS Lambda 一起用於無伺服器解決方案。客戶也可以在多個 AWS 區域中執行工作負載。透過[多站台主動/主動模式](#)，工作負載可以為來自多個區域的流量提供服務。使用多站台主動/被動模式時，工作負載可為來自主動區域的流量提供服務，同時將資料複製到次要區域，並在主要區域發生故障時變為作用中狀態。然後，Route 53 運作狀態檢查可用於控制從主要區域中任何端點到次要區域中的某個端點的 DNS 備援，從而確認您的工作負載是否可存取並可供使用者使用。

Amazon CloudFront 藉由使用全世界邊緣節點的網路為請求提供服務，以低延遲和高資料傳輸率提供簡易 API 來分發內容。內容交付網路 (CDN) 藉由在靠近使用者的位置提供放置或快取的內容來服務客戶。當內容負載從您的伺服器轉移到 CloudFront 的[邊緣節點](#)時，這也會改善應用程式的可用性。邊緣節點和區域邊緣快取會將您的內容快取複本保存在靠近您觀眾的位置，以便快速擷取並且增加您的工作負載的連線能力和可用性。

針對具有分散各地使用者工作負載，AWS Global Accelerator 可協助您改善應用程式的可用性和效能。AWS Global Accelerator 提供任播靜態 IP 位址，可做為一或多個 AWS 區域中託管之應用程式的固定進入點。這可讓流量盡可能輸入到使用者附近的 AWS 全球網路，改善您的工作負載的連線能力和可用性。AWS Global Accelerator 也會使用 TCP、HTTP 和 HTTPS 運作狀態檢查來監控您的應用程式端點的運作狀態。您的端點的運作狀態或組態的任何變更都允許將使用者流量重新導向到健康的端點，為您的使用者交付最佳效能和可用性。此外，AWS Global Accelerator 有故障隔離設計，使用由獨立網路區域提供服務的兩個靜態 IPv4 地址，增加您的應用程式的可用性。

為了幫助保護客戶免受 DDoS 攻擊，AWS 會提供 AWS Shield Standard。Shield Standard 會自動開啟並防止一般基礎架構 (第 3 層和第 4 層) 攻擊，例如 SYN/UDP 泛洪和反射攻擊，以支援 AWS 上應用程式的高可用性。針對更複雜和更大型攻擊 (例如 UDP 泛洪)、狀態耗盡攻擊 (例如 TCP SYN 泛洪) 的額外保護，以及協助保護您的應用程式在 Amazon Elastic Compute Cloud (Amazon EC2)、Elastic Load Balancing (ELB)、Amazon CloudFront、AWS Global Accelerator 以及 Route 53 上執行，您

可以考慮使用 AWS Shield Advanced。針對應用程式層攻擊的保護，例如 HTTP POST 或 GET 泛洪，請使用 AWS WAF。AWS WAF 可以使用 IP 位址、HTTP 標題、HTTP 本文、URI 字串、SQL injection 隱碼攻擊和跨網站指令碼條件來判斷應該封鎖或允許請求。

實作步驟

1. 設定高可用 DNS：Amazon Route 53 是一種可用性高、可擴展性強的[網域名稱系統 \(DNS\)](#) Web 服務。Route 53 會將使用者請求連接至在 AWS 上或內部部署中執行的網際網路應用程式。如需詳細資訊，請參閱 [configuring Amazon Route 53 as your DNS service](#)。
2. 設定運作狀態檢查：當使用 Route 53 時，請確認只有運作狀態良好的目標是可解析的。首先[建立 Amazon Route 53 運作狀態檢查和設定 DNS 備援](#)。以下是設定運作狀態檢查時要考慮的重要層面：
 - a. [Amazon Route 53 決定運作狀態檢查是否良好的方式](#)
 - b. [建立、更新和刪除運作狀態檢查](#)
 - c. [監控運作狀態檢查狀態和取得通知](#)
 - d. [Amazon Route 53 DNS 的最佳實務](#)
3. [將您的 DNS 服務連接到端點](#)。
 - a. 使用 Elastic Load Balancing 作為流量的目標時，請使用指向負載平衡器區域端點的 Amazon Route 53 建立[別名記錄](#)。建立別名記錄期間，將 [評估目標運作狀態] 選項設定為 [是]。
 - b. 對於使用 API Gateway 時的無伺服器工作負載或私有 API，請使用 [Route 53 將流量導向至 API Gateway](#)。
4. 決定內容交付網路。
 - a. 若要使用更接近使用者的邊緣節點交付內容，請先了解 [CloudFront 如何交付內容](#)。
 - b. [簡單的 CloudFront 分佈](#) 入門。CloudFront 接著會知道您想要從哪裡交付內容，以及如何追蹤和管理內容交付的詳細資料。以下是設定 CloudFront 分發時要了解 and 考慮的重要層面：
 - i. [快取如何與 CloudFront 邊緣節點搭配運作](#)
 - ii. [增加從 CloudFront 快取直接提供的請求比例 \(快取命中率\)](#)
 - iii. [使用 Amazon CloudFront Origin Shield](#)
 - iv. [透過 CloudFront 原始伺服器容錯移轉將高可用性最佳化](#)
5. 設定應用程式層保護：AWS WAF 可協助您保護免受常見 Web 漏洞和機器人的攻擊，這些攻擊會影響可用性、危及安全性或導致消耗過多資源。若要深入了解，請審核 [how AWS WAF works](#)，當您準備實作來自應用程式層 HTTP POST AND GET 泛洪的保護時，請參閱 [Getting started with AWS WAF](#)。也可以搭配使用 AWS WAF 與 CloudFront，請參閱有關 [how AWS WAF works with Amazon CloudFront features](#) 的文件。

6. 設定額外 DDoS 保護：根據預設，所有 AWS 客戶都能透過 AWS Shield Standard 獲得以您的網站或應用程式為目標，對於常見、最常發生網路和傳輸層 DDoS 攻擊的保護，不需額外費用。如需有關在 Amazon EC2、Elastic Load Balancing、Amazon CloudFront、AWS Global Accelerator 以及 Amazon Route 53 上執行的面向網際網路應用程式的額外保護，您可以考慮 [AWS Shield Advanced](#) 並審核 [DDoS 彈性架構的範例](#)。若要保護您的工作負載和公有端點免受 DDoS 攻擊，請參閱 [Getting started with AWS Shield Advanced](#)。

資源

相關的最佳實務：

- [REL10-BP01 將工作負載部署至多個位置](#)
- [REL11-BP04 復原期間需使用資料平面，而非控制平面](#)
- [REL11-BP06 當事件影響可用性時傳送通知](#)

相關文件：

- [APN 合作夥伴：可以幫助您規劃聯網的合作夥伴](#)
- [適用於網路基礎設施的 AWS Marketplace](#)
- [什麼是 AWS Global Accelerator？](#)
- [什麼是 Amazon CloudFront？](#)
- [什麼是 Amazon Route 53？](#)
- [什麼是 Elastic Load Balancing？](#)
- [網路連線能力 - 建立您的雲端基礎](#)
- [什麼是 Amazon API Gateway？](#)
- [什麼是 AWS WAF、AWS Shield 和 AWS Firewall Manager？](#)
- [什麼是 Amazon 應用程式復原控制器？](#)
- [設定 DNS 備援的自訂運作狀態檢查](#)

相關影片：

- [AWS re:Invent 2022 - 使用 AWS Global Accelerator 改善效能與可用性](#)
- [AWS re:Invent 2020：使用 Amazon Route 53 進行全球流量管理](#)
- [AWS re:Invent 2022 - 操作高可用性多可用區域應用程式](#)

- [AWS re:Invent 2022 – 深入了解 AWS 聯網基礎設施](#)
- [AWS re:Invent 2022 - 建置彈性網路](#)

相關範例：

- [使用 Amazon 應用程式復原控制器 \(ARC\) 進行災難復原](#)
- [AWS Global Accelerator 研討會](#)

REL02-BP02 在雲端中的私有網路與內部部署環境之間佈建冗餘連線能力

在雲端和內部部署環境中的私有網路之間的連線中實作備援，以實現連線恢復能力。這可以透過部署兩個或多個連結和流量路徑來實現，從而在發生網路故障時保持連接。

常見的反模式：

- 您只依賴一個網路連線，這會產生單點故障。
- 您只能使用一個 VPN 通道或結束於相同可用區域的多個通道。
- 您依賴一個 ISP 進行 VPN 連線，這可能會導致 ISP 中斷期間完全失敗。
- 不實作類似 BGP 的動態路由通訊協定，這對於在網路中斷期間重新路由流量至關重要。
- 忽略了 VPN 通道的頻寬限制，並高估了其備份功能。

建立此最佳實務的優勢：透過在您的雲端環境與您的公司或內部部署環境之間實作備援連線，即可確保兩個環境之間的相依服務能夠可靠地進行通訊。

未建立此最佳實務時的曝險等級：高

實作指引

使用 AWS Direct Connect 將內部部署網路連線至 AWS 時，您可以使用結束於多個內部部署位置和多個 AWS Direct Connect 位置的不同裝置上的不同連線，以實現最大的網路復原能力 (99.99% 的 SLA)。此拓撲可針對裝置故障、連線問題和完全的位置中斷提供復原能力。或者，可以使用多個位置的兩個單獨連線 (每個內部部署位置連線至單一 Direct Connect 位置) 實現高復原能力 (99.9% 的 SLA)。此方法可防止因光纖斷裂或裝置故障而造成的連線中斷，並有助於減輕完全的位置故障。Direct Connect 彈性工具組可協助您設計 AWS Direct Connect 拓撲。

您也可以考慮將結束於 AWS Transit Gateway 的 AWS Site-to-Site VPN 作為對主要 AWS Direct Connect 連線的具有成本效益的備份。此設定可在多個 VPN 通道中實現等價多路徑 (ECMP) 路由，即

使每個 VPN 通道的上限為 1.25 Gbps，也能達到高達 50Gbps 的輸送量。但是，請務必注意，AWS Direct Connect 仍然是最大程度地減少網路中斷並提供穩定連接的最有效選擇。

透過網際網路使用 VPN 將雲端環境連線至內部部署資料中心時，請將兩個 VPN 通道設定為單一站台對站台 VPN 連線的一部分。每個通道都應該在不同的可用區域結束，以獲得高可用性，並使用備援硬體來防止內部部署裝置故障。此外，請考慮內部部署位置來自不同網際網路服務供應商 (ISP) 的多個網際網路連線，以避免因單一 ISP 中斷而造成 VPN 連線完全中斷。選擇具有多種路由和基礎設施的 ISP，尤其是那些到 AWS 端點具有不同實體路徑到的 ISP，可提供高連線可用性。

除了具有多個 AWS Direct Connect 連線和多個 VPN 通道 (或兩者的組合) 的實體備援之外，實作邊界閘道協定 (BGP) 動態路由也很重要。動態 BGP 可根據即時網路條件和設定的政策，將流量從一個路徑自動重新路由到另一個路徑。這種動態行為對於在發生連結或網路故障時維持網路可用性和服務連續性特別有益。它可以快速選擇替代路徑，提高網路的彈性和可靠性。

實作步驟

- 獲得 AWS 與內部部署環境之間的高可用連線。
 - 在單獨部署的私有網路之間使用多個 AWS Direct Connect 連線或 VPN 通道。
 - 使用多個 Direct Connect 位置以實現高可用性。
 - 如果使用多個 AWS 區域，請至少在其中兩個區域中建立備援。
- 如果可能，請使用 AWS Transit Gateway 來結束 [VPN 連線](#)。
- 評估 AWS Marketplace 設備以結束 VPN 或將 [SD-WAN 擴展到 AWS](#)。如果您使用 AWS Marketplace 設備，可在不同的可用區域中部署冗餘執行個體以實現高可用性。
- 提供內部部署環境的備援連線。
 - 您可能需要多個 AWS 區域的冗餘連線才能滿足可用性需求。
 - 使用 [Direct Connect 彈性工具組](#) 以開始使用。

資源

相關文件：

- [AWS Direct Connect 恢復能力建議](#)
- [使用備援的站台對站台 VPN 連接來提供容錯移轉](#)
- [路由政策和 BGP 社群](#)
- [AWS Direct Connect 中的主動/主動和主動/被動組態](#)
- [APN 合作夥伴：可以幫助您規劃聯網的合作夥伴](#)

- [AWS Marketplace適用於網路基礎設施的](#)
- [Amazon Virtual Private Cloud 連線選項白皮書](#)
- [建立可擴展且安全的多個 VPC AWS 網路架構](#)
- [使用備援的站台對站台 VPN 連接來提供容錯移轉](#)
- [使用 Direct Connect 彈性工具組以開始使用](#)
- [VPC 端點和 VPC 端點服務 \(AWS PrivateLink\)](#)
- [什麼是 Amazon VPC ?](#)
- [什麼是傳輸閘道 ?](#)
- [什麼是 AWS Site-to-Site VPN ?](#)
- [使用 Direct Connect 閘道](#)

相關影片：

- [AWS re:Invent 2018：進階 VPC 設計及 Amazon VPC 的新功能](#)
- [AWS re:Invent 2019：適用於許多 VPC 的 AWS Transit Gateway 參考架構](#)

REL02-BP03 確保 IP 子網路分配帳戶具有擴展性和可用性

Amazon VPC IP 位址範圍必須足夠大，以因應工作負載的要求，包括考慮將來擴展 IP 位址以及跨可用區域將 IP 位址分配給子網路。這包括負載平衡器、EC2 執行個體和容器型應用程式。

規劃您的網路拓樸時，首先要定義 IP 位址空間。私有 IP 位址範圍 (依循 RFC 1918 指引) 應分配給各 VPC。在此流程中請滿足下列要求：

- 允許每個區域為多於一個 VPC 準備 IP 位址空間。
- 在 VPC 內，允許多個子網路，以便可涵蓋多個可用區域。
- 考慮在 VPC 內留下未用 CIDR 區塊空間，以供未來擴展。
- 確保有 IP 位址空間可以滿足您可能會用到之 Amazon EC2 執行個體的任何臨時機群需求，例如，用於機器學習的 Spot 機群、Amazon EMR 叢集或 Amazon Redshift 叢集。Kubernetes 叢集 (例如 Amazon Elastic Kubernetes Service (Amazon EKS)) 也應考慮類似的考量，因為預設情況下，每個 Kubernetes Pod 都會從 VPC CIDR 區塊指派一個可路由的位址。
- 請注意，在各子網路 CIDR 區塊中，前四個 IP 位址和最後一個 IP 位址均預留起來，無法供您使用。
- 請注意，雖然無法變更或刪除分配給 VPC 的最初 VPC CIDR 區塊，但您可以將不重疊的其他 CIDR 區塊新增至 VPC。無法變更子網路 IPv4 CIDR，但可以變更 IPv6 CIDR。

- 最大可能的 VPC CIDR 區塊是 /16，最小的是 /28。
- 考慮使用其他連線網路 (VPC、內部部署或其他雲端供應商)，並確保不重疊的 IP 位址空間。如需詳細資訊，請參閱 [REL02-BP05 在連線的所有私有位址空間中強制使用不重疊的私有 IP 位址範圍](#)。

預期成果：可擴展的 IP 子網路可以幫助您適應未來的成長，並避免不必要的浪費。

常見的反模式：

- 未能考慮未來的成長，導致 CIDR 區塊太小且需要重新設定，進而可能造成停機。
- 錯誤預估 Elastic Load Balancer 可以使用的 IP 位址數量。
- 在相同的子網路中部署許多高流量負載平衡器
- 使用自動擴展機制，同時無法監控 IP 位址使用情況。
- 定義過大的 CIDR 範圍遠遠超出了未來的成長預期，這可能會導致難以與具有重疊位址範圍的其他網路進行對等互連。

建立此最佳實務的優勢：如此可確保您可以適應工作負載的增長，並在向上擴展時繼續提供可用性。

未建立此最佳實務時的曝險等級：中

實作指引

規劃網路以適應增長、法規要求以及與其他網路整合。增長可能會被低估，法規要求可能會發生變化，並且如果沒有適當的規劃，採購或私有網路連線可能會難以實作。

- 根據您的服務要求、延遲、法規和災難復原 (DR) 要求，選取相關的 AWS 帳戶和區域。
- 確定您對區域 VPC 部署的需求。
- 確定 VPC 的大小。
 - 確定是否要部署多 VPC 連線。
 - [什麼是 Transit Gateway？](#)
 - [單區域多 VPC 連線](#)
 - 確定您是否需要區隔聯網以滿足法規要求。
 - 使用適當大小的 CIDR 區塊製作 VPC，以滿足當前和未來需求。
 - 如果您有未知的增長預測，您可能希望選擇較大的 CIDR 區塊，以減少未來重新配置的可能性
 - 請考慮使用子網路的 [IPv6 定址](#) 做為雙堆疊 VPC 的一部分。IPv6 非常適合在包含暫時性執行個體或容器的私有子網路中使用，否則需要大量 IPv4 位址。

資源

相關 Well-Architected 的最佳實務：

- [REL02-BP05 在連線的所有私有位址空間中強制使用不重疊的私有 IP 位址範圍](#)

相關文件：

- [APN 合作夥伴：可以幫助您規劃聯網的合作夥伴](#)
- [AWS Marketplace 適用於網路基礎設施的](#)
- [Amazon Virtual Private Cloud 連線選項白皮書](#)
- [多個資料中心 HA 網路連線能力](#)
- [單區域多 VPC 連線](#)
- [什麼是 Amazon VPC？](#)
- [上的 IPv6 AWS](#)
- [參考架構上的 IPv6](#)
- [Amazon Elastic Kubernetes Service 推出 IPv6 支援](#)
- [VPC 的建議 - Classic Load Balancer](#)
- [可用區域子網路 - Application Load Balancer](#)
- [可用區域 - Network Load Balancer](#)

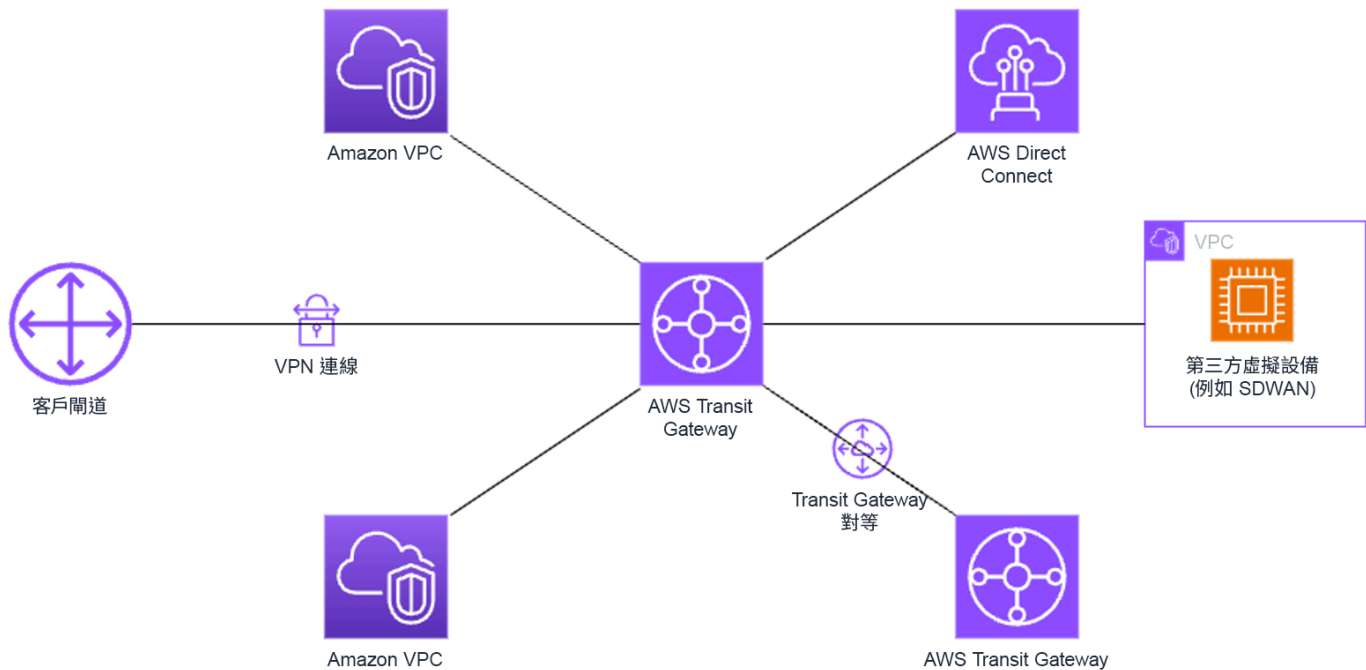
相關影片：

- [AWS re:Invent 2018：進階 VPC 設計及 Amazon VPC 的新功能 \(NET303\)](#)
- [AWS re:Invent 2019：適用於許多 VPC 的 AWS Transit Gateway 參考架構 \(NET406-R1\)](#)
- [AWS re:Invent 2023: AWS 是否為下一步做好準備？設計網路實現增長和靈活性 \(NET310\)](#)

REL02-BP04 偏好軸輻式拓撲而非多對多網狀拓撲

連接多個私有網路 (例如虛擬私有雲端 (VPC) 和內部部署網路時，請選擇軸輻式拓撲而非網狀拓撲。與網狀拓撲不同，其中每個網路都直接連接到其他網路並增加了複雜性和管理開銷，軸輻式架構會透過單一中樞集中連線。這種集中化簡化了網路結構，並增強了其可操作性、可擴展性和控制能力。

AWS Transit Gateway 是一項受管理、可擴展且高可用性的服務，專為在 AWS 上建構軸輻式網路而設計。它可作為網路的中心樞紐，提供網路分段、集中式路由以及與雲端和內部部署環境的簡化連線。下圖說明如何使用 AWS Transit Gateway 來建置軸輻式拓撲。



預期成果：您已透過中樞連接虛擬私有雲端 (VPC) 和內部部署網路。您能透過中樞設定對等連線，而此中樞可做為高度擴展的雲端路由器。您不需要處理複雜的對等關係，因此路由更簡單。網路之間的流量會經過加密，而且可以隔離網路。

常見的反模式：

- 您會建置複雜的網路對等規則。
- 您在彼此不應相互通訊的網路之間提供路由 (例如，彼此互不相依的個別工作負載)。
- 中樞執行個體的管控無效。

建立此最佳實務的優勢：隨著連線網路數量的增加，網狀連線的管理和擴充變得越來越具有挑戰性。網格架構會帶來額外的挑戰，例如額外的基礎設施元件、組態需求和部署考量。網格也會帶來額外的負荷，因為需要管理和監控資料平面和控制平面元件。您必須思考如何提供高度可用的網格架構、如何監控網格運作狀態和效能，以及如何處理網格元件的升級。

另一方面來說，軸輻式模型會在多個網路之間建立集中式流量路由。該模型提供更簡單的方法來管理和監控資料平面和控制平面元件。

未建立此最佳實務時的曝險等級：中

實作指引

若還沒有網路服務帳戶，請先建立帳戶。在組織的網路服務帳戶中設置中樞。此方法可讓網路工程師集中管理中樞。

軸輻式模型的中樞是做為虛擬路由器，可讓流量在您的虛擬私有雲端 (VPC) 與內部部署網路之間傳遞。這種方法降低了網路的複雜性，同時也更容易對網路問題進行故障診斷。

考慮您的網路設計，包括要互連的 VPC、AWS Direct Connect 和 Site-to-Site VPN 連線。

考慮針對每個傳輸閘道 VPC 連接使用個別子網路。對於每個子網路，請使用小型 CIDR (例如 /28)，以便擁有更多位址空間可供運算資源使用。此外，建立一個網路 ACL，並將它與中樞的所有關聯子網路建立關聯。在輸入和輸出方向上保持網路 ACL 開啟。

設計和實作您的路由表，以便僅在應通訊的網路之間提供路由。省略彼此不應相互通訊的網路之間的路由 (例如，彼此互不相依的個別工作負載之間)。

實作步驟

1. 規劃您的網路。決定要連線的網路，並確認這些網路提供的 CIDR 範圍不會重疊。
2. 建立 AWS Transit Gateway 並連接您的 VPC。
3. 如有需要，請建立 VPN 連線或 Direct Connect 閘道，並將其與 Transit Gateway 建立關聯。
4. 透過 Transit Gateway 路由表的組態，定義如何在連線的 VPC 和其他連線之間路由流量。
5. 使用 Amazon CloudWatch，視需要來監控和調整組態，以進行效能和成本最佳化。

資源

相關的最佳實務：

- [REL02-BP03 確保 IP 子網路配置帳戶具有擴展性和可用性](#)
- [REL02-BP05 在連線的所有私有位址空間中強制使用不重疊的私有 IP 位址範圍](#)

相關文件：

- [什麼是 Transit Gateway？](#)
- [傳輸閘道設計最佳實務](#)
- [建立可擴展且安全的多個 VPC AWS 網路架構](#)

- [使用 AWS Transit Gateway 區域間對等互連建置全球網路](#)
- [Amazon Virtual Private Cloud 連線選項](#)
- [APN 合作夥伴：可以幫助您規劃聯網的合作夥伴](#)
- [AWS Marketplace 適用於網路基礎設施的](#)

相關影片：

- [AWS re:Invent 2023 - AWS 聯網基礎](#)
- [AWS re:Invent 2023 - 進階 VPC 設計及新功能](#)

相關研討會：

- [AWS Transit Gateway 研討會](#)

REL02-BP05 在連線的所有私有地址空間中強制使用不重疊的私有 IP 位址範圍

在對等互連、透過 Transit Gateway 連線或透過 VPN 連線時，每個 VPC 的 IP 位址範圍不得重疊。避免 VPC 與內部部署環境或您所使用之其他雲端供應商之間出現 IP 位址衝突。您也必須有一種在需要時分配私有 IP 位址範圍的方法。IP 位址管理 (IPAM) 系統可以協助實現此自動化。

預期成果：

- VPC、內部部署環境或其他雲端供應商之間沒有 IP 位址範圍衝突。
- 適當的 IP 位址管理可以更輕鬆地擴展網路基礎設施，以適應網路需求的成長和變化。

常見的反模式：

- 在 VPC 中使用與內部部署、公司網路或其他雲端供應商相同的 IP 範圍
- 不追蹤用來部署工作負載之 VPC 的 IP 範圍。
- 仰賴手動 IP 位址管理程序，例如試算表。
- 大小過大或過小的 CIDR 區塊會導致 IP 位址浪費或位址空間不足以滿足您的工作負載。

建立此最佳實務的優勢：主動規劃網路，可確保在互連網路中不會出現多個相同的 IP 位址。這可防止使用不同應用程式的工作負載部分發生路由問題。

未建立此最佳實務時的曝險等級：中

實作指引

利用 IPAM，例如 [Amazon VPC IP Address Manager](#) 來監控和管理 CIDR 的使用情況。AWS Marketplace 也提供多套 IPAM 系統。評估您在 AWS 上的潛在使用情況，將 CIDR 範圍新增到現有 VPC，並建立 VPC 以允許計劃的用量增長。

實作步驟

- 擷取當前的 CIDR 消耗 (例如，VPC 和子網路等)。
 - 使用服務 API 操作來收集當前的 CIDR 消耗。
 - 使用 [Amazon VPC IP Address Manager 來探索資源](#)。
- 記錄當前的子網路用量。
 - 使用服務 API 操作來收集每個區域中每個 VPC 的 [子網路](#)。
 - 使用 [Amazon VPC IP Address Manager 來探索資源](#)。
- 記錄當前用量。
- 確定是否建立了任何重疊的 IP 範圍。
- 計算備用容量。
- 識別重疊的 IP 範圍。如果需要連線重疊的範圍，可以移轉至新的位址範圍，也可以考慮使用 [私有 NAT 閘道](#) 或 [AWS PrivateLink](#) 之類的技術。

資源

相關的最佳實務：

- [保護網路](#)

相關文件：

- [APN 合作夥伴：可以幫助您規劃聯網的合作夥伴](#)
- [適用於網路基礎設施的 AWS Marketplace](#)
- [Amazon Virtual Private Cloud 連線選項白皮書](#)
- [多個資料中心 HA 網路連線能力](#)
- [連線具有重疊 IP 範圍的網路](#)
- [什麼是 Amazon VPC？](#)
- [什麼是 IPAM？](#)

相關影片：

- [AWS re:Invent 2023 - 進階 VPC 設計及新功能](#)
- [AWS re:Invent 2019：適用於許多 VPC 的 AWS Transit Gateway 參考架構](#)
- [AWS re:Invent 2023 - 是否為下一步做好準備？設計網路實現增長和靈活性](#)
- [AWS re:Invent 2021 - {最新發佈} 在 AWS 上大規模管理 IP 位址](#)

工作負載架構

問題

- [REL 3. 如何設計您的工作負載服務架構？](#)
- [REL 4. 如何在分散式系統中設計防止失敗的互動？](#)
- [REL 5. 如何設計分散式系統中的互動以減輕或承受故障？](#)

REL 3. 如何設計您的工作負載服務架構？

使用服務導向架構 (SOA) 或微型服務架構，建置擴展性與可靠性高的工作負載。服務導向架構 (SOA) 是透過服務界面讓軟體元件可重複使用的做法。微型服務架構則進一步讓元件變得更小、更簡單。

最佳實務

- [REL03-BP01 選擇如何分割工作負載](#)
- [REL03-BP02 建置專注於特定業務領域和功能的服務](#)
- [REL03-BP03 提供每個的服務合約 API](#)

REL03-BP01 選擇如何分割工作負載

在確認應用程式的彈性要求時，工作負載劃分是很重要的。應盡可能避免整合型架構。您應審慎考量哪些應用程式元件可分解為微型服務。根據您的應用程式需求，這可能最終會盡可能結合服務導向架構 (SOA) 與微服務。可以無狀態的工作負載較有能力部署為微型服務。

預期成果：工作負載應可受支援、可擴展，並且盡可能地鬆散耦合。

在選擇如何劃分工作負載時，請在效益與複雜性之間取得平衡。讓新產品能率先推出的正確做法，不同於打造可從最初需求擴展的工作負載的做法。重構現有的整合型時，您必須考量應用程式如何能支援以無狀態為方向的解構。將服務細分為較小的服務，可讓明確定義的小型團隊加以開發及管理。但較小的服務可能會帶來複雜性，包括延遲可能增加、偵錯更複雜，以及運作負擔增加。

常見的反模式：

- [微服務 Death Star](#) 是一種特定情況：基本元件變得高度互相依賴，以致於只要有其中之一失敗，就會引發更加巨大的失敗，而導致元件像整合型一樣僵固且脆弱。

建立此實務的優勢：

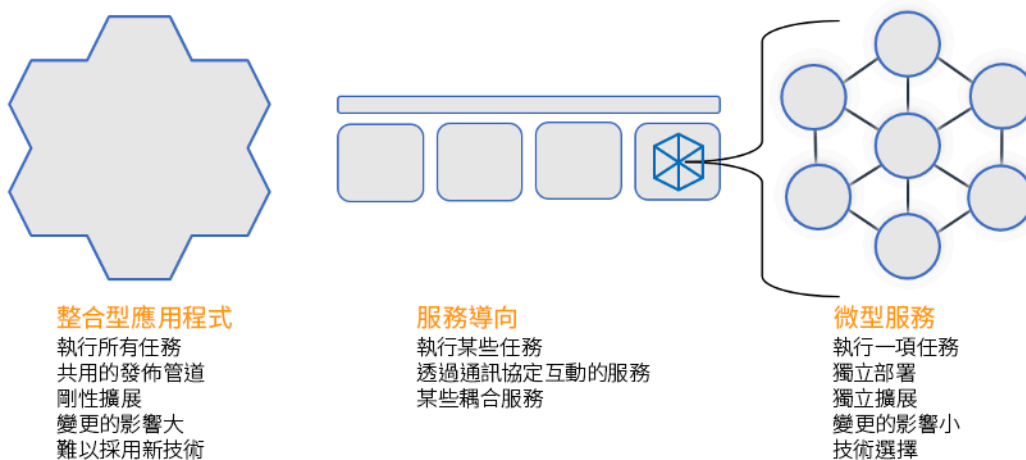
- 更明確的劃分可造就更高的靈活性、組織彈性及可擴展性。
- 降低服務中斷的影響。
- 應用程式元件可能會有不同的可用性要求，這一點可藉由更細微的劃分來支應。
- 為支援工作負載的團隊明確定義責任。

未建立此最佳實務時的曝險等級：高

實作指引

根據劃分工作負載的方式，選擇您的架構類型。選擇 SOA 或 微服務架構（或在極少數情況下，選擇單片架構）。即使您選擇從整體架構開始，仍必須確保其為模組化，並隨著產品隨著使用者採用而擴展，最終可以發展為 SOA 或 微服務。SOA 和 微服務分別提供較小的分割，這是現代可擴展且可靠的架構，但需要考慮權衡，特別是部署微服務架構時。

主要取捨之一，就是您現在擁有一種分散式運算架構，而其可能會增加您滿足使用者延遲要求的難度，並且在偵測和追蹤使用者互動方面還存在額外的複雜性。您可以利用 AWS X-Ray 來解決此問題。要考慮的另一個影響是，隨著您管理的應用程式數量增加，營運複雜性也隨之增加，因而需要部署多個獨立元件。



整合型、服務導向與微型服務架構

實作步驟

- 決定適當的架構以重構或建置您的應用程式。SOA 和 微服務分別提供較小的分割，這是現代可擴展且可靠的架構。SOA 對於實現更小的分割，同時避免微服務的一些複雜性，可能是一個很好的妥協。如需詳細資訊，請參閱 [Microservice Trade-Offs](#)。
- 如果您的工作負載適用於此類型，且您的組織可以提供支援，則應使用微型服務架構達成最佳的靈活性和可靠性。如需更多詳細資訊，請參閱[在上實作 Microservices AWS](#)。
- 考慮遵循 [Strangler Fig 模式](#)，將整體重構為較小的組件。這涉及逐步以新的應用程式和服務取代特定的應用程式元件。[AWS Migration Hub Refactor Spaces](#) 可充當增量重構的起點。如需詳細資訊，親參閱 [Seamlessly migrate on-premises legacy workloads using a strangler pattern](#)。
- 實作微服務可能需要服務探索機制，以允許這些分散式服務彼此通訊。[AWS App Mesh](#) 可與服務導向的架構搭配使用，以提供可靠的服務探索和存取。[AWS Cloud Map](#)也可用於動態DNS的 型服務探索。
- 如果您要從整體遷移至 SOA，[Amazon MQ](#) 可以在重新設計雲端中的舊版應用程式時，協助橋接差距，作為服務匯流排。
- 對於具有單一共用資料庫的現有整合型，請選擇如何將資料重新組織為較小的區段。此時可以按業務單位、存取模式或資料結構來劃分。在重構程序中，您應該選擇使用關聯式或非關聯式（否 SQL）類型的資料庫繼續。如需詳細資訊，請參閱[從 SQL到否 SQL](#)。

實作計劃的工作量：高

資源

相關的最佳實務：

- [REL03-BP02 建置專注於特定業務領域和功能的服務](#)

相關文件：

- [Amazon API Gateway：設定RESTAPI使用開啟API](#)
- [什麼是服務導向架構？](#)
- [有界限的環境 \(領域驅動型設計的集中模式\)](#)
- [在上實作 Microservices AWS](#)
- [微型服務權衡](#)

- [微型服務 - 此新架構術語的定義](#)
- [上的微服務 AWS](#)
- [什麼是 AWS App Mesh ?](#)

相關範例：

- [迭代應用程式現代化研討會](#)

相關影片：

- [在 上使用 Microservices 實現卓越 AWS](#)

REL03-BP02 建置專注於特定業務領域和功能的服務

服務導向架構 (SOA) 會定義服務，具有依商業需求定義的明訂功能。微型服務使用領域模型和有界限的環境，沿著業務環境界限繪製服務界限。專注於業務領域和功能，有助於團隊為其服務定義獨立的可靠性要求。有界限的環境可隔離和封裝商業邏輯，讓團隊更適切地推論如何處理失敗。

預期成果：工程師和業務利益相關者共同定義有界限的環境，並將其用來設計系統，作為滿足特定業務功能的服務。這些團隊使用既定的做法 (如事件風暴) 來定義要求。新的應用程式設計為服務妥善定義的界限和鬆散耦合。現有的整合型服務被分解為[有界限的環境](#)，系統設計轉向 SOA 或微服務架構。整合型服務重構時，會套用已建立的方法 (如 Bubble 環境) 和整合型分解模式。

領域導向服務會以一或多個不共用狀態的程序執行。它們會單獨回應需求的波動，並根據領域的特定要求來處理錯誤情境。

常見的反模式：

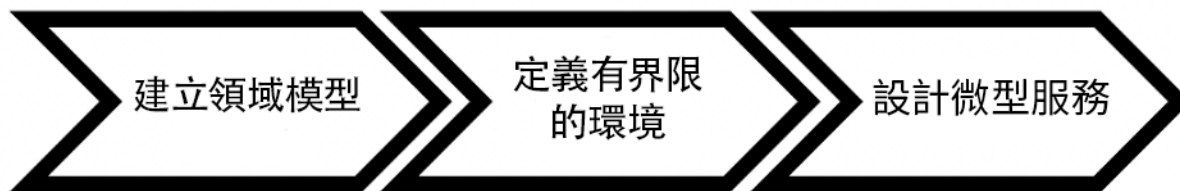
- 團隊是依據特定技術領域 (例如 UI 和 UX、中介軟體或資料庫) 組成的，而不是特定的業務領域。
- 應用程式跨多個領域責任。跨有界限環境的服務可能更難以維護，需要較大量的測試工作，且需要多個領域團隊參與軟體更新。
- 領域相依性 (例如領域實體程式庫) 會跨服務共用，因此一個服務領域出現變更時，需要變更其他服務領域
- 服務合約和商業邏輯無法以通用且一致的領域語言來表達實體，因此會導致翻譯層級使系統複雜化，並增加偵錯工作。

建立此最佳實務的優勢：應用程式設計為獨立的服務，受到業務領域限制，並使用共同的商務語言。服務可以單獨測試和部署。服務符合實作領域的特定恢復能力要求。

未建立此最佳實務時的曝險等級：高

實作指引

領域驅動設計 (DDD) 是依據業務領域設計和建置軟體的基礎方法。在建置專注於業務領域的服務時，使用現有架構將有所幫助。使用現有的整合型應用程式時，您可以利用分解模式提供已建立的技術，將應用程式現代化為服務。



領域驅動的設計

實作步驟

- 團隊可舉行[事件風暴](#)研討會，以便箋格式快速識別事件、命令、彙總和領域。
- 在網域環境中形成網域實體和函數之後，可以使用[有界限的環境](#)將網域劃分為服務，其中共用相似特徵和屬性的實體會分組在一起。隨著此模型劃分成多個環境，如何界定微型服務界限的範本便會浮現。
 - 例如，Amazon.com 網站實體可能包括包裝、交付、排程、價格、折扣和貨幣。
 - 包裝、交付和排程會分組到出貨環境中，而價格、折扣和貨幣則分組到訂價環境中。
- [將整合型服務分解為微型服務](#)概述了重構微型服務的模式。按業務功能、子領域或交易使用分解的模式，會與領域驅動的方法保持一致。
- 諸如 [bubble 環境](#) 等戰術可讓您在現有或舊版應用程式中導入 DDD，而無須預先重寫和對 DDD 完整承諾。在 bubble 環境方法中，使用服務映射和協調或[防損毀層](#)來建立一個小的有界限環境，可保護新定義的網域模型免受外部影響。

在團隊執行領域分析並定義實體和服務合約之後，他們可以利用 AWS 服務將其領域導向設計實作為雲端架構服務。

- 藉由定義執行領域商務規則的測試來起始您的開發。測試驅動的開發 (TDD) 和行為驅動的開發 (BDD) 可協助團隊將服務著重於解決業務問題上。

- 選取最符合您的企業網域需求和[微服務架構](#)的 [AWS 服務](#)：
 - [AWS Serverless](#) 讓您的團隊專注於特定的領域邏輯，而不是管理伺服器 and 基礎設施。
 - [AWS 的容器](#) 可簡化基礎設施的管理，讓您得以專注在您的網域要求上。
 - [專用資料庫](#) 協助您根據領域要求找出最適合的資料庫類型。
- [在 AWS 上建置六邊形架構](#)，它概述了一個框架，用以將業務邏輯建置到從業務領域回溯運作的服務中，以滿足功能要求，然後附加整合適配器。使用 AWS 服務將介面詳細資訊與商業邏輯分開的模式，可協助團隊專注於領域功能及改善軟體品質。

資源

相關的最佳實務：

- [REL03-BP01 選擇如何分割工作負載](#)
- [REL03-BP03 提供每個的服務合約 API](#)

相關文件：

- [AWS 微型服務](#)
- [實作 AWS 上的微型服務](#)
- [如何將整合型服務分成微型服務](#)
- [在置身於舊式系統時開始使用 DDD](#)
- [網域驅動設計：解決軟體核心的複雜性](#)
- [在 AWS 上建置六邊形架構](#)
- [將整合型服務分解為微型服務](#)
- [事件風暴](#)
- [有界限的環境之間的訊息](#)
- [微型服務](#)
- [測試驅動的開發](#)
- [行為驅動的開發](#)

相關範例：

- [在 AWS 上設計雲端原生微型服務 \(從 DDD/EventStormingWorkshop\)](#)

相關工具：

- [AWS 雲端 資料庫](#)
- [AWS 上的無伺服器](#)
- [AWS 上的容器](#)

REL03-BP03 提供每個的服務合約 API

服務合約是機器可讀取API定義中定義的API生產者和消費者之間的書面協議。合約版本控制策略可讓消費者繼續使用現有的，API並在準備好API時將其應用程式遷移至較新的應用程式。只要遵守合約，就隨時可執行生產者部署。服務團隊可以使用他們選擇的技術堆疊來滿足API合約。

預期結果：使用服務導向或微服務架構建置的應用程式可以獨立運作，同時具有整合的執行時間相依性。當雙方遵循共同API合約時，部署到API消費者或生產者的變更不會中斷整體系統的穩定性。透過服務通訊的元件APIs可以執行獨立的功能版本、升級至執行時間相依性，或容錯移轉至災難復原（DR）站台，彼此幾乎沒有影響。此外，離散服務能夠獨立擴展而滿足資源需求，不需要其他服務一起擴展。

常見的反模式：

- 建立APIs不含強烈輸入結構描述的服務。這會導致APIs無法用來產生無法以程式設計方式驗證的API繫結和承載。
- 不採用版本控制策略，這會強制API消費者在服務合約演進時更新和發行或失敗。
- 錯誤訊息會透露基礎服務實作的詳細資訊，而不是說明領域環境和語言中的整合失敗。
- 不使用API合約來開發測試案例和模擬API實作，以允許獨立測試服務元件。

建立此最佳實務的好處：由透過API服務合約通訊的元件組成的分散式系統可以提高可靠性。開發人員可以在開發程序的早期發現潛在的問題，並在編譯期間進行類型檢查，以確認請求和回應遵循API合約和必填欄位。API合約為APIs和提供者提供了明確的自我記錄介面，讓不同的系統和程式設計語言之間具有更好的互通性。

未建立此最佳實務時的曝險等級：中

實作指引

識別業務網域並確定工作負載分割後，您就可以開發服務APIs。首先，定義的機器可讀取服務合約APIs，然後實作API版本控制策略。當您準備好透過REST、GraphQL或非同步事件等常見通訊協定整合服務時，您可以將AWS服務整合到您的架構中，以將元件與強式API合約整合。

AWS 服務API對象的服務

將 [Amazon API Gateway](#)、[AWS AppSync](#)和 [Amazon EventBridge](#) 等 AWS 服務整合到您的架構中，以在應用程式中使用API服務合約。Amazon API Gateway 可協助您與直接原生 AWS 服務和其他 Web 服務整合。API Gateway 支援 [OpenAPI 規格](#)和 versioning。AWS AppSync 是您透過定義 [GraphQL](#) 結構描述來定義查詢、突變和訂閱的服務界面所設定的受管 GraphQL 端點。Amazon EventBridge 使用事件結構描述來定義事件，並為您的事件產生程式碼繫結。

實作步驟

- 首先，為您的 定義合約API。合約會表達 的功能API，並定義API輸入和輸出的強烈輸入資料物件和欄位。
- 當您APIs在 API Gateway 中設定 時，您可以匯入和匯出端點的開放API規格。
 - [匯入 OpenAPI 定義](#)可簡化 的建立，API並可整合 AWS 基礎設施作為程式碼工具，例如 [AWS Serverless Application Model](#)和 [AWS Cloud Development Kit \(AWS CDK\)](#)。
 - [匯出API定義](#)可簡化與API測試工具的整合，並為服務消費者提供整合規格。
- 您可以使用 APIs AWS AppSync 定義 GraphQL [結構描述檔案來定義和管理 GraphQL](#)，以產生合約界面，並簡化與複雜REST模型、多個資料庫資料表或舊版服務的互動。
- [AWS Amplify](#) 與 整合的專案 AWS AppSync 會產生強式 JavaScript 查詢檔案，用於您的應用程式，以及 [Amazon DynamoDB](#) 資料表的 AWS AppSync GraphQL 用戶端程式庫。
- 當您從 Amazon 取用服務事件時 EventBridge，事件會遵守已存在於結構描述登錄檔中的結構描述，或是您使用 OpenAPI Spec 定義的結構描述。使用登錄中定義的結構描述時，您也可以從結構描述合約產生用戶端繫結，以將程式碼與事件整合。
- 擴展或版本您的 API。新增可使用選用欄位或必要欄位的預設值設定的欄位時，延伸 API 是更簡單的選項。
 - JSON REST和 GraphQL 等通訊協定的 型合約非常適合合約延伸。
 - XML 等通訊協定的 型合約SOAP應與 服務消費者進行測試，以判斷合約延伸的可行性。
- 版本控制 時API，請考慮實作 Proxy 版本控制，其中使用外觀來支援版本，以便在單一程式碼庫中維護邏輯。
 - 使用 API Gateway，您可以使用[請求和回應映射](#)，透過建立外觀來提供新欄位的預設值，或從請求或回應中刪除的欄位，以簡化吸收合約變更。透過此方法，基礎服務可以維護單一程式碼基底。

資源

相關的最佳實務：

- [REL03-BP01 選擇如何分割工作負載](#)
- [REL03-BP02 建置專注於特定業務領域和功能的服務](#)
- [REL04-BP02 實作鬆散耦合相依性](#)
- [REL05-BP03 控制和限制重試呼叫](#)
- [REL05-BP05 設定用戶端逾時](#)

相關文件：

- [什麼是 API \(應用程式程式設計介面 \) ？](#)
- [在 上實作 Microservices AWS](#)
- [微型服務權衡](#)
- [微型服務 - 此新架構術語的定義](#)
- [上的微服務 AWS](#)
- [使用 API Gateway 擴充功能開啟API](#)
- [開啟 API-規格](#)
- [GraphQL：結構描述和類型](#)
- [Amazon EventBridge 程式碼繫結](#)

相關範例：

- [Amazon API Gateway：設定RESTAPI使用開啟API](#)
- [使用 Open 的 Amazon API Gateway 至 Amazon DynamoDB CRUD 應用程式API](#)
- [無伺服器時代的現代應用程式整合模式：APIGateway Service Integration](#)
- [使用 Amazon 實作標頭型API閘道版本控制 CloudFront](#)
- [AWS AppSync：建置用戶端應用程式](#)

相關影片：

- [使用在 中開啟API AWS SAM 管理API閘道](#)

相關工具：

- [Amazon API Gateway](#)

- [AWS AppSync](#)
- [Amazon EventBridge](#)

REL 4. 如何在分散式系統中設計防止失敗的互動？

分散式系統仰賴通訊網路將伺服器或服務等元件互相連線。儘管這些網路中出現資料遺失或延遲，但工作負載仍必須可靠地運作。分散式系統的元件必須以不會對其他元件或工作負載造成負面影響的方式運作。這些最佳實務可防止故障，並改善平均故障間隔時間 (MTBF)。

最佳實務

- [REL04-BP01 識別您依賴的分散式系統種類](#)
- [REL04-BP02 實作鬆散耦合相依性](#)
- [REL04-BP03 持續工作](#)
- [REL04-BP04 讓變異操作冪等](#)

REL04-BP01 識別您依賴的分散式系統種類

分散式系統可以是同步、非同步或批次處理。同步系統必須盡快處理要求，並透過使用 HTTP/S、REST 或遠端程序呼叫 (RPC) 通訊協定進行同步請求和回應呼叫，以便相互通訊。非同步系統透過中間服務以非同步方式交換資料來相互通訊，而不需要耦合個別系統。批處理系統接收大量輸入資料，無須人工介入即可執行自動化資料處理，並產生輸出資料。

預期成果：設計與同步、非同步和批次相依性有效互動的工作負載。

常見的反模式：

- 工作負載會無限期地等待來自其相依性的回應，這可能會導致工作負載用戶端逾時，而不知道是否已收到其請求。
- 工作負載使用可同步呼叫彼此的一系列相依系統。這需要每個系統都可供使用，並在整個系列成功之前成功處理請求，從而導致潛在的脆弱行為和整體可用性。
- 工作負載會以非同步方式與其相依性進行通訊，並依賴於僅保證傳遞訊息一次的概念，而且通常仍然可以接收重複的訊息。
- 工作負載不使用適當的批次排程工具，並允許同時執行相同的批次工作。

建立此最佳實務的優勢：對於特定的工作負載來說，在同步、非同步和批次之間實作一種或多種通訊方式很常見。此最佳實務可協助您識別與每種通訊方式相關的不同權衡，讓您的工作負載能夠容忍其任何相依性中斷。

未建立此最佳實務時的風險暴露等級：高

實作指引

下列各節包含每種相依性的一般和特定實作指引。

一般指引

- 請確定相依性所提供的效能與可靠性服務水準目標 (SLO) 符合工作負載的效能和可靠性需求。
- 使用 [AWS 可觀測性服務](#) 來 [監控回應時間和錯誤率](#)，以確保您的相依性在工作負載所需的層級提供服務。
- 識別工作負載在與其相依性通訊時可能面臨的潛在挑戰。分散式系統 [面臨各種挑戰](#)，可能會增加架構複雜性、營運負擔和成本。常見的挑戰包括延遲、網路中斷、資料遺失、擴展和資料複寫延遲。
- 實作強大的錯誤處理和 [日誌記錄](#)，以協助您在相依性遇到問題時對問題進行疑難排解。

同步相依性

在同步通訊中，工作負載會將請求傳送至其相依性，並封鎖等待回應的操作。當其相依性收到請求時，它會嘗試盡快處理它，並將回應傳送回您的工作負載。同步通訊的一個重大挑戰是它會導致時間耦合，這需要您的工作負載及其相依性同時可用。當您的工作負載需要與其相依性同步通訊時，請考慮下列指引：

- 您的工作負載不應該依賴多個同步相依性來執行單一函數。此相依性系列增加了整體的脆弱性，因為路徑中的所有相依性都必須可用，才能順利完成請求。
- 當相依性狀態不良或無法使用時，請確定處理錯誤並重試策略。避免使用雙模態行為。雙模態行為是指工作負載在正常和故障模式下呈現不同行為的情況。如需雙模態行為的詳細資訊，請參閱 [REL11-BP05 使用靜態穩定性來防止雙模態行為](#)。
- 請記住，快速檢錯好於讓工作負載等待。例如，[AWS Lambda 開發人員指南](#) 說明如何在呼叫 Lambda 函數時處理重試和失敗。
- 當工作負載呼叫其相依性時設定逾時。此技術可避免等待太長時間或無限期等待回應。如需此問題的有用討論，請參閱 [Tuning AWS Java SDK HTTP request settings for latency-aware Amazon DynamoDB applications](#)。
- 盡可能減少從工作負載到其相依性的呼叫次數，以滿足單一請求。在兩者之間進行聊天呼叫會增加耦合和延遲。

異步相依性

若要暫時將工作負載與其相依性分離，它們應該以非同步方式進行通訊。使用非同步方法，您的工作負載可以繼續執行任何其他處理，而不必等待其相依性或相依性系列來傳送回應。

當您的工作負載需要與其相依性異步通訊時，請考慮下列指引：

- 根據您的使用案例和需求來決定是使用訊息傳遞還是事件串流。[訊息傳遞](#)可讓您的工作負載透過訊息代理程式傳送和接收訊息，藉此與其相依性進行通訊。[事件串流](#)可讓您的工作負載及其相依性使用串流服務來發佈和訂閱事件 (以連續資料串流形式傳送)，這些事件需要盡快處理。
- 訊息傳遞和事件串流處理訊息的方式不同，因此您需要根據下列內容做出權衡決策：
 - 訊息優先級：訊息代理程式可以在正常訊息之前處理高優先級訊息。在事件串流中，所有訊息都有相同的優先級。
 - 訊息消耗：訊息代理程式確保消費者收到消息。事件串流消費者必須追蹤他們讀取的最後一則訊息。
 - 訊息排序：使用訊息傳遞時，不能保證以傳送的確切順序接收訊息，除非您使用先進先出 (FIFO) 方法。事件串流永遠會保留產生資料的順序。
 - 訊息刪除：使用訊息傳遞時，消費者必須在處理訊息後刪除它。事件串流服務會將訊息附加至串流，並保留在其中，直到訊息的保留期過期為止。此刪除政策使得事件串流適合重播訊息。
- 定義工作負載如何知道其相依性何時完成其工作。例如，當工作負載[以非同步方式調用 Lambda 函數](#)時，Lambda 會將事件放在佇列中，並傳回成功回應，但沒有額外資訊。處理完成後，Lambda 函數可以[將結果傳送至目的地](#)，並根據成功或失敗進行設定。
- 透過利用冪等性來構建工作負載以處理重複訊息。冪等性意味著即使為相同訊息產生多次工作負載，工作負載的結果也不會變更。重要的是要指出，如果發生網路故障或尚未收到確認，[訊息傳遞](#)或[串流](#)服務將重新傳遞訊息。
- 如果您的工作負載未從其相依性中取得回應，則需要重新提交請求。請考慮限制重試次數，以保留工作負載的 CPU、記憶體和網路資源來處理其他請求。[AWS Lambda 文件](#)說明了如何處理非同步調用的錯誤。
- 運用適當的可觀測性、偵錯和追蹤工具，管理和操作工作負載的非同步通訊及其相依性。可以使用 [Amazon CloudWatch](#) 來監控[訊息傳遞](#)和[事件串流](#)服務。還可以使用 [AWS X-Ray](#) 檢測工作負載，以快速獲得疑難排解問題的[洞見](#)。

批處理相依性

批處理系統會取得輸入資料，啟動一系列作業來處理它，並產生一些輸出資料，無須人工介入。視資料大小而定，工作可能會執行幾分鐘，在某些情況下執行數天。當您的工作負載需要與其批處理相依性通訊時，請考慮下列指引：

- 定義工作負載執行批次工作的時間範圍。工作負載可以設定週期性模式來調用批次系統，例如，每小時或每月結束時。
- 確定資料輸入和已處理的資料輸出的位置。選擇可讓您的工作負載大規模讀取和寫入檔案的儲存服務，例如 [Amazon Simple Storage Services \(Amazon S3\)](#)、[Amazon Elastic File System \(Amazon EFS\)](#) 和 [Amazon FSx for Lustre](#)。
- 如果您的工作負載需要調用多個批次工作，可以利用 [AWS Step Functions](#) 簡化在 AWS 或內部部署中執行的批次工作的協調作業。此[範例專案](#)示範使用 Step Functions、[AWS Batch](#) 和 Lambda 協調批次任務。
- 監控批次任務以尋找異常情況，例如任務所花費的時間超過應該完成的時間。可以使用 [CloudWatch Container Insights](#) 等工具來監控 AWS Batch 環境和任務。在這種情況下，工作負載將從一開始就停止下一個任務，並通知相關人員有關例外情況。

資源

相關文件：

- [AWS 雲端 操作：監控和可觀測性](#)
- [Amazon 建置者資料中心：分散式系統的挑戰](#)
- [REL11-BP05 使用靜態穩定性來防止雙模態行為](#)
- [AWS Lambda 開發人員指南：AWS Lambda 中的錯誤處理和自動重試](#)
- [針對延遲感知的 Amazon DynamoDB 應用程式調整 AWS Java SDK HTTP 請求設定](#)
- [AWS 訊息傳遞](#)
- [什麼是串流資料？](#)
- [AWS Lambda 開發人員指南：非同步調用](#)
- [Amazon Simple Queue Service 常見問答集：FIFO 佇列](#)
- [Amazon Kinesis Data Streams 開發人員指南：處理重複記錄](#)
- [Amazon Simple Queue Service 開發人員指南：適用於 Amazon SQS 的可用 CloudWatch 指標](#)
- [Amazon Kinesis Data Streams 開發人員指南：使用 Amazon CloudWatch 監控 Amazon Kinesis Data Streams 服務](#)

- [AWS X-Ray 開發人員指南：AWS X-Ray 概念](#)
- [GitHub 上的 AWS 範例：AWS Step 函數，複雜的協調器應用程式](#)
- [AWS Batch 使用者指南：AWS Batch CloudWatch Container Insights](#)

相關影片：

- [AWS Summit SF 2022 - 使用 AWS 獲得全堆疊可觀測性和應用程式監控 \(COP310\)](#)

相關工具：

- [Amazon CloudWatch](#)
- [Amazon CloudWatch Logs \(\)](#)
- [AWS X-Ray](#)
- [Amazon Simple Storage Services \(Amazon S3\)](#)
- [Amazon Elastic File System \(Amazon EFS\)](#)
- [Amazon FSx for Lustre](#)
- [AWS Step Functions](#)
- [AWS Batch](#)

REL04-BP02 實作鬆散耦合相依性

佇列系統、串流系統、工作流程和負載平衡器之間具有鬆散耦合的相依性。鬆耦合有助於將某個元件的行為與依賴它的其他元件隔離，進而提高彈性和敏捷性。

解除相依性 (例如佇列系統、串流系統和工作流程) 有助於將變更或失敗對系統造成的影響降到最低。這種分離使組件的行為不會影響依賴它的其他組件，從而提高了彈性和敏捷性。

在緊耦合的系統中，對某個元件進行變更時，可能必須變更其他依賴此元件的元件，從而導致所有元件的效能降低。鬆耦合會破壞此相依性，因此相依元件只需要知道受版本控制的和已發布的界面。在相依性之間實作鬆耦合，可避免一個元件中的故障影響另一個元件。

鬆耦合可讓您修改程式碼或新增功能至某個元件，同時將依賴該元件的其他元件的風險降至最低。其還能讓您在元件層級提供細微的恢復能力，您可以橫向擴充，甚至是變更相依性的基礎實作。

若要透過鬆耦合進一步改善彈性，請盡可能讓元件採用非同步互動。此模型適用於不需要立即回應的任何互動，以及確認已註冊請求便以足夠的狀況。它涉及產生事件的一個元件和取用事件的另一個元件。

這兩個元件不會透過直接 point-to-point 互動整合，而是通常透過中繼耐用儲存層整合，例如 Amazon SQS 佇列、串流資料平台，例如 Amazon Kinesis 或 AWS Step Functions。

圖 4：佇列系統和負載平衡器之間具有鬆散耦合的相依性

Amazon SQS 佇列和 AWS Step Functions 只是新增中繼層進行鬆散耦合的兩種方式。事件驅動的架構也可以 AWS 雲端使用 Amazon 建置在中 EventBridge，這可以從用戶端（事件生產者）仰賴的服務（事件消費者）中抽象用戶端（事件生產者）。當您需要高輸送量、推送型訊息時，many-to-many Amazon Simple Notification Service（Amazon SNS）是有效的解決方案。使用 Amazon SNS 主題，您的發佈者系統可以將訊息散播到大量訂閱者端點以進行平行處理。

雖然佇列提供多項優勢，但在大多數硬式即時系統中，超過閾值時間（通常為秒）的請求應視為過時（用戶端已放棄且不再等待回應）且未處理。這樣才可以處理較新的（且可能仍有效的）請求。

預期成果：實作鬆耦合的相依性可將元件層級故障的影響降到最低，這有助於診斷和解決問題。它還能簡化開發週期，讓團隊在模組化層級實作變更，而不會影響依賴此元件之其他元件的效能。這種方法可讓您根據資源需求，以及對成本效益有所貢獻之元件的使用情況，在元件層級進行橫向擴充。

常見的反模式：

- 部署整合型工作負載。
- 直接在工作負載層 APIs 之間調用，而沒有容錯移轉或非同步處理請求的能力。
- 使用共用資料的緊耦合。鬆耦合系統應避免透過共用資料庫或其他形式的緊耦合資料儲存共用資料，這可能會重新引入緊耦合並阻礙可擴展性。
- 忽略反壓。當元件無法以相同的速率處理傳入的資料時，工作負載應該要有能力減緩或停止傳入的資料。

建立此最佳實務的優勢：鬆耦合有助於將某個元件的行為與依賴它的其他元件隔離，進而提高彈性和敏捷性。避免一個元件中的失敗影響其他元件。

未建立此最佳實務時的曝險等級：高

實作指引

實作鬆耦合相依性。有各種解決方案可讓您建置鬆耦合的應用程式。其中包括實作完全受管佇列、自動化工作流程、對事件做出反應 APIs 等服務，這些服務可協助隔離元件與其他元件的行為，進而提高復原能力和靈活性。

- 建置事件驅動架構：[Amazon EventBridge](#) 可協助您建置鬆散耦合和分散式事件驅動架構。

- 在分散式系統中實作佇列：您可以使用 [Amazon Simple Queue Service \(Amazon SQS \)](#) 來整合和解耦分散式系統。
- 將元件容器化為微服務：[微服務](#)可讓團隊建置由小型獨立元件組成的應用程式，這些元件透過定義明確的 進行通訊APIs。[Amazon Elastic Container Service \(Amazon ECS \)](#) 和 [Amazon Elastic Kubernetes Service \(Amazon EKS \)](#) 可協助您更快開始使用容器。
- 使用 Step Functions 管理工作流程：[Step Functions](#) 可協助您將多項 AWS 服務協調為彈性工作流程。
- 利用發佈訂閱（ pub/sub ）訊息架構：[Amazon Simple Notification Service \(Amazon SNS \)](#) 提供從發佈者到訂閱者（也稱為生產者和消費者）的訊息傳遞。

實作步驟

- 事件驅動架構中的元件會由事件啟動。事件是系統中發生的動作，例如使用者將某個商品新增至購物車。動作成功時會產生可啟動系統下一個元件的事件。
 - [使用 Amazon 建置事件驅動應用程式 EventBridge](#)
 - [AWS re : Invent 2022 - 使用 Amazon 設計事件驅動整合 EventBridge](#)
- 分散式傳訊系統有三個需要針對佇列型架構來實作的主要部分。其中包括分散式系統的元件、用於解耦的佇列（在 Amazon SQS 伺服器上分佈），以及佇列中的訊息。典型的系統中有負責將訊息啟動至佇列的生產者，以及從佇列接收訊息的取用者。佇列會在多個 Amazon SQS 伺服器之間存放訊息，以進行備援。
 - [基本 Amazon SQS架構](#)
 - [使用 Amazon Simple Queue Service 在分散式應用程式之間傳送訊息](#)
- 充分利用的微型服務會增強可維護性並提高可擴展性，因為鬆耦合元件由獨立團隊管理。其還能夠在發生變更時隔離單一元件的行為。
 - [在上實作 Microservices AWS](#)
 - [開始建構吧！使用容器建構微型服務](#)
- AWS Step Functions 您可以透過 建置分散式應用程式、自動化程序、協調微服務等。將多個元件協同運作到自動化工作流程中可讓您解耦應用程式中的相依性。
 - [使用 AWS Step Functions 和 建立無伺服器工作流程 AWS Lambda](#)
 - [入門 AWS Step Functions](#)

資源

相關文件：

- [Amazon EC2：確保錯位](#)
- [Amazon 建置者資料中心：分散式系統的挑戰](#)
- [Amazon 建置者資料中心：可靠性、持續工作以及咖啡時刻](#)
- [什麼是 Amazon EventBridge？](#)
- [什麼是 Amazon Simple Queue Service？](#)
- [結束您的整合型架構](#)
- [使用 AWS Step Functions 和 Amazon 協調佇列型 Microservices SQS](#)
- [基本 Amazon SQS架構](#)
- [佇列式架構](#)

相關影片：

- [AWS 2019 年紐約高峰會：事件驅動架構和 Amazon 簡介 EventBridge \(MAD205 \)](#)
- [AWS re：Invent 2018：閉環和開場思維：如何控制系統，無論大小 ARC337 \(包括鬆散的耦合、固定工作、靜態穩定性 \)](#)
- [AWS re：Invent 2019：移至事件驅動的架構 \(SVS308 \)](#)
- [AWS re：Invent 2019：使用 Amazon SQS和 Lambda 可擴展的無伺服器事件驅動應用程式](#)
- [AWS re：Invent 2022 - 使用 Amazon 設計事件驅動整合 EventBridge](#)
- [AWS re：Invent 2017：Elastic Load Balancing Deep Dive 和最佳實務](#)

REL04-BP03 持續工作

負載大幅快速變更時，系統可能會發生故障。例如，如果您的工作負載正在執行運作狀態檢查，監控數千部伺服器的運作狀態，應該每次傳送相同大小的承載 (目前狀態的完整快照)。無論伺服器全無故障或全部出現故障，運作狀態檢查系統都會持續執行工作，而無大幅快速變更。

例如，如果運作狀態檢查系統正在監控 100,000 部伺服器，則在一般輕型伺服器失敗率下，其負載為額定值。不過，如果重大事件讓一半的伺服器運作狀況不良，則運作狀態檢查系統會因嘗試更新通知系統並向其用戶端溝通狀態，而承受不住負載。因此，運作狀態檢查系統應每次都傳送目前狀態的完整快照。100,000 個伺服器運作狀態 (每個以一位元表示) 只是 12.5 KB 的承載。無論沒有伺服器發生故障，還是全部發生故障，運作狀態檢查系統都會持續執行工作，而大型的快速變更也不會對系統穩定性造成威脅。實際上，這是 Amazon Route 53 處理端點運作狀態檢查 (例如 IP 位址) 的方式，以判斷最終使用者如何路由到端點。

未建立此最佳實務時的曝險等級：低

實作指引

- 執行持續工作，以便當負載大量快速變更時，系統不會發生故障。
- 實作鬆耦合相依性。佇列系統、串流系統、工作流程和負載平衡器之間具有鬆散耦合的相依性。鬆耦合有助於將某個元件的行為與依賴它的其他元件隔離，進而提高彈性和敏捷性。
- [Amazon 建置者資料中心：可靠性、持續工作以及咖啡時刻](#)
- [AWS re：Invent 2018：閉環和開場思維：如何控制大大小小的系統 ARC337（包括持續工作）](#)
 - 針對監控 100,000 部伺服器的運作狀態檢查系統範例，請設計工作負載，以便無論成功或失敗次數為何，承載大小都能保持不變。

資源

相關文件：

- [Amazon EC2：確保錯位](#)
- [Amazon 建置者資料中心：分散式系統的挑戰](#)
- [Amazon 建置者資料中心：可靠性、持續工作以及咖啡時刻](#)

相關影片：

- [AWS 2019 年紐約高峰會：事件驅動架構和 Amazon 簡介 EventBridge（MAD205）](#)
- [AWS re：Invent 2018：閉環和開場思維：如何控制大大小小的系統 ARC337（包括持續工作）](#)
- [AWS re：Invent 2018：閉環和開場思維：如何控制系統，大小 ARC337（包括鬆散的耦合、固定工作、靜態穩定性）](#)
- [AWS re：Invent 2019：移至事件驅動的架構（SVS308）](#)

REL04-BP04 讓變異操作冪等

冪等服務確保每個請求只處理一次，因此發出多個相同請求的效果，與發出單一請求的效果相同。如此一來，用戶端可更輕鬆地重試，而不用擔心多次錯誤地處理請求。為此，用戶端可以使用冪等性字元發出 API 請求，每次重複請求時，都會使用這個權杖。冪等服務 API 會使用字元傳回回應，而該回應與第一次完成請求時傳回的回應相同，即使系統的基礎狀態已改變也一樣。

在分散式系統中，相對容易的做法是最多執行一次動作（用戶端只發出一個請求）或至少執行一次動作（持續發出請求，直到用戶端確認成功）。但很難保證動作精準地執行一次，使得發出多個相同的請求與

發出單一請求效果相同。透過在 API 中使用冪等性字符，服務可以收到一次或多次變異請求，而不需建立重複的記錄或產生副作用。

預期成果：您擁有一致、完整記錄且廣泛採用的方法，以確保所有元件和服務之間的冪等性。

常見的反模式：

- 即使不需要，您仍一視同仁地套用冪等性。
- 您導入過於複雜的邏輯來實作冪等性。
- 您使用時間戳記做為冪等性的索引鍵。這樣可能會因為時鐘誤差或多個用戶端使用相同的時間戳記來套用變更，而造成不準確。
- 您為了冪等性而儲存整個承載。採用這個方法時，您會為每個請求儲存完整的資料承載，並在每個新請求中覆寫它。這可能會導致效能降低並影響可擴展性。
- 您在服務之間產生不一致的索引鍵。若索引鍵不一致，服務可能無法辨識重複的請求，進而導致意外的結果。

建立此最佳實務的優勢：

- 可擴展性更大：系統可以處理重試和重複的請求，而不需執行額外的邏輯或複雜的狀態管理。
- 加強可靠性：冪等性可協助服務以一致的方式處理多個相同的請求，進而降低意外副作用或重複記錄的風險。這在經常發生網路故障和重試的分散式系統中尤其重要。
- 改善資料一致性：由於相同的請求會產生相同的回應，因此冪等性有助於在分散式系統之間保持資料一致性。這對於維護交易和操作的完整性相當重要。
- 錯誤處理：冪等性字符讓錯誤處理更簡單。如果用戶端因發生問題而未收到回應，它可以使用相同的冪等性字符安全地重新傳送請求。
- 操作透明度：冪等性可實現更有效的監控與記錄。服務可以透過冪等性字符來記錄請求，這樣更容易追蹤問題並進行偵錯。
- 簡化的 API 合約：它可以簡化用戶端和伺服器端系統之間的合約，並減少對資料處理錯誤的擔憂。

未建立此最佳實務時的曝險等級：中

實作指引

在分散式系統中，最多執行一次動作 (用戶端只發出一個請求) 或至少執行一次動作 (用戶端持續發出請求直到確認成功) 是相對容易的做法。不過，實作精確一次的行為並不容易。若要達成此目的，您的用戶端應針對每個請求產生並提供冪等性字符。

透過使用冪等性字符，服務就能區分新請求和重複的請求。當服務收到具有冪等性字符的請求時，它會檢查字符是否已使用。如果已使用字符，服務就會擷取並傳回預存回應。如果字符是新的，則服務會處理請求、一併儲存回應與字符，然後傳回回應。此機制會讓所有回應變成冪等，進而提高分散式系統的可靠性和一致性。

冪等性也是事件驅動架構的重要行為。這些架構通常由訊息佇列支援，例如 Amazon SQS、Amazon MQ、Amazon Kinesis Streams 或 Amazon Managed Streaming for Apache Kafka (MSK)。在某些情況下，只發布一次的訊息可能會意外傳遞超過一次。當發布者在訊息中產生並包含冪等性字符時，它會要求處理收到的任何重複訊息時，不會對相同訊息產生重複的動作。取用者應追蹤收到的每個字符，並忽略包含重複字符的訊息。

服務與取用者也應將收到的冪等性字符傳遞給其呼叫的任何下游服務。處理鏈中的每個下游服務都同樣負責確保實作冪等性，以免發生處理訊息超過一次的副作用。

實作步驟

1. 識別冪等操作

確定哪些操作需要冪等性。這些通常包括 POST、PUT 和 DELETE HTTP 方法，以及資料庫插入、更新或刪除操作。不會改變狀態的操作 (例如唯讀查詢) 通常不需要冪等性，除非有副作用。

2. 使用唯一識別碼

在傳送者傳送的每個冪等操作請求中包含唯一的字符，無論是直接在請求中，還是做為其中繼資料的一部分 (例如 HTTP 標頭)。這可讓接收者辨識和處理重複的請求或操作。常用於字符的識別碼包括[通用唯一識別碼 \(UUID\)](#) 和 [K-Sortable Unique Identifiers \(KSUID\)](#)。

3. 追蹤和管理狀態

維護工作負載中每個操作或請求的狀態。這可以透過在資料庫、快取或其他持久性存放區中儲存冪等性字符及對應的狀態 (如待處理、已完成或失敗) 來達成。此狀態資訊可讓工作負載識別和處理重複的請求或操作。

如有需要，可使用適當的並行控制機制來維持一致性和單元性，例如鎖定、交易或積極並行控制。包括記錄冪等性字符和執行所有與處理請求相關聯的改變操作程序。這樣做有助於防止競爭條件，並確認冪等操作正確執行。

定期從資料儲存區移除舊的冪等性字符，以管理儲存空間和效能。如果您的儲存系統提供支援，請考慮針對資料使用過期時間戳記 (通常稱為存留時間或 TTL 值)。重複使用冪等性字符的可能性會隨著時間而降低。

通常用於儲存冪等性字符和相關狀態的常見 AWS 儲存選項包括：

- Amazon DynamoDB：DynamoDB 是一項 NoSQL 資料庫服務，可提供低延遲效能和高可用性，因此非常適合儲存冪等性相關資料。DynamoDB 的索引鍵值和文件資料模型具備高效率儲存和擷取冪等性字符與相關聯狀態資訊的能力。如果您的應用程式在插入時設定 TTL 值，則 DynamoDB 也會自動使冪等性字符過期。
- Amazon ElastiCache：ElastiCache 可以儲存高輸送量、低延遲且成本低廉的冪等性字符。如果您的應用程式在插入時設定了 TTL 值，則 ElastiCache (Redis) 和 ElastiCache (Memcached) 也會自動讓冪等性字符過期。
- Amazon Relational Database Service (RDS)：您可以使用 Amazon RDS 來存放冪等性字符和相關的狀態資訊，特別是您的應用程式已將關聯式資料庫用於其他用途時。
- Amazon Simple Storage Service (S3)：Amazon S3 是高度可擴展且耐用的物件儲存服務，可用於存放冪等性字符和相關的中繼資料。S3 的版本控制功能在維護冪等操作的狀態方面特別實用。選擇儲存服務時，通常取決於以下因素：與冪等性相關的資料量、所需的效能特性、所須耐用性和可用性，以及冪等性機制與整體工作負載架構整合的方式。

4. 實作冪等操作

將您的 API 和工作負載元件設計為冪等。將冪等性檢查納入工作負載元件中。在您處理請求或執行操作之前，請先檢查唯一識別碼是否已處理。若已處理，則傳回先前的結果，而非再次執行操作。例如，若用戶端傳送建立使用者的請求，則檢查具有相同唯一識別碼的使用者是否已存在。如果使用者已存在，則應傳回現有的使用者資訊，而非建立新使用者。為此，用戶端可以使用冪等性字符發出 API 請求，每次重複請求時，都會使用這個權杖。

建立全方位的測試套件，以確認請求的冪等性。這些套件應涵蓋各種不同的案例，例如成功請求、失敗請求和重複請求。

如果您的工作負載利用 AWS Lambda 函數，請考慮 Powertools for AWS Lambda。Powertools for AWS Lambda 是一套開發人員工具組，當您搭配使用 AWS Lambda 函數時，可協助實作無伺服器最佳實務，並加快開發人員速度。它特別提供公用程式將 Lambda 函數轉換為可安全重試的冪等操作。

5. 清楚傳達冪等性

記錄您的 API 和工作負載元件，以清楚傳達操作的冪等性質。這樣有助於用戶端了解預期的行為，以及如何與您的工作負載進行可靠的互動。

6. 監控和稽核

實作監控和稽核機制，以偵測任何與回應冪等性相關的問題，例如非預期的回應變化，或過多的重複請求處理。這可協助您偵測和調查工作負載中的任何問題或意外行為。

資源

相關的最佳實務：

- [REL05-BP03 控制和限制重試呼叫](#)
- [REL06-BP01 監控工作負載的所有元件 \(產生\)](#)
- [REL06-BP03 傳送通知 \(即時處理和警示\)](#)
- [REL08-BP02 將功能測試整合為部署的一部分](#)

相關文件：

- [Amazon 建置者資料中心：透過冪等 API 安全地進行重試](#)
- [Amazon 建置者資料中心：分散式系統的挑戰](#)
- [Amazon 建置者資料中心：可靠性、持續工作以及咖啡時刻](#)
- [Amazon Elastic Container Service：確保冪等性](#)
- [如何讓 Lambda 函數冪等？](#)
- [確保 Amazon EC2 API 請求中的冪等性](#)

相關影片：

- [使用事件驅動型架構建置分散式應用程式 - AWS 線上技術講座](#)
- [AWS re:Invent 2023 - 使用事件驅動型架構建置新一代應用程式](#)
- [AWS re:Invent 2023 - 鬆耦合系統的進階整合模式和權衡](#)
- [AWS re:Invent 2023 - 使用 Amazon EventBridge 的進階事件驅動型模式](#)
- [AWS re:Invent 2018 - 閉門造車與開放思維：如何取得大小型系統的控制權 \(ARC337\) \(包括鬆耦合、持續工作、靜態穩定性\)](#)
- [AWS re:Invent 2019 - 移至事件驅動型架構 \(SVS308\)](#)

相關工具：

- [AWS Lambda Powertools 的冪等性 \(Java\)](#)
- [AWS Lambda Powertools 的冪等性 \(Python\)](#)
- [AWS Lambda Powertools GitHub 頁面](#)

REL 5. 如何設計分散式系統中的互動以減輕或承受故障？

分散式系統倚賴通訊網路來互連元件 (例如，伺服器或服務)。即使這些網路上的資料遺失或延遲，您的工作負載仍必須可靠運作。分散式系統的元件必須以不會對其他元件或工作負載造成負面影響的方式運作。這些最佳實務讓工作負載能夠承受壓力或故障，更快速地從其中復原，並減輕這類受損的影響。最終縮短平均復原時間 (MTTR)。

最佳實務

- [REL05-BP01 實作適度降級，以將適用的硬相依性轉換為軟相依性](#)
- [REL05-BP02 限流請求](#)
- [REL05-BP03 控制和限制重試呼叫](#)
- [REL05-BP04 快速檢錯和限制佇列](#)
- [REL05-BP05 設定用戶端逾時](#)
- [REL05-BP06 盡可能讓系統處於無狀態](#)
- [REL05-BP07 實作緊急槓桿](#)

REL05-BP01 實作適度降級，以將適用的硬相依性轉換為軟相依性

即使相依性變得不可用，應用程式元件仍應繼續執行其核心功能。它們有可能提供稍微陳舊的資料、備用資料，甚至未提供任何資料。這可確保整體系統運作在本地化失敗時只會受到最低限度的阻礙，同時提供核心商業價值。

預期成果：當元件的相依性狀況不良，元件本身仍可運作，但以降級的方式運作。元件的失敗模式應被視為正常運作。工作流程應適當設計，使此類失敗不會導致完全失敗，或至少會進入可預測和可復原的狀態。

常見的反模式：

- 未識別所需的業務功能。即使在相依性失敗期間，也不測試元件是否正常運作。
- 發生錯誤時，或只有多個相依性的其中之一無法使用，且仍可傳回部分結果時，就不提供資料。
- 當交易部分失敗時建立不一致的狀態。
- 沒有替代方法可存取中央參數存放區。
- 因重新整理失敗而使本機狀態失效或清空，而未考量這麼做的後果。

建立此最佳實務的優勢：按正常程序降級可改善系統整體的可用性，並且讓最重要的功能保持運作，即使在失敗期間亦然。

未建立此最佳實務時的曝險等級：高

實作指引

按正常程序實作降級，有助於將相依性失敗對元件功能的影響降到最低。理想情況下，元件會偵測相依性失敗，並以對其他元件或客戶造成最小影響的方式解決這些問題。

按正常程序降級的架構，意味著在相依性設計期間會考量潛在的失敗模式。對於每種失敗模式，都有一種方法可至少將元件最關鍵的功能提供給呼叫者或客戶。這些考量可能會成為可供測試和驗證的其他要求。理想情況下，即使有一或多個相依性失敗，元件仍然能夠以可接受的方式執行其核心功能。

這在商業上和技術上都同樣值得討論。所有業務要求都很重要，都應盡可能地滿足。然而，若無法滿足各項要求將會如何，仍是值得提出的問題。一個系統可以設計成可用且一致的，但在必須放棄一項要求的情況下，何者較重要？對於付款處理，可能應選擇一致性。對於即時應用程式，可能應選擇可用性。對於面向客戶的網站，答案可能取決於客戶的期望。

這意味著什麼，取決於元件的要求，以及應將哪些內容視為其核心功能。例如：

- 電子商務網站可能會顯示來自多個不同系統的資料，例如個人化推薦、排名最高的產品，以及客戶訂單在登陸網頁上的狀態。當一個上游系統失敗時，顯示其他所有內容，而不是向客戶顯示錯誤頁面，仍然是合理的。
- 如果個別作業之一失敗，執行批次寫入的元件仍然可以繼續處理批次。實作重試機制應該要很簡單。為此，您可以向呼叫者傳回關於哪些操作成功、哪些操作失敗及其為何失敗的資訊，或將失敗的請求放入無效字母佇列以實作非同步重試。失敗操作的相關資訊也應記錄下來。
- 處理交易的系統必須確認是否執行了所有更新，或完全未執行更新。對於分佈式交易，可使用 Saga 模式在相同交易的後續操作失敗的情況下回復先前的操作。在此，核心功能保有一致性。
- 具時間性的系統應該能夠處理未及時回應的相依性。在這類情況下，可以使用斷路器模式。若來自相依性的回應開始逾時，系統可以切換到不會進行其他呼叫的關閉狀態。
- 應用程式可從參數存放區讀取參數。使用一組預設的參數建立容器映像，並在參數存放區無法使用時使用這些參數，會很有效用。

請注意，在元件失敗的情況下採取的路徑需進行測試，且應遠比主要途徑簡單。一般來說，[應避免使用備用策略](#)。

實作步驟

識別外部和內部相依性。請考量其中可能會發生什麼樣的失敗。思考在這類失敗期間，將上游和下游系統以及客戶受到的負面影響降到最低的方法。

以下列出相依性，並說明如何在其失敗時按正常程序降級：

1. 相依性的部分失敗：一個元件可能會向下游系統提出多個請求，可以是對一個系統的多個請求，或者對多個系統各提出一個請求。視業務環境而定，對此可能會有不同的適當處理方式 (如需詳細資訊，請參閱實作指引中的先前範例)。
2. 下游系統因高負載而無法處理請求：如果對下游系統的請求一直失敗，則繼續重試是沒有意義的。這樣可能會對已過載的系統產生額外的負載，並使復原變得更加困難。此時可以使用斷路器模式，以監控對下游系統的失敗呼叫。若有大量呼叫失敗，將會停止向下游系統傳送更多請求，且偶爾才會讓呼叫通過，以測試下游系統是否已恢復可用性。
3. 參數存放區無法使用：若要轉換參數存放區，可以使用容器或機器映像中包含的軟相依性快取或有效的預設值。請注意，這些預設值需要保持最新狀態，並包含在測試套件中。
4. 監控服務或其他非功能性相依性無法使用：如果元件間歇性地無法傳送記錄、指標或追蹤給中央監控服務，最好還是照常執行業務功能。一般而言，長時間不日誌記錄或推送指標且未顯示任何訊息，是不可接受的。此外，某些使用案例可能需要完整的稽核項目才能滿足合規要求。
5. 關聯式資料庫的主要執行個體可能無法使用：Amazon Relational Database Service 與幾乎所有關聯式資料庫一樣，只能有一個主要寫入器執行個體。這會對寫入工作負載造成單一失敗點，並使擴展變得更加困難。透過使用多可用區域組態以獲得高可用性，或使用 Amazon Aurora Serverless 以獲得更好的擴展性，可以減輕部分問題。對於非常高的可用性要求，完全不依賴主要寫入器是有效用的。對於唯讀的查詢可以使用讀取複本，以提供備援和橫向擴展的能力，而不僅僅是縱向擴展。寫入可以緩衝處理 (例如，在 Amazon Simple Queue Service 佇列中)，如此，即使主要寫入器暫時無法使用，仍然可以接受客戶的寫入請求。

資源

相關文件：

- [Amazon API Gateway：調節 API 請求以獲得最佳的輸送量](#)
- [CircuitBreaker \(摘要說明「Release It!」書籍中的斷路器\)](#)
- [AWS 中的錯誤重試與指數退避](#)
- [Michael Nygard “Release It! Design and Deploy Production-Ready Software”](#)
- [Amazon 建置者資料中心：避免分散式系統的備用](#)
- [Amazon 建置者資料中心：避免無法逾越的佇列待辦項目](#)
- [Amazon 建置者資料中心：快取挑戰和策略](#)
- [Amazon 建置者資料中心：逾時、重試、退避與抖動](#)

相關影片：

- [重試、退避和抖動：AWS re:Invent 2019：Amazon 建置者資料中心簡介 \(DOP328\)](#)

REL05-BP02 限流請求

限流請求以減輕因需求非預期地增加而耗盡資源。系統會處理低於限流速率的請求，並拒絕超過定義限制的請求，且傳回一則訊息，指出請求已遭到限流。

預期成果：由於客戶流量突然增加、洪水攻擊或重試風暴而造成的大量尖峰，可透過請求限流來緩解，讓工作負載能夠繼續正常處理支援的請求量。

常見的反模式：

- API 端點限流未實作或保留為預設值，而未考量預期的數量。
- API 端點未經過負載測試，或未測試限流限制。
- 限流請求率，而不考量請求大小或複雜性。
- 測試請求率上限或請求大小上限，但不同時測試兩者。
- 資源不會佈建在測試時建立的相同限制。
- 未針對應用程式對應用程式 (A2A) API 取用者設定或考量使用計畫。
- 水平擴展的佇列取用者未進行最大並行設定。
- 未實作個別 IP 位址的速率限制。

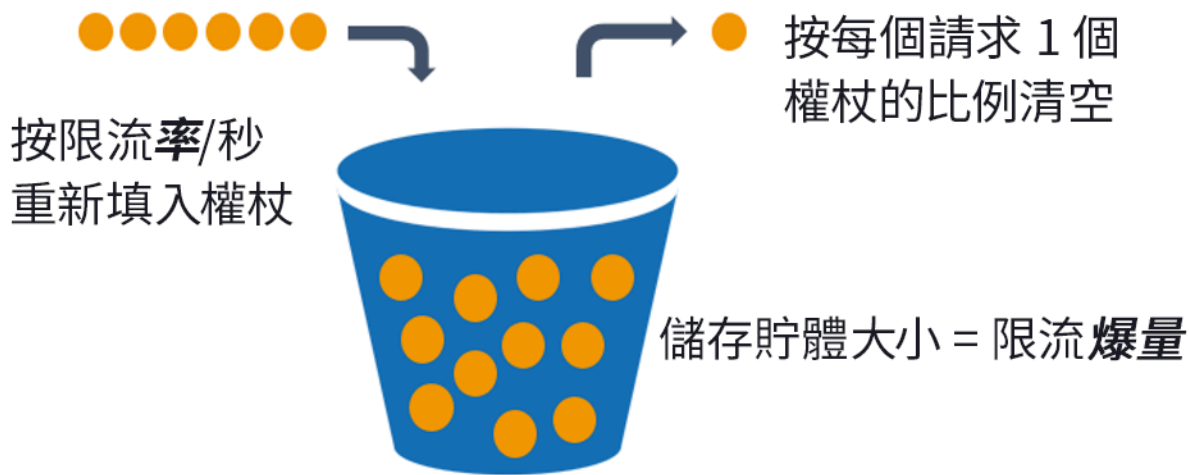
建立此最佳實務的優勢：設定限流限制的工作負載能夠在非預期的數量尖峰情況下正常運作，並成功處理已接受的請求負載。API 和佇列的請求若突然或持續激增將會受到限制，且不會耗盡請求處理資源。速率限制會限流個別請求者，使來自單一 IP 位址或 API 取用者的大量流量不會耗盡資源而影響到其他取用者。

未建立此最佳實務時的曝險等級：高

實作指引

服務應設計為處理已知的請求容量；此容量可透過負載測試來建立。如果請求到達率超過限制，會有適當的回應訊號指出請求已受到限流。這可讓取用者處理錯誤並於稍後重試。

當您的服務需要限流實作時，請考慮實作權杖儲存貯體演算法 (在此演算法中，權杖對於請求具重要性)。權杖會按每秒的限流率重新填入，並依照每個請求一個權杖的比例非同步清空。



權杖儲存貯體演算法。

[Amazon API Gateway](#) 會根據帳戶和區域限制實作權杖儲存貯體演算法，並且可根據使用計畫對個別用戶端進行設定。此外，[Amazon Simple Queue Service \(Amazon SQS\)](#) 和 [Amazon Kinesis](#) 可以緩衝請求以平滑請求速率，並為可解決的請求提供更高的限流。最後，可以使用 [AWS WAF](#) 實作速率限制，以針對會產生異常高負載的特定 API 取用者進行限流。

實作步驟

可以使用 API 的限流限制來設定 API Gateway，並在超出限制時傳回 429 Too Many Requests 錯誤。可以搭配使用 AWS WAF 與 AWS AppSync 和 API Gateway 端點，以按照 IP 位址啟用速率限制。此外，如果您的系統可容忍非同步處理，您可以將訊息放入佇列或串流中，以加快對服務用戶端的回應速度，進而提升到更高的限流率。

透過非同步處理，當您將 Amazon SQS 設定為 AWS Lambda 的事件來源時，可以[設定並行上限](#)，以避免高事件率消耗工作負載或帳戶中其他服務所需的可用帳戶並行執行配額。

雖然 API Gateway 提供了權杖儲存貯體的受管實作，在您無法使用 API Gateway 的情況下，您可以對服務使用權杖儲存貯體的語言特定開放原始碼實作 (請參閱「資源」中的相關範例)。

- 了解並設定每個區域帳戶層級的 [API Gateway 節流限制](#)、每個階段的 API 以及每個使用方案層級的 API 金鑰。
- 將 [AWS WAF 速率限制規則](#) 套用至 API Gateway 和 AWS AppSync 端點，以防止泛洪並封鎖惡意 IP。此外也可在 A2A 取用者的 AWS AppSync API 金鑰上設定速率限制規則。
- 考慮您是否需要比 AWS AppSync API 的速率限制更高的限流控制，如果需要，請在 AWS AppSync 端點前面設定 API Gateway。

- 將 Amazon SQS 佇列設定為 Lambda 佇列取用者的觸發器時，請將[並行上限](#)設定為足以符合服務層級目標的值，但不會消耗影響其他 Lambda 函數的並行限制。當您透過 Lambda 使用佇列時，請考慮在相同帳戶和區域中的其他 Lambda 函數上設定預留並行。
- 使用 API Gateway 與 Amazon SQS 或 Kinesis 的原生服務整合來緩衝請求。
- 如果您無法使用 API Gateway，請查看語言特定程式庫，為您的工作負載實作權杖儲存貯體演算法。查看範例區段並自行研究，以尋找合適的程式庫。
- 測試您預計要設定的限制，或您打算允許增加的限制，並記錄已測試的限制。
- 請勿將限制提高到您在測試時建立的範圍外。增加限制時，請先確認佈建的資源已等同於或大於測試情境中的資源，然後再套用增加。

資源

相關的最佳實務：

- [REL04-BP03 持續工作](#)
- [REL05-BP03 控制和限制重試呼叫](#)

相關文件：

- [Amazon API Gateway：調節 API 請求以獲得更佳的輸送量](#)
- [AWS WAF：以速率為基礎的規則陳述式](#)
- [導入使用 Amazon SQS 作為事件來源時的 AWS Lambda 並行上限](#)
- [AWS Lambda：並行上限](#)

相關範例：

- [三項最重要的 AWS WAF 以速率為基礎的規則](#)
- [Java Bucket4j](#)
- [Python 權杖儲存貯體](#)
- [節點權杖儲存貯體](#)
- [.NET 系統執行緒速率限制](#)

相關影片：

- [使用 AWS AppSync 實作 GraphQL API 安全最佳實務](#)

相關工具：

- [Amazon API Gateway](#)
- [AWS AppSync](#)
- [Amazon SQS](#)
- [Amazon Kinesis](#)
- [AWS WAF](#)
- [AWS 上的虛擬等候室](#)

REL05-BP03 控制和限制重試呼叫

使用指數退避，在每次重試之間的時間隔逐漸拉長後重試請求。在重試之間導入抖動以隨機化重試間隔。限制重試次數上限。

預期結果：分散式軟體系統中的典型元件包括伺服器、負載平衡器、資料庫和DNS伺服器。在正常操作期間，這些元件可以回應具有暫時性或有限錯誤的請求，以及無論是否重試都將持續存在的錯誤。當用戶端向服務發出請求時，請求會取用資源，包括記憶體、執行緒、連線、連接埠，或任何其他有限的資源。控制和限制重試是釋出資源並將資源耗用量降到最低的策略，可讓處於壓力下的系統元件不致不堪負荷。

當用戶端請求逾時或收到錯誤回應時，他們應該判斷是否要重試。如果執行重試，他們會使用具有抖動和最大重試值的指數退避來執行此作業。因此，後端服務和程序從負載中得到緩解並獲得自我修復的時間，進而更快速地復原和提供成功的請求服務。

常見的反模式：

- 在未新增指數退避、抖動和最大重試值的情況下實作重試。退避和抖動有助於避免因為在共用間隔內無意間進行協調重試而產生人為流量尖峰。
- 實作重試，而不測試其效果，或假設重試已內建到 中，SDK而不測試重試案例。
- 未能了解從相依性發布的錯誤代碼，因而重試了所有錯誤，包括有明確原因指出缺少權限的錯誤、組態錯誤，或其他預期必須要手動干預才能解決的狀況。
- 未解決可觀測性實務，包括對重複的服務失敗進行監控和提醒，使基礎問題廣為人知並且可以解決。
- 在內建或第三方重試功能堪用時，開發自訂重試機制。

- 以複合重試的方式在應用程式堆疊的多個層級重試，會嘗試在重試風暴中進一步耗用資源。請務必了解這些錯誤如何影響您所依賴的應用程式，然後僅在一個層級實作重試。
- 重試不是等冪的服務呼叫，導致非預期的副作用，例如重複的結果。

建立此最佳實務的優勢：重試可協助用戶端在請求失敗時獲得所需的結果，但也會耗用更多伺服器的時間來取得他們想要的成功回應。若失敗是罕見或暫時性的，重試可以有效運作。若失敗是由資源超載引起的，重試可能會使情況變得更糟。在用戶端重試中新增具有抖動的指數退避，可讓伺服器在資源超載導致失敗時進行復原。抖動可避免將請求對應到尖峰，而退避會減少將重試新增至正常請求負載所造成的負載上升。最後，請務必設定最大重試次數或經過時間，以避免建立會產生亞穩態失敗的積存。

未建立此最佳實務時的曝險等級：高

實作指引

控制和限制重試呼叫。使用指數退避以在逐漸延長間隔後重試。引進抖動來隨機化重試間隔，並限制重試次數上限。

有些 AWS SDKs 依預設會實作重試和指數退避。在工作負載中適用的情況下，使用這些內建 AWS 實作。在呼叫等冪的服務時，以及重試可改善用戶端可用性時，在您的工作負載中實作類似的邏輯。根據您的使用案例確定逾時時間以及何時停止重試。為那些重試使用案例建置和模擬演練測試情境。

實作步驟

- 確認應用程式堆疊中的最佳層級，以針對您應用程式所依賴的服務實作重試。
- 請注意，現有的 SDKs 會針對您選擇的語言實作經過驗證的重試策略，並具有指數退避和抖動，並且比編寫您自己的重試實作更有利。
- 在實作重試之前，請確認[服務是冪等的](#)。實作重試後，請務必在生產環境中加以測試和定期模擬演練。
- 呼叫 AWS 服務時 APIs，請使用 [AWS SDKs](#) 和 [AWS CLI](#) 並了解重試組態選項。確認預設值是否適用於您的使用案例，並視需要進行測試和調整。

資源

相關的最佳實務：

- [REL04-BP04 讓變異操作冪等](#)
- [REL05-BP02 限流請求](#)

- [REL05-BP04 快速檢錯和限制佇列](#)
- [REL05-BP05 設定用戶端逾時](#)
- [REL11-BP01 監控工作負載的所有元件以偵測故障](#)

相關文件：

- [中的錯誤重試和指數退避 AWS](#)
- [Amazon 建置者資料中心：逾時、重試、退避與抖動](#)
- [指數退避和抖動](#)
- [以無能方式確保重試安全 APIs](#)

相關範例：

- [Spring 重試](#)
- [Resilience4j 重試](#)

相關影片：

- [重試、退避和抖動：AWS re：Invent 2019：介紹 Amazon Builders 程式庫（DOP328）](#)

相關工具：

- [AWS SDKs 和 工具：重試行為](#)
- [AWS Command Line Interface：AWS CLI 重試](#)

REL05-BP04 快速檢錯和限制佇列

在服務無法成功回應請求時快速檢錯。如此將可釋出與請求關聯的資源，並且使服務可在資源用盡時復原。快速檢錯是一種完善的軟體設計模式，可用來在雲端中建置高度可靠的工作負載。佇列也是一種完善的企業整合模式，可以平滑負載，並且讓用戶端在可容忍非同步處理時釋出資源。如果服務在正常情況下能夠成功回應，但在請求速率太高時會失敗，請使用佇列來緩衝請求。不過，請勿允許建置長佇列積存，這可能導致用戶端已放棄的過時請求受到處理。

預期成果：當系統遇到資源爭用、逾時、例外狀況或灰色失敗而使服務水準目標無法達成時，快速檢錯的策略可加快系統復原速度。必須吸納流量尖峰且能支應非同步處理的系統，可讓用戶端使用佇列緩衝

處理後端服務的請求以快速釋出請求，藉此提升可靠性。緩衝處理要排入佇列的請求時，會實作佇列管理策略，以避免發生無法克服的積存。

常見的反模式：

- 在 DLQ 磁碟區上實作訊息佇列，但不設定無效字母佇列 (DLQ) 或警示，以偵測系統失敗。
- 未測量訊息在佇列中的存留期，這是一種延遲測量，用以了解佇列取用者何時進度落後或發生錯誤導致重試。
- 當處理這些訊息沒有任何價值，且業務需求已不存在時，未清除佇列中已積存的訊息。
- 在後進先出 (LIFO) 佇列可更妥善地滿足用戶端需求時設定了先進先出 (FIFO) 佇列，例如，當不需要嚴格排序，以及積存處理延遲了所有新的和時間敏感的請求，導致所有用戶端都經歷違反服務水準的狀況。
- 將內部佇列公開給用戶端，而不是公開管理工作接受以及將請求放入內部佇列的 API。
- 藉由將資源需求分攤到不同請求型態，將過多的工作請求類型合併到可能加劇積存條件的單一佇列中。
- 儘管需要不同的監控、逾時和資源分配，仍在同一佇列中處理複雜而簡單的請求。
- 不驗證輸入或使用判斷提示在軟體中實作快速檢錯的機制，以對可用正常程序處理錯誤的較高層級元件快顯例外狀況。
- 不會從請求路由中移除錯誤的資源，特別是在失敗處於灰色地帶，因損毀和重新啟動、間歇性相依性失敗、容量減少或網路封包遺失而導致成功與失敗並存。

建立此最佳實務的優勢：快速檢錯的系統更容易偵錯和修正，並且在發佈至生產環境之前，常會出現編碼和組態方面的問題。納入有效佇列策略的系統，可針對流量尖峰和間歇性系統失敗狀況提供更高的恢復能力和可靠性。

未建立此最佳實務時的曝險等級：高

實作指引

快速檢錯的策略可以編碼為軟體解決方案，並設定到基礎設施中。除了快速檢錯以外，佇列也是一種簡單而強大的架構技術，可將系統元件平滑負載分離。[Amazon CloudWatch](#) 提供監控失敗和發出相關警示的功能。已知系統失敗時，可以叫用緩解策略，包括背離受損的資源。當系統使用 [Amazon SQS](#) 和其他佇列技術來實作佇列以平滑負載時，必須考量如何管理佇列積存以及訊息取用失敗。

實作步驟

- 在軟體中實作程式化判斷提示或特定指標，並使用這些提示或指標來明確提醒系統問題。Amazon CloudWatch 可協助您根據應用程式日誌模式和 SDK 檢測來建立指標及提醒。

- 使用 CloudWatch 指標和警示背離受損的資源，這些資源會增加處理的延遲，或持續無法處理請求。
- 藉由設計 API 使用非同步處理來接受請求，並使用 Amazon SQS 將請求附加到內部佇列，然後透過成功訊息回應產生訊息的用戶端，讓用戶端可以釋出資源並繼續進行其他工作，同時讓後端佇列取用者處理請求。
- 藉由比較現在時間與訊息時間戳記，在每次從佇列中取出訊息時產生 CloudWatch 指標，來測量和監控佇列處理延遲。
- 因失敗而無法成功處理訊息，或無法在服務水準協議內處理磁碟區中的流量尖峰時，請將較舊或過多的流量排除至溢滿佇列。這可讓您優先處理新工作，並且等到有可用的容量時再處理較舊的工作。這種技術類似於 LIFO 處理，方便對所有新工作進行正常的系統處理。
- 使用無效字母或重新驅動佇列，將無法處理的訊息從積存移到可稍後再研究和解析的位置
- 進行重試，或在可接受的情況下，藉由比較目前時間與訊息時間戳記，捨棄與請求用戶端不再相關的訊息，將舊訊息捨棄。

資源

相關的最佳實務：

- [REL04-BP02 實作鬆散耦合相依性](#)
- [REL05-BP02 限流請求](#)
- [REL05-BP03 控制和限制重試呼叫](#)
- [REL06-BP02 定義和計算指標 \(彙總\)](#)
- [REL06-BP07 透過您的系統監控請求的端對端追蹤](#)

相關文件：

- [避免無法處理的佇列積存](#)
- [快速檢錯](#)
- [如何防止 Amazon SQS 佇列中不斷增加的積存訊息？](#)
- [Elastic Load Balancing：區域轉移](#)
- [Amazon 應用程式復原控制器：流量容錯移轉的路由控制](#)

相關範例：

- [企業整合模式：無效字母通道](#)

相關影片：

- [AWS re:Invent 2022 - 操作高可用性多可用區域應用程式](#)

相關工具：

- [Amazon SQS](#)
- [Amazon MQ](#)
- [AWS IoT Core](#)
- [Amazon CloudWatch](#)

REL05-BP05 設定用戶端逾時

在連線和請求上妥善設定逾時、有系統地對其進行驗證，並且不要依賴預設值，因為它們不知道工作負載具體細節。

預期成果：用戶端逾時應考量與等待需要花費異常時間才能完成的請求相關的用戶端、伺服器和工作負載成本。由於無法知道任何逾時的確切原因，用戶端必須使用服務知識來找出對可能原因和適當逾時的期望

用戶端連線根據設定的值逾時。經歷逾時後，用戶端決定退回並重試，或開啟[斷路器](#)。這些模式可避免發出可能使基礎錯誤情況惡化的請求。

常見的反模式：

- 不知道系統逾時或預設逾時。
- 不知道正常的請求完成時間。
- 不知道完成請求異常耗時的可能原因，或是與等待這些作業完成相關聯的用戶端、服務或工作負載效能成本。
- 不知道受損的網路只有在達到逾時後才會造成請求失敗的可能性，以及未採用較短逾時的用戶端和工作負載效能的成本。
- 不測試連線和請求的逾時情境。
- 將逾時設定得太高，這可能會導致較長的等待時間，並增加資源使用率。
- 將逾時設定得太低，導致人為失敗。
- 忽略模式以處理遠端呼叫 (例如斷路器和重試) 的逾時錯誤。

- 不考慮監控服務呼叫錯誤率、延遲的服務水準目標，以及延遲離群值。這些指標可提供對積極或寬鬆逾時的洞見

建立此最佳實務的優勢：遠端呼叫逾時已設定，且系統設計為按正常程序處理逾時，以便在遠端呼叫回應異常緩慢，而逾時錯誤由服務用戶端正常處理時，可以保留資源。

未建立此最佳實務時的曝險等級：高

實作指引

針對任何服務相依性呼叫和任何跨程序的呼叫，同時設定連線逾時和請求逾時。許多架構都提供內建的逾時功能，但請注意，對您的服務目標而言，有些架構具有無限或過高的預設值。太高的值會降低逾時的實用性，因為當用戶端等待逾時發生時，資源會持續耗用。太低值可能會增加後端流量和延遲，原因是重試的請求過多。在某些情況下，這可能導致完全停機，原因是正在重試所有請求。

決定逾時策略時，請考量下列事項：

- 由於請求的內容、目標服務受損或聯網分割失敗，處理請求的時間可能會比平常更長。
- 內容異常昂貴的請求可能會耗用不必要的伺服器 and 用戶端資源。在此情況下，讓這些請求逾時而不重試，可以保留資源。服務也應透過限流和伺服器端逾時，來保護自己免受異常昂貴的內容影響。
- 因服務受損而異常耗時的請求可能會逾時並重試。應考量請求和重試的服務成本，但如果原因是當地語系化的損害，則重試應該不會很昂貴，而且將可降低用戶端資源耗用量。逾時也可能會根據損害的性質釋出伺服器資源。
- 因網路傳遞請求或回應失敗而需要長時間才能完成的請求，可能會逾時並重試。由於請求或回應未傳遞，因此無論逾時長度為何，結果都是失敗。在此情況下，逾時不會釋出伺服器資源，但會釋出用戶端資源並改善工作負載效能。

利用如重試和斷路器等建立良好的設計模式，優雅地處理逾時，並支援快速失敗的方法。[AWS SDKs](#) 和 [AWS CLI](#) 允許設定連線和請求逾時，以及使用指數退避和抖動的重試。[AWS Lambda](#) 函數支援逾時組態，而使用 [AWS Step Functions](#)，您可以建置低程式碼斷路器，以利用與服務 AWS 和 預先建置的整合 SDKs。[AWS App Mesh](#) Envoy 提供逾時和斷路器功能。

實作步驟

- 設定遠端服務呼叫的逾時，並利用內建的語言逾時功能或開放原始碼逾時程式庫。
- 當您的工作負載使用 呼叫 時 AWS SDK，請檢閱文件，了解語言特定的逾時組態。
 - [Python](#)

- [PHP](#)
- [.NET](#)
- [Ruby](#)
- [Java](#)
- [Go](#)
- [Node.js](#)
- [C++](#)
- 在工作負載中使用或 AWS CLI 命令時 AWS SDKs，請設定 `connectTimeoutInMillis` 和 `tlsNegotiationTimeoutInMillis` 的 AWS [組態預設值](#)，以設定預設逾時值。
- 將 [命令列選項](#) `cli-connect-timeout` 和 `cli-read-timeout` 套用至 AWS 服務，以控制一次性 AWS CLI 命令。
- 監控遠端服務呼叫是否有逾時，並對持續性錯誤設定警示，以便您可以主動處理錯誤案例。
- 對呼叫錯誤率、延遲的服務層級目標和延遲異常值實作 [CloudWatch 指標](#) 和 [CloudWatch 異常偵測](#)，以深入了解如何管理過度激進或寬鬆的逾時。
- 設定 [Lambda 函數](#) 的逾時。
- API Gateway 用戶端在處理逾時時必須實作自己的重試。API Gateway 支援下游 [整合的 50 毫秒至 29 秒整合逾時](#)，在整合請求逾時時不會重試。
- 實作 [斷路器](#) 模式，以避免在逾時發生時進行遠端呼叫。開啟線路以避免呼叫失敗，並在呼叫正常回應時關閉線路。
- 對於基於容器的工作負載，請查看 [App Mesh Envoy](#) 功能以利用內建的逾時和斷路器。
- 使用 AWS Step Functions 建置用於遠端服務呼叫的低程式碼斷路器，特別是在呼叫 AWS 原生 SDKs 和支援的 Step Functions 整合時，以簡化工作負載。

資源

相關的最佳實務：

- [REL05-BP03 控制和限制重試呼叫](#)
- [REL05-BP04 快速檢錯和限制佇列](#)
- [REL06-BP07 透過您的系統監控請求的端對端追蹤](#)

相關文件：

- [AWS SDK：重試和逾時](#)
- [Amazon 建置者資料中心：逾時、重試、退避與抖動](#)
- [Amazon API Gateway 配額和重要備註](#)
- [AWS Command Line Interface：命令列選項](#)
- [AWS SDK for Java 2.x：設定API逾時](#)
- [AWS 使用組態物件和組態參考的 Botocore](#)
- [適用於 .NET 的 AWS SDK：重試與逾時](#)
- [AWS Lambda：設定 Lambda 函數選項](#)

相關範例：

- [搭配 AWS Step Functions 和 Amazon DynamoDB 使用斷路器模式](#)
- [Martin Fowler：CircuitBreaker](#)

相關工具：

- [AWS SDKs](#)
- [AWS Lambda](#)
- [Amazon SQS](#)
- [AWS Step Functions](#)
- [AWS Command Line Interface](#)

REL05-BP06 盡可能讓系統處於無狀態

系統不應要求狀態，或應該卸載狀態，以便在不同的用戶端請求之間，不依賴磁碟和記憶體中本機儲存的資料。這允許伺服器任意置換，而不會對可用性造成影響。

當使用者或服務與應用程式互動時，他們通常會執行形成工作階段的一系列互動。工作階段是使用者在使用應用程式時，在不同請求之間持續存在的唯一資料。無狀態應用程式是一種不需要了解先前互動，也不會儲存工作階段資訊的應用程式。

一旦設計為無狀態，您就可以使用無伺服器運算服務，例如 AWS Lambda 或 AWS Fargate。

除了伺服器替換之外，無狀態應用程式的另一個優點是他們可以水平擴展，因為任何可用的運算資源（例如 EC2 執行個體和 AWS Lambda 函數）都可以服務任何請求。

建立此最佳實務的優勢：設計為無狀態的系統更適合水平擴展，因此可以根據波動的流量和需求來新增或移除容量。其本質上也具有抵抗故障的能力，並在應用程式開發中提供靈活性和敏捷性。

未建立此最佳實務時的曝險等級：中

實作指引

讓您的應用程式無狀態。無狀態應用程式支援水平擴展，並且可以容忍單個節點的失敗。分析並了解在架構中維持狀態的應用程式元件。這可協助您評估轉換為無狀態設計的潛在影響。無狀態架構會分離使用者資料並卸載工作階段資料。這提供了獨立擴展每個元件的彈性，以滿足不同的工作負載需求，並最佳化資源使用率。

實作步驟

- 識別並了解應用程式中的有狀態元件。
- 透過將使用者資料與核心應用程式邏輯進行分離和管理來解耦資料。
 - [Amazon Cognito](#) 可以使用 [身分池](#)、[使用者集區](#) 和 [Amazon Cognito Sync](#) 等功能，將使用者資料與應用程式的程式碼分離。
 - 可以將密碼儲存在安全的集中位置，使用 [AWS Secrets Manager](#) 來分離使用者資料。這意味著應用程式的程式碼不需要存儲密碼，這使得它更安全。
 - 請考慮使用 [Amazon S3](#) 來存放大型非結構化資料，例如影像和文件。應用程式可以在需要時擷取此資料，而無需將其存儲在記憶體中。
 - 使用 [Amazon DynamoDB](#) 來存放使用者設定檔等資訊。應用程式可以近乎即時的速度查詢這些資料。
- 將工作階段資料卸載至資料庫、快取或外部檔案。
 - [Amazon ElastiCache](#)、Amazon DynamoDB、[Amazon Elastic File System](#)（Amazon EFS）和 [Amazon MemoryDB](#) 是可用來卸載工作階段資料 AWS 的服務範例。
- 在確定需要使用所選儲存解決方案維持哪些狀態和使用者資料之後，設計一個無狀態架構。

資源

相關的最佳實務：

- [REL11-BP03 在所有圖層上自動復原](#)

相關文件：

- [Amazon 建置者資料中心：避免分散式系統的備用](#)

- [Amazon 建置者資料中心：避免無法逾越的佇列待辦項目](#)
- [Amazon 建置者資料中心：快取挑戰和策略](#)
- [上的無狀態 Web 層最佳實務 AWS](#)

REL05-BP07 實作緊急槓桿

緊急控制桿是可緩解工作負載所受之可用性影響的快速程序。

緊急控制桿的運作方法是使用已知且經過測試的機制來停用、限流或變更元件或相依性的行為。這可以減輕因需求意外增加導致資源耗盡所造成的工作負載受損，並降低工作負載內非關鍵元件的故障影響。

預期成果：透過實作緊急控制桿，可以建立已知的良好流程，以維持工作負載中關鍵元件的可用性。在啟用緊急控制桿期間，工作負載應該會適度降級，並繼續執行其業務關鍵功能。如需優雅降級的詳細資訊，請參閱 [REL05-BP01 實作優雅降級，將適用的硬相依性轉換為軟相依性](#)。

常見的反模式：

- 非關鍵相依性失敗會影響核心工作負載的可用性。
- 未在非關鍵元件受損期間測試或驗證關鍵元件的行為。
- 沒有為啟用或停用緊急控制桿定義明確且決定性的準則。

建立此最佳實務的優勢：實作緊急控制桿可透過為解析程式提供已確立的程序來應對意外的需求峰值或非關鍵相依性失敗，從而提高工作負載中關鍵元件的可用性。

未建立此最佳實務時的曝險等級：中

實作指引

- 識別工作負載中的關鍵元件。
- 設計和建構工作負載中的關鍵元件，以承受非關鍵元件的故障。
- 進行測試以驗證非關鍵元件失敗期間您關鍵元件的行為。
- 定義和監控相關指標或觸發器以啟動緊急控制桿程序。
- 定義構成緊急控制桿的程序 (手動或自動)。

實作步驟

- 識別工作負載中的業務關鍵元件。

- 工作負載中的每個技術元件應對應到其相關業務功能，並將其排名為關鍵或非關鍵。如需 Amazon 關鍵和非關鍵功能的範例，請參閱 [Any Day Can Be Prime Day: How Amazon.com Search Uses Chaos Engineering to Handle Over 84K Requests Per Second](#)。
- 這同時是技術和業務方面的決策，並且會因組織和工作負載而異。
- 設計和建構工作負載中的關鍵元件，以承受非關鍵元件的故障。
- 在相依性分析期間，請考慮所有潛在的故障模式，並驗證您的緊急控制桿機制能為下游元件提供關鍵功能。
- 進行測試以驗證緊急控制桿啟動期間您關鍵元件的行為。
- 避免雙模式行為。如需更多詳細資訊，請參閱 [REL11-BP05 使用靜態穩定性來防止雙模式行為](#)。
- 定義、監控和警示相關指標，以啟動緊急控制桿程序。
- 尋找適合監控的指標取決於您的工作負載。一些範例指標是延遲或失敗的相依性請求次數。
- 定義構成緊急控制桿的程序 (手動或自動)。
- 這可能包括諸如[降載](#)、[限流請求](#)或實作[適度降級](#)等機制。

資源

相關的最佳實務：

- [REL05-BP01 實作優雅降級，將適用的硬相依性轉換為軟相依性](#)
- [REL05-BP02 節流請求](#)
- [REL11-BP05 使用靜態穩定性來防止雙模式行為](#)

相關文件：

- [自動化安全、無人為介入的部署](#)
- [任何一天都可能是黃金日：Amazon.com 搜尋功能如何使用混沌工程每秒處理 84K 以上的請求](#)

相關影片：

- [AWS re：Invent 2020：透過不可變性的可靠性、一致性和信心](#)

變更管理

問題

- [REL 6. 如何監控工作負載資源？](#)
- [REL 7. 如何設計工作負載以因應需求的變化？](#)
- [REL 8. 如何實作變更？](#)

REL 6. 如何監控工作負載資源？

日誌和指標是可深入洞察工作負載運作狀態的強大工具。您可以設定工作負載以監控日誌和指標，並在超過閾值或發生重大事件時傳送通知。監控可讓您的工作負載識別何時會超過低效能閾值或發生故障，以便自動復原來回應。

最佳實務

- [REL06-BP01 監控工作負載的所有元件 \(產生\)](#)
- [REL06-BP02 定義和計算指標 \(彙總\)](#)
- [REL06-BP03 傳送通知 \(即時處理和警示\)](#)
- [REL06-BP04 自動化回應 \(即時處理和警示\)](#)
- [REL06-BP05 分析日誌](#)
- [REL06-BP06 定期審查監控範圍和指標](#)
- [REL06-BP07 透過您的系統監控請求的端對端追蹤](#)

REL06-BP01 監控工作負載的所有元件 (產生)

使用 Amazon CloudWatch 或第三方工具監控工作負載的元件。使用 AWS Health 儀表板監控 AWS 服務。

工作負載的所有元件都應該受到監控，包括前端、商業邏輯和儲存層。定義關鍵指標，描述如何從日誌中擷取指標 (如果需要)，並設定調用對應警示事件的閾值。確保指標與工作負載的關鍵績效指標 (KPI) 相關，並使用指標和日誌來識別服務降級的早期警告訊號。例如，與業務成果相關的指標 (例如每分鐘成功處理的訂單數量) 可以比技術指標 (例如 CPU 使用率) 更快地指出工作負載問題。使用 AWS Health 儀表板可針對 AWS 資源下的 AWS 服務的效能和可用性，為您提供個人化檢視。

雲端監控提供新機遇。大多數雲端供應商都開發了可自訂的勾點，並且可以提供深入解析，協助您監控工作負載的多個層面。Amazon CloudWatch 等 AWS 服務會在使用者介入程度最低的情況下，套用統計和機器學習演算法來持續分析系統和應用程式的指標、判斷正常基準以及披露異常情況。異常偵測演算法會考慮指標的季節性和趨勢變化。

AWS 提供大量可供使用的監控和日誌資訊，可用於定義工作負載特定的指標、需求變更程序，並採用機器學習技術，而不論機器學習專業知識如何。

此外，監控所有外部端點，以確保它們獨立於基本實作。此主動監控可透過綜合交易 (有時稱為使用者 Canary，但請別與 Canary 部署混淆) 加以完成，它會定期運行工作負載的用戶端執行的許多常見任務匹配動作。在持續時間中讓這些任務保持簡單扼要，並確定在測試期間不會讓工作負載超載。Amazon CloudWatch Synthetics 讓您能夠[建立綜合性 Canary](#)，以監控您的端點和 API。您也可以將綜合性 Canary 用戶端節點與 AWS X-Ray 主控台結合，以指出綜合性 Canary 在所選時段內發生錯誤、故障或調節率等問題。

預期成果：

從工作負載的所有元件中收集並使用關鍵指標，以確保工作負載的可靠性和最佳的使用者體驗。偵測工作負載未達成業務成果，可讓您快速宣告災難並從事件中復原。

常見的反模式：

- 僅監控工作負載的外部界面。
- 不產生任何工作負載特定指標，只依賴工作負載使用的 AWS 服務提供給您的指標。
- 僅在工作負載中使用技術指標，而不監控與工作負載貢獻的非技術 KPI 相關的任何指標。
- 依賴生產流量和簡單的運作狀態檢查來監控和評估工作負載狀態。

建立此最佳實務的優勢：在工作負載中的所有層級進行監控，可讓您更快速地預測和解決構成工作負載的元件中的問題。

未建立此最佳實務時的曝險等級：高

實作指引

1. 在可用的地方開啟日誌記錄。應從工作負載的所有元件中取得監控資料。開啟其他日誌記錄，例如 S3 Access Logs，並允許您的工作負載記錄工作負載特定資料。從 Amazon ECS、Amazon EKS、Amazon EC2、Elastic Load Balancing、AWS Auto Scaling 和 Amazon EMR 等服務中收集 CPU、網路 I/O 和磁碟 I/O 平均值的指標。如需將指標發佈到 CloudWatch 的 AWS 服務清單，請參閱 [AWS Services That Publish CloudWatch Metrics](#)。
2. 審核所有預設指標，並探索任何資料收集漏洞。每個服務都會產生預設指標。收集預設指標可讓您進一步了解工作負載元件之間的相依性，以及元件可靠性和效能如何影響工作負載。您也可以使用 AWS CLI 或 API 建立自己的自訂指標並[發佈](#)到 CloudWatch。
3. 評估所有指標，以決定在工作負載中為每個 AWS 服務提醒哪些指標。您可以進行選擇，以選取對工作負載可靠性有重大影響的指標子集。專注於重要指標和閾值，可讓您調整[提醒](#)的數量，並協助將誤報率降至最低。

4. 在調用提醒後，定義工作負載的提醒和復原程序。定義提醒可讓您快速通知、呈報並遵循必要的步驟，以便從事件中復原並實現規定的復原時間點目標 (RTO)。可以使用 [Amazon CloudWatch 警示](#) 來調用自動化工作流程，並根據定義的閾值啟動復原程序。
5. 探索如何使用綜合交易來收集有關工作負載狀態的相關資料。綜合監控遵循相同的路由並執行與客戶相同的動作，即使您的工作負載沒有任何客戶流量，也能持續驗證您的客戶體驗。透過使用 [綜合交易](#)，您可以在客戶之前發現問題。

資源

相關的最佳實務：

- [REL11-BP03 將所有分層的修復自動化](#)

相關文件：

- [開始使用 AWS Health 儀表板 — 您的帳戶運作狀態](#)
- [發佈 CloudWatch 指標的 AWS 服務](#)
- [Network Load Balancer 的存取日誌](#)
- [Application Load Balancer 的存取日誌](#)
- [存取 AWS Lambda 的 Amazon CloudWatch Logs](#)
- [Amazon S3 伺服器存取日誌記錄](#)
- [啟用 Classic Load Balancer 的存取日誌](#)
- [將日誌資料匯出至 Amazon S3](#)
- [在 Amazon EC2 執行個體上安裝 CloudWatch Agent](#)
- [發佈自訂指標](#)
- [使用 Amazon CloudWatch 儀表板](#)
- [使用 Amazon CloudWatch 指標](#)
- [使用 Canary \(Amazon CloudWatch Synthetics\)](#)
- [什麼是 Amazon CloudWatch Logs？](#)

使用者指南：

- [建立追蹤](#)
- [監控 Amazon EC2 Linux 執行個體的記憶體和磁碟指標](#)

- [CloudWatch Logs 與容器執行個體搭配使用](#)
- [VPC 流量日誌](#)
- [什麼是 Amazon DevOps Guru ?](#)
- [什麼是 AWS X-Ray ?](#)

相關部落格：

- [使用 Amazon CloudWatch Synthetics 和 AWS X-Ray 進行偵錯](#)

相關範例：

- [Amazon 建置者資料中心：偵測分散式系統，以了解運作狀態](#)
- [可觀測性研討會](#)

REL06-BP02 定義和計算指標 (彙總)

從工作負載元件收集指標和日誌，並從中計算相關的彙總指標。這些指標可提供廣泛且深入的工作負載可觀測性，並可大幅改善您的彈性狀態。

可觀測性不只是從工作負載元件收集指標以及能夠檢視指標並設定提醒。重點在於全面了解工作負載的行為。此行為資訊來自工作負載中的所有元件，包括其所依賴的雲端服務、精心製作的日誌及指標。此資料可讓您全面監督工作負載的行為，並了解每個元件與每個工作單元的互動及其細節。

預期成果：

- 您可以從工作負載元件和 AWS 服務相依性收集日誌，並將其發布到方便存取和處理的集中位置。
- 您的日誌包含高保真度且準確的時間戳記。
- 您的日誌包含處理內容的相關資訊，例如追蹤識別碼、使用者或帳戶識別碼，以及遠端 IP 位址。
- 您可以從日誌建立彙總指標，以便從高層級的角度來呈現工作負載的行為。
- 您可以查詢彙總日誌，以取得有關工作負載的深入分析，並識別實際和潛在的問題。

常見的反模式：

- 您未從工作負載執行所在的運算執行個體或其使用的雲端服務收集相關的日誌或指標。
- 您忽略了收集與業務關鍵績效指標 (KPI) 相關的日誌和指標。

- 您單獨分析工作負載相關的遙測，而未彙總和建立相互關聯。
- 您太快讓指標和日誌過期，導致阻礙了趨勢分析和週期性問題識別。

建立這些最佳實務的優勢：您可以偵測更多異常情況，並將工作負載的不同元件之間的事件與指標相互關聯。您可以根據日誌中通常無法單獨用於指標的資訊，從工作負載元件建立深入分析。您可以大規模查詢日誌來加速判斷失敗原因。

未建立這些最佳實務時的曝險等級：高

實作指引

識別與您的工作負載及其元件相關的遙測資料來源。此資料不僅來自發布指標的元件 (例如作業系統 (OS) 和 Java 等應用程式執行時期)，也來自應用程式和雲端服務日誌。例如，Web 伺服器通常會記錄每個請求的詳細資訊，例如時間戳記、處理延遲、使用者 ID、遠端 IP 位址、路徑和查詢字串。這些日誌中的詳細資訊層級可協助您執行詳細的查詢，並產生其他方式無法提供的指標。

使用適當的工具和程序收集指標和日誌。Amazon EC2 執行個體上執行的應用程式所產生的日誌，可透過 [Amazon CloudWatch 代理程式](#) 等代理程式收集並發布至 [Amazon CloudWatch Logs](#) 等集中儲存服務。AWS 受管運算服務 (例如 [AWS Lambda](#) 和 [Amazon Elastic Container Service](#)) 會自動將日誌發布至 CloudWatch Logs。為 [Amazon CloudFront](#)、[Amazon S3](#)、[Elastic Load Balancing](#) 和 [Amazon API Gateway](#) 等工作負載所使用的 AWS 儲存和處理服務啟用日誌收集。

利用[維度](#)讓您的遙測資料更豐富，如此您就能更清楚地看見行為模式，並將相互關聯的問題隔離成相關元件的群組。新增後，您可以觀測更詳細的元件行為、偵測相互關聯的故障，並採取適當的補救措施。有用的維度範例包括可用區域、EC2 執行個體 ID，以及容器任務或 Pod ID。

收集指標和日誌後，您可以撰寫查詢並從中產生彙總指標，以針對正常和異常行為提供實用的深入分析。例如，您可以使用 [Amazon CloudWatch Logs Insights](#) 從您的應用程式日誌產生自訂指標、使用 [Amazon CloudWatch Metrics Insights](#) 大規模查詢您的指標、使用 [Amazon CloudWatch Container Insights](#) 從容器化應用程式和微服務收集、彙總及摘要整理指標與日誌，或者，如果您使用 AWS Lambda 函數，則可以使用 [Amazon CloudWatch Lambda 洞察](#)。若要建立彙總錯誤率指標，您可以在每次元件日誌中出現錯誤回應或訊息時讓計數器累進，或計算現有錯誤率指標的彙總值。您可以使用此資料來產生顯示結尾行為的長條圖，例如效能最差的請求或程序。您也可以使用 CloudWatch Logs [異常偵測](#) 等解決方案來即時掃描此資料中是否有異常模式。這些深入分析可以放置在儀表板上，並根據您的需求和偏好組織整理。

查詢日誌可協助您了解工作負載元件如何處理特定請求，並顯示請求模式或其他影響工作負載彈性的內容。根據您對應用程式和其他元件行為的了解，事先研究和準備查詢會很實用，讓您更方便視需要執行該程式或元件。舉例來說，若使用 [CloudWatch Logs Insights](#)，您能以互動方式搜尋和分析

CloudWatch Logs 中存放的日誌資料。您也可以使用 [Amazon Athena](#) 查詢多個來源的日誌，包括 PB 規模的許多 AWS 服務。

當您定義日誌保留政策時，請考慮歷史日誌的價值。歷史日誌有助於識別工作負載效能的長期使用情形及行為模式、迴歸與改善。日誌永久刪除後，即無法再供分析。然而，歷史日誌的價值往往會隨著時間而降低。選擇一項政策來適當平衡您的需求，並應付您可能須遵循的任何法律或合約要求。

實作步驟

1. 為您的可觀測性資料選擇收集、儲存、分析和顯示機制。
2. 在工作負載的適當元件上安裝並設定指標和日誌收集器 (例如，在 Amazon EC2 執行個體上和[附屬容器](#)中)。設定這些收集器，讓它們在意外停止時自動重新啟動。啟用收集器的磁碟或記憶體緩衝，不要讓臨時發布失敗影響您的應用程式或導致資料遺失。
3. 在做為工作負載的一部分使用的 AWS 服務上啟用記錄功能，並將這些日誌轉送到您選取的儲存服務 (如有需要)。如需詳細說明，請參閱個別服務的使用者或開發人員指南。
4. 根據您的遙測資料，定義與工作負載相關的營運指標。這些指標可以根據工作負載元件發出的直接指標，包括業務 KPI 相關指標，或彙總計算的結果，例如總和、速率、百分位數或長條圖。使用日誌分析器計算這些指標，並將其放在儀表板上適當的位置。
5. 視需要準備適當的日誌查詢，以分析工作負載元件、請求或交易行為。
6. 為您的元件日誌定義和啟用日誌保留政策。當日誌比政策允許的時間更早時，請定期刪除日誌。

資源

相關的最佳實務：

- [REL06-BP01 監控工作負載的所有元件 \(產生\)](#)
- [REL06-BP03 傳送通知 \(即時處理和警示\)](#)
- [REL06-BP04 自動化回應 \(即時處理和警示\)](#)
- [REL06-BP05 分析日誌](#)
- [REL06-BP06 定期審查監控範圍和指標](#)
- [REL06-BP07 透過您的系統監控請求的端對端追蹤](#)

相關文件：

- [Amazon CloudWatch 的運作方式](#)
- [Amazon Managed Prometheus](#)

- [Amazon Managed Grafana](#)
- [使用 CloudWatch Logs Insights 分析日誌資料](#)
- [Amazon CloudWatch Lambda 洞察](#)
- [Amazon CloudWatch Container Insights](#)
- [使用 CloudWatch Metrics Insights 查詢您的指標](#)
- [AWS Distro for OpenTelemetry](#)
- [Amazon CloudWatch Logs Insights 範例查詢](#)
- [使用 Amazon CloudWatch Synthetics 和 AWS X-Ray 進行偵錯](#)
- [搜尋和篩選日誌資料](#)
- [直接將日誌傳送至 Amazon S3](#)
- [Amazon 建置者資料中心：偵測分散式系統，以了解運作狀態](#)

相關研討會：

- [一個可觀測性研討會](#)

相關工具：

- [AWS Distro for OpenTelemetry \(GitHub\)](#)

REL06-BP03 傳送通知 (即時處理和警示)

當組織偵測到潛在問題時，他們會將即時通知和警示傳送給適當的人員和系統，以便快速有效地應對這些問題。

預期成果：根據服務和應用程式指標設定相關警示，就可以快速回應操作事件。違反警示閾值時，系統會通知適當的人員和系統，以便解決潛在問題。

常見的反模式：

- 將警示的閾值設得過高，會導致無法傳送重要通知。
- 將警示的閾值設得太低，導致使用者因通知過多的干擾而無法針對重要提醒採取行動。
- 當使用情況改變時，未更新警示及其閾值。
- 針對透過自動化動作解決的最佳警示，將通知傳送給人員而未引發自動化動作，會導致傳送過多的通知。

建立此最佳實務的優勢：將即時通知和警示傳送給適當的人員和系統，以便及早發現問題並快速回應操作事故。

未建立此最佳實務時的曝險等級：高

實作指引

工作負載應具備即時處理和警示功能，以改善可能影響應用程式可用性問題的可偵測性，並作為自動化回應的觸發程式。組織可以透過使用已定義的指標建立警示來執行即時處理和警示，以便在發生重大事件或指標超過閾值時收到通知。

[Amazon CloudWatch](#) 可讓您根據靜態閾值、異常偵測和其他條件，使用 CloudWatch 警示來建立[指標](#)和複合警示。如需有關可使用 CloudWatch 來設定的警示類型的詳細資訊，請參閱 [CloudWatch 文件的警示一節](#)。

您可以使用 [CloudWatch 儀表板](#) 為團隊建構 AWS 資源的指標和警示的自訂檢視。CloudWatch 主控台的可自訂首頁可讓您在單一檢視中監控多個區域的資源。

警示可執行一個或多個動作，例如向 [Amazon SNS 主題](#) 傳送通知、執行 [Amazon EC2](#) 動作或 [Amazon EC2 Auto Scaling](#) 動作，或在 AWS Systems Manager 中建立 [OpsItem](#) 或[事故](#)。

Amazon CloudWatch 使用 [Amazon SNS](#) 於警示變更狀態時傳送通知，將訊息傳遞從發布者 (生產者) 提供給訂閱用戶 (消費者)。如需有關設定 Amazon SNS 通知的詳細資訊，請參閱 [Configuring Amazon SNS](#)。

每當建立、更新、刪除 CloudWatch 警示或變更其狀態時，CloudWatch 就會傳送 [EventBridge 事件](#)。您可以使用 EventBridge 搭配這些事件來建立執行動作的規則，例如，當警示狀態變更時就通知您，或使用 [Systems Manager Automation](#) 自動觸發您帳戶中的事件。

利用 [AWS Health](#) 隨時掌握新知。AWS Health 是 AWS 雲端 資源運作狀態的權威資訊來源。使用 AWS Health 來接收任何已確認服務事件的通知，方便您快速採取行動來緩解任何影響。透過 [AWS 使用者通知](#) 建立符合用途的 AWS Health 事件通知，以利用電子郵件和聊天管道傳送，並透過 [Amazon EventBridge](#) 以程式設計方式與您的[監控和警示工具](#)整合。如果使用 AWS Organizations，則跨帳戶彙總 AWS Health 事件。

應何時使用 EventBridge 或 Amazon SNS？

EventBridge 和 Amazon SNS 可用於開發事件驅動的應用程式，將視具體需求來做出選擇。

當您想要建立應用程式以回應您自己的應用程式、SaaS 應用程式和 AWS 服務中的事件時，建議使用 Amazon EventBridge。EventBridge 是唯一可直接與第三方 SaaS 合作夥伴整合的事件型服

務。EventBridge 也會自動擷取來自 200 多個 AWS 服務的事件，而不需要開發人員在其帳戶中建立任何資源。

EventBridge 會將已定義的 JSON 架構用於事件，並協助您建立在整個事件內文中套用的規則，以選取要轉寄至目標的事件。EventBridge 目前支援超過 20 種 AWS 服務作為目標，包括 [AWS Lambda](#)、[Amazon SQS](#)、Amazon SNS、[Amazon Kinesis Data Streams](#) 和 [Amazon Data Firehose](#)。

針對需要高散發的應用程式 (數千或數百萬個端點)，建議使用 Amazon SNS。我們看到的常見模式是客戶使用 Amazon SNS 做為規則的目標，以篩選所需的事件並散發到多個端點。

訊息是非結構化的，可以是任何格式。Amazon SNS 支援將訊息轉寄到六種不同類型的目標，包括 Lambda、Amazon SQS、HTTP/S 端點、SMS、行動推送和電子郵件。Amazon SNS [一般的延遲時間短於 30 毫秒](#)。透過將服務設定為傳送 Amazon SNS 訊息，各種 AWS 服務就能做到這一點 (超過 30 個，包括 Amazon EC2、[Amazon S3](#) 和 [Amazon RDS](#))。

實作步驟

1. 使用 [Amazon CloudWatch 警示](#) 建立警示。
 - a. 指標警示會監控單一 CloudWatch 指標或與 CloudWatch 指標相依的表達式。與超過一段時間間隔的閾值相比，警示會根據指標或表達式的值起始一或多個動作。該動作可能包含將通知傳送至 [Amazon SNS 主題](#)、執行 [Amazon EC2](#) 動作或 [Amazon EC2 Auto Scaling](#) 動作，或在 AWS Systems Manager 中 [建立 OpsItem](#) 或 [事故](#)。
 - b. 複合警示由規則表達式組成，該規則表達式會將您已建立的其他警示條件納入考量。只有在符合所有規則條件時，複合警示才會進入警示狀態。在複合警示規則表達式中指定的警示可能會包括指標警示和其他複合警示。複合警示可以在變更狀態時傳送 Amazon SNS 通知，並且可以在進入警示狀態時建立 Systems Manager [OpsItems](#) 或 [事故](#)，但無法執行 Amazon EC2 動作或 Auto Scaling 動作。
2. 設定 [Amazon SNS 通知](#)。建立 CloudWatch 警示時，可以包含 Amazon SNS 主題，以便在警示變更狀態時傳送通知。
3. [在 EventBridge 中建立](#) 符合指定 CloudWatch 警示的規則。每個規則都支援多個目標，包括 Lambda 函數。例如，您可以定義在可用磁碟空間不足時啟動的警示，該警示會透過 EventBridge 規則觸發 Lambda 函數以清理空間。如需有關 EventBridge 目標的詳細資訊，請參閱 [EventBridge targets](#)。

資源

相關 Well-Architected 的最佳實務：

- [REL06-BP01 監控工作負載的所有元件 \(產生\)](#)
- [REL06-BP02 定義和計算指標 \(彙總\)](#)
- [REL12-BP01 使用程序手冊調查失敗](#)

相關文件：

- [Amazon CloudWatch](#)
- [CloudWatch Logs 洞察](#)
- [使用 Amazon CloudWatch 警示](#)
- [使用 Amazon CloudWatch 儀表板](#)
- [使用 Amazon CloudWatch 指標](#)
- [設定 Amazon SNS 通知](#)
- [CloudWatch 異常偵測](#)
- [CloudWatch Logs 資料保護](#)
- [Amazon EventBridge](#)
- [Amazon Simple Notification Service](#)

相關影片：

- [重塑 2022 年可觀測性影片](#)
- [AWS re:Invent 2022 - Amazon 的可觀測性最佳實務](#)

相關範例：

- [一個可觀測性研討會](#)
- [Amazon EventBridge 到 AWS Lambda，由 Amazon CloudWatch 警示進行回饋控制](#)

REL06-BP04 自動化回應 (即時處理和警示)

偵測到事件時，使用自動化以採取動作，例如取代故障的元件。

實作警示的自動即時處理，以便系統可以採取快速的糾正措施，並嘗試在觸發警示時防止故障或服務降級。對警示的自動回應可能包括替換故障元件、調整運算容量、將流量重新導向到運作狀態良好的主機、可用區域或其他區域，以及操作人員通知。

預期成果：識別即時警示，並設定警示的自動處理，以調用採取的適當動作，維護服務水準目標和服務水準協議 (SLA)。自動化的範圍從單一元件的自我修復活動到全站點的容錯移轉。

常見的反模式：

- 針對關鍵的即時警示沒有清晰的清單或目錄。
- 對關鍵警示沒有自動回應 (例如，當運算資源即將耗盡時，發生自動擴展)。
- 矛盾的警示回應動作。
- 操作人員在收到提醒通知時沒有可以遵循的標準作業程序 (SOP)。
- 未監控組態變更，因為未偵測到的組態變更可能會導致工作負載停機。
- 沒有復原意外組態變更的策略。

建立此最佳實務的優勢：自動化警示處理可改善系統復原能力。系統會自動採取糾正措施，減少人為介入時容易出錯的手動活動。工作負載運作符合可用性目標，並減少服務中斷。

未建立此最佳實務時的曝險等級：中

實作指引

為了有效管理提醒並自動化其回應，請根據提醒的重要性的影響來進行分類，記錄回應程序，並在為任務排名前規劃好回應。

識別需要特定動作的任務 (通常會在執行手冊中詳細說明)，並檢查所有執行手冊和程序手冊以確定哪些任務可以自動化。可以定義的動作通常也可以自動化。如果動作無法自動化，請在 SOP 中記錄手動步驟，並對操作人員進行相關培訓。持續挑戰手動程序以尋求自動化機會，以便您可以建立和維護用來自動化提醒回應的計畫。

實作步驟

1. 建立警示清單：若要取得所有警示的清單，可以透過 [Amazon CloudWatch](#) 命令 [describe-alarms](#) 來使用 [AWS CLI](#)。視設定的警示數量而定，您可能必須使用分頁來擷取每個呼叫的警示子集，或者也可以使用 AWS SDK [透過 API 呼叫](#) 來取得警示。
2. 記錄所有警報動作：更新執行手冊與所有警示及其動作，無論其為手動還是自動。[AWS Systems Manager](#) 可提供預先定義的執行手冊。如需有關執行手冊的詳細資訊，請參閱 [Working with runbooks](#)。如需有關如何檢視執行手冊內容的詳細資訊，請參閱[檢視執行手冊內容](#)。
3. 設定和管理警示動作：對於任何需要動作的警示，請[使用 CloudWatch SDK 指定自動化動作](#)。例如，您可以建立和啟用警示的動作，或停用警示的動作，以根據 CloudWatch 警示自動變更 Amazon EC2 執行個體的狀態。

也可使用 [Amazon EventBridge](#) 來自動回應系統事件，例如應用程式可用性問題或資源變動。您可建立規則來指示您在意的事件，以及當事件符合規則時執行的動作。可以自動啟動的動作包括調用 [AWS Lambda](#) 函數、調用 [Amazon EC2 Run Command](#)、將事件轉送至 [Amazon Kinesis Data Streams](#) 以及查看[使用 EventBridge 自動化 Amazon EC2](#)。

4. 標準操作程序 (SOP)：根據您的應用程式元件，[AWS Resilience Hub](#) 會建議多個 [SOP 範本](#)。您可以使用這些 SOP 記錄在發出提醒時操作員應遵循的所有程序。也可以根據 Resilience Hub 建議來[建構 SOP](#)，其中您需要具有相關彈性政策的 Resilience Hub 應用程式，以及針對該應用程式的歷史彈性評估。SOP 的建議會由彈性評估產生。

Resilience Hub 與 Systems Manager 合作，透過提供許多 [SSM 文件](#) 作為這些 SOP 的基礎來自動執行 SOP 的步驟。例如，Resilience Hub 可能會建議使用 SOP 來根據現有的 SSM 自動化文件新增磁碟空間。

5. 使用 Amazon DevOps Guru 執行自動化動作：可使用 [Amazon DevOps Guru](#) 來自動監控應用程式資源，以偵測異常行為並提供目標建議，以縮短問題識別和矯正時間。使用 DevOps Guru，可以近乎即時地監控多個來源的操作資料串流，包括 Amazon CloudWatch 指標、[AWS Config](#)、[AWS CloudFormation](#) 和 [AWS X-Ray](#)。也可以使用 DevOps Guru 在 OpsCenter 中自動建立 [OpsItems](#)，並將事件傳送至 [EventBridge](#) 以實現額外的自動化。

資源

相關的最佳實務：

- [REL06-BP01 監控工作負載的所有元件 \(產生\)](#)
- [REL06-BP02 定義和計算指標 \(彙總\)](#)
- [REL06-BP03 傳送通知 \(即時處理和警示\)](#)
- [REL08-BP01 將執行手冊用於部署等標準活動](#)

相關文件：

- [AWS Systems Manager Automation](#)
- [建立 EventBridge 規則，以透過 AWS 資源觸發事件](#)
- [一個可觀測性研討會](#)
- [Amazon 建置者資料中心：偵測分散式系統，以了解運作狀態](#)
- [什麼是 Amazon DevOps Guru？](#)
- [使用自動化文件 \(手冊\)](#)

相關影片：

- [AWS re:Invent 2022 - Amazon 的可觀測性最佳實務](#)
- [AWS re:Invent 2020：使用 AWS Systems Manager 自動執行任何操作](#)
- [AWS Resilience Hub 簡介](#)
- [為 Amazon DevOps Guru 通知建立自訂票證系統](#)
- [使用 Amazon DevOps Guru 啟用多帳戶洞察彙總](#)

相關範例：

- [Amazon CloudWatch 和 Systems Manager 研討會](#)

REL06-BP05 分析日誌

收集日誌檔和指標歷史記錄，並分析這些檔案和歷史記錄，以了解更廣泛的趨勢和工作負載洞見。

Amazon CloudWatch Logs Insights 支援[簡單但功能強大的查詢語言](#)，您可使用此語言來分析日誌資料。Amazon CloudWatch Logs 還支援訂閱，而這些訂閱允許資料無縫流至 Amazon S3，您可使用 Amazon S3 或 Amazon Athena 來查詢資料。它也支援對大量格式的查詢。如需詳細資訊，請參閱 Amazon Athena 使用者指南中 [Supported SerDes and Data Formats](#)。若要分析大型日誌檔集，可以執行 Amazon EMR 叢集來執行 PB 級分析。

AWS 合作夥伴和第三方提供了許多工具，可用於彙總、處理、儲存和分析。這些工具包含 New Relic、Splunk、Loggly、Logstash、CloudHealth 和 Nagios。但是，系統和應用程式日誌之外的產生對於每個雲端提供者都是唯一的，並且通常對於每個服務也都是唯一的。

資料管理是監控程序中常常被忽略的部分。您需要確定監控資料的保留要求，然後相應地套用生命週期政策。Amazon S3 可支援 S3 儲存貯體層級的生命週期管理。該生命週期管理能以不同方式套用至儲存貯體中的不同路徑。在生命週期即將結束時，您可以將資料傳輸到 Amazon Glacier 進行長期儲存，然後在保留期結束後到期。S3 智慧型分層儲存類別旨在透過自動將資料移至最經濟實惠的存取層來優化成本，而不會影響效能或營運開銷。

未建立此最佳實務時的曝險等級：中

實作指引

- CloudWatch Logs Insights 允許您以互動方式搜尋和分析 Amazon CloudWatch Logs 中的日誌資料。

- [使用 CloudWatch Logs Insights 分析日誌資料](#)
- [Amazon CloudWatch Logs Insights 範例查詢](#)
- 使用 Amazon CloudWatch Logs，將日誌傳送到您可以使用的 Amazon S3 或 Amazon Athena 以查詢資料。
- [我要如何使用 Athena 來分析 Amazon S3 伺服器存取日誌？](#)
 - 為您的伺服器存取日誌儲存貯體建立 S3 生命週期政策。設定生命週期政策以定期移除日誌檔案。這麼做可降低 Athena 分析每個查詢時的資料量。
 - [如何建立 S3 儲存貯體的生命週期政策？](#)

資源

相關文件：

- [Amazon CloudWatch Logs Insights 範例查詢](#)
- [使用 CloudWatch Logs Insights 分析日誌資料](#)
- [使用 Amazon CloudWatch Synthetics 和 AWS X-Ray 進行偵錯](#)
- [如何建立 S3 儲存貯體的生命週期政策？](#)
- [我要如何使用 Athena 來分析 Amazon S3 伺服器存取日誌？](#)
- [一個可觀測性研討會](#)
- [Amazon 建置者資料中心：偵測分散式系統，以了解運作狀態](#)

REL06-BP06 定期審查監控範圍和指標

經常檢閱工作負載監控的實作情形，並隨著工作負載及其架構的演進更新。定期稽核監控有助於降低遺漏或忽略問題指標的風險，並進一步協助工作負載達成其可用性目標。

有效的監控是以關鍵業務指標為基礎，這些指標會隨著業務優先事項的改變而演進。您的監控審查程序應強調服務層級指標 (SLI)，並納入基礎設施、應用程式、用戶端和使用者的深入分析。

預期成果：您擬訂一套有效的監控策略，它會定期審查和更新，以及在任何重大事件或變更之後更新。您確認隨著工作負載和業務需求的發展，關鍵應用程式運作狀態指標仍保持相關。

常見的反模式：

- 您只收集預設指標。
- 您擬訂了監控策略，但未曾檢閱它。

- 您未在部署重大變更時討論監控。
- 您相信過時的指標來判斷工作負載運作狀態。
- 由於指標和閾值過時，導致您的營運團隊因誤報而疲於奔命。
- 您缺乏未受監控之應用程式元件的可觀測性。
- 您只專注於低層級技術指標，並排除業務指標未加監控。

建立此最佳實務的優勢：若您定期審查監控，就可以預測潛在問題，並確認您能夠偵測到這些問題。它還能讓您發現在早期審查期間可能遺漏的盲點，藉此進一步改善您偵測問題的能力。

未建立此最佳實務時的曝險等級：中

實作指引

在您的[營運整備度審查 \(ORR\)](#) 程序中檢閱監控指標和範圍。依照一致的排程執行定期營運整備度審查，以評估目前工作負載與您設定的監控之間是否有任何差距。建立定期執行營運效能審查和知識共享的機制，以增強您從營運團隊獲得更高效能的能力。驗證現有的警示閾值是否仍適當，並確認是否發生營運團隊收到誤報，或未監控應監控之應用程式層面的情況。

[彈性分析架構](#)提供了實用的指引，可協助您進行整個程序。架構的重點在於識別潛在的失敗模式，以及您可採用哪些預防和修正控制來減輕其影響。這些知識可協助您確定要監控和警示的正確指標和事件。

實作步驟

1. 排程及定期審查工作負載儀表板。您對於檢查深度可能有不同規律。
2. 檢查指標中的趨勢。比較指標值與歷史值，以查看是否有趨勢可能指出某項需要調查的事務。這類範例包括：延遲增加、主要業務功能降低，以及失敗回應增加。
3. 檢查指標中是否有極端值和異常，這些可能被平均值或中位數掩蓋。查看時間範圍內的最高和最低值，並調查遠超出正常界限的觀測原因。當您繼續消除這些原因，您就可以設定更嚴格的預期指標界限，來回應獲得改善的工作負載效能一致性。
4. 尋找行為中的急劇變化。指標的數量或方向立即變更，可能表示應用程式有所變更，或您可能需要新增其他指標以追蹤的外部因素。
5. 檢閱目前的監控策略是否仍與應用程式相關。根據先前事件 (或彈性分析架構) 的分析，評估是否有其他應用程式層面應納入監控範圍。
6. 檢閱您的實際使用者監控 (RUM) 指標，以判斷應用程式功能涵蓋範圍是否有任何差距。
7. 檢閱您的變更管理程序。視需要更新您的程序，以納入應在核准變更之前執行的監控分析步驟。
8. 在營運整備度審查過程中確實檢閱監控並修正錯誤程序。

資源

相關的最佳實務

- [REL06-BP01 監控工作負載的所有元件 \(產生\)](#)
- [REL06-BP02 定義和計算指標 \(彙總\)](#)
- [REL06-BP07 透過您的系統監控請求的端對端追蹤](#)
- [REL12-BP02 執行事件後分析](#)
- [REL12-BP06 定期進行演練日](#)

相關文件：

- [為什麼您應該制定錯誤糾正 \(COE\)](#)
- [使用 Amazon CloudWatch 儀表板](#)
- [建置用於檢視營運狀況的儀表板](#)
- [進階多可用區域彈性模式 - 微小故障](#)
- [Amazon CloudWatch Logs Insights 範例查詢](#)
- [使用 Amazon CloudWatch Synthetics 和 AWS X-Ray 進行偵錯](#)
- [一個可觀測性研討會](#)
- [Amazon 建置者資料中心：偵測分散式系統，以了解運作狀態](#)
- [使用 Amazon CloudWatch 儀表板](#)
- [AWS 可觀測性最佳實務](#)
- [彈性分析架構](#)
- [彈性分析架構 - 可觀測性](#)
- [營運整備度審查 - ORR](#)

REL06-BP07 透過您的系統監控請求的端對端追蹤

在透過服務元件處理請求時追蹤請求，讓產品團隊可以更輕鬆地分析和偵錯問題，並改善效能。

預期成果：對所有元件進行全面追蹤的工作負載易於偵錯，透過簡化根本原因探查，提高了錯誤和延遲的[平均解決時間](#) (MTTR)。端對端追蹤可縮短探索受影響的元件時間，並詳細剖析錯誤或延遲的根本原因。

常見的反模式：

- 追蹤可用於某些元件，但不適用於所有元件。例如，在未追蹤 AWS Lambda 的情況下，團隊可能無法清楚了解尖峰工作負載中的冷啟動所造成的延遲。
- 未對追蹤設定綜合 Canary 或實際使用者監控 (RUM)。若沒有 Canary 或 RUM，追蹤分析就會省略用戶端互動遙測，而產生不完整的效能設定檔。
- 混合式工作負載同時包含雲端原生和第三方追蹤工具，但未採取相關步驟來選擇及完全整合單一追蹤解決方案。根據選擇的追蹤解決方案，應使用雲端原生追蹤 SDK 來檢測不是雲端原生的元件，或使用第三方工具來擷取雲端原生追蹤遙測。

建立此最佳實務的優勢：當開發團隊收到問題的提醒時，他們可以看到系統元件互動的全貌，包括個別元件與日誌記錄、效能和失敗的關聯性。由於追蹤可讓您輕鬆地以視覺化方式識別根本原因，調查根本原因的所需時間將可縮短。詳細了解元件互動的團隊，可在解決問題時做出更明智、更快速的決策。諸如何時應叫用災難復原 (DR) 容錯移轉，或何處最適合實作自我修復策略之類的決策，可藉由分析系統追蹤來改善，最終提升客戶對服務的滿意度。

未建立此最佳實務時的曝險等級：中

實作指引

操作分散式應用程式的團隊，可使用追蹤工具來建立關聯性識別碼、收集請求追蹤，以及建置連網元件的服務圖。所有應用程式元件均應包含在請求追蹤中，包括服務用戶端、中介軟體閘道和事件匯流排、運算元件和儲存體 (包括鍵值存放區和資料庫)。在端對端追蹤組態中包含綜合 Canary 和實際使用者監控，以測量遠端用戶端互動和延遲，以便您根據服務水準協議和目標正確評估系統效能。

您可以使用 [AWS X-Ray](#) 和 [Amazon CloudWatch 應用程式監控](#) 檢測服務，在請求透過您的應用程式傳送時提供完整檢視。X-Ray 會收集應用程式遙測，並可讓您跨承載、函數、追蹤、服務、API 進行視覺化和篩選，並可針對無程式碼或低程式碼的系統元件開啟。CloudWatch 應用程式監控包括 ServiceLens，可將您的追蹤與指標、記錄和警示整合。CloudWatch 應用程式監控也包含用來監控端點和 API 的綜合功能，以及用來檢測 Web 應用程式用戶端的實際使用者監控。

實作步驟

- 在所有受支援的原生服務 (例如 [Amazon S3](#)、[AWS Lambda](#) 和 [Amazon API Gateway](#)) 上使用 AWS X-Ray。這些 AWS 服務可使用基礎設施即程式碼、AWS SDK 或 AWS 管理主控台 來啟用具有組態切換的 X-Ray。
- 檢測應用程式 [AWS Distro for Open Telemetry](#) 和 [X-Ray](#) 或第三方收集代理程式。
- 審核 [AWS X-Ray 開發人員指南](#)，了解程式設計語言特定實作。這些文件章節會詳細說明如何檢測 HTTP 請求、SQL 查詢，以及應用程式設計語言特有的其他程序。

- 使用適用於 [Amazon CloudWatch Synthetics Canary](#) 和 [Amazon CloudWatch RUM](#) 的 X-Ray 追蹤，分析從最終使用者用戶端到下游 AWS 基礎設施的請求路徑。
- 根據資源運作狀態和 Canary 遙測來設定 CloudWatch 指標和警示，以便團隊快速收到問題的提醒，然後可使用 ServiceLens 深入探討追蹤和服務圖。
- 如果將第三方工具用於主要追蹤解決方案，請為第三方追蹤工具 (例如 [Datadog](#)、[New Relic](#) 或 [Dynatrace](#)) 啟用 X-Ray 整合。

資源

相關的最佳實務：

- [REL06-BP01 監控工作負載的所有元件 \(產生\)](#)
- [REL11-BP01 監控工作負載的所有元件以偵測故障](#)

相關文件：

- [什麼是 AWS X-Ray？](#)
- [Amazon CloudWatch：應用程式監控](#)
- [使用 Amazon CloudWatch Synthetics 和 AWS X-Ray 進行偵錯](#)
- [Amazon 建置者資料中心：偵測分散式系統，以了解運作狀態](#)
- [將 AWS X-Ray 與其他 AWS 服務整合](#)
- [AWS Distro for OpenTelemetry 和 AWS X-Ray](#)
- [Amazon CloudWatch：使用綜合監控](#)
- [Amazon CloudWatch：使用 CloudWatch RUM](#)
- [設定 Amazon CloudWatch Synthetics Canary 和 Amazon CloudWatch 警示](#)
- [可用性和超越各種可能：了解和改善 AWS 上分散式系統的恢復能力](#)

相關範例：

- [一個可觀測性研討會](#)

相關影片：

- [AWS re:Invent 2022 - 如何跨多個帳戶監控應用程式](#)

- [如何監控您的 AWS 應用程式](#)

相關工具：

- [AWS X-Ray](#)
- [Amazon CloudWatch](#)
- [Amazon Route 53](#)

REL 7. 如何設計工作負載以因應需求的變化？

可擴展的工作負載提供可自動新增或移除資源的彈性，讓資源能夠在任何特定時間點充分滿足目前的需求。

最佳實務

- [REL07-BP01 取得或擴展資源時使用自動化](#)
- [REL07-BP02 在偵測到工作負載受損時取得資源](#)
- [REL07-BP03 偵測到工作負載需要更多資源時取得資源](#)
- [REL07-BP04 Load 測試您的工作負載](#)

REL07-BP01 取得或擴展資源時使用自動化

雲端可靠性的基石，是基礎設施和資源的程式設計定義、佈建和管理。自動化可協助您簡化資源佈建、促進一致且安全的部署，以及在整個基礎設施中擴展資源。

預期成果：您管理自己的基礎設施即程式碼 (IaC)。您可以在版本控制系統 (VCS) 中定義和維護自己的基礎設施程式碼。您可以將佈建 AWS 資源的工作委派給自動化機制，並利用 Application Load Balancer (ALB)、Network Load Balancer (NLB) 和 Auto Scaling 群組等受管服務。您使用持續整合/持續交付 (CI/CD) 管道來佈建資源，以讓程式碼變更自動初始化資源更新，包括 Auto Scaling 組態的更新。

常見的反模式：

- 您使用命令列或在 AWS 管理主控台 (也稱為 click-ops) 手動部署資源。
- 您緊密耦合應用程式元件或資源，因而建立了缺乏彈性的架構。
- 您實作缺乏彈性的擴展政策，因而無法適應不斷變化的業務需求、流量模式或新的資源類型。
- 您手動預估容量來應付預測的需求。

建立此最佳實務的優勢：基礎設施即程式碼 (IaC) 可讓您以程式設計方式定義基礎設施。這可協助您透過與應用程式變更相同的軟體開發生命週期來管理基礎設施變更，進而提高一致性和可重複性，並降低手動易出錯任務的風險。您可以使用自動化交付管道實作 IaC，以進一步簡化佈建和更新資源的程序。您採取可靠且有效率的方式部署基礎設施更新，而無需手動介入。在擴展資源以應付不斷變化的需求時，這種敏捷性尤其重要。

您可以結合 IaC 和交付管道來實現動態的自動化資源擴展。透過監控關鍵指標並套用預先定義的擴展政策，Auto Scaling 就可視需要自動佈建或取消佈建資源，進而改善效能和成本效益。這可降低手動錯誤或延遲的可能性，以回應應用程式或工作負載需求的改變。

IaC、自動化交付管道與 Auto Scaling 相互結合之下，可協助組織安心佈建、更新和擴展其環境。這種自動化對於維護反應靈敏、具彈性且高效管理的雲端基礎設施至關重要。

未建立此最佳實務時的風險暴露等級：高

實作指引

若要為您的 AWS 架構設定具有 CI/CD 管道和基礎設施即程式碼 (IaC) 的自動化，請選擇 Git 這類版本控制系統來儲存您的 IaC 範本和組態。這些範本可以使用 [AWS CloudFormation](#) 等工具撰寫。請從在這些範本內定義您的基礎設施元件開始 (例如 AWS VPC、Amazon EC2 Auto Scaling 群組和 Amazon RDS 資料庫)。

接著將這些 IaC 範本與 CI/CD 管道整合，以自動化部署程序。[AWS CodePipeline](#) 提供了順暢的 AWS 原生解決方案，您也可以使用其他第三方 CI/CD 解決方案。建立管道，當您的版本控制儲存庫發生變更時，就會啟用該管道。設定管道以包含檢查和驗證 IaC 範本的階段、將基礎設施部署到預備環境、執行自動化測試，最後再部署到正式作業環境。必要時納入核准步驟，以維持對變更的控制。此自動化管道不僅能加快部署速度，還能促進環境之間的一致性和可靠性。

在 IaC 中設定 Amazon EC2 執行個體、Amazon ECS 任務和資料庫複本等資源的 Auto Scaling，以視需要提供自動橫向擴充和縮減。此方法可增強應用程式可用性和效能，並根據需求動態調整資源來最佳化成本。如需支援的資源清單，請參閱 [Amazon EC2 Auto Scaling](#) 和 [AWS Auto Scaling](#)。

實作步驟

1. 建立並使用原始程式碼儲存庫來存放控制基礎設施組態的程式碼。對此儲存庫執行變更，以反映您要進行的任何持續變更。
2. 選取基礎設施即程式碼解決方案 (例如 AWS CloudFormation) 讓您的基礎設施保持最新狀態，並偵測與您預期狀態不一致 (偏離) 的情況。
3. 將 IaC 平台與您的 CI/CD 管道整合，以自動化部署。

4. 確定並收集適當的指標，以自動擴展資源。
5. 使用適合您工作負載元件的橫向擴充和縮減政策來設定自動擴展資源。考慮針對可預測的用量模式使用排程擴展。
6. 監控部署以偵測失敗和迴歸。在 CI/CD 平台內實作回復機制，以在必要時還原變更。

資源

相關文件：

- [AWS Auto Scaling：擴展計畫的運作方式](#)
- [AWS Marketplace：可與 Auto Scaling 結合使用的產品](#)
- [使用 DynamoDB Auto Scaling 功能自動管理輸送容量](#)
- [將負載平衡器與 Auto Scaling 群組配合使用](#)
- [什麼是 AWS Global Accelerator？](#)
- [什麼是 Amazon EC2 Auto Scaling？](#)
- [什麼是 AWS Auto Scaling？](#)
- [什麼是 Amazon CloudFront？](#)
- [什麼是 Amazon Route 53？](#)
- [什麼是 Elastic Load Balancing？](#)
- [什麼是 Network Load Balancer？](#)
- [什麼是 Application Load Balancer？](#)
- [將 Jenkins 與 AWS CodeBuild 和 AWS CodeDeploy 整合](#)
- [使用 AWS CodePipeline 建立一個四階段的管道](#)

相關影片：

- [回歸基礎：將程式碼部署至 Amazon EC2](#)
- [AWS 支援您 | 使用 AWS CloudFormation 範本開始使用您的基礎設施即程式碼解決方案](#)
- [使用 AWS CodePipeline 簡化您的軟體版本程序](#)
- [使用 Amazon CloudWatch 儀表板監控 AWS 資源](#)
- [建立跨帳戶和跨區域 CloudWatch 儀表板 | Amazon Web Services](#)

REL07-BP02 在偵測到工作負載受損時取得資源

在可用性受到影響時視需要主動擴展資源，以還原工作負載可用性。

您必須先設定運作狀態檢查和這些檢查的條件，以指出可用性因資源不足而受到影響的時間。然後，通知適當的人員手動擴展資源，或啟動自動化以自動調整資源規模。

您可以針對工作負載手動調整規模 (例如，變更 Auto Scaling 群組中的 EC2 執行個體數量，或透過 AWS 管理主控台 或 AWS CLI 修改 DynamoDB 資料表的輸送量)。但是，應盡可能使用自動化 (請參閱取得或擴展資源時使用自動化)。

預期成果：啟動擴展活動 (自動或手動)，以在偵測到故障或客戶體驗降級時恢復可用性。

未建立此最佳實務時的曝險等級：中

實作指引

在工作負載中的所有元件實作可觀測性和監控，以監控客戶體驗並偵測故障。定義可調整所需資源的手動或自動化程序。o 如需詳細資訊，請參閱 [REL11-BP01 監控工作負載的所有元件以偵測故障](#)。

實作步驟

- 定義會擴展所需資源的手動或自動程序。
 - 擴展程序取決於工作負載內不同元件的設計方式。
 - 擴展程序也會根據所使用的基礎技術而有所不同。
 - 使用 AWS Auto Scaling 的元件可以使用擴展計劃來設定用於擴展資源的一組指示。如果使用 AWS CloudFormation 或新增標籤到 AWS 資源，您可以針對每個應用程式的不同資源組來設定擴展計畫。Auto Scaling 為針對每個資源自訂的擴展策略提供建議。建立擴展計畫之後，Auto Scaling 會將動態擴展和預測擴展方法結合在一起，以支援您的擴展策略。有關詳細資訊，請參閱 [How scaling plans work](#)。
 - Amazon EC2 Auto Scaling 會確認您有數量正確的 Amazon EC2 執行個體來處理應用程式的負載。您可以建立 EC2 執行個體的集合，此集合稱為「Auto Scaling 群組」。您可以在每個 Auto Scaling 群組中指定執行個體的最小和最大數量，而 Amazon EC2 Auto Scaling 可確保您的群組大小永遠不會低於或高於這些限制。如需詳細資訊，請參閱 [What is Amazon EC2 Auto Scaling?](#)
 - Amazon DynamoDB Auto Scaling 功能使用 Application Auto Scaling 服務代您動態調整佈建的輸送容量，以此回應實際流量模式。這可讓資料表或全域次要索引增加其佈建的讀取與寫入容量，以處理突然增加的流量，而不需限流。如需詳細資訊，請參閱 [Managing throughput capacity automatically with DynamoDB auto scaling](#)。

資源

相關的最佳實務：

- [REL07-BP01 取得或擴展資源時使用自動化](#)
- [REL11-BP01 監控工作負載的所有元件以偵測故障](#)

相關文件：

- [AWS Auto Scaling：擴展計劃的運作方式](#)
- [使用 DynamoDB Auto Scaling 功能自動管理輸送容量](#)
- [什麼是 Amazon EC2 Auto Scaling？](#)

REL07-BP03 偵測到工作負載需要更多資源時取得資源

雲端運算最有價值的功能之一，就是動態佈建資源的能力。

在傳統的內部部署運算環境中，您必須預先確定和佈建足夠的容量，以滿足尖峰需求。這點卻是問題所在，因為它很昂貴，而且如果您低估工作負載的尖峰容量需求，它可能會對可用性造成風險。

在雲端中，您就不需要這樣做。而是能夠視需要佈建運算、資料庫和其他資源容量，以滿足目前和預測的需求。Amazon EC2 Auto Scaling 和 Application Auto Scaling 等自動化解決方案可以根據您指定的指標，自動線上提供資源。這使得擴展程序更容易且可預測，同時可確保您隨時有足夠的資源可用，進而大幅提高工作負載的可靠性。

預期成果：您設定自動擴展運算和其他資源以滿足需求。您在擴展政策中提供足夠的預留空間，以便在其他資源上線時處理激增的流量。

常見的反模式：

- 您佈建固定數量的可擴展資源。
- 您選擇的擴展指標與實際需求無關。
- 您無法在擴展計畫中提供足夠的預留空間來應付需求激增的情況。
- 您的擴展政策太晚增加容量，以致於在其他資源上線時造成容量耗盡和服務降級的情況。
- 您未能正確設定資源計數的下限和上限，因而導致擴展失敗。

建立此最佳實務的優勢：擁有足夠的資源可滿足目前的需求，這點對於提供工作負載的高可用性並遵守您定義的服務層級目標 (SLO) 來說至關重要。自動擴展可讓您提供工作負載所需的確切運算、資料庫

和其他資源數量，以便處理目前和預測的需求。您不需要判斷尖峰容量需求，並靜態配置資源來處理這些需求。您可以改為隨著需求增加配置更多資源來應付這些需求，並且在需求減少之後停用資源以降低成本。

未建立此最佳實務時的曝險等級：中

實作指引

首先，確定工作負載元件是否適合自動擴展。這些元件稱為可水平擴展，因為它們提供相同的資源且行為相同。可水平擴展元件的範例包括相似設定的 EC2 執行個體、[Amazon Elastic Container Service \(ECS\)](#) 任務，以及在 [Amazon Elastic Kubernetes Service \(EKS\)](#) 上執行的 Pod。這些運算資源通常位於負載平衡器後方，稱為複本。

其他複寫資源可能包括資料庫讀取複本、[Amazon DynamoDB](#) 資料表和 [Amazon ElastiCache](#) (Redis OSS) 叢集。如需可支援資源的完整清單，請參閱[可與 Application Auto Scaling 搭配使用的 AWS 服務](#)。

對於容器型架構，您可能需要採用兩種不同的方式進行擴展。第一種方式是，您可能需要擴展提供可水平擴展服務的容器。第二種方式是，您可能需要擴展運算資源以騰出空間給新容器。每一層都有不同的自動擴展機制。若要擴展 ECS 任務，您可以使用 [Application Auto Scaling](#)。若要擴展 Kubernetes Pod，您可以使用 [Horizontal Pod Autoscaler \(HPA\)](#) 或 [Kubernetes Event-driven Autoscaling \(KEDA\)](#)。若要擴展運算資源，您可以針對 ECS 使用[容量提供者](#)，對於 Kubernetes，您可以使用 [Karpenter](#) 或 [Cluster Autoscaler](#)。

接著選取您要執行自動擴展的方式。有三個主要選項：指標型擴展、排程擴展和預測性擴展。

指標型擴展

指標型擴展會根據一或多個擴展指標的值佈建資源。擴展指標是指對應工作負載需求的指標。確定擴展指標是否適當的好方法，就是在非正式作業環境中執行負載測試。在負載測試期間，將可擴展資源的數量固定，並慢慢增加需求 (例如輸送量、並行或模擬的使用者)。然後尋找隨著需求增加 (或減少) 而增加的指標，以及隨著需求下降而減少 (或增加) 的指標。典型的擴展指標包括 CPU 使用率、工作佇列深度 (例如 [Amazon SQS](#) 佇列)、作用中使用者數量，以及網路輸送量。

Note

AWS 已觀測到，在大多數應用程式中，記憶體使用率會隨著應用程式暖機而增加，然後達到穩定值。當需求減少時，記憶體使用率通常會保持在升高處，而不會跟著減少。由於記憶體使用率不會隨著需求增加和減少而對應升降，因此在選取此指標用於自動擴展時，務必審慎考慮。

指標型擴展是延遲操作。利用率指標可能需要幾分鐘的時間才能傳播至自動擴展機制，而且這些機制通常會等待需求增加的明確訊號，然後才做出反應。於是，當自動擴展器建立新資源時，可能需要額外的時間才能提供完整服務。因此，請務必不要將擴展指標目標設定得太接近完全使用率 (例如 90% CPU 使用率)。這樣做可能產生在額外容量上線之前就耗盡現有資源容量的風險。典型的資源使用率目標範圍介於 50-70% 之間，此範圍可實現最佳可用性，具體取決於佈建其他資源的需求模式和所需的時間。

排程擴展

排程擴展會根據行事曆或一天當中的時間佈建或移除資源。它經常用於具有可預測需求的工作負載，例如工作日營業時間或銷售活動的尖峰使用率。[Amazon EC2 Auto Scaling](#) 和 [Application Auto Scaling](#) 都支援排程擴展。KEDA 的 [cron 擴展器](#) 支援 Kubernetes Pod 的排程擴展。

預測性擴展

預測性擴展使用機器學習，根據預期的需求自動擴展資源。預測性擴展會分析您提供的使用率指標的歷史值，並持續預測其未來值。然後使用預測的值來向上擴展資源或縮減其規模。[Amazon EC2 Auto Scaling](#) 可執行預測性擴展

實作步驟

1. 確定工作負載元件是否適合自動擴展。
2. 確定哪一種擴展機制最適合工作負載：指標型擴展、排程擴展或預測性擴展。
3. 為元件選取適當的自動擴展機制。對於 Amazon EC2 執行個體，使用 Amazon EC2 Auto Scaling。對於其他 AWS 服務，使用 Application Auto Scaling。對於 Kubernetes Pod (例如在 Amazon EKS 叢集中執行的 Pod)，請考慮 Horizontal Pod Autoscaler (HPA) 或 Kubernetes Event-driven Autoscaling (KEDA)。對於 Kubernetes 或 EKS 節點，請考慮 Karpenter 和 Cluster Auto Scaler (CAS)。
4. 對於指標或排程擴展，請執行負載測試，以確定適合工作負載的擴展指標和目標值。對於排程擴展，確定您選取的日期和時間所需的資源數量。確定滿足預期的尖峰流量所需的資源數量上限。
5. 根據上面收集的資訊設定自動擴展器。如需詳細資訊，請參閱自動擴展服務的文件。確認已正確設定擴展的上限和下限。
6. 確認擴展組態依預期運作。在非正式作業環境中執行負載測試，以及觀察系統如何反應，並視需要調整。在正式作業環境中啟用自動擴展時，請設定適當的警報來通知您任何非預期的行為。

資源

相關文件：

- [什麼是 Amazon EC2 Auto Scaling ?](#)

- [AWS 規範性指引：負載測試應用程式](#)
- [AWS Marketplace：可與 Auto Scaling 結合使用的產品](#)
- [使用 DynamoDB Auto Scaling 功能自動管理輸送容量](#)
- [EC2 的預測擴展，採用機器學習技術](#)
- [Amazon EC2 Auto Scaling 排程擴展](#)
- [說說「利特爾法則」的故事](#)

REL07-BP04 Load 測試您的工作負載

採用負載測試方法來衡量擴展活動是否滿足工作負載要求。

重要的是執行持續的負載測試。負載測試應探索突破點，並測試工作負載的效能。AWS 可讓您輕鬆設定臨時測試環境，以建立生產工作負載規模的模型。在雲端，您可隨需建立生產規模的測試環境、完成測試，然後再停用資源。因為您只為執行中的測試環境付費，所以能以與內部部署測試相較之下相當微小比例的成本來模擬即時環境。

在生產系統承受壓力的演練日，以及客戶使用量較低的時間內，您也應考慮在生產中進行負載測試，並且讓可用的所有人員解釋結果並解決所發生的任何問題。

常見的反模式：

- 在與生產組態不同的部署上執行負載測試。
- 僅對工作負載的個別部分執行負載測試，而非整個工作負載。
- 使用一部分請求而非代表性的一組實際請求來執行負載測試。
- 對高於預期負載的小型安全係數執行負載測試。

建立此最佳實務的優勢：您會知道架構中的哪些元件在負載時失敗，並能夠識別要監看哪些指標，指出您正在及時處理該負載來解決問題，避免受到該故障的影響。

未建立此最佳實務時的曝險等級：中

實作指引

- 執行負載測試，以識別工作負載的哪些層面指出您必須新增或移除容量。負載測試的代表性流量應與您在生產環境中收到的流量相似。在觀看您已檢測的指標時增加負載，以判斷哪些指標指出何時須新增或移除資源。
- [上的分散式負載測試 AWS：模擬數千個連線使用者](#)

- 識別請求混合。您可能會有不同的請求混合，因此您應該在識別流量混合時查看各種時間範圍。
- 實作載入驅動程式。您可以使用自訂程式碼、開放原始碼或商業軟體實作載入驅動程式。
- 最初使用小容量的負載測試。您將負載驅動到較小容量 (可能和單一執行個體或容器一樣小)，看到一些立即的影響。
- 針對較大容量的負載測試。在分散式負載上的效果會有所不同，因此您必須盡可能在接近產品環境的條件下進行測試。

資源

相關文件：

- [上的分散式負載測試 AWS：模擬數千個連線的使用者](#)
- [負載測試應用程式](#)

相關影片：

- [AWS Summit ANZ 2023：透過 AWS 分散式負載測試，放心加速](#)

REL 8. 如何實作變更？

變更須在受控的情況下，才能部署新功能，並確認工作負載和運作環境執行已知的軟體，且能夠以可預測的方式修補或取代。如果這些變更不受控制，那麼就很難預測這些變更的影響，也很難解決由於這些變更而產生的問題。

最佳實務

- [REL08-BP01 將執行手冊用於部署等標準活動](#)
- [REL08-BP02 將功能測試整合為部署的一部分](#)
- [REL08-BP03 將彈性測試整合為部署的一部分](#)
- [REL08-BP04 使用不可變基礎設施進行部署](#)
- [REL08-BP05 使用自動化部署變更](#)

REL08-BP01 將執行手冊用於部署等標準活動

執行手冊是實現特定結果的預定義程序。使用執行手冊執行手動或自動進行的標準活動。範例包括部署工作負載、修補工作負載或進行 DNS 修改。

例如，實施程序以[確保部署期間的回復安全性](#)。確保您可以回復部署，且不會對客戶造成任何中斷，這對於打造可靠的服務而言至為關鍵。

對於執行手冊程序，從有效的手動過程開始，以程式碼實作它，並在適當時調用它以自動執行。

即使是高度自動化的複雜工作負載，執行手冊仍然適用於[執行演練日](#)或滿足嚴格的報告和稽核要求。

請注意，程序手冊用於回應特定事件，而執行手冊用於實現特定成果。執行手冊通常用於例行活動，而程序手冊則用於回應非例行事件。

常見的反模式：

- 在生產環境中對組態執行非計劃中的變更。
- 為了更快速地部署而略過計畫中的步驟，會導致部署失敗。
- 在不測試變更反轉的情況下進行變更。

建立此最佳實務的優勢：有效的變更規劃可提高您成功執行變更的能力，因為您知道所有受影響的系統。在測試環境中驗證變更可增強信心。

未建立此最佳實務時的曝險等級：高

實作指引

- 透過在執行手冊中記錄程序，對熟知的事件提供一致且迅速的回應。
- 使用基礎設施即程式碼的原則來定義您的基礎設施。透過使用 AWS CloudFormation (或受信任的第三方) 來定義您的基礎設施，您可以使用版本控制軟體對變更進行版本控制和追蹤。
 - 使用 AWS CloudFormation (或受信任的第三方供應商) 來定義您的基礎設施。
 - [什麼是 AWS CloudFormation ?](#)
 - 使用良好的軟體設計原則，建立單一旦分離的範本。
 - 確定實作的許可、範本和負責方。
 - [使用 AWS Identity and Access Management 控制存取](#)
 - 使用以 Git 等常用技術為基礎的託管原始程式碼管理系統，來儲存原始程式碼和基礎設施即程式碼 (IaC) 組態。

資源

相關文件：

- [APN 合作夥伴](#)：可以幫助您建立自動化部署解決方案的**合作夥伴**
- [AWS Marketplace](#)：可用於自動化的**產品**
- [什麼是 AWS CloudFormation？](#)

相關範例：

- [使用程序手冊和執行手冊將操作自動化](#)

REL08-BP02 將功能測試整合為部署的一部分

使用可驗證所需功能的技術，例如單元測試和整合測試。

單元測試是您測試程式碼的最小功能單元以驗證其行為的程序。整合測試旨在驗證每一項應用程式功能是否根據軟體需求運作。單元測試著重於單獨測試應用程式的部分，整合測試則會考量副作用 (例如，透過改變操作變更資料的影響)。無論在哪一種情況下，測試都應整合到部署管道中，如果不符合成功條件，則管道會停止或復原。這些測試會在生產前環境中執行，而且會在生產前暫存於管道中。

當這些測試做為建置和部署動作的一部分自動執行時，您會獲得最佳成果。例如，使用 AWS CodePipeline 時，開發人員會將變更遞交至來源儲存庫，而 CodePipeline 會在該儲存庫中自動偵測變更。系統會建置應用程式並執行單元測試。通過單元測試後，建置的程式碼會部署至預備伺服器以進行測試。CodePipeline 會從預備伺服器執行更多測試，例如整合或負載測試。成功完成這些測試後，CodePipeline 會將已測試及已核准的程式碼部署至生產執行個體。

預期成果：您會使用自動化來執行單元和整合測試，以驗證程式碼的行為如預期。這些測試會整合到部署程序中，若測試失敗，則會中止部署。

常見的反模式：

- 您為了加快部署時間表，而忽略或略過在部署過程中的測試失敗和計畫。
- 可以在部署管道之外手動執行測試。
- 您透過手動緊急工作流程，跳過自動化中的測試步驟。
- 您執行自動化的測試環境與正式作業環境並非十分相似。
- 您建置的測試套件不夠靈活，且不易隨著應用程式發展進行維護、更新或擴展。

建立此最佳實務的優勢：在部署過程中進行自動化測試可及早發現問題，進而降低將有錯誤或非預期行為的版本發布至正式作業環境的風險。單元測試會驗證程式碼的行為確實依照要求，並遵循 API

合約。整合測試會驗證系統確實根據指定的需求運作。這些類型的測試會驗證元件 (例如使用者介面、API、資料庫和原始程式碼) 的預定工作順序。

未建立此最佳實務時的風險暴露等級：高

實作指引

採用測試驅動的開發 (TDD) 方法來撰寫軟體，藉此開發測試案例來指定和驗證程式碼。若要開始進行，請為每個函數建立測試案例。如果測試失敗，您會撰寫新的程式碼來通過測試。此方法可協助您驗證每個函數的預期結果。在將程式碼遞交至原始程式碼儲存庫之前，先執行單元測試並確認其通過測試。

在 CI/CD 管道的建置、測試和部署階段，實作單元和整合測試。自動化測試，並在應用程式有新版本可進行部署時自動初始化測試。如果未符合成功條件，則會終止或回復管道。

如果應用程式為 Web 或行動應用程式，則在多個桌面瀏覽器或實際的裝置上執行自動化整合測試。此方法對於驗證行動應用程式在各種不同裝置之間的相容性和功能特別有用。

實作步驟

1. 在撰寫功能程式碼 (測試驅動的開發或 TDD) 之前，先撰寫單元測試。建立程式碼指引，讓寫入和執行單元測試成為非功能性的編碼需求。
2. 建立自動化整合測試套件，當中涵蓋已識別的可測試功能。這些測試應模擬使用者互動並驗證預期結果。
3. 建立必要的測試環境以執行整合測試。當中可能包括與正式作業環境十分相似的預備或正式作業前環境。
4. 使用 AWS CodePipeline 主控台或 AWS Command Line Interface (CLI) 設定來源、建置、測試和部署階段。
5. 程式碼建置並測試完成後，部署應用程式。AWS CodeDeploy 可以將它部署到您的預備 (測試) 和正式作業環境。這些環境可包括 Amazon EC2 執行個體、AWS Lambda 函數或內部部署伺服器。您應使用相同的部署機制將應用程式部署到所有環境。
6. 監控管道的進度和每個階段的狀態。使用品質檢查，根據測試的狀態封鎖管道。也可以接收任何管道階段失敗或管線完成的通知。
7. 持續監控測試結果，並尋找較需要關注的模式、迴歸或區域。利用這些資訊改進測試套件、找出應用程式中需要更完善測試的區域，以及最佳化部署程序。

資源

相關的最佳實務：

- [REL07-BP04 對工作負載執行負載測試](#)
- [REL08-BP03 將彈性測試整合為部署的一部分](#)
- [REL12-BP04 使用混沌工程測試彈性](#)

相關文件：

- [AWS 規範性指引：測試自動化](#)
- [持續交付和持續整合](#)
- [功能測試指標](#)
- [監控管道](#)
- [搭配 AWS CodeBuild 使用 AWS CodePipeline 以測試程式碼並執行建置](#)
- [AWS Device Farm](#)

REL08-BP03 將彈性測試整合為部署的一部分

通過有意識地在系統中引入故障來整合彈性測試，以在破壞性情況下衡量其能力。彈性測試與通常整合在部署週期中的單元和功能測試不同，因為它們專注於識別系統中的意外故障。雖然從生產前的彈性測試整合開始是安全的，但應設定一個目標，在生產環境中實作這些測試，作為[演練日](#)的一部分。

預期成果：彈性測試有助於建立系統承受生產降級能力的信心。實驗可識別會導致故障的弱點，這有助於您改善系統，以自動且有效地緩解故障和降級。

常見的反模式：

- 在部署過程中缺乏可觀測性和監控
- 依賴人類解決系統故障
- 品質不佳的分析機制
- 專注於系統中的已知問題，缺乏識別任何未知因素的實驗
- 可識別故障，但沒有解決方案
- 沒有調查結果和執行手冊文件

建立最佳實務的優勢：在部署中整合的彈性測試有助於識別系統中未知的問題，否則這些問題會被忽視，從而導致生產中停機。在系統中識別這些未知因素可協助您記錄調查結果、將測試整合到您的 CI/CD 程序中以及建置執行手冊，透過高效率、可重複的機制簡化緩解作業。

未建立此最佳實務時的曝險等級：中

實作指引

可以整合到系統部署中的最常見彈性測試表單是災難復原和混沌工程。

- 在任何重要部署中包含災難復原計畫和標準操作程序 (SOP) 的更新。
- 將可靠性測試整合至您的自動化部署管道。諸如 [AWS Resilience Hub](#) 的服務可[整合到您的 CI/CD 管道](#)中，以建立持續的彈性評估，並在每次部署中自動進行評估。
- 在 AWS Resilience Hub 中定義您的應用程式。彈性評估會產生程式碼片段，協助您將復原程序建立為應用程式的 AWS Systems Manager 文件，並提供建議的 Amazon CloudWatch 監控和警示清單。
- DR 計畫和 SOP 更新後，請完成災難復原測試以驗證其是否有效。災難復原可協助您判斷是否可以在事件發生後還原系統並恢復正常操作。您可以模擬各種災難復原策略，並確定您的規劃是否足以滿足您的正常運行時間需求。常見的災難復原策略包括備份與還原、指示燈、冷待命、暖待命、熱待命和主動-主動式，而且成本和複雜性都有所不同。在災難復原之前，建議您定義復原時間點目標 (RTO) 和復原點目標 (RPO)，以簡化要模擬的策略選擇。AWS 提供災難復原工具，例如 [AWS Elastic Disaster Recovery](#)，協助您開始規劃和測試。
- 混沌工程實驗引入了系統中斷，例如網路中斷和服務故障。透過模擬受控故障，您可以發現系統的漏洞，同時控制注入故障的影響。就像其他策略一樣，在非生產環境中使用諸如 [AWS Fault Injection Service](#) 等服務執行受控故障模擬，以在部署到生產環境之前獲得信心。

資源

相關文件：

- [使用彈性測試進行故障實驗以建立復原備](#)
- [使用 AWS Resilience Hub 和 AWS CodePipeline 持續評估應用程式彈性](#)
- [AWS 上的災難復原 \(DR\) 架構，第 1 部分：雲端復原策略](#)
- [使用混沌工程確認工作負載的彈性](#)
- [混沌工程的原則](#)
- [混沌工程研討會](#)

相關影片：

- [AWS re:Invent 2020：使用混沌工程測試彈性](#)

- [利用 AWS 故障注入服務來提升應用程式彈性](#)
- [使用 AWS Resilience Hub 準備並保護您的應用程式免受中斷](#)

REL08-BP04 使用不可變基礎設施進行部署

不可變基礎設施是一種模式，要求在生產工作負載上不進行現場的更新、安全性修補或組態變更。需要進行變更時，會在新的基礎設施上建置架構並部署到生產環境。

請遵循不可變基礎設施的部署策略，以提高工作負載部署中的可靠性、一致性和可重複性。

預期成果：使用不可變基礎設施，不允許[就地修改](#)以執行工作負載內的基礎設施資源。相反地，在需要變更時，會以平行方式與現有資源一起部署新的、包含所有必要變更的更新後基礎設施資源集。此部署會自動進行驗證，如果成功，流量會逐漸轉移到新的資源集。

此部署策略適用於軟體更新、安全修補程式、基礎設施變更、組態更新和應用程式更新等。

常見的反模式：

- 對執行中的基礎設施資源實作就地變更。

建立此最佳實務的優勢：

- 提高跨環境的一致性：由於跨環境的基礎設施資源沒有差異，因此可以提高一致性並簡化測試。
- 降低組態偏移：透過將基礎設施資源取代為已知且版本控制的組態，可將基礎設施設為已知、經過測試且可信狀態，以避免組態偏移。
- 可靠的不可分割部署：部署可以順利完成，也可以保持不變，從而提高部署流程的一致性和可靠性。
- 簡化部署：部署不需要支援升級，因此會得到簡化。升級只是新的部署。
- 利用快速的回復及復原程序打造更安全的部署：前一個運作版本並未變更，因此部署變得更加安全。如果偵測到錯誤，您可以回復至該版本。
- 增強的安全狀態：透過不允許變更基礎設施，可以停用遠端存取機制 (例如 SSH)。這可減少攻擊媒介，從而改善組織的安全狀態。

未建立此最佳實務時的曝險等級：中

實作指引

Automation

定義不可變基礎設施部署策略時，建議盡可能使用[自動化](#)來提高可重複性並將人為錯誤的可能性降至最低。如需詳細資訊，請參閱 [REL08-BP05 透過自動化部署變更](#)和[安全且無需人為干預的自動化部署](#)。

使用[基礎設施即程式碼 \(IaC\)](#) 時，基礎設施佈建、協同運作和部署步驟會以程式設計、描述性和宣告式的方式定義，並儲存在原始檔控制系統中。利用基礎設施即程式碼可讓您更輕鬆地自動部署基礎設施，並協助您實現基礎設施不可變性。

部署模式

需要變更工作負載時，不可變基礎設施的部署策略會要求您部署一組新的基礎設施資源，包括所有必要的變更。這組新資源必須遵循推出模式，以最大程度地降低使用者所受到的影響。此部署有兩個主要策略：

[Canary 部署](#)：將少量客戶導向至新版本的實務，通常會在單一服務執行個體 (Canary) 上執行。之後，您可以仔細檢查所產生的任何行為變更或錯誤。如果遇到嚴重問題，可以從 Canary 中刪除流量，然後將使用者傳送回以前的版本。如果部署成功，則您可以繼續以期望的速度進行部署，同時監控變更是否有錯誤，直到完全部署為止。可以使用允許 Canary 部署的[部署組態](#)來設定 AWS CodeDeploy。

[藍/綠部署](#)：與 Canary 部署類似，不同之處在於整個應用程式須並行部署。您可在兩個堆疊 (藍色和綠色) 之間交替部署。再次強調，您可以將流量傳送到新版本，且如果發現部署問題，則可以回復到舊版本。通常，會一次切換所有流量，但您也可以將一小部分的流量用於每個版本，以使用 Amazon Route 53 的加權 DNS 路由功能，提高新版本的採用率。可以使用允許藍/綠部署的部署組態來設定 AWS CodeDeploy 和 [AWS Elastic Beanstalk](#)。

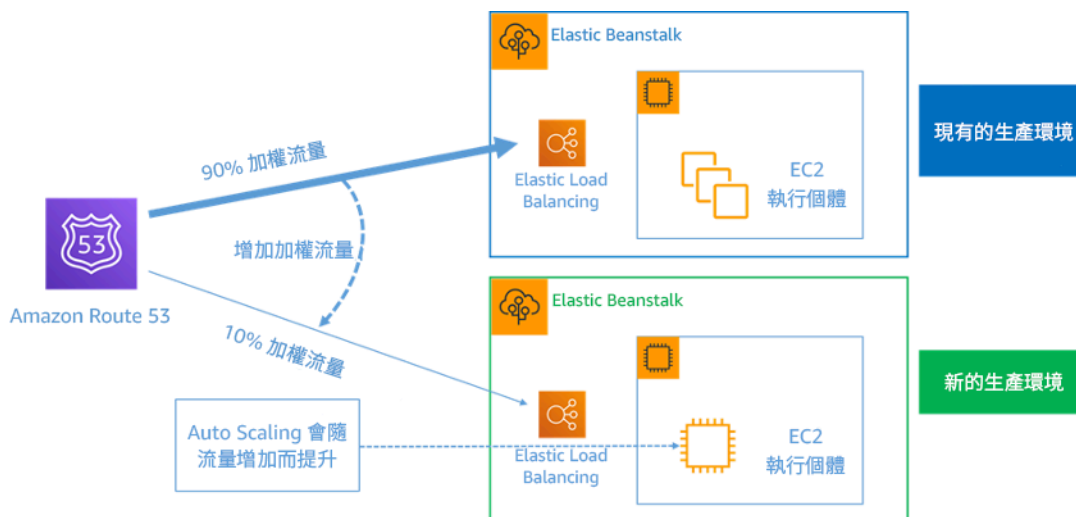


圖 8：使用 AWS Elastic Beanstalk 和 Amazon Route 53 進行藍/綠部署

漂移偵測

漂移被定義為導致基礎設施資源具有與預期不同的狀態或配置的任何變更。任何類型的未受管組態變更都違反了不可變基礎設施的概念，應加以偵測並修正，以便能成功實作不可變基礎設施。

實作步驟

- 禁止就地修改執行中的基礎設施資源。
- 您可以使用 [AWS Identity and Access Management \(IAM\)](#) 指定誰或哪些服務可以存取 AWS 中的服務和資源、集中管理精細權限，以及分析存取以完善 AWS 中的權限。
- 自動部署基礎設施資源以提高可重複性，並最大限度地減少發生人為錯誤的可能性。
- 如 [DevOps on AWS 簡介](#) 中所述，自動化是 AWS 服務的基石，並在所有服務、功能和產品中獲得內部支援。
- [預製](#)您的 Amazon Machine Image (AMI) 可以加快啟動它們的時間。[EC2 Image Builder](#) 是一項全受管 AWS 服務，可協助您自動建立、維護、驗證、共用和部署自訂、安全且最新的 Linux 或 Windows 自訂 AMI。
- 支援自動化的一些服務包括：
 - [AWS Elastic Beanstalk](#) 是一項服務，可在熟悉的伺服器 (例如 Apache、Nginx、Passenger 和 IIS) 上部署和擴展以 Java、.NET、PHP、Node.js、Python、Ruby、Go 和 Docker 開發的 Web 應用程式。
 - [AWS Proton](#) 協助平台團隊連接並協調開發團隊在基礎設施佈建、程式碼部署、監控和更新方面所需的所有不同工具。AWS Proton 可啟用無伺服器和容器型應用程式的自動化基礎設施即程式碼佈建和部署。
- 利用基礎設施即程式碼可讓您輕鬆地自動部署基礎設施，並協助實現基礎設施的不可變性。AWS 會提供能以程式化、描述性和宣告式的方式建立、部署和維護基礎設施的服務。
 - [AWS CloudFormation](#) 幫助開發人員以有序和可預測的方式建立 AWS 資源。資源會使用 JSON 或 YAML 格式以文字檔撰寫。範本需要特定語法和結構，而這取決於所建立和管理的資源類型。您可以使用任何程式碼編輯器以 JSON 或 YAML 撰寫資源、將其簽入版本控制系統，然後 CloudFormation 便會以安全、可重複的方式建置指定的服務。
 - [AWS Serverless Application Model \(AWS SAM\)](#) 是可用來在 AWS 上建置無伺服器應用程式的開放原始碼架構。AWS SAM 與其他 AWS 服務整合，是 CloudFormation 的擴展功能。
 - [AWS Cloud Development Kit \(AWS CDK\)](#) 是一個開放原始碼軟體開發架構，可用來建模，並使用熟悉的程式設計語言來佈建您的雲端應用程式資源。您可以使用 AWS CDK 透過 TypeScript、Python、Java 和 .NET 來建模應用程式基礎設施。AWS CDK 會在背景中使用 CloudFormation 以透過安全、可重複的方式佈建資源。

- [AWS 雲端控制 API](#) 引入一組通用的「建立」、「讀取」、「更新」、「刪除」和「清單」(CRUDL) API，協助開發人員以簡單且一致的方式管理雲端基礎設施。Cloud Control API 通用 API 可讓開發人員統一管理 AWS 和第三方服務的生命週期。
- 實作可將使用者所受到的影響降到最低的部署模式。
 - Canary 部署：
 - [設定 API Gateway Canary 發行部署](#)
 - [使用 AWS App Mesh 為 Amazon ECS 建立具有 Canary 部署的管道](#)
 - 藍/綠部署：[《AWS 上的藍/綠部署》白皮書](#)說明了實作藍/綠部署策略的範例技術。
- 偵測組態或狀態的偏移。如需詳細資訊，請參閱 [Detecting unmanaged configuration changes to stacks and resources](#)。

資源

相關的最佳實務：

- [REL08-BP05 使用自動化部署變更](#)

相關文件：

- [自動化安全、無人為介入的部署](#)
- [利用 AWS CloudFormation 在 Nubank 建立不可變基礎設施](#)
- [基礎設施即程式碼](#)
- [實作警示以自動檢測 AWS CloudFormation 堆疊中的漂移](#)

相關影片：

- [AWS re:Invent 2020：透過不變性實現可靠性、一致性和信心](#)

REL08-BP05 使用自動化部署變更

部署和修補經過自動化以消除負面影響。

改變生產系統是許多組織的最大風險領域之一。我們認為，相較於軟體要解決的業務問題，部署才是我們要解決的首要問題。如今，這表示在營運中實際可行的地方使用自動化，包括測試和部署變更，新增或刪除容量以及移轉資料。

期望成果：可以透過廣泛的生產前測試、自動復原和交錯的生產部署，在發佈過程中建置自動化部署安全性。這種自動化可將部署失敗所造成的潛在影響降到最低，而且開發人員不再需要主動觀察部署到生產環境的情況。

常見的反模式：

- 可以執行手動變更。
- 可以透過手動緊急工作流程，略過自動化中的步驟。
- 您不會遵循既定的計劃和流程，以加快時間表。
- 可以在不考慮封裝時間的情況下執行快速後續部署。

建立此最佳實務的優勢：當您使用自動化來部署所有變更時，可以移除導致人為錯誤的可能性，並能夠在變更生產之前進行測試。在生產推送之前執行此程序可驗證您的計劃是否已完成。此外，自動復原至您的發佈程序可以識別生產問題，並將工作負載恢復到先前的作業狀態。

未建立此最佳實務時的曝險等級：中

實作指引

自動化您的部署管道。部署管道讓您可以叫用自動測試、偵測異常，或者在生產部署之前的某個步驟中停止管道，或者自動回復變更。其中不可或缺的一部分就是採用[持續整合和持續交付/部署 \(CI/CD\)](#) 的文化，在這種文化中，提交或程式碼變更會經過各種自動化階段，從建置和測試階段到生產環境的部署。

儘管傳統觀點建議您將業內人員安排在營運程序中最困難的部分，但是出於這個原因，我們建議您能自動化最困難的程序。

實作步驟

可以依照下列步驟自動化部署以移除手動作業：

- 設定程式碼儲存庫以安全地儲存您的程式碼：使用以 Git 等常用技術為基礎的託管原始程式碼管理系統，來儲存原始程式碼和基礎設施即程式碼 (IaC) 組態。
- 設定持續整合服務以編譯原始程式碼、執行測試及建立部署成品：若要針對此目的設定建置專案，請參閱 [Getting started with AWS CodeBuild using the console](#)。
- 設定可自動化應用程式部署並處理應用程式更新複雜性的部署服務，而不需依賴容易出錯的手動部署：[AWS CodeDeploy](#) 可將軟體部署自動化至各種運算服務，例如 Amazon EC2、[AWS Fargate](#)、[AWS Lambda](#) 和內部部署伺服器。若要設定這些步驟，請參閱 [Getting started with CodeDeploy](#)。

- 設定持續交付服務，自動化您的發行管道，以便實現更快且更可靠的應用程式和基礎設施更新：請考慮使用 [AWS CodePipeline](#) 來協助您自動化發行管道。如需詳細資訊，請參閱 [CodePipeline tutorials](#)。

資源

相關的最佳實務：

- [OPS05-BP04 使用建置和部署管理系統](#)
- [OPS05-BP10 完全自動化整合和部署](#)
- [OPS06-BP02 測試部署](#)
- [OPS06-BP04 自動化測試和復原](#)

相關文件：

- [使用 AWS CodePipeline 連續交付巢狀 AWS CloudFormation 堆疊](#)
- [APN 合作夥伴：可以幫助您建立自動化部署解決方案的合作夥伴](#)
- [AWS Marketplace：可用於自動化部署的產品](#)
- [使用 Webhook 自動化聊天訊息。](#)
- [Amazon 建置者資料中心：確保部署期間的回復安全](#)
- [Amazon 建置者資料中心：使用持續交付加快腳步](#)
- [什麼是 AWS CodePipeline？](#)
- [什麼是 CodeDeploy？](#)
- [AWS Systems Manager Patch Manager](#)
- [什麼是 Amazon SES？](#)
- [什麼是 Amazon Simple Notification Service？](#)

相關影片：

- [2019 年 AWS 高峰會：AWS 上的 CI/CD](#)

故障管理

問題

- [REL 9. 如何備份資料？](#)
- [REL 10. 如何使用故障隔離來保護工作負載？](#)
- [REL 11. 如何設計工作負載以承受元件失敗？](#)
- [REL 12. 如何測試可靠性？](#)
- [REL 13. 如何規劃災難復原 \(DR\)？](#)

REL 9. 如何備份資料？

備份資料、應用程式和組態以滿足您對復原時間點目標 (RTO) 和復原點目標 (RPO) 的要求。

最佳實務

- [REL09-BP01 識別並備份所有需要備份的資料，或從來源複製資料](#)
- [REL09-BP02 保護和加密備份](#)
- [REL09-BP03 自動執行資料備份](#)
- [REL09-BP04 定期執行資料復原以驗證備份的完整性和程序](#)

REL09-BP01 識別並備份所有需要備份的資料，或從來源複製資料

了解和使用工作負載所使用的資料服務和資源的備份功能。大部分服務都會提供備份工作負載資料的功能。

預期成果：已根據重要性識別並分類資料來源。然後，根據 RPO 建立資料復原的策略。此策略涉及備份這些資料來源，或具有從其他來源重現資料的能力。若遺失資料，實作的策略可讓您在定義的 RPO 和 RTO 內復原或重現資料。

雲端成熟度階段：基礎

常見的反模式：

- 未注意工作負載的所有資料來源及其關鍵性。
- 未備份關鍵資料來源。
- 只備份某些資料來源，而未使用關鍵性做為準則。
- 沒有已定義的 RPO，或備份頻率無法符合 RPO。
- 未評估是否需要備份，或是否可從其他來源重現資料。

建立此最佳實務的優勢：確定需要備份的位置並實作建立備份的機制，或是能夠從外部來源重製資料，這可提升在中斷時還原及復原資料的能力。

未建立此最佳實務時的曝險等級：高

實作指引

所有 AWS 資料存放區都會提供備份功能。Amazon RDS 和 Amazon DynamoDB 等服務會額外地支援啟用時間點復原 (PITR) 的自動備份，這可讓您將備份還原到目前時間之前最多五分鐘或更短的任何時間。許多 AWS 服務都提供將備份複製到另一個 AWS 區域的功能。AWS Backup 是一種工具，可讓您跨 AWS 服務集中化和自動化資料保護。[AWS Elastic Disaster Recovery](#) 可讓您複製完整的伺服器工作負載，並維護內部部署、跨可用區域或跨區域的持續資料保護，並以秒為單位測量復原點目標 (RPO)。

Amazon S3 可以用作自行管理和 AWS 管理的資料來源的備份目的地。Amazon EBS、Amazon RDS 和 Amazon DynamoDB 等 AWS 服務具有建立備份的內建功能。也可以使用第三方備份軟體。

可使用 [AWS Storage Gateway](#) 或 [AWS DataSync](#)，將內部部署資料備份至 AWS 雲端。Amazon S3 儲存貯體可用來在 AWS 中存放此資料。Amazon S3 提供多種儲存層級，例如 [Amazon Glacier](#) 或 [Amazon Glacier Deep Archive](#)，以降低資料儲存成本。

您能夠從其他資源重現資料來符合資料復原需求。例如，如果主節點遺失，您可以使用 [Amazon ElastiCache 複本節點](#) 或 [Amazon RDS 讀取複本](#) 來重現資料。如果這類來源可以用於滿足您的 [復原點目標 \(RPO\)](#) 和 [復原時間點目標 \(RTO\)](#)，則可能不需要備份。另一個範例是，如果使用 Amazon EMR，則可能不需要備份 HDFS 資料存放區，只要您可以 [從 Amazon S3 將資料複製到 Amazon EMR](#) 即可。

選取備份策略時，請考慮復原資料所需的時間。復原資料所需的時間取決於備份的類型 (若有備份策略)，或資料重現機制的複雜性。此時間應該落在工作負載的 RTO 內。

實作步驟

1. 識別工作負載的所有資料來源。資料可以存儲在許多資源上，例如 [資料庫](#)、[磁碟區](#)、[檔案系統](#)、[日誌記錄系統](#) 和 [物件儲存](#)。請參閱資源部分，尋找儲存資料之不同 AWS 服務的相關文件，以及這些服務提供的備份功能。
2. 根據重要性對資料來源進行分類。不同的資料集對工作負載具有不同的關鍵性等級，因此對彈性具有不同的要求。例如，有些資料可能至關重要，且需要接近零的 RPO，而其他資料可能不太重要，且可以容忍更高的 RPO 和一些資料遺失。同樣地，不同的資料集也可能具有不同的 RTO 要求。
3. 使用 AWS 或第三方服務來建立資料的備份。[AWS Backup](#) 是一項託管服務，它允許在 AWS 上建立各種資料來源的備份。[AWS Elastic Disaster Recovery](#) 可處理 AWS 區域的自動次秒級資料複

寫。大部分 AWS 服務也具有建立備份的原生功能。AWS Marketplace 具有許多也提供這些功能的解決方案。有關如何從各種 AWS 服務建立資料備份的資訊，請參閱下面列出的資源。

4. 對於未備份的資料，請建立資料複製機制。您可能基於各種原因選擇不備份可從其他來源重現的資料。可能有一種情況，即在需要時從來源重現資料比建立備份更便宜，因為可能有與儲存備份相關聯的成本。另一個範例是從備份中還原比從來源重現資料需要更長的時間，因而導致 RTO 中出現缺口。在這類情況下，考慮取捨並建立一個妥善定義的流程，其中指出在需要資料復原時如何從這些來源重現資料。例如，如果您已將資料從 Amazon S3 載入資料倉儲 (如 Amazon Redshift) 或 MapReduce 叢集 (如 Amazon EMR) 來分析該資料，這可以成為從其他來源複製的資料之範例。只要這些分析的結果存放在某處或可複製，您就不會因為資料倉儲或 MapReduce 叢集故障而遺失資料。其他可從來源複製的範例包括快取 (如 Amazon ElastiCache) 或 RDS 的僅供讀取複本。
5. 建立備份資料的節奏。建立資料來源的備份是一種定期流程，而且頻率應取決於 RPO。

實作計劃的工作量：中

資源

相關的最佳實務：

[REL13-BP01 定義停機和資料遺失的復原目標](#)

[REL13-BP02 使用定義的復原策略來滿足復原目標](#)

相關文件：

- [什麼是 AWS Backup ?](#)
- [什麼是 AWS DataSync ?](#)
- [什麼是 Volume Gateway ?](#)
- [APN 合作夥伴：可以幫助備份的合作夥伴](#)
- [AWS Marketplace：可用於備份的產品](#)
- [Amazon EBS 快照](#)
- [備份 Amazon EFS](#)
- [備份 Amazon FSx for Windows File Server](#)
- [ElastiCache for Redis 的備份與還原](#)
- [在 Neptune 中建立資料庫叢集快照](#)
- [建立資料庫快照](#)
- [建立依照排程觸發的 EventBridge 規則](#)

- [使用 Amazon S3 進行跨區域複寫](#)
- [EFS 到 EFS AWS Backup](#)
- [將日誌資料匯出至 Amazon S3](#)
- [物件生命週期管理](#)
- [DynamoDB 的隨需備份與還原](#)
- [DynamoDB 的時間點復原](#)
- [使用 Amazon OpenSearch Service 索引快照](#)
- [什麼是 AWS Elastic Disaster Recovery？](#)

相關影片：

- [AWS re:Invent 2021 - 使用 AWS 進行備份、災難復原和勒索軟體保護](#)
- [AWS Backup 示範：跨帳戶和跨區域備份](#)
- [AWS re:Invent 2019：深入研究 AWS Backup, ft. Rackspace \(STG341\)](#)

REL09-BP02 保護和加密備份

使用身分驗證和授權控制並偵測對備份的存取。使用加密來防止並檢測是否危及備份的資料完整性。

實作安全控制，以防止未經授權存取備份資料。加密備份，以保護資料的機密性和完整性。

常見的反模式：

- 讓備份和還原自動化的存取權與資料的存取權相同。
- 不加密您的備份。
- 不實作不變性，以防止刪除或竄改。
- 針對生產和備份系統使用相同的安全網域。
- 不透過定期測試驗證備份完整性。

建立此最佳實務的優勢：

- 保護您的備份可防止資料遭到竄改，加密資料可防止意外暴露時存取該資料。
- 增強保護，防止遭到鎖定備份基礎結構為目標的勒索軟體和其他網路威脅攻擊。
- 透過經驗證的復原程序，縮短發生網路事件後的復原時間。

- 加強安全事件期間的業務持續性功能。

未建立此最佳實務時的曝險等級：高

實作指引

使用身分驗證和授權控制並偵測對備份的存取，例如 AWS Identity and Access Management (IAM)。使用加密來防止並檢測是否危及備份的資料完整性。

Amazon S3 支援多種靜態資料的加密方法。使用伺服器端加密時，Amazon S3 會以未加密資料的形式接受物件，然後在儲存這些物件之前將其加密。使用用戶端加密時，您的工作負載應用程式需負責加密資料，然後將資料傳送至 Amazon S3。這兩種方法都可讓您使用 AWS Key Management Service (AWS KMS) 來建立和存放資料金鑰，或者您也可以提供自己的金鑰，之後由您對其負責。使用 AWS KMS 時，您可以透過 IAM 設定政策，設定誰可以和誰無法存取您的資料金鑰和解密資料。

對於 Amazon RDS，如果您已選擇加密資料庫，則備份也會加密。DynamoDB 備份一律會加密。使用 AWS Elastic Disaster Recovery 時，所有傳輸中的資料和靜態資料都會加密。使用 Elastic Disaster Recovery，靜態資料可以使用預設 Amazon EBS 加密磁碟區加密金鑰或自訂客戶受管金鑰進行加密。

網路恢復能力考量事項

若要增強備份安全以防範網路威脅，除了加密之外，請考慮實作下列這些額外的控制：

- 使用 AWS Backup 文件庫鎖定或 Amazon S3 物件鎖定來實作不變性，以防止備份資料在保留期間遭到更改或刪除，以防範勒索軟體和惡意刪除。
- 針對關鍵系統使用 AWS Backup 邏輯氣隙隔離保存庫在正式執行環境和備份環境之間建立邏輯隔離，讓兩種環境彼此分隔，有助於防止其同時遭到入侵。
- 使用 AWS Backup 還原測試定期驗證備份完整性，以確認備份未損毀，且可在網路事件後成功還原。
- 針對關鍵復原操作使用 AWS Backup 多方核准實作多方核准，藉由要求多個指定核准者授權，以防止未經授權或惡意復原嘗試。

實作步驟

1. 在每個資料存放區使用加密。如果來源資料已加密，則備份也會加密。
 - [在 Amazon RDS 中使用加密](#)。您可以在建立 RDS 執行個體時，使用 AWS Key Management Service 設定靜態加密。
 - [在 Amazon EBS 磁碟區中使用加密](#)。您可以在建立磁碟區時設定預設加密或指定唯一金鑰。

- 使用必要的 [Amazon DynamoDB 加密](#)。DynamoDB 會加密所有靜態資料。您可以使用 AWS 自有的 AWS KMS 金鑰或 AWS 受管 KMS 金鑰，指定帳戶中儲存的金鑰。
 - [加密存放在 Amazon EFS 中的資料](#)。在建立檔案系統時設定加密。
 - 在來源和目的地區域設定加密。您可以使用 KMS 中存放的金鑰來設定 Amazon S3 中的靜態加密，但金鑰是區域限定的。您可以在設定複寫時指定目的地金鑰。
 - 選擇使用預設還是自訂 [Amazon EBS 加密進行彈性災難復原](#)。此選項會在模擬區域子網路磁碟和複寫磁碟上加密您的複寫靜態資料。
2. 實作存取備份的最低權限。遵循最佳實務，以根據[安全最佳實務](#)限制對備份、快照和複本的存取。
 3. 設定關鍵備份的不變性。針對關鍵資料實作 AWS Backup 文件庫鎖定或 S3 物件鎖定，以防止在指定的保留期間遭到刪除或修改。如需實作的詳細資訊，請參閱 [AWS Backup 文件庫鎖定](#)。
 4. 建立備份環境的邏輯分隔。針對要求增強保護防範網路威脅的關鍵系統，實作 AWS Backup 邏輯氣隙隔離保存庫。如需實作指引，請參閱[使用 AWS Backup 邏輯氣隙隔離保存庫建置網路恢復能力](#)。
 5. 實作備份驗證程序。設定 AWS Backup 還原測試以定期驗證備份未損毀，且可在網路事件後成功還原。如需詳細資訊，請參閱[使用 AWS Backup 還原測試驗證復原整備度](#)。
 6. 針對敏感復原操作設定多方核准。針對關鍵系統實作 AWS Backup 多方核准，以要求取得多個指定核准者授權，才能繼續復原。如需實作詳細資訊，請參閱[透過 AWS Backup 支援多方核准來加強復原能力](#)。

資源

相關文件：

- [AWS Marketplace：可用於備份的產品](#)
- [Amazon EBS 加密](#)
- [Amazon S3：使用加密保護資料](#)
- [CRR 其餘組態：複寫使用 AWS KMS 中儲存的金鑰，透過伺服器端加密 \(SSE\) 所建立的物件](#)
- [DynamoDB 靜態加密](#)
- [加密 Amazon RDS 資源](#)
- [在 Amazon EFS 中加密資料和中繼資料](#)
- [中的備份加密AWS](#)
- [管理加密資料表](#)
- [安全支柱 – AWS Well Architected Framework](#)
- [什麼是 AWS Elastic Disaster Recovery？](#)

- [FSISEC11：如何防範勒索軟體？](#)
- [在 AWS 上使用 NIST 網路安全架構進行勒索軟體風險管理](#)
- [使用 AWS Backup 邏輯氣隙隔離保存庫建置網路恢復能力](#)
- [使用 AWS Backup 還原測試驗證復原整備度](#)
- [透過 AWS Backup 支援多方核准來加強復原能力](#)

相關範例：

- [針對 Amazon S3 實作雙向跨區域複寫 \(CRR\)](#)

REL09-BP03 自動執行資料備份

設定備份以根據復原點目標 (RPO) 所通知的定期排程或資料集中的變更自動執行。資料遺失要求低的關鍵資料集需要經常自動備份，而可以接受一些遺失的不太重要資料可以較不頻繁地備份。

預期成果：一種自動化過程，可以按既定的節奏建立資料來源的備份。

常見的反模式：

- 手動執行備份。
- 使用具有備份功能的資源，但不包含您的自動化中的備份。

建立此最佳實務的優勢：自動化備份會驗證是否根據您的 RPO 定期進行備份，並在未進行備份時提醒您。

未建立此最佳實務時的曝險等級：中

實作指引

AWS Backup 可用於建立各種 AWS 資料來源的自動資料備份。Amazon RDS 執行個體幾乎可以持續每五分鐘備份一次，而且 Amazon S3 物件幾乎可以持續每十五分鐘備份一次，同時將時間點復原 (PITR) 提供至備份歷史記錄內的特定時間點。針對其他 AWS 資料來源，例如 Amazon EBS 磁碟區、Amazon DynamoDB 資料表或 Amazon FSx 檔案系統，AWS Backup 可以頻繁地每小時執行自動備份。這些服務也提供原生備份功能。AWS 服務提供具有時間點復原功能的自動備份，包括 [Amazon DynamoDB](#)、[Amazon RDS](#) 和 [Amazon Keyspaces \(適用於 Apache Cassandra\)](#) — 這些服務可以還原到備份歷史記錄中的特定時間點。大部分其他 AWS 資料儲存服務都會提供定期備份排程的能力，頻率為每小時備份一次。

Amazon RDS 和 Amazon DynamoDB 提供具有時間點復原功能的連續備份。Amazon S3 版本控制一旦開啟，便會自動執行。[Amazon Data Lifecycle Manager](#) 可用來自動化建立、複製和刪除 Amazon EBS 快照。其也可以自動建立、複製、棄用和取消註冊 Amazon EBS 支援的 Amazon Machine Image (AMI) 及其基礎 Amazon EBS 快照。

AWS Elastic Disaster Recovery 提供從來源環境 (內部部署或 AWS) 到目標復原區域的持續區塊層級複寫。時間點 Amazon EBS 快照會由服務自動建立及管理。

為了集中檢視備份自動化和歷史記錄，AWS Backup 提供全受管的、基於政策的備份解決方案。它使用 AWS Storage Gateway 在雲端和內部部署中跨多個 AWS 服務，自動集中進行資料備份。

除版本控制之外，Amazon S3 還具有複寫功能。整個 S3 儲存貯體可自動複寫至相同或不同 AWS 區域中的另一個儲存貯體。

實作步驟

1. 識別目前正在手動備份的資料來源。如需詳細資訊，請參閱[REL09-BP01 識別並備份所有需要備份的資料，或從來源複製資料](#)。
2. 確定工作負載的 RPO。如需詳細資訊，請參閱[REL13-BP01 定義停機和資料遺失的復原目標](#)。
3. 使用自動備份解決方案或受管服務。AWS Backup 是一項全受管的服務，可輕鬆在[雲端和內部部署環境內集中管理並自動化各 AWS 服務間的資料保護](#)。使用 AWS Backup 中的備份計劃建立規則，定義要備份的資源，以及應以何種頻率建立這些備份。此頻率應由步驟 2 中建立的 RPO 通知。如需有關如何使用 AWS Backup 建立自動化備份的實作指引，請參閱 [Testing Backup and Restore of Data](#)。大多數存放資料的 AWS 服務都會提供原生備份功能。例如，可以利用 RDS 搭配時間點復原 (PITR) 進行自動備份。
4. 對於自動備份解決方案或受管服務不支援的資料來源 (例如內部部署資料來源或訊息佇列)，請考慮使用受信任的第三方解決方案來建立自動備份。或者，您可以使用 AWS CLI 或 SDK 建立自動化來執行此動作。您可以使用 AWS Lambda Functions 或 AWS Step Functions，定義涉及建立資料備份的邏輯，並使用 Amazon EventBridge，以根據 RPO 的頻率調用它。

實作計劃的工作量：低

資源

相關文件：

- [APN 合作夥伴：可以幫助備份的合作夥伴](#)
- [AWS Marketplace：可用於備份的產品](#)

- [建立依照排程觸發的 EventBridge 規則](#)
- [什麼是 AWS Backup ?](#)
- [什麼是 AWS Step Functions ?](#)
- [什麼是 AWS Elastic Disaster Recovery ?](#)

相關影片：

- [AWS re:Invent 2019：深入研究 AWS Backup, ft. Rackspace \(STG341\)](#)

REL09-BP04 定期執行資料復原以驗證備份的完整性和程序

透過執行復原測試，驗證您的備份程序實作是否符合復原時間點目標 (RTO) 和復原點目標 (RPO)。

預期成果：備份中的資料會使用良好定義的機制定期復原，以驗證是否可以在工作負載的既定復原時間點目標 (RTO) 內進行復原。驗證從備份中還原是否會導致資源包含原始資料 (而其中沒有任何損壞或無法存取)，但在復原點目標 (RPO) 內發生資料遺失。

常見的反模式：

- 還原備份，但不查詢或擷取任何資料，以檢查還原可用。
- 假設備份存在。
- 假設系統的備份可以完全運作，而且可以從中復原資料。
- 假設從備份中還原或復原資料的時間落在工作負載的 RTO 內。
- 假設備份上包含的資料落在工作負載的 RPO 內
- 在不使用執行手冊的情況下，或在建立的自動化程序外部，視需要還原。

建立此最佳實務的優勢：測試備份復原可驗證資料是否可以在需要時還原，而不必擔心資料可能遺失或損毀，可在工作負載的 RTO 內進行還原和復原，而且任何資料遺失都屬於工作負載的 RPO 範圍。

未建立此最佳實務時的曝險等級：中

實作指引

測試備份和還原功能可以提高能夠在中斷期間執行這些動作的信心。定期將備份還原至新位置，並執行測試以驗證資料的完整性。某些應該執行的常用測試會檢查所有資料是否可用、未損毀、可存取，且任何資料遺失落在工作負載的 RPO 內。此類測試也可以協助確定，復原機制是否足夠快到適應工作負載的 RTO。

使用 AWS 時，您可以建立一個測試環境，還原備份來評估 RTO 和 RPO 功能，並針對資料內容和完整性執行測試。

此外，Amazon RDS 和 Amazon DynamoDB 允許時間點復原 (PITR)。使用持續備份時，您可以將資料集還原到指定日期和時間當時的狀態。

所有資料是否可用、未損壞、可存取，並且任何資料遺失都落在工作負載的 RPO 內。此類測試也可以協助確定，復原機制是否足夠快到適應工作負載的 RTO。

AWS Elastic Disaster Recovery 提供 Amazon EBS 磁碟區的持續時間點復原快照。複寫來源伺服器時，會根據設定的政策，隨著時間的推移記錄時間點狀態。彈性災難復原可透過啟動執行個體進行測試和演練，而無須重新導向流量，從而協助您驗證這些快照的完整性。

實作步驟

1. 識別目前正在備份的資料來源，以及這些備份的儲存位置。如需實作指引，請參閱 [REL09-BP01 識別並備份所有需要備份的資料，或從來源複製資料](#)。
2. 針對每個資料來源建立資料驗證條件。不同類型的資料將具有不同的屬性，可能需要不同的驗證機制。在您自信可於生產環境中使用此資料之前，請考慮如何驗證它。一些驗證資料的常用方法是使用資料和備份屬性，例如資料類型、格式、檢查總和、大小，或這些屬性與自訂驗證邏輯的組合。例如，這可能是建立備份時所還原資源與資料來源之間的檢查總和值比較。
3. 建立 RTO 和 RPO，以根據資料重要性還原資料。如需實作指引，請參閱 [REL13-BP01 定義停機和資料遺失的復原目標](#)。
4. 評估您的復原能力。審核您的備份和還原策略，以了解它是否可以符合您的 RTO 和 RPO，並視需要調整策略。使用 [AWS Resilience Hub](#)，可以執行工作負載的評估。此評定會針對彈性政策評估您的應用程式組態，並報告您的 RTO 和 RPO 目標是否可以實現。
5. 使用生產中用於資料還原的當前已建立的流程進行測試還原。這些程序取決於原始資料來源的備份方式、備份本身的格式和儲存位置，或是否已從其他源重現資料。例如，如果您使用的是諸如 [AWS Backup](#) 等受管服務，這可能就像將備份還原到新資源一樣簡單。如果使用 AWS Elastic Disaster Recovery，則可以 [啟動復原演練](#)。
6. 根據先前建立的資料驗證條件，從已還原的資源中驗證資料復原。還原和復原的資料是否包含備份時最新的記錄/項目？此資料是否落在工作負載的 RPO 內？
7. 測量還原和復原所需的時間，並將其與已確定的 RTO 進行比較。此程序是否落在工作負載的 RTO 內？例如，比較從還原程序開始到復原驗證完成的時間戳記，以計算此程序需要多長時間。所有 AWS API 呼叫都有時間戳記，而且此資訊在 [AWS CloudTrail](#) 中可用。儘管此資訊可以提供有關還原程序何時開始的詳細資訊，但驗證完成時的結束時間戳記應由驗證邏輯記錄。如果使用自動流程，則可以使用 [Amazon DynamoDB](#) 之類的服務來存放此資訊。此外，許多 AWS 服務會提供事件

歷史記錄，其中提供特定動作何時發生的時間戳記資訊。在 AWS Backup 中，備份和還原動作稱為工作，而這些工作包含時間戳記資訊做為其中繼資料的一部分，可用來測量還原和復原所需的時間。

8. 如果資料驗證失敗，或如果還原和復原所需的時間超過工作負載的既定 RTO，請通知利益相關者。實作自動化以執行此操作時，[例如在此實驗室中](#)，可以使用 Amazon Simple Notification Service (Amazon SNS) 等服務向利益相關者傳送推送通知，例如電子郵件或 SMS。[這些訊息也可以發佈到傳訊應用程式 \(例如 Amazon Chime、Slack 或 Microsoft Teams\)](#)，或用於[使用 AWS Systems Manager OpsCenter 將任務建立為 OpsItems](#)。
9. 將此流程自動化以定期執行。例如，諸如 AWS Lambda 或 AWS Step Functions 中的狀態機器等服務可以用來將還原和復原程序自動化，而且 Amazon EventBridge 可以用來定期調用此自動化工作流程，如下面架構圖所示。了解如何[使用 AWS Backup 自動化資料復原驗證](#)。此外，[此 Well-Architected 實驗室](#)為這裡的幾個步驟提供了一種自動化方法的實際操作體驗。

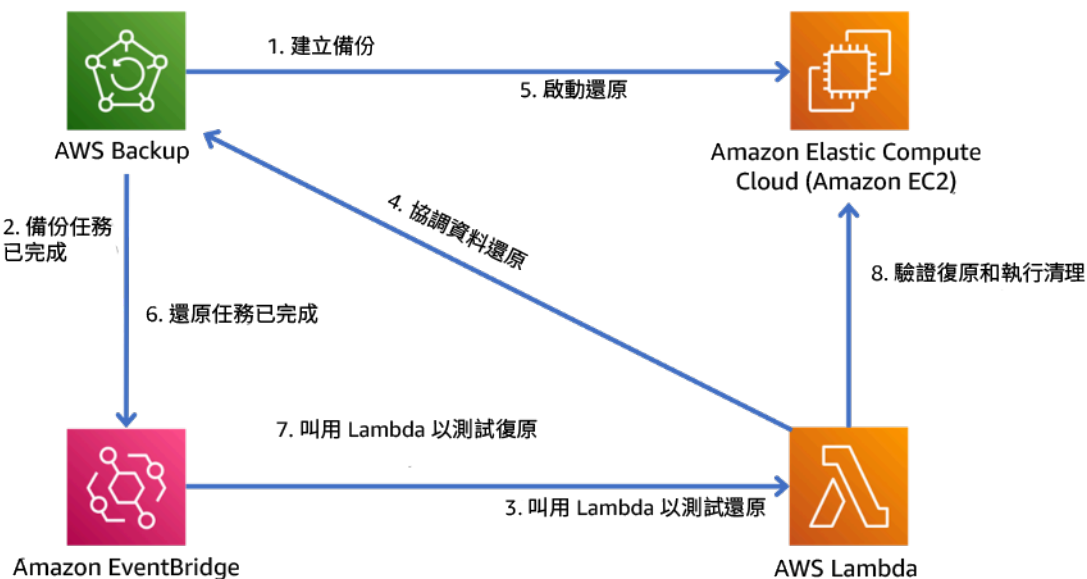


圖 9. 自動備份和還原流程

實施計劃的工作量：中到高，取決於驗證條件的複雜性。定。

資源

相關文件：

- [使用 AWS Backup 自動化資料復原驗證](#)
- [APN 合作夥伴：可以幫助備份的合作夥伴](#)
- [AWS Marketplace：可用於備份的產品](#)

- [建立依照排程觸發的 EventBridge 規則](#)
- [DynamoDB 的隨需備份與還原](#)
- [什麼是 AWS Backup ?](#)
- [什麼是 AWS Step Functions ?](#)
- [什麼是 AWS Elastic Disaster Recovery](#)
- [AWS Elastic Disaster Recovery](#)

REL 10. 如何使用故障隔離來保護工作負載？

故障隔離會將元件或系統故障的影響侷限在定義的界限內。藉由適當的隔離，界限外部的元件就不會受到故障影響。只要能跨多個故障隔離界限執行工作負載，就可更靈活地應對故障。

最佳實務

- [REL10-BP01 將工作負載部署至多個位置](#)
- [REL10-BP02 針對限制在單一位置的元件將復原自動化](#)
- [REL10-BP03 使用隔板架構限制影響範圍](#)

REL10-BP01 將工作負載部署至多個位置

跨多個可用區域或視需要跨 AWS 區域，分配工作負載資料和資源。

AWS 中服務設計的基本原則是避免單一故障點，包括基礎實體基礎設施。AWS 在全球多個地理位置 (稱為 [區域](#)) 提供雲端運算資源和服務。每個區域實際上和邏輯上都各自獨立，並且由三個以上的 [可用區域 \(AZ\)](#) 所組成。可用區域在地理位置上彼此接近，但實際上是分開且隔離的。若您將工作負載分配到不同的可用區域和區域，就可降低發生火災、洪水、天災、地震和人為錯誤等威脅的風險。

建立位置策略來提供適合您工作負載的高可用性。

預期成果：實際執行工作負載會分散到多個可用區域 (AZ) 或區域，以實現容錯能力和高可用性。

常見的反模式：

- 您的實際行工作負載只存在單一可用區域中。
- 當多可用區域架構可滿足業務需求時，您卻實作多區域架構。
- 您的部署或資料變得不同步，進而導致組態偏離或資料複寫不足。

- 您未考量在這些元件之間的彈性和多位置需求不同的情況下，應用程式元件之間的相依性。

建立此最佳實務的優勢：

- 您的工作負載對於影響 AZ 或整個區域的事件較有彈性，例如電力或環境控制故障、天災、上游服務故障或網路問題。
- 您可以存取更廣泛的 Amazon EC2 執行個體庫存，並降低啟動特定 EC2 執行個體類型時發生 `InsufficientCapacityExceptions` (ICE) 的可能性。

未建立此最佳實務時的曝險等級：高

實作指引

至少在區域中的兩個可用區域 (AZ) 中部署和操作所有實際執行工作負載。

使用多個可用區域

可用區域是託管資源的位置，這些位置實際上彼此分隔，以避免因火災、洪水和龍捲風等風險而導致相互關聯失敗。每個可用區域都有單獨的實體基礎設施，包括公用電源連接、備用電源、機械服務，以及網路連線。此配置可將上述任一元件中的故障限制在只有受影響的可用區域。例如，若 AZ 範圍的事件使得受影響的可用區域中的 EC2 執行個體無法使用，您在其他可用區域中的執行個體仍然可用。

雖然實際上彼此分隔，但相同 AWS 區域中的可用區域彼此接近，因此能提供高輸送量、低延遲 (僅幾毫秒) 聯網。您可以在可用區域之間同步複寫大多數工作負載的資料，而不會顯著影響使用者體驗。這表示，您可以在主動/主動或主動/待命組態中，與區域中使用可用區域。

所有與工作負載相關聯的運算都應分散到多個可用區域。這包括 [Amazon EC2](#) 執行個體、[AWS Fargate](#) 任務和 VPC 連接的 [AWS Lambda](#) 函數。包括 [EC2 Auto Scaling](#)、[Amazon Elastic Container Service \(ECS\)](#) 和 [Amazon Elastic Kubernetes Service \(EKS\)](#) 在內的 AWS 運算服務，為您提供了在不同可用區域之間啟動和管理運算的方法。將它們設定為在不同的可用區域中視需要自動取代運算，以維持可用性。若要將流量引導至可用的可用區域，請在運算前放置負載平衡器，例如 Application Load Balancer 或 Network Load Balancer。AWS 負載平衡器可以在可用區域受損時，將流量重新路由至可用的執行個體。

您也應該複寫工作負載的資料，並且在多個可用區域中提供。有些 AWS 受管資料服務預設會在多個可用區域中複寫資料，例如 [Amazon S3](#)、[Amazon Elastic File Service \(EFS\)](#)、[Amazon Aurora](#)、[Amazon DynamoDB](#)、[Amazon Simple Queue Service \(SQS\)](#) 和 [Amazon Kinesis Data Streams](#)，這些服務面對可用區域受損的情形時相當可靠。使用其他 AWS 受管資料服務，例如

[Amazon Relational Database Service \(RDS\)](#)、[Amazon Redshift](#) 和 [Amazon ElastiCache](#) 時，您必須啟用多可用區域複寫。啟用後，這些服務會自動偵測可用區域受損、將請求重新導向至可用的可用區域，並在復原後視需要重新複寫資料，而不需客戶介入。詳讀您使用的每一項 AWS 受管資料服務的使用者指南，以了解其多可用區域功能、行為和操作。

如果您使用的是自我管理儲存體，例如 [Amazon Elastic Block Store \(EBS\)](#) 磁碟區或 Amazon EC2 執行個體儲存體，您必須自行管理多可用去複寫。

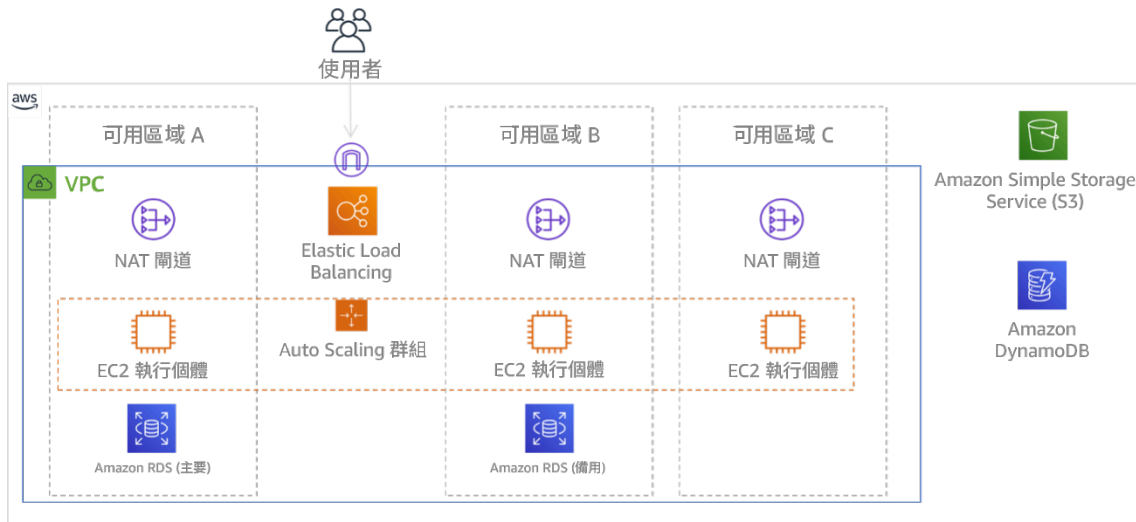


圖 9：跨三個可用區域部署的多層架構。請注意，Amazon S3 和 Amazon DynamoDB 一律自動採用異地同步備份策略。ELB 也會部署至全部三個區域。

使用多個 AWS 區域

如果您的工作負載需要相當大的彈性 (例如，關鍵基礎設施、運作狀態相關應用程式，或是有嚴格的客戶或強制性可用性需求的服務)，則您可能需要額外的可用性，而這並非單一 AWS 區域 能夠提供。在此情況下，您應該至少在兩個 AWS 區域 (假設您的資料落地要求允許) 中部署和操作工作負載。

AWS 區域 位於全球不同的地理區域和多個不同的大陸上。單是 AWS 區域 的實際分隔和隔離距離甚至就比可用區域還大。AWS 服務 (除了少數例外) 利用此設計在不同區域 (也稱為區域服務) 之間完全獨立運作。AWS 區域 服務的故障設計為，不會影響不同區域中的服務。

當您在多個區域中操作工作負載時，應考慮其他需求。由於不同區域中的資源彼此分隔且獨立，因此您必須在每個區域中複寫工作負載的元件。除了運算和資料服務之外，還包括基本的基礎設施，例如 VPC。

注意：當您考慮多區域設計時，務必確認您的工作負載能夠在單一區域中執行。如果您在區域之間建立相依性，讓某一個區域中的元件依賴另一個區域中的服務或元件，則可能會使得失敗的風險升高，並顯著降低可靠性狀態。

為了簡化多區域部署並維持一致性，[AWS CloudFormation StackSets](#) 可將整個 AWS 基礎設施複寫到多個區域。[AWS CloudFormation](#) 還可以偵測組態偏離，並且在區域中的 AWS 資源不同步時通知您。許多 AWS 服務都為重要的工作負載資產提供多區域複寫。例如，[EC2 Image Builder](#) 可以在每次建置之後，將 EC2 機器映像 (AMI) 發布到您使用的每個區域。[Amazon Elastic Container Registry \(ECR\)](#) 可以將容器映像複寫到您選取的區域。

您也必須將資料複寫到您選擇的每一個區域。有許多 AWS 受管資料服務都提供跨區域複寫功能，包括 Amazon S3、Amazon DynamoDB、Amazon RDS、Amazon Aurora、Amazon Redshift、Amazon ElastiCache 和 Amazon EFS。[Amazon DynamoDB 全域資料表](#) 接受寫入任何支援的區域中，並且會將資料複寫到您設定的所有其他區域中。使用其他服務時，您必須指定用於寫入的主要區域，因為其他區域包含唯讀複本。對於工作負載使用的每個 AWS 受管資料服務，請參閱其使用者指南和開發人員指南，以了解其多區域功能和限制。請特別注意寫入的方向、交易功能和限制、複寫執行方式，以及如何監控區域之間的同步作業。

AWS 還提供了將請求流量路由至區域部署的能力，帶來了極大的靈活度。例如，您可以使用 [Amazon Route 53](#) 設定 DNS 記錄，將流量導向距離使用者最近的可用區域。您也可以在主動/待命組態中設定您的 DNS 記錄，在其中指定一個區域做為主要區域，並且僅在主要區域運作狀態不佳時，容錯移轉至另一個區域複本。您可以設定 [Route53 運作狀態檢查](#)，以偵測運作狀態不佳的端點，並執行自動容錯移轉，並視需要額外使用 [Amazon Application Recovery Controller \(ARC\)](#) 來提供高度可用的路由控制，以手動重新路由流量。

即使您選擇不在多個區域中操作以實現高可用性，仍務必將多個區域納入您的災難復原 (DR) 策略中。如果可能，請在次要區域中，於暖待命或指示燈組態中複寫工作負載的基礎設施元件和資料。在此設計中，您會從主要區域複寫基準基礎設施，例如 VPC、Auto Scaling 群組、容器協調器和其他元件，但您會將待命區域中的可變大小的元件 (例如 EC2 執行個體和資料庫複本的數量) 設定為最小可操作大小。您也可以安排從主要區域持續複寫資料至待命區域。如果有事件發生，您可以橫向擴充或擴大待命區域中的資源，然後將其提升為主要區域。

實作步驟

1. 與業務利害關係人和資料落地專家合作，確定哪些 AWS 區域可以用來託管您的資源和資料。
2. 與業務和技術利害關係人合作，評估您的工作負載，並判斷多可用區域方法 (單一 AWS 區域) 是否可滿足其彈性需求，或者需要多區域方法 (若允許多個區域)。使用多個區域可以實現更高的可用性，但可能也會帶來額外的複雜性和成本。評估時，請考慮下列因素：
 - a. 業務目標和客戶需求：在可用區域或區域中發生影響工作負載的事件時，允許多久的停機時間？如 [REL13-BP01 定義停機時間和資料遺失的復原目標](#) 中所述，評估復原點目標。
 - b. 災難復原 (DR) 需求：您希望確實防範哪種可能的災難？請考慮從單一可用區域到整個區域的不同影響範圍下，資料遺失或長期無法使用的可能性。如果您將資料和資源複寫到不同的可用區

域，而某一個可用區域發生持續故障，您可以在另一個可用區域中復原服務。如果您將資料和資源複製到不同的區域，就可以在另一個區域中復原服務。

3. 將運算資源部署到多個可用區域中。

- a. 在 VPC 中，於不同可用區域建立多個子網路。設定每個子網路，使其足以容納處理工作負載所需的資源，即使在發生事件時也一樣。如需詳細資訊，請參閱 [REL02-BP03 確保 IP 子網路配置帳戶具有擴展性和可用性](#)。
- b. 如果您使用 Amazon EC2 執行個體，請使用 [EC2 Auto Scaling](#) 來管理執行個體。建立 Auto Scaling 群組時，請指定您在上一個步驟中選擇的子網路。
- c. 如果您針對 [Amazon ECS](#) 或 [Amazon EKS](#) 使用 AWS Fargate 運算，請在建立 ECS Service、啟動 ECS 任務，或是為 EKS 建立 [Fargate 設定檔](#) 時，選取您在第一個步驟中選擇的子網路。
- d. 如果您使用的是需要在 VPC 中執行的 AWS Lambda 函數，請在建立 Lambda 函數時選取您在第一個步驟中選擇的子網路。對於沒有 VPC 組態的任何函數，AWS Lambda 會自動為您管理可用性。
- e. 將負載平衡器等流量導向器放在運算資源前面。如果啟用跨區域負載平衡，[AWS Application Load Balancer](#) 和 [Network Load Balancer](#) 會偵測 EC2 執行個體和容器等目標何時因可用區域受損而無法連線，並將流量重新路由至運作狀態良好的可用區域中的目標。如果您停用跨區域負載平衡，請使用 Amazon Application Recovery Controller (ARC) 來提供區域轉移功能。如果您使用第三方負載平衡器，或已實作自己的負載平衡器，請在不同的可用區域中為它們設定多個前端。

4. 在多個可用區域中複製工作負載的資料。

- a. 如果您使用 AWS 受管資料服務，例如 Amazon RDS、Amazon ElastiCache 或 Amazon FSx，請詳讀其使用者指南，以了解其資料複製和復原功能。視需要啟用跨可用區複製和容錯移轉。
- b. 如果您使用 AWS 受管儲存服務，例如 Amazon S3、Amazon EFS 和 Amazon FSx，請避免針對需要高耐久性的資料使用單一可用區域或單一區域組態。針對這些服務使用多可用區域組態。詳閱個別服務的使用者指南，判斷多可用區域複製為預設啟用，或是您必須啟用它。
- c. 如果您執行自我管理的資料庫、佇列或其他儲存服務，請根據應用程式的說明或最佳實務安排多可用區域複製。熟悉應用程式的容錯移轉程序。

5. 設定您的 DNS 服務，使其偵測 AZ 受損情形，並將流量重新路由至運作狀態良好的可用區域。Amazon Route 53 與 Elastic Load Balancer 搭配使用時，可以自動執行此操作。Route 53 也可以設定為擁有容錯移轉記錄，這些記錄使用運作狀態檢查來回應查詢，並且只提供運作狀態良好的 IP 位址。對於用於容錯移轉的任何 DNS 記錄，請指定短存留時間 (TTL) 值 (例如 60 秒或更短)，這樣做有助於防止記錄快取阻礙復原 (Route 53 別名記錄會為您提供適當的 TTL)。

使用多個 AWS 區域 時的其他步驟

1. 複寫工作負載在所選區域中使用的所有作業系統 (OS) 和應用程式程式碼。視需要使用 Amazon EC2 Image Builder 等解決方案複寫 EC2 執行個體所使用的 Amazon Machine Image (AMI)。使用 Amazon ECR 跨區域複寫等解決方案複寫存放在登錄檔中的容器映像。為任何用於儲存應用程式資源的 Amazon S3 儲存貯體啟用區域複寫。
2. 將運算資源和組態中繼資料 (例如存放在 AWS Systems Manager 參數存放區的參數) 部署到多個區域。使用先前步驟所述的相同程序，但是複寫您用於工作負載的每個區域的組態。使用基礎設施即程式碼解決方案 (例如 AWS CloudFormation)，在區域之間統一重複產生組態。如果您在災難復原的指示燈組態中使用次要區域，您可以將運算資源的數量減少到最小值以節省成本，並相應地增加復原時間。
3. 將您的資料從主要區域複寫到次要區域。
 - a. Amazon DynamoDB 全域資料表提供可從任何支援的區域寫入的全域資料複本。使用其他 AWS 受管資料服務時 (例如 Amazon RDS、Amazon Aurora 和 Amazon ElastiCache)，您可以指定主要 (讀取/寫入) 區域和複本 (唯讀) 區域。如需區域複寫的詳細資訊，請參閱個別服務的使用者指南和開發人員指南。
 - b. 如果您執行自我管理的資料庫，請根據應用程式的說明或最佳實務安排多區域複寫。熟悉應用程式的容錯移轉程序。
 - c. 如果您的工作負載使用 AWS EventBridge，您可能需要將所選事件從主要區域轉送至次要區域。若要這麼做，請將次要區域中的事件匯流排指定為主要區域中相符事件的目標。
4. 考慮是否要在不同區域使用相同的加密金鑰，以及使用的程度。在安全性和易用性之間取得平衡的典型方法，是針對區域本機資料和身分驗證使用區域範圍金鑰，以及使用全域範圍金鑰來加密在不同區域之間複寫的資料。[AWS Key Management Service\(KMS\)](#) 支援[多區域金鑰](#)，它可安全地分配和保護在區域之間共用的金鑰。
5. 考慮使用 AWS Global Accelerator，透過將流量導向包含運作狀態良好之端點的區域來改善應用程式的可用性。

資源

相關的最佳實務：

- [REL02-BP03 確保 IP 子網路配置帳戶具有擴展性和可用性](#)
- [REL11-BP05 使用靜態穩定性來防止雙模態行為](#)
- [REL13-BP01 定義停機時間和資料遺失的復原目標](#)

相關文件：

- [AWS 全球基礎設施](#)
- [白皮書：AWS 故障隔離界限](#)
- [Amazon EC2 Auto Scaling 的恢復能力](#)
- [Amazon EC2 Auto Scaling：範例：將執行個體發布到不同的可用區域](#)
- [EC2 Image Builder 的運作方式](#)
- [Amazon ECS 如何將任務放置在容器執行個體上 \(包括 Fargate\)](#)
- [AWS Lambda 中的恢復能力](#)
- [Amazon S3：複寫物件概觀](#)
- [Amazon ECR 中的私有映像複寫](#)
- [全域資料表：使用 DynamoDB 進行多區域複寫](#)
- [Amazon ElastiCache for Redis OSS：使用全域資料儲存在 AWS 區域之間複寫](#)
- [Amazon RDS 的恢復能力](#)
- [使用 Amazon Aurora 全球資料庫](#)
- [AWS Global Accelerator 開發人員指南](#)
- [AWS KMS 中的多區域金鑰](#)
- [Amazon Route 53：設定 DNS 備援](#)
- [Amazon Application Recovery Controller \(ARC\) 開發人員指南](#)
- [在 AWS 區域之間傳送和接收 Amazon EventBridge 事件](#)
- [使用 AWS Services 部落格系列建立多區域應用程式](#)
- [AWS 上的災難復原 \(DR\) 架構，第一部分：雲端復原策略](#)
- [AWS 上的災難復原 \(DR\) 架構，第 III 部分：指示燈和暖待命](#)

相關影片：

- [AWS re:Invent 2018：適用於多區域主動-主動應用程式的架構模式](#)
- [AWS re:Invent 2019：AWS 全球網路基礎設施的創新和營運](#)

REL10-BP02 針對限制在單一位置的元件將復原自動化

如果工作負載的元件只能在單一可用區域或內部部署資料中心執行，在定義的復原目標內實作完整重建工作負載的功能。

未建立此最佳實務時的曝險等級：中

實作指引

如果因為技術限制而無法實作將工作負載部署至多個位置的最佳實務，您必須實作彈性的替代路徑。您必須將以下能力自動化：重新建立必要基礎設施、重新部署應用程式，以及針對這些案例重新建立必要資料。

例如，Amazon EMR 會在相同可用區域中啟動指定叢集的所有節點，因為在相同區域執行叢集可以提供更高的資料存取速率，從而能提高任務流程的效能。如果為實現工作負載彈性而需要此元件，您必須要有方法重新部署叢集及其資料。此外，對於 Amazon EMR，您還應以異地同步備份以外的方式佈建冗餘。可以佈建[多個節點](#)。使用 [EMR 檔案系統 \(EMRFS\)](#) 時，EMR 中的資料可存放在 Amazon S3 中，然後可複寫至多個可用區域或 AWS 區域。

同樣地，對於 Amazon Redshift，它預設會將叢集佈建在您所選 AWS 區域內隨機選取的可用區域中。所有叢集節點將佈建在相同的區域中。

針對部署到內部部署資料中心的有狀態的伺服器型工作負載，您可以使用 AWS Elastic Disaster Recovery 在 AWS 中保護您的工作負載。如果已經在 AWS 中託管，則可以使用 Elastic Disaster Recovery 將工作負載保護到備選可用區域或區域。Elastic Disaster Recovery 使用輕量型暫存區的持續區塊層級複寫，以提供內部部署應用程式和雲端應用程式的快速且可靠的復原。

實作步驟

1. 實作自我修復。盡可能使用 Automatic Scaling 來部署執行個體或容器。如果無法使用 Automatic Scaling，請對 EC2 執行個體使用自動復原，或者根據 Amazon EC2 或 ECS 容器生命週期事件實作自我修復自動化。
 - 對於不需要單個執行個體 IP 位址、私有 IP 位址、彈性 IP 位址和執行個體中繼資料的執行個體和容器工作負載，使用 [Amazon EC2 Auto Scaling 群組](#)。
 - 啟動範本使用者資料可用於實現自動自我修復大多數工作負載。
 - 對於需要單個執行個體 IP 位址、私有 IP 位址、彈性 IP 位址和執行個體中繼資料的工作負載，使用 [Amazon EC2 執行個體的自動復原](#)。
 - 在偵測到執行個體失敗時，自動復原會將提醒傳送到 SNS 主題。
 - 在無法使用 Auto Scaling 或 EC2 復原的情況下，使用 [Amazon EC2 執行個體生命週期事件](#) 或 [Amazon ECS 事件](#) 自動執行自我修復。
 - 使用事件來叫用自動化，以根據您所需的過程邏輯來修復您的元件。
 - 使用 [AWS Elastic Disaster Recovery](#) 保護僅限於單一位置的有狀態工作負載。

資源

相關文件：

- [Amazon ECS 事件](#)
- [Amazon EC2 Auto Scaling lifecycle hooks](#)
- [復原您的執行個體。](#)
- [服務自動擴展](#)
- [什麼是 Amazon EC2 Auto Scaling ?](#)
- [AWS Elastic Disaster Recovery](#)

REL10-BP03 使用隔板架構限制影響範圍

實作隔板架構 (也稱為小組型架構) 將工作負載內的失敗效應限制為有限數量的元件。

預期成果：小組型架構使用工作負載的多個獨立執行個體，其中每個執行個體稱為小組。每個小組都是獨立的，不會與其他小組共用狀態，並且處理整體工作負載請求的子集。這會對個別小組和它處理的請求降低失敗的潛在影響，例如不良的軟體更新。如果工作負載使用 10 個小組為 100 個請求提供服務，發生失敗時，整體請求中 90% 不會受到失敗影響。

常見的反模式：

- 允許小組成長，沒有界限。
- 將程式碼更新或部署同時套用到所有小組。
- 在小組之間共用狀態或元件 (路由器層例外)。
- 將複雜商業或路由邏輯新增至路由器層。
- 不將跨小組互動降至最低。

建立此最佳實務的優勢：使用小組型架構時，小組本身包含許多常見的故障類型，從而提供額外的故障隔離。這些故障界限可針對難以控制的失敗類型提供復原能力，例如失敗的程式碼部署或已損毀或叫用特定失敗模式 (也稱為毒藥請求) 的請求。

未建立此最佳實務時的曝險等級：高

實作指引

在船上，隔板可確保船體破口包含在船體的其中一個區段內。在複雜的系統中，通常會複寫這個模式以實現故障隔離。故障隔離界限會在工作負載內將失敗影響限制為有限數量的元件。界限外部的元件不會

受到故障影響。您可以使用多個錯誤隔離界限來限制對工作負載的影響。在 AWS 上，客戶可以使用多個可用區域或區域來提供故障隔離，但是故障隔離的概念也可以延伸為您的工作負載的架構。

整體工作負載是依分割區索引鍵的分割區小組。此索引鍵需要與服務的精細度保持一致，否則服務的工作負載會自然地透過最小的跨小組互動進行細分。分割區索引鍵的範例為客戶 ID、資源 ID 或可在大部分 API 呼叫中輕易存取的其他任何參數。小組路由層會根據分割區索引鍵將請求分散到個別小組，並且對用戶端呈現單一端點。

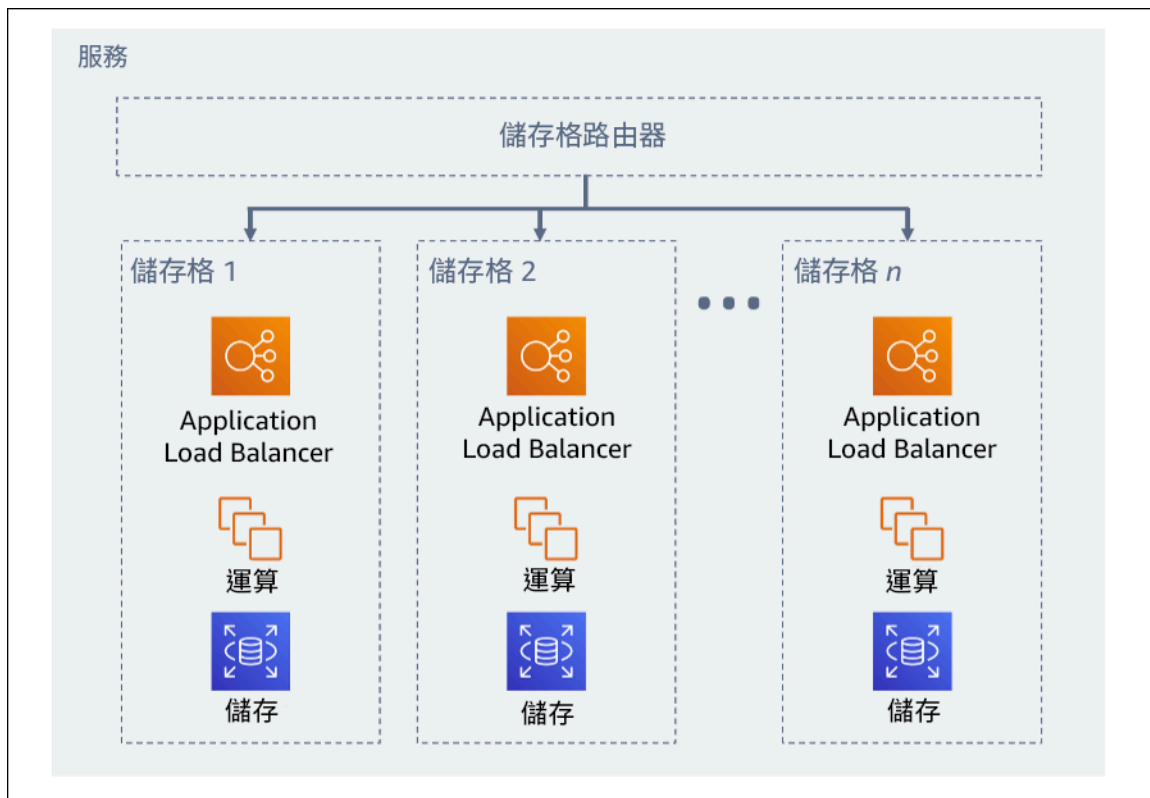


圖 11：小組型架構

實作步驟

設計小組型架構時，要考慮數個設計考量：

1. 分割區索引鍵：選擇分割區索引鍵時，應特別考慮。
 - 應該與服務的精細度保持一致，否則服務的工作負載會自然地透過最小跨小組互動進行細分。例如 customer ID 或 resource ID。
 - 分割區索引鍵必須在所有請求中都可供使用，無論是直接或由其他參數確定性地推斷。
2. 持久性小組映射：上游服務應該僅在其資源的生命週期內與單個小組進行互動。

- 依據工作負載而定，可能需要小組遷移策略，以便從其中一個小組將資料遷移到另一個小組。可能需要小組遷移的可能情境是，如果您的工作負載中特定使用者或資源變得太大並且要求它具備專有小組。
 - 小組不應該在小組之間共用狀態或元件。
 - 因此，應該避免跨小組互動或保持在最低程度，因為這些互動會建立小組之間的相依性，因而消滅故障隔離改善。
3. 路由器層：路由器層是儲存格之間的共用元件，因此無法遵循與儲存格相同的區隔策略。
- 建議路由器層以有效率運算的方式使用分割區對應演算法將請求分發到個別小組，例如結合加密雜湊函數和模組化算術以將分割區索引鍵對應至小組。
 - 若要避免多小組影響，路由層必須保持簡單並且盡可能水平擴展，如此才能避免此層級內的複雜商業邏輯。這樣有增加的優點，隨時都容易了解其預期行為，以獲得徹底的可測試性。正如 Colm MacCárthaigh 在[可靠性、持續工作以及一杯好咖啡](#)中所述，簡單的設計和持續的工作模式會產生可靠的系統並降低抗脆弱性。
4. 儲存格大小：儲存格應具有最大的大小，且不允許超過它。
- 最大大小應該藉由執行徹底測試來識別，直到觸及中斷點並且建立安全的操作邊距。如需如何實作測試實務的詳細資訊，請參閱 [REL07-BP04 Load 測試您的工作負載](#)
 - 整體工作負載應該透過新增額外小組來成長，讓工作負載隨著需求的增加而擴展。
5. 多可用區域或多區域策略：應利用多層彈性來防範不同的故障網域。
- 對於彈性，您應該使用建置防禦層的方法。一層透過使用多個 AZ 建置高度可用的架構來防範更小、更常見的中斷。另一防禦層旨在防範發生罕見事件，例如廣泛的自然災害和區域級中斷。這第二層涉及架構您的應用程式以跨越多個 AWS 區域。針對您的工作負載實作多區域策略有助於其防範影響國家一大片地理區域的廣泛自然災害，或整個區域範圍的技術失敗。請注意，實作多區域架構可能相當複雜，並且通常對於大多數工作負載而言不是必要的。如需詳細資訊，請參閱[REL10-BP01 將工作負載部署至多個位置](#)。
6. 程式碼部署：交錯的程式碼部署策略應優先於同時將程式碼變更部署到所有單元格。
- 這樣可協助將多個小組由於不良部署或人為錯誤的潛在失敗降至最低。如需詳細資訊，請參閱[安全且無需人為干預的自動化部署](#)。

資源

相關的最佳實務：

- [REL07-BP04 Load 測試您的工作負載](#)
- [REL10-BP01 將工作負載部署至多個位置](#)

相關文件：

- [可靠性、持續工作以及一杯好咖啡](#)
- [AWS 和分隔](#)
- [使用隨機切換分區隔離工作負載](#)
- [自動化安全、無人為介入的部署](#)

相關影片：

- [AWS re:Invent 2018：閉環與開放思維：如何取得大小型系統的控制權](#)
- [AWS re:Invent 2018：AWS 如何最大程度地減小故障的影響範圍 \(ARC338\)](#)
- [隨機切換分區：AWS re:Invent 2019：Amazon 建置者資料中心簡介 \(DOP328\)](#)
- [AWS Summit ANZ 2021 - 故障總在情理之中、意料之外 \(Everything fails, all the time\)：專為彈性而設計](#)

REL 11. 如何設計工作負載以承受元件失敗？

須架構具高可用性和低平均復原時間 (MTTR) 需求的工作負載以實現彈性。

最佳實務

- [REL11-BP01 監控工作負載的所有元件以偵測故障](#)
- [REL11-BP02 容錯移轉至運作良好的資源](#)
- [REL11-BP03 將所有分層的修復自動化](#)
- [REL11-BP04 復原期間需使用資料平面，而非控制平面](#)
- [REL11-BP05 使用靜態穩定性來防止雙模態行為](#)
- [REL11-BP06 當事件影響可用性時傳送通知](#)
- [REL11-BP07 建立您的產品架構以符合可用性目標和運行時間服務水準協議 \(SLA\)](#)

REL11-BP01 監控工作負載的所有元件以偵測故障

持續監控工作負載的運作狀態，讓您和自動化系統在發生故障或效能降低時能夠察覺。根據商業價值監控關鍵績效指標 (KPI)。

所有復原和修復機制首先都必須能夠快速偵測問題。應該先偵測技術故障，以便解決問題。不過，可用性取決於工作負載提供商業價值的能力，因此測量此需求的關鍵績效指標 (KPI) 必須成為偵測和修復策略的一部分。

預期成果：工作負載的基本元件會單獨監控，以偵測故障發生的時機和位置並發出警示。

常見的反模式：

- 未設定任何警報，因此會在未發出通知的情況下發生中斷。
- 警示存在，但在此閾值下無法提供足夠的回應時間。
- 收集的指標經常不足以符合復原時間點目標 (RTO)。
- 只主動監控面對客戶的工作負載介面。
- 只收集技術指標，未收集業務功能指標。
- 無測量工作負載使用者體驗的指標。
- 建立了太多監控。

建立此最佳實務的優勢：在各層級內進行適當的監控，可讓您減少偵測時間，進而減少復原時間。

未建立此最佳實務時的曝險等級：高

實作指引

確定將要審核以進行監控的所有工作負載。確定需要監控的所有工作負載元件之後，您現在需要確定監控間隔。根據偵測故障所需的時間而定，監控間隔會直接影響復原的速度。平均偵測時間 (MTTD) 是指從發生故障到開始修復作業經過的時間。服務清單應盡可能廣泛且完整。

監控必須涵蓋應用程式堆疊的所有層級，包括應用程式、平台、基礎設施和網路。

您的監控策略應考慮微小故障的影響。如需微小故障的詳細資訊，請參閱《進階多可用區域彈性模式》白皮書中的 [Gray failures](#)。

實作步驟

- 您的監控間隔取決於復原必須多快完成。您的復原時間取決於所需的復原時間，因此您必須考量此時間和復原時間點目標 (RTO)，藉以決定收集頻率。
- 設定元件和受管服務的詳細監控。
 - 判斷 [EC2 執行個體](#) 和 [Auto Scaling](#) 是否需要詳細監控。詳細監控提供 1 分鐘的間隔指標，預設監控則提供 5 分鐘的間隔指標。

- 判斷 RDS 是否需要[增強型監控](#)。增強型監控使用 RDS 執行個體上的代理程式，以取得不同處理程序或執行緒的實用資訊。
- 判斷 [Lambda](#)、[API Gateway](#)、[Amazon EKS](#)、[Amazon ECS](#) 和所有類型的[負載平衡器](#)的關鍵無伺服器元件的監控需求。
- 確定 [Amazon S3](#)、[Amazon FSx](#)、[Amazon EFS](#) 和 [Amazon EBS](#) 的儲存元件的監控需求。
- 建立[自訂指標](#)以測量業務關鍵績效指標 (KPI)。工作負載會實作重要的業務功能，這些功能應做為 KPI，以利確定間接問題發生的時間。
- 以使用者 Canary 監控使用者的故障體驗。可執行和模擬客戶行為的[綜合交易測試](#) (也稱為 Canary 測試，但請別與 Canary 部署混淆)，是最重要的測試程序之一。針對來自不同遠端位置的工作負載端點持續執行這些測試。
- 建立追蹤使用者體驗的[自訂指標](#)。如果您可以檢測客戶的體驗，則可以判斷消費者體驗何時變差。
- [設定警示](#)以偵測工作負載的任何部分何時未正常運作，並指示何時自動擴展資源。警示會在儀表板上以視覺化方式顯示、透過 Amazon SNS 或電子郵件傳送提醒，以及搭配使用 Auto Scaling 來擴展或縮減工作負載資源。
- 建立[儀表板](#)以視覺化指標。儀表板可以讓您以視覺化方式查看趨勢、極端值和其他潛在問題的指標，或指出您可能想要調查的問題。
- 為您的服務建立[分散式追蹤監控](#)。透過分散式監控，您可以了解應用程式及其基礎服務的執行方式，以確定和疑難排解效能問題與錯誤的根本原因。
- 在單獨的區域和帳戶中建立監控系統 (使用 [CloudWatch](#) 或 [X-Ray](#)) 儀表板和資料收集。
- 透過 [AWS Health](#) 隨時掌握服務降級的相關資訊。[透過 AWS 使用者通知 建立符合用途的 AWS Health 事件通知](#)，以利用電子郵件和聊天管道傳送，並透過 [Amazon EventBridge 以程式設計方式與您的監控和警示工具整合](#)。

資源

相關的最佳實務：

- [可用性定義](#)
- [REL11-BP06 當事件影響可用性時傳送通知](#)

相關文件：

- [Amazon CloudWatch Synthetics 可讓您建立使用者 Canary](#)
- [為執行個體啟用或停用詳細監控](#)

- [Enhanced Monitoring \(增強型監控\)](#)
- [使用 Amazon CloudWatch 監控 Auto Scaling 群組和執行個體](#)
- [發佈自訂指標](#)
- [使用 Amazon CloudWatch 警示](#)
- [使用 CloudWatch 儀表板](#)
- [使用跨區域跨帳戶 CloudWatch 儀表板](#)
- [使用跨區域跨帳戶 X-Ray 追蹤](#)
- [了解可用性](#)

相關影片：

- [減少微小故障](#)

相關範例：

- [一個可觀測性研討會：探索 X-Ray](#)

相關工具：

- [CloudWatch](#)：
- [CloudWatch X-Ray](#)

REL11-BP02 容錯移轉至運作良好的資源

如果發生資源失敗，運作良好的資源應繼續處理請求。對於位置受損 (例如可用區域或 AWS 區域)，請確保您的系統已就位，可容錯移轉至未受影響位置中運作良好的資源。

設計服務時，請將負載分散到各個資源、可用區域或區域。因此，可以透過將流量轉移到剩餘運作狀態良好的資源來減輕個別資源故障或損害的影響。請考慮發生故障時，如何找到服務及其路由。

設計服務時，務必考慮故障復原。在 AWS，我們設計服務以盡可能減少從故障復原的時間並減輕對資料的影響。我們的服務主要使用的資料存放區，會在請求持久儲存於區域內的多個複本中之後，才確認請求。經過建構後，它們會使用以儲存格為基礎的隔離，以及使用可用區域提供的故障隔離。我們在營運程序中廣泛使用自動化。我們還將取代-重啟功能最佳化，以期從中斷快速復原。

允許容錯移轉的模式和設計會隨著各 AWS 平台服務而有所不同。許多 AWS 原生受管服務本身就是多個可用區域 (例如 Lambda 或 API Gateway)。其他 AWS 服務 (例如 EC2 和 EKS) 需要特定的最佳實務設計，以支援在 AZ 的各資源或資料儲存容錯移轉。

監控應設定為確認容錯移轉資源是否正常運作、追蹤資源容錯移轉的進度，以及監控業務程序復原。

預期成果：系統能夠自動或手動使用新資源，以從降級恢復。

常見的反模式：

- 故障計畫不是規劃和設計階段的一部分。
- 未建立 RTO 和 RPO。
- 監控不足，無法偵測出失敗的資源。
- 正確隔離故障網域。
- 未考慮多區域容錯移轉。
- 決定進行容錯移轉時，失敗偵測太過敏感或積極。
- 未測試或驗證容錯移轉設計。
- 進行自動修復自動化，但未通知需要修復。
- 缺少緩衝期，以避免過早容錯恢復。

建立此最佳實務的優勢：您可以建置更具彈性的系統，在發生故障時透過適當降級並快速復原來維持可靠性。

未建立此最佳實務時的曝險等級：高

實作指引

諸如 [Elastic Load Balancing](#) 和 [Amazon EC2 Auto Scaling](#) 等 AWS 服務可協助跨資源和可用區域分配負載。因此，可以透過將流量轉移到剩餘運作狀態良好的資源來緩解個別資源 (例如 EC2 執行個體) 的失敗或可用區域的損害。

對於多區域工作負載，設計會更複雜。例如，跨區域僅供讀取複本可讓您將資料部署到多個 AWS 區域。不過仍需要容錯移轉，才能將僅供讀取複本提升為主要複本，然後將流量指向新端點。Amazon Route 53、[Amazon 應用程式復原控制器 \(ARC\)](#)、Amazon CloudFront 和 AWS Global Accelerator 可協助在 AWS 區域之間路由流量。

諸如 Amazon S3、Lambda、API Gateway、Amazon SQS、Amazon SNS、Amazon SES、Amazon Pinpoint、Amazon ECR、AWS Certificate Manager、EventBridge 或 Amazon DynamoDB 等 AWS

服務由 AWS 自動部署到多個可用區域。如果發生故障，這些 AWS 服務會自動將流量路由到運作良好的位置。資料以冗餘方式存放在多個可用區域中，並且仍然可用。

對於 Amazon RDS、Amazon Aurora、Amazon Redshift、Amazon EKS 或 Amazon ECS 而言，多可用區域是一種組態選項。如果啟動容錯移轉，則 AWS 可將流量導向運作良好的執行個體。此容錯移轉動作可由 AWS 執行，或依客戶要求執行

對於 Amazon EC2 執行個體、Amazon Redshift 或 Amazon ECS 任務或 Amazon EKS Pod，您可以選擇要部署到哪個可用區域。對於某些設計，Elastic Load Balancing 會提供解決方案，以偵測運作狀態不佳區域中的執行個體，並將流量路由至運作良好的區域。Elastic Load Balancing 也可將流量路由至內部部署資料中心內的元件。

對於多區域流量容錯移轉，重新路由可利用 Amazon Route 53、Amazon 應用程式復原控制器、AWS Global Accelerator、適用於 VPC 的 Route 53 私有 DNS 或 CloudFront 來提供定義網際網路網域和指派路由政策 (包括運作狀態檢查) 的方法，以便將流量路由到運作狀態良好的區域。AWS Global Accelerator 提供靜態 IP 位址，做為應用程式端點的固定進入點，然後使用 AWS 全球網路 (而不是網際網路) 路由至您所選 AWS 區域中的端點，以獲得更好的效能和可靠性。

實作步驟

- 為所有適當的應用程式和服務建立容錯移轉設計。隔離每個架構元件，並為每個元件建立符合 RTO 和 RPO 的容錯移轉設計。
- 設定較低的環境 (例如開發或測試)，且其中所有服務都需要有容錯移轉計畫。使用基礎設施即程式碼 (IaC) 來部署解決方案，以確保可重複性。
- 設定復原站台 (例如第二個區域)，以實作和測試容錯移轉設計。如有必要，可以臨時設定測試的資源，以限制額外的成本。
- 判斷哪些容錯移轉計畫是由 AWS 自動執行、哪些可由 DevOps 程序自動執行，以及哪些可能要手動執行。記錄並測量每一項服務的 RTO 和 RPO。
- 建立容錯移轉程序手冊，並包括容錯移轉每個資源、應用程式和服務的所有步驟。
- 建立容錯恢復程序手冊，並包括容錯恢復 (含時程) 每個資源、應用程式和服務的所有步驟
- 制定計畫來啟動和演練程序手冊。使用模擬和混亂測試來測試程序手冊的步驟和自動化。
- 對於位置受損 (例如可用區域或 AWS 區域)，請確保您的系統已就位，可容錯移轉至未受影響位置中運作良好的資源。在容錯移轉測試之前，檢查配額、自動擴展層級和執行的資源。

資源

相關 Well-Architected 的最佳實務：

- [REL13 - 災難復原 \(DR\) 計畫](#)
- [REL10 - 使用故障隔離來保護您的工作負載](#)

相關文件：

- [設定 RTO 和 RPO 目標](#)
- [使用 Route 53 加權路由進行容錯移轉](#)
- [使用 Amazon 應用程式復原控制器進行災難復原](#)
- [具有自動擴展的 EC2](#)
- [EC2 部署 - 多可用區域](#)
- [ECS 部署 - 多可用區域](#)
- [使用 Amazon 應用程式復原控制器切換流量](#)
- [具有 Application Load Balancer 和容錯移轉的 Lambda](#)
- [ACM 複寫和容錯移轉](#)
- [參數存放區複寫和容錯移轉](#)
- [ECR 跨區域複寫和容錯移轉](#)
- [Secrets Manager 跨區域複寫組態](#)
- [針對 EFS 和容錯移轉啟用跨區域複寫](#)
- [EFS 跨區域複寫和容錯移轉](#)
- [聯網容錯移轉](#)
- [使用 MRAP 的 S3 端點容錯移轉](#)
- [為 S3 建立跨區域複寫](#)
- [AWS 上的跨區域容錯移轉和寬限容錯恢復指南](#)
- [使用多區域 Global Accelerator 進行容錯移轉](#)
- [透過 DRS 進行容錯移轉](#)

相關範例：

- [上的災難復原AWS](#)
- [上的彈性災難復原AWS](#)

REL11-BP03 將所有分層的修復自動化

偵測到失敗時，使用自動化功能執行動作來進行修復。降級可能透過內部服務機制自動修復，或需要透過矯正動作重新啟動或移除資源。

對於自我管理的應用程式和跨區域修復，復原設計和自動修復程序可從[現有最佳實務](#)中提取。

重新啟動或移除資源是修復故障的重要工具。最佳實務是盡可能讓服務無狀態。這可防止資源重新啟動時遺失資料或可用性。在雲端，您可以 (且通常應該) 在重新啟動時取代整個資源 (例如，運算執行個體或無伺服器函數)。重新啟動本身是從故障中復原的一個簡單、可靠方法。工作負載中會發生許多不同類型的故障。硬體、軟體、通訊和營運可能會發生故障。

重新啟動或重試也適用於網路請求。對網路逾時和相依系統故障 (其中相依系統會返回錯誤) 套用相同的復原方法。這兩個事件對系統具有類似的影響，因此，不要嘗試讓任何一個事件成為特殊情況，而是藉由指數退避和抖動來採用類似的限制重試策略。重新啟動的能力是復原導向運算和高可用性叢集架構中的一種復原機制。

預期成果：執行自動化動作來矯正錯誤偵測。

常見的反模式：

- 佈建資源，但無自動擴展。
- 個別部署執行個體或容器中的應用程式。
- 部署不透過自動復原就無法部署到多個位置的應用程式。
- 手動復原自動擴展和自動復原無法修復的應用程式。
- 未自動化資料庫容錯移轉。
- 缺乏自動化方法可將流量重新路由至新端點。
- 沒有儲存複寫。

建立此最佳實務的優勢：自動修復可減少您的平均復原時間，並提高可用性。

未建立此最佳實務時的曝險等級：高

實作指引

Amazon EKS 或其他 Kubernetes 服務的設計應包括最小和最大複本或有狀態的集合，以及最小叢集和節點群組規模調整。這些機制提供了最少量的連續可用處理資源，同時會使用 Kubernetes 控制平面自動修復任何失敗。

透過使用運算叢集的負載平衡器存取的设计模式應利用 Auto Scaling 群組。Elastic Load Balancing (ELB) 會自動將傳入的應用程式流量分配到一或多個可用區域 (AZ) 中的多個目標和虛擬設備。

未使用負載平衡的叢集式運算設計，其大小設計應考量至少遺失一個節點。這可讓服務在復原新節點的同時，維持在可能減少的容量中自行執行。範例服務包括 Mongo、DynamoDB Accelerator、Amazon Redshift、Amazon EMR、Cassandra、Kafka、MSK-EC2、Couchbase、ELK 和 Amazon OpenSearch Service。其中許多服務都可以設計為納入額外的自動修復功能。某些叢集技術必須在節點遺失時產生警示，才能觸發自動或手動工作流程來重新建立新節點。此工作流程可以使用 AWS Systems Manager 自動化，以快速修復問題。

Amazon EventBridge 可用來監控及篩選事件，例如 CloudWatch 警示，或其他 AWS 服務的狀態變更。根據事件資訊，它接著可以調用 AWS Lambda、Systems Manager Automation 或其他目標，在您的工作負載上執行自訂修復邏輯。Amazon EC2 Auto Scaling 可設定為檢查 EC2 執行個體的運作狀態。如果執行個體處於執行中以外的任何狀態，或系統狀態為受損，Amazon EC2 Auto Scaling 會將執行個體視為運作狀態不佳，並啟動替代執行個體。對於大規模替換 (例如遺失整個可用區域)，靜態穩定性是高可用性的首選。

實作步驟

- 使用 Auto Scaling 群組在工作負載中部署分層。[Auto Scaling](#) 可以對無狀態應用程式進行自我修復，並新增或移除容量。
- 對於先前提及的運算執行個體，請使用[負載平衡](#)並選擇適當的負載平衡器類型。
- 考慮修復 Amazon RDS。對於待命執行個體，請設定待命執行個體的[自動容錯移轉](#)。對於 Amazon RDS 僅供讀取複本，須有自動化工作流程才能將僅供讀取複本設為主要。
- 對已部署無法在多個位置中部署之應用程式且可以容忍失敗後重新開機的 EC2 執行個體，實作[自動復原](#)。無法將應用程式部署到多個位置時，自動復原可以用來取代失敗的硬體並重新啟動執行個體。執行個體中繼資料和相關聯的 IP 位址，以及 [EBS 磁碟區](#) 和 [Amazon Elastic File System](#) 或 [File Systems for Lustre](#) 以及 [Windows](#) 的掛載點皆會保留。使用 [AWS OpsWorks](#)，可在層級中設定 EC2 執行個體的自動修復功能。
- 當您無法使用自動擴展或自動復原，或自動復原失敗時，則使用 [AWS Step Functions](#) 和 [AWS Lambda](#) 實作自動復原。當您無法使用自動擴展，且無法使用自動復原或自動復原失敗時，則可以使用 AWS Step Functions 和 AWS Lambda 將修復作業自動化。
- [Amazon EventBridge](#) 可用來監控及篩選事件，例如 [CloudWatch 警示](#) 或其他 AWS 服務的狀態變更。根據事件資訊，它接著可以調用 AWS Lambda (或其他目標)，在您的工作負載上執行自訂修復邏輯。

資源

相關的最佳實務：

- [可用性定義](#)
- [REL11-BP01 監控工作負載的所有元件以偵測故障](#)

相關文件：

- [AWS Auto Scaling 的運作方式](#)
- [Amazon EC2 自動復原](#)
- [Amazon Elastic Block Store \(Amazon EBS\)](#)
- [Amazon Elastic File System \(Amazon EFS\)](#)
- [什麼是 Amazon FSx for Lustre ?](#)
- [什麼是 Amazon FSx for Windows File Server ?](#)
- [AWS OpsWorks：使用自動修復來替換出現故障的執行個體](#)
- [什麼是 AWS Step Functions ?](#)
- [什麼是 AWS Lambda ?](#)
- [什麼是 Amazon EventBridge ?](#)
- [使用 Amazon CloudWatch 警示](#)
- [Amazon RDS 容錯移轉](#)
- [SSM - Systems Manager Automation](#)
- [彈性架構最佳實務](#)

相關影片：

- [自動佈建及擴展 OpenSearch Service](#)
- [Amazon RDS 自動容錯移轉](#)

相關範例：

- [Amazon RDS 容錯移轉研討會](#)

相關工具：

- [CloudWatch](#) :
- [CloudWatch X-Ray](#)

REL11-BP04 復原期間需使用資料平面，而非控制平面

控制平面提供的管理 API 適用於建立、讀取和描述、更新、刪除和列出 (CRUDL) 資源，而資料平面則處理日常服務流量。對可能影響彈性的事件實作復原或緩解回應時，請盡量使用最少數量的控制平面操作來復原、重新擴展、還原、修復或容錯移轉服務。資料平面動作應取代這些降級事件期間的任何活動。

例如，以下全都是控制平面動作：啟動新的運算執行個體、建立區塊儲存，以及說明佇列服務。啟動運算執行個體時，控制平面必須執行多項工作，例如尋找具有容量的實體主機、配置網路介面、準備本機區塊儲存磁碟區、產生憑證，以及新增安全規則。控制平面往往是複雜的協同運作。

預期成果：當資源進入受損狀態時，系統能夠將流量從受損資源轉移到健康狀況良好的資源，來自動或手動復原。

常見的反模式：

- 依賴變更 DNS 記錄來重新路由流量。
- 依賴控制平面擴展操作來取代因佈建資源不足而受損的元件。
- 依靠大量、多服務、多 API 的控制平面動作來修復任何類別的損害。

建立此最佳實務的優勢：提高自動化修復的成功率可減少平均復原時間，並改善工作負載的可用性。

未建立此最佳實務時的風險暴露等級：中。對於某些類型的服務降級，則會影響控制平面。若倚賴大量使用控制平面來進行修復，可能會增加復原時間 (RTO) 和平均復原時間 (MTTR)。

實作指引

若要限制資料平面動作，請評估每一項服務還原時所需的動作。

利用 Amazon 應用程式復原控制器來轉移 DNS 流量。這些功能會持續監控應用程式從失敗中復原的功能，讓您在多個 AWS 區域、可用區域和內部部署上控管應用程式復原。

Route 53 路由政策使用控制平面，因此不要依賴它進行復原。Route 53 資料平面會答覆 DNS 查詢，以及執行並評估運作狀態檢查。它們遍佈全球，專為 [100% 可用性服務水準協議 \(SLA\)](#) 而設計。

您在其中建立、更新和刪除 Route 53 資源的 Route 53 管理 API 和主控台在控制平面上執行，這些控制平面的設計旨在優先考慮您在管理 DNS 時所需的強大一致性和耐久性。為了實現此目標，控制平面

位於單一區域中：美國東部 (維吉尼亞北部)。儘管這兩個系統都建置得非常可靠，但控制平面未包含在 SLA 中。在極少數情況下，資料平面的彈性設計允許它保持可用性，而控制平面則不允許。對於災難復原和容錯移轉機制，使用資料平面功能提供可能最好的可靠性。

將您的運算基礎設施設計為靜態穩定，以避免在發生事件期間使用控制平面。例如，若您使用的是 Amazon EC2 執行個體，請避免手動佈建新執行個體，或避免指示 Auto Scaling 群組在回應中新增執行個體。為獲得最高層級的彈性，請在用於容錯移轉的叢集中佈建足夠的容量。如果必須限制此容量閾值，請對整體端對端系統設定節流，以安全地限制總流量達到所限制的資源集。

對於像是 Amazon DynamoDB、Amazon API Gateway、負載平衡器和 AWS Lambda 無伺服器服務，使用這些服務會利用資料平面。不過，建立新功能、負載平衡器、API 閘道或 DynamoDB 資料表是控制平面動作，應在降級前完成，以準備進行事件和容錯移轉動作的演練。對於 Amazon RDS，資料平面動作允許存取資料。

如需關於資料平面、控制平面以及 AWS 如何建置服務以滿足高可用性目標的資訊，請參閱[使用可用區域的靜態穩定性](#)。

了解哪些作業位於資料平面，哪些位於控制平面。

實作步驟

針對需要在降級事件之後還原的每個工作負載，評估容錯移轉執行手冊、高可用性設計、自動修復設計，或 HA 資源還原計畫。找出可能視為控制平面動作的每個動作。

考慮將控制動作變更為資料平面動作：

- Auto Scaling (控制平面) 至預先擴展的 Amazon EC2 資源 (資料平面)
- Amazon EC2 執行個體擴展 (控制平面) 到 AWS Lambda 擴展 (資料平面)
- 使用 Kubernetes 評估任何設計，以及控制平面動作的性質。新增 Pod 是 Kubernetes 中的資料平面動作。動作應限於新增 Pod 而不是新增節點。使用[過度佈建的節點](#)是限制控制平面動作的慣用方法

請考慮可讓資料平面動作影響相同修復措施的替代方法。

- Route 53 記錄變更 (控制平面) 或 Amazon 應用程式復原控制器 (資料平面)
- [Route 53 運作狀態檢查以進行更多自動化更新](#)

如果服務具任務關鍵性，請考慮次要區域中的某些服務，以便在未受影響的區域中執行更多控制平面和資料平面動作。

- 將主要區域中的 Amazon EC2 Auto Scaling 或 Amazon EKS 與次要區域中的 Amazon EC2 Auto Scaling 或 Amazon EKS 相比較，並將流量路由到次要區域 (控制平面動作)
- 將僅供讀取複本設為主要，或在主要區域中嘗試相同的動作 (控制平面動作)

資源

相關的最佳實務：

- [可用性定義](#)
- [REL11-BP01 監控工作負載的所有元件以偵測故障](#)

相關文件：

- [APN 合作夥伴：可以幫助您實現容錯自動化的合作夥伴](#)
- [AWS Marketplace：可用於容錯的產品](#)
- [Amazon 建置者資料中心：控管較小服務，避免分散式系統過載](#)
- [Amazon DynamoDB API \(控制平面和資料平面\)](#)
- [AWS Lambda 執行 \(分割成控制平面和資料平面\)](#)
- [AWS Elemental MediaStore 資料平面](#)
- [使用 Amazon 應用程式復原控制器建置高彈性應用程式，第 1 部分：單一區域堆疊](#)
- [使用 Amazon 應用程式復原控制器建置高彈性應用程式，第 2 部分：多區域堆疊](#)
- [使用 Amazon Route 53 建立災難復原機制](#)
- [什麼是 Amazon 應用程式復原控制器？](#)
- [Kubernetes 控制平面和資料平面](#)

相關影片：

- [回歸基礎 - 使用靜態穩定性](#)
- [使用 AWS 全球服務建置彈性的多站點工作負載](#)

相關範例：

- [Amazon 應用程式復原控制器簡介](#)
- [Amazon 建置者資料中心：控管較小服務，避免分散式系統過載](#)

- [使用 Amazon 應用程式復原控制器建置高彈性應用程式，第 1 部分：單一區域堆疊](#)
- [使用 Amazon 應用程式復原控制器建置高彈性應用程式，第 2 部分：多區域堆疊](#)
- [使用可用區域實現靜態穩定性](#)

相關工具：

- [Amazon CloudWatch](#)
- [AWS X-Ray](#)

REL11-BP05 使用靜態穩定性來防止雙模態行為

工作負載應該是靜態穩定的，且只在單一正常模式下運作。雙模態行為是指工作負載在正常和故障模式下呈現不同行為的情況。

例如，您可能在不同的可用區域中啟動新的執行個體，嘗試回復可用區域故障。這可能會導致在故障模式期間產生雙模態回應。您應改為建置靜態穩定且僅以一種模式操作的工作負載。在此範例中，這些執行個體應該在發生故障之前已佈建在第二個可用區域。此靜態穩定設計可以確保工作負載僅在單一模式下運作。

預期成果：工作負載不會在正常和故障模式出現雙模態行為。

常見的反模式：

- 假設無論故障範圍，一律可以佈建資源。
- 嘗試在故障期間動態取得資源。
- 在發生故障之前，請勿在多個區域佈建適度的資源。
- 僅考慮運算資源的靜態穩定設計。

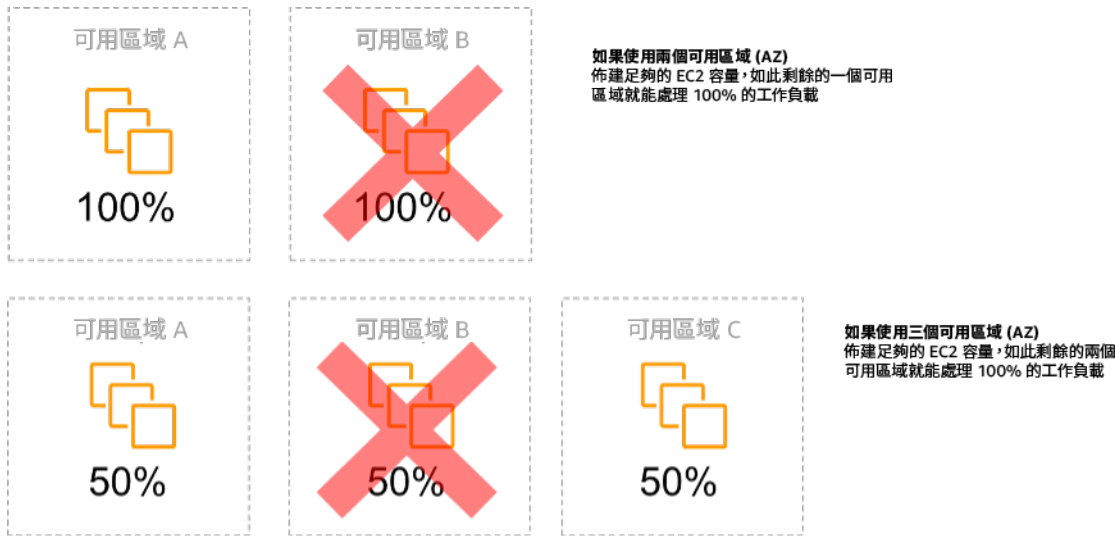
建立此最佳實務的優勢：使用靜態穩定設計執行的工作負載，能夠在正常和故障事件發生時產生可預測的結果。

未建立此最佳實務時的曝險等級：中

實作指引

雙模態行為是指您的工作負載在正常和故障模式下展現出不同的行為，例如，當可用區域故障時，仰賴啟動新的執行個體。其中一個雙模態行為範例是，如果移除一個可用區域，則穩定的 Amazon EC2 設計會在每個可用區域佈建足夠的執行個體來處理工作負載。Elastic Load Balancing 或

Amazon Route 53 運作狀態會進行檢查，將負載從受損的執行個體中移出。流量轉移後，使用 AWS Auto Scaling 以非同步方式取代故障區域的執行個體，並在運作良好的區域中啟動這些執行個體。運算部署 (例如 EC2 執行個體或容器) 的靜態穩定性可提供最高的可靠性。



跨可用區域 EC2 執行個體的靜態穩定性

這必須在所有彈性情況下，與此模型的成本以及維護工作負載的商業價值互相衡量。佈建較少運算容量並在故障時啟動新執行個體的成本較低，但是對於大規模故障 (例如可用區域損壞)，這種方法的效率較低，因為它同時仰賴作業平面，以及未受影響區域中的足夠資源。

您的解決方案也應該權衡可靠性與工作負載的成本需求。靜態穩定架構適用於多種架構，包括在多個可用區域的運算執行個體、資料庫僅供讀取複本設計、Kubernetes (Amazon EKS) 叢集設計，以及多區域容錯移轉架構。

若在每個區域使用更多資源，也可以實施更靜態的穩定設計。透過新增更多區域，您可以降低靜態穩定性所需的額外運算量。

雙模態行為範例之一是網路逾時，網路逾時可能導致系統嘗試重新整理整個系統的組態狀態。這樣一來，即會給另一個元件新增意外負載，且可能導致其發生故障，從而引發其他意外後果。這種負面意見回饋迴圈會影響工作負載的可用性。反之，您可以建置靜態穩定且僅以一種模式操作的系統。靜態穩定的設計是執行持續工作，並始終以固定的規律重新整理組態狀態。叫用失敗時，工作負載會使用先前的快取數值，並啟動警示。

另一個雙模態行為範例是允許用戶端在發生失敗時繞過您的工作負載快取。這看起來可能是滿足用戶端需求的解決方案，但會大幅變更工作負載的需求，且可能導致故障。

評估關鍵工作負載，決定哪些工作負載需要此類彈性設計。針對關鍵工作負載，必須檢視每個應用程式元件。需要靜態穩定性評估的服務類型範例如下：

- 運算：Amazon EC2、EKS-EC2、ECS-EC2、EMR-EC2
- 資料庫：Amazon Redshift、Amazon RDS、Amazon Aurora
- 儲存：Amazon S3 (單一區域)、Amazon EFS (掛載)、Amazon FSx (掛載)
- 負載平衡器：在某些設計下

實作步驟

- 建置靜態穩定且僅以一種模式操作的系統。在此情況下，請在每個可用區域佈建足夠的執行個體，以處理移除一個可用區域時的工作負載容量。許多服務皆可用於路由到運作狀態良好的資源，例如：
 - [跨區域 DNS 路由](#)
 - [MRAP Amazon S3 多區域路由](#)
 - [AWS Global Accelerator](#)
 - [Amazon 應用程式復原控制器說明](#)
- 設定[資料庫讀取複本](#)以考慮單一主要執行個體或讀取複本的遺失情況。若僅供讀取複本為流量提供服務，則每個可用區域中的數量應等同於區域故障時的整體需求。
- 在 Amazon S3 儲存中設定重要資料，以便可用區域故障時，能針對所儲存的資料保持靜態穩定。如果使用 [Amazon S3 One Zone-IA](#) 儲存類別，則不應將其視為靜態穩定，因為該區域的遺失會最小化此儲存資料的存取權。
- [Load balancers](#) 有時會設定錯誤，或本來就設定為供特定可用區域使用。在這種情況下，靜態穩定設計可能是在更複雜的設計中將工作負載分散到多個可用區域。出於安全性、延遲或成本考量，可以使用原始設計來減少區域間流量。

資源

相關 Well-Architected 的最佳實務：

- [可用性定義](#)
- [REL11-BP01 監控工作負載的所有元件以偵測故障](#)
- [REL11-BP04 復原期間需使用資料平面，而非控制平面](#)

相關文件：

- [在災難復原計畫中盡可能減少相依關係](#)
- [Amazon 建置者資料中心：使用可用區域實現靜態穩定性](#)

- [故障隔離界限](#)
- [使用可用區域實現靜態穩定性](#)
- [多區域 RDS](#)
- [在災難復原計畫中盡可能減少相依關係](#)
- [跨區域 DNS 路由](#)
- [MRAP Amazon S3 多區域路由](#)
- [AWS Global Accelerator](#)
- [Amazon 應用程式復原控制器說明](#)
- [單區域 Amazon S3](#)
- [跨區域負載平衡](#)

相關影片：

- [AWS 中的靜態穩定性：AWS re:Invent 2019：Amazon 建置者資料中心簡介 \(DOP328\)](#)

REL11-BP06 當事件影響可用性時傳送通知

當偵測到閾值超標時傳送通知，即使問題造成的事件已自動解決。

自動修復功能可讓您的工作負載變得可靠。不過，也可能會遮蔽需要解決的潛在問題。實作適當的監控和事件，讓您能夠偵測到問題模式 (包括自動修復功能處理的問題模式)，以解決根本原因問題。

具有韌性的系統可將降級事件立即傳達給權責團隊。這些通知應該透過一個或多個通訊管道傳送。

預期成果：當超過閾值 (例如錯誤率、延遲或其他關鍵績效指標 (KPI)) 時，營運團隊會立即收到警示，以盡快解決問題，避免或將使用者負面影響降至最低。

常見的反模式：

- 傳送太多警示。
- 傳送不可採取行動的警示。
- 警示閾值設置太高 (太敏感) 或太低 (太遲鈍)。
- 不傳送外部相依性的警示。
- 在設計監控和警示時，不考慮[微小故障](#)。

- 進行修復自動化，但不通知權責團隊需要修復。

建立此最佳實務的優勢：回復通知可讓營運和業務團隊注意到服務降級，讓他們可以立即反應，將平均偵測時間 (MTTD) 和平均復原時間 (MTTR) 降至最低。回復事件的通知也會確認您不會忽略不常發生的問題。

未建立此最佳實務時的風險暴露等級：中。若無法實作適當的監控和事件通知機制，您可能就無法偵測到問題模式 (包括自動修復功能處理的問題模式)。只有當使用者聯絡客服或偶然情況下，團隊才會注意到系統降級。

實作指引

定義監控策略時，觸發警示是常見的事件。此事件可能包含警示的識別碼、警示狀態 (例如 IN ALARM 或 OK) 以及觸發原因詳情。在許多情況下，系統應檢測到警示事件並傳送電子郵件通知。這是警示動作範例。警示通知對於可觀測性至關重要，因為它會通知權責人員有問題發生。然而，當可觀測性解決方案對事件的回應措施夠熟練後，便可以自動修復問題，無須人為介入。

建立 KPI 監控警示後，閾值超過時就應會向權責團隊傳送警示。這些警示也可用於觸發嘗試修復降級的自動化程序。

針對更複雜的閾值監控，則應考慮使用複合警示。複合警示會使用數個 KPI 監控警示，根據作業商務邏輯建立警示。CloudWatch 警示可設定為傳送電子郵件，或使用 Amazon SNS 整合或 Amazon EventBridge 在第三方事件追蹤系統中記錄事件。

實作步驟

根據監控工作負載的方式建立各種警示類型，例如：

- 應用程式警示可用來偵測工作負載任何無法正常運作的部分。
- [基礎設施警示](#)會指出何時擴展資源。警示會在儀表板上以視覺化方式顯示、透過 Amazon SNS 或電子郵件傳送提醒，以及搭配使用 Auto Scaling 來擴展或縮減工作負載資源。
- 可建立簡單的[靜態指示](#)，以監控指標在指定評估期間內超過靜態閾值的時間。
- [複合警示](#)可以涵蓋來自多個來源的複雜警示。
- 建立警示後，請建立適當的通知事件。可以直接調用 [Amazon SNS API](#) 來傳送通知，並連結任何自動化以進行修復或通訊。
- 透過 [AWS Health](#) 隨時掌握服務降級的相關資訊。[透過 AWS 使用者通知 建立符合用途的 AWS Health 事件通知](#)，以利用電子郵件和聊天管道傳送，並透過 [Amazon EventBridge 以程式設計方式與您的監控和警示工具整合](#)。

資源

相關 Well-Architected 的最佳實務：

- [可用性定義](#)

相關文件：

- [根據靜態閾值建立 CloudWatch 警示](#)
- [什麼是 Amazon EventBridge？](#)
- [什麼是 Amazon Simple Notification Service？](#)
- [發佈自訂指標](#)
- [使用 Amazon CloudWatch 警示](#)
- [設定 CloudWatch 複合警示](#)
- [re:Invent 2022 中 AWS Observability 的最新消息](#)

相關工具：

- [CloudWatch](#)：
- [CloudWatch X-Ray](#)

REL11-BP07 建立您的產品架構以符合可用性目標和運行時間服務水準協議 (SLA)

建立您的產品架構以符合可用性目標和運行時間服務水準協議 (SLA)。如果您發佈或私下同意可用性目標或運行時間 SLA，請確認您的架構和操作程序的設計可以支援。

預期成果：每個應用程式都有一個已定義的可用性目標和效能指標的 SLA，可以進行監控和維護，以達到業務成果。

常見的反模式：

- 設計和部署工作負載，而未設定任何 SLA。
- SLA 指標設定為高，而沒有合理或業務要求。
- 設定 SLA 但未考慮相依性及其基礎 SLA。
- 建立應用程式設計而未考慮彈性的共同責任模型。

建立此最佳實務的優勢：根據關鍵彈性目標設計應用程式，可協助您達成業務目標和客戶期望。這些目標可協助推動應用程式設計程序，評估不同的技術和考慮各種權衡。

未建立此最佳實務時的曝險等級：中

實作指引

應用程式設計必須將多元的要求納入考慮，這些要求是從業務、營運和財務目標衍生而來。在營運要求內，工作負載必須有特定彈性指標目標，才能適當地監控和支援。彈性指標不應該在部署工作負載之後設定或衍生。它們應該在設計階段期間定義，協助引導各種決策和權衡。

- 每個工作負載都應該有自己的一組彈性指標。這些指標可能與其他業務應用程式不同。
- 降低相依性對可用性有正面影響。每個工作負載都應該考慮其相依性及其 SLA。一般而言，選取可用性目標等於或大於工作負載目標的相依性。
- 請考慮鬆散耦合設計，讓您的工作負載在可行時不論是否有相依性受損，都可以正確操作。
- 減少控制平面相依性，特別是復原或降級期間。評估針對任務關鍵性工作負載靜態穩定的設計。使用資源節省來增加工作負載中這些相依性的可用性。
- 可觀測性和檢測對於透過降低平均偵測時間 (MTTD) 和平均修復時間 (MTTR) 來達成 SLA 相當關鍵。
- 低頻率失敗 (MTBF 較長)、較短的失敗偵測時間 (較短 MTTD) 和較短的修復時間 (較短 MTTR)，是用來在分散式系統中改善可用性的三個因素。
- 建立和符合工作負載的彈性指標，是任何有效設計的基礎。這些設計必須考慮到設計複雜性、服務相依性、效能、擴展和成本的權衡。

實作步驟

- 請考慮下列問題，審核和記載工作負載設計：
 - 控制平面用於工作負載的哪個地方？
 - 工作負載如何實作容錯能力？
 - 擴展、自動擴展、備援和高可用性元件的設計模式是什麼？
 - 資料一致性和可用性的要求是什麼？
 - 資源節省或資源靜態穩定性是否有任何考慮？
 - 服務相依性是什麼？
- 與利益相關者合作時根據工作負載架構定義 SLA 指標。請考慮工作負載所使用所有相依性的 SLA。
- 一旦設定 SLA 目標，最佳化架構以符合 SLA。

- 一旦設定可符合 SLA 的設計，實作營運變更、處理自動化以及也會著重在降低 MTDD 和 MTTR 的執行手冊。
- 部署之後，監控和報告 SLA。

資源

相關的最佳實務：

- [REL03-BP01 選擇如何分割工作負載](#)
- [REL10-BP01 將工作負載部署至多個位置](#)
- [REL11-BP01 監控工作負載的所有元件以偵測故障](#)
- [REL11-BP03 將所有分層的修復自動化](#)
- [REL12-BP04 使用混沌工程測試彈性](#)
- [REL13-BP01 定義停機和資料遺失的復原目標](#)
- [了解工作負載運作狀態](#)

相關文件：

- [備援的可用性](#)
- [可靠性支柱 - 可用性](#)
- [測量可用性](#)
- [AWS 故障隔離界限](#)
- [彈性的共同責任模型](#)
- [使用可用區域實現靜態穩定性](#)
- [AWS 服務水準協議 \(SLA\)](#)
- [AWS 上小組型架構指南](#)
- [AWS 基礎設施](#)
- [《進階多可用區域彈性模式》白皮書](#)

相關服務：

- [Amazon CloudWatch](#)
- [AWS Config](#)

- [AWS Trusted Advisor](#)

REL 12. 如何測試可靠性？

在將工作負載設計為可彈性應對生產壓力之後，進行測試是確認其依設計運作並提供預期之彈性的唯一方法。

最佳實務

- [REL12-BP01 使用程序手冊調查失敗](#)
- [REL12-BP02 執行事件後分析](#)
- [REL12-BP03 測試可擴展性和效能需求](#)
- [REL12-BP04 使用混沌工程測試彈性](#)
- [REL12-BP05 定期進行演練日](#)

REL12-BP01 使用程序手冊調查失敗

藉由在程序手冊中記錄調查程序，對無法充分理解的失敗情境進行快速一致的回應。程序手冊是為識別造成失敗情境的因素所執行的預先定義步驟。在識別或呈報問題之前，任何程序步驟的結果都會用來決定要採取的後續步驟。

程序手冊是您必須進行的主動規劃，以便能夠有效地採取回應動作。在生產環境中遇到程序手冊未涵蓋的故障情境時，請先解決問題 (解決燃眉之急)。然後返回並查看您為解決問題所採取的步驟，並使用這些步驟在程序手冊中新增新的項目。

請注意，程序手冊用於回應特定事件，而執行手冊則用於實現特定成果。執行手冊通常用於例行活動，而程序手冊則用於回應非例行事件。

常見的反模式：

- 在不知道診斷問題或回應事件的程序之情況下，規劃部署工作負載。
- 調查事件時，未規劃即決定要向哪些系統收集日誌和指標。
- 指標和事件的保留時間過短，無法用以擷取資料。

建立此最佳實務的優勢：擷取程序手冊可確保流程得到一致遵循。有系統地編纂您的程序手冊可限制手動活動引入錯誤。程序手冊自動化可免除團隊成員介入的需要，或在介入開始時提供其他資訊，從而縮短事件回應時間。

未建立此最佳實務時的曝險等級：高

實作指引

- 使用程序手冊識別出問題。程序手冊是調查問題的書面程序。透過在程序手冊中記錄程序，對失敗情境做出一致且迅速的回應。程序手冊包含的資訊和指南必須能夠讓技能嫺熟的人員得以收集適用資訊、識別潛在的失敗來源、隔離故障，以及判斷成因 (執行事件後分析)。
- 將程序手冊實做為程式碼。透過撰寫程序手冊指令碼，以程式碼形式執行操作，確保一致性並限制和減少手動程序引起的錯誤。程序手冊可由多個指令碼組成，這些指令碼代表識別成因時可能需要的不同步驟。執行手冊活動可以做為程序手冊活動的一部分被調用或執行，或者可能提示執行程序手冊，以回應已識別的事件。
 - [透過 AWS Systems Manager 自動化您的操作程序手冊](#)
 - [AWS Systems Manager 執行命令](#)
 - [AWS Systems Manager Automation](#)
 - [什麼是 AWS Lambda？](#)
 - [什麼是 Amazon EventBridge？](#)
 - [使用 Amazon CloudWatch 警示](#)

資源

相關文件：

- [AWS Systems Manager Automation](#)
- [AWS Systems Manager 執行命令](#)
- [透過 AWS Systems Manager 自動化您的操作程序手冊](#)
- [使用 Amazon CloudWatch 警示](#)
- [使用 Canary \(Amazon CloudWatch Synthetics\)](#)
- [什麼是 Amazon EventBridge？](#)
- [什麼是 AWS Lambda？](#)

相關範例：

- [使用程序手冊和執行手冊將操作自動化](#)

REL12-BP02 執行事件後分析

審查影響客戶的事件，並識別成因和預防性行動項目。使用此資訊來開發緩解措施，以限制或防止事件再次發生。制定可快速有效回應的程序。適當地傳達成因和為目標受眾量身打造的糾正措施。建立一種可以根據需要將這些原因傳達給其他人的方法。

評估現有測試找不到問題的原因。如果測試尚未存在，請為此案例新增測試。

預期成果：您的團隊擁有一致且商定的方法來處理事件後分析。一種機制是[錯誤糾正 \(COE\) 流程](#)。COE 程序可幫助您的團隊識別、了解和解決事件的根本原因，同時還能建置機制和防護機制，以限制相同事件再次發生的可能性。

常見的反模式：

- 尋找成因，但未繼續深入尋找其他潛在問題和減輕方法。
- 僅確定人為錯誤原因，而未嘗試可防止人為錯誤發生的任何培訓或自動化。
- 專注於追究責任，而不是了解根本原因，造成恐懼文化並阻礙開放的溝通
- 無法分享見解，只讓一小群人知道事件分析調查結果，讓其他人無法從學到的教訓中受益
- 沒有機制可擷取機構知識而失去寶貴的見解，因為組織不會以更新過的最佳實務形式保存所學到的教訓，並導致重複發生相同或類似根本原因的事件

建立此最佳實務的優勢：進行事件後分析並分享結果，以讓其他實作了相同成因的工作負載減輕風險，並讓工作負載能夠在事件發生前實作減輕措施或自動復原。

未建立此最佳實務時的曝險等級：高

實作指引

良好的事件後分析提供了機會，為系統中其他地方使用的架構模式問題提出通用解決方案。

COE 程序的基石是記錄和解決問題。建議您定義標準化方式來記錄關鍵的根本原因，並確保加以檢視和解決。為事件後分析程序指派明確的擁有權。指定負責監督事件調查和後續跟進的團隊或個人。

鼓勵專注於學習和改進的文化，而不是追究責任的文化。強調目標是預防未來的事件，而不是懲罰個人。

開發用於進行事件後分析的明確定義程序。這些程序應概述要採取的步驟、要收集的資訊，以及要在分析期間解決的關鍵問題。徹底調查事件，跳脫出直接原因以找出根本原因和成因。使用諸如[五個為什麼](#)等技巧深入研究潛在問題。

維護事件分析所學教訓的儲存庫。此機構知識可以做為未來事件和預防工作的參考。分享事件後分析的調查結果和見解，並考慮舉行公開邀請的事件後檢討會議，以討論學到的教訓。

實作步驟

- 在進行事件後分析時，請確保事件後分析不會讓相關人員受到責備。這可讓事件中的相關人員平心靜氣看待建議的糾正措施，並促進誠實地自我評估與跨團隊合作。
- 定義標準化方式來記錄重要問題。這類文件的範例結構如下：
 - 發生了什麼？
 - 對客戶和您的業務有什麼影響？
 - 根本原因是什麼？
 - 您擁有什麼可以提供支援的資料？
 - 例如，指標和圖表
 - 對關鍵支柱的影響有哪些 (尤其是安全性)？
 - 建立工作負載的架構時，您可依照業務環境，在各支柱之間作出權衡。這些業務決策可以讓您了解工程設計的優先順序。您可以選擇在開發環境中以可靠性做為代價最佳化成本，或者針對關鍵任務解決方案，以較高成本達到可靠性的最佳化。安全始終是首要工作，因為您必須保護客戶。
 - 您獲得了什麼教訓？
 - 您正在採取什麼糾正措施？
 - 動作項目
 - 相關項目
- 建立用於進行事件後分析的明確定義標準作業程序。
- 設定標準化的事件報告程序。全面記錄所有事件，包括初始事件報告、日誌、通訊，以及事件期間採取的行動。
- 請記住，發生事件時不見得會有中斷情形。事件也可能是幾乎錯過的情況，或是系統雖以意想不到的方式執行，卻仍可履行其業務功能。
- 請根據意見回饋和學到的教訓，持續改善事件後分析程序。
- 擷取知識管理系統中的關鍵調查結果，並考慮任何應新增至開發人員指南或部署前檢查清單的模式。

資源

相關文件：

- [為什麼您應該制定錯誤糾正 \(COE\)](#)

相關影片：

- [Amazon 成功失敗的方法](#)
- [AWS re:Invent 2021 - Amazon 建置者資料中心：Amazon 的卓越營運](#)

REL12-BP03 測試可擴展性和效能需求

使用諸如負載測試等技術來驗證工作負載符合擴展和效能需求。

在雲端，您可以隨需為工作負載建立正式作業規模的測試環境。您不必依賴縮減規模的測試環境，這可能會導致正式作業行為的預測不準確，而是可以使用雲端來佈建測試環境，以確切反映您預期的正式作業環境。此環境可協助您以更準確模擬您的應用程式面臨的真實狀況來進行測試。

除了效能測試之外，務必確認您的基本資源、擴展設定、服務配額和彈性設計能夠在負載下如預期運作。這種整體方法可確認您的應用程式即使在最嚴苛的條件下，仍能可靠地擴展和執行。

預期成果：您的工作負載即使在承受尖峰負載的情況下，仍維持其預期行為。您主動解決可能隨應用程式發展和演進而出現的任何效能相關問題。

常見的反模式：

- 您使用的測試環境與正式作業環境相似程度不高。
- 您將負載測試視為單獨的一次性活動，而不是部署持續整合 (CI) 管道整體的一部分。
- 您未定義明確且可衡量的效能需求，例如回應時間、輸送量和可擴展性目標。
- 您在負載情況不切實際或不足的情況下執行測試，而且未能測試尖峰負載、突發性峰值和持續高負載。
- 您未透過超過預期的負載限制來對工作負載進行壓力測試。
- 您使用不適當或不適合的負載測試和效能分析工具。
- 您缺乏全方位的監控和警示系統可用來追蹤效能指標並偵測異常。

建立此最佳實務的優勢：

- 負載測試可協助您在付諸正式作業之前，找出系統中潛在的效能瓶頸。當您模擬正式作業層級流量和工作負載時，就可找出系統處理負載時可能遇到困難層面，例如，回應時間緩慢、資源限制或系統故障。
- 當您在各種負載條件下測試系統時，您可以更了解支援工作負載所需的資源需求。此資訊可協助您做出有關資源分配的明智決策，並防止資源過度佈建或佈建不足。

- 若要找出潛在的故障點，您可以觀察工作負載在高負載條件下的運作情形。此資訊可協助您透過適時實作容錯機制、容錯移轉策略和備援措施，來改善工作負載的可靠性和彈性。
- 您可及早找出和解決效能問題，這有助於避免系統中斷、回應時間緩慢和使用者不滿意所帶來代價高昂的後果。
- 測試期間收集的詳細效能資料和分析資訊，可協助您解決正式作業時可能出現的效能相關問題。這樣就能加快回應和解決事件，進而降低對使用者和組織營運的影響。
- 在某些產業中，主動效能測試可協助讓工作負載符合合規標準，進而降低遭受懲罰或產生法律問題的風險。

未建立此最佳實務時的曝險等級：高

實作指引

第一步是定義全面的測試策略，涵蓋擴展和效能需求的所有層面。首先，請根據您的業務需求，例如輸送量、延遲長條圖和錯誤率，清楚定義工作負載的服務層級目標 (SLO)。接著設計一套測試，可模擬從平均用量到突發尖峰和持續尖峰負載的各種負載案例，並確認工作負載的行為符合您的 SLO。這些測試應該自動化，並整合到您的持續整合和部署管道中，以及早發現開發過程中的效能迴歸。

為了有效測試擴展和效能，請投資正確的工具和基礎設施。這包括可產生實際使用者流量的負載測試工具、識別瓶頸的效能分析工具，以及追蹤關鍵指標的監控解決方案。重要的是，您應確認您的測試環境在基礎設施和環境條件方面盡量與正式作業環境相符，以使測試結果盡可能準確。為了更容易可靠地複寫和調整類似正式作業的設定，請使用基礎設施即程式碼和容器型應用程式。

擴展和效能測試是持續進行的程序，而不是一次性的活動。實作全面監控和警示，以追蹤應用程式正式作業時的效能，並使用這些資料持續改進您的測試策略和最佳化工作。定期分析效能資料，以找出新興問題、測試新的擴展策略並實作最佳化，以提高應用程式的效率和可靠性。只要您採用迭代方法並持續從正式作業資料中學習，就可以確認您的應用程式能夠適應不斷變化的使用者需求，並隨著時間維持彈性和最佳效能。

實作步驟

1. 確立明確且可衡量的效能需求，例如回應時間、輸送量和可擴展性目標。這些需求應以工作負載的使用模式、使用者期望和業務需求為基礎。
2. 選取並設定負載測試工具，以準確地模擬正式作業環境中的負載模式和使用者的行為。
3. 設定盡量與正式作業環境相符的測試環境，包括基礎設施和環境條件，以提高測試結果的準確性。
4. 建立測試套件，涵蓋從平均用量模式到尖峰負載、快速尖峰和持續高負載等各種案例。將這些測試整合到您的持續整合和部署管道中，以及早發現開發過程中的效能迴歸。

5. 執行負載測試以模擬實際使用者流量，並了解應用程式在不同負載條件下的行為。若要對您的應用程式進行壓力測試，請超過預期的負載並觀察其行為，例如回應時間變慢、資源耗盡或系統故障，這樣做有助於識別應用程式的中斷點並納入擴展策略。透過逐步增加負載來評估工作負載的可擴展性，並衡量效能影響以了解擴展限制並規劃未來容量需求。
6. 實作全面的監控和警示，以追蹤效能指標、偵測異常情況，並在超出閾值時初始化擴展動作或通知。
7. 持續監控和分析效能資料，以識別需要改進的層面。對您的測試策略和最佳化工作採取迭代方法。

資源

相關的最佳實務：

- [REL01-BP04 監控和管理配額](#)
- [REL06-BP01 監控工作負載的所有元件 \(產生\)](#)
- [REL06-BP03 傳送通知 \(即時處理和警示\)](#)

相關文件：

- [負載測試應用程式](#)
- [AWS 上的分散式負載測試](#)
- [應用程式效能監控](#)
- [Amazon EC2 測試政策](#)

相關範例：

- [AWS 上的分散式負載測試 \(GitHub\)](#)

相關工具：

- [Amazon CodeGuru Profiler](#)
- [Amazon CloudWatch RUM](#)
- [Apache JMeter](#)
- [K6](#)
- [Vegeta](#)

- [Hey](#)
- [ab](#)
- [wrk](#)
- [AWS 上的分散式負載測試](#)

REL12-BP04 使用混沌工程測試彈性

定期在位於或盡可能鄰近生產環境的環境中執行混沌試驗，以了解您的系統因應不良狀況的能力。

預期成果：

除了以彈性測試驗證您的工作負載在某事件期間的已知預期行為以外，還可以藉由在故障注入試驗中套用混沌工程或注入非預期的負載，來定期驗證工作負載的彈性。結合混沌工程與彈性測試，您將可確信工作負載在經歷元件失敗後仍可存留，並且可在 (幾乎) 不受影響的情況下從非預期的中斷復原。

常見的反模式：

- 針對彈性進行設計，但未確認工作負載在錯誤發生時的整體運作情形。
- 未曾在真實的情況和預期的負載下試驗。
- 未將試驗視為程式碼或透過開發週期加以維護。
- 未在 CI/CD 管道中與部署以外執行混沌試驗。
- 在決定要以哪些錯誤進行試驗時，未使用過去的事故後分析。

建立此最佳實務的優勢：注入錯誤以驗證工作負載的彈性，可讓您確信在發生真正的錯誤時，彈性設計的復原程序將可發揮作用。

未建立此最佳實務時的曝險等級：中

實作指引

混沌工程可讓您的團隊有能力以受控的方式，持續在服務供應商、基礎架構、工作負載和元件層級注入真實的中斷 (模擬)，且對客戶 (幾乎) 不會造成影響。它可讓您的團隊從錯誤中學習，並且觀察、測量及改善工作負載的彈性，以及驗證在事件發生時會引發提醒，且團隊會收到通知。

持續執行時，混沌工程可能會凸顯您工作負載中的缺陷，且若未解決，可能會對可用性與操作產生負面影響。

Note

混沌工程是在系統中進行試驗的專業領域，旨在建立對系統承受生產環境中紊亂情況的能力的信心。 – [混沌工程的原則](#)

如果系統能夠承受這些中斷，則應將混沌試驗視為自動化迴歸測試來維護。此時，您應在系統開發生命週期 (SDLC) 和 CI/CD 管道中執行混沌試驗。

若要確定您的工作負載可以承受元件失敗，請在試驗中注入真實事件。例如，進行遺失 Amazon EC2 執行個體或容錯移轉主要 Amazon RDS 資料庫執行個體的試驗，並驗證您的工作負載未受影響 (或僅受到些微影響)。使用元件錯誤的組合，模擬可用區域的中斷可能導致的事件。

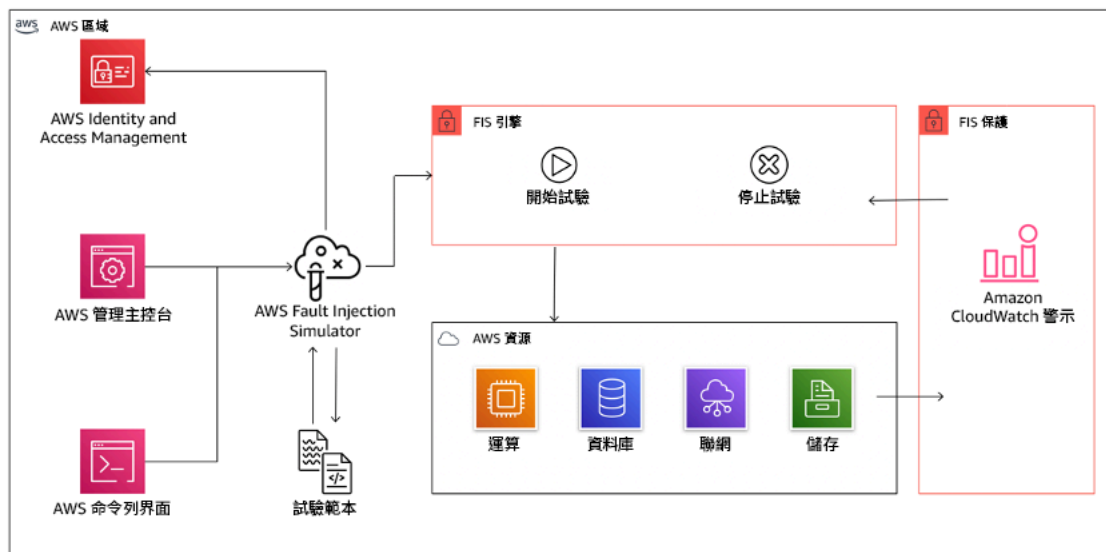
對於應用程式層級的錯誤 (例如當機)，您可以從記憶體和 CPU 用盡之類的壓力源開始著手。

為了驗證由於間歇性網路中斷導致的外部相依性的[備用或容錯移轉機制](#)，您的元件應在指定的時間內 (可延續數秒到數小時) 阻止對第三方供應商的存取來模擬這類事件。

其他降級模式可能會導致功能降低和回應速度緩慢，而往往會導致服務中斷。這種降級的常見原因是關鍵服務的延遲增加和不可靠的網路通訊 (丟包)。以這些錯誤 (包括延遲、已捨棄訊息和 DNS 失敗等聯網影響) 進行的試驗，可包含無法解析名稱、無法連線到 DNS 服務，或無法建立相依服務的連線等情境。

混沌工程工具：

AWS Fault Injection Service (AWS FIS) 是用來執行故障注入試驗的全受管服務，這些試驗可作為 CD 管道的一部分，或未於管道以外。AWS FIS 很適合在混沌工程演練日期間使用。它支援跨不同類型的資源同時引入故障，包括 Amazon EC2、Amazon Elastic Container Service (Amazon ECS)、Amazon Elastic Kubernetes Service (Amazon EKS) 和 Amazon RDS。這些錯誤包括終止資源、強制執行容錯移轉、施壓於 CPU 或記憶體、限流，以及封包遺失。由於它與 Amazon CloudWatch 警示整合，因此您可以將停止條件設定為防護機制，以在試驗導致非預期的影響時將其回復。



AWS Fault Injection Service 與 AWS 資源整合，讓您可對工作負載執行故障注入試驗。

故障注入試驗也有數個第三方選項。其中包括開放原始碼工具，例如 [Chaos Toolkit](#)、[Chaos Mesh](#) 和 [Litmus Chaos](#)，以及諸如 Gemlin 之類的商業選項。為了擴大可以在 AWS 上注入的故障範圍，AWS FIS 會與 [Chaos Mesh](#) 和 [Litmus Chaos](#) 整合，使您可以在多種工具之間協調故障注入工作流程。例如，您可以在使用 AWS FIS 錯誤動作終止隨機選定百分比的叢集節點時，使用 Chaos Mesh 或 Litmus 錯誤對 Pod 的 CPU 執行壓力測試。

實作步驟

1. 決定要將哪些錯誤用於試驗。

評估您的工作負載設計是否有彈性。這類設計 (使用 [Well-Architected Framework](#) 的最佳實務建立) 會根據關鍵相依性、過去的事件、已知問題和合規性要求來考量風險。列出要用來維護彈性的每個設計元素，及其依設計要減輕的錯誤。如需有關建立此類清單的詳細資訊，請參閱 [Operational Readiness Review 白皮書](#)，指導您如何建立程序以防止先前事件再次發生。Failure Modes and Effects Analysis (FMEA) 程序提供了相關架構，讓您執行失敗的元件層級分析，並說明失敗對於工作負載有何影響。Adrian Cockcroft 在 [Failure Modes and Continuous Resilience](#) 中詳盡地闡述了 FMEA。

2. 指派每個錯誤的優先順序。

請從粗略的分類開始著手，例如高、中或低。若要評估優先順序，請考量錯誤的頻率，以及失敗對整體工作負載的影響。

考量特定錯誤的頻率時，請分析此工作負載過去的資料 (如果可用)。如果無法使用，請使用在類似環境中執行的其他工作負載所包含的資料。

考量特定錯誤的影響時，錯誤的範圍愈大，通常影響就愈大。另請考量工作負載的設計和用途。例如，對執行資料轉換和分析的工作負載而言，存取來源資料存放區的能力至關重要。在此情況下，您應優先執行存取錯誤以及限流存取和延遲注入的試驗。

事故後分析是您了解失敗模式的頻率與影響的理想資料來源。

請使用指派的優先順序，決定要先以哪些錯誤進行試驗，以及要以何種順序開發新的故障注入試驗。

3. 對於您所執行的每個試驗，均應依循下圖中的混沌工程和連續彈性飛輪操作。



混沌工程和連續彈性飛輪，採用 Adrian Hornsby 的科學方法。

- a. 將穩定狀態定義為顯示出正常行為之工作負載的某種可測量輸出。

工作負載的運作若可靠且符合預期，就會呈現穩定狀態。因此，在定義穩定狀態前，請先驗證工作負載的運作狀態良好。穩定狀態不一定表示在錯誤發生時完全不會影響到工作負載，因為有特

定百分比的錯誤可能會在可接受的限制內。穩定狀態是您在試驗期間將觀察到的基準，如果您在下一步定義的假設未符合預期，就會凸顯異常。

例如，支付系統的穩定狀態可定義為 300 TPS、成功率 99%、且來回時間為 500 毫秒的處理。

b. 形成關於工作負載將如何回應錯誤的假設。

良好的假設奠基於工作負載應如何減輕錯誤以維護穩定狀態。假設指出，在發生特定類型的錯誤時，系統或工作負載將持續保有穩定狀態，因為工作負載設有特定緩解機制。特定類型的錯誤和緩解機制應指定於假設中。

以下是可用於假設的範本 (但也接受其他措辭)：

 Note

如果發生####，#####工作負載將#####，以維持#####。

例如：

- 若 Amazon EKS 節點群組中有 20% 的節點遭到關閉，Transaction Create API 會在 100 毫秒以內繼續提供第 99 個百分位數的請求 (穩定狀態)。Amazon EKS 節點將在五分鐘內復原，而 Pod 將在試驗起始後的八分鐘內進入排程並處理流量。提醒將在三分鐘內引發。
- 單一 Amazon EC2 執行個體失敗發生時，訂單系統的 Elastic Load Balancing 運作狀態檢查將使 Elastic Load Balancing 僅將請求傳送至其餘運作狀態良好的執行個體，而 Amazon EC2 Auto Scaling 會取代失敗的執行個體，將伺服器端 (5xx) 錯誤的增量維持在 0.01% 以內 (穩定狀態)。
- 主要 Amazon RDS 資料庫執行個體失敗時，供應鏈資料收集工作負載將進行容錯移轉並連線至待命 Amazon RDS 資料庫執行個體，以維持不到 1 分鐘的資料庫讀取或寫入錯誤 (穩定狀態)。

c. 藉由注入錯誤來執行試驗。

試驗依預設應處於安全模式，並獲得工作負載的容許。如果您確知工作負載將失敗，請不要執行試驗。混沌工程應該用來尋找已知的未知或未知的未知。已知的未知是您知道但不完全理解的事情，未知的未知是您不知道也不完全理解的事情。對您確知已失效的工作負載執行試驗，將不會為您帶來新的見解。試驗應經過審慎規劃、具有明確的影響範圍，並且提供在非預期的錯亂發生時可供套用的回復機制。如果您的盡職調查顯示工作負載應可承受試驗，請繼續執行試驗。有數種選項可用來注入錯誤。對於 AWS 上的工作負載，[AWS FIS](#) 提供許多預先定義的錯誤模擬，稱為[動作](#)。您也可以定義在 AWS FIS 中執行的自訂動作 (使用 [AWS Systems Manager 文件](#))。

我們不鼓勵使用自訂指令碼來執行混沌試驗，除非指令碼有能力理解工作負載目前的狀態、能夠發出日誌，並且提供回復機制和停止條件 (若情況允許)。

支援混沌工程的有效架構或工具集，應追蹤試驗目前的狀態、發出日誌，並提供回復機制以支援受控制的試驗執行。請先從 AWS FIS 這類已建立的服務著手，以便能以明確定義的範圍執行試驗，並且有安全機制可在試驗導入非預期的錯亂時回復試驗。要了解更多有關使用 AWS FIS 的實驗，請參閱 [Resilient and Well-Architected Apps with Chaos Engineering lab](#)。此外，[AWS Resilience Hub](#) 會分析您的工作負載，並建立可供您選擇在 AWS FIS 中實作並執行的試驗。

Note

對於每一項試驗，您都應明確了解其範圍與影響。我們建議，錯誤應先在非生產環境中模擬，再於生產環境中執行。

在可行的情況下，實驗應該使用 [Canary 部署](#) 在實際負載下在生產環境中執行，這可加速控制和實驗系統部署。在非尖峰時段執行試驗是很好的做法，可以減少首次在生產環境中試驗時的潛在影響。此外，如果使用實際的客戶流量會伴隨太高的風險，您可以對控制和試驗部署使用生產基礎架構上的綜合流量，來執行試驗。無法使用生產環境時，請在盡可能接近生產環境的生產前環境中執行試驗。

您必須建立防護機制並加以監控，以確定試驗不會超出可接受的限制而影響到生產流量或其他系統。請建立停止條件，以在試驗達到您定義的防護機制指標閾值時，將試驗停止。其中應包括工作負載的穩定狀態指標，以及您對其注入錯誤的元件所適用的指標。[綜合監測](#) (也稱為使用者 Canary)，是您在一般情況下應納入作為使用者代理的指標之一。[AWS FIS 的停止條件](#) 被視為試驗範本的一部分受到支援，每個範本最多可啟用五個停止條件。

混沌的準則之一，是盡可能縮小試驗的範圍與影響：

儘管容許某些短期負面影響是必要的，但混沌工程師有責任和義務將試驗的副作用控制在最低限度。

驗證範圍和潛在影響的方法之一，是先在非生產環境中執行試驗，驗證停止條件的閾值在試驗期間會依預期啟動，且有可觀測性會捕捉例外狀況，而不是直接在生產環境中試驗。

執行故障注入試驗時，請驗證所有的責任方都會及時獲得通知。請與營運團隊、服務可靠性團隊和客戶支援等適當的團隊通訊，讓他們知道試驗將於何時執行，且預期會有何情況。請為這些團隊提供通訊工具，以便他們在試驗執行期間發現任何不利影響時發出通知。

您必須將工作負載及其基礎系統還原為原始的已知良好狀態。工作負載的彈性設計通常具有自愈能力。但某些錯誤設計或失敗的試驗可能會使您的工作負載處於非預期的失敗狀態。試驗結束時，您必須察覺到這一點，並還原工作負載和系統。透過 AWS FIS，您可以在動作參數內設定回復組態（也稱為後置動作）。後置動作會將目標回復為動作執行前原有的狀態。無論是自動（例如使用 AWS FIS）還是手動，這些後續動作皆應為程序手冊的一部分，以說明如何偵測及處理失敗。

d. 驗證假設。

[混沌工程的原則](#)提供了下列關於如何驗證工作負載穩定狀態的指引：

著重於可測量的系統輸出，而不是系統的內部屬性。這類輸出在一段時間內的測量，會構成系統穩定狀態的代理。整體系統的輸送量、錯誤率和延遲百分位數，全都可能成為呈現穩定狀態行為的相關指標。著重於試驗期間的系統行為模式，混沌工程會驗證系統是否可運作，而非試著驗證其運作情形。

在先前的兩個範例中，我們納入了伺服器端 (5xx) 錯誤的增量低於 0.01% 的穩定狀態指標，以及資料庫讀取和寫入錯誤不到一分鐘的穩定狀態指標。

5xx 錯誤是工作負載的用戶端在失敗模式下將直接經歷的結果，因此可說是良好的指標。資料庫錯誤測量是錯誤的直接產物，因此有其效用，但應同時輔以用戶端影響測量，例如失敗的客戶請求或用戶端遇到的錯誤。此外，請在工作負載的用戶端直接存取的任何 API 或 URI 納入綜合監控（也稱為使用者 Canary）。

e. 改善工作負載設計的彈性。

如果未維持穩定狀態，請採用 [AWS Well-Architected 可靠性支柱](#) 的最佳實務，調查如何改進工作負載設計來緩解故障。可以在 [AWS 建置者資料中心](#) 中找到其他指引和資源，其中包含有關如何[改善運作狀態檢查](#)或[在應用程式碼中透過輪詢進行重試](#)等文章。

這些變更實作完成後，請再次執行試驗（在混沌工程飛輪中以虛線表示）以判斷其有效性。若驗證步驟指出假設成立，則工作負載將處於穩定狀態，且週期會繼續。

4. 請定期執行試驗。

混沌試驗是一個週期，而試驗應被視為混沌工程的一部分定期執行。當工作負載符合試驗的假設後，即應將試驗自動化，以將其視為 CI/CD 管道的迴歸部分持續執行。要了解如何執行此操作，請參閱有關[如何使用 AWS CodePipeline 執行 AWS FIS 實驗](#)的部落格。這個關於 [CI/CD 管道中的經常性 AWS FIS 實驗](#)的實驗室可讓您親手操作。

故障注入試驗也是演練日的一部分（請參閱 [REL12-BP05 定期進行演練日](#)）。演練日會模擬失敗或事件，以驗證系統、程序和團隊的應變。目的是實際執行在異常事件發生時團隊將要執行的動作。

5. 擷取並儲存試驗結果。

故障注入試驗的結果必須擷取並保存。請納入所有必要資料 (例如時間、工作負載和條件)，以便後續能分析試驗結果和趨勢。舉例來說，結果可包括儀表板的螢幕擷取畫面、指標的資料庫產生的 CSV 傾印，或是試驗中的事件與觀察的手寫記錄。[使用 AWS FIS 試驗日誌記錄](#)可以是此資料擷取的一部分。

資源

相關的最佳實務：

- [REL08-BP03 將彈性測試整合為部署的一部分](#)
- [REL13-BP03 測試災難復原實作以驗證實作](#)

相關文件：

- [什麼是 AWS Fault Injection Service ?](#)
- [什麼是 AWS Resilience Hub ?](#)
- [混沌工程的原則](#)
- [混沌工程：規劃您的第一個試驗](#)
- [彈性工程：學習接受故障](#)
- [混沌工程案例](#)
- [避免分散式系統的備用](#)
- [混沌試驗的 Canary 部署](#)

相關影片：

- [AWS re:Invent 2020：使用混沌工程測試彈性 \(ARC316\)](#)
- [AWS re:Invent 2019：透過混沌工程提升彈性 \(DOP309-R1\)](#)
- [AWS re:Invent 2019：在無伺服器環境中執行混沌工程 \(CMY301\)](#)

相關工具：

- [AWS Fault Injection Service](#)
- AWS Marketplace：[Gremlin 混沌工程平台](#)

- [Chaos Toolkit](#)
- [Chaos Mesh](#)
- [Litmus](#)

REL12-BP05 定期進行演練日

舉辦演練日，以定期執执行程序來回應影響工作負載的事件和損害。讓可能負責處理正式作業案例的相同團隊參與。這些練習有助於強制執行措施，以防止使用者因正式作業事件而受到影響。若您在實際情況下練習回應程序，就可以在實際事件發生之前識別和解決任何差距或弱點。

演練日會在類似正式作業環境中模擬事件，以測試系統、程序和團隊回應。目的是實際執行在異常事件發生時，團隊會執行的動作。這些練習可幫助您了解何處有改善空間，並能協助發展組織處理事件和損害的經驗。這些練習應該定期進行，讓您的團隊能夠培養應對的肌肉記憶。

演練日可讓團隊更有信心地處理實際發生的事件。熟悉演練的團隊更能夠快速偵測和回應各種案例。這可大幅改善整備度和彈性狀態。

預期成果：您定期一致地進行彈性演練日。這些演練日會做為預期的正常業務營運。您的組織培養出整備文化，當正式作業問題發生時，您的團隊已準備好有效地回應、有效率地解決問題，並減輕對客戶的影響。

常見的反模式：

- 您記載您的程序，但未曾進行演練。
- 您未讓業務決策者參與測試練習。
- 您進行演練日，但未通知所有相關利害關係人。
- 您只專注於技術失敗，但未納入業務利害關係人。
- 您未將演練日所學到的經驗納入復原程序中。
- 您在發生失敗或錯誤時責怪團隊。

建立此最佳實務的優勢：

- 增強回應技能：在演練日，團隊會在模擬事件的過程中練習其職責並測試其溝通機制，從而在實際情境中建立更協調且更有效率的回應。
- 識別和解決相依性：複雜的環境通常涉及各種系統、服務和元件之間繁複的相依性。演練日可協助您識別並解決這些相依性，以及確認您的關鍵系統和服務確實涵蓋在您的執行手冊程序內，並且能夠及時向上擴展或復原。

- 培養彈性的文化：演練日有助於培養組織內的彈性思維。當您納入跨職能團隊和利害關係人時，這些練習就能提升整個組織對彈性重要性的意識、協作和同理。
- 持續改進和適應：定期演練日可協助您持續評估和調整彈性策略，以便在面對不斷變化的情況時，保持關聯性和有效性。
- 提高對系統的信心：成功的演練日可協助您建立對系統承受中斷並從中復原能力的信心。

未建立此最佳實務時的曝險等級：中

實作指引

設計並實作必要的彈性措施後，請舉行演練日來確認正式作業時，一切是否如規畫運作。特別是第一次的演練日，應納入所有團隊成員，且所有利害關係人和參與者都應事先收到有關日期、時間和模擬案例的通知。

在演練日，參與的團隊會根據既定程序模擬各種事件和可能的案例。參與者會密切監控和評估這些模擬事件的影響。如果系統依設計運作，則自動偵測、擴展和自我修復機制應會啟動，而且對使用者的影響極少或無影響。如果團隊觀察到任何負面影響，則會透過自動化的方式或適用的執行手冊中記錄的手動介入方式，來復原測試並補救找到的問題。

為了持續改善彈性，記錄並納入學到的經驗至關重要。此程序是一種回饋循環，採取系統化的方式從演練日獲得洞察，並利用它們來增強系統、程序和團隊功能。

為了協助您重現系統元件或服務可能意外失敗的真實案例，請將模擬錯誤植入演練日練習中。團隊可以測試其系統的彈性和容錯能力，並在受控的環境中模擬其事件回應和復原程序。

在 AWS 中，您的演練日可以使用基礎設施即程式碼，透過正式作業環境的複本來執行。透過此程序，您可以在類似正式作業環境的安全環境中進行測試。考慮使用 [AWS Fault Injection Service](#) 來建立不同的失敗案例。使用 [Amazon CloudWatch](#) 和 [AWS X-Ray](#) 等服務來監控演練日的系統行為。使用 [AWS Systems Manager](#) 來管理和執程序手冊，並使用 [AWS Step Functions](#) 來協調週期性演練日工作流程。

實作步驟

- 擬訂演練日計畫：擬訂結構化的計畫，以定義演練日的頻率、範圍和目標。規畫和執行這些練習時，讓主要利害關係人和主題專家參與。
- 準備演練日：
 1. 指定做為演練日重點的重要關鍵業務服務。將支援這些服務的人員、程序和技術造冊並對應。

2. 設定演練日的流程，並讓參與的團隊準備好參與活動。備妥自動化服務，以模擬規劃的案例並執行適當的復原程序。[AWS Fault Injection Service](#)、[AWS Step Functions](#) 和 [AWS Systems Manager](#) 等 AWS 服務可協助您自動進行演練日的各個層面，例如植入錯誤和啟動復原動作。
- 執行模擬：在演練日當天，執行規劃的案例。觀察並記錄人員、程序和技術對模擬事件的反應。
- 進行練習後檢討：在演練日結束後，舉行回顧會議來檢討學到的經驗。識別需要改進的層面，以及改善營運彈性所需的任何動作。記錄您的調查結果，並追蹤任何必要的變更，以增強彈性策略並完成準備。

資源

相關的最佳實務：

- [REL12-BP01 使用程序手冊調查失敗](#)
- [REL12-BP04 使用混沌工程測試彈性](#)
- [OPS04-BP01 識別關鍵績效指標](#)
- [OPS07-BP03 使用執行手冊執执行程序](#)
- [OPS10-BP01 使用程序進行事件、事故和問題管理](#)

相關文件：

- [什麼是 AWS GameDay？](#)

相關影片：

- [AWS re：Invent 2023 - 寓教於樂：Amazon 如何將彈性擴展到新境界](#)

相關範例：

- [AWS 研討會 - 駕馭風暴：充分利用彈性系統的受控混沌](#)
- [打造自己的演練日以支援營運彈性](#)

REL 13. 如何規劃災難復原 (DR)？

備妥備份和冗餘工作負載元件是 DR 策略的開始。[RTO 和 RPO](#) 是您還原工作負載的目標。根據業務需求設定這些目標。實作策略以滿足這些目標，考量工作負載資源和資料的位置和功能。發生中斷的可能性和復原成本也是重要因素，可反映為工作負載提供災難復原的商業價值。

最佳實務

- [REL13-BP01 定義停機和資料遺失的復原目標](#)
- [REL13-BP02 使用定義的復原策略來滿足復原目標](#)
- [REL13-BP03 測試災難復原實作以驗證實作](#)
- [REL13-BP04 管理 DR 站點或區域的組態偏移](#)
- [REL13-BP05 自動化回復](#)

REL13-BP01 定義停機和資料遺失的復原目標

故障可能會以多種形式影響您的業務。首先，故障可能會導致服務中斷 (停機時間)。其次，故障可能會導致資料遺失、不一致或過時。為了引導您回應故障並從故障中復原，請為每個工作負載定義復原時間點目標 (RTO) 和復原點目標 (RPO)。復原時間點目標 (RTO) 是服務中斷與服務還原之間的可接受延遲上限。復原點目標 (RPO) 是從上次資料復原點之後起算，可接受的最長時間。

預期成果：每個工作負載都根據技術考量和業務影響指定了 RTO 和 RPO。

常見的反模式：

- 您尚未指定復原目標。
- 您選取任意復原目標。
- 您選取的復原目標過於寬鬆且不符合業務目標。
- 您尚未評估停機時間和資料遺失的影響。
- 您選取的復原目標不切實際，例如零復原時間或零資料損失，這對於您的工作負載組態而言可能無法實現。
- 您選取的復原目標比實際業務目標更嚴格。這會強制進行比工作負載所需更昂貴和更複雜的復原實作。
- 您選取的復原目標與相依的工作負載不相容。
- 您未將法規和合規要求納入考量。

建立此最佳實務的優勢：設定工作負載的 RTO 和 RPO 之後，您就可以根據業務需求訂定清楚且可衡量的復原目標。設定這些目標之後，您就可以制定專為實現這些目標所打造的災難復原 (DR) 計畫。

未建立此最佳實務時的曝險等級：高

實作指引

建構對照表或工作表，以協助引導您的災難復原規劃。在對照表中，根據個別的業務影響 (例如嚴重、高、中和低) 以及要做為個別目標的相關聯 RTO 和 RPO，建立不同的工作負載類別或分級。下列對照表為您提供可遵循的範例 (請注意，您的 RTO 和 RPO 值可能不同)：

災難復原方法						
		復原點目標				
		< 1 分鐘	< 1 小時	< 6 小時	< 1 天	+ 1 天
復原時間目標	< 10 分鐘	嚴重	嚴重	高	中	中
	< 2 小時	嚴重	高	中	中	低
	< 8 小時	高	中	中	低	低
	< 24 小時	中	中	低	低	低
	24 + 小時	中	低	低	低	低

災難復原對照表範例

針對每個工作負載，調查並了解停機時間和資料遺失對業務造成的影響。影響通常會隨著停機時間和資料遺失而增強，但影響的程度會根據工作負載類型而有所不同。例如，長達一小時的停機時間可能會產生低程度的影響，但隨後影響可能會迅速增強。影響可能有多種形式，包括財務影響 (例如收益損失)、聲譽影響 (包括失去客戶信任)、營運影響 (例如發不出薪資或生產力降低)，以及法規風險。完成後，將工作負載指派至適當的分級。

當您分析故障的影響時，請考慮下列問題：

1. 在對業務造成無法接受的影響之前，可以容忍的最長工作負載停機時間是多久？
2. 工作負載中斷會對業務產生多大程度的影響以及何種影響？請考慮各種影響，包括財務、聲譽、營運和法規。
3. 在對業務造成無法接受的影響之前，可以容忍的遺失或無法復原的最大資料量是多少？
4. 是否可從其他來源重新建立遺失的資料 (也稱為衍生資料)？如果可以，請同時考量用來重新建立工作負載資料的所有來源資料的 RPO。

5. 此工作負載所依賴的工作負載 (下游) 的復原目標和可用性期望為何？工作負載的目標必須在其下游相依的工作負載具有復原能力的情況下才能達成。請考慮可能改善此工作負載復原能力的下游相依性因應措施或緩解措施。
6. 依賴此工作負載的工作負載 (上游) 的復原目標和可用性期望為何？上游工作負載目標可能要求此工作負載具有比首次出現更嚴格的復原能力。
7. 是否取決於事件類型而有不同的復原目標？例如，根據事件影響的是可用區域或整個區域而定，您可能有不同的 RTO 和 RPO。
8. 您的復原目標是否在一年中的某些事件或某段時間改變？例如，您可能在節日購物季節、體育賽事、特別促銷和新產品推出時，採用不同的 RTO 和 RPO。
9. 復原目標如何與您可能擁有的任何業務範圍和組織災難復原策略保持一致？
10. 是否需要考慮任何法律或合約後果？例如，您是否在合約方面有義務要提供具有特定 RTO 或 RPO 的服務？若未履行義務，可能會受到哪些懲罰？
11. 是否需要維護資料完整性以符合法規或合規要求？

下列工作表可協助您評估每個工作負載。您可以修改此工作表以滿足您的特定需求，例如新增其他問題。

步驟 2: 主要問題	是否適用於工作負載？	工作負載 RTO	工作負載 RPO	RTO 調整。	RPO 調整。	簡介
[1] 工作負載可以關閉的最長時間						以開始中斷到復原的時間進行測量
[2] 可以遺失的資料數量上限						以自從上次已知良好的可還原資料集後的時間進行測量
[3a] 上游相依性						輸入最嚴格的上游復原目標
[3b] 下游相依性						輸入最不嚴格的下流復原目標
[3a] 達成一致的上游相依性						如果上游值小於目前值，而下游值更大，則使用相依性來達成一致，
[3b] 達成一致的下游相依性						並在這裡輸入達成一致的值
[3] 相依性						請降低值以符合上游相依性，或根據下游相依性功能提高這些值
步驟 2: 其他問題						指出問題是否適用。如果不適用，則略過它
基底 RTO/RPO						將上面的 RTO 和 RPO 值帶至這裡
[4] 中斷類型	[] Y / [] N					為具有最嚴格需求的事件類型輸入復原目標
[5] 特定時間型目標	[] Y / [] N					為具有最嚴格需求的時間輸入復原目標
[6] 顛覆客戶	[] Y / [] N					透過圖表以停機時間或資料遺失的函數表示受影響的客戶。使用該函數，根據客戶影響輸入最大允許的 RTO 和 RPO
[7] 信譽影響	[] Y / [] N					與企業合作，根據對信譽的影響決定最大 RTO 和 RPO
[8] 營運影響	[] Y / [] N					根據營運影響輸入最大 RTO 和 RPO
[9] 組織遵循	[] Y / [] N					根據 LOB 和組織需求輸入此類型的最大工作負載 RTO 和 RPO
[10] 合約義務	[] Y / [] N					根據合約義務輸入最大 RTO 和 RPO
[11] 法規合規	[] Y / [] N					根據適用的法規合規輸入最大 RTO 和 RPO
以其他問題為基礎的目標						從 Q' s 4-11 取得並在這裡輸入最小值 (更嚴格的值)
調整後的目標						如果無法滿足上行的目標，請與利害關係人合作放寬限制，並在這裡輸入新的最小值
調整後的 RTO/RPO						輸入基底 RPO/RTO 值或調整後的目標，以較低者為準
步驟 3						
對應至預先定義的類別或層級						向下調整這兩個值 (更嚴格) 以與最接近的定義層級一致

工作表

實作步驟

1. 確定負責每個工作負載的業務利害關係人和技術團隊，並與他們互動。
2. 在組織中建立工作負載影響的關鍵性類別或分級。範例類別包括：嚴重、高、中和低。為每個類別選擇合乎您業務目標和需求的 RTO 和 RPO。
3. 將您在上一個步驟中建立的其中一個影響類別指派給每個工作負載。若要決定工作負載對應至類別的方式，請考慮工作負載對業務的重要性，以及中斷或資料遺失的影響，並利用上述問題來引導您。這樣就能找到每個工作負載的 RTO 和 RPO。
4. 考慮上一個步驟中為每個工作負載確定的 RTO 和 RPO。讓工作負載的業務和技術團隊參與，以判斷是否應調整目標。例如，業務利害關係人可判斷是否需要更嚴格的目標。或者，技術團隊可以判斷應修改目標，以透過可用的資源和技術限制來實現目標。

資源

相關的最佳實務：

- [REL09-BP04 定期執行資料復原以驗證備份的完整性和程序](#)
- [REL12-BP01 使用程序手冊調查失敗](#)
- [REL13-BP02 使用定義的復原策略來滿足復原目標](#)
- [REL13-BP03 測試災難復原實作以驗證實作](#)

相關文件：

- [AWS 架構部落格：災難復原系列](#)
- [AWS 上工作負載的災難復原：在雲端中復原 \(AWS 白皮書\)](#)
- [使用 AWS Resilience Hub 管理彈性政策](#)
- [APN 合作夥伴：可協助進行災難復原的合作夥伴](#)
- [AWS Marketplace：可用於災難復原的產品](#)

相關影片：

- [AWS re:Invent 2018：適用於多區域主動-主動應用程式的架構模式](#)
- [上工作負載的災難復原AWS](#)

REL13-BP02 使用定義的復原策略來滿足復原目標

定義一個符合工作負載復原目標的災難復原 (DR) 策略。選擇如下策略：備份與還原；待命 (主動/被動)；或是主動/主動。

預期成果：對於每個工作負載，都有已定義並實作的 DR 策略，可讓工作負載實現災難復原目標。工作負載之間的 DR 策略會利用可重複使用模式 (例如上述策略)，

常見的反模式：

- 針對具有類似 DR 目標的工作負載實作不一致的復原程序。
- 災難發生時臨時實作 DR 策略。
- 沒有災難復原的計劃。
- 復原期間依賴控制平面操作。

建立此最佳實務的優勢：

- 使用定義的復原策略可讓您使用常用的工具和測試程序。
- 使用定義的復原策略可改善在團隊之間分享知識，並更輕鬆地在他們擁有的工作負載上實作 DR。

未建立此最佳實務時的風險暴露等級：高。若沒有事先規劃、實作和測試災難復原策略，您就不可能在發生災難時實現復原目標。

實作指引

如果您的主要位置變成無法執行工作負載，則災難復原策略會依賴在復原站點中支持您工作負載的能力。最常見的復原目標為 RTO 和 RPO，如 [REL13-BP01 定義停機和資料遺失的復原目標](#) 中所討論。

單一 AWS 區域內跨多個可用區域 (AZ) 的 DR 策略可以緩解火災、洪水和重大停電等災難事件。如果需要實作保護，以防範阻止您的工作負載在給定 AWS 區域中執行且不太可能發生的事件，您可以使用一個使用多個區域的 DR 策略。

在跨多個區域架構 DR 策略時，您應該選擇下列其中一個策略。其按成本和複雜性的升序以及 RTO 和 RPO 的降序排列。復原區域是指用於工作負載的主要區域以外的 AWS 區域。

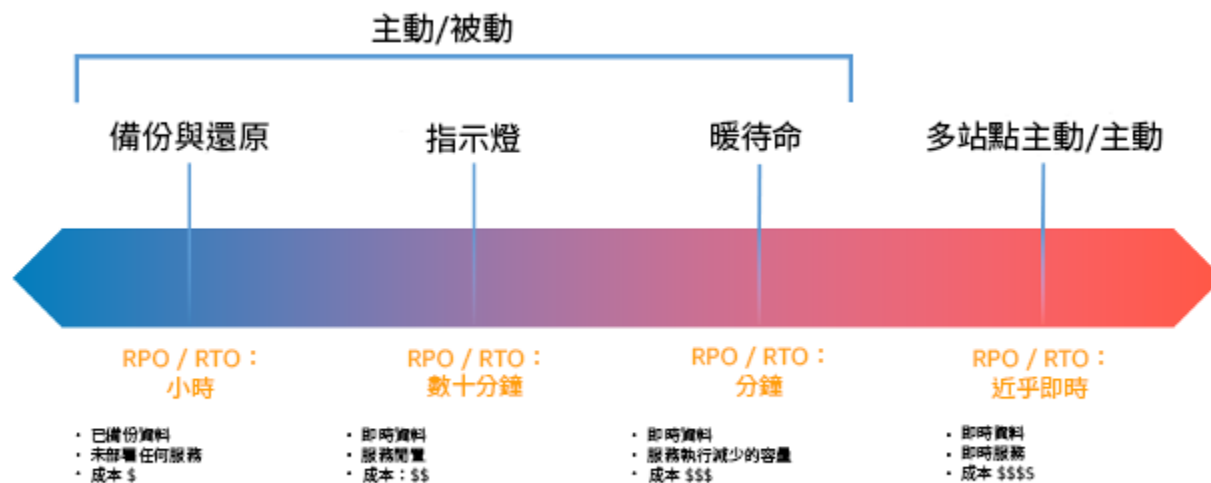


圖 17：災難復原 (DR) 策略

- **備份與還原** (RPO 以小時為單位，24 小時以內的 RTO)：將您的資料和應用程式備份至復原區域。使用自動或連續備份將允許時間點復原 (PITR)，在某些情況下可以將 RPO 降低至 5 分鐘。如果發生災難，您將部署您的基礎設施 (使用基礎設施架構即程式碼來減少 RTO)、部署您的程式碼，並還原備份的資料以從復原區域中的災難中復原。
- **指示燈** (RPO 以分鐘為單位，RTO 以十分鐘為單位)：在復原區域佈建核心工作負載基礎設施的副本。將您的資料複製到復原區域並在該處建立其備份。支援資料複製和備份所需的資源 (例如資料庫和物件儲存) 始終處於開啟狀態。其他元素 (例如應用程式伺服器或無伺服器運算) 未部署，但可在需要時使用必要的組態和應用程式碼建立。
- **暖待命** (RPO 以秒為單位，RTO 以分鐘為單位)：維持始終在復原區域中執行，規模縮減但功能完整的工作負載版本。業務關鍵系統會完全複製且持續開啟，但叢集會縮小。資料會被複製並存在於復原區域中。當需要復原時，系統會迅速擴展以處理生產負載。暖待命的縱向擴增越多，對 RTO 和控制平面的依賴就越低。當完全擴展時，這稱為熱待命。
- **多區域 (多站台) 主動-主動式** (RPO 接近零，RTO 可能為零)：您的工作負載會部署到多個 AWS 區域，並可主動為其流量提供服務。此策略需要您跨區域同步資料。必須避免或處理在兩個不同區域複本中寫入同一記錄所引起的可能衝突，這可能很複雜。資料複製對於資料同步很有用，而且可以保護您防範某些類型的災難，但它不能保護您防範資料損毀或破壞，除非您的解決方案也包括時間點復原的選項。

Note

指示燈和暖待命之間的差異有時可能很難理解。這兩者都在您的復原區域中包含一個環境，其中具有主要區域資產的副本。區別在於，若未先採取額外動作，指示燈無法處理請求，而暖待命可以立即處理流量 (容量層級降低)。指示燈將需要您開啟伺服器，可能會部署額外 (非核心) 基礎設施並向上擴展，而暖待命只需要您向上擴展 (一切都已部署並執行中)。根據您的 RTO 和 RPO 需求在這兩者之間進行選擇。

當成本是一大顧慮時，且想要達到與暖待命策略所定義類似的 RPO 和 RTO 目標，您可以考慮雲端原生解決方案，例如 AWS Elastic Disaster Recovery，它會採取指示燈方法並且提供改善的 RPO 和 RTO 目標。

實作步驟**1. 決定將滿足此工作負載復原需求的 DR 策略。**

選擇 DR 策略是在減少停機時間和資料遺失 (RTO 和 RPO) 與實作策略的成本和複雜性之間進行取捨。您應該避免實作比其所需更嚴格的策略，因為這會產生不必要的成本。

例如，在下圖中，企業已確定其最大允許的 RTO 以及其可以在服務還原策略上花費的限制。鑑於業務目標，DR 策略指示燈或暖待命將同時滿足 RTO 和成本準則。

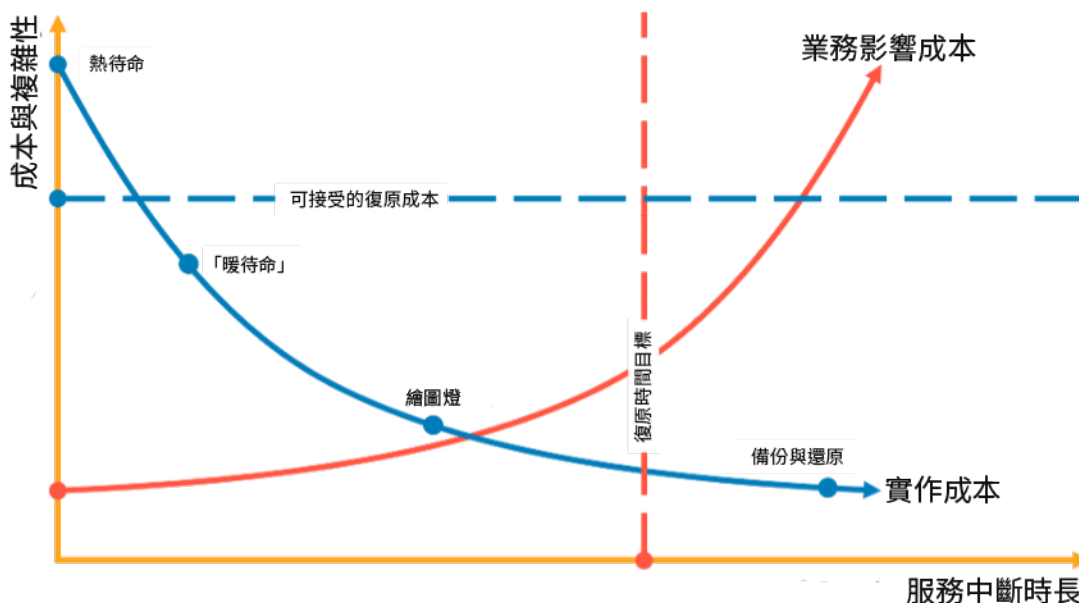


圖 18：根據 RTO 和成本選擇 DR 策略

如需進一步了解，請參閱[業務持續性計劃 \(BCP\)](#)。

2. 審核如何實作所選 DR 策略的模式。

此步驟在於了解您將如何實作所選策略。使用 AWS 區域 做為主要站點和復原站點來解釋這些策略。不過，您也可以選擇使用單一區域內的可用區域，做為您的 DR 策略，這會利用其中多個策略的元素。

在下列步驟中，您可以將策略套用到您的特定工作負載。

備份與還原

備份和還原是最不複雜的實作策略，但還原工作負載需要更多時間和精力，從而導致更高的 RTO 和 RPO。始終對資料進行備份並將其複製到另一個站點 (例如另一個 AWS 區域) 是一種很好的做法。

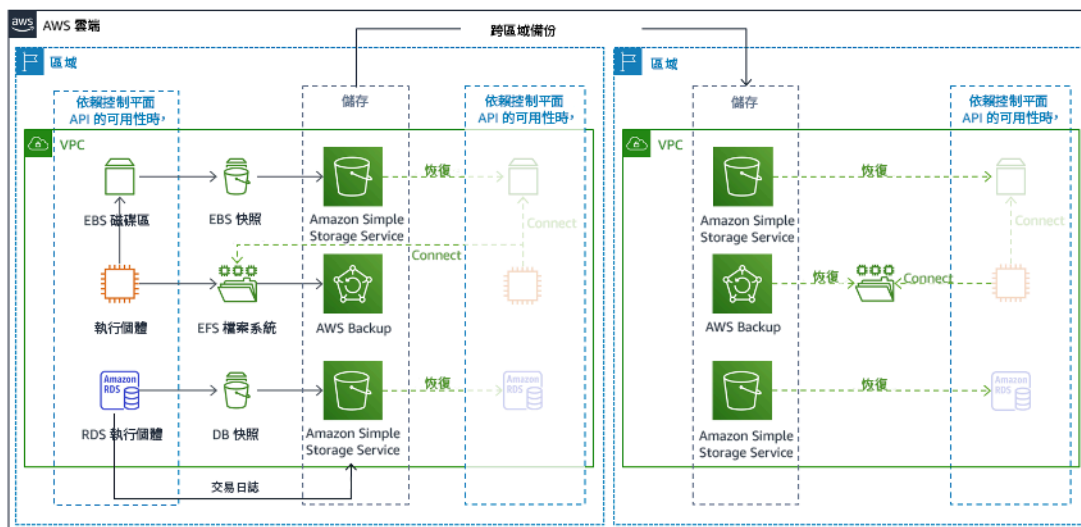


圖 19：備份和還原架構

如需此策略的詳細資訊，請參閱 [Disaster Recovery \(DR\) Architecture on AWS, Part II: Backup and Restore with Rapid Recovery](#)。

指示燈

使用指示燈方法，可以將資料從主要區域複製到復原區域。用於工作負載基礎設施的核心資源會部署在復原區域中，不過，仍需要額外的資源和任何相依性，才能使其成為功能堆疊。例如，在圖 20 中，未部署任何運算執行個體。

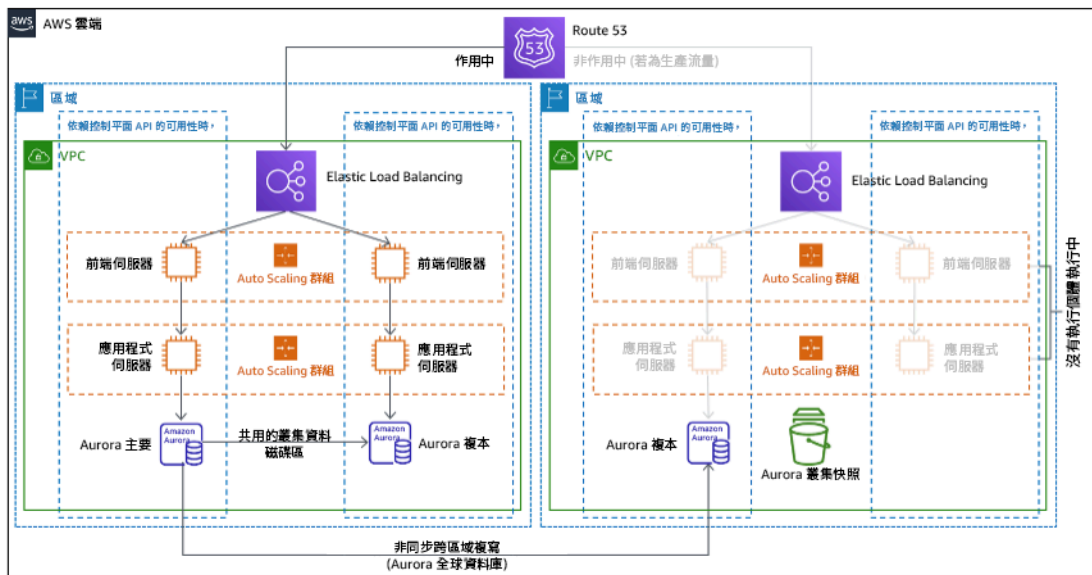


圖 20：指示燈架構

有關此策略的詳細資訊，請參閱 [AWS 上的災難復原 \(DR\) 架構，第 III 部分：指示燈和暖待命](#)。

暖待命。

暖待命方法包括確保在另一個區域中有規模縮減但功能完整的生產環境副本。這種方法擴充了指示燈概念並減少了復原時間，因為您的工作負載始終在另一個區域中開啟。如果復原區域已部署全部容量，則稱為熱待命。

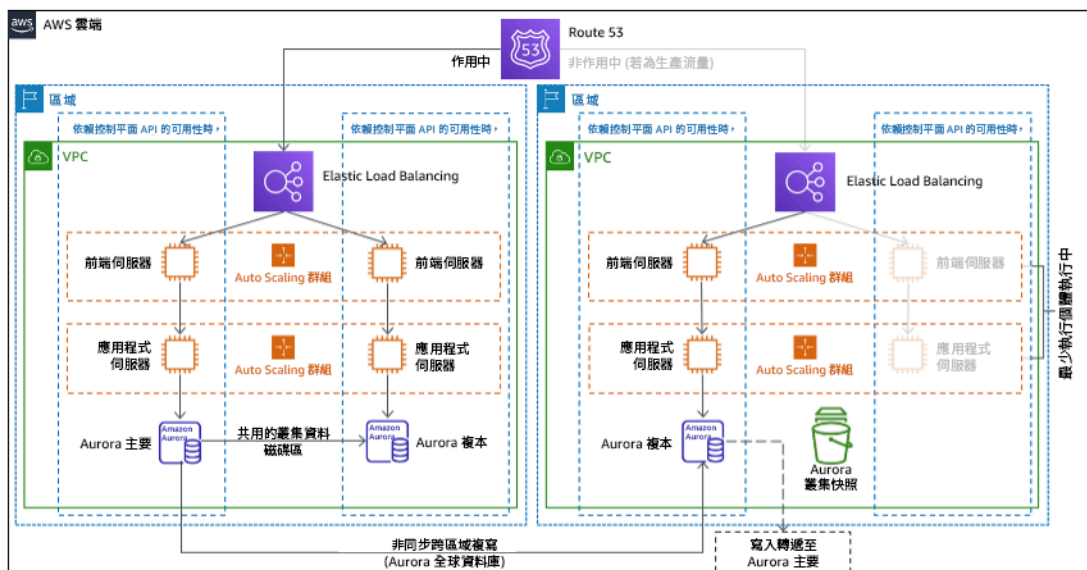


圖 21：暖待命架構

使用暖待命或指示燈需要縱向擴展復原區域中的資源。若要確認容量在需要時可用，請考慮使用 EC2 執行個體的[容量保留](#)。如果使用 AWS Lambda，[佈建並行](#)可以提供執行時期環境，以便立即回應函數的調用。

有關此策略的詳細資訊，請參閱 [AWS 上的災難復原 \(DR\) 架構](#)，第 III 部分：指示燈和暖待命。

多站點主動/主動

作為多站台主動/主動策略的一部分，您可以在多個區域同時執行工作負載。多站點主動/主動會為來自其部署至的所有區域的流量提供服務。客戶可以出於 DR 以外的原因選擇此策略。它可以用來提高可用性，或在將工作負載部署至全球對象（使端點更靠近使用者和/或將本地化的堆疊部署到該區域的對象）時使用它。作為 DR 策略，如果工作負載無法在其部署至的其中一個 AWS 區域中得到支援，則會撤離該區域，而其餘區域則會用來維護可用性。多站點主動/主動是災難復原策略中操作最複雜的策略，因此只有在業務要求有此需要時才應選取它。

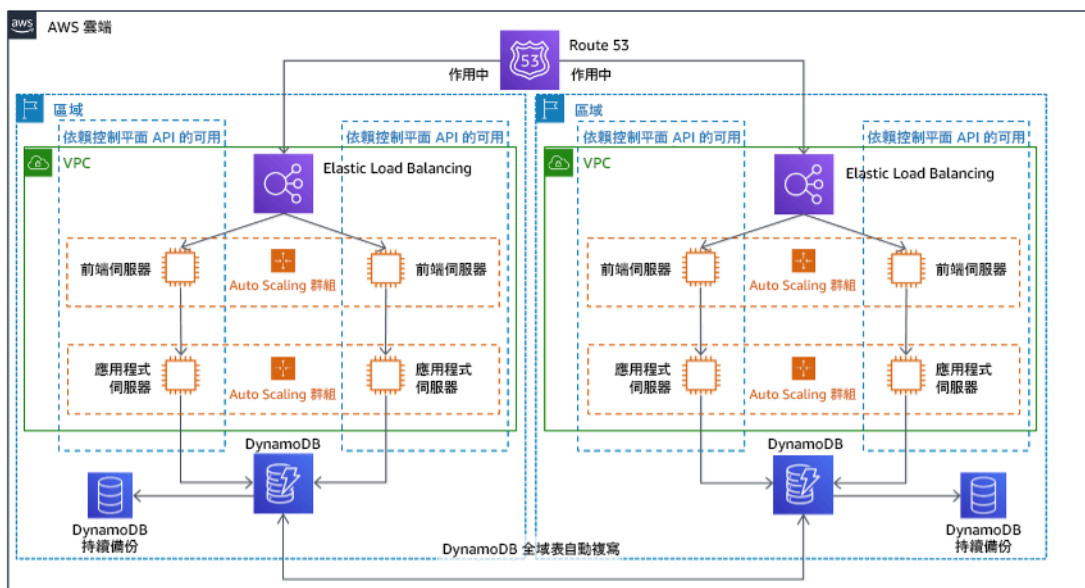


圖 22：多站點主動/主動架構

如需此策略的詳細資訊，請參閱 [Disaster Recovery \(DR\) Architecture on AWS, Part IV: Multi-site Active/Active](#)。

AWS Elastic Disaster Recovery

如果您正在考慮用於災難復原的指示燈或暖待命策略，AWS Elastic Disaster Recovery 可以提供具有改進效益的替代方法。Elastic Disaster Recovery 可以提供類似於暖待命的 RPO 和 RTO 目標，

但維持指示燈的低成本方法。Elastic Disaster Recovery 會使用持續的資料保護功能，實現以秒為單位測量的 RPO，以及以幾分鐘為單位測量的 RTO，將您的資料從主要區域複製到復原區域。只有複寫資料所需的資源會在復原區域中部署，保持低成本，類似於指示燈策略。使用 Elastic Disaster Recovery 時，服務會在容錯移轉或練習過程中啟動時進行協調。

AWS 彈性災難復原 (AWS DRS) 一般架構

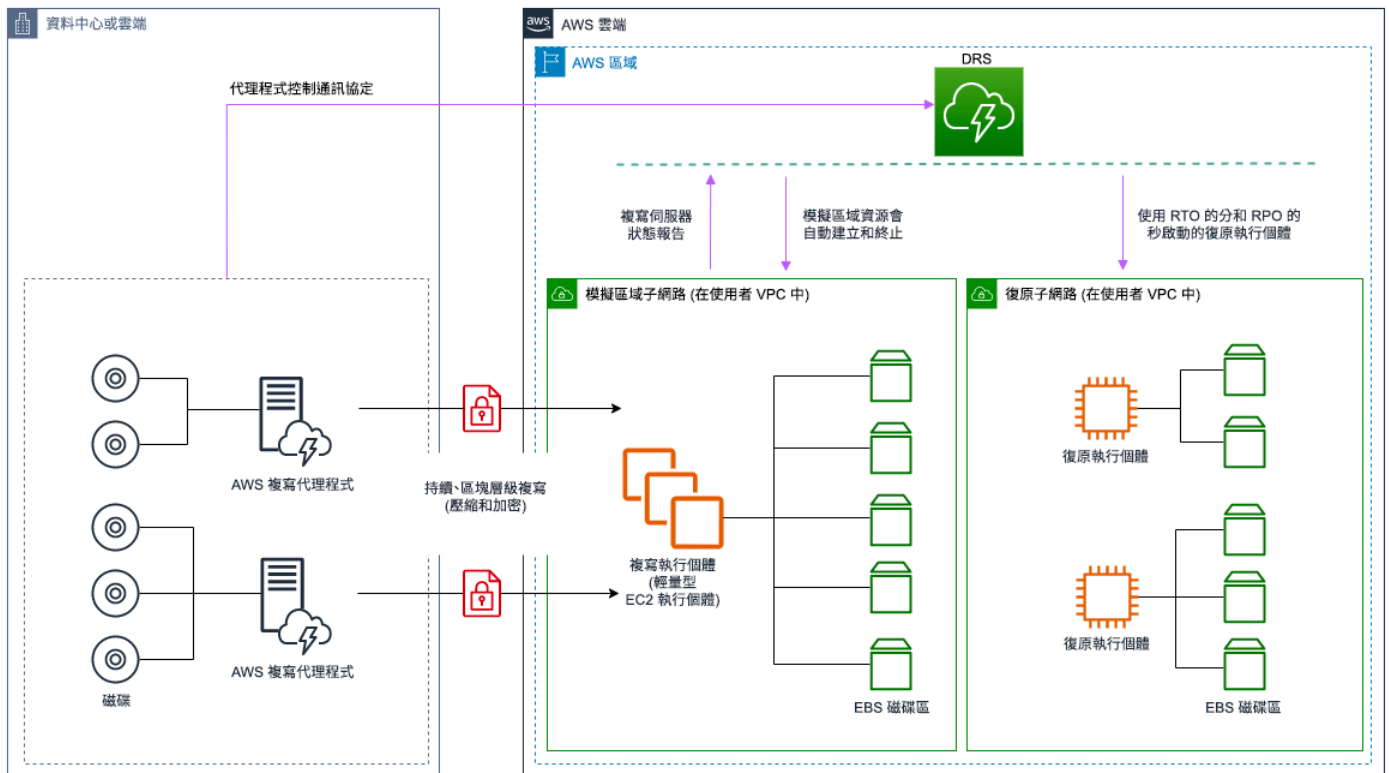


圖 23 : AWS Elastic Disaster Recovery 架構

保護資料的其他實務

使用所有策略時，您還必須緩解資料災難。持續資料複寫可以保護您防範某些類型的災難，但它不能保護您防範資料損毀或破壞，除非您的策略也包括所存放資料的版本控制，或時間點復原的選項。除了複本之外，您還必須備份復原站點中的複寫資料，以建立時間點備份。

在單一 AWS 區域內使用多個可用區域 (AZ)

在單一區域內使用多個 AZ 時，您的 DR 實作會使用上述策略的多個元素。首先，您必須建立高可用性 (HA) 架構，使用多個 AZ，如圖 23 所示。由於 [Amazon EC2 執行個體](#) 和 [Elastic Load Balancer](#) 已在多個 AZ 中部署了資源，因此架構可使用多站點主動/主動方法，以積極處理請求。該

架構也示範熱待命，如果主要 [Amazon RDS](#) 執行個體發生故障 (或 AZ 本身失敗)，則待命執行個體會升級為主執行個體。

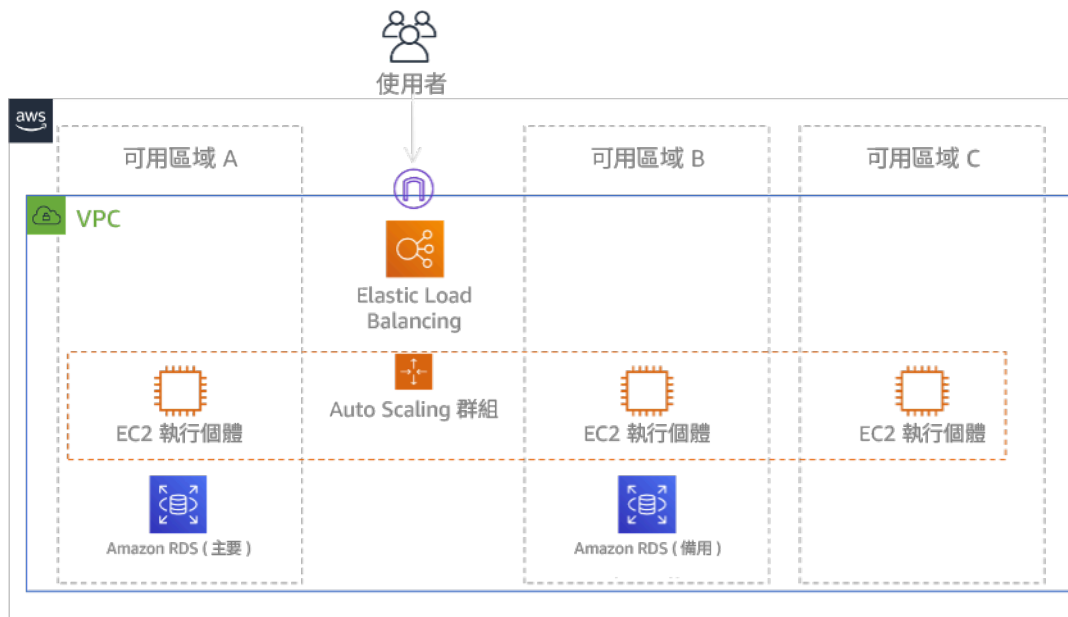


圖 24：多可用區域架構

除了這種 HA 架構之外，您還需要新增執行工作負載所需之所有資料的備份。這對於受限於單一區域的資料 (例如 [Amazon EBS 磁碟區](#) 或 [Amazon Redshift 叢集](#)) 尤其重要。如果 AZ 失敗，您需要將此資料還原至另一個 AZ。可能的話，您也應該將資料備份複製到另一個 AWS 區域，做為額外的保護層。

部落格文章[使用 Amazon 應用程式復原控制器建置高彈性應用程式，第 1 部分：單一區域堆疊](#)中說明了單一區域、多可用區域災難復原的不常見替代方法。在這裡，策略是盡可能地在 AZ 之間保持隔離，就像區域的操作方式一樣。使用這種替代策略，您可以選擇主動/主動或主動/被動方法。

Note

某些工作負載具有法規資料落地要求。如果在目前只有一個 AWS 區域的區域性中，這適用於您的工作負載，則多區域將不適合您的業務需求。異地同步備份策略提供良好的保護，可防範大部分災難。

3. 在容錯移轉之前 (在正常操作期間)，評估工作負載的資源，以及其在復原區域中的組態。

對於基礎設施和 AWS 資源，請使用基礎設施即程式碼 (例如 [AWS CloudFormation](#)) 或諸如 Hashicorp Terraform 之類的第三方工具。若要透過單一操作在多個帳戶和區域中進行部署，可以使用 [AWS CloudFormation StackSets](#)。對於多站點主動/主動和熱待命策略，您的復原區域中部署的

基礎設施具有與您主要區域相同的資源。對於指示燈和暖待命策略，部署的基礎設施將需要額外的動作，才能為生產做好準備。使用 CloudFormation [參數](#)和[條件邏輯](#)，可透過[單一範本](#)來控制已部署的堆疊是處於作用中還是待命。使用 Elastic Disaster Recovery 時，服務會複寫和協調應用程式組態和運算資源的還原。

所有 DR 策略都需要在 AWS 區域中備份資料來源，然後將這些備份複製到復原區域。[AWS Backup](#) 提供集中式檢視，您可以在其中設定、排程和監控這些資源的備份。對於指示燈、暖待命和多站點主動/主動式，您還應該將資料從主要區域複寫到復原區域中的資料資源，例如 [Amazon Relational Database Service \(Amazon RDS\)](#) 資料庫執行個體或 [Amazon DynamoDB](#) 資料表。因此，這些資料資源是即時的，而且可以為復原區域中的請求提供服務。

若要深入了解 AWS 服務如何跨區域運作，請參閱 [Creating a Multi-Region Application with AWS Services](#) 部落格系列。

4. 確定並實作如何讓復原區域在需要時 (在災難事件發生時) 做好容錯移轉的準備。

對於多站點主動/主動，容錯移轉意味著撤離一個區域，並依賴剩餘的主動區域。通常，這些區域已準備好接受流量。對於指示燈和暖待命策略，您的復原動作將需要部署遺漏的資源，例如圖 20 中的 EC2 執行個體，以及任何其他遺漏的資源。

對於上述所有策略，您可能需要提升資料庫的唯讀執行個體，以變成主要讀取/寫入執行個體。

對於備份和還原，從備份還原資料會為該資料建立資源，例如 EBS 磁碟區、RDS 資料庫執行個體和 DynamoDB 資料表。您也需要還原基礎設施和部署程式碼。您可以使用 AWS Backup，還原復原區域中的資料。如需詳細資訊，請參閱[REL09-BP01 識別並備份所有需要備份的資料，或從來源複製資料](#)。重建基礎設施包括建立 EC2 執行個體等資源，以及所需的 [Amazon Virtual Private Cloud \(Amazon VPC\)](#)、子網路和安全群組。您可以將大部分還原程序自動化。若要了解如何操作，請參閱[此部落格文章](#)。

5. 確定並實作如何在需要時 (在災難事件發生時) 重新路由流量以進行容錯移轉。

此容錯移轉作業可以自動或手動啟動。應謹慎使用根據運作狀態檢查或警示自動啟動的容錯移轉，因為不必要的容錯移轉 (誤報) 會產生非可用性和資料遺失等成本。因此通常使用手動啟動的容錯移轉。在此情況下，您仍應將容錯移轉的步驟自動化，讓手動啟動就像按下按鈕一樣簡易。

使用 AWS 服務時，有數個流量管理選項需要考慮。一種選擇是使用 [Amazon Route 53](#)。使用 Amazon Route 53，您可以將一個或多個 AWS 區域中的 IP 端節與一個 Route 53 網域名稱建立關聯。若要實作手動啟動的容錯移轉，您可以使用 [Amazon 應用程式復原控制器](#)，它可提供高可用性的資料平面 API，將流量重新路由到復原區域。實作容錯移轉時，使用資料平面操作並避免控制平面操作，如 [REL11-BP04 復原期間需使用資料平面，而非控制平面](#) 中所述。

若要進一步了解此選項和其他選項，請參閱 [《災難復原白皮書》中的此章節](#)。

6. 設計一個工作負載故障恢復計畫。

容錯恢復是指在災難事件減弱後將工作負載操作回復到主要區域。將基礎設施和程式碼佈建到主要區域通常遵循最初使用的相同步驟，依賴基礎設施即程式碼和程式碼部署管道。容錯恢復的挑戰是還原資料存放區，並確保它們與操作中的復原區域保持一致。

在容錯移轉狀態下，復原區域中的資料庫是即時的並具有最新資料。後續目標是從復原區域重新同步到主要區域，確保它是最新的。

有些 AWS 服務將會自動執行此動作。如果使用 [Amazon DynamoDB 全域表](#)，即使主要區域中的資料表變得無法使用，當它重新上線時，DynamoDB 會繼續傳播任何擱置的寫入。如果使用 [Amazon Aurora 全球資料庫](#)並使用[受管計劃容錯移轉](#)，則維護 Aurora 全球資料庫的現有複寫拓撲。因此，主要區域中先前的讀取/寫入執行個體將成為複本，並從復原區域中接收更新。

如果這不是自動的，您將需要在主要區域中重建資料庫，做為復原區域中資料庫的複本。在許多情況下，這將涉及刪除舊的主要資料庫並建立新的複本。

容錯移轉後，如果您可以繼續在復原區域中執行，請考慮使其成為新的主要區域。您仍會執行上述所有步驟，使先前的主要區域成為復原區域。有些組織會執行排程輪換，定期 (例如每三個月) 交換其主要區域和復原區域。

容錯移轉和復原所需的所有步驟都應保持在可供所有團隊成員使用的程序手冊中，並定期進行審查。

使用 Elastic Disaster Recovery 時，該服務會協助協調和自動化容錯恢復程序。如需詳細資訊，請參閱 [Performing a failback](#)。

實作計畫的工作量：高

資源

相關的最佳實務：

- [the section called “REL09-BP01 識別並備份所有需要備份的資料，或從來源複製資料”](#)
- [the section called “REL11-BP04 復原期間需使用資料平面，而非控制平面”](#)
- [the section called “REL13-BP01 定義停機時間和資料遺失的復原目標”](#)

相關文件：

- [AWS 架構部落格：災難復原系列](#)
- [AWS 上工作負載的災難復原：在雲端中復原 \(AWS 白皮書\)](#)
- [雲端中的災難復原選項](#)
- [一小時建置無伺服器端的多區域、主動-主動後端解決方案](#)
- [多區域無伺服器後端 - 重新載入](#)
- [RDS：跨區域複寫僅供讀取複本](#)
- [Route 53：設定 DNS 備援](#)
- [S3：跨區域複寫](#)
- [什麼是 AWS Backup？](#)
- [什麼是 Amazon 應用程式復原控制器？](#)
- [AWS 彈性災難復原](#)
- [HashiCorp Terraform：入門 - AWS](#)
- [APN 合作夥伴：可以幫助災難復原的合作夥伴](#)
- [AWS Marketplace：可用於災難復原的產品](#)

相關影片：

- [AWS 上工作負載的災難復原](#)
- [AWS re:Invent 2018：適用於多區域主動-主動應用程式的架構模式 \(ARC209-R2\)](#)
- [AWS Elastic Disaster Recovery 入門 | Amazon Web Services](#)

REL13-BP03 測試災難復原實作以驗證實作

定期測試容錯移轉到您的復原站點以確認它正常操作，並符合 RTO 和 RPO。

常見的反模式：

- 切勿在生產環境中執行容錯移轉。

建立此最佳實務的優勢：定期測試您的災難復原計畫，可驗證該計畫能在需要時運作，也能讓您的團隊知道如何執行策略。

未建立此最佳實務時的曝險等級：高

實作指引

要避免的模式是：開發鮮少執行的復原路徑。例如，您可能有一個次要資料存放區，只供唯讀查詢之用。當您寫入資料存放區而主資料存放區發生故障時，您可能需要容錯移轉到次要資料存放區。如果您不經常測試此容錯移轉，則可能會發現您對次要資料存放區的功能的假設不正確。次要資料存放區的容量 (在您上次測試時可能已經足夠) 在這種情況下可能無法再容忍負載。我們的經驗顯示，唯一能發揮功用的錯誤復原，是您經常測試的路徑。因此，最好擁有少量的復原路徑。您可建立復原模式，並定期進行測試。若擁有複雜或關鍵復原路徑，您還是需要定期在生產環境中執行該故障，說服自己該復原路徑能發揮功用。在我們剛剛討論的範例中，無論是否需要，您都應定期容錯移轉到備用資料庫。

實作步驟

1. 為復原設計您的工作負載。定期測試您的復原路徑。復原導向運算可識別系統中能增強復原能力的特性：隔離和備援，系統範圍內的回復變更能力，監控和確定運行狀態的能力，提供診斷、自動復原和模組化設計的能力，以及重新啟動的能力。練習復原路徑，以確認您可以在指定時間內完成復原到指定狀態。在復原過程中使用您的執行手冊，以記錄問題並在下一次測試前找出其解決方案。
2. 對於以 Amazon EC2 為基礎的工作負載，請使用 [AWS Elastic Disaster Recovery](#) 為 DR 策略實作和啟動演練執行個體。AWS Elastic Disaster Recovery 提供有效執行演練的功能，協助您準備容錯移轉事件。您也可以使用 Elastic Disaster Recover 頻繁啟動您的執行個體進行測試和演練，而不需要重新導向流量。

資源

相關文件：

- [APN 合作夥伴：可以幫助災難復原的合作夥伴](#)
- [AWS 架構部落格：災難復原系列](#)
- [AWS Marketplace：可用於災難復原的產品](#)
- [AWS Elastic Disaster Recovery](#)
- [AWS 上工作負載的災難復原：在雲端中復原 \(AWS 白皮書\)](#)
- [AWS Elastic Disaster Recovery 準備容錯移轉](#)
- [柏克萊加州大學/史丹佛大學復原導向的運算專案](#)
- [什麼是 AWS Fault Injection Simulator？](#)

相關影片：

- [AWS re:Invent 2018：適用於多區域主動-主動應用程式的架構模式](#)

- [AWS re:Invent 2019：使用 AWS 的備份與還原和災難復原解決方案](#)

REL13-BP04 管理 DR 站點或區域的組態偏移

若要執行成功的災難復原 (DR) 程序，您的工作負載必須能夠及時恢復正常運作，而且在 DR 環境上線後，不會發生相關的功能喪失或資料損失。若要實現此目標，請務必在 DR 環境與主要環境之間維護一致的基礎設施、資料和組態。

預期成果：您的災難復原站點的組態和資料與主要站點相同，這有助於在需要時快速且完整的復原。

常見的反模式：

- 當主要位置發生變更時，您未能更新復原位置，這會導致組態過時，進而阻礙復原工作。
- 您未考量潛在的限制，例如，主要和復原位置之間的服務差異，這可能會導致容錯移轉期間發生非預期的失敗。
- 您依賴手動程序來更新和同步 DR 環境，這會增加人為錯誤和不一致的風險。
- 您未能偵測到組態偏移，這會導致在事件發生之前對 DR 站點整備度產生錯誤的認知。

建立此最佳實務的優勢：DR 環境與主要環境之間的一致性可大幅改善事件發生後成功復原的可能性，並降低復原程序失敗的風險。

未建立此最佳實務時的曝險等級：高

實作指引

一套完整的組態管理和容錯移轉整備方法，可協助您確認 DR 站點持續更新，並準備好在主要站點故障時接管。

若要實現主要與災難復原 (DR) 環境之間的一致性，請確定您的交付管道會將應用程式同時分配到主要和 DR 網站。經過適當的評估期 (也稱為交錯部署) 之後，將變更推展至 DR 站點，以偵測主要站點的問題，並在問題擴大之前停止部署。實作監控來偵測組態偏移，並追蹤整個環境的變更和合規性。在 DR 站點中執行自動修復，使其完全一致，並準備好在事件發生時接管。

實作步驟

1. 確定 DR 區域包含成功執行 DR 計畫所需的 AWS 服務和功能。
2. 使用基礎設施即程式碼 (IaC)。保持正式作業基礎設施和應用程式組態範本的準確性，並定期將其套用至災難復原環境。[AWS CloudFormation](#) 可偵測 CloudFormation 範本所指定內容與實際部署內容之間的偏移。

3. 設定 CI/CD 管道，將應用程式和基礎設施更新部署到所有環境，包括主要和 DR 站點。CI/CD 解決方案 (例如 [AWS CodePipeline](#)) 可以自動化部署程序，進而降低組態偏移的風險。
4. 在主要環境和 DR 環境之間交錯部署。此方法允許一開始在主要環境中部署和測試更新，這樣會隔離主要站點中的問題，避免後續傳播到 DR 站點。此方法可防止同時將瑕疵推送至正式作業環境和 DR 站點，並維護 DR 環境的完整性。
5. 同時主要和 DR 環境中持續監控資源組態。[AWS Config](#) 這類解決方案可協助強制執行組態合規性並偵測偏移，這有助於在環境之間維持組態的一致性。
6. 實作警示機制，以追蹤任何組態偏移或資料複寫中斷或延遲的情形，並發出通知。
7. 自動修復偵測到的組態偏移。
8. 排程定期稽核和合規檢查，以確認主要和 DR 組態之間持續保持一致。定期審查可協助您保持符合既定規則，並識別任何需要解決的差異。
9. 檢查 AWS 佈建的容量、服務配額、限流限制是否不相符，以及組態和版本的差異。

資源

相關的最佳實務：

- [REL01-BP01 了解服務配額和限制](#)
- [REL01-BP02 管理帳戶與區域之間的服務配額](#)
- [REL01-BP04 監控和管理配額](#)
- [REL13-BP03 測試災難復原實作以驗證實作](#)

相關文件：

- [依 AWS Config 規則 修補不合規的 AWS 資源](#)
- [AWS Systems Manager Automation](#)
- [AWS CloudFormation：偵測堆疊和資源的未受管組態變更](#)
- [AWS CloudFormation：在整個 CloudFormation 堆疊上偵測偏差](#)
- [AWS Systems Manager Automation](#)
- [AWS 上工作負載的災難復原：在雲端中復原 \(AWS 白皮書\)](#)
- [如何在 AWS 上實作基礎設施組態管理解決方案？](#)
- [依 AWS Config 規則 修補不合規的 AWS 資源](#)

相關影片：

- [AWS re:Invent 2018：適用於多區域主動-主動應用程式的架構模式 \(ARC209-R2\)](#)

相關範例：

- [CloudFormation 登錄](#)
- [適用於 AWS 的配額監控](#)
- [使用 Amazon CloudWatch 和 AWS Lambda 實作自動修復 AWS CloudFormation 的偏移](#)
- [AWS 架構部落格：災難復原系列](#)
- [AWS Marketplace：可用於災難復原的產品](#)
- [自動化安全、無人為介入的部署](#)

REL13-BP05 自動化回復

實作經測試、可靠、可觀測且可重現的自動復原機制，以減少故障的風險和業務影響。

預期成果：您已針對復原程序實作記錄完整、標準化且經過徹底測試的自動化工作流程。您的復原自動化會自動修正資料遺失或無法使用屬於低風險的次要問題。您可針對嚴重事件快速調用復原程序、觀察執行修復過程時的行為，並且在發現危險情況或故障時結束程序。

常見的反模式：

- 您依賴處於失敗或降級狀態的元件或機制做為復原計畫的一部分。
- 您的復原程序需要手動介入，例如主控台存取 (也稱為點按操作)。
- 您在資料遺失或無法使用處於高風險的情況下，自動初始化復原程序。
- 您未納入中止無法運作或會產生其他風險的復原程序的機制 (例如 Andon cord 或 大型紅色停止按鈕)。

建立此最佳實務的優勢：

- 復原操作擁有更高的可靠性、可預測性和一致性。
- 能夠實現更嚴格的復原目標，包括復原時間點目標 (RTO) 和復原點目標 (RPO)。
- 降低在事件期間復原失敗的可能性。
- 降低容易發生人為錯誤的手動復原程序伴隨的失敗風險。

未建立此最佳實務時的曝險等級：中

實作指引

若要實作自動復原程序，您需要一套使用 AWS 服務和最佳實務的全方位方法。若要開始進行，請找出工作負載中的關鍵元件和潛在的失敗點。制定可在不需要人為介入的情況下復原您的工作負載和資料的自動化程序。

使用基礎設施即程式碼 (IaC) 原則來制定復原自動化。這可讓復原環境與來源環境保持一致，並允許復原程序的版本控制。若要協調複雜的復原工作流程，請考慮 [AWS Systems Manager Automations](#) 或 [AWS Step Functions](#) 等解決方案。

自動化復原程序可提供顯著的優勢，並可協助您更輕鬆地達成復原時間點目標 (RTO) 和復原點目標 (RPO)。不過，這些程序可能會遇到非預期的情況，而導致程序失敗或本身產生新風險，例如額外的停機時間和資料遺失。為了降低此風險，請提供可快速停止進行中復原自動化的功能。停止後，您就可以調查並採取修正步驟。

對於支援的工作負載，請考慮採用 AWS 彈性災難復原 (AWS DRS) 這類解決方案來提供自動容錯移轉。AWS DRS 會持續將您的機器 (包括作業系統、系統狀態組態、資料庫、應用程式和檔案) 複寫至您的目標 AWS 帳戶和慣用區域中的預備區域。如果事件發生，AWS DRS 會自動將複寫的伺服器轉換到 AWS 上復原區域中完整佈建的工作負載。

自動復原的維護和改善是一個持續進行的過程。根據學到的經驗持續測試和精進復原程序，並隨時掌握可增強復原功能的新 AWS 服務和功能。

實作步驟

1. 規劃自動復原

- a. 徹底檢閱工作負載架構、元件和相依性，以識別和規劃自動復原機制。將工作負載的相依性分類為硬和軟相依性。硬相依性是指工作負載無法在沒有此相依性的情況下運作，且此相依性無法替代。軟相依性是指工作負載平常使用的相依性，此相依性可使用臨時替代系統或程序取代，也可以透過[正常降級](#)處理。
- b. 建立程序以識別和復原缺少或損毀的資料。
- c. 定義在復原動作完成後，用來確認復原穩定狀態的步驟。
- d. 考慮讓復原的系統就緒可提供完整服務所需的任何動作，例如預熱和填入快取。
- e. 考慮在復原過程中可能遇到的問題，以及如何偵測和修復這些問題。
- f. 考慮無法存取主要站點及其控制平面的情況。確認復原動作可以在不依賴主要站點的情況下單獨執行。考慮採用 [Amazon 應用程式復原控制器 \(ARC\)](#) 這類解決方案來重新導向流量，而不需要手動改變 DNS 記錄。

2. 制定自動復原程序

- a. 實作自動故障偵測和容錯移轉機制，不需手動介入即可自動復原。建置像是 [Amazon CloudWatch](#) 中的儀表板，用來報告自動復原程序的進度和運作狀態。納入確認成功復原的程序。提供中止進行中復原程序的機制。
- b. 針對無法自動復原的故障，製作[程序手冊](#)做為備援程序，並考量您的[災難復原計畫](#)。
- c. 依照 [REL13-BP03](#) 中所述，測試復原程序。

3. 準備復原

- a. 評估復原站點的狀態，並提前部署關鍵元件。如需詳細資訊，請參閱 [REL13-BP04](#)。
- b. 定義明確的角色、責任，以及復原操作的決策流程，並且讓整個組織的利害關係人和團隊參與其中。
- c. 定義初始化復原程序的條件。
- d. 建立計畫來還原復原程序，並視需要或在安全時恢復您的主要站點。

資源

相關的最佳實務：

- [REL07-BP01 取得或擴展資源時使用自動化](#)
- [REL11-BP01 監控工作負載的所有元件以偵測故障](#)
- [REL13-BP02 使用定義的復原策略來滿足復原目標](#)
- [REL13-BP03 測試災難復原實作以驗證實作](#)
- [REL13-BP04 管理 DR 站點或區域的組態偏移](#)

相關文件：

- [AWS 架構部落格：災難復原系列](#)
- [AWS 上工作負載的災難復原：在雲端中復原 \(AWS 白皮書\)](#)
- [使用 Amazon Route 53 ARC 和 AWS Step Functions 協調災難復原自動化](#)
- [使用 AWS CDK 建置 AWS Systems Manager Automation 執行手冊](#)
- [AWS Marketplace：可用於災難復原的產品](#)
- [AWS Systems Manager Automation](#)
- [AWS 彈性災難復原](#)
- [使用彈性災難復原進行容錯移轉和容錯恢復](#)

- [AWS 彈性災難復原資源](#)
- [APN 合作夥伴：可協助進行災難復原的合作夥伴](#)

相關影片：

- [AWS re:Invent 2018：適用於多區域主動-主動應用程式的架構模式 \(ARC209-R2\)](#)
- [AWS re:Invent 2022：AWS On Air 搭配 AWS 彈性災難復原的 AWS 容錯恢復](#)

效能效率

效能效率支柱包括能夠有效率地使用雲端資源，以滿足效能需求，並隨著需求變更與技術發展來保持該效率需求。可以在[效能達成效率支柱白皮書](#)中找到實作指引。

最佳實務領域

- [架構選擇](#)
- [運算與硬體](#)
- [資料管理](#)
- [聯網與內容交付](#)
- [程序和文化](#)

架構選擇

問題

- [PERF 1. 如何為工作負載選取合適的雲端資源和架構？](#)

PERF 1. 如何為工作負載選取合適的雲端資源和架構？

適用於特定工作負載的最佳解決方案各不相同，而解決方案通常會結合多種方法。Well-Architected 工作負載會使用多種解決方案，並採用不同的功能以提升效能。

最佳實務

- [PERF01-BP01 了解並理解可用的雲端服務和功能](#)
- [PERF01-BP02 使用雲端供應商或適當合作夥伴提供的指引，了解架構模式和最佳實務](#)
- [0PERF01-BP03 將成本納入架構決策](#)

- [PERF01-BP04 評估權衡如何影響客戶和架構效率](#)
- [PERF01-BP05 使用政策和參考架構](#)
- [PERF01-BP06 使用基準測試來推動架構決策](#)
- [PERF01-BP07 使用資料驅動型方法來選擇架構](#)

PERF01-BP01 了解並理解可用的雲端服務和功能

持續了解並探索可用的服務和組態，有助您做出更完善的架構決策，並提升工作負載架構的效能效率。

常見的反模式：

- 您可以使用雲端作為並置資料中心。
- 移轉到雲端後，您不會將應用程式現代化。
- 對於需要保留的所有項目，您只使用一種儲存類型。
- 您使用的執行個體類型與目前標準最相符，但大於需求。
- 您會部署和管理可做為受管服務的技術。

建立此最佳實務的優勢：透過考慮新服務和設定，您可以大幅提升效能、降低成本並最佳化維護工作負載所需的工作量。這麼做還可幫助您縮短具有雲端功能之產品的價值實現時間。

未建立此最佳實務時的曝險等級：高

實作指引

AWS 持續推出可提升效能並降低雲端工作負載成本的新服務和功能。保持這些新服務和功能的最新狀態對於維持雲端效能有效性至關重要。將工作負載架構現代化也可協助您提升生產力、推動創新並釋放更多成長機會。

實作步驟

- 清查工作負載軟體和架構以存放相關服務。決定要深入了解的產品類別。
- 探索 AWS 供應項目，以識別並了解相關服務和組態選項，這些選項可協助您改善效能，並降低成本和操作複雜性。
 - [Amazon Web Services 雲端](#)
 - [AWS Academy](#)
 - [AWS 最新消息](#)
 - [AWS 部落格](#)

- [AWS Skill Builder](#)
- [AWS 活動和研討會](#)
- [AWS 培訓 和認證](#)
- [AWS Youtube 頻道](#)
- [AWS 研討會](#)
- [AWS 社群](#)
- 使用 [Amazon Q](#) 取得有關服務的相關資訊和建議。
- 使用沙盒 (非生產) 環境來學習和試驗新服務，而不會產生額外成本。
- 持續了解新雲端服務和功能。

資源

相關文件：

- [Amazon Web Services 概觀](#)
- [Amazon EC2 功能](#)
- [透過 AWS 合作夥伴學習計劃逐步學習](#)
- [AWS 培訓和認證](#)
- [我成為 AWS 解決方案架構師的學習路徑](#)
- [AWS 架構中心](#)
- [AWS Partner Network](#)
- [AWS 解決方案程式庫](#)
- [AWS 知識中心](#)
- [在 AWS 上建置現代化應用程式](#)

相關影片：

- [AWS re:Invent 2023 - Amazon EC2 最新消息](#)
- [AWS re:Invent 2022 - 使用 Amazon ECS 降低您的營運和基礎設施成本](#)
- [AWS re:Invent 2023 - 使用 AWS，利用雲端的效率、敏捷性和創新進行建置](#)
- [AWS re:Invent 2022 - 以高效能和低成本部署機器學習 \(ML\) 模型以進行推論](#)
- [This is my Architecture](#)

相關範例：

- [AWS 範例](#)
- [AWS SDK 範例](#)

PERF01-BP02 使用雲端供應商或適當合作夥伴提供的指引，了解架構模式和最佳實務

使用文件、解決方案架構師、專業服務或適當的合作夥伴等雲端公司資源，來引導您做出架構決策。這些資源可協助審核和改善架構，以實現最佳效能。

常見的反模式：

- 使用 AWS 作為常見的雲端供應商。
- 以非設計宗旨的方式使用 AWS 服務。
- 遵循所有指引，但未考量自身的業務環境。

建立此最佳實務的優勢：使用雲端供應商或適當合作夥伴的指引，可協助您針對工作負載做出正確的架構選擇，並讓您對決策充滿信心。

未建立此最佳實務時的曝險等級：中

實作指引

AWS 提供廣泛的指引、文件和資源，可協助您建置和管理高效雲端工作負載。AWS 文件提供程式碼範例、教學課程和詳細的服務說明。除了說明文件外，AWS 還提供訓練與認證計畫、解決方案架構師和專業服務，協助客戶探索雲端服務的不同層面，並在 AWS 上實作高效的雲端架構。

利用這些資源來深入了解寶貴的知識和最佳實務、節省時間並在 AWS 雲端 中取得更好的成果。

實作步驟

- 審核 AWS 文件和指引，並遵循最佳實務。這些資源可協助您有效選擇和設定服務，並取得更好的效能。
 - [AWS 文件](#) (例如使用者指南和白皮書)
 - [AWS 部落格](#)
 - [AWS 培訓 和認證](#)
 - [AWS Youtube 頻道](#)

- 參加 AWS 合作夥伴活動 (例如 AWS 全球高峰會、AWS re:Invent、使用者群組和研討會)，向 AWS 專家學習使用 AWS 服務的最佳實務。
 - [透過 AWS 合作夥伴學習計劃逐步學習](#)
 - [AWS 活動和研討會](#)
 - [AWS 研討會](#)
 - [AWS 社群](#)
- 當您需要其他指引或產品資訊時，請聯絡 AWS 尋求協助。AWS 解決方案架構師和 [AWS 專業服務](#) 會為解決方案實作提供指引。[AWS 合作夥伴](#) 提供 AWS 專業知識，協助您提升業務的靈活性和創新性。
- 如果您需要技術支援才能有效使用服務，請使用 [支援](#)。[我們的支援計劃](#) 旨在為您提供適當的工具組合和專業知識，以便您在最佳化效能、管理風險以及控制成本的同時，在 AWS 上取得成功。

資源

相關文件：

- [AWS 架構中心](#)
- [AWS Partner Network](#)
- [AWS 解決方案程式庫](#)
- [AWS 知識中心](#)
- [AWS 企業支援](#)

相關影片：

- [This is my Architecture](#)
- [AWS re:Invent 2023 - 使用 Amazon EventBridge 提升事件驅動型模式](#)
- [AWS re:Invent 2023 - 在 AWS 上實作分散式設計模式](#)
- [AWS re:Invent 2023 - 應用程式架構即程式碼](#)

相關範例：

- [AWS 範例](#)
- [AWS SDK 範例](#)
- [AWS 分析參考架構](#)

OPERF01-BP03 將成本納入架構決策

將成本納入架構決策中，以提高雲端工作負載的資源使用率和效能效率。當您意識到雲端工作負載的成本影響時，就更有可能利用有效的資源並減少浪費的做法。

常見的反模式：

- 您只能使用一個執行個體系列。
- 您不會針對開放原始碼解決方案評估授權解決方案。
- 您不會定義儲存區生命週期政策。
- 您不會檢閱 的新服務和功能 AWS 雲端。
- 您只能使用區塊儲存。

建立此最佳實務的優勢：將成本納入到決策中可讓您使用更有效率的資源並探索其他投資。

未建立此最佳實務時的曝險等級：中

實作指引

優化工作負載成本可以提高資源利用率並避免雲端工作負載中的浪費。將成本納入架構決策中，通常包括適當調整工作負載元件大小以及啟用彈性，進而提高雲端工作負載效能的效率。

實作步驟

- 確立成本目標，例如雲端工作負載的預算限制。
- 找出造成工作負載成本增加的關鍵元件 (例如執行個體和儲存)。可使用 [AWS 定價計算工具](#) 和 [AWS Cost Explorer](#) 找出工作負載中的關鍵成本驅動因素。
- 了解雲端中的 [定價模式](#)，例如隨需執行個體、預留執行個體、Savings Plans 和 Spot 執行個體。
- 使用 [Well-Architected 成本最佳實務](#)，針對成本最佳化這些關鍵元件。
- 持續監控和分析成本，以找出工作負載中成本最佳化的機會。
 - 使用 [AWS Budgets](#) 取得不可接受成本的警示。
 - 使用 [AWS Compute Optimizer](#) 或 [AWS Trusted Advisor](#) 得成本最佳化建議。
 - 使用 [AWS Cost Anomaly Detection](#) 取得自動化成本異常偵測和根本原因分析。

資源

相關文件：

- [什麼是 AWS Billing and Cost Management ?](#)
- [使用 進行成本最佳化 AWS](#)
- [選擇 AWS 成本管理策略](#)
- [AWS 成本管理入門指南](#)
- [成本智慧儀表板的詳細概要](#)
- [AWS 架構中心](#)
- [AWS 解決方案程式庫](#)
- [AWS 知識中心](#)

相關影片：

- [This is my Architecture](#)
- [AWS re : Invent 2023 - AWS 成本最佳化的新功能](#)
- [AWS re : Invent 2023 - 最佳化成本和效能，並追蹤緩解進度](#)
- [AWS re : Invent 2023 - AWS 儲存成本最佳化最佳實務](#)
- [AWS re : Invent 2023 - 最佳化多帳戶環境中的成本](#)

相關範例：

- [AWS Compute Optimizer 示範程式碼](#)
- [成本最佳化研討會](#)
- [雲端財務管理技術實作說明手冊](#)
- [啟動最佳化：調整應用程式效能以實現最高效率](#)
- [無伺服器最佳化研討會 \(效能與成本\)](#)
- [擴充經濟高效的架構](#)

PERF01-BP04 評估權衡如何影響客戶和架構效率

在評估與效能相關的改進時，判斷哪些選擇會影響客戶和工作負載效率。例如，如果使用鍵值資料存放區可提高系統效能，請務必評估此變更最終一致性本質對客戶的影響。

常見的反模式：

- 即使實作過程中有所取捨，您都假設應實作所有效能增益。

- 您只會在效能問題達到臨界點時才會評估工作負載變更。

建立此最佳實務的優勢：評估潛在的效能相關改善項目時，必須判斷技術變更的權衡是否符合工作負載要求。在某些情況下，您可能需要實作其他控制來彌補權衡。

未建立此最佳實務時的曝險等級：高

實作指引

根據效能和客戶影響，識別架構中的關鍵領域。確定如何進行改進、這些改進帶來的權衡，以及它們如何影響系統和使用者的體驗。例如，實作快取資料有助於大幅提升效能，但需要明確的策略來確定更新或使快取資料失效的方式和時間，以防止不正確的系統行為。

實作步驟

- 了解工作負載需求和 SLA。
- 清楚定義評估因素。因素可能與工作負載的成本、可靠性、安全性和效能有關。
- 選擇可滿足需求的架構和服務。
- 進行實驗和概念驗證 (POC)，以評估權衡因素以及對客戶和架構效率的影響。通常，高可用性、高效能且安全的工作負載會耗用更多雲端資源，但能夠提供更完善的客戶體驗。了解工作負載複雜性、效能和成本的權衡。通常情況下，優先考慮其中兩個因素會以犧牲第三個因素為代價。

資源

相關文件：

- [Amazon 建置者資料中心](#)
- [Quick KPI](#)
- [Amazon CloudWatch RUM](#)
- [X-Ray 文件](#)
- [了解恢復模式和權衡取捨以便在雲端中高效進行架構](#)

相關影片：

- [透過 Amazon CloudWatch RUM 優化應用程式](#)
- [AWS re:Invent 2023 - 容量、可用性、成本效率：挑選三項](#)
- [AWS re:Invent 2023 - 鬆耦合系統的進階整合模式和權衡](#)

相關範例：

- [使用 Amazon CloudWatch Synthetics 測量頁面載入時間](#)
- [Amazon CloudWatch RUM Web 用戶端](#)

PERF01-BP05 使用政策和參考架構

選擇服務和組態時，使用內部政策和現有的參考架構，以便在設計和實作工作負載提高效率。

常見的反模式：

- 您允許各種可能會影響公司管理開銷的技術。

建立此最佳實務的優勢：為架構、技術和供應商選擇制定政策，可讓您快速做出決策。

未建立此最佳實務時的曝險等級：中

實作指引

在選擇資源和架構方面擁有內部政策，提供在選擇架構時要遵循的標準和準則。這些準則可簡化在選擇合適的雲端服務時的決策過程，並有助於提高效能效率。使用政策或參考架構來部署工作負載。將服務整合到您的雲端部署，然後使用效能測試以確認您可以繼續滿足效能需求。

實作步驟

- 清楚了解雲端工作負載的需求。
- 檢閱內部和外部政策，以識別最相關的政策。
- 使用 AWS 提供的適當參考架構或您的業界最佳實務。
- 針對常見情況，建立包含政策、標準、參考架構和規範指引的連續體。這樣做可以讓您的團隊更快地行動。如果適用，為您的垂直發展量身打造資產。
- 針對沙盒環境中的工作負載，驗證這些政策和參考架構。
- 隨時 up-to-date 掌握業界標準和 AWS 更新，確保您的政策和參考架構有助於最佳化您的雲端工作負載。

資源

相關文件：

- [AWS 架構中心](#)

- [AWS Partner Network](#)
- [AWS 解決方案程式庫](#)
- [AWS 知識中心](#)
- [AWS 架構部落格](#)

相關影片：

- [This is my Architecture](#)
- [AWS re : Invent 2022 - 使用 SAP & AWS reference 架構為您的企業加速價值](#)

相關範例：

- [AWS 範例](#)
- [AWS SDK 範例](#)

PERF01-BP06 使用基準測試來推動架構決策

對現有工作負載的效能進行基準化分析，以了解工作負載在雲端的效能，並根據該資料推動架構決策。

常見的反模式：

- 您倚賴不代表工作負載特性的常見基準。
- 您將客戶的意見回饋和看法作為唯一基準。

建立此最佳實務的優勢：對目前的實作進行基準測試可讓您衡量效能改進。

未建立此最佳實務時的曝險等級：中

實作指引

使用基準化分析搭配綜合測試，以評估工作負載元件的效能。與負載測試相比，基準化分析通常速度更快；要評估特定元件的技術時，會使用基準化分析。當您缺少執行負載測試的完整解決方案時，通常可在新專案開始時使用基準化分析。

您可以建置自己的自訂基準化分析測試，也可以使用產業標準測試，例如 [TPC-DS](#)，對工作負載進行基準化分析。比較環境時，產業基準化分析很有幫助。對於確定您希望在架構中進行的特定營運類型，自訂基準化分析非常實用。

基準化分析時，務必要預熱測試環境，以獲得有效結果。多次執行相同的基準化分析，以確認您已擷取到隨時間推移出現的任何變化。

由於基準化分析的速度通常比負載測試要快，因此可以在部署管道中盡早使用基準化分析，以便能更快提供有關效能偏差的回饋。當您評估元件或服務中的重大變更時，藉助基準化分析，您可以更快速地查看所做的變更是否合理。請務必使用基準化分析搭配負載測試，因為負載測試將告訴您工作負載在生產中的效能。

實作步驟

- 規劃和定義：
 - 為基準化分析定義目標、基準、測試案例、指標 (例如 CPU 使用率、延遲或輸送量) 以及 KPI。
 - 關注使用者體驗方面的使用者需求，以及回應時間和可存取性等因素。
 - 找出工作負載適用的基準化分析工具。可以使用 AWS 服務 (例如 [Amazon CloudWatch](#))，或與工作負載相容的第三方工具。
- 配置並檢測：
 - 設定環境並配置資源。
 - 實作監控和日誌記錄以擷取測試結果。
- 基準化分析和監控：
 - 在測試期間執行基準化分析並監控指標。
- 分析並記錄：
 - 記錄基準化分析過程和調查結果。
 - 分析結果以找出瓶頸、趨勢和需要改善的領域。
 - 使用測試結果做出架構決策並調整工作負載。這可能包括變更服務或採用新功能。
- 最佳化並重複：
 - 根據您的基準化分析來調整資源配置和分配。
 - 調整後重新測試您的工作負載，以驗證改進。
 - 記錄您的學習，並重複此過程以確定其他有待改進的領域。

資源

相關文件：

- [AWS 架構中心](#)
- [AWS Partner Network](#)

- [AWS 解決方案程式庫](#)
- [AWS 知識中心](#)
- [Amazon CloudWatch RUM](#)
- [Amazon CloudWatch Synthetics](#)
- [基因體工作流程，第 5 部分：自動化基準測試](#)
- [對 Amazon SageMaker AI JumpStart 中的端點部署進行基準測試和最佳化](#)

相關影片：

- [AWS re:Invent 2023 - 對 AWS Lambda 冷啟動進行基準測試](#)
- [在雲端中對有狀態服務進行基準測試](#)
- [This is my Architecture](#)
- [透過 Amazon CloudWatch RUM 優化應用程式](#)
- [Amazon CloudWatch Synthetics 的示範](#)

相關範例：

- [AWS 範例](#)
- [AWS SDK 範例](#)
- [分散式負載測試](#)
- [使用 Amazon CloudWatch Synthetics 測量頁面載入時間](#)
- [Amazon CloudWatch RUM Web 用戶端](#)

PERF01-BP07 使用資料驅動型方法來選擇架構

為架構選擇定義清晰、資料驅動型方法，以確認是否使用正確的雲端服務和組態，來滿足特定業務需求。

常見的反模式：

- 您假設目前的架構是靜態的，且不應隨著時間而更新。
- 您的架構選擇基於猜測和假設。
- 您會隨時間導入架構變更，而且無須理由佐證。

建立此最佳實務的優勢：透過採用明確定義的方法來做出架構選擇，您可以使用資料來影響工作負載設計，並隨著時間的推移做出明智的決策。

未建立此最佳實務時的曝險等級：中

實作指引

使用雲端或外部資源 (例如已發佈的使用案例或白皮書) 的內部經驗和知識，在架構中選擇資源和服務。您應擁有一個明確定義的流程，鼓勵對工作負載中可能使用的服務進行實驗和基準化分析。

關鍵工作負載的待辦項目不僅應包括可提供與業務和使用者相關的功能的使用者故事，還包括構成工作負載架構跑道的技術故事。這條跑道了解科技和新服務的新進展，並根據資料和適當理由採用這些技術和新服務。這證明該架構仍然面向未來，不會停滯不前。

實作步驟

- 與關鍵利益相關者互動，以定義工作負載需求，包括效能、可用性和成本考量。考慮工作負載的使用者數量和使用模式等因素。
- 建立架構跑道或技術待辦項目，系統會優先處理這些項目與功能待辦事項。
- 評價和評估不同的雲端服務 (如需詳細資訊，請參閱 [PERF01-BP01 了解並理解可用的雲端服務和功能](#))。
- 探索符合效能需求的不同架構模式，例如微型服務或無伺服器 (如需詳細資訊，請參閱 [PERF01-BP02 使用雲端供應商或適當合作夥伴提供的指引，了解架構模式和最佳實務](#))。
- 諮詢其他團隊、架構圖和資源，例如 AWS Solution Architects、[AWS Architecture Center](#) 和 [AWS Partner Network](#)，以協助您選擇適合工作負載的正確架構。
- 定義輸送量和回應時間等效能指標，以協助您評估工作負載的效能。
- 實驗並使用定義的指標來驗證所選架構的效能。
- 視需要持續監控並進行調整，以維持架構的最佳效能。
- 記錄您選擇的架構和決策，作為未來更新和學習的參考。
- 根據學習、新技術和指標 (其指出目前方法中需要的變更或問題)，持續審核和更新架構選擇方法。

資源

相關文件：

- [AWS 解決方案程式庫](#)

- [AWS 知識中心](#)
- [在 AWS 上建置端對端資料驅動型應用程式的架構模式](#)

相關影片：

- [This is my Architecture](#)
- [AWS re:Invent 2021 - 資料驅動型企業：從願景走向價值](#)
- [AWS re:Invent 2022 - 提供可持續、高效能的架構](#)
- [AWS re:Invent 2023 - 優化成本和效能並追蹤緩解措施的進度](#)
- [AWS re:Invent 2022 - AWS 優化：立即見效的可操作步驟](#)

相關範例：

- [AWS 範例](#)
- [AWS SDK 範例](#)

運算與硬體

問題

- [PERF 2. 如何在工作負載中選取和使用運算資源？](#)

PERF 2. 如何在工作負載中選取和使用運算資源？

特定工作負載的最佳運算選擇會根據應用程式設計、使用模式和組態設定而有所不同。架構會針對不同元件使用不同運算選擇，並採用不同功能以提升效能。若選錯運算資源，可能使架構的效能達成效率降低。

最佳實務

- [PERF02-BP01 選取最適合您工作負載的運算選項](#)
- [PERF02-BP02 了解可用的運算組態和功能](#)
- [PERF02-BP03 收集運算相關指標](#)
- [PERF02-BP04 設定運算資源並適當調整大小](#)
- [PERF02-BP05 動態擴展您的運算資源](#)
- [PERF02-BP06 使用最佳化的硬體型運算加速器](#)

PERF02-BP01 選取最適合您工作負載的運算選項

為工作負載選擇最合適的運算選項，可讓您改善效能、減少不必要的基礎設施成本，並降低維護工作負載所需的作業工作量。

常見的反模式：

- 您使用曾用於內部部署的同一個運算選項。
- 缺乏對雲端運算選項、特徵以及解決方案，以及那些解決方案可以如何改善運算效能的認識。
- 您在替代運算選項更精確地符合工作負載特性時，過度佈建現有運算選項以符合擴展或效能需求。

建立此最佳實務的優勢：可以透過找出運算需求並根據可用選項進行評估，提高工作負載的資源效率。

未建立此最佳實務時的曝險等級：高

實作指引

為了最佳化雲端工作負載以提高效能效率，請務必根據使用案例和效能需求選擇最合適的運算選項。AWS 提供多種運算選項，以滿足雲端中不同工作負載的需求。例如，您可以使用 [Amazon EC2](#) 來啟動和管理虛擬伺服器，使用 [AWS Lambda](#) 來執行程式碼，而不必佈建或管理伺服器，使用 [Amazon ECS](#) 或 [Amazon EKS](#) 執行和管理容器，或者使用 [AWS Batch](#) 平行處理大量資料。根據擴展和運算需求，您應該根據自己的情況選擇並設定最佳的運算解決方案。也可以考慮在單一工作負載中使用多種類型的運算解決方案，因為每種運算解決方案都有自己的優點和缺點。

下列步驟會引導您選取正確的運算選項，以符合您的工作負載特性和效能需求。

實作步驟

- 了解工作負載運算需求。需要考慮的關鍵需求包括處理需求、流量模式、資料存取模式、擴展需求和延遲需求。
- 了解適用於工作負載的不同 [AWS 運算服務](#)。如需更多詳細資訊，請參閱 [PERF01-BP01 了解並理解可用的雲端服務和功能](#)。以下是一些關鍵的 AWS 運算選項、其特性和常見使用案例：

AWS 服務	重要特性	常用案例
Amazon Elastic Compute Cloud (Amazon EC2)	擁有專為硬體、授權要求、大規模選取的不同執行個體系列、處理器類型與運算加速器設計的選項	平移遷移、整合型應用程式、混合環境、企業應用程式

AWS 服務	重要特性	常用案例
Amazon Elastic Container Service (Amazon ECS) 、 Amazon Elastic Kubernetes Service (Amazon EKS)	輕鬆的部署、一致的環境、可擴展	微型服務、混合環境
AWS Lambda	無伺服器運算 服務可執行程式碼以回應事件，並自動管理基礎運算資源。	微型服務、事件驅動型應用程式
AWS Batch	有效且動態地佈建和擴展 Amazon Elastic Container Service (Amazon ECS) 、 Amazon Elastic Kubernetes Service (Amazon EKS) 和 AWS Fargate 運算資源，並可根據您的任務需求選擇使用隨需執行個體或 Spot 執行個體	高效能運算 (HPC)，訓練機器學習 (ML) 模型
Amazon Lightsail	預先設定用於執行小型工作負載的 Linux 和 Windows 應用程式	簡易網路應用程式、自訂的網站

- 評估與每個運算選項相關聯的成本 (例如每小時費用或資料傳輸) 和管理開銷 (例如修補和擴展)。
- 在非生產環境中執行實驗和基準測試，以確定哪個運算選項最能滿足您的工作負載需求。
- 在您試驗和找出新的運算解決方案，請規劃遷移並驗證效能指標。
- 使用諸如 [Amazon CloudWatch](#) 的 AWS 監控工具，和諸如 [AWS Compute Optimizer](#) 的最佳化服務，根據實際使用模式持續最佳化運算資源。

資源

相關文件：

- [使用 AWS 進行雲端運算](#)

- [Amazon EC2 執行個體類型](#)
- [Amazon EKS 容器：Amazon EKS 工作節點](#)
- [Amazon ECS 容器：Amazon ECS 容器執行個體](#)
- [函數：Lambda 函數組態](#)
- [容器的規範性指引](#)
- [無伺服器的規範性指引](#)

相關影片：

- [AWS re:Invent 2023 - AWS Graviton：AWS 工作負載的最佳性價比](#)
- [AWS re:Invent 2023 - 在 AWS 中新建 Amazon Elastic Compute Cloud 生成式 AI 功能](#)
- [AWS re:Invent 2023 - Amazon Elastic Compute Cloud 的最新消息](#)
- [AWS re:Invent 2023 - 智慧型節約：Amazon Elastic Compute Cloud 成本最佳化策略](#)
- [AWS re:Invent 2021 - 為新一代 Amazon Elastic Compute Cloud 提供支援：深入研究 Nitro 系統](#)
- [AWS re:Invent 2019 - 最佳化 AWS 運算的效能和成本](#)
- [AWS re:Invent 2019 - Amazon Elastic Compute Cloud 基礎](#)
- [AWS re:Invent 2022 - 以高效能和低成本部署機器學習 \(ML\) 模型以進行推論](#)
- [AWS re:Invent 2019 - 最佳化 AWS 運算的效能和成本](#)
- [Amazon EC2 基礎](#)
- [以高效能和低成本部署機器學習 \(ML\) 模型以進行推論](#)

相關範例：

- [遷移 Web 應用程式至容器](#)
- [執行 Serverless Hello World](#)
- [Amazon EKS 研討會](#)
- [Amazon EC2 研討會](#)
- [使用 Amazon Elastic Compute Cloud 自動擴展實現高效且彈性的工作負載](#)
- [透過容器服務遷移至 AWS Graviton](#)

PERF02-BP02 了解可用的運算組態和功能

了解運算服務的可用組態選項和特徵，有助您佈建適量的資源並提高效能效率。

常見的反模式：

- 您沒有根據工作負載特性，評估運算選項或可用的執行個體系列。
- 過度佈建運算資源以符合尖峰需求。

建立此最佳實務的優勢：熟悉 AWS 運算功能和組態，以便使用最佳化的運算解決方案，以符合您的工作負載特性和需求。

未建立此最佳實務時的曝險等級：中

實作指引

每個運算解決方案都有獨特的組態和功能，以支援不同的工作負載特性和需求。了解這些選項如何與您的工作負載互補，並確定哪種組態選項最適合您的應用程式。這些選項的範例包括執行個體系列、大小、特徵 (GPU、I/O)、爆量、逾時、函數大小、容器執行個體，以及並行。如果您的工作負載使用相同的運算選項超過四週，而且您預期這些特性未來將保持不變，則可以使用 [AWS Compute Optimizer](#) 來確定目前的運算選項是否適合 CPU 和記憶體的工作負載。

實作步驟

- 了解工作負載需求 (例如 CPU 需求、記憶體和延遲)。
- 審核 AWS 文件和最佳實務，以了解可協助改善運算效能的建議組態選項。以下是一些需要考慮的關鍵組態選項：

組態選項	範例
執行個體類型	• 運算最佳化的執行個體非常適合需要較高的 vCPU 與記憶體比率的工作負載。 • 記憶體最佳化執行個體提供大量記憶體，以支援記憶體密集型工作負載。 • 儲存最佳化執行個體專為需要對本機儲存進行高序列讀取及寫入存取 (IOPS) 的工作負載所設計。
定價方式	• 隨需執行個體允許您按秒數或時數來使用運算能力，無須長期承諾。這些執行個體適合於超過效能基準需求的爆量。

組態選項	範例
	• Savings Plans 可大幅節省隨需執行個體，以換取在一年或三年內使用特定運算能力的承諾。 • Spot 執行個體 可讓您以折扣價利用未使用的執行個體容量，用於無狀態、容錯的工作負載。
Auto Scaling	使用 Auto Scaling 設定，使運算資源與流量模式相符。
規模調整	• 使用 Compute Optimizer 取得機器學習支援的建議，了解哪些運算組態最符合您的運算特性。 • 使用 AWS Lambda Power Tuning 為您的 Lambda 函數選擇最佳組態。
硬體型運算加速器	• 加速運算執行個體 比起 CPU 型替代品而言，可更有效地執行圖形處理或資料模式比對等功能。 • 針對機器學習工作負載，請利用專供工作負載使用的專用硬體，例如 AWS Trainium、AWS Inferentia 和 Amazon EC2 DL1

資源

相關文件：

- [使用 AWS 進行雲端運算](#)
- [Amazon EC2 執行個體類型](#)
- [Amazon EC2 執行個體的處理器狀態控制](#)
- [Amazon EKS 容器：Amazon EKS 工作節點](#)
- [Amazon ECS 容器：Amazon ECS 容器執行個體](#)
- [函數：Lambda 函數組態](#)

相關影片：

- [AWS re:Invent 2023 – AWS Graviton：AWS 工作負載的最佳價格效能](#)
- [AWS re:Invent 2023 – AWS 管理主控台 中的最新 Amazon EC2 生成式 AI 功能](#)
- [AWS re:Invent 2023 - Amazon EC2 最新消息](#)
- [AWS re:Invent 2023 - 智慧型節省：Amazon EC2 成本優化策略](#)
- [AWS re:Invent 2021 – 為新一代 Amazon EC2 提供支援：深入探索 Nitro 系統](#)
- [AWS re:Invent 2019 – Amazon EC2 基礎](#)
- [AWS re:Invent 2022 - 在 AWS 上針對效能和成本最佳化 Amazon EKS](#)

相關範例：

- [運算最佳化工具示範程式碼](#)
- [Amazon EC2 Spot 執行個體研討會](#)
- [使用 Amazon EC2 AWS Auto Scaling 提供有效且彈性的工作負載](#)
- [Graviton 開發人員研討會](#)
- [AWS for Microsoft workloads immersion day](#)
- [AWS for Linux workloads immersion day](#)
- [AWS Compute Optimizer 示範程式碼](#)
- [Amazon EKS 研討會](#)

PERF02-BP03 收集運算相關指標

記錄並追蹤與運算相關的指標，進一步了解運算資源的效能，並改善效能及使用率。

常見的反模式：

- 您只使用手動日誌檔案來搜尋指標。
- 您只會使用監控軟體記錄的預設指標。
- 您只會在有問題時審查指標。

建立此最佳實務的優勢：收集效能相關指標有助於使應用程式效能與業務需求保持一致，確保符合工作負載需求。這麼做也可以協助您持續改善工作負載中的資源效能和使用率。

未建立此最佳實務時的曝險等級：高

實作指引

雲端工作負載可以產生大量資料，例如指標、日誌和事件。在 中 AWS 雲端，收集指標是改善安全性、成本效益、效能和永續性的重要步驟。使用 [Amazon CloudWatch](#) 等監控服務 AWS，提供廣泛的效能相關指標，為您提供寶貴的洞見。CPU 使用率、記憶體使用率、磁碟 I/O 和網路傳入和傳出等指標可以提供使用率層級或效能瓶頸的洞察。將這些指標納入資料驅動的方法，以主動調整和優化工作負載的資源。在理想的情況下，應該在單一平台中收集與運算資源相關的所有指標，並實作保留政策以支援成本和營運目標。

實作步驟

- 識別與您的工作負載相關的效能相關指標。您應該收集與資源使用率和雲端工作負載運作方式有關的指標 (例如回應時間和輸送量)。
 - [Amazon EC2 預設指標](#)
 - [Amazon ECS 預設指標](#)
 - [Amazon EKS 預設指標](#)
 - [Lambda 預設指標](#)
 - [Amazon EC2記憶體和磁碟指標](#)
- 為工作負載選擇並設定合適的日誌記錄和監控解決方案。
 - [AWS 原生可觀測性](#)
 - [AWS Distro for OpenTelemetry](#)
 - [Amazon Managed Service for Prometheus](#)
- 根據工作負載需求，為指標定義必要的篩選條件和彙總。
 - [使用 Amazon CloudWatch Logs 和指標篩選條件量化自訂應用程式指標](#)
 - [使用 Amazon CloudWatch 策略標記收集自訂指標](#)
- 為指標設定資料保留政策，以符合安全性和營運目標。
 - [CloudWatch 指標的預設資料保留](#)
 - [CloudWatch 日誌的預設資料保留](#)
- 如有必要，為指標建立警示和通知，可協助您主動回應效能相關問題。
 - [使用 Amazon CloudWatch 異常偵測建立自訂指標的警示](#)
 - [使用 Amazon 為特定網頁建立指標和警示 CloudWatch RUM](#)
- 使用自動化來部署指標和記錄彙總代理程式。

- [AWS Systems Manager 自動化](#)
- [OpenTelemetry 收集器](#)

資源

相關文件：

- [監控與可觀測性](#)
- [最佳實務：使用 實作可觀測性 AWS](#)
- [Amazon CloudWatch 文件](#)
- [使用 CloudWatch 代理程式從 Amazon EC2執行個體和內部部署伺服器收集指標和日誌](#)
- [存取 的 Amazon CloudWatch Logs AWS Lambda](#)
- [將 CloudWatch 日誌與容器執行個體搭配使用](#)
- [發佈自訂指標](#)
- [AWS Answers：集中式日誌記錄](#)
- [AWS 發佈 CloudWatch 指標的服務](#)
- [在 EKS上監控 Amazon AWS Fargate](#)

相關影片：

- [AWS re：Invent 2023 – 【LAUNCH】 現代工作負載的應用程式監控](#)
- [AWS re：Invent 2023 – 實作應用程式可觀測性](#)
- [AWS re：Invent 2023 – 建立有效的可觀測性策略](#)
- [AWS re：Invent 2023 – AWS Distro for 的無縫可觀測性 OpenTelemetry](#)
- [上的應用程式效能管理 AWS](#)

相關範例：

- [AWS 適用於 Linux Workloads Immersion Day - Amazon CloudWatch](#)
- [監控 Amazon ECS叢集和容器](#)
- [使用 Amazon CloudWatch 儀表板進行監控](#)
- [Amazon EKS研討會](#)

PERF02-BP04 設定運算資源並適當調整大小

設定運算資源及適當調整其大小，以符合工作負載的效能需求，並避免未充分使用資源或過度使用資源的情況。

常見的反模式：

- 您忽略工作負載效能需求，導致過度佈建或佈建不足的運算資源。
- 您只選擇適用於所有工作負載的最大或最小執行個體。
- 為了方便管理，只能使用一個執行個體系列。
- 您可以忽略來自 AWS Cost Explorer 或 Compute Optimizer 的建議，以適當調整大小。
- 您未重新評估工作負載是否適用於新執行個體類型。
- 您只驗證組織的少量執行個體組態。

建立此最佳實務的優勢：透過避免資源的過度佈建和佈建不足，適當調整運算資源的大小可確保雲端中的最佳操作。適當調整運算資源的大小，通常可以提高效能和增強客戶體驗，同時降低成本。

未建立此最佳實務時的曝險等級：中

實作指引

適當調整大小可讓組織以有效率且符合成本效益的方式操作雲端基礎架構，同時滿足其業務需求。過度佈建雲端資源可能會導致額外成本，而佈建不足可能會導致效能不佳和負面的客戶體驗。AWS 可提供 [AWS Compute Optimizer](#) 和 [AWS Trusted Advisor](#) 等工具，這些工具使用歷史資料提供建議，以適當調整運算資源大小。

實作步驟

- 選擇最適合您需求的執行個體類型：
 - [如何為工作負載選擇適當的 Amazon EC2 執行個體類型？](#)
 - [Amazon EC2 Fleet 的屬性型執行個體類型選取](#)
 - [使用屬性型執行個體類型選取範圍來建立 Auto Scaling 群組](#)
 - [利用 Karpenter 整合來最佳化 Kubernetes 運算成本](#)
- 分析工作負載的各種效能特性，以及這些特性與記憶體、網路和 CPU 用量的關係。使用此資料，可以選擇最適合您工作負載描述檔和效能目標的資源。
- 使用 Amazon CloudWatch 之類的 AWS 監控工具，監控資源使用情況。
- 為運算資源選取適合的組態。

- 對於臨時工作負載，請評估[執行個體 Amazon CloudWatch 指標](#)，例如 CPUUtilization，來確定執行個體是否未充分使用或使用過度。
- 對於穩定的工作負載，請定期檢查 AWS 適當調整大小的工具 (例如 AWS Compute Optimizer 和 AWS Trusted Advisor)，以找出對運算資源進行最佳化和適當調整大小的機會。
- 在即時環境中實作之前，先測試非生產環境中的組態變更。
- 持續重新評估新的運算供應項目，並且根據工作負載需求進行比較。

資源

相關文件：

- [使用 AWS 進行雲端運算](#)
- [Amazon EC2 執行個體類型](#)
- [Amazon ECS 容器：Amazon ECS 容器執行個體](#)
- [Amazon EKS 容器：Amazon EKS 工作節點](#)
- [函數：Lambda 函數組態](#)
- [您的 Amazon EC2 執行個體的處理器狀態控制](#)

相關影片：

- [Amazon EC2 基礎](#)
- [AWS re:Invent 2023 – AWS Graviton：AWS 工作負載的最佳價格效能](#)
- [AWS re:Invent 2023 – AWS 管理主控台 中的最新 Amazon EC2 生成式 AI 功能](#)
- [AWS re:Invent 2023 – Amazon EC2 的新功能](#)
- [AWS re:Invent 2023 - 智慧型節省：Amazon EC2 成本優化策略](#)
- [AWS re:Invent 2021 – 為新一代 Amazon EC2 提供支援：深入探索 Nitro 系統](#)
- [AWS re:Invent 2019 – Amazon EC2 基礎](#)

相關範例：

- [AWS Compute Optimizer 示範程式碼](#)
- [Amazon EKS 研討會](#)
- [適當調整大小的建議](#)

PERF02-BP05 動態擴展您的運算資源

為滿足需求，請使用雲端的彈性，來動態擴充或縮減運算資源，並避免為工作負載佈建過多或過少的容量。

常見的反模式：

- 您可以手動增加容量，對警示做出反應。
- 使用與內部部署相同的大小規模準則 (通常是靜態基礎設施)。
- 您在擴展事件之後維持增加容量，而不是縮減規模。

建立此最佳實務的優勢：設定和測試運算資源的彈性可協助您節省成本、維持效能基準，並隨著流量變化提升可靠性。

未建立此最佳實務時的曝險等級：高

實作指引

AWS 透過各種擴展機制，提供了動態擴展或縮減資源的彈性，以滿足需求的變化。結合與運算相關的指標，動態擴展允許工作負載自動回應變更，並使用最佳運算資源集來實現目標。

您可以使用多種不同的方法達到資源的供需平衡。

- 目標追蹤法：監控您的擴展指標，並視需要自動增加或減少容量。
- 預測擴展：縮減每日和每週趨勢的預期。
- 基於排程的方法：按照排程來擴展可讓您根據可預測的負載變化來設定自己的擴展排程。
- 服務擴展：選擇可根據設計自動擴展的服務 (例如無伺服器)。

您必須確保工作負載部署可以同時處理向上擴展和縮減規模事件。

實作步驟

- 運算執行個體、容器和函數提供了彈性機制，可與自動擴展功能結合使用，或是作為服務功能提供。以下是自動擴展機制的幾個範例：

自動擴展機制	在哪裡使用
Amazon EC2 Auto Scaling	確保您有正確的 Amazon EC2 執行個體數量可應付應用程式的使用者負載。

自動擴展機制	在哪裡使用
Application Auto Scaling	自動將個別 AWS 服務的資源擴展到 Amazon EC2 以外，例如 AWS Lambda 函數或 Amazon Elastic Container Service (Amazon ECS) 服務。
Kubernetes Cluster Autoscaler/Karpenter	自動擴展 Kubernetes 叢集。

- 我們常將擴展與 Amazon EC2 執行個體或 AWS Lambda 函數等運算服務一起討論。請務必同時考慮非運算服務的組態 (例如 [AWS Glue](#)) 以符合需求。
- 確認用於擴展的指標符合要部署之工作負載的特性。如果您要部署影片轉碼應用程式，則預期為 100% CPU 使用率，且不應做為您的主要指標。請改用轉碼任務佇列的深度。您可以將[自訂指標](#)用於擴展政策 (如有必要)。若要選擇正確的指標，請考量 Amazon EC2 的下列指引：
 - 指標應為有效的使用率指標，並說明執行個體的忙碌程度。
 - 指標值必須與 Auto Scaling 群組中的執行個體數成比例增加或減少。
- 對於 Auto Scaling 群組，確保使用[動態擴展](#)，而非[手動擴展](#)。我們也建議您在動態擴展中使用[目標追蹤擴展政策](#)。
- 確認工作負載部署可同時處理擴展事件 (擴充和縮減)。例如，您可以使用[活動歷史記錄](#)來驗證 Auto Scaling 群組的擴展活動。
- 評估工作負載以取得可預測模式，並在預計發生預測中的變化和隨需規劃變化時主動擴展。透過預測性擴展，可以消除過度佈建容量的需求。如需詳細資訊，請參閱 [Predictive Scaling with Amazon EC2 Auto Scaling](#)。

資源

相關文件：

- [使用 AWS 進行雲端運算](#)
- [Amazon EC2 執行個體類型](#)
- [Amazon ECS 容器：Amazon ECS 容器執行個體](#)
- [Amazon EKS 容器：Amazon EKS 工作節點](#)
- [函數：Lambda 函數組態](#)
- [您的 Amazon EC2 執行個體的處理器狀態控制](#)
- [深入探討 Amazon ECS 叢集自動擴展](#)

- [介紹 Karpenter - 一個開放原始碼的高效能 Kubernetes Cluster Autoscaler](#)

相關影片：

- [AWS re:Invent 2023 – AWS Graviton：AWS 工作負載的最佳價格效能](#)
- [AWS re:Invent 2023 – AWS 管理主控台的全新 Amazon EC2 生成式 AI 功能](#)
- [AWS re:Invent 2023 – Amazon EC2 的新功能](#)
- [AWS re:Invent 2023 - 智慧型節省：Amazon EC2 成本優化策略](#)
- [AWS re:Invent 2021 – 為新一代 Amazon EC2 提供支援：深入探索 Nitro 系統](#)
- [AWS re:Invent 2019 – Amazon EC2 基礎](#)

相關範例：

- [Amazon EC2 Auto Scaling 群組範例](#)
- [Amazon EKS 研討會](#)
- [透過在 IPv6 上執行來擴展您的 Amazon EKS 工作負載](#)

PERF02-BP06 使用最佳化的硬體型運算加速器

使用硬體加速器執行特定功能，比以 CPU 為基礎的替代方案更有效率。

常見的反模式：

- 在工作負載中，您尚未基準化分析一般用途執行個體和專用執行個體，而專用執行個體可以改善效能和降低成本。
- 您使用硬體型運算加速器來執行任務，比起使用以 CPU 為基礎的替代方案更有效率。
- 未監控 GPU 使用率。

建立此最佳實務的優勢：透過使用硬體型加速器，例如圖形處理單元 (GPU) 和現場可程式化閘道陣列 (FPGA)，您就可以更有效率地執行特定處理功能。

未建立此最佳實務時的曝險等級：中

實作指引

加速運算執行個體可讓您存取硬體型運算加速器，例如 GPU 和 FPGA。這些硬體加速器比基於 CPU 的替代品更有效地執行某些功能，例如圖形處理或資料模式匹配。許多加速的工作負載 (例如轉譯、轉

碼和機器學習) 在資源使用方面變化很大。只在需要時執行此硬體，並在不需要時自動停用它們，以提高整體效能的效率。

實作步驟

- 確定哪些[加速運算執行個體](#)可以滿足您的需求。
- 針對機器學習工作負載，請利用專供工作負載使用的專用硬體，例如 [AWS Trainium](#)、[AWS Inferentia](#) 和 [Amazon EC2 DL1](#)。AWS 與同類 Amazon EC2 執行個體相比，Inferentia 執行個體 (例如 Inf2 執行個體) [所提供的效能功耗比要高出 50%](#)。
- 收集加速運算執行個體的用量指標。例如，可以使用 CloudWatch 代理程式為您的 GPU 收集 utilization_gpu 和 utilization_memory 等指標，如[使用 Amazon CloudWatch 收集 NVIDIA GPU 指標](#)中所示。
- 優化硬體加速器的程式碼、網路運作和設定，以確保系統會充分利用基礎硬體。
 - [最佳化 GPU 設定](#)
 - [深度學習 AMI 中的 GPU 監控和最佳化](#)
 - [將 I/O 最佳化以針對 Amazon SageMaker AI 中的深度學習訓練進行 GPU 效能調校](#)
- 使用最新的高效能程式庫和 GPU 驅動程式。
- 使用自動化來釋出不使用的 GPU 執行個體。

資源

相關文件：

- [在 Amazon Elastic Container Service 上使用 GPU](#)
- [GPU 執行個體](#)
- [AWS Trainium 的執行個體](#)
- [AWS Inferentia 的執行個體](#)
- [開始建構吧！使用自訂晶片和加速器來進行建構](#)
- [加速運算](#)
- [Amazon EC2 VT1 執行個體](#)
- [如何為工作負載選擇適當的 Amazon EC2 執行個體類型？](#)
- [選擇最佳 AI 加速器和模型編譯來以 Amazon SageMaker AI 推斷電腦視覺](#)

相關影片：

- [AWS re:Invent 2021 - 如何選擇 Amazon Elastic Compute Cloud GPU 執行個體進行深度學習](#)
- [AWS re:Invent 2022 - \[最新發佈！\] 介紹基於 AWS Inferentia2 的 Amazon EC2 Inf2 執行個體](#)
- [AWS re:Invent 2022 - 使用 AWS Trainium 加速深度學習並加快創新速度](#)
- [AWS re:Invent 2022 - 透過 NVIDIA 在 AWS 上進行深度學習：從訓練到部署](#)

相關範例：

- [Amazon SageMaker AI 和 NVIDIA GPU Cloud \(NGC\)](#)
- [搭配使用 SageMaker AI 與 Trainium 和 Inferentia，進行最佳化的深度學習訓練和推論工作負載](#)
- [使用 Amazon SageMaker AI 中的 Amazon Elastic Compute Cloud Inf1 執行個體最佳化 NLP 模型](#)

資料管理

問題

- [PERF 3. 如何在工作負載中儲存、管理和存取資料？](#)

PERF 3. 如何在工作負載中儲存、管理和存取資料？

特定系統的最佳資料管理解決方案會根據資料類型 (區塊、檔案或物件)、存取模式 (隨機或循序)、所需輸送量、存取頻率 (線上、離線、封存)、更新頻率 (WORM、動態) 及可用性和耐用性限制而有所不同。Well-Architected 工作負載會使用專用資料存放區，這些存放區採用不同的功能以提升效能。

最佳實務

- [PERF03-BP01 使用最能滿足資料存取和儲存需求的專用資料存放區](#)
- [PERF03-BP02 評估資料存放區的可用組態選項](#)
- [PERF03-BP03 收集並記錄資料存放區效能指標](#)
- [PERF03-BP04 實作策略以提高資料存放區中的查詢效能](#)
- [PERF03-BP05 實作利用快取的資料存取模式](#)

PERF03-BP01 使用最能滿足資料存取和儲存需求的專用資料存放區

了解資料特性 (例如可共用、大小、快取大小、存取模式、延遲、輸送量和資料的持續性)，為工作負載選擇適合的專用資料存放區 (儲存或資料庫)。

常見的反模式：

- 由於具備某種特定類型資料庫解決方案的內部經驗和知識，您堅持使用某個資料存取區。
- 您假設所有工作負載都有類似的資料儲存和存取需求。
- 您未實作資料目錄以清查資料資產。

建立此最佳實務的優勢：了解資料特性和需求，可協助您判斷能滿足工作負載需求的最有效率且效能最高的儲存技術。

未建立此最佳實務時的曝險等級：高

實作指引

選取和實作資料儲存時，請確定查詢、擴展和儲存特性支援工作負載資料需求。AWS 提供多種資料儲存和資料庫技術，包括區塊儲存、物件儲存、串流儲存、檔案系統、關聯式、鍵值、文件、記憶體內、圖形、時間序列和總帳資料庫。每個資料管理解決方案都有為您提供的選項和組態，以支援您的使用案例和資料模型。透過了解資料特性和需求，您可以擺脫整體式儲存技術和限制性、一刀切的方法，以專注於適當地管理資料。

實作步驟

- 對您工作負載現有的各種資料類型執行清查。
- 了解並記錄資料特性和需求，包括：
 - 資料類型 (非結構化、半結構化、關聯式)
 - 資料量與成長
 - 資料耐用性：持續性、暫時性、臨時
 - ACID (原子性、一致性、隔離性、持久性) 要求
 - 資料存取模式 (大量讀取或大量寫入)
 - 延遲
 - 輸送量
 - IOPS (每秒輸入/輸出操作次數)
 - 資料保留期間
- 了解 AWS 上可用於工作負載的不同資料存放區 ([儲存體](#)和[資料庫](#)服務)，這些資料存放區可以滿足資料特性，詳情請參閱 [PERF01-BP01 了解並理解可用的雲端服務和功能](#)。AWS 儲存技術及其重要特性的一些範例包含：

類型	AWS 服務	重要特性
物件儲存	Amazon S3	具有不受限的可擴展性、高可用性，以及多個可存取性選項。對 Amazon S3 輸入和存取物件時，可以使用 Transfer Acceleration 或 Access Points 之類的服務來支援您的位置、安全需求和存取模式。
封存儲存	Amazon Glacier	專為資料封存而打造。
串流儲存空間	Amazon Kinesis Amazon Managed Streaming for Apache Kafka (Amazon MSK)	快速地擷取和儲存串流資料。
共用檔案系統	Amazon Elastic File System (Amazon EFS)	可供多種類型的運算解決方案存取的可掛載檔案系統。
共用檔案系統	Amazon FSx	建置於最新的 AWS 運算解決方案之上，用以支援四個常用的檔案系統：NetApp ONTAP、OpenZFS、Windows File Server 和 Lustre。Amazon FSx 的 延遲 、 輸送量 和 IOPS 會隨著檔案系統而不同，當您為工作負載需求選取適當的檔案系統時，應予以考量。

類型	AWS 服務	重要特性
區塊儲存	Amazon Elastic Block Store (Amazon EBS)	專為 Amazon Elastic Compute Cloud (Amazon EC2) 設計的可擴展、高效能區塊儲存服務。Amazon EBS 包含支援 SSD 的儲存，適用於交易型、IOPS 密集型工作負載，以及適用於輸送量密集型工作負載的支援 HDD 的儲存。
關聯式資料庫	Amazon Aurora 、 Amazon RDS 、 Amazon Redshift 。	旨在支援 ACID (單元性、一致性、隔離行為、持續性) 交易，並維護參考完整性和強大的資料一致性。許多傳統應用程式、企業資源規劃 (ERP)、客戶關係管理 (CRM) 和電子商務都使用關聯式資料庫來儲存資料。
鍵值資料庫	Amazon DynamoDB	已針對常見的存取模式進行最佳化，通常用於儲存和擷取大量資料。高流量 Web 應用程式、電子商務系統和遊戲應用程式是鍵值資料庫的典型使用案例。
文件資料庫	Amazon DocumentDB	旨在將半結構化資料儲存為 JSON 類文件。這些資料庫可協助開發人員快速建置和更新應用程式，例如內容管理、目錄和使用者設定檔。

類型	AWS 服務	重要特性
記憶體資料庫	Amazon ElastiCache 、 Amazon MemoryDB for Redis	適用於需要即時存取資料、最低延遲和最高輸送量的應用程式。您可以將記憶體資料庫用於應用程式快取、工作階段管理、遊戲排行榜、低延遲 ML 特徵存放區、微型服務簡訊系統，以及高輸送量串流機制
圖形資料庫	Amazon Neptune	適用於此類應用程式：必須在高度連線圖形資料集之間，大規模導覽和查詢數百萬個關係，並且在過程中僅有毫秒延遲。許多公司使用圖形資料庫進行詐騙偵測、社交聯網和推薦引擎。
時間序列資料庫	Amazon Timestream	可快速地從隨時間變化的資料收集、合成和衍生洞見。IoT 應用程式、DevOps 和工業遙測可以利用時間序列資料庫。
寬欄	Amazon Keyspaces (適用於 Apache Cassandra)	可使用表格、列和欄，但與關聯式資料庫不同，在同一個表格中，欄的名稱和格式會因列而異。您通常會在大規模工業應用程式中看到寬欄存放區，用於設備維護、叢集管理和路由優化。

類型	AWS 服務	重要特性
總帳	Amazon Quantum Ledger Database (Amazon QLDB)	可提供集中化且受信任的機構，為每個應用程式維護可擴展、不可變且以密碼編譯方式驗證的交易記錄。我們看到用於記錄、供應鏈、註冊甚至銀行交易系統的總帳資料庫。

- 如果您要建置資料平台，請利用 AWS 上的[現代資料架構](#)來整合資料湖、資料倉儲和專用資料存放區。
- 為工作負載選擇資料存放區時，需要考慮的關鍵問題如下：

問題	重要考慮事項
如何建構資料？	• 如果資料是非結構化的，請考慮諸如 Amazon S3 等物件存放區，或諸如 Amazon DocumentDB 等 NoSQL 資料庫 • 對於索引鍵值資料，請考慮使用 DynamoDB、Amazon ElastiCache (Redis OSS) 或 Amazon MemoryDB
需要哪種層級的參考完整性？	• 對於外部索引鍵限制，Amazon RDS 和 Aurora 等關聯式資料庫可以提供此等級的完整性。 • 通常情況下，在 NoSQL 資料模型中，您會將資料去正規化到單一文件或文件集中，以在單個請求中擷取，而不是跨文件或資料表聯結。
ACID (單元性、一致性、隔離行為、持續性) 是否需要合規？	• 如果需要與關聯式資料庫相關聯的 ACID 屬性，請考慮關聯式資料庫，例如 Amazon RDS 和 Aurora。

問題	重要考慮事項
	如果 NoSQL 資料庫 需要強大的一致性，您可以搭配使用高度一致性讀取與 DynamoDB。
儲存要求如何隨時間變更？這如何影響可擴展性？	DynamoDB 和 Amazon Quantum Ledger Database (Amazon QLDB) 等無伺服器資料庫將動態擴展。 - 關聯式資料庫在佈建的儲存體上有上限，而且一旦達到這些限制，通常必須使用碎片等機制進行水平分割。
讀取查詢與寫入查詢的比例是多少？快取可能改善效能嗎？	- 如果資料庫為 DynamoDB，則讀取量繁重的工作負載可以受益於快取層，例如 ElastiCache 或 DAX。 - 讀取也可以卸載至具有關聯式資料庫的讀取複本，例如 Amazon RDS。
儲存和修改 (OLTP - 線上交易處理) 或擷取和報告 (OLAP - 線上分析處理) 是否具有更高的優先順序？	- 對於高輸送量的按原樣讀取交易處理，請考慮使用 NoSQL 資料庫，例如 DynamoDB。 - 對於具有一致性的高輸送量和複雜讀取模式 (例如聯結)，請使用 Amazon RDS。 - 對於分析查詢，請考慮使用 Amazon Redshift 之類的單欄式資料庫，或將資料匯出到 Amazon S3，然後使用 Athena 或 Amazon Quick 執行分析。

問題	重要考慮事項
資料需要哪種層級的耐久性？	• Aurora 會自動跨區域內的三個可用區域複寫資料，這表示資料高度耐用且資料遺失的機會較低。 • DynamoDB 會自動跨多個可用區域複寫，具有高可用性和資料耐久性。 • Amazon S3 提供 11 個九的耐用性。許多資料庫服務 (例如 Amazon RDS 和 DynamoDB) 支援將資料匯出至 Amazon S3，進行長期保留和封存。
是否希望擺脫商務資料庫引擎、或授權成本？	• 考慮開放原始碼引擎，例如 Amazon RDS 或 Aurora 上的 PostgreSQL 和 MySQL。 • 利用 AWS Database Migration Service 和 AWS Schema Conversion Tool，從商務資料庫引擎遷移至開放原始碼
對資料庫的操作期望是什麼？移至受管服務是否為主要問題？	• 利用 Amazon RDS 而非 Amazon EC2，以及利用 DynamoDB 或 Amazon DocumentDB 而非自行託管 NoSQL 資料庫，可以減少營運開銷。
目前如何存取資料庫？它是否只是應用程式存取，或是否有商業智能 (BI) 使用者和其他連網的現成應用程式？	• 如果您依賴於外部工具，則可能必須保持與其所支援之資料庫的相容性。Amazon RDS 與它支援的差異引擎版本完全相容，包括 Microsoft SQL Server、Oracle、MySQL 和 PostgreSQL。

- 在非生產環境中執行實驗和基準測試，以確定哪個資料存放區最能滿足您的工作負載需求。

資源

相關文件：

- [Amazon EBS 磁碟區類型](#)
- [Amazon EC2 儲存](#)

- [Amazon EFS : Amazon EFS 效能](#)
- [Amazon FSx for Lustre 效能](#)
- [Amazon FSx for Windows File Server 效能](#)
- [Amazon Glacier : Amazon Glacier 文件](#)
- [Amazon S3 : 請求率和效能考量](#)
- [AWS 的雲端儲存](#)
- [Amazon EBS I/O 特性](#)
- [的雲端資料庫AWS](#)
- [AWS 資料庫快取](#)
- [DynamoDB Accelerator](#)
- [Amazon Aurora 最佳實務](#)
- [Amazon Redshift 效能](#)
- [Amazon Athena 10 大效能秘訣](#)
- [Amazon Redshift Spectrum 最佳實務](#)
- [Amazon DynamoDB 最佳實務](#)
- [在 Amazon EC2 和 Amazon RDS 之間進行選擇](#)
- [實作 Amazon ElastiCache 的最佳實務](#)

相關影片：

- [AWS re:Invent 2023 : 提高 Amazon Elastic Block Store 效率並更具成本效益](#)
- [AWS re:Invent 2023 : 使用 Amazon Simple Storage Service 最佳化儲存價格和效能](#)
- [AWS re:Invent 2023 : 在 Amazon Simple Storage Service 上建置和最佳化資料湖](#)
- [AWS re:Invent 2022 : 在 AWS 上建置現代資料架構](#)
- [AWS re:Invent 2022 : 在 AWS 上建置資料網格架構](#)
- [AWS re:Invent 2023 : 深入探索 Amazon Aurora 及其創新](#)
- [AWS re:Invent 2023 : 使用 Amazon DynamoDB 的進階資料建模](#)
- [AWS re:Invent 2022 : 使用專用資料庫將應用程式現代化](#)
- [深入探討 Amazon DynamoDB : 進階設計模式](#)

相關範例：

- [AWS 專用資料庫研討會](#)
- [專為開發人員打造的資料庫](#)
- [AWS 現代資料架構 Immersion Day](#)
- [在 AWS 上建置資料網格](#)
- [Amazon S3 範例](#)
- [使用 Amazon Redshift 資料共用來最佳化資料模式](#)
- [資料庫遷移](#)
- [MS SQL Server - AWS Database Migration Service \(AWS DMS\) 複寫示範](#)
- [資料庫現代化實際操作研討會](#)
- [Amazon Neptune 範例](#)

PERF03-BP02 評估資料存放區的可用組態選項

了解並評估資料存放區可用的各種功能和組態選項，以最佳化工作負載的儲存空間和效能。

常見的反模式：

- 所有工作負載只能使用一種儲存類型，例如 Amazon EBS。
- 您為所有工作負載使用已佈建的 IOPS，卻未針對所有儲存層進行實際測試。
- 您不知道所選資料管理解決方案的組態選項。
- 您完全依賴於增加執行個體大小，而不查看其他可用的組態選項。
- 您並不測試資料存放區的擴展特性。

建立此最佳實務的優勢：藉由探索和試驗資料存放區組態，您能夠降低基礎架構成本、改善效能，以及減少維護工作負載所需的工作量。

未建立此最佳實務時的曝險等級：中

實作指引

工作負載可以根據資料儲存和存取需求，使用一個或多個資料存放區。要優化效能達成效率和成本，您必須評估資料存取模式，以判斷適當的資料存放區組態。在探索資料存放區選項時，請考量各種層面，例如儲存選項、記憶體、運算、讀取複本、一致性要求、連線集區以及快取選項。嘗試使用這些不同的組態選項來改善效能達成效率指標。

實作步驟

- 了解資料存放區的目前組態 (例如執行個體類型、儲存體大小或資料庫引擎版本)。
- 審核 AWS 文件和最佳實務，以了解可協助改善資料存放區效能的建議組態選項。要考慮的關鍵資料存放區選項如下：

組態選項	範例
卸載讀取 (例如讀取複本和快取)	• 對於 DynamoDB 資料表，可以使用 DAX 卸載讀取以進行快取。 • 您可以建立 Amazon ElastiCache (Redis OSS) 叢集，並將應用程式設定為先從快取讀取，如果請求的項目不存在，則會退回到資料庫。 • 關聯式資料庫 (例如 Amazon RDS 和 Aurora) 以及佈建的 NoSQL 資料庫 (例如 Neptune 和 Amazon DocumentDB) 都支援新增讀取複本，以卸載工作負載的讀取部分。 • 諸如 DynamoDB 等無伺服器資料庫將自動擴展。確定已佈建足夠的讀取容量單位 (RCU) 來處理工作負載。
擴展寫入 (例如分區金鑰碎片或引進佇列)	• 對於關聯式資料庫，可以增加執行個體的大小以容納增加的工作負載，或增加已佈建的 IOP 以增加基礎儲存體的輸送量。 • 也可以在資料庫前面引入佇列，而不是直接寫入資料庫。此模式可讓您將擷取與資料庫分離並控制流速，這樣資料庫就不會不堪重負。 • 批次處理寫入請求，而不是建立許多短期交易，這有助於改善高寫入量關聯式資料庫的輸送量。 • DynamoDB 等無伺服器資料庫可以自動調整寫入輸送量，或根據容量模式調整已佈建的寫入容量單位 (WCU) 來調整寫入輸送量。

組態選項	範例
	當您達到指定分割區索引鍵的輸送量限制時，仍然可能會遇到熱分割區的問題。透過選擇更均勻分佈的分割區索引鍵，或分片寫入分割區索引鍵，來緩解此問題。
使用政策來管理資料集的生命週期	- 可以使用 Amazon S3 生命週期，以在整個生命週期中管理物件。如果存取模式不明、會變化或是無法預測，則可以使用 Amazon S3 Intelligent-Tiering，讓其監控存取模式，並自動將未存取的物件移至成本較低的存取層。可以利用 Amazon S3 Storage Lens 指標，找出生命週期管理中的最佳化機會和差距。 Amazon EFS 生命週期管理 會自動管理檔案系統的檔案儲存。
連線管理與集區	- Amazon RDS Proxy 可以與 Amazon RDS 和 Aurora 一起使用，以管理資料庫的連線。 - 無伺服器資料庫 (例如 DynamoDB) 沒有與其相關聯的連線，但請考慮已佈建的容量和自動擴展政策來處理負載中的高峰。

- 在非生產環境中執行實驗和基準測試，以確定哪個組態選項能滿足您的工作負載需求。
- 完成試驗之後，請規劃遷移並確認效能指標。
- 使用 AWS 監控 (例如 [Amazon CloudWatch](#)) 和最佳化 (例如 [Amazon S3 Storage Lens](#)) 工具，以透過實際使用模式持續最佳化資料存放區。

資源

相關文件：

- [AWS 的雲端儲存](#)
- [Amazon EBS 磁碟區類型](#)
- [Amazon EC2 儲存](#)

- [Amazon EFS : Amazon EFS 效能](#)
- [Amazon FSx for Lustre 效能](#)
- [Amazon FSx for Windows File Server 效能](#)
- [Amazon Glacier : Amazon Glacier 文件](#)
- [Amazon S3 : 請求率和效能考量](#)
- [Amazon EBS I/O 特性](#)
- [AWS 的雲端資料庫](#)
- [AWS 資料庫快取](#)
- [DynamoDB Accelerator](#)
- [Amazon Aurora 最佳實務](#)
- [Amazon Redshift 效能](#)
- [Amazon Athena 10 大效能秘訣](#)
- [Amazon Redshift Spectrum 最佳實務](#)
- [Amazon DynamoDB 最佳實務](#)

相關影片：

- [AWS re:Invent 2023 : 提高 Amazon Elastic Block Store 效率並更具成本效益](#)
- [AWS re:Invent 2023 : 使用 Amazon Simple Storage Service 最佳化儲存價格和效能](#)
- [AWS re:Invent 2023 : 在 Amazon Simple Storage Service 上建置和最佳化資料湖](#)
- [AWS re:Invent 2023 : AWS 檔案儲存最新消息](#)
- [AWS re:Invent 2023 : 深入了解 Amazon DynamoDB](#)

相關範例：

- [AWS 專用資料庫研討會](#)
- [專為開發人員打造的資料庫](#)
- [AWS 現代資料架構 Immersion Day](#)
- [Amazon EBS 自動擴展](#)
- [Amazon S3 範例](#)
- [Amazon DynamoDB 範例](#)

- [AWS 資料庫遷移範例](#)
- [資料庫現代化研討會](#)
- [使用 Amazon RDS for Postgress 資料庫上的參數](#)

PERF03-BP03 收集並記錄資料存放區效能指標

追蹤並記錄資料存放區的相關績效指標，以了解資料管理解決方案的成效。這些指標可協助您最佳化資料存放區、確認是否符合工作負載需求，並提供工作負載執行方式的清晰概觀。

常見的反模式：

- 您只使用手動日誌檔案來搜尋指標。
- 您只會將指標發佈到團隊使用的內部工具，而不會全面了解您的工作負載。
- 您只會使用所選監控軟體記錄的預設指標。
- 您只會在有問題時審查指標。
- 您只監控系統層級指標，而沒有擷取資料存取或用量指標。

建立此最佳實務的優勢：建立效能基準可協助您了解工作負載的正常行為和需求。異常模式可以更快地識別和偵錯，進而改善資料存放區的效能和可靠性。

未建立此最佳實務時的曝險等級：高

實作指引

要監控資料存放區的效能，您必須記錄一段時間的多個效能指標。這可讓您偵測異常情況，並根據業務指標衡量效能，以確認您是否滿足工作負載需求。

指標應包括支援資料存放區的基礎系統和資料庫指標。基礎系統指標可能包括 CPU 使用率、記憶體、可用磁碟儲存體、磁碟 I/O、快取命中率以及網路輸入和輸出指標，而資料存放區指標可能包括每秒交易數、常用查詢、平均查詢速率、回應時間、索引使用狀況、表格鎖定、查詢逾時以及開啟的連線數目。此資料對於了解工作負載的執行情況以及資料管理解決方案的使用方式至關重要。將這些指標納入資料驅動的方法，以調整和優化工作負載的資源。

使用工具、程式庫和系統來記錄與資料庫效能有關的效能測量值。

實作步驟

- 找出要追蹤的資料存放區關鍵效能指標。

- [Amazon S3 指標和維度](#)
- [監控 Amazon RDS 執行個體中的指標](#)
- [在 Amazon RDS 上使用 Performance Insights 監控資料庫負載](#)
- [增強型監視概觀](#)
- [DynamoDB 指標和維度](#)
- [監控 DynamoDB Accelerator](#)
- [使用 Amazon CloudWatch 監控 Amazon MemoryDB](#)
- [應監控哪些指標？](#)
- [監控 Amazon Redshift 叢集效能](#)
- [Timestream 指標和維度](#)
- [Amazon Aurora 的 Amazon CloudWatch 指標](#)
- [Amazon Keyspaces \(適用於 Apache Cassandra\) 中的日誌記錄和監控](#)
- [監控 Amazon Neptune 資源](#)
- 使用核准的日誌記錄和監控解決方案來收集這些指標。[Amazon CloudWatch](#) 可以收集架構中各種資源的指標。您還可以收集和發佈自訂指標以顯示業務或衍生指標。使用 CloudWatch 或第三方解決方案，設定可指出何時超過閾值的警示。
- 檢查資料存放區監控是否能從可偵測效能異常的機器學習解決方案中獲益。
 - [適用於 Amazon RDS 的 Amazon DevOps Guru](#) 可檢視效能問題，並提出修正動作的建議。
- 在監控和日誌記錄解決方案中設定資料保留，以符合安全性和營運目標。
 - [CloudWatch 指標的預設資料保留](#)
 - [CloudWatch Logs 的預設資料保留](#)

資源

相關文件：

- [AWS 資料庫快取](#)
- [Amazon Athena 10 大效能秘訣](#)
- [Amazon Aurora 最佳實務](#)
- [DynamoDB Accelerator](#)
- [Amazon DynamoDB 最佳實務](#)
- [Amazon Redshift Spectrum 最佳實務](#)

- [Amazon Redshift 效能](#)
- [AWS 的雲端資料庫](#)
- [Amazon RDS Performance Insights](#)

相關影片：

- [AWS re:Invent 2022 - 使用 Amazon RDS 和 Aurora 進行效能監控，採用 Autodesk](#)
- [使用 Amazon DevOps Guru for Amazon RDS 進行資料庫效能監控和調整](#)
- [AWS re:Invent 2023 - AWS 檔案儲存最新消息](#)
- [AWS re:Invent 2023 - 深入了解 Amazon DynamoDB](#)
- [AWS re:Invent 2023 - 在 Amazon S3 上建置和最佳化資料湖](#)
- [AWS re:Invent 2023 - AWS 檔案儲存最新消息](#)
- [AWS re:Invent 2023 - 深入了解 Amazon DynamoDB](#)
- [在 Amazon ElastiCache 上監控 Redis 工作負載的最佳實務](#)

相關範例：

- [AWS 資料集擷取指標收集架構](#)
- [Amazon RDS 監控研討會](#)
- [AWS 專用資料庫研討會](#)

PERF03-BP04 實作策略以提高資料存放區中的查詢效能

實作策略以最佳化資料並改善資料查詢，以便為工作負載提供更高的可擴展性和更高效的效能。

常見的反模式：

- 您沒有分割資料存放區中的資料。
- 您在資料存放區中僅以一種檔案格式儲存資料。
- 您沒有在資料存放區中使用索引。

建立此最佳實務的優勢：最佳化資料和查詢效能可提高效率、降低成本並改善使用者體驗。

未建立此最佳實務時的曝險等級：中

實作指引

資料最佳化和查詢調整是資料存放區中效能效率的關鍵層面，因為其會影響整個雲端工作負載的效能和回應能力。未經過最佳化的查詢可能會使用更多資源和造成更大的瓶頸，進而降低資料存放區的整體效率。

資料最佳化包括數個技術，以確保高效的資料儲存和存取。這也有助於提高資料存放區中的查詢效能。關鍵策略包括資料分割、資料壓縮和資料去常規化，這些都有助最佳化資料的儲存和存取。

實作步驟

- 了解和分析在資料存放區中執行的重要資料查詢。
- 找出資料存放區中執行速度緩慢的查詢，並使用查詢計畫了解其目前狀態。
 - [分析 Amazon Redshift 中的查詢計畫](#)
 - [在 Athena 中使用 EXPLAIN 和 EXPLAIN ANALYZE](#)
- 實作策略以改善查詢效能。有些關鍵策略包括下列情況：
 - 使用[單欄式檔案格式](#) (例如 Parquet 或 ORC)。
 - 壓縮資料存放區中的資料以減少儲存空間和 I/O 作業。
 - 資料分割可將資料拆分為較小的部分並縮短資料掃描時間。
 - [在 Athena 中分割資料](#)
 - [分割區與資料分配](#)
 - 在查詢中對共同欄進行資料索引編制。
 - 針對頻繁查詢使用具體化視觀表。
 - [了解具體化視觀表](#)
 - [在 Amazon Redshift 中建立具體化視觀表](#)
 - 選擇正確的聯結作業以進行查詢。當您聯結兩個資料表時，請在聯結左側指定較大的資料表，並在聯結右側指定較小的資料表。
 - 分散式快取解決方案可改善延遲並減少資料庫 I/O 操作的次數。
 - 定期維護，例如[清空](#)、重新索引以及[執行統計](#)。
- 在非生產環境中實驗和測試策略。

資源

相關文件：

- [Amazon Aurora 最佳實務](#)
- [Amazon Redshift 效能](#)
- [Amazon Athena 10 大效能秘訣](#)
- [AWS 資料庫快取](#)
- [實作 Amazon ElastiCache 的最佳實務](#)
- [在 Athena 中分割資料](#)

相關影片：

- [AWS re:Invent 2023 - AWS 儲存成本最佳化最佳實務](#)
- [AWS re:Invent 2022 - 使用 Amazon RDS 和 Aurora 進行效能監控，採用 Autodesk](#)
- [使用新的查詢分析工具最佳化 Amazon Athena 查詢](#)

相關範例：

- [AWS 專用資料庫研討會](#)

PERF03-BP05 實作利用快取的資料存取模式

實作可受益於快取資料的存取模式，以便快速擷取經常存取的資料。

常見的反模式：

- 快取頻繁變更的資料。
- 您依賴快取資料，就好像它是持久存儲並始終可用一樣。
- 您不考慮快取資料的一致性。
- 您不監控快取實作的效率。

建立此最佳實務的優勢：將資料儲存在快取中可改善讀取延遲、讀取輸送量、使用者體驗和整體效率，並降低成本。

未建立此最佳實務時的風險暴露等級：中

實作指引

快取是旨在存儲資料的軟體或硬體組件，以便更快或更有效地滿足未來對相同資料的請求。如果存儲在快取中的資料丟失，可以透過重複之前的計算或從另一個資料存放區中擷取來進行重建。

資料快取可能是改善整體應用程式效能並減輕基礎主要資料來源負擔的最有效策略之一。可以在應用程式的多個層級快取資料，例如在進行遠端呼叫的應用程式內 (稱為用戶端快取)，或使用快速次要服務來儲存資料 (稱為遠端快取)。

用戶端快取

透過用戶端快取，每個用戶端 (查詢後端資料儲存的應用程式或服務) 都可以在指定的時間內，在本機儲存其唯一查詢的結果。這可以先檢查本機用戶端快取，來減少網路對資料存放區的請求數量。如果結果不存在，應用程式便可查詢資料存放區，並將這些結果儲存在本機。此模式允許每個用戶端將資料儲存在最接近的位置 (用戶端本身)，從而達到最低的延遲。當後端資料存放區無法使用時，用戶端也可以繼續提供某些查詢，從而提高整體系統的可用性。

這種方法的一個缺點是，當涉及多個用戶端時，它們可能會在本地存儲相同的快取資料。這會導致這些用戶端之間的重複儲存用量和資料不一致。一個用戶端可能會快取查詢結果，一分鐘後，另一個用戶端可以執行相同查詢並獲得不同結果。

遠端快取

為了解決用戶端之間的重複資料問題，可以使用快速外部服務或遠端緩存來存儲查詢的資料。每個用戶端都會在查詢後端資料存放區之前檢查遠端快取，而非檢查本機資料存放區。此策略可實現用戶端之間更一致的回應、更好的儲存資料效率以及更高的快取資料量，因為儲存空間會獨立於用戶端進行擴展。

遠端快取的缺點是整個系統可能會遇到較高延遲，因為需要額外的網路跳轉來檢查遠端快取。用戶端快取可以與遠端快取一起用於多層級快取，以改善延遲。

實作步驟

- 識別可受益於快取的資料庫、API 和網路服務。具有大量讀取工作負載、高讀寫比率或擴展成本較高的服務都是快取的候選者。
 - [資料庫快取](#)
 - [啟用 API 快取以提升回應能力](#)
- 找出最適合您的存取模式的適當快取策略類型。
 - [快取策略](#)
 - [AWS 快取解決方案](#)

- 遵循資料存放區的[快取最佳實務](#)。
- 為所有資料設定快取失效策略，例如存留時間 (TTL)，以平衡資料新鮮度並降低後端資料存放區壓力。
- 在用戶端中啟用自動連線重試、指數退避、用戶端逾時和連線集區等功能 (如果可用)，因為它們可以改善效能和可靠性。
 - [最佳實務：Redis 用戶端和 Amazon ElastiCache \(Redis OSS\)](#)
- 監控快取命中率，目標為 80% 或更高。較低的值可能表示快取大小不足，或者無法從快取中受益的存取模式。
 - [應監控哪些指標？](#)
 - [在 Amazon ElastiCache 上監控 Redis 工作負載的最佳實務](#)
 - [使用 Amazon CloudWatch 搭配 Amazon ElastiCache \(Redis OSS\) 進行監控的最佳實務](#)
- 實作[資料複寫](#)，將讀取卸載至多個執行個體，並提高資料讀取效能和可用性。

資源

相關文件：

- [使用 Amazon ElastiCache Well-Architected Lens](#)
- [使用 Amazon CloudWatch 搭配 Amazon ElastiCache \(Redis OSS\) 進行監控的最佳實務](#)
- [應監控哪些指標？](#)
- [《利用 Amazon ElastiCache 大規模提高效能》白皮書](#)
- [快取挑戰和策略](#)

相關影片：

- [Amazon ElastiCache 學習路徑](#)
- [使用 Amazon ElastiCache 最佳實務打造邁向成功的設計](#)
- [AWS re:Invent 2020 - 使用 Amazon ElastiCache 最佳實務打造邁向成功的設計](#)
- [AWS re:Invent 2023 - \[發佈\] Amazon ElastiCache 無伺服器簡介](#)
- [AWS re:Invent 2022 - 使用 Redis 重塑資料層級的 5 個好方法](#)
- [AWS re:Invent 2021 - 深入了解 Amazon ElastiCache \(Redis OSS\)](#)

相關範例：

- [使用 Amazon ElastiCache \(Redis OSS\) 提升 MySQL 資料庫效能](#)

聯網與內容交付

問題

- [PERF 4. 如何在工作負載中選取和設定聯網資源？](#)

PERF 4. 如何在工作負載中選取和設定聯網資源？

工作負載的最佳聯網解決方案會根據延遲、輸送量需求、抖動和頻寬而有所不同。實體限制 (例如使用者或內部部署資源) 會決定位置選項。這些限制可能隨著邊緣節點或資源位置而有所差異。

最佳實務

- [PERF04-BP01 了解聯網如何影響效能](#)
- [PERF04-BP02 評估可用的聯網功能](#)
- [PERF04-BP03 為工作負載選擇適當的專用連線或 VPN](#)
- [PERF04-BP04 使用負載平衡將流量分配到多個資源](#)
- [PERF04-BP05 選擇網路通訊協定以提高效能](#)
- [PERF04-BP06 根據網路需求選擇工作負載的位置](#)
- [PERF04-BP07 根據指標最佳化網路組態](#)

PERF04-BP01 了解聯網如何影響效能

分析並了解網路相關決策如何影響您的工作負載，以提供高效的效能並改善使用者體驗。

常見的反模式：

- 所有流量都流經現有資料中心。
- 可以透過中央防火牆路由所有流量，而非使用雲端原生網路安全工具。
- 佈建 AWS Direct Connect 連線，無須了解實際使用需求。
- 定義聯網解決方案時，不會考慮工作負載特性和加密開銷。
- 對於雲端中的聯網解決方案，您可以使用內部部署概念和策略。

建立此最佳實務的優勢：了解聯網如何影響工作負載效能，有助於您識別潛在瓶頸、改善使用者體驗、提高可靠性並在工作負載變更時減少操作維護。

未建立此最佳實務時的風險暴露等級：高

實作指引

網路負責處理應用程式元件、雲端服務、邊緣網路和內部部署資料之間的連線，因此對工作負載效能可能有嚴重影響。除了工作負載效能外，使用者體驗也會受到網路延遲、頻寬、通訊協定、位置、網路擁塞、抖動、輸送量和路由規則的影響。

取得工作負載的已記錄在案的聯網需求清單，包括延遲、封包大小、路由規則、通訊協定和支援的流量模式。審核可用的聯網解決方案，確定哪些服務符合您的工作負載網路特性。雲端型網路可以快速重建，因此隨著時間演進您的網路架構是提高效能達成效率的必要條件。

實作步驟：

- 定義並記錄聯網效能需求，包括網路延遲、頻寬、通訊協定、位置、流量模式 (尖峰和頻率)、輸送量、加密、檢查和路由規則等指標。
- 了解 [VPC](#)、[AWS Direct Connect](#)、[Elastic Load Balancing \(ELB\)](#) 和 [Amazon Route 53](#) 等關鍵 AWS 聯網服務。
- 擷取下列主要聯網特性：

特性	工具與指標
基礎聯網特性	• VPC 流程日誌 • AWS Transit Gateway 流量日誌 • AWS Transit Gateway 指標 • AWS PrivateLink 指標
應用程式聯網特性	• Elastic Fabric Adapter • AWS App Mesh 指標 • Amazon API Gateway 指標
邊緣聯網特性	• Amazon CloudFront 指標 • Amazon Route 53 指標 • AWS Global Accelerator 指標
混合聯網特性	• Direct Connect 指標 • AWS Site-to-Site VPN 指標

特性	工具與指標
	• AWS Client VPN 指標 • AWS 雲端 WAN 指標
安全聯網特性	• AWS Shield、AWS WAF 和 AWS Network Firewall 指標
追蹤特性	• AWS X-Ray • VPC Reachability Analyzer • 網路存取分析器 • Amazon Inspector • Amazon CloudWatch RUM

- 基準測試並測試網路效能：
 - [基準測試](#)網路輸送量，因為當執行個體位於相同 VPC 時，某些因素可能會影響 Amazon EC2 網路效能。測量相同 VPC 中 Amazon EC2 Linux 執行個體之間的網路頻寬。
 - 執行[負載測試](#)，試驗聯網解決方案和選項。

資源

相關文件：

- [Application Load Balancer](#)
- [Linux 上的 EC2 增強型聯網](#)
- [Windows 上的 EC2 增強型聯網](#)
- [EC2 置放群組](#)
- [在 Linux 執行個體上啟用搭配彈性網路介面卡 \(ENA\) 的增強型聯網](#)
- [Network Load Balancer](#)
- [AWS 的聯網產品](#)
- [轉換閘道](#)
- [轉換到 Amazon Route 53 中的以延遲為基礎的路由](#)
- [VPC 端點](#)

相關影片：

- [AWS re:Invent 2023 - AWS 聯網基礎](#)
- [AWS re:Invent 2023 - 聯網可以為您的應用程式做些什麼？](#)
- [AWS re:Invent 2023 - 進階 VPC 設計及新功能](#)
- [AWS re:Invent 2023 - 雲端聯網開發人員指南](#)
- [AWS re:Invent 2019 - 與 AWS 和混合 AWS 網路架構的連線](#)
- [AWS re:Invent 2019 - 最佳化 Amazon EC2 執行個體的網路效能](#)
- [AWS Summit Online - 改善應用程式的全球網路效能](#)
- [AWS re:Invent 2020 - Well-Architected Framework 的聯網最佳實務和秘訣](#)
- [AWS re:Invent 2020 - 大規模遷移中的 AWS 聯網最佳實務](#)

相關範例：

- [AWS Transit Gateway 以及可擴展的安全性解決方案](#)
- [AWS 聯網研討會](#)
- [網路防火牆實際操作研討會](#)
- [在 AWS 中觀察並診斷您的網路](#)
- [在 AWS 中查找並解決網路設定錯誤](#)

PERF04-BP02 評估可用的聯網功能

評估雲端中可提升效能的聯網功能。透過測試、指標和分析來測量這些功能的影響。例如，利用可用的網路層級功能來降低延遲、網路距離或抖動。

常見的反模式：

- 您只在單一區域中活動，這是因為該區域是您總部的所在區域。
- 可以使用防火牆而非安全群組來篩選流量。
- 您會中斷 TLS 以進行流量檢查，而不是仰賴安全群組、端點政策和其他雲端原生功能。
- 您只會使用子網路來分隔，而非採用安全群組的方式。

建立此最佳實務的優勢：評估所有服務功能和選項可提高工作負載效能、降低基礎架構成本、減少維護工作負載所需的人力，以及提升整體安全狀態。您可以使用全球 AWS 骨幹，為客戶提供最佳的聯網體驗。

未建立此最佳實務時的風險暴露等級：高

實作指引

AWS 提供 [AWS Global Accelerator](#) 和 [Amazon CloudFront](#) 等服務，可協助改善網路效能，而大多數 AWS 服務則具有可最佳化網路流量的產品功能 (例如 [Amazon S3 Transfer Acceleration](#) 功能)。

審核您可以使用哪些網路相關組態選項，及其對工作負載可能有何影響。效能最佳化取決於了解這些選項如何與您的架構互動，以及它們對衡量的效能與使用者體驗的影響。

實作步驟

- 建立工作負載元件清單。
 - 建立統一的全球網路時，請考慮使用 [AWS 雲端 WAN](#) 來建置、管理及監控組織的網路。
 - 使用 [Amazon CloudWatch Logs 指標](#) 監控全球核心網路。利用 [Amazon CloudWatch RUM](#)，它提供了有助於識別、了解和增強使用者數位體驗的洞見。
 - 檢視 AWS 區域 和可用區域之間以及每個可用區域內的彙總網路延遲，使用 [AWS Network Manager](#) 深入了解應用程式效能與基礎 AWS 網路效能的關聯性。
 - 使用現有的組態管理資料庫 (CMDB) 工具或服務，例如 [AWS Config](#)，建立工作負載的詳細目錄及其設定方式。
- 如果這是現有的工作負載，則請識別並記錄效能指標的基準，並著重於瓶頸和要改善的領域。效能相關聯網指標會依據業務需求和工作負載特性而有所不同。首先，這些指標對於審核工作負載可能很重要：頻寬、延遲、封包遺失、抖動和重新傳輸。
- 如果這是新的工作負載，則請執行[負載測試](#)以識別效能瓶頸。
- 對於您找出的效能瓶頸，請審核您解決方案的組態選項，以找出改善效能的機會。查看下列主要聯網選項和功能：

改進機會	解決方案
網路的路徑或路由	使用 網路存取分析器 來識別路徑或路由。
網路通訊協定	請參閱 PERF04-BP05 選擇網路通訊協定以提高效能
網路拓撲	連接多個帳戶時，評估 VPC 對等互連 和 AWS Transit Gateway 之間的操作和效能權衡。AWS Transit Gateway 可簡化互連所有 VPC 的方式，這些 VPC 可跨越數千個 AWS 帳戶和內部部署網路。使用 AWS Resource

改進機會	解決方案
	Access Manager 在多個帳戶間共用您的 AWS Transit Gateway : 請參閱 PERF04-BP03 為工作負載選擇適當的專用連線或 VPN
網路服務	AWS Global Accelerator 是一項聯網服務，可使用 AWS 全域網路基礎架構將使用者流量效能提升最多達 60%。 Amazon CloudFront 可以在全球範圍改善工作負載內容交付和延遲的效能。 使用 Lambda @edge 執行函數，這些函數可自訂 CloudFront 提供的更接近使用者的內容，減少延遲並改善效能。 Amazon Route 53 提供以延遲為基礎的路由、地理位置路由、地理位置鄰近性路由和以 IP 為基礎的路由選項，可協助您為全球使用者改善工作負載的效能。當工作負載分佈在全球範圍時，請審核工作負載流量和使用者位置，找出能夠最佳化工作負載效能的路由選項。
儲存功能資源	Amazon S3 Transfer Acceleration 是一個功能，它可讓外部使用者從 CloudFront 的聯網優化中受益，以將資料上傳到 Amazon S3。這樣就可以更輕易地從與 AWS 雲端 沒有專用連線的遠端位置輸送大量資料。 Amazon S3 多區域存取點可將內容複製到多個區域，並透過提供一個存取點來簡化工作負載。使用多區域存取點時，您可以使用可識別最低延遲儲存貯體的服務，來要求資料或將資料寫入 Amazon S3。

改進機會	解決方案
運算資源功能	Amazon EC2 執行個體、容器和 Lambda 函式所使用的彈性網路介面 (ENI) 會按個別流程受到限制。審核置放群組以最佳化 EC2 聯網輸送量。若要避免個別流程的瓶頸，請將應用程式設計為使用多個流程。若要監控及檢視您的運算相關聯網指標，請使用 CloudWatch Metrics 和 ethtool。ethtool 命令包含在 ENA 驅動程式中，並公佈了其他網路相關指標，這些指標可作為自訂指標發佈到 CloudWatch。 Amazon 彈性網路介面卡 (ENA) 可為叢集置放群組中的執行個體提供更好的輸送量，從而提供進一步優化。 Elastic Fabric Adapter (EFA) 是 Amazon EC2 執行個體的網路介面，讓您能夠在 AWS 上大規模執行需要高階節點間通訊的工作負載。 經 Amazon EBS 優化的執行個體使用優化的組態堆疊，可提供更多專用容量以提高 Amazon EBS I/O。

資源

相關文件：

- [Application Load Balancer](#)
- [Linux 上的 EC2 增強型聯網](#)
- [Windows 上的 EC2 增強型聯網](#)
- [EC2 置放群組](#)
- [在 Linux 執行個體上使用彈性網路介面卡 \(ENA\) 啟用增強型聯網](#)
- [Network Load Balancer](#)
- [的聯網產品AWS](#)
- [轉換到 Amazon Route 53 中的以延遲為基礎的路由](#)

- [VPC 端點](#)
- [VPC 流量日誌](#)

相關影片：

- [AWS re:Invent 2023 – 是否為下一步做好準備？設計網路實現增長和靈活性](#)
- [AWS re:Invent 2023 - 進階 VPC 設計及新功能](#)
- [AWS re:Invent 2023 - 雲端聯網開發人員指南](#)
- [AWS re:Invent 2022 – 深入了解 AWS 聯網基礎設施](#)
- [AWS re:Invent 2019 – AWS 和混合式 AWS 網路架構的連線](#)
- [AWS re:Invent 2018 - 最佳化 Amazon EC2 執行個體的網路效能](#)
- [AWS Global Accelerator](#)

相關範例：

- [AWS Transit Gateway 以及可擴展的安全性解決方案](#)
- [AWS 聯網研討會](#)
- [觀察並診斷您的網路](#)
- [在 AWS 中查找並解決網路設定錯誤](#)

PERF04-BP03 為工作負載選擇適當的專用連線或 VPN

需要混合式連線來連接內部部署資源和雲端資源時，請佈建足夠的頻寬以滿足您的效能要求。預估混合工作負載的頻寬和延遲需求。這些數字將促進調整大小需求。

常見的反模式：

- 您只會針對網路加密需求評估 VPN 解決方案。
- 不會評估備份或備援連線選項。
- 您無法識別所有工作負載需求 (加密、通訊協定、頻寬和流量需求)。

建立此最佳實務的優勢：選擇和設定適當的連線解決方案將提高工作負載的可靠性並將效能最大化。藉由識別工作負載要求、提前規劃和評估混合解決方案，您將最大限度地減少昂貴的實體網路變更和營運負擔，同時延長上市時間。

未建立此最佳實務時的曝險等級：高

實作指引

根據頻寬需求開發混合式聯網架構。[Direct Connect](#) 允許您私下將內部部署網路與 AWS 連線。需要高頻寬、低延遲同時可達到一致效能時，適合這個選項。VPN 連線會建立透過網際網路的安全連線。在以下情況使用它：當只需要臨時連線時、當成本是一個考慮因素時、或者在使用 Direct Connect 時等待建立彈性物理網路連線作為應急措施時。

如果您的頻寬需求很高，可以考慮使用多個 Direct Connect 或 VPN 服務。流量可以跨服務進行負載平衡，儘管由於延遲和頻寬差異，我們不建議在 Direct Connect 和 VPN 之間進行負載平衡。

實作步驟

- 預估現有應用程式的頻寬和延遲要求。
 - 針對移至 AWS 的現有工作負載，利用來自您的內部網路監控系統的資料。
 - 針對您沒有監控資料的新工作負載或現有工作負載，請諮詢產品擁有者以確定足夠的效能指標，並且提供良好的使用者體驗。
- 選取專用連線或 VPN 做為您的連線選項。根據所有工作負載要求 (加密、頻寬和流量需求)，您可以選擇 AWS Direct Connect 或 [Site-to-Site VPN](#) (或兩者)。下圖可協助您選擇適當的連線類型。
 - [AWS Direct Connect](#) 使用專用連線或託管連線，提供 AWS 環境的專用連線，範圍從 50 Mbps 到 100 Gbps。這可為您提供受管和受控的延遲以及佈建頻寬，因此您的工作負載可以有效率地連線到其他環境。使用 AWS Direct Connect 合作夥伴，您可以擁有來自多個環境的端對端連線能力，提供擴充的網路和一致的效能。AWS 提供使用原生 100 Gbps、連結彙總群組 (LAG) 或 BGP 等價多路徑 (ECMP) 的擴展直接連線連線頻寬。
 - AWS [Site-to-Site VPN](#) 提供受管 VPN 服務，支援網際網路通訊協定安全性 (IPsec)。建立 VPN 連線時，每個 VPN 連線都包含兩個通道以獲得高可用性。
- 依照 AWS 說明文件選擇適當的連線選項：
 - 如果決定使用 Direct Connect，請為您的連線選取適當的頻寬。
 - 如果跨多個位置使用 AWS Site-to-Site VPN 以連線到 AWS 區域，請使用 [已加速的 Site-to-Site VPN 連線](#)，以提高網路效能。
 - 如果您的網路設計包含 [AWS Direct Connect](#) 上的 IPsec VPN 連線，請考慮使用私有 IP VPN 來改善安全性並實現分段。[AWS Site-to-Site Private IP VPN](#) 會部署在傳輸虛擬介面 (VIF) 之上。
 - [AWS Direct Connect SiteLink](#) 透過繞過 AWS 區域在 [AWS Direct Connect 地點](#) 之間的最快路徑上傳送資料，允許在您的全球資料中心間建立低延遲和備援連線。

- 在部署到生產環境之前驗證連線設定。執行安全性和效能測試，以確保其符合您的頻寬、可靠性、延遲和合規性要求。
- 定期監控您的連線效能和使用情況，並視需要進行最佳化。

確定性效能流程圖

資源

相關文件：

- [AWS 的聯網產品](#)
- [AWS Transit Gateway](#)
- [VPC 端點](#)
- [建立可擴展且安全的多個 VPC AWS 網路架構](#)
- [Client VPN](#)

相關影片：

- [AWS re:Invent 2023 – 使用 AWS 建立混合式網路連線](#)
- [AWS re:Invent 2023 – AWS 的安全遠端連線](#)
- [AWS re:Invent 2022 – 利用 Amazon CloudFront 最佳化效能](#)
- [AWS re:Invent 2019 – AWS 和混合式 AWS 網路架構的連線](#)
- [AWS re:Invent 2020 – AWS Transit Gateway Connect](#)

相關範例：

- [AWS Transit Gateway 以及可擴展的安全性解決方案](#)
- [AWS 聯網研討會](#)

PERF04-BP04 使用負載平衡將流量分配到多個資源

在多個資源或服務之間分配流量，以讓您的工作負載能夠利用雲端提供的彈性。您也可以使用負載平衡來卸載加密終止，以提升效能、可靠性，以及有效管理和路由流量。

常見的反模式：

- 您在選擇負載平衡器類型時未考慮工作負載需求。
- 您不利用負載平衡器功能來進行效能最佳化。
- 工作負載在不使用負載平衡器的情況下，直接公開到網際網路。
- 您可以透過現有的負載平衡器路由所有網際網路流量。
- 您可以使用一般 TCP 負載平衡，並讓每個運算節點處理 SSL 加密。

建立此最佳實務的優勢：負載平衡器會處理單一可用區域中或跨多個可用區域的應用程式流量不同的負載，並實現高可用性、自動擴展及更充分利用您的工作負載。

未建立此最佳實務時的曝險等級：高

實作指引

負載平衡器會做為您的工作負載的進入點，從那裡將您的流量分散到後端目標，例如運算執行個體或容器，以提高利用率。

選擇正確的負載平衡器類型是最佳化架構的第一步。從列出您的工作負載特性開始，例如通訊協定 (例如 TCP、HTTP、TLS 或 WebSockets)、目標類型 (例如執行個體、容器或無伺服器)、應用程式要求 (例如長時間執行連線、使用者身分驗證或黏性) 和置放 (例如 Region、Local Zone、Outpost 或區域隔離)。

AWS 為您的應用程式提供了多種模型來使用負載平衡。[Application Load Balancer](#) 最適合 HTTP 和 HTTPS 流量的負載平衡，並提供了針對現代應用程式架構 (包括微型服務和容器) 交付的進階請求路由。

[Network Load Balancer](#) 最適合需要極高效能的 TCP 流量的負載平衡。它能夠每秒處理數百萬個請求，同時保持超低延遲性，並且還進行優化，可處理突發的和不穩定的流量模式。

[Elastic Load Balancing](#) 提供整合的憑證管理和 SSL/TLS 解密，讓您能夠靈活地集中管理負載平衡器的 SSL 設定，並從工作負載中卸載 CPU 密集型工作。

選擇正確的負載平衡器之後，您可以開始利用其功能來減少後端為流量提供服務所需投入的工作量。

例如，同時使用 Application Load Balancer (ALB) 和 Network Load Balancer (NLB)，您可以執行 SSL/TLS 加密卸載，這是避免 CPU 密集型 TLS 交握由您的目標完成，並且改善憑證管理的機會。

在您的負載平衡器中設定 SSL/TLS 卸載時，它會負責往返用戶端的流量的加密，同時將未加密的流量交付給您的後端，釋放您的後端資源並且改善用戶端的回應時間。

Application Load Balancer 也可以為 HTTP/2 流量提供服務，不需要在您的目標上支援它。這個簡單的決策可以改善您的應用程式回應時間，因為 HTTP/2 更有效率地使用 TCP 連線。

定義架構時，應該考慮您的工作負載延遲要求。例如，如果您有對延遲敏感的應用程式，您可能會決定使用 Network Load Balancer，以獲得極低的延遲。另外，您可能會決定讓工作負載更靠近您的客戶，也就是利用 [AWS Local Zones](#) 或甚至 [AWS Outposts](#) 中的 Application Load Balancer。

對延遲敏感的工作負載的另一個考慮是跨區域負載平衡。使用跨區域負載平衡，每個負載平衡器節點會將已註冊目標之間的流量分散到所有允許的可用區域中。

使用與您的負載平衡器整合的 Auto Scaling。效能效率系統的其中一個關鍵層面與適當調整後端資源大小有關。若要完成此操作，您可以利用後端目標資源的負載平衡器整合。使用與 Auto Scaling 群組整合的負載平衡器，目標會視需要從負載平衡器新增或移除，以因應傳入流量。負載平衡器也可以針對容器化工作負載與 [Amazon ECS](#) 和 [Amazon EKS](#) 整合。

- [Amazon ECS - 服務負載平衡](#)
- [Amazon EKS 上的應用程式負載平衡](#)
- [Amazon EKS 上的網路負載平衡](#)

實作步驟

- 定義您的負載平衡需求，包括流量、可用性和應用程式可擴展性。
- 為您的應用程式選擇正確的負載平衡器類型。
 - 針對 HTTP/HTTPS 工作負載使用 Application Load Balancer。
 - 針對在 TCP 或 UDP 上執行的非 HTTP 工作負載使用 Network Load Balancer。
 - 如果要利用這兩種產品的功能，請使用兩者的組合 ([ALB 作為 NLB 的目標](#))。例如，如果您想要搭配使用 NLB 的靜態 IP 與來自 ALB 的 HTTP 標題型路由，或者如果您想要將您的 HTTP 工作負載公開到 [AWS PrivateLink](#)。
 - 如需負載平衡器的完整比較，請參閱 [ELB 產品比較](#)。
- 盡可能使用 SSL/TLS 卸載。
 - 使用與 [AWS Certificate Manager](#) 整合的 [Application Load Balancer](#) 和 [Network Load Balancer](#) 來設定 HTTPS/TLS 接聽程式。
 - 請注意，基於合規理由，某些工作負載可能需要端對端加密。在此情況下，必須允許在目標進行加密。
 - 如需安全性最佳實務，請參閱 [SEC09-BP02 強制執行傳輸中加密](#)。
- 選取正確的路由演算法 (僅 ALB)。

- 路由演算法可以造成您的後端目標的妥善使用程度和它們影響效能程度的差異。例如，ALB [為路由演算法提供兩個選項](#)：
- 最低未解決請求：針對應用程式的請求因複雜性而異，或目標因處理功能而異的情況，用來讓負載更妥善地分散到您的後端目標。
- 循環配置：當請求和目標類似，或是如果您需要在目標之間平均分散請求時使用。
- 考慮跨區域或區域隔離。
 - 針對延遲改善和區域失敗網域使用跨區域關閉 (區域隔離)。在 NLB 中預設關閉它，在 [ALB 中，您可以根據每個目標群組將其關閉](#)。
 - 使用跨區域開啟來增加可用性和彈性。預設情況下，針對 ALB 關閉跨區域，[在 NLB 中，您可以根據每個目標群組將其關閉](#)。
- 為您的 HTTP 工作負載開啟 HTTP keep-alives (僅 ALB)。使用這項功能，負載平衡器可以重複使用後端連線，直到 keep-alive 逾時到期，改善您的 HTTP 請求和回應時間，同時減少您的後端目標上的資源使用率。有關如何為 Apache 和 Nginx 執行此操作的詳細資訊，請參閱[使用 Apache 或 NGINX 作為 ELB 的後端伺服器的最佳設定是什麼？](#)
- 開啟負載平衡器的監控功能。
 - 開啟 [Application Load Balancer](#) 和 [Network Load Balancer](#) 的存取記錄。
 - 針對 ALB 要考慮的主要欄位是 request_processing_time、request_processing_time 和 response_processing_time。
 - 針對 NLB 要考慮的主要欄位是 connection_time 和 tls_handshake_time。
 - 請準備好在您需要日誌時進行查詢。您可以使用 Amazon Athena 查詢 [ALB 日誌](#)和 [NLB 日誌](#)。
 - 建立效能相關指標的警示，例如 [ALB 的 TargetResponseTime](#)。

資源

相關文件：

- [ELB 產品比較](#)
- [AWS 全球基礎設施](#)
- [使用可用區域親和性改善效能並且降低成本](#)
- [使用 Amazon Athena 逐步執行日誌分析](#)
- [查詢 Application Load Balancer 日誌](#)
- [監控 Application Load Balancer](#)
- [監控 Network Load Balancer](#)

- [使用 Elastic Load Balancing 在 Auto Scaling 群組的執行個體中分配流量](#)

相關影片：

- [AWS re:Invent 2023：聯網可以為您的應用程式做些什麼？](#)
- [AWS re:Inforce 20：如何使用 Elastic Load Balancing 大規模增強您的安全態勢](#)
- [AWS re:Invent 2018：Elastic Load Balancing：深入探討和最佳實務](#)
- [AWS re:Invent 2021 - 如何為您的 AWS 工作負載選擇正確的負載平衡器](#)
- [AWS re:Invent 2019：針對不同工作負載充分發揮 Elastic Load Balancing](#)

相關範例：

- [Gateway Load Balancer](#)
- [使用 Amazon Athena 進行日誌分析的 CDK 和 CloudFormation 範例](#)

PERF04-BP05 選擇網路通訊協定以提高效能

根據對工作負載效能的影響，做出系統和網路間通訊協定的決策。

實現輸送量的延遲和頻寬之間存在關係。如果您的檔案傳輸使用傳輸控制協定 (TCP)，較高的延遲很可能會降低整體輸送量。有一些方法可以使用 TCP 調校和最佳化的傳輸通訊協定來解決這個問題，但有一個解決方案是利用使用者資料包通訊協定 (UDP)。

常見的反模式：

- 無論效能需求為何，您都可以將 TCP 用於所有工作負載。

建立此最佳實務的優勢：確認針對使用者與工作負載元件之間的通訊使用適當的通訊協定，可協助改善您的應用程式的整體使用者體驗。例如，無連線 UDP 雖然達到高速，但卻失去重新傳輸能力或高可靠性。TCP 是功能完整的通訊協定，但需要更大的額外負荷來處理封包。

未建立此最佳實務時的曝險等級：中

實作指引

如果您有能力為應用程式選擇不同的通訊協定，而且您具備此領域的專業知識，請使用不同的通訊協定來最佳化應用程式和使用者體驗。請注意，這種方法有很大的困難，如果您已先用其他方法最佳化應用程式，才可嘗試。

改善您的工作負載效能的主要考慮是了解延遲和輸送量需求，然後選擇可最佳化效能的網路通訊協定。

何時考慮使用 TCP

TCP 提供可靠的資料交付，並且可用於可靠性和保證資料交付很重要的工作負載元件間通訊。許多 Web 式應用程式仰賴 TCP 型通訊協定，例如 HTTP 和 HTTPS，針對應用程式元件之間的通訊開啟 TCP 通訊端。電子郵件和檔案資料傳輸也使用 TCP 的常見應用程式，因為它是應用程式元件之間的簡單且可靠的傳輸機制。使用 TLS 與 TCP 會增加一些通訊負擔，導致提高延遲和降低輸送量，但它具有安全性優勢。負擔主要來自交握處理的增加負擔，需要數個往返才能完成。一旦交握完成，加密和解密資料的負擔相對小。

何時考慮使用 UDP

UDP 是無連線導向的通訊協定，因此適用於需要快速、有效傳輸的應用程式，例如日誌、監控和 VoIP 資料。此外，如果您有會回應來自大量用戶端之小型查詢的工作負載元件，請考慮使用 UDP，以確保最佳工作負載效能。資料包傳輸層安全性 (DTLS) 等同於 Transport Layer Security (TLS) 的 UDP。使用 DTLS 與 UDP 時，負擔是來自加密和解密資料，因為交握處理已簡化。DTLS 也會對 UDP 封包增加小量負擔，因為它包含額外欄位以指出安全參數以及偵測竄改。

何時考慮使用 SRD

Scalable Reliable Datagram (SRD) 是針對高輸送量工作負載最佳化的網路傳輸通訊協定，因為它能夠在多個路徑之間負載平衡流量，並且快速從封包捨棄或連結失敗復原。因此，SRD 最適合用於高效能運算 (HPC) 工作負載，這些工作負載需要運算節點之間的高輸送量和低延遲通訊。這可能包含平行處理任務，例如牽涉到在節點之間大量資料傳輸的模擬、建模和資料分析。

實作步驟

- 使用 [AWS Global Accelerator](#) 和 [AWS Transfer Family](#) 服務來改善線上檔案傳輸應用程式的輸送量。AWS Global Accelerator 服務可協助您達成用戶端裝置與 AWS 上工作負載之間的較低延遲。使用 AWS Transfer Family，您可以使用 TCP 型通訊協定，例如 Secure Shell File Transfer Protocol (SFTP) 和 File Transfer Protocol over SSL (FTPS)，安全地擴展和管理您對於 AWS 儲存服務的檔案傳輸。
- 使用網路延遲來判斷 TCP 是否適合工作負載元件之間的通訊。如果您的用戶端應用程式與伺服器之間的網路延遲高，則 TCP 三向交握會耗費一些時間，因此會影響您的應用程式的回應能力。例如到第一個位元組的時間 (TTFB) 和往返時間 (RTT) 等指標可用來測量網路延遲。如果您的工作負載為使用者提供動態內容，請考慮使用 [Amazon CloudFront](#)，它會為動態內容的每個來源建立持續連線，以消除連線設定時間，否則會減慢每個用戶端請求的速度。
- 使用 TLS 與 TCP 或 UDP 會由於加密和解密的影響，導致對您的工作負載增加延遲和減少輸送量。對於這類工作負載，請考慮在 [Elastic Load Balancing](#) 上卸載 SSL/TLS，透過允許負載平衡器處理

SSL/TLS 加密和解密流程，而不是讓後端執行個體執行此操作，來提高工作負載效能。這可協助減少後端執行個體上的 CPU 使用率，可以改善效能和增加容量。

- 使用 [Network Load Balancer \(NLB\)](#) 部署依賴於 UDP 通訊協定的服務，例如驗證和授權、日誌記錄、DNS、IoT 和串流媒體，以改善工作負載的效能和可靠性。NLB 會在多個目標之間分散傳入 UDP 流量，讓您水平地擴展工作負載、增加容量以及減少單一目標的負擔。
- 對於高效能運算 (HPC) 工作負載，請考慮使用 [Elastic Network Adapter \(ENA\) Express](#) 功能，該功能使用 SRD 通訊協定，透過為 EC2 執行個體之間的網路流量提供更高的單一流量頻寬 (25 Gbps) 和更低的尾延遲 (99.9%) 來改善網路效能。
- 可以使用 [Application Load Balancer \(ALB\)](#)，在工作負載元件之間或 gRPC 用戶端與服務之間對 gRPC (遠端程序呼叫) 進行路由和負載平衡。gRPC 使用 TCP 型 HTTP/2 通訊協定進行傳輸，並提供效能優勢，例如更少的網路佔用空間、壓縮、高效的二進位序列化、支援多種語言以及雙向串流。

資源

相關文件：

- [如何將 UDP 流量路由到 Kubernetes](#)
- [Application Load Balancer](#)
- [Linux 上的 EC2 增強型聯網](#)
- [Windows 上的 EC2 增強型聯網](#)
- [EC2 置放群組](#)
- [在 Linux 執行個體上啟用搭配彈性網路介面卡 \(ENA\) 的增強型聯網](#)
- [Network Load Balancer](#)
- [AWS 的聯網產品](#)
- [轉換到 Amazon Route 53 中的以延遲為基礎的路由](#)
- [VPC 端點](#)

相關影片：

- [AWS re:Invent 2022 – 在新一代 Amazon Elastic Compute Cloud 執行個體上擴展網路效能](#)
- [AWS re:Invent 2022 – 應用程式聯網基礎](#)

相關範例：

- [AWS Transit Gateway 以及可擴展的安全性解決方案](#)
- [AWS 聯網研討會](#)

PERF04-BP06 根據網路需求選擇工作負載的位置

評估資源置放的選項以減少網路延遲和提高輸送量，藉由減少頁面載入和資料傳輸時間來提供最佳的使用者體驗。

常見的反模式：

- 您可以將所有工作負載資源合併到單一地理位置。
- 您選擇的區域最接近您的位置，但不是最接近工作負載最終使用者。

建立此最佳實務的優勢：使用者體驗因使用者與您的應用程式之間的延遲而大受影響。透過使用適當的 AWS 區域 AWS 私有全域網路，您可以減少延遲，並為遠端使用者提供更好的體驗。

未建立此最佳實務時的風險暴露等級：中

實作指引

資源，例如 Amazon EC2 執行個體，會放置在 [AWS 區域](#)、[AWS Local Zones](#)、[AWS Outposts](#) 或 [AWS Wavelength](#) 區域中的可用區域中。此位置的選擇會影響來自特定使用者位置的網路延遲和輸送量。[Amazon CloudFront](#) 和 等邊緣服務 [AWS Global Accelerator](#) 也可用於透過快取邊緣位置的內容，或透過 AWS 全球網路為使用者提供工作負載的最佳路徑來改善網路效能。

Amazon EC2 提供置放群組以進行聯網。置放群組是執行個體的邏輯分組，用於減少延遲。搭配支援的執行個體類型和彈性網路轉接器（ENA）使用置放群組，可讓工作負載參與低延遲、減少抖動的 25 Gbps 網路。建議將置放群組用於受益於低網路延遲、高網路輸送量或兩者兼而有之的工作負載。

延遲敏感服務會使用 AWS 全球網路在邊緣位置提供，例如 [Amazon CloudFront](#)。這些邊緣位置通常提供內容交付網路（CDN）和網域名稱系統（）等服務 DNS。透過在邊緣使用這些服務，工作負載可以對內容或 DNS 解決方案的請求做出低延遲的回應。這些服務還提供地理服務，例如內容的地理定位（根據最終使用者的位置提供不同的內容），或以延遲為基礎的路由，將最終使用者定向到最近區域的（最小延遲）。

使用邊緣服務來減少延遲及啟用內容快取。正確設定 DNS 和 HTTP/HTTPS 的快取控制，以從這些方法中獲益最多。

實作步驟

- 擷取與往返網路介面的 IP 流量有關的資訊。
 - [使用VPC流程日誌記錄 IP 流量](#)
 - [如何保留用戶端 IP 地址 AWS Global Accelerator](#)
- 分析您工作負載中的網路存取模式，以識別使用者如何使用您的應用程式。
 - 使用監控工具，例如 [Amazon CloudWatch](#) 和 [AWS CloudTrail](#)，收集網路活動的資料。
 - 分析資料以識別網路存取模式。
- 根據下列關鍵元素，為您的工作負載部署選取區域：
 - 資料所在位置：對於資料密集型應用程式 (例如大數據和機器學習)，應用程式碼執行時應盡可能接近資料。
 - 使用者所在位置：對於面向使用者的應用程式，請選擇接近工作負載使用者的一或多個區域。
 - 其他限制：考慮諸如成本和合規性之類的限制，如[為工作負載選取區域時應考慮的事項](#)中所述。
- 使用 [AWS Local Zones](#) 執行諸如影片轉譯等工作負載。Local Zones 可讓您因運算和儲存資源更接近最終使用者而獲益。
- [AWS Outposts](#) 適用於需要保持內部部署的工作負載，而您希望該工作負載能夠與 AWS 中的其他工作負載無縫執行。
- 5G 裝置需要 ultra-low-latency 高解析度即時影片串流、高擬真度音訊和擴增實境或虛擬實境 (AR/VR) 等應用程式。對於此類應用程式，請考慮在 5G 網路內[AWS Wavelength](#) AWS Wavelength 內嵌 AWS 運算和儲存服務，提供用於開發、部署和擴展 ultra-low-latency 應用程式的行動邊緣運算基礎設施。
- 針對常用資產，使用本機快取或 [AWS 快取解決方案](#) 以提升效能、減少資料移動以及降低環境影響。

服務	使用情況
Amazon CloudFront	使用 快取靜態內容，例如影像、指令碼和影片，以及動態內容，例如 API 回應或 Web 應用程式。
Amazon ElastiCache	用來快取 Web 應用程式的內容。
DynamoDB Accelerator	用來將記憶體內加速新增至 DynamoDB 資料表。

- 使用可協助您在更接近工作負載使用者的位置執行程式碼的服務，如下所示：

服務	使用情況
Lambda@Edge	用於在物件未經快取時起始的大量運算作業。
Amazon CloudFront Functions	用於可由短期函數啟動的簡單使用案例，例如 HTTP (s) 請求或回應操作。
AWS IoT Greengrass	用來為連線的裝置執行本機運算、傳訊和資料快取。

- 某些應用程式需要藉由減少第一個位元組延遲和抖動並且增加輸送量，來獲得固定的進入點或較高的效能。這些應用程式可以受益於提供靜態廣播 IP 地址和在邊緣位置TCP終止的網路服務。[AWS Global Accelerator](#) 可以為您的應用程式提升效能達 60%，並為多區域架構提供快速容錯移轉。AWS Global Accelerator 為您提供靜態廣播 IP 地址，作為託管於一或多個的應用程式固定進入點 AWS 區域。這些 IP 地址允許流量盡可能接近您的使用者傳入 AWS 全域網路。透過在用戶端與最接近用戶端的 AWS 邊緣位置之間建立TCP連線，AWS Global Accelerator 以減少初始連線設定時間。檢閱 [使用 AWS Global Accelerator](#)，以改善 TCP/UDP 工作負載的效能，並為多區域架構提供快速容錯移轉。

資源

相關的最佳實務：

- [COST07-BP02 根據成本實作區域](#)
- [COST08-BP03 實作 服務以降低資料傳輸成本](#)
- [REL10-BP01 將工作負載部署到多個位置](#)
- [REL10-BP02 為您的多位置部署選取適當的位置](#)
- [SUS01-BP01 根據業務需求和永續性目標選擇區域](#)
- [SUS02-BP04 根據其聯網需求最佳化工作負載的地理定位](#)
- [SUS04-BP07 將網路之間的資料移動降至最低](#)

相關文件：

- [AWS 全球基礎設施](#)
- [AWS 本機區域 和 AWS Outposts，為您的邊緣工作負載選擇正確的技術](#)

- [置放群組](#)
- [AWS 本機區域](#)
- [AWS Outposts](#)
- [AWS Wavelength](#)
- [Amazon CloudFront](#)
- [AWS Global Accelerator](#)
- [AWS Direct Connect](#)
- [AWS Site-to-Site VPN](#)
- [Amazon Route 53](#)

相關影片：

- [AWS Local Zones 解說器影片](#)
- [AWS Outposts：概觀和運作方式](#)
- [AWS re：Invent 2023 - 邊緣和內部部署工作負載的遷移策略](#)
- [AWS re：Invent 2021 - AWS Outposts：將 AWS 體驗帶到內部部署](#)
- [AWS re：Invent 2020：AWS Wavelength：在 5G 邊緣以超低延遲執行應用程式](#)
- [AWS re：Invent 2022 - AWS Local Zones：為分散式邊緣建置應用程式](#)
- [AWS re：Invent 2021 - 使用 Amazon 建置低延遲網站 CloudFront](#)
- [AWS re：Invent 2022 - 透過 改善效能和可用性 AWS Global Accelerator](#)
- [AWS re：Invent 2022 - 使用 建置您的全球廣域網路 AWS](#)
- [AWS re：Invent 2020：使用 Amazon Route 53 進行全域流量管理](#)

相關範例：

- [AWS Global Accelerator 自訂路由研討會](#)
- [使用邊緣函數處理重新撰寫和重新導向](#)

PERF04-BP07 根據指標最佳化網路組態

使用收集和分析的資料來做出有關優化網路組態的明智決策。

常見的反模式：

- 您假設所有效能相關問題都與應用程式有關。
- 您只能從靠近已部署工作負載的位置測試網路效能。
- 將預設組態用於所有網路服務。
- 您過度佈建網路資源來提供足夠的容量。

建立此最佳實務的優勢：收集必要的 AWS 網路指標並實作網路監控工具，可讓您了解網路效能並最佳化網路組態。

未建立此最佳實務時的曝險等級：低

實作指引

監控往返 VPC、子網路或網路介面的流量，對於了解如何利用 AWS 網路資源並最佳化網路組態而言非常重要。藉由使用下列 AWS 聯網工具，您可以進一步檢查流量用量、網路存取和日誌的相關資訊。

實作步驟

- 識別要收集的關鍵效能指標，例如延遲或封包遺失。AWS 會提供數種工具，可協助您收集這些指標。藉由使用下列工具，您可以進一步檢查流量用量、網路存取和日誌的相關資訊：

AWS 工具	在哪裡使用
Amazon VPC IP Address Manager。	使用 IPAM 來規劃、追蹤和監控您的 AWS 和內部部署工作負載的 IP 位址。這是最佳化 IP 位址用量和分配的最佳實務。
VPC 流程日誌	使用 VPC 流程日誌來擷取有關往返您的 VPC 中網路界面的詳細資訊。使用 VPC 流程日誌，您可以診斷過於嚴格或寬鬆的安全群組規則，並且判斷往返網路界面的流量方向。
AWS Transit Gateway 流量日誌	使用 AWS Transit Gateway Flow Logs 可擷取傳入及傳出傳輸閘道的 IP 流量的相關資訊。
DNS 查詢日誌記錄	記錄 Route 53 所收到的公有或私有 DNS 查詢的相關資訊。使用 DNS 日誌，您可以藉由了解請求的網域或子網域或是回應 DNS 查詢的 Route 53 邊緣節點，最佳化 DNS 組態。

AWS 工具	在哪裡使用
Reachability Analyzer	Reachability Analyzer 可幫助您分析和偵錯網路可達性。Reachability Analyzer 是一種組態分析工具，可讓您在 VPC 中的來源資源和目的地資源之間執行連線測試。此工具可協助您確認您的組態符合您預期的連線能力。
網路存取分析器	網路存取分析器可協助您了解對資源的網路存取。您可以使用網路存取分析器來指定您的網路存取需求，並識別未符合您的指定需求的潛在網路路徑。藉由最佳化您的對應網路組態，您可以了解及確認網路的狀態，並且示範 AWS 上的網路是否符合您的合規需求。
Amazon CloudWatch ()	使用 Amazon CloudWatch 並為網路選項開啟適當的指標。請確定為您的工作負載選擇正確的網路指標。例如，您可以開啟 VPC 網路地址用量、VPC NAT 閘道、AWS Transit Gateway、VPN 通道、AWS Network Firewall、Elastic Load Balancing 和 AWS Direct Connect 的指標。持續監控指標是觀察和了解您的網路狀態和用量的良好實務，並且可以協助您根據您的觀察來最佳化網路組態。
AWS Network Manager	使用 AWS Network Manager，可以監控 AWS 全球網路 的即時和歷史效能，以作營運和規劃之用。Network Manager 提供 AWS 區域和可用區域之間以及每個可用區域內的彙總網路延遲，讓您更好地了解應用程式效能與基礎 AWS 網路效能的關聯性。
Amazon CloudWatch RUM	使用 Amazon CloudWatch RUM 收集為您提供洞見的指標，協助您識別、了解和改善使用者體驗。

- 使用 VPC 和 AWS Transit Gateway Flow Logs 來識別最受歡迎者和應用程式流量模式。

- 評估並最佳化目前的網路架構，包括 VPC、子網路和路由。例如，您可以評估不同的 VPC 對等互連或 AWS Transit Gateway 如何協助您改善架構中的聯網。
- 評估網路中的路由路徑，以確認始終使用目的地之間的最短路徑。網路存取分析器可協助您執行此操作。

資源

相關文件：

- [公有 DNS 查詢日誌記錄](#)
- [什麼是 IPAM？](#)
- [什麼是 Reachability Analyzer？](#)
- [什麼是網路存取分析器？](#)
- [VPC 的 CloudWatch 指標](#)
- [使用 Apache Parquet 格式的 VPC 流程日誌，最佳化效能並降低網路分析成本](#)
- [使用 Amazon CloudWatch 指標監控全球核心網路](#)
- [持續監控網路流量和資源](#)

相關影片：

- [AWS re:Invent 2023 - 雲端聯網開發人員指南](#)
- [AWS re:Invent 2023 - 是否為下一步做好準備？設計網路實現增長和靈活性](#)
- [AWS re:Invent 2023 - 進階 VPC 設計及新功能](#)
- [AWS re:Invent 2022 – 深入了解 AWS 聯網基礎設施](#)
- [AWS re:Invent 2020 – AWS Well-Architected Framework 的聯網最佳實務和秘訣](#)
- [AWS re:Invent 2020 – 監控網路流量並進行疑難排解](#)

相關範例：

- [AWS 聯網研討會](#)
- [AWS 網路監控](#)
- [在 AWS 中觀察並診斷您的網路](#)
- [在 AWS 中查找並解決網路設定錯誤](#)

程序和文化

問題

- [PERF 5. 您的組織實務和文化如何促進工作負載的效能達成效率？](#)

PERF 5. 您的組織實務和文化如何促進工作負載的效能達成效率？

在架構工作負載時，您可以採取一些原則和實務來協助您更有效率地執行高效能雲端工作負載。為了培養高效能雲端工作負載的文化，請考慮下列重要原則和實務：

最佳實務

- [PERF05-BP01 建立用於測量工作負載運作狀態和效能的關鍵績效指標 \(KPI\)](#)
- [PERF05-BP02 使用監控解決方案了解效能最關鍵的領域](#)
- [PERF05-BP03 定義提高工作負載效能的程序](#)
- [PERF05-BP04 Load 測試工作負載](#)
- [PERF05-BP05 使用自動化主動修復效能相關問題](#)
- [PERF05-BP06 保留工作負載和服務 up-to-date](#)
- [PERF05-BP07 定期審查指標](#)

PERF05-BP01 建立用於測量工作負載運作狀態和效能的關鍵績效指標 (KPI)

識別定量和定性衡量工作負載效能的 KPI。KPI 有助於測量與業務目標相關的工作負載的運作狀態和效能。

常見的反模式：

- 您只能監控系統層級指標，以深入了解工作負載，而不了解這些指標的業務影響。
- 假設 KPI 已做為標準指標資料發佈和共用。
- 沒有定義定量、可衡量的 KPI。
- 沒有將 KPI 與業務目標或策略保持一致。

建立此最佳實務的優勢：找出代表工作負載健康狀態和效能的特定 KPI，有助團隊以一致的標準排定優先事項並定義成功的業務成果。與所有部門共用這些指標可提供閾值、期望和業務影響的可見性和一致性。

未建立此最佳實務時的曝險等級：高

實作指引

KPI 可讓業務和工程團隊以一致的標準衡量目標和策略，以及將這些因素結合以產生商業成果的方式。例如，網站工作負載可能使用頁面載入時間，作為整體效能的指示。此指標將是衡量使用者體驗的多個資料點之一。除了找出頁面載入時間閾值外，您還應該記錄未符合理想效能時預期的成果或業務風險。較長的頁面載入時間會直接影響您的使用者，降低其使用者體驗等級，並可能導致客戶流失。當您定義 KPI 閾值時，請同時結合業界基準和最終使用者期望。例如，如果目前業界基準是網頁在兩秒內載入，但最終使用者期望網頁在一秒內載入，則您在建立 KPI 時應將這兩個資料點列入考慮。

團隊必須使用即時精密資料和歷史資料作為參考，來評估工作負載 KPI，並建立儀表板，針對 KPI 資料執行指標數學，以衍生營運和使用率洞察。KPI 應該加以記錄，並包含支援業務目標和策略的 KPI 和閾值，並應映射至受監控的指標。當業務目標、策略或最終使用者需求變更時，應該重新檢視 KPI。

實作步驟

- 識別利益相關者：識別和記錄關鍵業務利益相關者，包括開發和運營團隊。
- 定義目標：與利益相關者合作，以定義和記錄工作負載的目標。考慮工作負載的關鍵效能層面，例如輸送量、回應時間和成本，以及業務目標，例如使用者滿意度。
- 審核業界最佳實務：審核業界最佳實務，以找出符合您工作負載目標的相關 KPI。
- 識別指標：找出符合您工作負載目標的指標，可協助您衡量效能和業務目標。建立以這些指標為基礎的 KPI。範例指標是諸如平均回應時間或並發使用者數量等測量值。
- 定義並記錄 KPI：使用業界最佳實務和工作負載目標，為工作負載 KPI 設立目標。使用此資訊，來設定嚴重性或警示層級的 KPI 閾值。找出並記錄未達到 KPI 的風險和影響。
- 實作監控：使用監控工具 (例如 [Amazon CloudWatch](#) 或 [AWS Config](#)) 來收集指標並測量 KPI。
- 以視覺化方式傳達 KPI：使用 [Amazon Quick](#) 等儀表板工具，視覺化 KPI 並與利益相關者溝通。
- 分析和最佳化：定期審核和分析 KPI，以找出需要改善的工作負載領域。與利益相關者合作以實作這些改進。
- 重新檢視和調整：定期審核指標和 KPI 以評估其有效性，尤其是在業務目標或工作負載效能變更時。

資源

相關文件：

- [CloudWatch 文件](#)
- [監控、日誌記錄和效能 AWS Partner](#)
- [AWS 可觀測性工具](#)
- [大規模雲端遷移關鍵績效指標 \(KPI\) 的重要性](#)
- [如何使用 KPI 儀表板追蹤成本最佳化 KPI](#)
- [X-Ray 文件](#)
- [使用 Amazon CloudWatch 儀表板](#)
- [Quick KPI](#)

相關影片：

- [AWS re:Invent 2023 - 優化成本和效能並追蹤緩解措施的進度](#)
- [AWS re:Invent 2023 - 使用 AWS Health 大規模管理資源生命週期事件](#)
- [AWS re:Invent 2023 - Pinterest 的效能與效率：最佳化最新的執行個體](#)
- [AWS re:Invent 2022 - AWS 優化：立即見效的可操作步驟](#)
- [AWS re:Invent 2023 - 建立有效的可觀測性策略](#)
- [AWS Summit SF 2022 - 使用 AWS 獲得全堆疊可觀測性和應用程式監控](#)
- [AWS re:Invent 2023 - 針對前 1000 萬個使用者在 AWS 上進行擴展](#)
- [AWS re:Invent 2022 - Amazon 如何使用更好的指標來提高網站效能](#)
- [為您的企業建立有效的指標策略 | AWS 活動](#)

相關範例：

- [使用 Quick 建立儀表板](#)

PERF05-BP02 使用監控解決方案了解效能最關鍵的領域

了解並找出提高工作負載效能將對效率或客戶體驗產生正面影響的地方。例如，具有大量客戶互動的網站可受益於邊緣服務的使用，因為這樣可以將內容交付移至更接近客戶的地方。

常見的反模式：

- 您假設標準運算指標 (例如 CPU 使用率或記憶體壓力) 足以找出效能問題。
- 您只會使用所選監控軟體記錄的預設指標。

- 您只會在有問題時審查指標。

建立此最佳實務的優勢：了解效能的關鍵領域，有助於工作負載擁有者監控 KPI 和優先處理具有高影響力的待改善之處。

未建立此最佳實務時的曝險等級：高

實作指引

設定端對端追蹤，以識別流量模式、延遲和關鍵效能領域。監控您的資料存取模式，以確定是否有緩慢的查詢或分段及分割不佳的資料。使用負載測試或監控來找出工作負載受限領域。

透過了解架構、流量模式和資料存取模式，來提高效能效率，並確定延遲和處理時間。找出隨著工作負載的成長，可能會影響客戶體驗的潛在瓶頸。調查這些領域後，請審視自己可以部署哪個解決方案，來消除這些效能疑慮。

實作步驟

- 設置端到端監控，來擷取所有工作負載組成部分和指標。以下是 AWS 上的監控解決方案範例。

服務	在哪裡使用
Amazon CloudWatch 實際使用者監控 (RUM)	擷取來自實際使用者用戶端和前端工作階段的應用程式效能指標。
AWS X-Ray	透過應用程式層追蹤流量，並找出組成部分和相依性之間的延遲。使用 X-Ray 服務地圖，查看工作負載組成部分之間的關係和延遲。
Amazon Relational Database Service 效能洞見	檢視資料庫效能指標並找出效能待改善之處。
Amazon RDS 增強型監控	檢視資料庫 OS 效能指標。
Amazon DevOps Guru	偵測異常作業模式，以便在營運問題影響客戶之前識別。

- 執行測試，來產生指標、確定流量模式、瓶頸和關鍵效能區域。以下是如何進行測試的一些範例：
 - 設定 [CloudWatch SyntheticCanary](#) 以程式設計方式使用 Linux Cron 任務或評分運算式，模擬以瀏覽器為基礎的使用者活動，以產生長期一致的指標。

- 使用 [AWS 分散式負載測試](#) 解決方案，來產生尖峰流量或以預期成長速率測試工作負載。
- 評估指標和遙測，來找出關鍵的效能領域。與您的團隊一起審核這些領域，討論監控和解決方案，以避免瓶頸。
- 進行效能改善的實驗，並透過資料來衡量這些變更。例如，可以使用 [CloudWatch Evidently](#) 測試對工作負載的新改進和效能影響。

資源

相關文件：

- [re:Invent 2023 中 AWS Observability 的最新消息](#)
- [Amazon 建置者資料中心](#)
- [X-Ray 文件](#)
- [Amazon CloudWatch RUM](#)
- [Amazon DevOps Guru](#)

相關影片：

- [AWS re:Invent 2023 – \[發佈\] 針對現代工作負載的應用程式監控](#)
- [AWS re:Invent 2023 – 實作應用程式可觀測性](#)
- [AWS re:Invent 2023 - 建立有效的可觀測性策略](#)
- [AWS Summit SF 2022 - 使用 AWS 獲得全堆疊可觀測性和應用程式監控](#)
- [AWS re:Invent 2022 - AWS 優化：立即見效的可操作步驟](#)
- [AWS re:Invent 2022 - Amazon 建置者資料中心：25 年 Amazon 卓越營運](#)
- [AWS re:Invent 2022 - Amazon 如何使用更好的指標來提高網站效能](#)
- [使用 Amazon CloudWatch Synthetics 進行應用程式的視覺化監控](#)

相關範例：

- [使用 Amazon CloudWatch Synthetics 測量頁面載入時間](#)
- [Amazon CloudWatch RUM Web 用戶端](#)
- [適用於 Python 的 X-Ray 開發套件](#)
- [AWS 上的分散式負載測試](#)

PERF05-BP03 定義提高工作負載效能的程序

定義一個程序，以在新的服務、設計模式、資源類型和組態可用時對其進行評估。例如，對新的執行個體方案執行現有的效能測試，以判斷其是否可能改善工作負載。

常見的反模式：

- 您假設目前的架構是靜態的，且不會隨著時間而更新。
- 您會隨時間導入架構變更，而且無須指標佐證。

建立此最佳實務的優勢：定義進行架構變更的程序後，您就能使用收集的資料，以隨著時間影響工作負載。

未建立此最佳實務時的曝險等級：中

實作指引

工作負載的效能有一些關鍵限制。記錄這些內容，以便您知道哪種創新可以改善工作負載的效能。當新服務或技術可用時，請使用此資訊來找出緩解限制或瓶頸的方法。

識別工作負載的關鍵效能限制。記錄工作負載的效能限制，讓您知道哪些類型的創新可能會改善工作負載的效能。

實作步驟

- 識別 KPI：識別 [PERF05-BP01 建立用於測量工作負載運作狀態和效能的關鍵績效指標 \(KPI\)](#) 中所述的工作負載效能 KPI，以設立工作負載基準。
- 實作監控：使用 [AWS 可觀測性工具](#) 收集績效指標並衡量 KPI。
- 執行分析：執行深入分析，以找出工作負載中效能不佳的區域 (例如組態和應用程式的程式碼)，步驟請參閱 [PERF05-BP02 使用監控解決方案了解效能最關鍵的領域](#)。使用分析和效能工具，來確定效能改進策略。
- 驗證改進：使用沙盒或生產前環境，來驗證改進策略的有效性。
- 實作變更：實作生產中的變更，並持續監控工作負載的效能。記錄改進項目並與利益相關者溝通這些變更。
- 重新檢視和完善：定期檢視您的績效改善程序，以找出需要提高的領域。

資源

相關文件：

- [AWS 部落格](#)
- [AWS 最新消息](#)
- [AWS Skill Builder](#)

相關影片：

- [AWS re:Invent 2022 - 提供可持續、高效能的架構](#)
- [AWS re:Invent 2023 - 優化成本和效能並追蹤緩解措施的進度](#)
- [AWS re:Invent 2022 - AWS 優化：立即見效的可操作步驟](#)
- [AWS re:Invent 2022 - 使用最佳實務指引來最佳化 AWS 工作負載](#)

相關範例：

- [AWS Github](#)

PERF05-BP04 Load 測試工作負載

對工作負載執行負載測試，以確認它可以處理生產負載並識別任何效能瓶頸。

常見的反模式：

- 可以對工作負載的個別部分進行負載測試，而非整個工作負載。
- 可以在與生產環境不同的基礎設施中進行負載測試。
- 您只對預期的 (而非超標) 負載進行負載測試，以協助預測未來可能發生問題的位置。
- 您可以在未諮詢 [Amazon 測試政策的情況下執行負載EC2測試](#)，並提交模擬事件提交表單。這會導致您的測試無法執行，因為它看起來像事件 denial-of-service。

建立此最佳實務的優勢：在負載測試過程中測量效能時，會顯示您將在負載增加到何種程度時受到影響。這可讓您能夠在工作負載受到影響之前預測所需的變更。

未建立此最佳實務時的曝險等級：低

實作指引

雲端中的負載測試是在實際條件下，以預期的使用者負載來衡量雲端工作負載效能的程序。此程序包括佈建類似生產環境的雲端環境、使用負載測試工具產生負載，以及分析指標以評估工作負載處理實際負載的能力。必須使用生產資料的綜合或處理過的版本 (移除敏感或可識別身分的資訊) 執行負載測試。

自動執行負載測試作為交付管道的一部分，並將結果與預先定義的KPIs閾值進行比較。此程序有助於您持續達到所需的效能。

實作步驟

- 定義測試目標：確定您要評估的工作負載效能層面，例如輸送量和回應時間。
- 選擇測試工具：選擇並設定適合您工作負載的負載測試工具。
- 設定您的環境：根據生產環境設定測試環境。您可以使用 AWS 服務來執行生產規模環境，以測試您的架構。
- 實作監控：使用 [Amazon CloudWatch](#) 等監控工具，收集架構中各項資源的指標。也可以收集和發布自訂指標。
- 定義方案：定義負載測試方案和參數 (如測試持續時間和使用者數量)。
- 進行負載測試：大規模執行測試方案。利用 AWS 雲端 來測試工作負載，以探索其無法擴展的位置，或它是否以非線性方式擴展。例如，使用 Spot 執行個體以低成本產生負載，並在生產中遇到瓶頸之前發現瓶頸。
- 分析測試結果：分析結果以找出效能瓶頸和需要改善的區域。
- 記錄和分享調查結果：記錄並報告調查結果和建議。與利益相關者分享此資訊，協助他們做出有關效能最佳化策略的明智決策。
- 不斷反覆執行：負載測試應定期執行，尤其是在系統更新變更之後。

資源

相關文件：

- [Amazon CloudWatch RUM](#)
- [Amazon CloudWatch Synthetics](#)
- [上的分散式負載測試 AWS](#)

相關影片：

- [AWS Summit ANZ 2023：透過 AWS 分散式負載測試，放心加速](#)
- [AWS re：Invent 2022 - AWS 為前 1,000 萬使用者擴展](#)
- [使用 AWS 解決方案解決：分散式負載測試](#)
- [AWS re：Invent 2021 - 透過使用 Amazon 的終端使用者洞察最佳化應用程式 CloudWatch RUM](#)

- [Amazon CloudWatch Synthetics 示範](#)

相關範例：

- [上的分散式負載測試 AWS](#)

PERF05-BP05 使用自動化主動修復效能相關問題

使用關鍵績效指標 (KPI) 搭配監控和提醒系統，主動處理效能相關的問題。

常見的反模式：

- 您只讓操作人員有能力對工作負載進行操作變更。
- 您讓所有警示篩選到操作團隊，無須主動修復。

建立此最佳實務的優勢：主動修復警示動作能夠讓支援人員專注在無法自動採取行動的項目上。這有助於操作人員無須疲於處理所有警示，而僅專注於關鍵警示。

未建立此最佳實務時的曝險等級：低

實作指引

使用警示觸發自動化動作，盡可能修復問題。如果無法自動回應，則將警示上報給能夠回應的人員。例如，您可能有一個可以預測關鍵績效指標 (KPI) 預期值並在超過特定閾值時發出警示的系統，或者在 KPI 超出預期值時可以自動停止或回復部署的工具。

實作可在工作負載執行時提供效能可見度的程序。建置監控儀表板並建立效能預期的基準規範，以確定工作負載是否以最佳狀態執行。

實作步驟

- 識別修復工作流程：識別並了解可自動修復的效能問題。使用 AWS 監控解決方案 (例如 [Amazon CloudWatch](#) 或 AWS X-Ray)，協助您更好地了解問題的根本原因。
- 定義自動化程序：制定可用來自動修正問題的逐步修復程序。
- 設定啟動事件：將事件設定為自動啟動修復程序。例如，您可以定義觸發程式，在執行個體達到特定 CPU 使用率閾值時自動重新啟動執行個體。
- 自動化修復：使用 AWS 服務和技術，自動化修復程序。例如，[AWS Systems Manager Automation](#) 提供安全且可擴展的方式，來自動化修復程序。如果變更無法成功解決問題，則請務必使用自我修復邏輯來還原變更。

- 測試工作流程：在生產前環境中測試自動修復程序。
- 實作工作流程：在生產環境中實作自動修復。
- 制定說明手冊：制定並記錄說明手冊，其中概述了補救計劃的步驟，包括啟動事件、補救邏輯和採取的動作。確保培訓利益相關者，以協助他們有效地應對自動補救事件。
- 審查和完善：定期評估自動補救工作流程的有效性。如有必要，請調整啟動事件和補救邏輯。

資源

相關文件：

- [CloudWatch 文件](#)
- [監控、日誌記錄和效能 AWS Partner Network 合作夥伴](#)
- [X-Ray 文件](#)
- [使用 CloudWatch 中的警示和警示動作](#)
- [建立雲端自動化實務以實現卓越營運：AWS Managed Services 的最佳實務](#)
- [使用自動表格優化來自動調整您的 Amazon Redshift 效能](#)

相關影片：

- [AWS re:Invent 2023 - 自動擴展、補救和智慧自我修復的策略](#)
- [AWS re:Invent 2023 – \[發佈\] 針對現代工作負載的應用程式監控](#)
- [AWS re:Invent 2023 – 實作應用程式可觀測性](#)
- [AWS re:Invent 2021 - 智慧自動化雲端作業](#)
- [AWS re:Invent 2022 - 在 AWS 環境中大規模設定控制項](#)
- [AWS re:Invent 2022 - 使用 AWS 自動化修補程式管理和合規性](#)
- [AWS re:Invent 2022 - Amazon 如何使用更好的指標來提高網站效能](#)
- [AWS re:Invent 2023 - 減輕負擔：使用 Amazon RDS 診斷和解決效能問題](#)
- [AWS re:Invent 2021 - {最新發佈} 使用 Amazon DevOps Guru 自動偵測並解決問題](#)
- [AWS re:Invent 2023 - 將操作集中化](#)

相關範例：

- [CloudWatch Logs 自訂警示](#)

PERF05-BP06 保留工作負載和服務 up-to-date

繼續使用 up-to-date 新的雲端服務和功能，以採用有效率的功能、移除問題，並改善工作負載的整體效能效率。

常見的反模式：

- 假設您目前的架構為靜態，且不會隨著時間的推移而更新。
- 您沒有任何系統或定期規律可評估更新的軟體與套件是否與您的工作負載相容。

建立此最佳實務的優點：透過建立程序以持續 up-to-date 使用新服務和產品，您可以採用新功能和功能、解決問題並改善工作負載效能。

未建立此最佳實務時的曝險等級：低

實作指引

在新服務、設計模式和產品功能推出時，評估提升效能的方法。透過評估、內部討論或外部分析，確定哪些方法可以提高工作負載效能或效率。定義程序來評估與工作負載相關的更新、新功能和服務。例如，建立使用新技術的概念證明或與內部小組協商。嘗試新的想法或服務時，執行效能測試以衡量其對工作負載效能的影響。

實作步驟

- 清查工作負載：清查工作負載軟體和架構，並識別需要更新的元件。
- 識別更新來源：找出與工作負載組成部分相關的新聞和更新來源。例如，您可以訂閱符合您工作負載元件的產品的 [AWS 部落格最新消息](#)。您可以訂閱 RSS 摘要或管理您的 [電子郵件訂閱](#)。
- 定義更新排程：定義排程以評估工作負載的新服務和特徵。
 - 您可以使用 [AWS Systems Manager Inventory](#) 從您的 Amazon EC2 執行個體收集作業系統 (OS)、應用程式和執行個體中繼資料，並快速了解哪些執行個體正在執行軟體政策所需的軟體和組態，以及哪些執行個體需要更新。
- 評估最新更新：了解如何更新工作負載的元件。利用雲端的靈活性快速測試新特徵對工作負載有何改善，藉以提高效能效率。
- 使用自動化：使用更新程序自動化，以減少部署新功能的工作量，並避免手動程序引起的錯誤。
 - 您可以使用 [CI/CD](#) 自動更新 AMIs、容器映像，以及與雲端應用程式相關的其他成品。
 - 可以使用 [AWS Systems Manager Patch Manager](#) 之類的工具來自動化系統更新流程，並使用 [AWS Systems Manager Maintenance Windows](#) 來排程活動。

- 記錄過程：記錄用於評估更新和新服務的過程。向擁有者提供所需的時間和空間，來研究、測試、試驗和驗證更新及新服務。參考文件化的業務需求KPIs，並協助排定哪些更新將對業務產生正面影響的優先順序。

資源

相關文件：

- [AWS 部落格](#)
- [新功能 AWS](#)
- [使用自動化映像建置器管道實作 up-to-date EC2 映像](#)

相關影片：

- [AWS re : Inforce 2022 - 使用 自動化修補程式管理和合規 AWS](#)
- [所有事項修補程式：AWS Systems Manager | AWS Events](#)

相關範例：

- [庫存和修補程式管理](#)
- [一個可觀測性研討會](#)

PERF05-BP07 定期審查指標

作為日常維護的一部分或對事件或事故的回應，審查收集了哪些指標。透過這些審查來識別哪些指標是解決問題的關鍵，以及哪些其他指標 (如果被追蹤) 有助於識別、解決或預防問題。

常見的反模式：

- 您讓指標長時間持續處於警示狀態。
- 您建立自動化系統無法採取行動的警示。

建立此最佳實務的優勢：持續審查正在收集的指標，以確認指標正確識別、處理或防止問題發生。如果讓指標長時間持續處於警示狀態，指標也會變得過時。

未建立此最佳實務時的曝險等級：中

實作指引

不斷改進指標收集和監控。作為對事故或事件的回應的一部分，評估哪些指標有助於解決問題，哪些指標可以幫助解決問題但未被追蹤。使用此方法提高所收集指標的品質，從而可以防止事故發生或更快地解決將來的事務。

作為對事故或事件的回應的一部分，評估哪些指標有助於解決問題，哪些指標可以幫助解決問題但未被追蹤。使用此方法提高所收集指標的品質，從而可以防止事故發生或更快地解決將來的事務。

實作步驟

- 定義指標：定義與您的工作負載目標一致的關鍵效能指標以進行監控，包括回應時間和資源使用率等指標。
- 建立基準：設定各指標的基準和期望值。基準應提供參考點以識別偏差或異常。
- 設定規律：設定規律 (例如每週或每月一次) 以審核重要指標。
- 識別效能問題：每次審查期間都會評估趨勢，以及與基準值的偏差。查看是否有任何效能瓶頸或異常情況。對於已確認的問題，請展開深入根本原因分析，以了解問題背後的主要原因。
- 識別修正動作：使用您的分析來識別修正動作。這可能包括參數調整、修正錯誤和擴展資源。
- 記錄調查結果：記錄您的調查結果，包括已識別的問題、根本原因和修正動作。
- 反覆執行並改善：持續評估並改善指標審核過程。使用從以前的審核中學到的經驗教訓，隨著時間的推移提升程序。

資源

相關文件：

- [CloudWatch 文件](#)
- [使用 CloudWatch 代理程式從 Amazon EC2 執行個體和內部部署伺服器收集指標和日誌](#)
- [使用 CloudWatch Metrics Insights 查詢您的指標](#)
- [監控、日誌記錄和效能 AWS Partner Network 合作夥伴](#)
- [X-Ray 文件](#)

相關影片：

- [AWS re:Invent 2022 - 在 AWS 環境中大規模設定控制項](#)

- [AWS re:Invent 2022 - Amazon 如何使用更好的指標來提高網站效能](#)
- [AWS re:Invent 2023 - 建立有效的可觀測性策略](#)
- [AWS Summit SF 2022 - 使用 AWS 獲得全堆疊可觀測性和應用程式監控](#)
- [AWS re:Invent 2023 - 減輕負擔：使用 Amazon RDS 診斷和解決效能問題](#)

相關範例：

- [使用 Quick 建立儀表板](#)
- [CloudWatch 儀表板](#)

成本最佳化

成本最佳化支柱包括以最低價格執行系統來產生商業價值的能力。您可以在[成本最佳化支柱白皮書](#)中找到有關實作的規定性指引。

最佳實務領域

- [實作雲端財務管理](#)
- [了解支出和用量](#)
- [具有經濟效益的資源](#)
- [管理需求與供應資源](#)
- [隨時間優化](#)

實作雲端財務管理

問題

- [COST 1. 如何實作雲端財務管理？](#)

COST 1. 如何實作雲端財務管理？

透過實作雲端財務管理，就可幫助組織最佳化成本和用量，並且在 AWS 上進行規模調整，進而實現商業價值和財務上的成功。

最佳實務

- [COST01-BP01 建立成本最佳化的擁有權](#)

- [COST01-BP02 在財務與技術之間建立合作夥伴關係](#)
- [COST01-BP03 建立雲端預算和預測](#)
- [COST01-BP04 在組織程序中實作成本感知](#)
- [COST01-BP05 報告並通知成本最佳化](#)
- [COST01-BP06 主動監控成本](#)
- [COST01-BP07 up-to-date保留新的服務版本](#)
- [COST01-BP08 建立成本感知文化](#)
- [COST01-BP09 從成本最佳化量化商業價值](#)

COST01-BP01 建立成本最佳化的擁有權

建立團隊（Cloud Business Office、Cloud Center of Excellence 或 FinOps 團隊），負責建立和維護整個組織的成本意識。成本最佳化的負責人可以是個人或是團隊，條件是必須是來自財務、技術或業務團隊，且了解整個組織和雲端財務的人員。

未建立此最佳實務時的曝險等級：高

實作指引

這是 Cloud Business Office（CBO）或 Cloud Center of Excellence（CCOE）函數或團隊的推出，負責建立和維護雲端運算的成本意識文化。這個職能可以是現有個人、組織內的團隊，或是由組織內主要財務、技術和組織利益相關者組成的新團隊。

此職能部門（個人或團隊）會優先並花費一定比例的時間，進行成本管理和成本最佳化活動。相較於大型企業的全職職能部門，小型組織的此職能部門花費的時間比例可能較少。

此職能部門必須採行跨領域合作的方法，並要具備專案管理、資料科學、財務分析和軟體或基礎架構開發等能力。藉此，在三種不同的所有權下執行成本最佳化，以改善工作負載效率：

- 集中：透過 FinOps 團隊、雲端財務管理（CFM）團隊、雲端商業辦公室（CBO）或 Cloud Center of Excellence（CCoE）等指定團隊，客戶可以設計和實作治理機制，並推動全公司最佳實務。
- 分散式：影響技術團隊，進行成本最佳化。
- 混合式：結合集中式與去中心化方法，讓團隊互相合作，進行成本最佳化。

可以根據成本最佳化目標（例如工作負載效率指標）來衡量此職能部門的執行和交付能力。

您必須設法讓高層支持此職能部門，這是成功的關鍵因素。高層支持者會成為運用雲端服務節省成本的推動者，並替團隊提報支援，確保成本最佳化活動獲組織認定為優先要務。否則，相關的方針可能不會受到重視，且節省成本將不會被列為優先要務。高層支持者和這個團隊共同協助您的組織，讓其得以聰明高效地使用雲端，並提供商業價值。

如果您有 業務或企業[支援計劃](#)，Enterprise-On-Ramp且需要協助建置此團隊或函數，請透過您的 帳戶團隊聯絡您的 Cloud Financial Management (CFM) 專家。

實作步驟

- 定義關鍵成員：貴組織的所有相關人員都必須貢獻己力，進一步了解成本管理。組織內的常見團隊通常包括：財務、應用程式或產品擁有者、管理和技術團隊 (DevOps)。有些團隊必須全職參與 (財務或技術)，有些團隊則可視需要定期參與。執行的個人或團隊CFM需要以下一組技能：
 - 軟體開發：如果正在建構指令碼和自動化。
 - 基礎架構工程：用以部署指令碼、自動化程序，並理解服務或資源的佈建方式。
 - 操作敏銳度：CFM透過測量、監控、修改、規劃和擴展雲端的有效使用，來有效率地在雲端上操作。
- 定義目標和指標：該職能部門需要以不同的方式提供價值給組織。定義的目標會隨著組織的發展而不斷演變。常見的活動包括：建立和執行整個組織成本最佳化的教育計畫，制定整個組織的標準 (例如成本最佳化的監控和報告)，以及設定工作負載最佳化目標。此職能部門也需要定期向組織報告其成本最佳化的能力。

您可以定義值型或成本型金鑰效能指標 (KPIs)。當您定義 時KPIs，您可以根據效率和預期業務結果來計算預期成本。以價值為基礎的將成本和用量指標與商業價值驅動因素KPIs聯繫起來，並協助合理化 AWS 支出的變化。衍生以價值為基礎的第一步KPIs是跨組織合作，以選取並商定一組標準 KPIs。

- 確立定期規律：各群組 (財務、技術和業務團隊) 應定期會談，並審查其目標和指標。一般的規律包括審查組織的狀態、審查目前執行的任何計畫、整體財務和最佳化指標。然後，再更詳細地報告關鍵工作負載。

在這類定期會談中，您可以審查工作負載效率 (成本) 和商業成果。例如，工作負載成本上升 20% 與增加的客戶用量，是相對應的。在此案例中，這 20% 的成本上升可被視為投資。這些定期節奏呼叫可協助團隊識別為整個組織KPIs提供意義的值。

資源

相關文件：

- [AWS CCOE 部落格](#)
- [建立雲端商業辦公室](#)
- [CCOE - Cloud Center of Excellence](#)

相關影片：

- [Vanguard CCOE 成功案例](#)

相關範例：

- [使用 Cloud Center of Excellence \(CCOE \) 轉換整個企業](#)
- [建置 CCOE 以轉換整個企業](#)
- [建置時要避免的 7 個陷阱 CCOE](#)

COST01-BP02 在財務與技術之間建立合作夥伴關係

讓財務和技術團隊參與討論雲端之旅各個階段的成本和用量。各團隊定期碰面並討論相關主題，例如，組織總目標和具體目標、成本和用量的目前狀態，以及財務和會計實務。

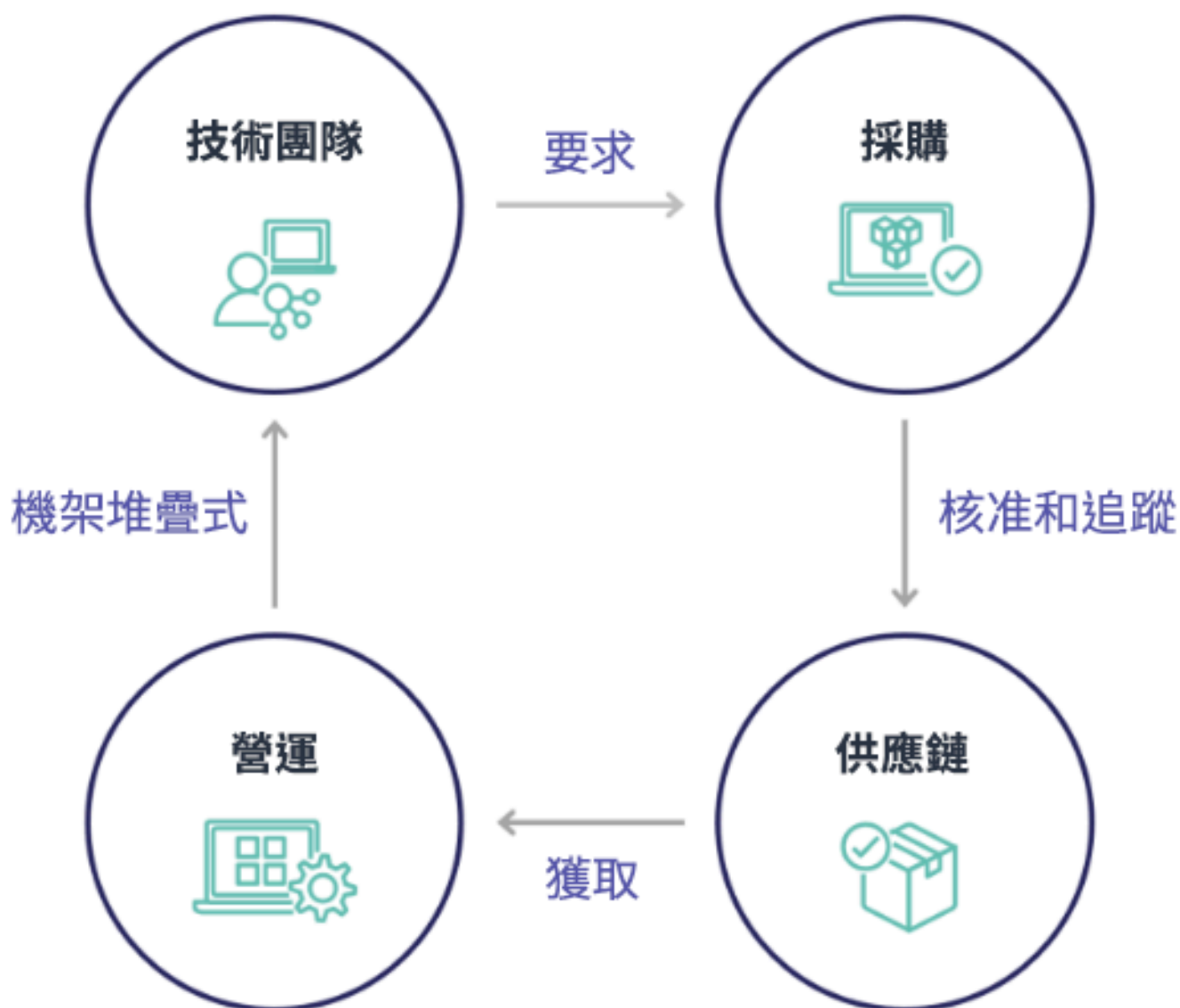
未建立此最佳實務時的曝險等級：高

實作指引

由於核准、採購和基礎設施部署週期縮短，技術團隊可在雲端提高創新速度。對於之前習慣於執行耗時且資源密集型程序，以便採購資料中心和內部部署環境，並且只在核准專案時才分配成本的財務組織來說，這是一項調整。

就金融與採購組織的觀點而言，資本預算、資金要求、核准、採購和安裝實體基礎架構的流程，在過去數十年來早已廣為人知並標準化：

- 工程或 IT 團隊通常是要求者
- 核准者和採購者由不同的財務團隊擔任
- 營運團隊負責建構、堆疊及交付現成可用的基礎架構



採用雲端後，基礎架構的採購和取用不再受制於一連串的相依性。在雲端模型中，技術及產品小組不再只是建置者，而是產品的操作人員和擁有者，負責處理在過去與財務與營運團隊相關聯的多數活動，包括採購和部署。

要佈建雲端資源，所需的其實就是一個帳戶以及一組適當的權限。IT 和財務風險也因而降低；這意味著，團隊只需按幾下滑鼠或執行 API 呼叫，即可終止閒置或非必要的雲端資源。這也讓技術團隊得以加速創新 – 基於建立和推翻試驗的靈活性與能力。儘管使用雲端的本質是多變的，就資本預算和預測的角度而言可能會影響到可預測性，但雲端仍讓組織得以降低過度佈建的成本，以及降低因保守的佈建不足而伴隨的機會成本。



在關鍵財務和技術利益相關者之間建立合作夥伴關係，以形成對組織目標的共識，並建立可在雲端運算可變支出模型中取得財務成功的機制。組織內的相關團隊必須參與雲端之旅各個階段的成本和用量討論，包括：

- 財務主管：財務長、財務總監、財務規劃師、商業分析師、採購和應付帳款必須了解雲端消費模式、購買選項和每月發票開立程序。財務部門需要與技術團隊合作，來建立 IT 價值故事並加以傳播，以協助業務團隊了解技術支出與業務成果之間的連結。這樣，技術支出就不再被視為成本，而是投資。由於雲端與內部部署營運存在基本差異（例如，用量改變速率、依用量計費定價、分級定價、定價模式以及詳細帳單和用量資訊），財務組織必須了解雲端用量如何影響商業層面，包括採購程序、激勵追蹤、成本分配和財務報表。
- 技術主管：技術主管（包括產品和應用程式擁有者）必須了解財務需求（例如，預算限制）以及業務需求（例如，服務水準協議）。如此可允許實作工作負載，達成組織希望的目標。

財務與科技的合作夥伴關係可帶來下列好處：

- 財務和技術團隊可近乎即時地檢視成本和用量。
- 財務和技術團隊建立標準操作程序來處理雲端支出變化。
- 在資本如何用於購買承諾折扣 (例如，預留執行個體或 AWS Savings Plans)，以及如何使用雲端來發展組織方面，財務利益相關者會擔任策略顧問。
- 現有的應付帳款和採購程序會與雲端搭配使用。
- 財務和技術團隊共同預測未來的 AWS 成本和用量，以評估並擬定組織預算。
- 透過共同的語言以及對財務概念的一致理解，促進跨組織溝通。

組織內應參與成本和用量討論的其他利益相關者包括：

- 業務單位擁有者：業務單位擁有者必須了解雲端業務模式，以便對業務單位和全公司提供指引。當有需要預測成長和工作負載用量，以及需要評估長期購買選項，例如預留執行個體或 Savings Plans 時，此項雲端知識相當重要。
- 工程團隊：在財務與技術團隊之間建立合作夥伴關係至關重要，這是培養成本感知文化，鼓勵工程師對雲端財務管理 (CFM) 採取行動，所不可或缺的。CFM 或財務營運從業人員與財務團隊的常見問題之一，是不易讓工程師了解雲端業務的全貌、遵循最佳實務，以及執行建議的動作。
- 第三方：如果您的組織使用第三方 (例如，顧問或工具)，請確保他們符合您的財務目標，並能透過其參與模式和投資報酬率 (ROI) 證實符合。通常第三方會報告和分析其管理的一切工作負載，並且提供所設計一切工作負載的成本分析。

要實作 CFM 並取得成功，需要財務、技術和業務團隊之間進行協作，並且需要轉變整個組織傳達和評估雲端支出的方式。請納入工程團隊，使他們在各階段都能加入這些成本與用量的討論中，並鼓勵他們遵循最佳實務，並據以執行已達成共識的動作。

實作步驟

- 定義關鍵成員：確認您的財務和技術團隊中的所有相關成員都參與此合作夥伴關係。相關財務成員會處理雲端帳單。涉及人員通常包括財務總監、財務控制者、財務規劃師、商業分析師、採購和採購專員。技術成員通常是產品與應用程式擁有者、技術經理以及在雲端進行建置的所有團隊的代表。其他成員可能包括業務單位擁有者，例如，顧問等會影響產品用量的行銷單位，以及實現與目標和機制保持一致並協助報告的第三方人員。
- 定義討論主題：確定團隊中常見的主題，或需要有共識的主題。從建立時開始追蹤成本，直到帳單支付為止。請記下所有參與的成員，以及需要應用的組織程序。了解採用的每個步驟或程序及相關資訊，例如可用的定價模式、分級定價、折扣模式、預算編列和財務要求。

- 建立定期規律：若要建立財務與技術的合作夥伴關係，請建立定期通訊規律，以樹立並維持一致性。該群組需要針對他們的目標和指標定期聚會進行討論。一般的規律包括審查組織的狀態、審查目前執行的任何計畫、整體財務和最佳化指標。然後，會更詳細地報告關鍵工作負載。

資源

相關文件：

- [AWS 新聞部落格](#)

COST01-BP03 建立雲端預算和預測

調整現有的組織預算編列和預測程序，使其與本質會高度變動的雲端成本和用量相容。程序必須是動態的，並使用以趨勢為基礎和/或以業務驅動因素為基礎的演算法。

未建立此最佳實務時的曝險等級：高

實作指引

在傳統的內部部署 IT 設定中，客戶通常會面臨規劃固定成本的挑戰，這些成本只是偶爾變化，通常是購買新的 IT 硬體和服務以滿足尖峰需求。相比之下，AWS 雲端採用不同的方法，客戶根據其實際 IT 和業務需求，支付他們使用的資源費用。在雲端環境中，需求可能會每月、每天甚至每小時波動。

使用雲端可帶來了效率、速度和敏捷性，進而產生高度變化的成本和使用模式。成本會隨著工作負載效率的增加或部署新的工作負載和功能而降低或增加。隨著工作負載擴展以滿足不斷擴大的客戶群，由於資源的可存取性增加，雲端用量和成本也隨之增加。雲端服務的這種靈活性延伸到成本和預測，這創造了一定程度的彈性。

與這些不斷變化的業務需求和需求驅動因素緊密保持一致至關重要，並盡可能準確地規劃。傳統的組織預算流程需要適應這種變化。

在預測新工作負載的成本時，請考慮成本建模。成本建模可建立對預期雲端成本的基準理解，協助您執行總體擁有成本 (TCO)、投資報酬率 (ROI) 和其他財務分析，與利害關係人設立目標和期望，並找出成本最佳化的機會。

您的組織應了解成本定義和接受的分組。您預測的詳細程度會根據組織的結構與內部工作流程而有所不同。選取符合您特定需求和組織設定的精細度。了解在什麼層級執行預測非常重要：

- 管理帳戶或 AWS Organizations 層級：管理帳戶是您用於建立 AWS Organizations 的帳戶。組織預設有一個管理帳戶。

- **連結帳戶或成員帳戶：**組織中的帳戶是標準 AWS 帳戶，其中包含您的 AWS 資源，以及可存取這些資源的身分。
- **環境：**環境是執行某個應用程式版本的 AWS 資源的集合。可以使用多個連結帳戶或成員帳戶建立一個環境。
- **專案：**專案是指固定期間內要完成的既定目標或任務的組合。在預測期間考慮專案生命週期非常重要。
- **AWS 服務：**群組或類別 (例如運算或儲存服務)，您可以在其中針對您的預測將 AWS 服務分組。
- **自訂分組：**您可以根據組織的需求建立自訂群組，例如業務單位、成本中心、團隊、成本分配標籤、成本類別、連結帳戶或這些項目的組合。

識別出會影響用量成本的業務驅動因素，並分別預測每個因素，以預先計算預期用量。部分驅動因素可能與組織內的 IT 和產品團隊相關。您的銷售、行銷和業務主管已經熟知行銷活動、促銷、地理擴張、合併與收購等其他業務驅動因素，進行協作並考慮所有這些需求驅動因素也很重要。

可使用 [AWS Cost Explorer](#) 根據您過去的支出預測已定義的未來時間範圍內的趨勢成本。AWS Cost Explorer 的預測引擎會根據收費類型 (例如，預留執行個體) 對您的歷史資料進行細分，並結合使用機器學習和基於規則的模型，來分別預測所有收費類型的支出。

建立預測流程並建置模型後，可使用 [AWS Budgets](#) 更精細地設定自訂預算，方法是指定時段、重複週期或金額 (固定或可變)，並新增篩選條件 (例如服務、AWS 區域 和標籤)。預算通常以一年為期，且保持固定不變，所有參與者必須嚴格遵守預算計畫。相較之下，預測更加靈活，也可以全年隨時調整，並提供一年、兩年或三年的動態預測。在技術和商業利益相關者之間建立財務期望時，預算和預測至關重要。準確的預測和實作，不僅讓直接負責佈建成本的利益相關者更能掌握狀況，還能夠提高整體成本感知。

若要及時了解現有預算的執行情況，您可以建立和排程 AWS Budgets 報告，以定期向您和利益相關者傳送電子郵件。您還可以根據實際成本 (為主動式) 或預測成本建立 AWS Budgets 提醒，從而為採取措施緩解潛在成本超支提供了時間。當您的成本或用量實際超出某個級別，或預計超出預算額度時，系統會提醒您。

使用基於趨勢的演算法 (使用歷史成本作為輸入) 和基於驅動因素的演算法 (例如：新產品推出、區域擴展或工作負載的新環境) 調整現有預算並預測流程，使其更具動態性，這是動態和可變支出環境的理想選擇。使用 Cost Explorer 或任何其他工具確定基於趨勢的預測後，請使用 [AWS 定價計算工具](#) 根據預期的用量 (流量、每秒請求數或必要的 Amazon EC2 執行個體) 來估算您的 AWS 使用案例和未來的成本。

追蹤預測的準確度，因為可以根據這些預測計算和估計來設定預算。監控整合式雲端成本預測的準確性和有效性。定期檢查與預測相比的實際支出，並根據需要進行調整以提高預測精確度。追蹤預測差異，並對報告的差異執行根本原因分析，以採取行動並調整預測。

如 [COST01-BP02 在財務與技術之間建立合作夥伴關係](#) 中所述，重要的是在 IT、財務和其他利益相關者之間建立合作關係和規律，才能確認所有人以一致的方式使用相同的工具或程序。如果預算可能需要改變，提高接觸頻率可提升對這些變化的因應速度。

實作步驟

- 定義組織內的成本語言：在組織中建立具有多個維度和分組的通用 AWS 成本語言。確保利益相關者了解預測精細度、定價模型和成本預測的水平。
- 分析基於趨勢的預測：使用基於趨勢的預測工具，例如 AWS Cost Explorer 和 Amazon Forecast。從服務、帳戶、標籤和成本類別等多個角度分析您的使用成本。
- 分析基於驅動因素的預測：識別出業務驅動因素對雲端使用情況的影響，並分別預測每個因素，以預先計算預期使用成本。與業務單位主管和利益相關者密切合作，了解對新驅動因素的影響，並計算預期成本變動，以準確編列預算。
- 更新現有預測與預算流程：根據所採用的預測方法 (例如基於趨勢、基於業務驅動因素、或兩種預測方法的組合)，定義您的預測和預算流程。預算應經過計算、切合實際並基於您的預測。
- 設定提醒和通知：使用 AWS Budgets 提醒和成本異常偵測來獲取提醒和通知。
- 與利益相關者一起執行定期審查：例如，與 IT、財務、平台團隊和其他業務部門的利益相關者一起商討如何因應經營方向與用量的變化。

資源

相關文件：

- [AWS Cost Explorer](#)
- [AWS Cost and Usage Report](#)
- [使用 Cost Explorer 進行預測](#)
- [Quick 預測](#)
- [AWS Budgets](#)

相關影片：

- [如何使用 AWS Budgets 追蹤我的支出和用量](#)

- [AWS 成本最佳化系列：AWS Budgets](#)

相關範例：

- [了解並建置基於驅動程式的預測](#)
- [如何建立和推動預測文化](#)
- [如何改善雲端成本預測](#)
- [使用適當的工具進行雲端成本預測](#)

COST01-BP04 在組織程序中實作成本感知

在會影響用量的全新或現有程序中實作成本感知、建立成本的透明度與權責劃分，並利用現有程序落實成本感知。在員工培訓中實作成本感知。

未建立此最佳實務時的曝險等級：高

實作指引

必須在新的和現有的組織程序中實作成本感知。對於其他最佳實務而言，這是基本的必備能力之一。建議盡可能重複使用和修改現有程序，這樣可將對靈活性和速度的影響降到最低。向技術團隊以及業務與財務團隊的決策者報告雲端成本，不僅可增強成本感知，也可為財務與業務利益相關者建立效率的關鍵績效指標 (KPI)。下列建議有助於在您的工作負載中實作成本感知：

- 確認變更管理包含成本測量，以量化變更所帶來的財務影響。這有助於主動解決成本相關疑慮，並提供成本節省資訊。
- 確認成本優化是您營運能力的核心部分。例如，您可以利用現有的事件管理程序，調查並找出成本和用量異常或成本超支的根本原因。
- 透過自動化或工具加速節省成本和實現商業價值。考慮實作的成本時，將投資報酬率 (ROI) 部分納入對話中，以證明投入時間或金錢的合理性。
- 藉由實作雲端支出的回報 (showback) 或計費 (chargeback) 來分配雲端成本 (包括以承諾為基礎的購買選項、共用服務和市場購買的支出)，以實現最具成本感知力的雲端使用。
- 擴展現有的培訓和發展計畫，納入整個組織的成本感知培訓。建議包含持續培訓和認證。這將建立一個能夠自我管理成本和用量的組織。
- 充分利用免費的 AWS 原生工具，例如 [AWS Cost Anomaly Detection](#)、[AWS Budgets](#) 和 [AWS Budgets Reports](#)。

當組織持續採用 [雲端財務管理](#) (CFM) 做法時，這些行為會在工作和決策的方式中變得根深蒂固。結果會產生更注重成本的文化，無論是開發人員設計新的雲端原生應用程式，還是財務經理分析這些新的雲端投資的投資報酬率，皆注重成本。

實作步驟

- 識別相關的組織程序：每個組織單位均審查其程序，並識別影響成本和用量的程序。任何導致資源建立或終止的程序都需要納入審查。尋找能夠在業務上支援成本感知的程序，例如事件管理和培訓。
- 建立自主的成本感知文化：確定所有相關的利益相關者都認同成本的改變原因和影響，因此都了解雲端成本。這將可讓您的組織針對創新建立自主的成本感知文化。
- 以成本感知更新程序：每個程序都會經過修改，以提高成本感知。程序可能需要額外的預先檢查，例如評估成本的影響，或進行後置檢查以驗證成本和用量預期的變更是否發生。可以擴展培訓和事件管理等支援程序，以包含成本和用量的項目。

如需協助，請透過客戶團隊洽詢 CFM 專家，或瀏覽下方的資源和相關文件。

資源

相關文件：

- [AWS 雲端財務管理](#)

相關範例：

- [高效雲端成本管理的策略](#)
- [成本控制部落格系列 3：如何處理成本衝擊](#)
- [AWS Cost Management 入門指南](#)

COST01-BP05 報告並通知成本最佳化

設定雲端預算及相關機制，偵測使用期間的異常情況。針對預先定義的目標設定成本和用量警示的相關工具，並於用量超過目標時接收通知。舉辦定期會議，分析工作負載的成本效益並提升成本感知。

未建立此最佳實務時的曝險等級：低

實作指引

您必須定期在組織內報告成本和用量最佳化。您可以舉辦專門的會議討論成本效益，或在工作負載的定期營運報告週期中包含優化成本的內容。使用服務和工具定期監控您的成本效益，並實施能夠節省成本的措施。

使用 [AWS Cost Explorer](#)，透過多種篩選條件和精細度來檢視成本和用量時，這會提供儀表板和報告，例如依服務或帳戶分類的成本、每日成本或市場成本。使用 [AWS Budgets Reports](#)，根據設定的預算追蹤成本使用和用量狀況時可使用。

使用 [AWS Budgets](#) 設定自訂預算以追蹤成本和用量，並在超出閾值時快速回應從電子郵件或 Amazon Simple Notification Service (Amazon SNS) 通知收到的提醒。[將您偏好的預算](#) 期間設定為每日、每月、每季或每年，並建立特定的預算限制，以隨時了解實際或預測的成本和用量是否接近預算閾值。也可根據這些提醒來設定[提醒](#)和[動作](#)以自動執行，或在超出預算目標時透過核准程序執行。

實作成本和用量通知，以確保能夠快速處理非預期的成本和用量變化。[AWS Cost Anomaly Detection](#) 可避免成本超出預期，並強化控制且不影響創新速度。AWS Cost Anomaly Detection 可識別異常支出與根本原因，有助於降低帳單超出預期的風險。只需簡單的三個步驟，您即可建立自己的情境化監視器，並且在偵測到任何異常支出時收到提醒。

您也可以使用 [Quick](#) 及 AWS Cost and Usage Report (CUR) 資料來提供包含更精細資料的高度自訂報告。Quick 可讓您排程報告，並定期接收成本報告電子郵件，了解歷史成本和用量或節省成本的機會。查看在 Quick 上建立的[成本智慧儀表板](#) (CID) 解決方案，可讓您更詳細地了解。

使用 [AWS Trusted Advisor](#) 作為指引，以確認已佈建的資源是否符合 AWS 的成本優化最佳實務。

根據您的成本細項和用量，透過視覺化圖表查看您的 Savings Plans 建議。每小時呈現的圖表顯示隨需支出以及建議的 Savings Plans 承諾，提供估計成本節省、Savings Plans 涵蓋範圍和 Savings Plans 使用率的深入分析。這些資訊能協助組織了解 Savings Plans 如何在無須投入時間資源建立模型來分析支出的條件下，應用於每小時的支出。

定期建立相關報告，納入 Savings Plans、預留執行個體和 Amazon EC2 適當調整大小的建議 (來自 AWS Cost Explorer)，以開始降低與穩定狀態的工作負載、閒置和低利用率資源相關聯的成本。識別並收回與已部署資源的雲端浪費相關聯的支出。若建立了大小不當的資源，或是發現非預期的不同用量模式時，就會發生雲端浪費。遵循 AWS 最佳實務來減少浪費，或請您的客戶團隊和合作夥伴協助您[最佳化和節省](#)雲端成本。

定期產生報告以找出更好的資源採購選項，進而降低工作負載的單位成本。Savings Plans、預留執行個體或 Amazon EC2 Spot 執行個體等購買選項可為容錯工作負載提供最大的成本節省效益，並且可讓利益相關者 (企業擁有者、財務和技術團隊) 參與這些承諾討論。

將報告分享給相關各方，使其了解可能有助於降低雲端總體擁有成本 (TCO) 的機會或新版本公告。採用新的服務、區域、功能、解決方案或新方法來實現進一步的成本降低。

實作步驟

- 設定 AWS Budgets：針對您的工作負載在所有帳戶上設定 AWS Budgets。使用標籤來設定整體帳戶支出的預算，以及工作負載的預算。
 - [Well-Architected 實驗室：成本與管控用量](#)
- 成本優化報告：設置定期週期來討論和分析工作負載的效率。使用建立的指標來報告所達到的指標和達到指標所需的成本。識別並修正任何負面趨勢，並識別您可以在整個組織中推廣的正面趨勢。報告參與者應包含應用程式團隊和擁有者、財務和雲端成本相關重要決策者的代表。

資源

相關文件：

- [AWS Cost Explorer](#)
- [AWS Trusted Advisor](#)
- [AWS Budgets](#)
- [AWS Cost and Usage Report](#)
- [AWS Budgets 最佳實務](#)
- [Amazon S3 分析](#)

相關範例：

- [開始優化 AWS 雲端成本的關鍵方法](#)

COST01-BP06 主動監控成本

實作工具和儀表板以主動監控工作負載的成本。定期使用已設定的工具或現成可用的工具來審查成本。不要只在收到通知時才查看成本和類別。主動監控和分析成本有助於識別正面趨勢，並讓您在整個組織中加以推廣。

未建立此最佳實務時的曝險等級：中

實作指引

建議監控組織內的成本與用量，而不只是在發生例外狀況或異常狀況時。在所有辦公室或工作環境中均可以使用高度可見的儀表板，確保了關鍵人員可存取所需的資訊，並且這些儀表板指出組織專注於成本優化的程度。可見的儀表板可讓您主動推廣成功的成果，並在整個組織中加以實作。

建立每日或頻繁的例行作業以使用 [AWS Cost Explorer](#) 或任何其他儀表板 (例如 [Amazon Quick](#))，查看成本並主動分析。在 AWS 帳戶層級、工作負載層級或特定 AWS 服務層級使用群組和篩選來分析 AWS 服務用量與成本，並驗證是否符合預期。使用每小時和資源層級精細度與標籤，來篩選及識別最高排名資源所產生的成本。也可以使用 [成本智慧儀表板](#) (由 AWS Solutions Architects 建立的 [Amazon Quick](#) 解決方案) 建立自己的報告，並將您的預算與實際成本和用量進行比較。

實作步驟

- 成本優化報告：設置定期週期來討論和分析工作負載的效率。使用建立的指標來報告所達到的指標和達到指標所需的成本。識別並修正任何負面趨勢，並識別要在整個組織中推廣的正面趨勢。報告應讓應用程式團隊和擁有者、財務和管理層的代表參與。
- 針對成本和用量建立並啟動每日精細度 [AWS Budgets](#)，及時採取行動以防止任何潛在的成本超支：AWS Budgets 可讓您設定警示通知，因此如果任何預算類型超出預先設定的閾值，您可以隨時了解情況。AWS Budgets 的最佳運用方式是將您預期的成本與用量設為限制，如此即可將任何超過預算的部分視為超支。
- 建立 AWS Cost Anomaly Detection 作為成本監視器：[AWS Cost Anomaly Detection](#) 會使用進階機器學習技術來識別異常支出與根本原因，以便您迅速做出因應。它可讓您設定成本監視器以定義您要評估的支出區段 (例如個別 AWS 服務、成員帳戶、成本分配標籤和成本類別)，並且可讓您設定接收提醒通知的時間、位置和方式。每個監視器可以為企業擁有者和技術團隊連結多個提醒訂閱，包括每個訂閱的名稱、成本影響閾值和提醒頻率 (個別提醒、每日摘要、每週摘要)。
- 使用 AWS Cost Explorer 或整合您的 AWS Cost and Usage Report (CUR) 資料與 Amazon Quick 儀表板，以視覺化方式呈現組織的成本：AWS Cost Explorer 具有易於使用的介面，可讓您視覺化、了解和管理一段時間內的 AWS 成本和用量。[成本智慧儀表板](#) 是一個可自訂且可供存取的儀表板，可協助您建立自身成本管理和優化工具的基礎。

資源

相關文件：

- [AWS Budgets](#)
- [AWS Cost Explorer](#)

- [每日成本與用量預算](#)
- [AWS Cost Anomaly Detection](#)

相關範例：

- [Slack 的 AWS Cost Anomaly Detection 提醒](#)

COST01-BP07 up-to-date保留新的服務版本

定期諮詢專家或 AWS 合作夥伴，以考慮哪些服務和功能提供較低的成本。檢閱 AWS 部落格和其他資訊來源。

未建立此最佳實務時的曝險等級：中

實作指引

AWS 不斷增加新功能，因此您可以利用最新的技術更快地進行實驗和創新。您可以實作新的 AWS 服務和功能，以提高工作負載的成本效率。定期檢閱 [AWS 成本管理](#)、[AWS 新聞部落格](#)、[AWS 成本管理部落格](#)、[AWS 最新消息](#)，以取得新的服務和功能版本的相關資訊。新文章提供發佈所有 AWS 服務、功能和區域擴展公告的簡短概觀。

實作步驟

- 訂閱部落格：前往 AWS 部落格頁面並訂閱 What's New Blog 和其他相關部落格。可以使用您的電子郵件地址在[通訊偏好](#)頁面註冊。
- 訂閱 AWS 新聞：定期檢閱[AWS 新聞部落格](#)和 [What's New with AWS](#)，以取得新服務和功能版本的相關資訊。訂閱RSS摘要，或使用電子郵件追蹤公告和版本。
- 遵循 AWS 折價：我們所有服務的一般降價，對於從規模中獲得的客戶 AWS 來說，傳遞經濟效率是標準的方式。截至 2023 年 9 月 20 日，自 2006 年以來 AWS，價格已降低 134 次。如果您有任何商業決策因價格考量而未定，您可以在降價和新的服務整合之後再次加以審查。您可以在[AWS 新聞部落格的降價類別](#)中了解先前的降價工作，包括 Amazon Elastic Compute Cloud（Amazon EC2）執行個體。
- AWS 事件和會議：參加您的本機 AWS 高峰會，以及與您所在區域的其他組織的任何本機會議。如果您無法親自出席，請嘗試參加虛擬活動，以聽取 AWS 專家和其他客戶商業案例的更多訊息。
- 與您的客戶團隊會面：與您的客戶團隊排定一個定期規律，與他們開會並討論產業趨勢和 AWS 服務。與您的客戶經理、解決方案架構師和支援團隊進行討論。

資源

相關文件：

- [AWS 成本管理](#)
- [新功能 AWS](#)
- [AWS 新聞部落格](#)

相關範例：

- [Amazon EC2 – 15 年的 IT 成本最佳化和節省](#)
- [AWS 新聞部落格 - 折價](#)

COST01-BP08 建立成本感知文化

在您的組織中實作變更或計畫，以建立成本感知文化。建議從較小的計畫開始，然後隨著能力的增強和使用雲端的增加，再實作更大和更廣泛的計畫。

未建立此最佳實務時的曝險等級：低

實作指引

成本感知文化可讓您透過在組織中以系統和分散的方式執行最佳實務，擴展成本優化和雲端財務管理 (財務營運、智慧雲端中心、雲端維運團隊等等)。相較於嚴格的由上而下、集中式方法，成本感知可讓您輕鬆地在整個組織建立高水準的能力。

建立雲端運算的成本感知 (尤其是對於雲端運算的主要成本動因)，可讓團隊了解成本方面的任何變更預期會產生的結果。存取雲端環境的團隊應了解定價模型，以及傳統內部部署資料中心與雲端運算之間的差異。

成本感知文化的主要優點是，技術團隊可主動且持續地優化成本 (例如，在建構新的工作負載，或對現有的工作負載進行變更時，會將其視為非功能性需求)，而不是等到必要時才被動執行成本優化。

文化中的小幅變化可以對目前和未來工作負載的效率產生很大的影響。這些範例包括：

- 在工程團隊中提供可見性和建立感知以了解其工作性質，及其對成本方面有何影響。
- 在您的組織中對成本和用量進行遊戲化。這可以透過公開可見的儀表板，或比較各團隊的標準化成本和用量報告來實現 (例如，每一工作負載的成本和每一交易的成本)。

- 認識成本效益。公開或私下獎勵自願或未經要求完成的成本優化成就，並從錯誤中學習，以避免重蹈覆轍。
- 建立由上而下的組織要求，讓工作負載依預先定義的預算執行。
- 探究企業的變更需求，以及要求的變更對於基礎架構或工作負載組態的成本影響，以確保您只須就需要的部分付費。
- 確定變更的規劃師了解預期的變更有何成本影響，且已經過利益相關者的確認，應以符合成本效益的方式提供商業成果。

實作步驟

- 向技術團隊報告雲端成本：增強成本感知，並且為財務與業務利益相關者建立效率 KPI。
- 通知利益相關者或團隊成員有已規劃的變更：在每週變更會議期間建立議程項目來討論已規劃的變更，以及對於工作負載的成本效益影響。
- 與您的客戶團隊會面：安排與客戶團隊的定期會面，與他們討論產業趨勢和 AWS 服務。與您的客戶經理、架構師和支援團隊進行討論。
- 分享成功案例：分享關於任何工作負載、AWS 帳戶 或組織降低成本的成功案例，以建立成本優化方面的正面態度與鼓勵。
- 培訓：確定技術團隊或其成員已受過 AWS 雲端 資源成本感知的相關培訓。
- AWS 活動和會議：參加當地的 AWS 高峰會，以及與您當地區域的其他組織一同參加當地會議。
- 訂閱部落格：前往 AWS 部落格頁面並訂閱[最新消息部落格](#)和其他相關部落格，以遵循 AWS 分享的新版本、實作、範例和變更。

資源

相關文件：

- [AWS 部落格](#)
- [AWS 成本管理](#)
- [AWS 新聞部落格](#)

相關範例：

- [AWS 雲端財務管理](#)

COST01-BP09 從成本最佳化量化商業價值

量化成本優化所帶來的商業價值，可讓您了解給組織提供的全部效益。由於成本優化是一項必要的投資，因此量化商業價值可讓您向利益相關者解釋投資報酬率。量化商業價值有助於您在未來就成本優化投資獲得利益相關者更多的支持，並提供一個框架來衡量組織成本優化活動的成果。

未建立此最佳實務時的曝險等級：中

實作指引

量化商業價值意味著衡量企業從採取的行動和決策中所獲得的好處。商業價值可以有形的 (例如費用降低或利潤增加)，也可以是無形的 (例如品牌信譽提升或客戶滿意度變高)。

量化成本最佳化所帶來的商業價值意味著判斷您在更有效率地支出成本上所做的努力，可以讓您獲得多少價值或收益。例如，如果公司花費 100,000 美元在 上部署工作負載，AWS 然後對其進行最佳化，則新成本只會變成 80,000 美元，而不會犧牲品質或輸出。在這種情況下，成本最佳化所帶來的量化商業價值會是節省了 20,000 美元。不過，除了節省成本外，公司還可以從更快的交貨時間、提高的客戶滿意度或成本最佳化努力所產生的其他指標等方面來量化價值。利益相關者需要就成本最佳化的潛在價值、工作負載的最佳化成本和回報價值做出決策。

除了報告成本優化所帶來的節省之外，建議您量化提供的額外價值。成本優化效益通常根據每個業務成果所較低的成本進行量化。例如，您可以在購買 Savings Plans 時量化 Amazon Elastic Compute Cloud (Amazon EC2) 的成本節省，這可降低成本並維持工作負載輸出層級。您可以在移除閒置的 Amazon EC2 執行個體，或刪除未連接的 Amazon Elastic Block Store (Amazon EBS) 磁碟區時，量化 AWS 支出的成本降低。

不過，成本優化的消費絕非僅限於成本降低或避免。考慮擷取額外資料，以測量效率改善和商業價值。

實作步驟

- 評估商業利益：這是分析和調整 AWS 雲端 成本的程序，以最大限度地提高從每一美元花費中獲得的利益。請不要不顧商業價值，一味地降低成本，而是要考慮成本最佳化所帶來的商業效益和投資回報，這樣才有可能從支出的成本中獲得更多價值。重點在於聰明地支出，以及在能產生最佳回報的領域進行投資和支出。
- 分析預測 AWS 成本：預測有助於財務利益相關者與其他內部和外部組織利益相關者設定期望，並可以改善組織的財務可預測性。[AWS Cost Explorer](#) 可用於執行成本和用量的預測。

資源

相關文件：

- [AWS 雲端 經濟效益](#)
- [AWS 部落格](#)
- [AWS 成本管理](#)
- [AWS 新聞部落格](#)
- [Well-Architected 可靠性支柱白皮書](#)
- [AWS Cost Explorer](#)

相關影片：

- [在 Windows 開啟時解除鎖定商業價值 AWS](#)

相關範例：

- [測量並最大化 Customer 360 的商業價值](#)
- [採用 Amazon Web Services 受管資料庫的商業價值](#)
- [Amazon Web Services 對於獨立軟體供應商的商業價值](#)
- [雲端現代化的商業價值](#)
- [遷移到 Amazon Web Services 的商業價值](#)

了解支出和用量

問題

- [COST 2. 如何管控用量？](#)
- [COST 3. 如何監控您的成本和用量？](#)
- [COST 4. 如何停用資源？](#)

COST 2. 如何管控用量？

制訂政策和機制以確認產生合理的成本，同時達成目標。您可以運用相互制衡的方法，在不超支的情況下進行創新。

最佳實務

- [COST02-BP01 根據貴組織的需求制定政策](#)
- [COST02-BP02 實作總目標和具體目標](#)

- [COST02-BP03 實作帳戶結構](#)
- [COST02-BP04 實作群組和角色](#)
- [COST02-BP05 實作成本控制](#)
- [COST02-BP06 追蹤專案生命週期](#)

COST02-BP01 根據貴組織的需求制定政策

制定定義組織如何管理資源的政策，並定期加以檢查。政策應涵蓋資源和工作負載的成本面向，包括資源生命週期中的建立、修改和停用。

未建立此最佳實務時的曝險等級：高

實作指引

了解組織的成本和動因對於有效管理成本和用量，以及識別降低成本的機會至關重要。組織通常會營運由多個團隊執行的多個工作負載。這些團隊可能分屬不同組織單位，各有本身的收入流。將資源成本歸因至工作負載、個別組織或產品擁有者的能力，能夠帶動高效使用的行為模式，並且有助於減少浪費。精確的成本和用量監控可協助您了解工作負載的優化程度，以及組織單位和產品的獲利程度。這項知識可讓您更明智地決定應將資源分配到組織內的何處。讓組織內所有層級建立用量意識，這是推動變革的關鍵，因為用量變化會帶來成本變化。請考慮採行多面向的方法以了解您的用量和開支。

執行管控的第一步是使用組織的要求來制定雲端使用政策。這些政策定義您的組織如何使用雲端以及如何管理資源。政策應涵蓋資源和工作負載的成本或用量的各面向，包括在資源生命週期中資源的建立、修改和停用。確認已遵循政策和程序，並已實作雲端環境中的任何變更。在 IT 變更管理會議中提出問題，以釐清計畫性變更對成本的影響 (無論是增加還是減少)、商務理由和預期成果。

政策應該簡單易懂，以便有效地在整個組織中實作。政策還需要易於遵循和解釋 (以方便使用) 並且明確 (團隊間不會產生誤解)。此外，必須定期加以檢查 (如我們的機制)，並隨著客戶業務狀況或優先權的變化 (政策會因而過時) 進行更新。

從廣泛的高階政策開始，例如應使用哪個地理區域，或一天中應該執行資源的時間。逐步為各組織單位和工作負載優化政策。常用政策包括可以使用哪些服務和功能 (例如，測試和開發環境中較低效能的儲存體)、不同群組可以使用哪些類型的資源 (例如，開發帳戶中最大的資源大小是中型)，以及這些資源的使用期間長短 (暫時、短期還是一段特定期間)。

政策範例

以下是範例政策，可供您審核以建立自己的雲端治理政策，其重點為成本優化。確實根據組織的要求和利益相關者的請求來調整政策。

- 政策名稱：定義明確的政策名稱，例如「資源優化」和「成本降低」政策。
- 用途：解釋為何應使用此政策，以及預期的結果為何。此政策的目標是要確認部署和執行所需的工作負載以符合業務需求時的最低成本。
- 範圍：明確定義誰應使用此政策，以及何時應使用此政策，例如 DevOps X Team 在 X 環境 (生產或非生產) 將此政策用於美國東部客戶。

政策聲明

1. 根據工作負載的環境和業務要求 (開發、使用者接受度測試、生產前或生產)，選取美國東部 1 或多個美國東部區域。
2. 將 Amazon EC2 和 Amazon RDS 執行個體排程在早上六點到晚上八點之間執行 (東部標準時間 (EST))。
3. 在八小時後停止所有未使用的 Amazon EC2 執行個體，並在閒置 24 小時後停止未使用的 Amazon RDS 執行個體。
4. 在非生產環境中閒置 24 小時後，終止所有未使用的 Amazon EC2 執行個體。提醒 Amazon EC2 執行個體擁有者 (根據標籤) 審核其生產環境中已停止的 Amazon EC2 執行個體，並通知他們如果 Amazon EC2 執行個體未使用，將在 72 小時內終止。
5. 使用一般執行個體系列和大小 (例如 m5.large)，然後根據 CPU 和記憶體使用率，使用 AWS Compute Optimizer 調整執行個體大小。
6. 使用自動擴展根據流量動態調整執行中的執行個體數量，以訂定優先順序。
7. 對非關鍵工作負載使用 Spot 執行個體。
8. 審核容量要求，以認可可預測工作負載的 Savings Plans 或預留執行個體，並通知雲端財務管理團隊。
9. 使用 Amazon S3 生命週期政策將不常存取的資料移至成本較低的儲存層。若未定義保留政策，請使用 Amazon S3 Intelligent Tiering 將物件自動移至封存層。
10. 使用 Amazon CloudWatch 監控資源使用率並設定警示以觸發擴展事件。
11. 針對每個 AWS 帳戶，使用 AWS Budgets 根據成本中心和業務單位設定帳戶的成本及用量預算。
12. 使用 AWS Budgets 設定帳戶的成本和用量預算，可協助您掌握支出並避免出現非預期的帳單，進而讓您更有效地控制成本。

程序：提供實作此政策的詳細程序，或參閱說明如何實作每項政策聲明的其他文件。本節應提供執行政策要求的逐步指示。

若要實作此政策，您可以使用各種第三方工具或 AWS Config 規則來檢查是否符合政策聲明，並使用 AWS Lambda 函數觸發自動修復動作。您也可以使用 AWS Organizations 來強制執行政策。此外，您應定期審核資源用量，並視需要調整政策，以確認政策持續符合您的商業需求。

實作步驟

- **與利益相關者會面：**若要制定政策，請要求組織內的利益相關者 (雲端業務辦公室、工程師或執行政策的功能決策者) 指定其要求，並將其記錄下來。採取反復的方法，從廣泛討論開始，然後在每個步驟持續細化至最小的單位。團隊成員包括對工作負載有直接關係的人員，例如組織單位或應用程式擁有者，以及支援群組，例如安全和財務團隊。
- **獲取確認：**確定團隊成員均同意誰可對 AWS 雲端 進行存取及部署的政策。請確定成員遵循組織的政策，並確認其資源建立符合議定的政策和程序。
- **建立上線培訓課程：**要求新進的組織成員完成上線培訓課程，以建立對成本的掌握度和組織要求。他們可以根據自身過往的經驗採行不同的政策，也可以完全不列入考量。
- **定義工作負載的位置：**定義工作負載運作的位置，包括國家和國家中的區域。這項資訊用來對應至 AWS 區域 和可用區域。
- **定義並分組服務和資源：**定義工作負載所需的服務。針對每項服務，指定所需的類型、大小和資源數量。依職能定義資源群組，例如應用程式伺服器或資料庫儲存體。資源可屬於多個群組。
- **依職能定義並分組使用者：**定義與工作負載互動的使用者，專注於使用者執行的操作以及他們如何使用工作負載，而不是專注於他們的身分或他們在組織中的位置。將類似的使用者或職能分組在一起。您可以使用 AWS 受管政策做為指南。
- **定義動作：**使用先前識別的位置、資源和使用者的，定義每個項目在其生命週期內 (開發、營運和停用) 達成工作負載結果所需的動作。根據每個位置中的群組 (不是群組中的個別元素) 來識別動作。從廣泛地讀取或寫入開始，然後縮小精細至每項服務的特定動作。
- **定義審查期間：**工作負載和組織需求可能會隨時間變更。定義工作負載審查排程，以確保其與組織優先事項保持一致。
- **記錄政策：**確認組織可視需要存取已定義的政策。這些政策用於實作、維護和稽核環境的存取權。

資源

相關文件：

- [雲端中的變更管理](#)
- [工作職能的 AWS 受管政策](#)
- [AWS 多帳戶帳單策略](#)

- [AWS 服務的動作、資源及條件索引鍵](#)
- [AWS 管理與管控](#)
- [使用 IAM 政策控制對 AWS 區域的存取](#)
- [全球基礎架構區域和可用區域](#)

相關影片：

- [大規模的 AWS 管理與管控](#)

COST02-BP02 實作總目標和具體目標

為您的工作負載實作成本與用量的總目標和具體目標。總目標可為您的組織提供預期成果的方向，具體目標則可提供要為您的工作負載達成的特定可測量成果。

未建立此最佳實務時的曝險等級：高

實作指引

為您的組織制定成本與用量總目標和具體目標。作為 AWS 上成長中的組織，設定和追蹤成本優化的總目標是非常重要的。這些目標或[關鍵績效指標 \(KPI\)](#) 可包含隨需支出的百分比，或是否採用特定優化服務 (例如 AWS Graviton 執行個體或 gp3 EBS 磁碟區類型) 等等。設定可衡量和可實現的總目標有助於衡量效率的改善情況，這對於業務營運非常重要。總目標可為您的組織提供預期結果的指引和方向。

具體目標是要實現的具體可衡量成果。簡言之，總目標是您的努力方向，具體目標則表示該方向有多遠，以及多久可以達成總目標 (其原則為具體 (Specific)、可測量 (Measurable)、可指派 (Assignable)、務實 (Realistic) 和及時 (Timely)，簡稱 SMART)。舉例來說，平台用量大幅增加，而成本僅稍微增加 (非線性)，即為總目標。平台用量增加 20%，成本增加少於百分之五，則是具體目標範例。另一個常見的總目標是工作負載每六個月必須更有效率。相關的具體目標是每個業務指標的成本每六個月需要減少百分之五。使用正確的指標，並為組織設定已計算好的 KPI。可以從基礎 KPI 開始，並在稍後根據業務需求進行發展。

成本優化的總目標是提高工作負載效率，這對應於工作負載的每個業務成果的成本隨著時間而降低。為所有工作負載實作這個總目標，並設定具體目標，例如每六個月至一年將效率提高百分之五。在雲端中，可以透過建立成本最佳化功能以及發行新服務和功能來達成此目標。

具體目標是您希望達到以達到的可量化基準，以實現總體目標，而基準則會將您的實際結果與具體目標進行比較。使用 KPI 針對每個運算單位服務的成本 (例如 Spot 採用、Graviton 採用、最新執行個體類型和隨需涵蓋範圍)、儲存服務 (例如 EBS GP3 採用、過時的 EBS 快照和 Amazon S3 標準儲存) 或資

料庫服務用量 (例如 RDS 開放原始碼引擎、Graviton 採用和隨需涵蓋範圍) 建立基準。這些基準和 KPI 可協助您確認是否以最具成本效益的方式使用 AWS 服務。

下表列出標準 AWS 指標，以供參考。每個組織可以針對這些 KPI 設定不同的目標值。

類別	KPI (%)	描述
運算	EC2 用量涵蓋範圍	使用 SP+RI+Spot 的 EC2 執行個體 (以成本或小時為單位) 與 EC2 執行個體的總計 (以成本或小時為單位) 的比較
運算	計算 SP/RI 使用率	與總體可用的 SP 或 RI 小時數相比，已使用的 SP 或 RI 小時數
運算	EC2/小時成本	EC2 成本除以該小時內執行的 EC2 執行個體數量
運算	vCPU 成本	所有執行個體的每個 vCPU 成本
運算	最新一代執行個體	Graviton (或其他新一代執行個體類型) 上的執行個體百分比
資料庫	RDS 涵蓋範圍	使用 RI 的 RDS 執行個體 (以成本或小時為單位) 與 RDS 執行個體的總計 (以成本或小時為單位) 的比較
資料庫	RDS 使用率	與總體可用的 RI 小時數相比，已使用的 RI 小時數
資料庫	RDS 執行時間	RDS 成本除以該小時內執行的 RDS 執行個體數量
資料庫	最新一代執行個體	Graviton (或其他現代執行個體類型) 上的執行個體百分比

類別	KPI (%)	描述
儲存	儲存使用率	最佳化儲存成本 (例如 Glacier、Deep Archive 或 Infrequent Access) 除以總儲存成本
標記	未標記資源	Cost Explorer : 1. 篩選掉抵用金、折扣、稅金、退款、市場，並複製最新的每月成本 2. 在 Cost Explorer 中選取僅顯示未標記的資源 3. 將未標記資源中的金額除以您的每月成本。

使用此表格，包括目標或基準值，應根據組織目標計算這些值。您需要衡量業務的某些指標，並了解該工作負載的業務成果，以定義準確且切合實際的 KPI。當您評估組織內的績效指標時，請區分服務於不同目的之不同類型的指標。這些指標主要衡量技術基礎設施的效能和效率，而不是直接衡量整體業務影響。例如，它們可能會追蹤伺服器響應時間、網路延遲或系統正常運行時間。這些指標對於評估基礎設施如何支援組織的技術操作至關重要。但是，它們不能直接洞察更廣泛的業務目標，例如客戶滿意度，收入增長或市場份額。為了全面了解業務績效，請使用與業務成果直接相關的策略性業務指標來補充這些效率指標。

要能夠近乎即時地檢視 KPI 和相關節省機會，並追蹤一段時間內的進度。若要開始定義和追蹤 KPI 總目標，建議您使用[雲端智慧儀表板](#) (CID) 中的 KPI 儀表板。根據來自成本和用量報告 (CUR) 的資料，KPI 儀表板會提供一系列建議的成本優化 KPI，讓您能夠設定自訂總目標以及追蹤一段時間內的進度。

如果您有其他解決方案可以設定和追蹤 KPI 總目標，請確定組織中的所有雲端財務管理利益相關者都採用這些方法。

實作步驟

- **定義預期的用量等級：**首先，請關注用量等級。與應用程式擁有者、行銷團隊和更大的業務團隊互動，以了解工作負載的預期用量等級。客戶需求如何隨著時間而變更，以及因季節性增加或行銷活動會發生哪些變更？
- **定義工作負載資源與成本：**定義用量等級後，量化達成這些用量等級所需的工作負載資源變更。您可能需要為工作負載元件增加資源的大小或數量、增加資料傳輸，或將工作負載元件變更為特定等級的不同服務。指定每個要點的成本，並預測當用量發生變化時成本會有什麼變化。
- **定義業務總目標：**從預期用量和成本變更中取得輸出，將此項目與預期的技術變更或任何您正在執行的計畫結合，並制定工作負載的總目標。總目標必須涵蓋用量和成本，以及兩者之間的關係。總目標必須簡單具體，以協助大家了解企業預期的成果 (例如，確保將未使用的資源控制在特定成本水位以下)。無須為每個未使用的資源類型定義總目標，也不需要為總目標和具體目標定義造成損失的成本。如果預期有成本變更但用量不變，請確認制定有組織計畫 (例如培訓和教育等能力打造計畫)。
- **定義具體目標：**對於定義的每個總目標，指定可測量的具體目標。如果總目標是要提高工作負載的效率，具體目標將會量化改善的程度 (通常是所有經費所獲得的業務輸出)，及其達成時間。例如，可設定一個總目標，以盡量減少因過度佈建而造成的浪費。有了這個總目標後，您的具體目標可能是生產工作負載第一層中的運算過度佈建產生的浪費不應超過分層運算成本的 10%。此外，第二個具體目標可能是生產工作負載第二層中的運算過度佈建產生的浪費不應超過分層運算成本的 5%。

資源

相關文件：

- [工作職能的 AWS 受控政策](#)
- [AWS 多帳戶帳單策略](#)
- [使用 IAM 政策控制對 AWS 區域 的存取](#)
- [S.M.A.R.T. 目標](#)
- [如何使用 CID KPI 儀表板追蹤成本最佳化 KPI](#)

相關影片：

- [Well-Architected 實驗室：總目標和具體目標 \(Level 100\)](#)

相關範例：

- [什麼是單位指標？](#)

- [選擇單位指標以支援您的業務](#)
- [實務中的單位指標 — 經驗教訓](#)
- [單位指標如何幫助在業務職能之間建立一致性](#)

COST02-BP03 實作帳戶結構

實作與您的組織對應的帳戶結構。這有助於在整個組織中分配和管理成本。

未建立此最佳實務時的曝險等級：高

實作指引

AWS Organizations 可讓您建立多個 AWS 帳戶，當您在 AWS 上擴展工作負載時，這可協助您集中管控環境。您可以採用組織單位 (OU) 結構來為 AWS 帳戶進行分組，再於每個組織單位下建立多個 AWS 帳戶，藉此塑造組織階層的模型。若要建立帳戶結構，您必須先決定要以哪個 AWS 帳戶作為管理帳戶。之後，您可以遵循[管理帳戶最佳實務](#)和[成員帳戶最佳實務](#)，根據您設計的帳戶結構來建立新的 AWS 帳戶或選擇現有帳戶作為成員帳戶。

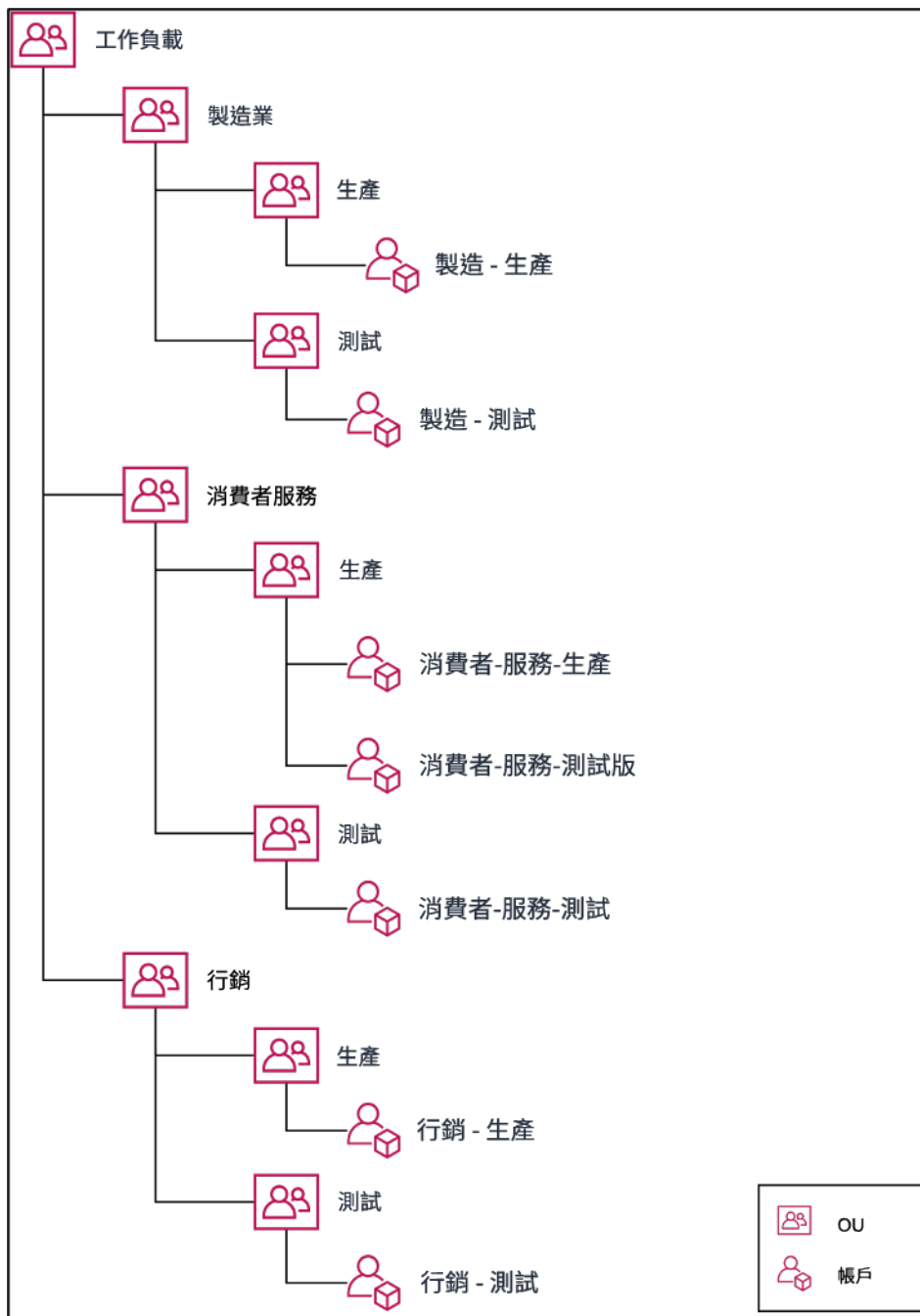
無論您的組織規模或用量為何，都建議您一律要有至少一個管理帳戶，以及一個與管理帳戶連結的成員帳戶。所有工作負載資源都只應位於成員帳戶內，請勿在管理帳戶內建立任何資源。關於應該擁有多少個 AWS 帳戶，並沒有一體適用的答案。請評估您目前和未來的運作與成本模式，確保 AWS 帳戶的結構呼應組織的目標。有些公司基於業務原因建立多個 AWS 帳戶，例如：

- 組織單位、成本中心或特定工作負載之間需要行政管理或會計年度和帳單上的區隔。
- AWS 服務限制是依照特定工作負載區分所設定。
- 工作負載和資源之間需要區隔和隔離。

在 [AWS Organizations](#) 中，[合併帳單](#)會在一個或多個成員帳戶與管理帳戶之間建立結構。成員帳戶可讓您依群組隔離和區分成本和用量。常見實務是各組織單位分別有成員帳戶 (例如財務、行銷和銷售)，或是各個環境生命週期分立 (例如開發、測試和生產)，或是各工作負載分立 (工作負載 a、b 和 c)，再使用合併帳單彙總這些連結帳戶。

合併帳單可讓您將多個 AWS 帳戶的款項合併至單一管理帳戶之下，同時仍為各連結帳戶的活動提供可見度。由於成本和用量的在管理帳戶中彙總，這可讓您獲得最大的服務容量折扣以及最大的使用承諾折扣 (Savings Plans 和預留執行個體)，以享受最高折扣。

下圖顯示如何使用 AWS Organizations 與組織單位 (OU) 來將多個帳戶分組，並將多個 AWS 帳戶放到每個 OU 底下。建議您使用 OU 來處理各種使用案例和工作負載，以便提供用於整理帳戶的模式。



在組織單位下對多個 AWS 帳戶進行分組的範例。

[AWS Control Tower](#) 可以快速建立和設定多個 AWS 帳戶，確保管控符合組織的要求。

實作步驟

- 定義分隔要求：分隔要求是多個因素的組合，包括安全性、可靠性和財務結構。依序處理每個因素，並指定工作負載或工作負載環境是否應與其他工作負載分開。為了安全，我們必須遵守存取和資料要求。為求可靠，我們必須有所限制，以免環境和工作負載影響其他資源。請定期審核 Well-

Architected 架構的安全性和可靠性支柱，並遵循其中所提供的最佳實務。財務結構會建立嚴格的財務分隔 (不同的成本中心、工作負載擁有權和責任)。常見的分隔範例是生產和測試工作負載會在不同的帳戶開始執行，或使用單獨的帳戶，以便將發票和帳單資料提供給組織內的個別業務單位或部門，或是擁有帳戶的利益相關者。

- 定義分組要求：分組要求不會覆寫分隔要求，而是用來協助管理。將不需要分隔的類似環境或工作負載分成同一組。例如，將來自一或多個工作負載的多個測試或開發環境分組在一起。
- 定義帳戶結構：使用這些分隔和分組，為每個群組指定一個帳戶，並維護分隔要求。這些帳戶是您的成員帳戶或連結帳戶。透過將這些成員帳戶分組到單一管理帳戶或付款人帳戶下，您可以結合用量，以讓所有帳戶獲得更多數量折扣，而為所有帳戶提供單一帳單。您可以分隔帳單資料，以便在每個成員帳戶中檢視單獨的帳單資料。如果不允許透過任何其他帳戶查看某個成員帳戶中的用量或帳單資料，或是需要與 AWS 分開的帳單，請定義多個管理帳戶或付款人帳戶。在這種情況下，每個成員帳戶都有自己的管理帳戶或付款人帳戶。資源應一律放在成員或連結帳戶中。管理帳戶或付款人帳戶只能用於管理。

資源

相關文件：

- [使用成本分配標籤](#)
- [工作職能的 AWS 受控政策](#)
- [AWS 多帳戶帳單策略](#)
- [使用 IAM 政策控制對 AWS 區域 的存取](#)
- [AWS Control Tower](#)
- [AWS Organizations](#)
- [管理帳戶](#)和[成員帳戶](#)的最佳實務
- [使用多個帳戶組織您的 AWS 環境](#)
- [開啟共享的預留執行個體和 Savings Plans 折扣](#)
- [合併帳單](#)
- [合併帳單](#)

相關範例：

- [分割 CUR 和共用存取權](#)

相關影片：

- [AWS Organizations 簡介](#)
- [設定使用 AWS Organizations 最佳實務的多帳戶 AWS 環境](#)

相關範例：

- [定義電信公司的 AWS 多帳戶策略](#)
- [最佳化 AWS 帳戶的最佳實務](#)
- [組織單位的 AWS Organizations 最佳實務](#)

COST02-BP04 實作群組和角色

實作符合您政策的群組和角色，並控制哪些人員可以建立、修改或停用每個群組中的執行個體和資源。例如，實作開發、測試和生產群組。這適用於 AWS 服務和第三方解決方案。

未建立此最佳實務時的曝險等級：低

實作指引

使用者角色和群組是設計和實作安全高效系統的基礎建置組塊。角色和群組可協助組織在控制需求與靈活性和生產力的要求兩方面取得平衡，從而最終能支援組織目標和使用者的需求。如同 AWS Well-Architected Framework 安全支柱的[身分識別與存取管理](#)部分所建議，您需要強大的身分識別管理和權限，才能在適當的條件下為適當的人員提供適當資源的存取權。使用者只會獲得要完成其任務所需的存取權。這可將未經授權存取或濫用的相關風險降至最低。

在制定政策後，您可以在組織內建立邏輯群組和使用者角色。這可讓您指派許可、控制使用情況，並協助實作強大的存取控制機制，防止有人未經授權存取敏感資訊。從簡要的人員分組開始。通常這與組織單位和工作角色 (例如 IT 部門的系統管理員、財務控制者或商業分析師) 相符。這些群組會將執行類似任務且需要類似存取權限的人員進行分類。角色定義群組必須執行的工作。管理群組和角色的許可會比管理個別使用者的許可容易。角色和群組能以一致且有系統的方式為所有使用者指派許可，以避免錯誤和不一致。

當使用者的角色變更時，管理員可以調整角色或群組層級的存取權，而不是重新設定個別使用者帳戶。例如，IT 的系統管理員需要建立所有資源的存取權限，但分析團隊成員只需要建立分析資源的權限。

實作步驟

- 實作群組：使用組織政策中定義的使用者群組，視需要實作對應的群組。如需有關使用者、群組和驗證的最佳實務，請參閱 AWS Well-Architected Framework 的[安全支柱](#)。
- 實作角色和政策：使用組織政策中定義的動作，建立所需的角色和存取政策。如需有關角色和政策的最佳實務，請參閱 AWS Well-Architected Framework 的[安全支柱](#)。

資源

相關文件：

- [工作職能的 AWS 受控政策](#)
- [AWS 多帳戶帳單策略](#)
- [AWS Well-Architected Framework 安全支柱](#)
- [AWS Identity and Access Management \(IAM\)](#)
- [AWS Identity and Access Management 政策](#)

相關影片：

- [為何使用 Identity and Access Management](#)

相關範例：

- [使用 IAM 政策控制對 AWS 區域 的存取](#)
- [開始您的雲端財務管理之旅：雲端成本操作](#)

COST02-BP05 實作成本控制

根據組織政策以及定義的群組和角色實作控制。這些控制措施可證明成本的發生始終符合組織要求：例如，控制對區域或資源類型的存取。

未建立此最佳實務時的曝險等級：中

實作指引

實作成本控制常見的第一步設定在發生偏離政策的成本或用量事件時發出通知。您可以快速採取動作，並驗證是否需要採取糾正措施，而不會限制或對工作負載或新的活動造成負面影響。了解工作負載和環境限制之後，就可以強制執行管控。[AWS Budgets](#) 可讓您針對 AWS 成本、用量和承諾折扣 (Savings

Plans 和預留執行個體) 設定通知並定義每月預算。您可以在彙總成本層級 (例如, 所有成本) 或更精細的層級建立預算, 其中只包含特定維度, 例如連結的帳戶、服務、標籤或可用區域。

透過 AWS Budgets 設定預算限制後, 可使用 [AWS Cost Anomaly Detection](#) 來減少意外成本。AWS Cost Anomaly Detection 是一項成本管理服務, 其使用機器學習來持續監控您的成本和用量, 以偵測異常支出。其可協助您識別異常支出與根本原因, 以便您迅速因應。請先在 AWS Cost Anomaly Detection 中建立成本監視器, 然後設定美元閾值以選擇提醒偏好 (例如, 針對影響金額大於 1,000 美元的異常設定提醒)。收到提醒後, 便能分析異常背後的原因, 以及其對成本的影響。您也可以使用 AWS Cost Explorer 中監控和執行您自己的異常分析。

透過 [AWS Identity and Access Management](#) 和 [AWS Organizations 服務控制政策 \(SCP\)](#), 在 AWS 中強制執行管控政策。IAM 可讓您安全地管理對 AWS 服務和資源的存取。使用 IAM, 您可以控制誰可以建立或管理 AWS 資源、可建立的資源類型以及建立資源的位置。這可以最大程度地降低在所定義的政策外建立資源的可能性。使用先前建立的角色和群組, 並指派 [IAM 政策](#) 以執行正確的用量。SCP 可集中控制組織中所有帳戶的最大可用許可, 讓您的帳戶符合您的存取控制指引。SCP 只能在開啟所有功能的組織中使用, 而且您可以設定 SCP, 為成員帳戶設定預設拒絕或允許的動作。如需實作存取管理的詳細資訊, 請參閱 [Well-Architected 安全支柱白皮書](#)。

亦可透過管理 [AWS 服務配額](#) 來實作管控。藉由確保服務配額設定為冗餘最低並且正確維護, 可盡量避免建立超出組織要求的資源。為達成此目的, 您必須了解要求的變更速度能有多快、了解進行中的專案 (包括資源的建立與停用) 並將變更配額的實作速度能有多快列入作為考量因素。[服務配額](#) 可在需要時用來增加您的配額。

實作步驟

- 實作支出通知：使用您定義的組織政策, 建立 [AWS Budgets](#) 以在支出超出政策時通知您。設定多個成本預算 (每個帳戶一個), 各帳戶會通知您整體帳戶支出。請針對帳戶中的較小單位, 為每個帳戶設定額外的成本預算。這些單位會根據您的帳戶結構而有所不同。一些常見的範例是 AWS 區域、工作負載 (使用標籤) 或 AWS 服務。請將電子郵件分發清單設定為通知收件人, 而非個人的電子郵件帳戶。您可以設定超過數量時的實際預算, 或使用預測預算來通知預測用量。您也可以預先設定 AWS 預算操作, 以實施特定的 IAM 或 SCP 政策, 或停止目標 Amazon EC2 或 Amazon RDS 執行個體。預算操作可以開始, 也可以要求工作流程核准。
- 實施異常支出的通知：使用 [AWS Cost Anomaly Detection](#) 減少組織中的意外成本, 並分析潛在異常支出的根本原因。在建立成本監視器以識別指定精細度的不尋常支出, 並在 AWS Cost Anomaly Detection 中設定通知後, 其便會在偵測到不尋常支出時向您發出提醒。這可讓您分析異常背後的原因, 並了解其對成本的影響。在設定 AWS Cost Anomaly Detection 時使用 AWS Cost Categories, 可識別哪個專案團隊或業務單位團隊能夠分析非預期成本的根本原因, 並及時採取必要動作。

- **實作用量控制：**使用您定義的組織政策實作 IAM 政策和角色，以指定使用者可以執行的動作，以及他們無法執行的動作。一項 AWS 政策中可包含多項組織政策。使用與您定義政策相同的方式，一開始廣泛定義，然後在每個步驟中套用更精細的控制措施。服務限制也能有效控制用量。在您所有帳戶中實作正確的服務限制。

資源

相關文件：

- [工作職能的 AWS 受控政策](#)
- [AWS 多帳戶帳單策略](#)
- [使用 IAM 政策控制對 AWS 區域 的存取](#)
- [AWS Budgets](#)
- [AWS Cost Anomaly Detection](#)
- [控制您的 AWS 成本](#)

相關影片：

- [如何使用 AWS Budgets 追蹤我的支出和用量](#)

相關範例：

- [IAM 存取管理政策範例](#)
- [服務控制政策範例](#)
- [AWS 預算動作](#)
- [建立 IAM 政策以使用標籤控制對 Amazon EC2 資源的存取](#)
- [限制 IAM 身分對特定 Amazon EC2 資源的存取](#)
- [使用聊天應用程式中的 Amazon Q Developer 進行成本異常偵測的 Slack 整合](#)

COST02-BP06 追蹤專案生命週期

追蹤、測量和稽核專案、團隊和環境的生命週期，以避免使用不必要的資源並節省成本。

未建立此最佳實務時的曝險等級：低

實作指引

透過有效追蹤專案生命週期，組織可以透過強化的規劃、管理和資源最佳化來實現更好的成本控制。透過追蹤所獲得的見解十分寶貴，可讓您做出有助於專案成本效益和整體成功率的明智決策。

追蹤工作負載的整個生命週期可協助您了解何時不再需要工作負載或工作負載元件。現有的工作負載和元件可能正在使用中，但 AWS 當發行新的服務或功能時，它們可以停用或採用。检查工作負載的先前階段。工作負載進入生產環境後，之前的環境可能會停用或大幅降低容量，直到再次需要這些環境為止。

您可以使用時間範圍或提醒來標記資源，以固定審核工作負載的時間。舉例來說，如果開發環境上次是在幾個月前進行審核，那麼現在是時候再次審核，以探索是否可以採用新的服務，或是環境是否正在使用中。您可以在 [myApplications](#) 上將應用程式分組和標記 AWS，以管理和追蹤重要性、環境、上次檢閱和成本中心等中繼資料。可以追蹤工作負載的生命週期，監控和管理應用程式的成本、運作狀態、安全狀態和效能。

AWS 提供各種管理和治理服務，可用於實體生命週期追蹤。您可以使用 [AWS Config](#) 或 [AWS Systems Manager](#) 提供 AWS 資源和組態的詳細清查。建議與您現行的專案或資產管理系統整合，與持續追蹤您的組織進行中的專案和產品。將您目前的系統與提供的豐富事件和指標集相結合，AWS 可讓您建立重大生命週期事件的檢視，並主動管理資源以減少不必要的成本。

與 [Application Lifecycle Management \(ALM\)](#) 類似，追蹤專案生命週期應涉及多個程序、工具和團隊合作，例如設計和開發、測試、生產、支援和工作負載備援。

透過仔細監控專案生命週期的每個階段，組織可以獲得重要的洞見和增強控制，促進成功的專案規劃、實作和完成。這種仔細的監督會驗證專案不僅符合品質標準，而且會準時地在預算內交付，從而提高整體成本效率。

如需有關實作實體生命週期追蹤的詳細資訊，請參閱 [AWS Well-Architected 卓越營運支柱白皮書](#)。

實作步驟

- 建立專案生命週期監控程序：[雲端卓越中心團隊](#)必須建立專案生命週期監控程序。建立結構化與系統化的方法來監控工作負載，以改善專案的控制、可見性和效能。讓監控流程透明、協作並專注於持續改進，以最大程度地提高其有效性和價值。
- 執行工作負載審核：根據組織政策所定義，設定一個定期節奏以稽核現有專案並執行工作負載審核。在稽核上付出的工作量應與組織的大致風險、價值或成本成正比。要納入稽核的關鍵領域包括事件或中斷給組織帶來的風險、對組織的價值或貢獻 (以收入或品牌聲譽來衡量)、工作負載成本 (以資源總成本和營運成本來衡量)，以及工作負載用量 (以每單位時間的組織結果數量來衡量)。如果這些領域在生命週期內發生變化，則需要調整工作負載，例如完整或部分停用。

資源

相關文件：

- [上的標記指南 AWS](#)
- [什麼是 ALM \(應用程式生命週期管理 \) ？](#)
- [工作職能的AWS 受控政策](#)

相關範例：

- [AWS 區域 使用IAM政策控制對 的存取](#)

相關工具

- [AWS Config](#)
- [AWS Systems Manager](#)
- [AWS Budgets](#)
- [AWS Organizations](#)
- [AWS CloudFormation](#)

COST 3. 如何監控您的成本和用量？

制訂政策和程序以監控並適當分配成本。這樣能夠讓您衡量並改善此工作負載的成本效益。

最佳實務

- [COST03-BP01 設定詳細資訊來源](#)
- [COST03-BP02 將組織資訊新增至成本與用量](#)
- [COST03-BP03 識別成本歸因類別](#)
- [COST03-BP04 建立組織指標](#)
- [COST03-BP05 設定帳單和成本管理工具](#)
- [COST03-BP06 根據工作負載指標分配成本](#)

COST03-BP01 設定詳細資訊來源

設定成本管理和報告工具，以增強分析以及成本和用量資料的透明度。設定您的工作負載以建立日誌項目，以便追蹤和隔離成本和用量。

未建立此最佳實務時的曝險等級：高

實作指引

詳細的帳單資訊 (例如成本管理工具中的每小時精細度) 可讓組織更詳細地追蹤其耗用量，並協助他們找出一些成本增加的原因。這些資料來源提供整個組織最準確的成本和用量的檢視。

可以使用 AWS 資料匯出 來建立 AWS Cost and Usage Report (CUR) 2.0 的匯出。這是從 AWS 接收詳細成本和用量資料的最新建議方法。它提供所有 AWS 收費服務的每日或每小時使用精細度、費率、成本和使用屬性 (與 CUR 的資訊相同)，以及一些改進。CUR 中包含所有可能的維度，例如：標記、位置、資源屬性和帳戶 ID。

根據您要建立的匯出類型，有三種匯出類型：標準資料匯出、透過 Quick 整合匯出至成本和用量儀表板，或是舊式資料匯出。

- 標準資料匯出：表格的自訂匯出，可定期交付給 Amazon S3。
- 成本和用量儀表板：匯出和整合至 Quick，以部署預先建置的成本和用量儀表板。
- 舊式資料匯出：舊式 AWS Cost and Usage Report (CUR) 的匯出。

可以使用下列自訂來建立資料匯出：

- 包含資源 ID
- 分割成本分配資料
- 每小時的精細程度
- 版本控制
- 壓縮類型和檔案格式

對於在 Amazon ECS 或 Amazon EKS 上執行容器的工作負載，請啟用分割成本分配資料，以便根據容器工作負載消耗共用運算和記憶體資源的方式，將容器成本分配給個別業務單位和團隊。分割成本分配資料會將新容器層級資源的成本和用量資料引入 AWS Cost and Usage Report。分割成本分配資料是透過運算在叢集上執行的個別 ECS 服務和任務的成本來計算的。

成本和用量儀表板會定期將成本和用量儀表板資料表匯出到 S3 儲存貯體，並將預先建立的成本和用量儀表板部署到 Quick。如果想要在不進行自訂的情況下快速部署成本和用量資料的儀表板，請使用此選項。

如有需要，仍可以使用舊版模式匯出 CUR，在這裡您可以整合其他處理服務，例如 [AWS Glue](#)，以準備要分析的資料，並使用 SQL 來查詢資料，從而透過 [Amazon Athena](#) 執行資料分析。

實作步驟

- 建立資料匯出：使用您想要的資料建立自訂匯出，並控制匯出結構描述。使用基礎 SQL 建立帳單和成本管理資料匯出，並透過與 Quick 整合，將您的帳單和成本管理資料視覺化。也可以使用標準模式匯出資料，以使用 Amazon Athena 等其他處理工具來分析資料。
- 設定成本和用量報告：使用帳單主控台，設定至少一個成本和用量報告。用含所有識別符與資源 ID 的每小時精細度設定報告。您也可以使用不同的精細度建立其他報告，以提供較高層級的摘要資訊。
- 在 Cost Explorer 中設定每小時精細度：若要存取過去 14 天內每小時精細度的成本和用量資料，請考慮在帳單主控台中啟用每小時和資源層級資料。
- 設定應用程式日誌記錄：確認您的應用程式會記錄其交付的每個業務成果，以便追蹤和衡量相應成果。確保此資料的精細度至少為每小時，以便與成本和用量資料相符。如需有關日誌記錄和監控的詳細資訊，請參閱 [Well-Architected 卓越營運支柱](#)。

資源

相關文件：

- [AWS 資料匯出](#)
- [AWS Glue](#)
- [快速：](#)
- [AWS 成本管理定價](#)
- [標記 AWS 資源](#)
- [使用 Cost Explorer 分析成本](#)
- [管理 AWS Cost and Usage Report](#)

相關範例：

- [AWS 帳戶設定](#)
- [用於 AWS 帳單和成本管理的資料匯出](#)

- [AWS Cost Explorer 常用案例](#)

COST03-BP02 將組織資訊新增至成本與用量

根據您的組織、工作負載屬性和成本分配類別來定義標記結構描述，以便您在成本管理工具中篩選及搜尋資源，或監控成本與用量。情況允許時，依據目的、團隊、環境，或其他與您的業務有關的條件，在所有資源間實作一致的標記。

未建立此最佳實務時的曝險等級：中

實作指引

在 [AWS 中實作標記](#)，將組織資訊新增到您的資源，然後將這些資訊新增至您的成本與用量資訊。標籤是鍵/值對；鍵已定義，且在組織中必須是唯一的，而值對於一組資源是唯一的。鍵值對的範例：鍵為 Environment，其值為 Production。生產環境中的所有資源都會有此鍵/值對。標記可讓您使用有意義且相關的組織資訊，來分類和追蹤成本。您可以套用代表組織類別 (例如成本中心、應用程式名稱、專案或擁有者) 的標籤，並識別工作負載及其特性 (例如，測試或生產)，以在整個組織中劃分成本和用量歸屬。

您套用標籤至 AWS 資源 (例如 Amazon Elastic Compute Cloud 執行個體或 Amazon Simple Storage Service 儲存貯體) 並啟用標籤時，AWS 會將此資訊加入至成本和用量報告。您可以對已標記和未標記的資源執行報告和分析，以便更符合內部成本管理政策，並確保準確劃分歸屬。

在組織的各帳戶建立和實作 AWS 標記標準，有助於您以一致統一的方式管理和管控 AWS 環境。在 AWS Organizations 中使用 [標籤政策](#)，定義如何在 AWS Organizations 中將標籤用於帳戶中 AWS 資源的規則。標籤政策可讓您輕鬆採用標準化方法來標記 AWS 資源

[AWS Tag Editor](#) 讓您可以新增、刪除和管理多個資源的標籤。利用標籤編輯器，您會搜尋要加標籤的資源，然後為搜尋結果中的資源管理標籤。

[AWS Cost Categories](#) 讓您可以為成本指派組織的意義，無須在資源上加上標籤。您可以將成本和用量資訊對應到唯一的內部組織結構。您可以定義類別規則，使用帳單維度 (例如帳戶和標籤) 來映射和分類成本。除了標記，這可提供另一個層級的管理功能。您也可以將特定帳戶和標籤對應到多個專案。

實作步驟

- 定義標記結構描述：收集業務中的所有利益相關者，以定義結構描述。這通常包括屬於技術、財務和管理角色的人員。定義所有資源必須具備的標籤清單，以及資源應該具備的標籤清單。確認標籤名稱和值在整個組織中保持一致。

- 標記資源：使用您定義的成本屬性類別，根據類別在工作負載中的所有資源上[放置標籤](#)。使用 CLI、Tag Editor 或 AWS Systems Manager 等工具來提高效率。
- 實作 AWS Cost Categories：您可以在不實作標記的情況下建立 [Cost Categories](#)。Cost Categories 會使用現有的成本和用量維度。從您的結構描述建立類別規則，並將其實作至 Cost Categories。
- 自動化標記：為驗證您在所有資源中保持高層級標記，請自動化標記，以便在建立資源時自動對其進行標記。使用諸如 [AWS CloudFormation](#) 等服務來驗證資源在建立時是否已加上標記。您也可以使用 Lambda 函數建立自訂解決方案以進行自動標記，或者使用可定期掃描工作負載並移除任何未標記資源的微型服務，這非常適合用於測試和開發環境。
- 監控和報告標記：為驗證您可在整個組織中保有高層級標記，請報告並監控工作負載間的標籤。您可以使用 [AWS Cost Explorer](#) 檢視已標記和未標記資源的成本，或使用 [Tag Editor](#) 等服務。定期審查未標記資源的數量，並採取措施來新增標籤，直至達到所需的標記層級。

資源

相關文件：

- [標記最佳實務](#)
- [AWS CloudFormation 資源標籤](#)
- [AWS Cost Categories](#)
- [標記 AWS 資源](#)
- [使用 AWS Budgets 分析成本](#)
- [使用 Cost Explorer 分析成本](#)
- [管理 AWS 成本與用量報告](#)

相關影片：

- [如何標記 AWS 資源以按成本中心或專案劃分帳單](#)
- [標記 AWS 資源](#)

COST03-BP03 識別成本歸因類別

識別組織分類 (例如業務單位、部門或專案)，這些分類可以將組織內的成本分配給內部取用實體。使用這些分類來強制執行支出權責劃分、建立成本感知並推動有效的取用行為。

未建立此最佳實務時的曝險等級：高

實作指引

成本分類的程序對預算、會計、財務報告、決策制定、基準和專案管理至關重要。透過對費用進行分類，團隊可更加了解他們在整個雲端之旅中將產生的成本類型，從而做出明智的決策並有效管理預算。

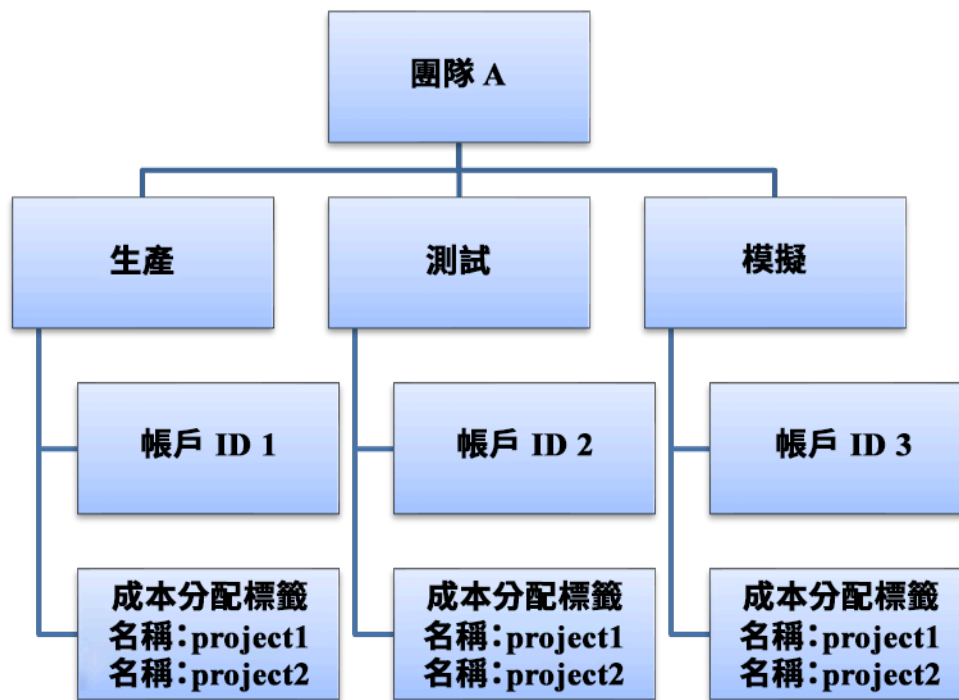
雲端支出權責劃分為有紀律的需求和成本管理建立了有力的誘因。對於將大部分雲端支出分配給取用業務單位或團隊的組織，這樣可以大幅節省雲端成本。此外，分配雲端支出有助於組織採用更多集中式雲端控管的最佳實務。

在定期會議中與財務團隊和其他相關利益相關者合作，了解在組織內分配成本的要求。工作負載成本必須在整個生命週期中分配，包括開發、測試、生產和停用。了解組織內學習、員工發展和創意成本的狀況。這有助於將用於此目的的帳戶正確分配到培訓和開發預算，而不是籠統的 IT 成本預算。

與組織中的利益相關者一起定義成本歸因類別之後，請在 AWS 雲端 中使用 [AWS Cost Categories](#) 將成本與用量資訊分組為有意義的類別，例如特定專案的成本或者部門或業務單位的 AWS 帳戶。您可以建立自訂類別，並使用各種不同的維度 (例如帳戶、標籤、服務或費用類型)，根據您定義的規則將成本與用量資訊對應至這些類別中。設定成本類別後，您就能依據這些類別檢視成本與用量資訊，進而讓組織能制定更好的策略與購買決策。這些類別也會顯示在 AWS Cost Explorer、AWS Budgets 和 AWS Cost and Usage Report 中。

例如，假設您為業務單位 (DevOps 團隊) 建立成本類別，並根據您所定義的群組，在每個類別下建立多個規則 (每個子類別的規則)，分別具有多個維度 (AWS 帳戶、成本分配標籤、服務或收費類型)。有了成本類別，即可使用以規則為基礎的引擎來整理成本。您設定的規則會將您的成本整理至各個類別。在這些規則中，您可以使用多個維度篩選每個類別，例如特定 AWS 帳戶、AWS 服務或費用類型。然後，您就可以在 [AWS 帳單與成本管理](#) 和 [成本管理主控台](#) 中使用多個產品中的這些類別。其中包括 AWS Cost Explorer、AWS Budgets、AWS Cost and Usage Report 和 AWS Cost Anomaly Detection。

例如，下圖顯示您可以有多個團隊 (成本類別)、多個環境 (規則)，且每個環境有多個資源或資產 (維度)，進而分組您組織中的成本與用量資訊。



成本與用量組織圖表

您也可以使用成本類別建立成本的群組。在您建立成本類別後 (您的用量記錄可在成本類別建立後的 24 小時內更新為新值)，這些類別會出現在 [AWS Cost Explorer](#)、[AWS Budgets](#)、[AWS Cost and Usage Report](#) 和 [AWS Cost Anomaly Detection](#) 中。在 AWS Cost Explorer 和 AWS Budgets 中，成本類別會顯示為額外的帳單維度。您可以使用該值來篩選特定的成本類別值，或依成本類別分組。

實作步驟

- 定義您的組織類別：與內部利益相關者和業務單位會談，定義可反映組織結構和要求的類別。這些類別應該直接對應至現有財務類別的結構，例如業務單位、預算、成本中心或部門。查看雲端服務為您的業務帶來的成果，例如培訓或教育，因為這些也是屬於組織類別。
- 定義您的功能類別：與內部利益相關者和業務單位會談，定義可反映您在企業內具有之職能的類別。這可能是工作負載或應用程式名稱，以及環境類型，例如生產、測試或開發。
- 定義 AWS Cost Categories：建立成本類別，使用 [AWS Cost Categories](#) 來組織成本與用量資訊，並將 AWS 成本與用量對應至 [有意義的類別](#) 中。您可以將多個類別指派給一個資源，而資源可以位於多個不同的類別中，因此請視需要定義任意數量的類別，以便可使用 AWS Cost Categories 在分類的結構中 [管理您的成本](#)。

資源

相關文件：

- [標記 AWS 資源](#)
- [使用成本分配標籤](#)
- [使用 AWS Budgets 分析您的成本](#)
- [使用 Cost Explorer 分析成本](#)
- [管理 AWS Cost and Usage Report](#)
- [AWS Cost Categories](#)
- [使用 AWS Cost Categories 管理成本](#)
- [建立成本類別](#)
- [標記成本類別](#)
- [在成本類別中拆分費用](#)
- [AWS Cost Categories 功能](#)

相關範例：

- [使用 AWS Cost Categories 組織您的成本與用量資料](#)
- [使用 AWS Cost Categories 管理成本](#)

COST03-BP04 建立組織指標

建立此工作負載所需的組織指標。工作負載的指標範例包括產生的客戶報告或向客戶提供的網頁。

未建立此最佳實務時的曝險等級：高

實作指引

了解工作負載的輸出是否算得上業務成功。每個工作負載通常都有少數幾個能夠指出效能的主要輸出。如果您有包含許多元件的複雜工作負載，則可以排定清單的優先順序，或定義和追蹤每個元件的指標。與您的團隊合作，以了解要使用哪些指標。此單位將用於了解工作負載的效率，或每個業務輸出的成本。

實作步驟

- 定義工作負載成果：與業務中的利益相關者會面，並定義工作負載的成果。這些是客戶用量的主要衡量方式，並且必須是業務指標而非技術指標。每個工作負載應該有少量的高層級指標 (少於五個)。如果工作負載為不同的使用案例產生多個成果，請將它們分組為單一指標。
- 定義工作負載元件成果：或者，如果您有大型且複雜的工作負載，或者可以用明確定義的輸入和輸出，輕鬆地將工作負載分成元件 (例如微型服務)，則請為每個元件定義指標。工作應該反映元件的價值和成本。從最大的元件開始，並向較小的元件運行。

資源

相關文件：

- [標記 AWS 資源](#)
- [使用 AWS Budgets 分析您的成本](#)
- [使用 Cost Explorer 分析成本](#)
- [管理 AWS 成本與用量報告](#)

COST03-BP05 設定帳單和成本管理工具

設定符合組織政策的成本管理工具，以管理及優化雲端支出。其中包括以服務、工具和資源來組織及追蹤成本與用量資料、透過整合的帳單和存取許可加強控制、透過預算制定與預測提升規劃效能、接收通知或提醒，以及藉由資源與定價優化降低成本。

未建立此最佳實務時的風險暴露等級：高

實作指引

為了建立健全的權責劃分，應先將帳戶策略視為成本分配策略的一部分。正確做到這一點，應該就夠了。否則，後續會發生意料之外和棘手的問題。

為了鼓勵雲端支出的權責劃分，使用者應有權存取可讓他們檢視成本與用量的工具。AWS 建議您基於下列目的設定所有工作負載和團隊：

- 組織：使用您自己的標記策略和分類法，來建立成本分配與管控基準。使用 AWS Control Tower 或 AWS Organization 等工具建立多個 AWS 帳戶。標記支援的 AWS 資源，並根據您的組織結構 (業務單位、部門或專案) 進行有意義的分類。標記特定成本中心的帳戶名稱，並與 AWS Cost Categories 對應，以將其成本中心的業務單位的帳戶分組，讓業務單位擁有者可在一個位置查看多個帳戶的耗用量。
- 存取：在合併帳單中追蹤整個組織的帳單資訊。確認適當的利益相關者和企業擁有者具有存取權。

- **控制：**使用適當防護機制建立有效控管機制，以避免在使用服務控制政策 (SCP)、標記政策、IAM 政策和預算警示時發生意外狀況。例如，您可以允許團隊只使用有效的控制機制在首選區域中建立特定資源，並防止在沒有特定標籤的情況下建立資源 (例如成本中心)。
- **目前狀態：**設定儀表板，顯示目前的成本和用量級別。儀表板應在工作環境中顯眼的位置提供，類似於營運儀表板。可以匯出資料，並使用 AWS 成本最佳化中心的成本和用量儀表板或任何支援產品來建立此可見性。您可能需要為不同的角色建立不同的儀表板。例如，管理員儀表板可能與工程儀表板不同。
- **通知：**使用「AWS 預算」或「AWS 成本異常偵測」，在成本或用量超過定義的限制並發生異常情況時提供通知。
- **報告：**彙總所有成本和用量資訊。利用詳細的可歸因成本資料，提高雲端支出的意識和責任。建立與使用這些報告的團隊相關且包含建議的報告。
- **追蹤：**顯示相對於設定的總目標或具體目標目前成本和用量的狀況。
- **分析：**允許團隊成員使用不同的篩選條件 (資源、帳戶、標籤等) 執行自訂和深度分析，精確到每小時、每日或每月。
- **檢查：**隨時掌握資源部署和成本最佳化商機的最新資訊。使用 Amazon CloudWatch、Amazon SNS 或 Amazon SES 接收通知，以便在組織層級進行資源部署。使用 AWS Trusted Advisor 或 AWS Compute Optimizer 審核成本最佳化建議。
- **趨勢報告：**以所需的精細度顯示所需期間內的成本與用量變化。
- **預測：**使用您建立的預測儀表板顯示預估的未來成本，以及預估您的資源用量和支出。

您可以使用 [AWS 成本最佳化中心](#)，了解從中央位置整合的潛在成本節省機會，並建立資料匯出以便與 Amazon Athena 整合。您也可以使用 AWS 成本最佳化中心來部署成本和用量儀表板，它利用 Quick 進行互動式成本分析和安全的成本洞察分享。

如果您的組織中沒有基本技能或頻寬，可以使用 [AWS ProServ](#)、[AWS Managed Services \(AMS\)](#) 或 [AWS 合作夥伴](#)。您也可以使用第三方工具，但請務必驗證價值主張。

實作步驟

- **允許以團隊為基礎的工具存取權：**設定您的帳戶並建立群組，以存取所需的成本和用量報告以供其使用，並使用 [AWS Identity and Access Management](#) 來控制諸如 AWS Cost Explorer 等工具的[存取權](#)。這些群組必須包含擁有或管理應用程式的所有團隊中的代表。這可證明每個團隊都能存取其成本和用量資訊以追蹤取用情形。
- **管理成本標籤和類別：**跨團隊、業務單位、應用程式、環境和專案來管理成本。使用資源標籤依成本分配標籤來管理成本。使用標籤、帳戶、服務等，依據維度建立成本類別以對應您的成本。

- 設定 AWS Budgets：針對您的工作負載在所有帳戶上[設定 AWS Budgets](#)。使用標籤和成本類別，設定整體帳戶支出的預算以及工作負載的預算。在 AWS Budgets 中設定通知，以接收超出預算金額時的提醒，或預估成本超出預算時的提醒。
- 設定 AWS 成本異常偵測：針對您的帳戶、核心服務或您所建立的成本類別使用 [AWS 成本異常偵測](#)，以監控成本與用量並偵測異常支出。您可以在彙總報告中個別接收提醒，也可以透過電子郵件或 Amazon SNS 主題接收提醒，以便分析和判斷發生異常的根本原因，並找出導致成本上升的因素。
- 使用成本分析工具：針對您的工作負載和帳戶來設定 [AWS Cost Explorer](#)，將成本資料視覺化以進行深入分析。根據歷史成本資料建立工作負載的儀表板，以追蹤整體支出、工作負載的關鍵用量指標，以及未來成本的預測。
- 使用成本節省分析工具：使用 AWS 成本最佳化中心，透過量身訂做的建議來識別節省機會，包括刪除未使用的資源、調整大小、Savings Plans、保留以及運算最佳化工具建議。
- 設定進階工具：您可以選擇性地建立視覺效果以促進交互式分析和成本洞見分享。透過 AWS 成本最佳化中心上的資料匯出，可以為您的組織建立由 Quick 提供支援的成本和用量儀表板，以提供更多詳細資訊和精細度。您也可以使用 [Amazon Athena](#) 中使用資料匯出實作進階分析功能，以進行進階查詢，並在 [Quick](#) 上建立儀表板。與 [AWS 合作夥伴](#) 進行合作，採用雲端管理解決方案，進行整合式雲端帳單監控與最佳化。

資源

相關文件：

- [什麼是 AWS 帳單與成本管理 和成本管理？](#)
- [建立您的最佳實務 AWS 環境](#)
- [標記 AWS 資源的最佳實務](#)
- [標記您的 AWS 資源](#)
- [AWS Cost Categories](#)
- [使用 AWS Budgets 分析成本](#)
- [使用 AWS Cost Explorer 分析您的成本](#)
- [什麼是 AWS 資料匯出？](#)

相關影片：

- [部署雲端智慧儀表板](#)
- [取得任何 FinOps 或成本優化指標或 KPI 的提醒](#)

相關範例：

- 由 Quick 提供支援的[成本和用量儀表板](#)
- [AWS 成本與用量管控研討會](#)

COST03-BP06 根據工作負載指標分配成本

依據用量指標或商業成果分配工作負載的成本，以衡量工作負載的成本效率。實作程序以透過分析服務 (可提供洞見和退款功能) 來分析成本和用量資料。

未建立此最佳實務時的曝險等級：低

實作指引

成本優化意味著以最低的價格提供業務成果，只有依工作負載指標 (按工作負載效率測量) 來分配工作負載成本才能達成。透過日誌檔案或其他應用程式監控，監控已定義的工作負載指標。結合此資料與工作負載成本，您可以透過查看具有特定標籤值或帳戶 ID 的成本來取得成本資料。每小時執行一次此分析。如果您擁有靜態成本元件 (例如，持續執行的後端資料庫) 且請求率不同 (例如，用量尖峰在早上九到晚上五點，晚上只有少量請求)，您的效率通常會有所改變。了解靜態成本與可變成本之間的關係，有助於您確定優化活動的重點。

相較於 Amazon Elastic Container Service (Amazon ECS) 和 Amazon API Gateway 上的容器化應用程式等資源，為共用資源建立工作負載指標可能具有挑戰性。但是，可以透過某些方法對使用情況進行分類並追蹤成本。如果您需要追蹤 Amazon ECS 和 AWS Batch 共用資源，則可以在 AWS Cost Explorer 中啟用分割成本分配資料。使用分割成本分配資料，您可以了解並優化容器化應用程式的成本和用量，並根據共用運算和記憶體資源的使用情形，將應用程式成本分配給個別業務實體。

實作步驟

- 將成本分配到工作負載指標：使用定義的指標和設定的標籤，建立結合工作負載輸出和工作負載成本的指標。使用諸如 Amazon Athena 和 Amazon Quick 等分析服務，為整體工作負載和任何元件建立效率儀表板。

資源

相關文件：

- [標記 AWS 資源](#)
- [使用 AWS Budgets 分析成本](#)

- [使用 Cost Explorer 分析成本](#)
- [管理 AWS 成本與用量報告](#)

相關範例：

- [透過 AWS 分割成本分配資料提升 Amazon ECS 和 AWS Batch 的成本可見性](#)

COST 4. 如何停用資源？

從啟動到結束專案期間，控制變更並管理資源。這樣做可確保您關閉或終止未使用的資源，以減少浪費。

最佳實務

- [COST04-BP01 追蹤其生命週期的資源](#)
- [COST04-BP02 實作停用程序](#)
- [COST04-BP03 停用資源](#)
- [COST04-BP04 自動停用資源](#)
- [COST04-BP05 強制執行資料保留政策](#)

COST04-BP01 追蹤其生命週期的資源

定義並實作一種方法，在資源的生命週期內追蹤資源及其與系統的關聯。您可以使用標記來識別資源的工作負載或功能。

未建立此最佳實務時的曝險等級：高

實作指引

停用不再需要的工作負載資源。常見的範例是用於測試的資源：測試完成後，便可移除資源。使用標籤來追蹤資源 (並針對這些標籤執行報告) 可協助您識別要停用的資產 (不會再使用這些資產，或是其授權將到期時)。使用標籤是追蹤資源的有效方法，方法是使用資源的功能標記資源，或標記停用日期。然後，即可對這些標籤執行報告。功能標記的範例值可以是 feature-X testing，可識別資源在工作負載生命週期的用途。另一個範例是TTL針對 資源使用 LifeSpan或，例如 to-be-deleted標籤金鑰名稱和值，以定義停用的時段或特定時間。

實作步驟

- 實作標記結構描述：實作識別資源所屬工作負載的標記結構描述，確保工作負載內的所有資源都已相應地加上標籤。標記可協助您依用途、團隊、環境或其他與您業務相關的準則，來將資源分類。有關標記使用案例、策略和技巧的更多詳細資訊，請參閱 [AWS 標記最佳實務](#)。
- 實作工作負載輸送量或輸出監控：實作工作負載輸送量監控或警示，在輸入請求或輸出完成時啟動。將其設定為在工作負載請求或輸出降至零時提供通知，指示不再使用工作負載資源。如果工作負載在正常條件下定期下降到零，則併入時間因素。如需有關未使用或未充分利用資源的詳細資訊，請參閱 [AWS Trusted Advisor 成本最佳化檢查](#)。
- 群組 AWS 資源：建立 AWS 資源的群組。您可以使用 [AWS Resource Groups](#) 來組織和管理位於相同中的 AWS 資源 AWS 區域。可以針對大多數的資源新增標籤，以便識別和排序組織內的資源。使用 [標籤編輯器](#) 將標籤大量新增至支援的資源。考慮使用 [AWS Service Catalog](#) 來建立、管理並向最終使用者分發批准的產品組合，並管理產品生命週期。

資源

相關文件：

- [AWS Auto Scaling](#)
- [AWS Trusted Advisor](#)
- [AWS Trusted Advisor 成本最佳化檢查](#)
- [標記 AWS 資源](#)
- [發佈自訂指標](#)

相關影片：

- [如何使用 最佳化成本 AWS Trusted Advisor](#)

相關範例：

- [組織 AWS 資源](#)
- [使用 最佳化成本 AWS Trusted Advisor](#)

COST04-BP02 實作停用程序

實作識別和停用未使用資源的程序。

未建立此最佳實務時的曝險等級：高

實作指引

在您的組織中實作標準化程序，以識別並移除未使用的資源。此程序應該要定義執行搜尋的頻率，以及移除資源的程序，以便驗證是否有符合組織的所有要求。

實作步驟

- 建立並實作停用程序：與工作負載開發人員和擁有者合作，為工作負載及其資源建置停用程序。此程序應該涵蓋用於驗證工作負載是否正在使用的方法，以及用於驗證每個工作負載資源是否正在使用的方法。詳述停用資源的必要步驟，將它們從服務中移除，同時確保符合任何的法規要求。應包含任何關聯的資源，例如授權或連接的儲存。發出通知讓工作負載擁有者知道停用程序已經執行。

使用下列停用步驟來引導您了解過程中應檢查的事項：

- 識別要停用的資源：識別在 AWS 雲端 中有資格停用的資源。記錄所有必要資訊，並排定停用時間。在規劃時間表時，請務必考慮到過程中可能會發生沒預期到的問題。
- 協調與溝通：與工作負載擁有者合作，確認要停用的資源
- 記錄中繼資料並建立備份：記錄中繼資料 (例如公有 IP、區域、AZ、VPC、子網路和安全群組)，並建立備份 (例如 Amazon Elastic Block Store 快照或擷取 AMI、金鑰匯出和憑證匯出)，如果生產環境中的資源需要，或者如果它們是關鍵資源。
- 驗證基礎設施即程式碼：判斷資源是使用 CloudFormation、Terraform、AWS Cloud Development Kit (AWS CDK) 或任何其他基礎設施即程式碼部署工具來部署，以便在必要時重新部署這些資源。
- 防止存取：在一段時間內套用限制性控制，以防止在判斷是否需要資源時使用資源。確認資源環境可在必要時恢復為原始狀態。
- 遵循您的內部停用程序：遵循組織的管理任務和停用程序，例如從組織網域中移除資源、移除 DNS 記錄，以及從組態管理工具、監控工具、自動化工具和安全性工具中移除資源。

如果資源是 Amazon EC2 執行個體，請參考下列清單。[如需詳細資訊，請參閱「如何刪除或終止 Amazon EC2 資源？」](#)

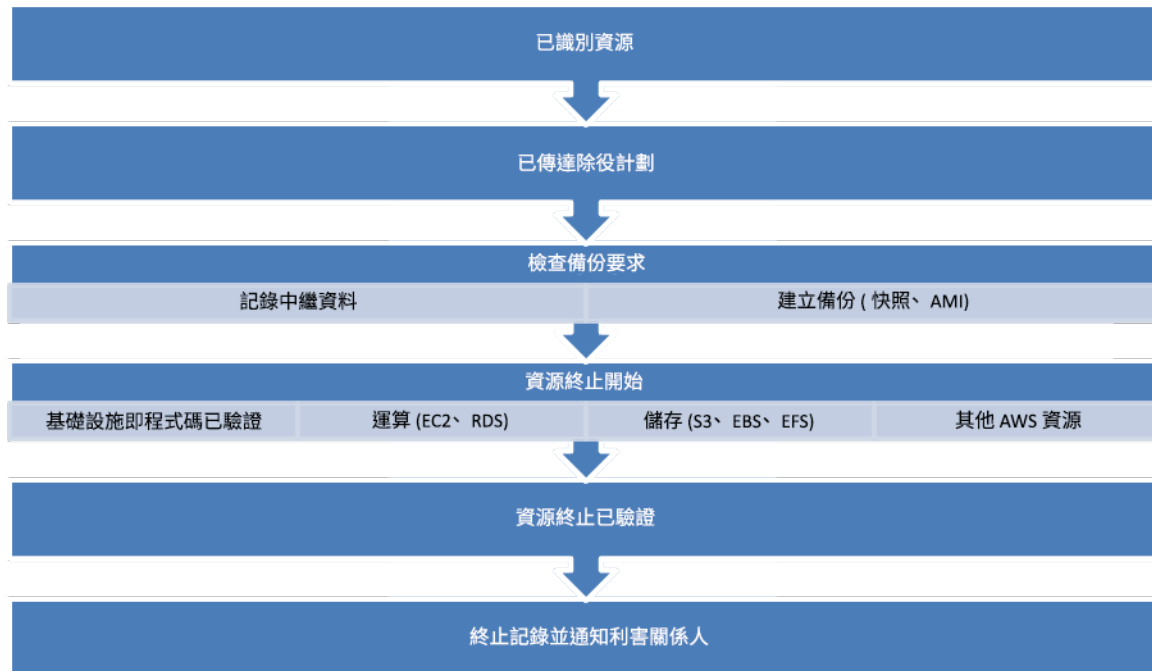
- 停止或終止所有 Amazon EC2 執行個體和負載平衡器。Amazon EC2 執行個體終止之後可短時間內在主控制台看到。您不需要為任何未處於執行中狀態的執行個體付費
- 刪除 Auto Scaling 基礎設施。
- 釋放所有專用執行個體。
- 刪除所有 Amazon EBS 磁碟區和 Amazon EBS 快照。
- 釋放所有彈性 IP 位址。

• 取消註冊所有 Amazon Machine Image (AMI)。

- 終止所有 AWS Elastic Beanstalk 環境。

如果資源是 Amazon Glacier 儲存中的物件，而且在封存未達最低儲存持續時間之前就將其刪除，則會按比例向您收取過早刪除費。Amazon Glacier 的最短儲存持續時間取決於所使用的儲存類別。如需每個儲存類別的最短儲存持續時間摘要，請參閱 [Amazon S3 儲存類別的效能](#)。如需有關如何計算提前刪除費用的詳細資訊，請參閱 [Amazon S3 定價](#)。

下面的簡單停用程序流程圖會概述停用步驟。在停用資源之前，請先確認您確定要停用的資源沒有被組織使用。



資源停用流程。

資源

相關文件：

- [AWS Auto Scaling](#)
- [AWS Trusted Advisor](#)
- [AWS CloudTrail](#)

相關影片：

- [刪除 CloudFormation 堆疊但保留一些資源](#)

- [找出哪個使用者啟動了 Amazon EC2 執行個體](#)

相關範例：

- [刪除或終止 Amazon EC2 資源](#)
- [找出哪個使用者啟動了 Amazon EC2 執行個體](#)

COST04-BP03 停用資源

停用由諸如定期稽核或用量變更等事件觸發的資源。通常會定期執行停用，其執行方式可以手動，也可以自動。

未建立此最佳實務時的曝險等級：中

實作指引

搜尋未使用資源的頻率和努力應該反映潛在節省的成本，因此較低成本的帳戶的分析頻率應該比較高成本帳戶低。搜尋和停用事件可由工作負載的狀態變更觸發，例如產品壽命結束或被取代。搜尋和停用事件也可由外部事件啟動，例如市場條件變化或產品終止。

實作步驟

- 停用資源：這是不再被需要或授權協議結束的 AWS 資源的折舊階段。請先完成所有最終檢查，再移至處置階段並停用資源，以防止發生任何不需要的中斷，例如擷取快照或備份。使用停用程序，停用已識別為未使用的每個資源。

資源

相關文件：

- [AWS Auto Scaling](#)
- [AWS Trusted Advisor](#)

COST04-BP04 自動停用資源

設計工作負載，在識別和停用非關鍵資源、不需要的資源或低利用率資源時，妥善處理資源終止。

未建立此最佳實務時的曝險等級：低

實作指引

使用自動化來降低或消除停用程序的相關成本。將工作負載設計為執行自動停用，可降低工作負載生命週期內的整體成本。可以使用 [Amazon EC2 Auto Scaling](#) 或 [Application Auto Scaling](#) 來執行停用程序。也可以使用 [API 或 SDK](#) 來實作自訂程式碼，以自動停用工作負載資源。

[現代化應用程式](#)以無伺服器為優先，這是一種優先採用無伺服器服務的策略。AWS 為堆疊的全部三個層級開發了[無伺服器服務](#)：運算、整合和資料存放區。使用無伺服器架構可讓您透過自動縱向擴展和縮減規模，在低流量期間節省成本。

實作步驟

- 實作 Amazon EC2 Auto Scaling 或 Application Auto Scaling：對於受支援的資源，請使用 Amazon EC2 Auto Scaling 或 Application Auto Scaling 進行設定。這些服務可協助您在使用 AWS 服務時優化使用率和成本效益。當需求下降時，這些服務會自動移除超額的資源容量，以免您超支。
- 設定 CloudWatch 以終止執行個體：可將執行個體設定為使用 [CloudWatch 警示](#)來終止執行個體。使用來自於停用程序的指標，透過 Amazon Elastic Compute Cloud 動作實作警示。推出之前，確認非生產環境中的操作。
- 在工作負載內實作程式碼：可以使用 AWS SDK 或 AWS CLI 來停用工作負載資源。在整合 AWS 的應用程式內實作程式碼，並終止或移除不再使用的資源。
- 使用無伺服器服務：優先在 AWS 上建置[無伺服器架構](#)和[事件驅動架構](#)，以建置並執行應用程式。AWS 提供多種無伺服器技術服務，本質上可提供自動最佳化的資源使用率和自動停用 (縮減和擴充)。在使用無伺服器應用程式時，系統會自動為您提供最佳化的資源使用率，您永遠不會因為過度佈建而支付費用。

資源

相關文件：

- [Amazon EC2 Auto Scaling](#)
- [Amazon EC2 Auto Scaling 入門](#)
- [Application Auto Scaling](#)
- [AWS Trusted Advisor](#)
- [AWS 上的無伺服器](#)
- [建立警示以停止、終止、重新啟動或復原執行個體](#)
- [將終止動作新增至 Amazon CloudWatch 警示](#)

相關範例：

- [排程自動刪除 AWS CloudFormation 堆疊](#)

COST04-BP05 強制執行資料保留政策

對支援的資源定義資料保留政策，以根據組織的要求處理物件刪除。識別並刪除不再需要的非必要或孤立資源與物件。

未建立此最佳實務時的曝險等級：中

使用資料保留政策和生命週期政策，降低已識別資源的停用程序相關成本和儲存成本。定義資料保留政策和生命週期政策以執行自動化儲存類別遷移和刪除，可降低生命週期內的整體儲存成本。您可以使用 Amazon Data Lifecycle Manager 自動建立和刪除 Amazon Elastic Block Store 快照與 Amazon EBS 支援的 Amazon Machine Image (AMI)，並且可使用 Amazon S3 Intelligent-Tiering 或 Amazon S3 生命週期組態來管理 Amazon S3 物件的生命週期。也可以使用 [API 或 SDK](#) 實作自訂程式碼，為要自動刪除的物件建立生命週期政策和政策規則。

實作步驟

- 使用 Amazon Data Lifecycle Manager：使用 Amazon Data Lifecycle Manager 上的生命週期政策來自動刪除 Amazon EBS 快照和 Amazon EBS 支援的 AMI。
- 在儲存貯體上設定生命週期組態：在儲存貯體上使用 Amazon S3 生命週期組態，定義 Amazon S3 在物件生命週期中採取的動作，以及根據您的業務需求在物件生命週期結束時進行刪除。

資源

相關文件：

- [AWS Trusted Advisor](#)
- [Amazon Data Lifecycle Manager](#)
- [如何在 Amazon S3 儲存貯體上設定生命週期組態](#)

相關影片：

- [使用 Amazon Data Lifecycle Manager 來自動化 Amazon EBS 快照](#)
- [使用生命週期組態規則來清空 Amazon S3 儲存貯體](#)

相關範例：

- [使用生命周期組態規則來清空 Amazon S3 儲存貯體](#)

具有經濟效益的資源

問題

- [COST 5. 如何在選取服務時評估成本？](#)
- [COST 6. 如何在選取資源類型、大小和數量時達成成本目標？](#)
- [COST 7. 如何使用定價模式降低成本？](#)
- [COST 8. 如何規劃資料傳輸費？](#)

COST 5. 如何在選取服務時評估成本？

Amazon EC2、Amazon EBS 和 Amazon S3 是基礎的 AWS 服務。Amazon RDS 和 Amazon DynamoDB 等受管服務為更高等級或應用程式等級的 AWS 服務。選取適當的基礎和受管服務，就可最佳化此工作負載的成本。舉例來說，您可以使用受管服務減少或省下大部分管理和營運開銷，讓您從事應用程式或企業相關活動。

最佳實務

- [COST05-BP01 識別組織的成本要求](#)
- [COST05-BP02 分析工作負載的所有元件](#)
- [COST05-BP03 對每個元件執行徹底的分析](#)
- [COST05-BP04 選取具成本效益授權的軟體](#)
- [COST05-BP05 選取此工作負載的元件，以按照組織優先事項來優化成本](#)
- [COST05-BP06 對不同用量執行一段時間內的成本分析](#)

COST05-BP01 識別組織的成本要求

與團隊成員一起為此工作負載定義成本最佳化與其他支柱 (例如效能和可靠性) 之間的平衡。

未建立此最佳實務時的曝險等級：高

實作指引

大多數組織的資訊技術 (IT) 部門會由多個小型團隊組成，每個團隊都有自己的議程和重點領域，而這會反映出其團隊成員的專業和技能。您需要了解組織的整體目標、優先順序、目標，以及每個部門或專案如何為這些目標做出貢獻。對於實現組織目標和全面預算規劃來說，將所有重要資源進行分類至關重要，這些資源包括人員、設備、技術、材料和外部服務。採用這種系統化方法來識別和了解成本，是為組織建立實際、可靠成本計畫的基礎。

為工作負載選取服務時，關鍵是了解組織的優先事項。在成本最佳化和其他 AWS Well-Architected Framework 支柱之間建立平衡，例如效能和可靠性。此流程應有系統且定期地進行，以反映組織目標、市場條件和營運動態的變化。完全成本優化的工作負載是最符合您組織需求的解決方案，不一定是成本最低的解決方案。與組織中的所有團隊 (例如產品、業務、技術和財務團隊) 會面以收集資訊。評估在相互衝突的利益或替代方法之間做出權衡的影響，以協助您在確認工作重點或選擇行動方案時做出明智的決定。

例如，新功能加速上市可能是成本優化所強調的重點，或您可為非關聯式資料選擇關聯式資料庫，以便更輕鬆地遷移系統，而非遷移至針對您的資料類型優化的資料庫並更新您的應用程式。

實作步驟

- 確定組織的成本要求與您組織的團隊成員開會，包括產品管理人員、應用程式擁有者、開發和營運團隊、管理層和財務部人員。排定此工作負載及其元件的 Well-Architected 支柱優先順序。輸出應為依序列出的支柱清單。您也可以為每個支柱新增加權，以指出相應支柱有多少個額外焦點，或兩個支柱之間的焦點有多相似。
- 解決技術債務並將其記錄在案：在工作負載檢閱期間，處理技術債務。記錄積存項目以在將來重新檢視工作負載，目標是重構或重新架構以將工作負載進一步最佳化。向其他利益相關者清楚傳達所做出的權衡至關重要。

資源

相關的最佳實務：

- [REL11-BP07 建構您的產品以滿足可用性目標和運作時間服務等級協議 \(SLAs\)](#)
- [OPS01-BP06 評估權衡](#)

相關文件：

- [AWS 總擁有成本 \(TCO\) 計算器](#)

- [Amazon S3 儲存類別](#)
- [雲端產品](#)

COST05-BP02 分析工作負載的所有元件

確認會分析每個工作負載元件，無論目前大小或目前成本為何。審查工作應反映潛在的效益，例如目前和預計的成本。

未建立此最佳實務時的曝險等級：高

實作指引

旨在為組織提供商業價值的工作負載元件可能包含各種服務。對於每個元件，可以選擇特定的 AWS 雲端服務來滿足業務需求。這個選擇可能會受到熟悉與否或之前使用這些服務的經驗等因素所影響。

在確定 [COST05-BP01 確定組織的成本要求](#) 中所述的組織要求之後，請對工作負載中的所有元件執行徹底分析。考慮當前和預測的成本與大小來分析每個元件。針對工作負載生命週期中的任何潛在工作負載節省來考慮分析成本。在分析此工作負載的所有元件上所花費的努力應與最佳化該特定元件所預期的潛在節省或改進相當。例如，如果所提議資源的成本是每月 10 美元，而低於預測的負載不會超過每月 15 美元，則努力一天以減少 50% 成本 (每月 5 美元) 可能會超過系統生命週期內的潛在利益。使用更快速且更有效率的資料型估算，會為此元件建立最佳整體結果。

工作負載可能會隨時間改變，而且如果工作負載架構或用量變化，適當的服務組合可能並非最佳。選擇服務的分析必須納入目前和未來的工作負載狀態以及用量水平。為未來的工作負載狀態或用量實作服務，可減少或消除未來變更所需的工作量，藉此降低整體成本。例如，使用 EMR Serverless 最初可能是合適的選擇。但是，隨著該服務的取用量增加，轉換到 EMR on EC2 可以降低工作負載中該元件的成本。

[AWS Cost Explorer](#) 和 AWS Cost and Usage Report ([CUR](#)) 可分析概念驗證 (PoC) 或執行環境的成本。也可以使用 [AWS 定價計算工具](#) 來估算工作負載成本。

撰寫工作流程，供技術團隊審核其工作負載。讓此工作流程保持簡單，同時也涵蓋所有必要步驟，以確保團隊了解工作負載的每個元件及其定價。然後，您的組織可以根據每個團隊的特定需求來遵循和自訂此工作流程。

1. 列出工作負載使用的每個服務：這是一個很好的起點。確定目前使用的所有服務以及成本來源。
2. 了解這些服務的定價方式：了解每項服務的 [定價模式](#)。根據用量、資料傳輸和特定功能定價等因素，不同的 AWS 服務會有不同的定價模式。

3. 專注於具有非預期工作負載成本且與預期用量和業務結果不符的服務：使用 AWS Cost Explorer 或 AWS Cost and Usage Report 識別成本與價值或用量不成比例的異常值或服務。將成本與業務成果相互關聯以優先考慮最佳化工作至關重要。
4. 使用 AWS Cost Explorer、CloudWatch Logs、VPC Flow Logs 和 Amazon S3 Storage Lens，了解這些高成本的根本原因：這些工具有助於高成本的診斷。每項服務都可提供不同的視角來檢視和分析使用情況和成本。例如，Cost Explorer 可協助判斷整體成本趨勢，CloudWatch Logs 可提供營運洞察，VPC Flow Logs 可顯示 IP 流量，而 Amazon S3 Storage Lens 則適用於儲存分析。
5. 使用 AWS Budgets 為服務或帳戶的某些金額設定預算：設定預算是管理成本的有效方式。使用 AWS Budgets 設定自訂預算閾值，並在成本超過這些閾值時接收提醒。
6. 設定 Amazon CloudWatch 警示以傳送帳單和用量提醒：設定成本和用量指標的監控和提醒。CloudWatch 警示可在超出特定閾值時通知您，從而縮短干預回應時間。

透過對所有工作負載元件進行策略審查 (無論其目前屬性為何)，可隨著時間的推移帶來顯著的改進和財務方面的節省。在這個審查流程中所投入的努力應經過深思熟慮，並仔細考慮可能實現的潛在優勢。

實作步驟

- 列出工作負載元件：建立工作負載元件清單。使用此清單可確認是否已分析每個元件。所做的工作應反映貴組織優先事項所定義之工作負載的關鍵性。按功能將資源分組在一起以提高效率 (例如，生產資料庫儲存 (若有多個資料庫的話))。
- 設定元件清單的優先順序：取得元件清單並按照工作順序排列其優先順序。這通常是依最昂貴到最便宜的元件成本排序，或依貴組織優先事項所定義的關鍵性排序。
- 執行分析：對於清單上的每個元件，審核可用的選項和服務並選擇最適合您組織優先事項的選項。

資源

相關文件：

- [AWS 定價計算工具](#)
- [AWS Cost Explorer](#)
- [Amazon S3 儲存類別](#)
- [AWS 雲端 產品](#)

相關影片：

- [AWS 成本最佳化系列：CloudWatch](#)

COST05-BP03 對每個元件執行徹底的分析

查看每個元件的組織整體成本。考量營運和管理成本以計算總體擁有成本，尤其是在使用雲端供應商的受管服務時。審查工作應反映潛在的效益 (例如，用於分析的時間與元件成本成正比)。

未建立此最佳實務時的曝險等級：高

實作指引

考量如何節省時間，讓您的團隊能夠專注於淘汰技術負債、創新和附加價值功能，以及創造企業與眾不同之處。例如，您可能需要將內部部署環境中的資料庫盡快「平移」至雲端 (也稱為主機轉換)，然後進行優化。能否使用 AWS 上的受管服務以去除或降低授權成本，進而獲得節省的效益，是值得探討的。AWS 上的受管服務免除了維護服務的營運和管理重擔 (例如修補或升級作業系統)，讓您得以專注於創新和業務。

因為受管服務以雲端規模運作，可使每次交易或服務的成本較低。您可以進行可能的優化以獲得實際的好處，且無須變更應用程式的核心架構。例如，您可能希望透過遷移到諸如 [Amazon Relational Database Service \(Amazon RDS\)](#) 等資料庫即服務平台，或將應用程式遷移到諸如 [AWS Elastic Beanstalk](#) 等全受管平台，以減少管理資料庫執行個體所花費的時間。

通常受管服務具有屬性，您可設定以確保備充足容量。您必須設定並監控這些屬性，使得額外的容量保持最低程度，並且獲得最大效能。您可使用 AWS 管理主控台 或 AWS API 和 SDK 來修改 AWS Managed Services 的屬性，使資源需求與持續變動的需求保持一致。例如，可將 Amazon EMR 叢集 (或 Amazon Redshift 叢集上) 節點的數量增加或減少，以便擴展或者縮減。

您也可將多個執行個體裝填到一項 AWS 資源上，啟動密度更高的使用。例如，可將多個小資料庫佈建至單一 Amazon Relational Database Service (Amazon RDS) 資料庫執行個體。隨著用量增長，可使用快照和恢復程序，將其中一個資料庫遷移至專用 Amazon RDS 資料庫執行個體。

將工作負載佈建至受管服務上時，您必須了解調整服務容量的要求。這些要求通常是時間、心力和對一般工作負載運作的影響。佈建的資源必須允許發生任何變更，佈建必要的額外開銷來實現。為了修改服務所需持續投注的心力，利用與系統和監控工具例如 Amazon CloudWatch 相整合的 API 和 SDK，可降低為幾乎是零。

[Amazon RDS](#)、[Amazon Redshift](#) 和 [Amazon ElastiCache](#) 可提供受管資料庫服務。[Amazon Athena](#)、[Amazon EMR](#) 和 [Amazon OpenSearch Service](#) 可提供受管分析服務。

[AMS](#) 是代表企業客戶和合作夥伴營運 AWS 基礎設施的服務。它提供安全且合規的環境，您可以將工作負載部署至其中。AMS 使用企業雲端營運模型與自動化，讓您符合組織需求、更快速地遷移至雲端，以及降低持續管理成本。

實作步驟

- 執行徹底的分析：使用元件清單，從最高優先順序到最低優先順序處理每個元件。對於優先順序更高且成本更高的元件，請執行額外的分析並評估所有可用選項及其長期影響。對於優先順序較低的元件，評估用量的變更是否會變更元件的優先順序，然後執行適當的工作分析。
- 比較受管和非受管資源：考慮您管理的資源的營運成本，並將其與 AWS 受管資源進行比較。例如，審查在 Amazon EC2 執行個體上執行的資料庫，並且與 Amazon RDS 選項 (AWS 受管服務) 比較，或將 Amazon EMR 相較於在 Amazon EC2 上執行 Apache Spark。從自我管理工作負載移轉至 AWS 全受管工作負載時，請仔細研究您的選項。要考慮的三個最重要的因素是您要使用的[受管服務類型](#)、將用來[遷移資料](#)的程序，以及了解[AWS 共同責任模型](#)。

資源

相關文件：

- [AWS 總體擁有成本 \(TCO\) 計算器](#)
- [Amazon S3 儲存類別](#)
- [AWS 雲端 產品](#)
- [AWS 共同責任模型](#)

相關影片：

- [為什麼要移至受管資料庫？](#)
- [什麼是 Amazon EMR 以及我該如何使用它來處理資料？](#)

相關範例：

- [為什麼要移至受管資料庫](#)
- [使用 AWS DMS 將來自相同 SQL Server 資料庫的資料整合到單一的 Amazon RDS for SQL Server 資料庫](#)
- [將資料大規模交付到 Amazon Managed Streaming for Apache Kafka \(Amazon MSK\)](#)
- [將 ASP.NET Web 應用程式遷移至 AWS Elastic Beanstalk](#)

COST05-BP04 選取具成本效益授權的軟體

開放原始碼軟體會剔除對工作負載增加大量成本的軟體授權費用。請在需要授權軟體時，避免繫結至任意屬性 (例如 CPU) 的授權，尋找繫結至輸出或成果的授權。這些授權的成本會更接近其提供的效益。

未建立此最佳實務時的曝險等級：低

實作指引

開放原始碼源於軟體開發的背景，以指出該軟體符合某些免費發行條件。開放原始碼軟體會由任何人都可以檢查、修改和增強的原始程式碼組成。根據業務要求、工程師的技能、預測用量或其他技術相依性，組織可以考慮使用 AWS 上的開放原始碼軟體，以最大程度地降低其授權成本。換句話說，使用[開放原始碼軟體](#)可降低軟體授權的成本。隨著工作負載的大小擴展，這可能會對工作負載成本產生重大影響。

請根據總成本來測量授權軟體的效益，以將工作負載最佳化。模擬授權的任何變更以及這些變更對工作負載成本的影響。如果廠商變更資料庫授權的成本，調查這會如何影響工作負載的整體效率。考慮廠商的歷史定價公告，以了解其產品授權變更趨勢。授權成本也可能獨立於輸送量或用量，例如依硬體擴展的授權 (CPU 綁定授權)。應該避免這些授權，因為成本可能會快速增加，而不會帶來相應結果。

例如，相較於執行另一個在 Windows 上執行的 Amazon EC2 執行個體，使用 Linux 作業系統在 us-east-1 中操作 Amazon EC2 執行個體可讓您削減大約 45% 的成本。

[AWS 定價計算工具](#) 提供了一種全面的方法來比較具有不同授權選項的各種資源的成本，例如 Amazon RDS 執行個體和不同的資料庫引擎。此外，AWS Cost Explorer 還為現有工作負載的成本提供了寶貴的觀點，尤其是具有不同授權的工作負載的成本。對於許可證管理，[AWS License Manager](#) 提供一種簡化的方法來監督和處理軟體授權。客戶可以在 AWS 雲端 中部署和操作自己喜歡的開放原始碼軟體。

實作步驟

- 分析授權選項：審核可用軟體的授權條款。尋找具有所需功能的開放原始碼版本，以及授權軟體的效益是否超過成本。有利條款會使軟體成本符合其提供的效益。
- 分析軟體供應商：審核來自於廠商的任何歷史定價或授權變更。尋找不符合成果的任何變更，例如，在特定廠商硬體或平台上執行的懲罰性條款。此外，尋找他們執行稽核和可能施加的懲罰的方式。

資源

相關文件：

- [AWS 的開放原始碼](#)
- [AWS 總體擁有成本 \(TCO\) 計算器](#)
- [Amazon S3 儲存類別](#)

- [雲端產品](#)

相關範例：

- [開放原始碼部落格](#)
- [AWS 開放原始碼部落格](#)
- [最佳化和授權評定](#)

COST05-BP05 選取此工作負載的元件，以按照組織優先事項來優化成本

選取工作負載的所有元件時均應考量成本。這包括使用應用程式層級和受管服務或無伺服器、容器或事件驅動架構，以降低整體成本。使用開放原始碼軟體、無需授權費用的軟體或替代方案，藉以將授權成本降至最低。

未建立此最佳實務時的曝險等級：中

實作指引

選取所有元件時均應考量服務和選項的成本。這包括使用應用程式層級和受管服務，例如 [Amazon Relational Database Service](#) (Amazon RDS)、[Amazon DynamoDB](#)、[Amazon Simple Notification Service](#) (Amazon SNS) 以及 [Amazon Simple Email Service](#) (Amazon SES)，以降低整體組織成本。

使用無伺服器和容器進行運算，例如 [AWS Lambda](#) 及針對靜態網站的 [Amazon Simple Storage Service](#) (Amazon S3)。如果可能，將您的應用程式容器化，並使用 AWS 受管容器服務，例如 [Amazon Elastic Container Service](#) (Amazon ECS) 或 [Amazon Elastic Kubernetes Service](#) (Amazon EKS)。

使用開放原始碼軟體或沒有授權費用的軟體，將授權成本降到最低 (例如，用於運算工作負載的 Amazon Linux，或將資料庫遷移到 Amazon Aurora)。

您可以使用無伺服器或應用程式層級服務，例如 [Lambda](#)、[Amazon Simple Queue Service \(Amazon SQS\)](#)、[Amazon SNS](#) 以及 [Amazon SES](#)。這些服務讓您無須管理資源，並提供程式碼執行、佇列服務和訊息傳遞功能。另一個好處是，這些服務可隨用量擴展效能和成本，因此能夠有效率地分配成本和劃分歸屬。

無伺服器服務也可以使用[事件驅動型架構](#)。事件驅動型架構是推送架構，因此一切都會在事件呈現於路由器時隨需進行。如此，您就無須付費持續進行輪詢以檢查事件。這表示網路頻寬耗用量、CPU 使用率、閒置機群容量和 SSL/TLS 交握都可降低。

如需有關無伺服器架構的詳細資訊，請參閱 [Well-Architected 無伺服器應用程式聚焦白皮書](#)。

實作步驟

- 選取每個服務以最佳化成本：使用您的優先順序清單和分析，選取最符合您組織優先事項的每個選項。與其增加容量以符合需求，您應考慮使用其他選項，以較低的成本獲得更好的效能。例如，您應審查資料庫在 AWS 上的預期流量，並考慮增加執行個體大小，或使用 Amazon ElastiCache 服務 (Redis 或 Memcached) 為資料庫提供快取的機制。
- 評估事件驅動型架構：使用無伺服器架構也可讓您為分散式微型服務應用程式建置事件驅動架構，以利設計可擴展、彈性、敏捷且符合成本效益的解決方案。

資源

相關文件：

- [AWS 總體擁有成本 \(TCO\) 計算器](#)
- [AWS Serverless](#)
- [什麼是事件驅動型架構](#)
- [Amazon S3 儲存類別](#)
- [雲端產品](#)
- [Amazon ElastiCache \(Redis OSS\)](#)

相關範例：

- [事件驅動型架構入門](#)
- [事件驅動型架構](#)
- [Statsig 如何使用 Amazon ElastiCache \(Redis OSS\) 以 100 倍的成本效益運行](#)
- [使用 AWS Lambda 函數的最佳實務](#)

COST05-BP06 對不同用量執行一段時間內的成本分析

工作負載可能隨時間變更。某些服務或功能在不同的用量層級上更具成本效益。按預計用量對每個元件執行一段時間內的分析，讓工作負載在其生命週期內保持成本效益。

未建立此最佳實務時的曝險等級：中

實作指引

隨著 AWS 發佈新的服務和功能，工作負載的最佳服務可能會改變。所需的努力應與潛在效益相符。工作負載審核頻率取決於您的組織需求。如果成本高昂，則更快實作新的服務可節省最多成本，因此更頻繁的審核是有利的。另一個需要審核的方面是使用模式的變更。用量的重大變更可能表示替代服務更理想。

如果需要將資料移至 AWS 雲端中，您可以選取 AWS 所提供的各種服務以及合作夥伴工具，以便遷移您的資料集，無論是檔案、資料庫、機器映像、區塊磁碟區甚或磁帶備份均可。例如，若要對 AWS 移入或移出大量資料，或是在邊緣處理資料，您可以使用其中一項 AWS 專用裝置，以符合成本效益的方式離線移動數以 PB 計的資料。另一個範例是，在資料傳輸速率較高時，直接連線服務可能會比 VPN 更便宜，為您的企業提供所需的連線能力。

根據對不同用量在一段時間內的成本分析，審查您的擴展活動。分析結果，確認是否可以調整擴展政策，以使用多個執行個體類型和購買選項新增執行個體。審查您的設定，確認是否可以降低最小值，以較小的機群大小處理使用者要求，以及新增更多資源以符合預期的高需求。

透過與組織中的利益相關者討論，針對不同使用情況執行成本分析，並使用 [AWS Cost Explorer](#) 的預測功能來預測服務變更的潛在影響。使用 AWS Budgets、CloudWatch 帳單警示和 AWS Cost Anomaly Detection 來監控用量等級發佈，以快速識別及實作最符合成本效益的服務。

實作步驟

- 定義預測使用模式：與您的組織 (例如行銷和產品擁有者) 合作，記錄工作負載的預期和預測使用模式。與利益相關者討論關於歷史和預測成本與用量增加的議題，並確定這類增加符合業務要求。識別您預期會有較多使用者使用 AWS 資源的日曆日、週或月，這表示您應增加現有資源的容量或採用其他服務，以降低成本並提升效能。
- 根據預測用量執行成本分析：使用定義的使用模式，在上述每個點執行分析。分析工作應反映潛在成果。例如，如果用量變化很大，則應執行徹底的分析以驗證任何成本和變化。換句話說，當成本增加時，企業的用量也應增加。

資源

相關文件：

- [AWS 總體擁有成本 \(TCO\) 計算器](#)
- [Amazon S3 儲存類別](#)
- [雲端產品](#)
- [Amazon EC2 Auto Scaling](#)

- [雲端資料遷移](#)
- [AWS Snow Family](#)

相關影片：

- [AWS OpsHub for Snow Family](#)

COST 6. 如何在選取資源類型、大小和數量時達成成本目標？

確認您為手邊的任務選取適當的資源大小和數量。選取最具成本效益的類型、大小和數量，就能盡量減少浪費。

最佳實務

- [COST06-BP01 執行成本建模](#)
- [COST06-BP02 根據資料選取資源類型、大小及數目](#)
- [COST06-BP03 根據指標自動選取資源類型、大小和數目](#)
- [COST06-BP04 考慮使用共用資源](#)

COST06-BP01 執行成本建模

識別組織要求 (例如業務需求和現有承諾)，並對工作負載及其每個元件執行成本建模 (整體成本)。在不同預測負載下對工作負載執行基準測試活動，並比較成本。建模工作應反映潛在效益。例如，花費的時間與元件成本成正比。

未建立此最佳實務時的曝險等級：高

實作指引

為您的工作負載及其每個元件執行成本建模，以了解資源之間的平衡，並根據特定效能等級，找出工作負載中每個資源的合適大小。了解成本考量，可在評估計劃性工作負載部署的價值實現成果時，傳達組織的商業案例和決策程序。

在不同預測負載下對工作負載執行基準測試活動，並比較成本。建模工作應反映潛在效益；例如，花費的時間與元件成本或預測的節省成正比。如需最佳實務，請參閱 [AWS Well-Architected 架構的效能達成效率支柱白皮書中的「審查」一節](#)。

例如，若要為包含運算資源的工作負載建立成本建模，[AWS Compute Optimizer](#) 可協助執行中工作負載的成本建模。它根據歷史用量，提供運算資源的合適大小建議。請確定 CloudWatch Agent 已部署至

Amazon EC2 執行個體以收集記憶體指標，可在 AWS Compute Optimizer 內為您提供更精確的建議。這是運算資源的理想資料來源，因為它是免費服務，並使用機器學習根據風險等級提出多個建議。

您可以將[多種服務](#)搭配自訂日誌作為資料來源使用，以最適化調整其他服務和工作負載元件 (例如 [AWS Trusted Advisor](#)、[Amazon CloudWatch](#) 和 [Amazon CloudWatch Logs](#)) 的操作。AWS Trusted Advisor 會檢查資源並標記低使用率資源，以協助您最適化調整資源並建立成本建模。

以下是成本建模資料和指標的建議：

- 監控必須精確反映使用者體驗。為時段選擇正確的精細度，並悉心選擇最大或 99%，而非平均值。
- 為分析的時段選擇涵蓋任何工作負載週期所需的正確精細度。例如，假設所執行的是為期兩週的分析，您可能會忽略高利用率的每月週期，導致佈建不足。
- 考量您現有的承諾、為其他工作負載選取的定價模式，以及加速創新和專注於核心業務價值的能力，藉此為您的計劃性工作負載選擇正確的 AWS 服務。

實作步驟

- 針對資源執行成本建模：將工作負載或概念驗證部署到具有要測試之特定資源類型和大小的獨立帳戶。使用測試資料執行工作負載，並記錄輸出結果以及測試執行時的成本資料。然後，重新部署工作負載或變更資源類型和大小，並再次執行測試。納入可能用於這些資源之任何產品的授權費用，以及在建立成本模型時部署和管理這些資源的預估營運 (勞工或工程師) 成本。考慮建立一段時間 (每小時、每日、每月、每月或三年) 的成本模型。

資源

相關文件：

- [AWS Auto Scaling](#)
- [找出機會進行適當調整](#)
- [Amazon CloudWatch 功能](#)
- [成本最佳化：Amazon EC2 調整大小](#)
- [AWS Compute Optimizer](#)
- [AWS 價格計算器](#)

相關範例：

- [執行資料驅動的成本建模](#)

- [預估計劃的 AWS 資源組態的成本](#)
- [選擇正確的 AWS 工具](#)

COST06-BP02 根據資料選取資源類型、大小及數目

根據有關工作負載和資源特性的資料來選擇資源大小或類型。例如，運算、記憶體、輸送量或寫入密集。通常使用工作負載的先前 (內部部署) 版本、文件或其他有關工作負載的資訊來源來進行此選擇。

未建立此最佳實務時的曝險等級：中

實作指引

Amazon EC2 提供各種執行個體類型，其各自具有不同等級的 CPU、記憶體、儲存和聯網容量，適合不同的使用案例。這些執行個體類型具有 CPU、記憶體、儲存和聯網功能的不同組合，可讓您在選取適合專案的資源組合時獲得多樣選擇。每個執行個體類型都有多種大小，因此您可以根據工作負載的需求調整資源。若要判斷您需要的執行個體類型，請收集有關您計劃在執行個體上執行之應用程式或軟體系統要求的詳細資訊。這些詳細資訊應包括以下內容：

- 作業系統
- CPU 核心數量
- GPU 核心
- 系統記憶體 (RAM) 數量
- 儲存類型和空間
- 網路頻寬要求

確定運算要求的目的以及需要的執行個體，然後探索各種 Amazon EC2 執行個體系列。Amazon 提供下列執行個體類型系列：

- 一般用途
- 運算最佳化
- 記憶體最佳化
- 儲存優化
- 加速運算
- HPC 最佳化

如需深入了解特定 Amazon EC2 執行個體系列可實現的特定目的和使用案例，請參閱 [AWS 執行個體類型](#)。

收集系統要求對於您選取最適合需求的特定執行個體系列和執行個體類型來說非常重要。執行個體類型的名稱由系列名稱和執行個體大小組成。例如，t2.micro 執行個體來自 T2 系列，並且是微型大小。

根據工作負載和資源特性選擇資源大小或類型 (例如，運算、記憶體、輸送量或寫入密集)。通常使用成本建模、工作負載的先前版本 (例如內部部署版本)、文件或其他有關工作負載的資訊來源 (白皮書或已發佈的解決方案) 來進行此選擇。使用 AWS 定價計算器或成本管理工具可協助您對執行個體類型、大小和組態做出明智的決策。

實作步驟

- 根據資料選取資源：使用成本建模資料來選取預期的工作負載用量層級，然後選擇指定的資源類型和大小。依據成本建模資料，決定虛擬 CPU 數目、總記憶體 (GiB)、本機執行個體儲存體磁碟區 (GB)、Amazon EBS 磁碟區和網路效能等級，並將執行個體所需的資料傳輸速率納入考量。一律根據詳細分析和準確的資料進行選取，以最佳化效能，同時有效地管理成本。

資源

相關文件：

- [AWS 執行個體類型](#)
- [AWS Auto Scaling](#)
- [Amazon CloudWatch 功能](#)
- [成本優化：EC2 調整大小](#)

相關影片：

- [為您的工作負載選擇合適的 Amazon EC2 執行個體](#)
- [調整您的服務](#)

相關範例：

- [探索和比較 Amazon EC2 執行個體類型變得更加容易](#)

COST06-BP03 根據指標自動選取資源類型、大小和數目

使用目前執行的工作負載中的指標來選擇正確的大小和類型，以優化成本。為運算、儲存、資料和聯網服務適當地佈建輸送量、大小和儲存。這可透過回饋迴圈 (例如自動調整規模) 或工作負載中的自訂程式碼來完成。

未建立此最佳實務時的曝險等級：低

實作指引

在工作負載中建立意見回饋迴圈，使用執行中工作負載的作用中指標來變更該工作負載。可以使用受管服務，例如 [AWS Auto Scaling](#)，您可以設定這項服務，以便執行精簡化操作。AWS 還提供 [API](#)、[SDK](#) 和功能，讓您以最小的努力修改資源。您可以設定工作負載來停止和啟動 Amazon EC2 執行個體，以允許變更執行個體大小或執行個體類型。這不僅帶來精簡化的效益，同時消除變更所需的幾乎所有營運成本。

有些 AWS 服務內建自動類型或大小選擇，例如 [Amazon Simple Storage Service Intelligent-Tiering](#)。Amazon S3 Intelligent-Tiering 會根據您的使用模式，自動在兩個存取層 (經常存取和不常存取) 之間移動您的資料。

實作步驟

- 透過設定工作負載指標來提高您的可觀測性：擷取工作負載的關鍵指標。這些指標提供客戶體驗 (例如工作負載輸出) 的指示，並符合資源類型和大小 (例如 CPU 和記憶體用量) 之間的差異。針對運算資源，請分析效能資料以將 Amazon EC2 執行個體調整到適當大小。識別閒置的執行個體，以及未充分使用的執行個體。要查找的關鍵指標是 CPU 使用率和記憶體利用率 (例如，90% 的時間為 40% 的 CPU 利用率，如[使用 AWS Compute Optimizer 正確調整大小和啟用記憶體使用率](#)中所述)。識別在四週期間內，CPU 使用率達到最大且記憶體使用率小於 40% 的執行個體。這些便是需要適當調整大小以降低成本的執行個體。對於 Amazon S3 等儲存資源，您可以使用 [Amazon S3 Storage Lens](#)，預設情況下，您可以在儲存貯體層級查看各種類別的 28 個指標，以及在儀表板中查看 14 天的歷史資料。您可以依摘要和成本最佳化或事件來篩選 Amazon S3 Storage Lens 儀表板，以分析特定指標。
- 檢視適當調整大小的建議：使用 AWS Compute Optimizer 中的適當調整大小的建議和成本管理主控台內的 Amazon EC2 適當調整大小工具，或審核 AWS Trusted Advisor 適當調整資源大小以對工作負載進行調整。不論是 Amazon EC2 執行個體、AWS 儲存類別還是 Amazon RDS 執行個體類型，在適當調整不同資源的大小時，請務必使用[正確的工具](#)，並遵循[適當調整大小準則](#)。針對儲存資源，您可以使用 Amazon S3 Storage Lens，以便能夠檢視物件儲存用量、活動趨勢並提出可行建議，以將成本最佳化並套用資料保護最佳實務。使用 [Amazon S3 Storage Lens](#) 從整個組織的指標分析衍生出來的情境式建議，您可以立即採取步驟來最佳化儲存。

- 根據指標自動選取資源類型和大小：使用工作負載指標，手動或自動選取工作負載資源。針對運算資源，在應用程式內設定 AWS Auto Scaling 或實作程式碼，可在需要頻繁變更時減少所需的工作量，而且它可能比手動程序更快地實作變更。您可以在單一 Auto Scaling 群組內啟動和自動擴展隨需執行個體和 Spot 執行個體組成的機群。除了獲得使用 Spot 執行個體的折扣之外，您還可以使用預留執行個體或 Savings Plan，以獲得定期隨需執行個體定價的折扣費率。這些因素合在一起可協助您將 Amazon EC2 執行個體所能節省的成本最佳化，並確定應用程式所需的規模和效能。也可以在 [Auto Scaling 群組 \(ASG\)](#) 中使用 [屬性型執行個體類型選取 \(ABS\)](#) 策略，以透過一組屬性 (例如 vCPU、記憶體和儲存) 來表達您的執行個體要求。您可以自動使用新發行的較新一代執行個體類型，並使用 Amazon EC2 Spot 執行個體來存取更大範圍的容量。Amazon EC2 Fleet 和 Amazon EC2 Auto Scaling 會選取和啟動符合指定屬性的執行個體，您不必再手動挑選執行個體類型。對於儲存資源，您可以使用 [Amazon S3 Intelligent Tiering](#) 和 [Amazon EFS Infrequent Access](#) 功能，這些功能可讓您自動選取儲存類別，在資料存取模式變更時自動節省儲存成本，而不會影響效能或營運開銷。

資源

相關文件：

- [AWS Auto Scaling](#)
- [AWS 適當調整大小](#)
- [AWS Compute Optimizer](#)
- [Amazon CloudWatch 功能](#)
- [CloudWatch 開始設定](#)
- [CloudWatch 發佈自訂指標](#)
- [Amazon EC2 Auto Scaling 入門](#)
- [Amazon S3 Storage Lens](#)
- [Amazon S3 Intelligent-Tiering](#)
- [Amazon EFS Infrequent Access](#)
- [使用 SDK 來啟動 Amazon EC2 執行個體](#)

相關影片：

- [適當調整服務的大小](#)

相關範例：

- [Amazon EC2 Fleet 的 Auto Scaling 屬性型執行個體類型選取](#)
- [使用已排程的擴展，針對成本最佳化 Amazon Elastic Container Service](#)
- [Amazon EC2 Auto Scaling 的預測擴展](#)
- [使用 Amazon S3 Storage Lens 將成本最佳化並了解使用情況](#)

COST06-BP04 考慮使用共用資源

對於已在組織層級針對多個業務單位部署的服務，請考慮使用共用資源來提高使用率並降低總體擁有成本 (TCO)。使用共用資源可能是一個具成本效益的選項，可透過使用現有解決方案、共用元件或兩者來集中化管理和成本。在帳戶界限內或專用帳戶中管理常用功能，例如監控、備份和連線。還可以透過實作標準化、減少重複及降低複雜性來降低成本。

未建立此最佳實務時的曝險等級：中

實作指引

當多個工作負載導致相同的功能時，請使用現有的解決方案和共用元件來改善管理並最佳化成本。請考慮使用現有資源 (尤其是共用資源)，例如非生產資料庫伺服器或目錄服務，透過遵循安全性最佳實務和組織法規來降低雲端成本。為了實現最佳價值和效率，將成本 (使用回報 (showback) 和計費 (chargeback)) 分配到推動消費的相關業務領域至關重要。

回報 (showback) 是指將雲端成本分解為可歸因類別的報告，例如消費者、業務單位、總賬帳戶或其他負責實體。回報的目標是向團隊、業務單位或個人展示其所用雲端資源的成本。

計費 (chargeback) 是指根據適合特定財務管理程序的策略，將中央服務支出分配給成本單位。對於客戶而言，計費 (chargeback) 會將一個共用服務帳戶產生的成本計入適合客戶報告流程的不同財務成本類別。透過建立計費機制，可以報告不同業務單位、產品和團隊所產生的成本。

工作負載可以分類為關鍵和非關鍵。根據此分類，對於較不重要的工作負載，使用具有一般組態的共用資源。為了進一步最佳化成本，請僅為關鍵工作負載預留專用伺服器。共用資源或在多個帳戶之間佈建資源，以便有效地管理它們。即使在不同的開發、測試和生產環境中，安全共用也是可行的，而且不會影響組織結構。

為了提高您對容器化應用程式的了解並最佳化其成本和用量，請使用分割成本分配資料，它可幫助您根據應用程式使用共用運算和記憶體的方式，將成本分配給個別業務實體。分割成本分配資料可協助您在 Amazon Elastic Container Service (Amazon ECS) 或 Amazon Elastic Kubernetes Service (Amazon EKS) 上執行的容器工作負載中實現任務層級的回報和計費。

對於分散式架構，請建立共用服務 VPC，以集中存取每個 VPC 中工作負載所需的共用服務。這些共用服務可以包含目錄服務或 VPC 端點等資源。為了減少管理開銷和成本，請從中央位置共用資源，而不是在每個 VPC 中建置資源。

當您使用共用資源時，可以節省營運成本、最大化資源利用率並提高一致性。在多帳戶設計中，可以集中託管某些 AWS 服務，並在一個中心位置使用多個應用程式和帳戶來存取它們，以節省成本。您可以使用 [AWS Resource Access Manager \(AWS RAM\)](#) 共用其他常用資源，例如 [VPC 子網路](#)和 [AWS Transit Gateway](#) 附件、[AWS Network Firewall](#) 或 [Amazon SageMaker AI 管道](#)。在多帳戶環境中，使用 AWS RAM 建立一次資源並與其他帳戶共用。

組織應有效地標記共用成本，並確認大部分成本已標記或分配。如果未有效地分配共用成本，而且沒有人負責共用成本管理，則共用雲端成本可能會螺旋式上升。您應該知道在資源、工作負載、團隊或組織層級產生了哪些成本，因為相較於達成的業務成果，這項知識可增強您對適用層級所提供的價值的了解。最終，組織可以從共用雲端基礎設施的成本節約中獲益。鼓勵共用雲端資源的成本分配，以最佳化雲端支出。

實作步驟

- 評估現有資源：審核針對工作負載使用類似服務的現有工作負載。視工作負載的元件而定，如果業務邏輯或技術需求允許，請考慮現有平台。
- 在 AWS RAM 中使用資源共用並進行相應限制：使用 AWS RAM 與組織內的其他 AWS 帳號共用資源。共用資源時，無需在多個帳戶中重複資源，這樣可將資源維護的作業負擔降到最低。此流程也可協助您安全地與帳戶中的角色和使用者以及其他 AWS 帳戶共用您所建立的資源。
- 標記資源：標記屬於成本報告的候選資源，並在成本分類中將其分類。啟動這些與成本相關的資源標籤以進行成本分配，從而了解 AWS 資源使用情況。專注於在成本和用量可見性方面建立適當的精細度級別，並透過成本分配報告和 KPI 追蹤來影響雲端消費行為。

資源

相關的最佳實務：

- [SEC03-BP08 在組織內安全地共用資源](#)

相關文件：

- [什麼是 AWS Resource Access Manager？](#)
- [可與 AWS Organizations 搭配使用的 AWS 服務](#)
- [可共用的 AWS 資源](#)

- [AWS 成本和用量 \(CUR\) 查詢](#)

相關影片：

- [AWS Resource Access Manager - 具有受管權限的精細存取控制](#)
- [如何設計 AWS 成本分配策略](#)
- [AWS Cost Categories](#)

相關範例：

- [如何計費共用服務：AWS Transit Gateway 範例](#)
- [如何使用 CUR 為 Savings Plans 建立計費/回報模式](#)
- [使用 VPC 共用建立具成本效益的多帳戶微服務架構](#)
- [透過 AWS 分割成本分配資料提升 Amazon EKS 的成本可見性](#)
- [透過 AWS 分割成本分配資料提升 Amazon ECS 和 AWS Batch 的成本可見性](#)

COST 7. 如何使用定價模式降低成本？

使用最適合您資源的定價模式，就能盡量減少支出。

最佳實務

- [COST07-BP01 執行定價模式分析](#)
- [COST07-BP02 根據成本選擇區域](#)
- [COST07-BP03 選取具成本效益條款的第三方協議](#)
- [COST07-BP04 針對此工作負載的所有元件實作定價模式](#)
- [COST07-BP05 在管理帳戶層級執行定價模式分析](#)

COST07-BP01 執行定價模式分析

分析工作負載的每個元件。判斷元件與資源會執行較長期間 (針對承諾折扣)，還是動態短期執行 (針對 Spot 或隨需)。使用成本管理工具中的建議對工作負載執行分析，並且對這些建議套用商業規則，以達到高報酬。

未建立此最佳實務時的曝險等級：高

實作指引

AWS 具有多種[定價模式](#)，可讓您以最經濟實惠的方式為資源付費，以滿足組織需求並取決於產品。請與您的團隊合作，確認最適當的定價模式。定價模式常會包含多種選項的組合，這取決於您的可用性

隨需執行個體可讓您按照時數或秒數 (最少 60 秒) 支付運算或資料庫容量的費用，視您執行的執行個體而定，且不需長期承諾或預付款。

Savings Plans 是一種靈活的定價模型，為 Amazon EC2、Lambda 和 AWS Fargate 用量提供較低價格，以換取一年或三年期限內一致用量 (以每小時美元計算) 的承諾。

Spot 執行個體是一種 Amazon EC2 定價機制，可讓您以折扣的小時費率 (最高可達隨需價格的 90%) 請求備用運算容量，而無須預先承諾。

透過預付容量，預留執行個體可讓您獲得高達 75% 的折扣。如需詳細資訊，請參閱[透過預留來最佳化成本](#)。

您可能會選擇為生產、品質和開發環境的相關資源納入 Savings Plans。或者，由於沙盒資源僅在需要時開啟，因此您可以為該環境中的資源選擇隨需模型。使用 Amazon [Spot 執行個體](#) 降低 Amazon EC2 成本，或使用 [Compute Savings Plans](#) 降低 Amazon EC2、Fargate 和 Lambda 成本。[AWS Cost Explorer](#) 建議工具透過 Savings Plans 提供承諾折扣的機會。

如果您過去曾為 Amazon EC2 購買[預留執行個體](#)，或者已在組織內部建立成本分配實務，則可以暫時繼續使用 Amazon EC2 預留執行個體。但我們建議應擬定相關策略，在未來使用 Savings Plans 作為更具彈性的節省成本機制。您可以隨時重新整理 AWS Cost Management 中的 Savings Plans (SP) 建議，以重新產生新的 Savings Plans 建議。使用預留執行個體 (RI) 來降低 Amazon RDS、Amazon Redshift、Amazon ElastiCache、Amazon OpenSearch Service 成本。有三個選項提供 Saving Plans 和預留執行個體：全額預付款、部分預付款和無預付款。使用 AWS Cost Explorer RI 和 SP 購買建議中提供的建議。

若要尋找 Spot 工作負載的機會，可使用整體用量的每小時檢視，並尋找定期出現用量或彈性變化的時段。您可以將 Spot 執行個體用於具備容錯能力和靈活性的各種應用程式。範例包括無狀態 Web 伺服器、API 端點、大數據和分析應用程式、容器化工作負載、CI/CD 與其他彈性工作負載。

分析您的 Amazon EC2 和 Amazon RDS 執行個體是否可在未使用時 (下班時間和週末) 關閉。相較於全年無休地使用，此方法可讓您降低成本達 70% 甚或更高。如果您有僅需在特定時間啟用的 Amazon Redshift 叢集，您可以暫停叢集，等稍後再繼續執行。當 Amazon Redshift 叢集或 Amazon EC2 和 Amazon RDS 執行個體停止時，運算計費也會隨之停止，而只會計算儲存費用。

請注意，[隨需容量保留](#) (ODCR) 不是定價折扣。無論您是否以預留容量執行執行個體，都需要支付按隨需費率計算的容量保留費用。若需要為預計要執行的資源提供足夠的容量，就必須考量這些因

素。ODCR 無須綁定長期承諾，您不再需要時即可取消，但也可適用 Savings Plans 或預留執行個體所提供的折扣。

實作步驟

- 分析工作負載彈性：使用 Cost Explorer 中的每小時精細度或自訂儀表板，分析工作負載的彈性。尋找正在執行的執行個體數量的定期變更。短期執行個體是 Spot 執行個體或 Spot 機群的候選項目。
 - [Well-Architected 實驗室：Cost Explorer](#)
 - [Well-Architected 實驗室：成本視覺化](#)
- 審核現有定價合約：針對長期需求，審核目前的合約或承諾。分析您目前擁有的項目，以及有多少承諾正在使用中。運用既有的合約折扣或企業協議。[企業協議](#)為客戶提供最符合其需求的專屬協議選項。針對長期承諾，請考慮特定執行個體類型、執行個體系列、AWS 區域 和可用區域的預留定價折扣、預留執行個體或 Savings Plans。
- 執行承諾折扣分析：在您的帳戶中使用 Cost Explorer，審核 Savings Plans 和預留執行個體建議。要驗證您是否以所需的折扣和風險實作了正確的建議，請遵循 [Well-Architected 實驗室](#)。

資源

相關文件：

- [存取預留執行個體建議](#)
- [執行個體購買選項](#)
- [AWS Enterprise](#)

相關影片：

- [節省高達 90% 的成本並在 Spot 執行生產工作負載](#)

相關範例：

- [Well-Architected 實驗室：Cost Explorer](#)
- [Well-Architected 實驗室：成本視覺化](#)
- [Well-Architected 實驗室：定價模型](#)

COST07-BP02 根據成本選擇區域

每個區域的資源定價可能不同。識別區域成本差異，並僅部署於具有較高成本的區域，以符合延遲、資料落地和資料主權要求。考量區域成本，有助於讓您針對此工作負載支付最低的總價。

未建立此最佳實務時的曝險等級：中

實作指引

[AWS 雲端 基礎設施](#)是全球性的，託管在[世界各地的多個地點](#)，並圍繞 AWS 區域、可用區域、Local Zones、AWS Outposts 和 Wavelength Zones 而建置。區域是世界上的實體位置，每個區域各有一個地理區域，而 AWS 在其中有多個可用區域。可用區域是每個區域內的多個隔離位置，由一或多個分散的資料中心組成，各自有其備援電力、聯網和連線能力。

每個 AWS 區域 各在當地市場條件之下運作，且各區域的資源定價因土地、光纖設施、電力和稅賦等因素而有所差異。您可以選擇特定區域以操作解決方案的元件或全部，以便以最低價格於全球執行。使用 [AWS Calculator](#)，按位置類型 (區域、Wavelength Zone 和 Local Zone) 和區域搜尋服務，以預估您的工作負載在不同區域中的成本。

當您建構解決方案時，一項最佳實務是盡量將運算資源置於接近使用者之處，以提供較低延遲和強大的資料主權。根據您的業務、資料隱私權、效能和安全要求，選取適當的地理位置。對於全球各地都有使用者的應用程式，請使用多個位置。

如果您在資料隱私權、安全和業務要求方面不受約束，請使用提供較低 AWS 服務價格的區域來部署工作負載。例如，如果您的預設區域是亞太地區 (雪梨) (ap-southwest-2)，且沒有使用其他區域方面的限制 (例如資料隱私權、安全)，則將非關鍵性 (開發和測試) Amazon EC2 執行個體部署在美國東部 (維吉尼亞北部) (us-east-1) 區域，將可降低成本。

	合規	延遲	成本	服務/功能
區域 1	✓	15 毫秒	\$\$	✓
區域 2	✓	20 毫秒	\$\$\$	X
區域 3	✓	80 毫秒	\$	✓
區域 4	✓	15 毫秒	\$\$	✓
區域 5	✓	20 毫秒	\$\$\$	X
區域 6	✓	15 毫秒	\$	✓
區域 7	✓	80 毫秒	\$	✓
區域 8	✓	15 毫秒	\$	X

區域功能矩陣表

上方的矩陣表顯示區域 6 是這種情況下的最佳選擇，因為與其他區域相比，其延遲很低、服務可供使用，並且是成本最低的區域。

實作步驟

- 審核 AWS 區域 定價：分析目前區域的工作負載成本。依服務和用量類型，從最高成本開始，計算其他可用區域的成本。如果預測儲存超過移動元件或工作負載的成本，請遷移至新區域。
- 審核多區域部署的要求：分析您的業務要求和義務 (資料隱私權、安全或效能)，確認是否有任何限制使您無法使用多個區域。如果沒有使用單一區域的限制，請使用多個區域。
- 分析所需的資料傳輸：選取區域時，請考慮資料傳輸成本。將資料存放在接近客戶與資源之處。選取資料流動成本較低、且資料傳輸最少的 AWS 區域。根據您對資料傳輸的業務需求，可以使用 [Amazon CloudFront](#)、[AWS PrivateLink](#)、[AWS Direct Connect](#) 和 [AWS Virtual Private Network](#) 來降低聯網成本、改善效能並增強安全性。

資源

相關文件：

- [存取預留執行個體建議](#)

- [Amazon EC2 定價](#)
- [執行個體購買選項](#)
- [區域表](#)

相關影片：

- [節省高達 90% 的成本並在 Spot 執行生產工作負載](#)

相關範例：

- [常見架構的資料傳輸成本概觀](#)
- [全球部署的成本考量](#)
- [為工作負載選取區域時應考慮的事項](#)

COST07-BP03 選取具成本效益條款的第三方協議

具成本效益的協議和條款可確保這些服務的成本隨其提供的優勢而擴展。選擇可在為您的組織提供額外優勢時擴展的協議和定價。

未建立此最佳實務時的曝險等級：中

實作指引

市場上有多種產品可以幫助您管理雲端環境的成本。它們在功能方面可能會有一些差異，而這取決於客戶要求，例如有些客戶專注於成本管控或成本可見性，其他客戶則專注於成本最佳化。有效成本最佳化和管控的一個關鍵因素是使用具有必要功能和合適定價模式的合適工具。這些產品具有不同的定價模式。有些產品會向您收取每月賬單的一定百分比，有些產品則收取所實現節省金額的百分比。理想情況下，請只為您需要的功能付費。

當您在雲端中使用第三方解決方案或服務時，定價結構務必要符合您想要的成果。定價應根據其提供的結果和價值進行擴展。例如，在會從節省的成本中提取一定比例的軟體中，節省的成本 (成果) 越多，收費就越高。會隨著開支增加而要支付更多費用的授權協議可能不會永遠對您的成本最佳化目標有利。但是，如果供應商能為您帳單的所有部分提供明確的效益，則此擴展費用可能是合理的。

例如，如果您使用其他無效益的服務，則會提供 Amazon EC2 建議並收取整個帳單一定比例的解決方案可能會變得更加昂貴。另一個範例是受管服務，其會依受管資源成本的一定百分比計費。較大的執行個體大小不一定需要更多的管理工作，但收費會更高。請確認這些服務定價安排在其服務中包含成本最佳化計劃或功能，以提升效率。

客戶可能會發現市場上的這些產品更先進或更易於使用。您需要考慮這些產品的成本，並考慮長遠的潛在成本最佳化成果。

實作步驟

- 分析第三方協議與條款：審核第三方協議中的定價。針對不同的用量等級執行建模，並將新成本納入考量，例如新服務用量，或因工作負載成長而產生的目前服務增加量。決定額外成本是否為您的企業提供所需的優勢。

資源

相關文件：

- [存取預留執行個體建議](#)
- [執行個體購買選項](#)

相關影片：

- [節省高達 90% 的成本並在 Spot 執行生產工作負載](#)

COST07-BP04 針對此工作負載的所有元件實作定價模式

永久執行的資源應使用預留容量，例如 Savings Plans 或預留執行個體。設定短期容量以使用 Spot 執行個體或 Spot 機群。隨需執行個體僅用於無法中斷且執行時間不夠長，以及不適合使用預留容量的短期工作負載 (介於 25% 到 75% 之間的時間，視資源類型而定)。

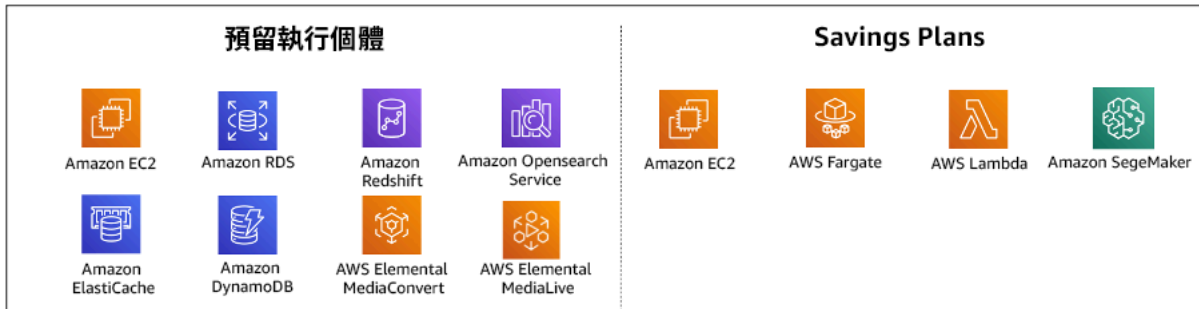
未建立此最佳實務時的曝險等級：低

實作指引

為了提高成本效率，AWS 會根據您過去的用量提供多個承諾建議。您可以使用這些建議來了解您可以節省的成本，以及如何使用承諾。您可以將這些服務作為隨需服務、Spot 服務，也可以承諾一定時間，並使用預留執行個體 (RI) 和 Savings Plans (SP) 降低隨需成本。您不僅需要了解每個工作負載元件和多項 AWS 服務，還需要了解這些服務的承諾折扣、購買選項和 Spot 執行個體，才能將工作負載最佳化。

考慮工作負載元件的要求，並了解這些服務的不同定價模式。定義這些元件的可用性要求。判斷是否有多個獨立資源在工作負載中執行相同功能，以及隨時間工作負載需求的變化。比較使用預設隨需定價模式和其他適用的模式的資源成本。考量資源或工作負載元件的任何潛在變更。

例如，讓我們看看 AWS 上的這個 Web 應用程式架構。此範例工作負載包含多種 AWS 服務，例如 Amazon Route 53、AWS WAF、Amazon CloudFront、Amazon EC2 執行個體、Amazon RDS 執行個體、負載平衡器、Amazon S3 儲存和 Amazon Elastic File System (Amazon EFS)。您需要審核這些服務中的每一項，並透過不同的定價模式找出潛在的成本節省機會。其中有些服務可能符合 RI 或 SP 的資格，有些則只能隨需提供。如下圖所示，部分 AWS 服務可以使用 RI 或 SP 來重諾。



使用預留執行個體和 Savings Plans 承諾的 AWS 服務

實作步驟

- 實作定價模式：使用分析結果，購買 Savings Plans、預留執行個體或實作 Spot 執行個體。如果是第一次購買承諾，請選擇清單中的前五項或前十項建議，然後監控和分析未來一兩個月的結果。AWS Cost Management Console 會引導您完成該過程。從主控台審核 RI 或 SP 建議、自訂建議 (類型、付款和期限)，並審核每小時承諾 (例如每小時 20 美元)，然後加入到購物車。折扣會自動套用到符合資格的用量。定期購買少量承諾折扣 (例如每 2 週或每月)。針對可能中斷或無狀態的工作負載，實作 Spot 執行個體。最後，選取隨需 Amazon EC2 執行個體，並為其餘要求配置資源。
- 工作負載審查週期：實作工作負載的審查週期，特別分析定價模型涵蓋範圍。一旦工作負載達到所需的涵蓋範圍，請部分購買額外的承諾折扣 (每隔幾個月)，或隨著組織用量的變更進行購買。

資源

相關文件：

- [了解您的 Savings Plans 建議](#)
- [存取預留執行個體建議](#)
- [如何購買預留執行個體](#)
- [執行個體購買選項](#)
- [Spot 執行個體](#)
- [其他 AWS 服務的預留模型](#)

- [Savings Plans 支援的服務](#)

相關影片：

- [節省高達 90% 的成本並在 Spot 執行生產工作負載](#)

相關範例：

- [購買 Savings Plans 前應考量哪些事項？](#)
- [如何使用 Cost Explorer 來分析我的支出和用量？](#)

COST07-BP05 在管理帳戶層級執行定價模式分析

查看計費和成本管理工具，並檢視承諾和保留的建議折扣，在管理帳戶層級執行定期分析。

未建立此最佳實務時的曝險等級：低

實作指引

執行定期成本建模可讓您有機會進行多個工作負載間的優化。例如，如果多個工作負載使用隨需執行個體，則在彙總層級變更的風險會更低，而且實作以承諾為基礎的折扣能獲得更低的整體成本。建議以兩週到一個月的頻率定期執行分析。這可讓您進行小幅的調整，因此定價模式的涵蓋範圍會隨著不斷變化的工作負載及其元件不斷演變。

使用 [AWS Cost Explorer](#) 建議工具，在您的管理帳戶中尋找承諾折扣的機會。管理帳戶層級的建議在計算過程中會考量您的 AWS 組織中已啟用預留執行個體 (RI) 或 Savings Plans (SP) 折扣分享的帳戶。計算過程也會在折扣分享啟用時啟動，以推薦可盡量節省整體帳戶成本的承諾。

雖然在許多情況下，在管理帳戶層級購買可省下最多成本，但在某些情況下，您可以考慮在連結帳戶層級購買 SP，例如，您希望先將折扣套用至該連結帳戶中的用量時。成員帳戶建議會在個別帳戶層級上進行計算，以盡可能節省各個獨立帳戶的成本。如果您的帳戶同時擁有 RI 和 SP 承諾，則會按以下順序套用這些承諾：

1. 區域 RI
2. 標準 RI
3. 可轉換 RI
4. Instance Savings Plan

5. Compute Savings Plan

如果您在管理帳戶層級購買 SP，則將根據最高到最低的折扣百分比來套用節省的金額。管理帳戶層級的 SP 會查看所有連結帳戶，並以最高的折扣套用節省的金額。如果您希望限定節省金額的套用項目，您可以在連結的帳戶層級購買 Savings Plan，如此，每當該帳戶執行符合資格的運算服務時，就會先為該項目套用折扣。當帳戶未執行符合資格的運算服務時，折扣將會分享到相同管理帳戶下的其他連結帳戶。折扣分享預設為開啟，但可視需要關閉。

在合併帳單系列中，Savings Plans 會先套用至擁有者帳戶的用量，然後套用至其他帳戶的用量。只有在折扣分享啟用時，才會執行此模式。您的 Savings Plans 會先套用至您最高的節省金額百分比。如果有多種用量具有相同的節省金額百分比，則 Savings Plans 會套用至第一個具有最低 Savings Plans 費率的用量。Savings Plans 將繼續套用，直至沒有剩餘用量或您的承諾用盡為止。任何剩餘用量均按隨需費率收費。您可以隨時重新整理 AWS Cost Management 中的 Savings Plans 建議，以重新產生新的 Savings Plans 建議。

分析執行個體的彈性後，您可以採納建議的承諾。使用可能的不同資源選項分析工作負載的短期成本、分析 AWS 定價模型，並使其符合您的業務要求，以找出總體擁有成本和[成本最佳化](#)機會。

實作步驟

執行承諾折扣分析：在您的帳戶中使用 Cost Explorer，審核 Savings Plans 和預留執行個體建議。請確實了解 Saving Plan 建議，並估計您的每月支出和每個月節省的成本。審查管理帳戶層級的建議；其計算過程中考量到您的 AWS 組織中已啟用 RI 或 Savings Plans 折扣分享，以盡可能節省帳戶成本的所有成員帳戶間的整體用量。您可以依照 Well-Architected 實驗室的指示，確定在所需的折扣與風險方面，採用了正確的建議。

資源

相關文件：

- [AWS 定價如何運作？](#)
- [執行個體購買選項](#)
- [Saving Plan 概觀](#)
- [Saving Plan 建議](#)
- [存取預留執行個體建議](#)
- [了解 Savings Plans 建議](#)
- [Savings Plans 如何套用於 AWS 使用情況](#)

- [具有合併帳單功能的 Savings Plans](#)
- [開啟共享的預留執行個體和 Savings Plans 折扣](#)

相關影片：

- [節省高達 90% 的成本並在 Spot 執行生產工作負載](#)

相關範例：

- [在購買 Savings Plan 前，我應考量哪些事項？](#)
- [如何利用滾動 Savings Plans 降低承諾風險？](#)
- [何時應使用 Spot 執行個體](#)

COST 8. 如何規劃資料傳輸費？

確實規劃和監控資料傳輸費，以便做出盡量減少成本的架構決策。小規模而有效的架構變更能夠隨時間大幅減少營運成本。

最佳實務

- [COST08-BP01 執行資料傳輸建模](#)
- [COST08-BP02 選取元件以將資料傳輸成本最佳化](#)
- [COST08-BP03 實作可降低資料傳輸成本的服務](#)

COST08-BP01 執行資料傳輸建模

收集組織要求並執行工作負載及其每個元件的資料傳輸建模。這可確定其目前資料傳輸要求的最低成本點。

未建立此最佳實務時的曝險等級：高

實作指引

在設計雲端解決方案時，由於習慣使用內部部署資料中心來設計架構或缺乏知識，通常會忽略掉資料傳輸費用。AWS 中的資料傳輸費用會由來源、目的地和流量的數量來決定。在設計階段考慮這些費用能夠讓您省下成本。了解資料傳輸在工作負載中的發生位置、傳輸成本及其相關效益，對於準確估算總體擁有成本 (TCO) 來說非常重要。這可讓您做出明智的決策，以修改或接受架構決策。例如，您可能有一個多個可用區域組態，您在可用區域之間複寫資料。

您要為會在工作負載中傳輸資料的服務元件建模，並決定這是實現所需可靠性和彈性可接受的成本 (類似於在兩個可用區域中支付運算和儲存費用)。針對不同用量等級建立成本模型。工作負載用量會隨時間改變，在不同等級，不同的服務可能更經濟實惠。

在為資料傳輸建模時，請考慮所擷取的資料量以及資料的來源。此外，也請考慮所處理的資料量以及需要的儲存或運算容量。在建模期間，請遵循工作負載架構的聯網最佳實務，以將潛在的資料傳輸成本最佳化。

AWS 定價計算工具 可以幫助您查看特定 AWS 服務的預估成本和預期的資料傳輸。如果您的工作負載已經在執行 (用於測試目的或在生產前環境中)，請使用 [AWS Cost Explorer](#) 或 [AWS Cost and Usage Report](#) (CUR) 來了解資料傳輸成本並建模。設定概念驗證 (PoC) 或測試工作負載，並以逼真的模擬負載執行測試。您可以根據不同的工作負載需求建立成本模型。

實作步驟

- 確定需求：在來源與目的地之間規劃的資料傳輸的主要目標和業務需求是什麼？所預期的最終業務成果是什麼？收集業務要求並定義預期的成果。
- 識別來源和目的地：資料傳輸的資料來源和目的地是什麼，例如在 AWS 區域 內部、傳輸至 AWS 服務或傳出到網際網路？
 - [AWS 區域 內的資料傳輸](#)
 - [AWS 區域 之間的資料傳輸](#)
 - [資料傳出到網際網路](#)
- 識別資料分類：此資料傳輸的資料分類為何？這是什麼種類的資料？資料有多大？資料必須以何種頻率進行傳輸？資料敏感嗎？
- 識別要使用的 AWS 服務或工具：此資料傳輸使用哪些 AWS 服務？是否可將已佈建的服務用於其他工作負載？
- 計算資料傳輸成本：使用先前建立的資料傳輸模型 [AWS Pricing](#) 來計算工作負載的資料傳輸成本。針對工作負載用量的增加和減少，計算不同用量等級的資料傳輸成本。如果工作負載架構具有多個選項，請計算每個選項的成本進行比較。
- 將成本與結果連結：對於產生的每筆資料傳輸成本，請指定其為工作負載達到的結果。如果在元件之間傳輸，可能是用於解耦，如果在可用區域之間傳輸，則可能是用於備援。
- 建立資料傳輸模型：收集所有資訊後，為多個使用案例和不同的工作負載建立概念性基礎資料傳輸模型。

資源

相關文件：

- [AWS 快取解決方案](#)
- [AWS 定價](#)
- [Amazon EC2 定價](#)
- [Amazon VPC 定價](#)
- [了解資料傳輸費用](#)

相關影片：

- [監控和最佳化您的資料傳輸成本](#)
- [S3 Transfer Acceleration](#)

相關範例：

- [常見架構的資料傳輸成本概觀](#)
- [用於聯網的 AWS 規範指引](#)

COST08-BP02 選取元件以將資料傳輸成本最佳化

選擇所有元件，並設計架構以降低資料傳輸成本。這包括使用廣域網路 (WAN) 最佳化和多可用區域 (AZ) 組態等元件

未建立此最佳實務時的曝險等級：中

實作指引

資料傳輸建構可將資料傳輸成本降至最低。這可能涉及使用內容交付網路以將資料靠近使用者放置，或從您內部至 AWS 使用專用網路連結。您也可以使用 WAN 優化和應用程式優化，來減少元件之間傳輸的資料量。

將資料傳輸到 AWS 雲端 或於其中傳輸資料時，重要的是根據不同的使用案例來了解目的地、資料性質和可用的網路資源，以便選取合適的 AWS 服務來將資料傳輸最佳化。AWS 提供了一系列針對各種資料遷移要求量身打造的資料傳輸服務。根據組織內的業務需求，選擇正確的[資料儲存](#)和[資料傳輸](#)選項。

在計劃或審核工作負載架構時，請考慮下列事項：

- 在 AWS 中使用 VPC 端點：VPC 端點允許您的 VPC 和支援的 AWS 服務之間建立私有連線。這可讓您避免使用可能會產生資料傳輸成本的公用網際網路。

- 使用 NAT 閘道：使用 [NAT 閘道](#)，以便私有子網路中的執行個體可以連線至網際網路或 VPC 外部的服務。檢查 NAT 閘道後方傳送最多流量的資源是否與 NAT 閘道位於相同的可用區域。如果沒有，請在與該資源相同的可用區域中建立新的 NAT 閘道，以降低跨 AZ 資料傳輸費用。
- 使用 AWS Direct Connect Direct Connect 繞過公用網際網路，並在內部部署網路與 AWS 之間建立直接、私有連線。這可能會比透過網際網路傳輸大量資料更具成本效益和一致性。
- 避免跨區域界限傳輸資料：在 AWS 區域之間 (從一個區域到另一個區域) 傳輸資料通常會產生費用。請深思熟慮後再決定是否追求多區域路徑。如需詳細資訊，請參閱[多區域案例](#)。
- 監控資料傳輸：使用 Amazon CloudWatch 和 [VPC 流程日誌](#)擷取有關資料傳輸和網路使用情況的詳細資訊。分析 VPC 中擷取到的網路流量資訊，例如進出網路介面的 IP 位址或範圍。
- 分析網路用量：使用計量和報告工具 (例如 AWS Cost Explorer、CUDOS Dashboards 或 CloudWatch) 來了解工作負載的資料傳輸成本。

實作步驟

- 選擇用於資料傳輸的元件：使用 [COST08-BP01 執行資料傳輸建模](#) 中所述的資料傳輸模型，專注於資料傳輸成本最高的位置或工作負載用量變更時資料傳輸成本最高的位置。尋找替代架構或其他元件，以消除或降低資料傳輸需求 (或降低其成本)。

資源

相關的最佳實務：

- [COST08-BP01 執行資料傳輸建模](#)
- [COST08-BP03 實作可降低資料傳輸成本的服務](#)

相關文件：

- [雲端資料遷移](#)
- [AWS 快取解決方案](#)
- [使用 Amazon CloudFront 更快地交付內容](#)

相關範例：

- [常見架構的資料傳輸成本概觀](#)
- [AWS 網路最佳化秘訣](#)

- [使用 Apache Parquet 格式的 VPC 流程日誌，最佳化效能並降低網路分析成本](#)

COST08-BP03 實作可降低資料傳輸成本的服務

實作服務以減少資料傳輸。例如，使用邊緣節點或內容交付網路 (CDN) 將內容提供給終端使用者、在應用程式伺服器或資料庫前面建置快取層，以及使用專用網路連線而非 VPN 來連線至雲端。

未建立此最佳實務時的曝險等級：中

實作指引

有許多 AWS 服務可以協助您最佳化網路資料傳輸用量。根據您的工作負載元件、類型和雲端架構，這些服務可以協助您在雲端上壓縮、快取、共用和分配流量。

- [Amazon CloudFront](#) 是全球內容交付網路，在低延遲和高傳輸速度之下遞送資料。其快取位於全球節點的資料，能減輕您的資源所受的負載。藉由 CloudFront，在最低延遲之下交付內容給全球大量使用者方面，您可減少管理所費的心力。如果您計劃隨著時間的推移增加用量，[安全節省服務包](#) 可以幫助您節省高達 30% 的 CloudFront 使用率。
- [AWS Direct Connect](#) 可讓您建立連接至 AWS 的專用網路連線。如此可降低網路成本，增加頻寬，並且比網際網路連線提供更一致的網路體驗。
- [Site-to-Site VPN](#) 可讓您在私有網路和 AWS 全球網路之間建立安全且私有的連線。它非常適合小型辦公室或商業合作夥伴，因為它提供簡便的連線，而且是全受管的彈性服務。
- [VPC 端點](#) 允許透過私有網路連接各 AWS 服務，可用於降低公有聯網的資料傳輸量和 [NAT 閘道](#) 的成本。[閘道 VPC 端點](#) 不收取小時費用，且支援 Amazon S3 和 Amazon DynamoDB。[介面 VPC 端點](#) 由 [AWS PrivateLink](#) 提供，收取小時費用和每 GB 使用費。
- [NAT 閘道](#) 提供內建擴展和管理功能，與獨立 NAT 執行個體相比，成本更低。將 NAT 閘道放置在與高流量執行個體相同的可用區域中，並考慮為需要存取 Amazon DynamoDB 或 Amazon S3 的執行個體使用 VPC 端點，來降低資料傳輸和處理成本。
- 使用具有運算資源的 [AWS Snow Family](#) 裝置在邊緣收集和處理資料。AWS Snow Family 裝置 ([Snowball Edge](#)、[Snowball Edge](#) 和 [Snowmobile](#)) 讓您能夠以成本效益高且離線的方式將 PB 規模的資料移至 AWS 雲端。

實作步驟

- 實作服務：使用資料傳輸建模並審核 VPC Flow Logs，根據您的服務工作負載類型選擇適用的 AWS 網路服務。查看成本最高和磁碟區流量最大的情況。審核 AWS 服務，並評估是否有可減少或移除傳輸的服務，特別是聯網和內容交付方面。另請尋找可重複存取資料或大量資料的快取服務。

資源

相關文件：

- [AWS Direct Connect](#)
- [AWS 探索我們的產品](#)
- [AWS 快取解決方案](#)
- [Amazon CloudFront](#)
- [AWS Snow Family](#)
- [Amazon CloudFront 安全防護優惠組合](#)

相關影片：

- [監控和最佳化您的資料傳輸成本](#)
- [AWS 成本最佳化系列：CloudFront](#)
- [如何降低 NAT 閘道的資料傳輸費用？](#)

相關範例：

- [如何退款共享服務：AWS Transit Gateway 範例](#)
- [使用 Athena 查詢和 QuickSight，從成本和用量報告深入了解 AWS 資料傳輸詳細資訊](#)
- [常見架構的資料傳輸成本概觀](#)
- [使用 AWS Cost Explorer 分析資料傳輸成本](#)
- [利用 Amazon CloudFront 功能，針對您的 AWS 架構進行成本最佳化](#)
- [如何降低 NAT 閘道的資料傳輸費用？](#)

管理需求與供應資源

問題

- [COST 9. 如何管理需求和供應資源？](#)

COST 9. 如何管理需求和供應資源？

針對支出和效能達到平衡的工作負載，確認您購買的每一個項目都用到，並避免極少使用執行個體。往任一端傾斜的使用指標，對您組織在營運成本 (因過度使用而降低效能) 或浪費的 AWS 花費 (因過度佈建) 方面會造成負面影響。

最佳實務

- [COST09-BP01 對工作負載需求進行分析](#)
- [COST09-BP02 實作緩衝或調節機制來管理需求](#)
- [COST09-BP03 動態提供資源](#)

COST09-BP01 對工作負載需求進行分析

分析工作負載隨時間的需求。確認分析涵蓋季節性趨勢，並準確反映整個工作負載生命週期內的運作狀況。分析工作應反映潛在效益：例如，花費的時間與工作負載成本成正比。

未建立此最佳實務時的曝險等級：高

實作指引

要分析工作負載對雲端運算的需求，就必須了解雲端環境中啟動的運算工作模式和特性。這類分析可協助使用者優化資源配置、管理成本，並確保效能符合所需等級。

了解工作負載的需求。組織要求應指出請求的工作負載回應時間。回應時間可用來判斷需求是否已得到滿足，或是資源供應是否需要改變以符合需求。

分析應包含需求的可預測性和重複性、需求的變化速率，以及需求的變化量。針對足夠長的時間執行分析，以納入任何季節變化，例如月底處理或節假日尖峰。

分析工作應反映實作擴展的潛在效益。查看元件的預期總成本，以及工作負載生命週期內用量和成本的任何增加或減少。

以下是執行雲端運算的工作負載需求分析時需要考慮的一些關鍵事項：

1. 資源使用率和效能指標：分析 AWS 資源在一段時間內的使用方式。確認尖峰和離峰使用模式，以最佳化資源配置和擴展策略。監控效能指標，例如回應時間、延遲、輸送量和錯誤率。這些指標有助於評估雲端基礎架構的整體運作狀態和效率。
2. 使用者和應用程式擴展行為：了解使用者行為及其對工作負載需求的影響。檢查使用者流量的模式，有助於提高交付內容的完整性和應用程式的回應能力。分析工作負載如何隨著需求增加而擴展。判斷是否已正確、有效地設定自動擴展參數，以處理負載波動。

3. 工作負載類型：識別在雲端中執行的不同工作負載類型，例如批次處理、即時資料處理、Web 應用程式、資料庫或機器學習。每種工作負載類型可能有不同的資源需求和效能資料。
4. 服務水準協議 (SLA)：將實際效能與 SLA 進行比較，以確保合規性並找出需要改進的部分。

您可以使用 [Amazon CloudWatch](#) 收集和追蹤指標、監控日誌檔、設定警示，以及自動對 AWS 資源的變更做出反應。您也可以使用 Amazon CloudWatch 全面了解系統的資源使用率、應用程式效能和運作狀態。

使用 [AWS Trusted Advisor](#)，您可以根據最佳實務佈建資源，以改善系統效能和可靠性、提高安全性，並尋找節省成本的機會。您也可以關閉非生產執行個體，並使用 Amazon CloudWatch 和 Auto Scaling 來因應需求增加或減少。

最後，搭配 AWS Cost and Usage Report (CUR) 檔案或應用程式日誌使用 [AWS Cost Explorer](#) 或 [Quick](#)，以執行工作負載需求的進階分析。

整體而言，全面的工作負載需求分析可讓組織在資源佈建、擴展和最佳化方面做出明智決策，進而提高效能、成本效益和使用者滿意度。

實作步驟

- 分析現有工作負載資料：分析現有工作負載、舊版工作負載或預測使用模式中的資料。使用 Amazon CloudWatch、日誌檔和監控資料來深入了解工作負載的使用情況。分析工作負載的完整週期，並收集所有季節性變更的資料，例如月末或年末事件。分析中所反映的工作應反映工作負載特性。應將工作重點放在需求變更最大的高價值工作負載上。針對需求變更最少的低價值工作負載，應將投入的工作量降到最低。
- 預測外部影響：與整個組織中的團隊成員面談，這些成員可能會影響或變更工作負載的需求。常見的團隊是銷售團隊、行銷團隊或業務開發團隊。與這些團隊合作以了解其作業週期，以及是否有任何事件會改變工作負載需求。利用此資料來預測工作負載需求。

資源

相關文件：

- [Amazon CloudWatch](#)
- [AWS Trusted Advisor](#)
- [AWS X-Ray](#)
- [AWS Auto Scaling](#)

- [AWS Instance Scheduler](#)
- [Amazon SQS 入門](#)
- [AWS Cost Explorer](#)
- [快速](#)：

相關範例：

- [監控、追蹤和分析以實現成本最佳化](#)
- [搜尋和分析 CloudWatch 中的日誌](#)

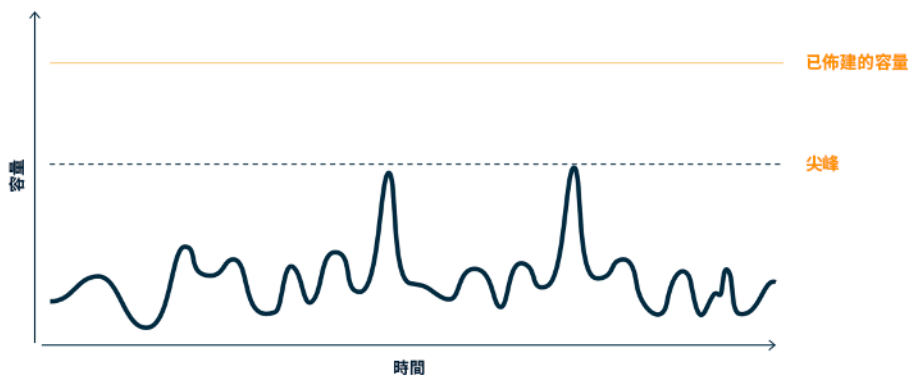
COST09-BP02 實作緩衝或調節機制來管理需求

緩衝和限流機制會修改工作負載的需求，以消除任何尖峰時段。在用戶端執行重試時實作限流機制。實作緩衝機制以儲存請求，並將處理的時間往後延遲。確認調節和緩衝機制經過設計，以便讓用戶端在所需時間內收到回應。

未建立此最佳實務時的曝險等級：中

實作指引

在雲端運算中實作緩衝或調節機制至關重要，如此才能管理需求並降低工作負載所需的佈建容量。為了獲得最佳效能，請務必評估總需求，包括峰值、請求變更速度以及必要的回應時間。當用戶端能夠重新發送他們的請求時，套用限流就變得很實用。相反地，對於缺少重試功能的用戶端，最理想的方法是實作緩衝解決方案。這類緩衝機制簡化了請求的湧入作業，並且會將有不同操作速度之應用程式的互動最佳化。



需求曲線圖，內含兩個需要大量佈建容量的相異尖峰

假設某個工作負載的需求曲線如上圖所示。此工作負載有兩個尖峰，為了處理這些尖峰，已佈建了資源容量 (以橙色線顯示)。用於此工作負載的資源和能源並非由需求曲線底下的區域表示，而是已佈建的容量底下的區域，因為這兩個尖峰必須用已佈建的容量處理。使工作負載需求曲線扁平化，有助於減少工作負載所需的已佈建容量，以及降低對環境造成的影響。若要消除尖峰時段，請考慮實作限流或緩衝解決方案。

為了深入了解，讓我們探索一下限流和緩衝機制。

限流：如果需求來源具有重試功能，則您可以實作限流。限流會告知來源，如果目前無法服務請求，則應稍後再試。來源會等待一段時間，然後重試請求。實作限流的優點是限制最大資源量和工作負載成本。在 AWS 中，您可以使用 [Amazon API Gateway](#) 實作限流。

基於緩衝區：基於緩衝區的方法會使用生產者 (將訊息傳送至佇列的元件)、取用者 (從佇列接收訊息的元件) 和佇列 (保留訊息) 來儲存訊息。消費者可讀取訊息並進行處理，允許以符合取用者業務要求的速度運作訊息。透過使用緩衝為主的方法，生產者的訊息會儲存在佇列或串流中，隨時可供取用者以符合其操作需求的速度來存取。

在 AWS 中，有多個服務可供選擇以實作緩衝方法。[Amazon Simple Queue Service \(Amazon SQS\)](#) 是一個受管服務，可提供佇列，允許單一取用者讀取個別訊息。[Amazon Kinesis](#) 可提供串流，允許許多取用者讀取相同訊息。

緩衝和限流可透過修改工作負載的需求來消除任何尖峰時段。當用戶端會重試動作時請使用限流，並使用緩衝機制來保存請求以供稍後處理。使用緩衝為主的方法時，請將工作負載建構為可在所需的時間內為請求提供服務，並確認您能夠處理重複的工作請求。分析整體需求、變更率及所需的回應時間，以適當調整所需的調節或緩衝區大小。

實作步驟

- **分析用戶端要求：**分析用戶端請求以判斷是否能夠執行重試。針對無法執行重試的用戶端，則需要實作緩衝機制。分析整體需求、變更率及所需的回應時間，以便判斷所需的調節或緩衝區大小。
- **實作緩衝區或限流：**在工作負載中實作緩衝區或限流。Amazon Simple Queue Service (Amazon SQS) 這類佇列可為工作負載元件提供緩衝區。Amazon API Gateway 可以為您的工作負載元件提供限流。

資源

相關的最佳實務：

- [SUS02-BP06 實作緩衝或限流使需求曲線趨於扁平化](#)
- [REL05-BP02 限流請求](#)

相關文件：

- [AWS Auto Scaling](#)
- [AWS Instance Scheduler](#)
- [Amazon API Gateway](#)
- [Amazon Simple Queue Service \(\)](#)
- [Amazon SQS 入門](#)
- [Amazon Kinesis](#)

相關影片：

- [為分散式應用程式選擇正確的訊息傳遞服務](#)

相關範例：

- [管理和監控工作負載中的 API 限流](#)
- [使用 API Gateway 大規模限流分層、多租用戶 REST API](#)
- [使用 Amazon API Gateway 在租用戶 Amazon EKS SaaS 解決方案中啟用分層和限流](#)
- [使用佇列與訊息進行應用程式整合](#)

COST09-BP03 動態提供資源

資源會按計劃進行佈建。這可以是以需求為基礎 (例如，透過自動調整規模)，或是以時間為基礎，其中需求可預測，並且根據時間提供資源。這些方法可盡量減少過度佈建或佈建不足的數量。

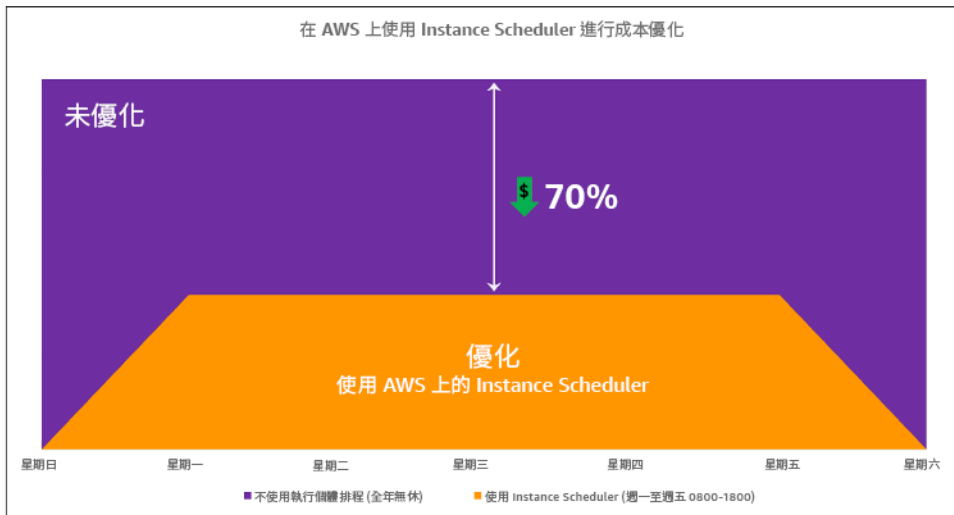
未建立此最佳實務時的曝險等級：低

實作指引

AWS 客戶有數種方法可以增加應用程式的可用資源並提供資源，以滿足需求。其中一個選項是使用 AWS Instance Scheduler，它可自動啟動和停止 Amazon Elastic Compute Cloud (Amazon EC2) 和 Amazon Relational Database Service (Amazon RDS) 執行個體。另一個選項是使用 AWS Auto Scaling，這可讓您根據應用程式或服務的需求自動擴展運算資源。根據需求提供資源可讓您僅為自己使用的資源付費，以及在需要時啟動資源，並在不需要資源時將其終止，藉以降低成本。

[AWS Instance Scheduler](#) 可讓您將 Amazon EC2 和 Amazon RDS 執行個體設定為在已定義的時間停止及啟動，以便在一致的時間模式下達到相同資源的需求，例如，使用者在每天早上八點存取 Amazon

EC2 執行個體，而晚上六點後則不需存取。此解決方案可停止非使用中的資源，並在需要時才加以啟動，藉以降低營運成本。



使用 AWS Instance Scheduler 優化成本。

您也可以使用 AWS Systems Manager 快速設定，透過簡單的使用者介面 (UI) 輕鬆設定跨帳戶和區域的 Amazon EC2 執行個體排程。您可以使用 AWS Instance Scheduler 來排程 Amazon EC2 或 Amazon RDS 執行個體，也可以停止和啟動現有的執行個體。但是，您無法停止和啟動屬於 Auto Scaling 群組 (ASG) 的執行個體，或管理 Amazon Redshift 或 Amazon OpenSearch Service 等服務的執行個體。Auto Scaling 群組對群組中的執行個體有自己的排程，並且會建立這些執行個體。

[AWS Auto Scaling](#) 可協助您調整容量，盡可能以最低的成本維持穩定、可預測的效能，以因應持續變動的需求。它是免費的全受管服務以擴展與 Amazon EC2 執行個體和 Spot 機群、Amazon ECS、Amazon DynamoDB 和 Amazon Aurora 整合的應用程式的能力。Auto Scaling 提供自動資源探索，以協助尋找工作負載中可設定的資源，它具有內建的擴展策略以優化效能、成本或兩者之間的平衡，並提供預測擴展以協助處理定期發生的尖峰。

有多個擴展選項可用來擴展您的 Auto Scaling 群組：

- 隨時維持目前執行個體層級
- 手動擴展
- 依據排程擴展
- 依據需求擴展
- 使用預測擴展

Auto Scaling 政策有所不同，可分類為動態和排程擴展政策。動態政策是手動或動態擴展，屬於排程或預測擴展。您可以使用擴展政策來進行動態、排程和預測擴展。也可以使用 [Amazon CloudWatch](#) 中的指標和警示，為您的工作負載觸發擴展事件。建議您使用[啟動範本](#)，它允許您存取最新功能和改善項目。即便使用啟動組態，也並非所有 Auto Scaling 功能都可用。例如：您無法建立同時啟動 Spot 及隨需執行個體的 Auto Scaling 群組或指定多個執行個體類型的群組。您必須使用啟動範本來設定這些功能。使用啟動範本時，建議您對每個範本進行版本控制。藉由啟動範本的版本控制，您可以建立一組完整的參數子集，之後可以重複使用子集來建立相同啟動範本的其他版本。

您可以使用 AWS Auto Scaling，或使用 [AWS API 或 SDK](#) 在程式碼中加入擴展功能。透過消除手動變更環境所需的營運成本，這可讓您降低整體工作負載成本，且變更的執行速度更快。這也可讓您隨時依據需求做出相應的工作負載資源配置。為了遵循此最佳實務，並且為組織動態提供資源，您應了解 AWS 雲端 中的水平和垂直擴展，以及在 Amazon EC2 執行個體中執行的應用程式有何性質。建議讓您的雲端財務管理團隊與技術團隊相互合作，以遵循此最佳實務。

[彈性負載平衡 \(Elastic Load Balancing\)](#) 可跨多個資源分配需求以協助您進行擴展。藉由使用 ASG 和 Elastic Load Balancing，您可以用最佳方式路由流量以管理傳入請求，讓 Auto Scaling 群組中的所有執行個體都不會不堪負荷。請求會以循環方式散佈在目標群組的所有目標之間，而不考量容量或使用率。

典型的指標可以是標準 Amazon EC2 指標，例如 CPU 使用率、網路輸送量，以及 Elastic Load Balancing 觀察到的請求與回應延遲。若可行的話，您應該使用可指示客戶體驗的指標，這通常是自訂指標，可能源自您工作負載內的應用程式程式碼。為了在本文件中詳細說明如何動態滿足需求，我們將 Auto Scaling 分類為需求為主和時間為主的供應模式，並深入探討這兩種模式。

需求為主的供應：依賴幾近即時的需求狀態，充分利用雲端的彈性來供應資源，以滿足不斷變化的需求。對於需求為主的供應，請使用 API 或服務功能，以程式設計方式更動架構中的雲端資源量。這樣可讓您增減架構中元件的規模，在需求激增時增加資源數量以維持效能，待需求消退時減少容量以降低成本。

需求為主的供應 (動態擴展政策)



需求為主的動態擴展政策

- 簡單/階段式擴展：根據客戶手動定義的步驟，監控指標及新增/移除執行個體。
- 目標追蹤：類似恆溫器的控制機制，可自動新增或移除執行個體，以在客戶定義的目標上維護指標。

以需求為主的方法進行建構時，請牢記兩大考量要點。第一，了解必須多迅速地佈建起新的資源。第二，了解供應與需求之間差距的大小會改變。您必須隨時因應需求的改變速度，並為資源失敗做好準備。

時間為主的供應：時間為主方法能使資源容量符合可預測或依照時間定義完善的需求。這種方法通常不依存於資源的利用率。時間為主方法能確保需要資源的特定時間有資源可用，並且因為啟動程序和系統或一致性檢查的緣故，能在毫無延遲之下提供。採用時間為主方法，您可在忙碌期提供更多資源或增加容量。

時間為主的供應 (排程和預測擴展政策)



時間為主的擴展政策

您可以使用排程或預測自動擴展來實作時間為主的方法。可排定工作負載於定義的時間橫向擴展或縮減 (例如在營業時段開始時)，以便在使用者到來或需求增加時有資源可用。預測擴展會使用模式進行橫向擴展，而排程的擴展則使用先定義的時間進行橫向擴展。您也可以在 Auto Scaling 群組中使用[屬性型執行個體類型選取 \(ABS\) 策略](#)，以透過一組屬性 (例如 vCPU、記憶體和儲存) 來表達您的執行個體要求。這也可讓您自動使用新發行的新世代執行個體類型，並使用 Amazon EC2 Spot 執行個體來存取更大範圍的容量。Amazon EC2 Fleet 和 Amazon EC2 Auto Scaling 會選取和啟動符合指定屬性的執行個體，您不必再手動挑選執行個體類型。

您也可利用 [AWS API 和 SDK](#) 以及 [AWS CloudFormation](#) 以視需要自動佈建和停用整個環境。這種方法十分適合僅在定義的營業時段或時期執行的開發或測試環境。您可使用 API 縮放環境之內的資源大小 (垂直縮放)。例如，可變更執行個體的大小或類別，以擴展生產工作負載。作法是將執行個體停止再啟動，選擇不同的執行個體大小或類別。此技法亦可套用至其他資源，例如 Amazon EBS 彈性磁碟區，在使用中時經過修改可增加大小、調整效能 (IOPS) 或變更磁碟區類型。

以時間為主的方法進行建構時，請牢記兩大考量要點。首先，用量模式的一致性有多高？第二，若是模式改變會有何影響？您可藉由監控工作負載和使用商業智慧來提高預測的準確性。若看出用量模式有明顯變化，可調整時間以確保涵蓋。

實作步驟

- 設定排程擴展：針對可預測的需求變更，以時間為主的擴展機制可以及時提供正確的資源數目。此外，當資源建立和設定的速度不夠快，不足以回應隨需變更時，此機制也能派上用場。透過 AWS Auto Scaling，使用工作負載分析來設定排程的擴展。若要設定以時間為基礎的排程，您可以根據預期或可預測的負載變更，事先使用排程擴展的預測擴展來增加 Auto Scaling 群組中的 Amazon EC2 執行個體數目。
- 設定預測擴展：預測擴展允許您在流量的每日和每週模式之前增加 Auto Scaling 群組中的 Amazon EC2 執行個體數量。如果您有定期流量尖峰和啟動耗時的應用程式，則應考慮使用預測擴展。預測擴展可在預估的負載之前初始化容量，協助您以優於單純動態擴展 (本質上是被動的) 的速度進行擴展。例如，如果使用者在營業時間開始時開始使用您的工作負載，且在營業時間結束後不使用，則預測擴展可在營業時間之前新增容量，以消除動態擴展為了回應變動的流量而產生的延遲。
- 設定動態自動擴展：若要根據作用中的工作負載指標來設定擴展，請使用 Auto Scaling。使用分析並設定 Auto Scaling 以在正確的資源層級上啟動，並確認工作負載在所需的時間內擴展。您可以在單一 Auto Scaling 群組內啟動和自動擴展隨需執行個體和 Spot 執行個體組成的機群。除了獲得使用 Spot 執行個體的折扣之外，您還可以使用預留執行個體或 Savings Plan，以獲得定期隨需執行個體定價的折扣費率。這些因素組合起來，可協助您將 Amazon EC2 執行個體的費用節省最佳化，並協助您獲得應用程式所需的規模與效能。

資源

相關文件：

- [AWS Auto Scaling](#)
- [AWS Instance Scheduler](#)
- 擴展 Auto Scaling 群組的大小
- [Amazon EC2 Auto Scaling 入門](#)
- [Amazon SQS 入門](#)
- [Amazon EC2 Auto Scaling 排程擴展](#)
- [Amazon EC2 Auto Scaling 的預測擴展](#)

相關影片：

- [Auto Scaling 的目標追蹤擴展政策](#)
- [AWS Instance Scheduler](#)

相關範例：

- [Amazon EC2 Fleet 的 Auto Scaling 屬性型執行個體類型選取](#)
- [使用已排程的擴展，針對成本最佳化 Amazon Elastic Container Service](#)
- [Amazon EC2 Auto Scaling 的預測擴展](#)
- [如何搭配使用 Instance Scheduler 與 CloudFormation 來排程 Amazon EC2 執行個體？](#)

隨時間優化

問題

- [COST 10. 如何評估新服務？](#)
- [COST 11. 如何評估工作的成本？](#)

COST 10. 如何評估新服務？

隨著 AWS 推出新服務和功能，最佳實務是檢視現有架構決策，以確認其持續發揮最大成本效益。

最佳實務

- [COST10-BP01 開發工作負載審核程序](#)
- [COST10-BP02 定期檢閱和分析此工作負載](#)

COST10-BP01 開發工作負載審核程序

制定一個程序，用於定義工作負載審核的標準和程序。審查工作應反映潛在的效益。例如，核心工作負載或價值超過帳單 10% 的工作負載每季或每六個月審查一次，而低於 10% 的工作負載則每年審查一次。

未建立此最佳實務時的曝險等級：高

實作指引

為了擁有最符合成本效益的工作負載，您必須定期審查工作負載，以了解是否有機會實作新的服務、功能和元件。若要實現較低的整體成本，程序必須與可能的節省金額成正比。例如，相較於佔整體支出 5% 的工作負載，您應更頻繁且更徹底地審查佔整體支出 50% 的工作負載。考量任何外部因素或波動性。如果工作負載服務特定的地理或市場區隔，並且預測該區域會發生改變，則更頻繁的檢閱可能會帶來成本節省。需要檢閱的另一個因素是實作變更的工作量。如果測試與驗證變更需要付出大量成本，則應降低檢閱頻率。

考量維護過時和舊版元件和資源的長期成本，以及無法在其中實作新的功能。目前的測試和驗證成本可能會超過提議的效益。不過，隨著時間推移，工作負載與目前技術之間的差距增大，從而變更的成本可能會大幅增加，進而產生更高的成本。例如，移至新的程式設計語言目前看來可能並非具有成本效益之舉。不過，在五年後，該語言熟練人員的成本可能會增加，而且由於工作負載的成長，您會將更大的工作負載轉移到新的語言，此時需要付出的努力會比以前更多。

將您的工作負載細分成多個元件，指派元件的成本 (估算值就足夠)，然後在每個元件旁列出因素 (例如，工作量和外部市場)。使用這些指標來決定每個工作負載的檢閱頻率。例如，您可能會將 Web 伺服器視為高成本、變更所需工作量低和受外部因素影響高，因此檢閱頻率高。中央資料庫可能是中等成本、變更所需工作量高，以及受外部因素影響低，因此檢閱頻率中等。

定義一個程序，以在新的服務、設計模式、資源類型和組態可用時對其進行評估，進而優化您的工作負載。與[績效支柱審查](#)和[可靠性支柱審查](#)流程類似，識別、驗證及優先處理最佳化和改進活動以及問題修復，並將其納入您的待辦項目中。

實作步驟

- **定義審查頻率：**定義工作負載及其元件的審查頻率。配置時間和資源給持續性改進與審查頻率，以改進工作負載的效率和優化。這結合了許多因素，可能隨著組織內的工作負載而異，也可能隨著工作負載中的元件而異。常見的因素包括，在收入或品牌方面對組織的重要性、執行工作負載的總成本 (包括營運和資源成本)、工作負載的複雜性、實作變更的簡易性、任何軟體授權合約，以及因懲罰性授權，變更會導致授權成本大幅增加。元件可在功能或技術上進行定義，例如 Web 伺服器 and 資料庫，或運算和儲存資源。相應平衡這些因素，並為工作負載及其元件制定一個期間。您可以決定每 18 個月審查一次完整工作負載、每 6 個月審查一次 Web 伺服器、每 12 個月審查一次資料庫、每 6 個月審查一次運算和短期儲存，以及每 12 個月審查一次長期儲存。
- **定義審查徹底性：**定義審查工作負載或工作負載元件所需的工作量。與審查頻率類似，這需在多個因素之間取得平衡。評估改進機會並制定其優先順序，以將精力集中在可以帶來最大收益的機會上，同時預估這些活動需要多少工作量。如果預期成果未能達到目標，且所需的工作量成本較高，請使用替代行動方案重複進行。您的審查程序應包含專用的時間和資源，用於持續的漸進式改善。例如，您可以決定花費一週分析來資料庫元件、一週分析運算資源，以及花費四小時進行儲存審查。

資源

相關文件：

- [AWS 新聞部落格](#)
- [雲端運算的類型](#)
- [AWS 最新消息](#)

相關範例：

- [AWS 支援主動式服務](#)
- [工作負載的定期SAP工作負載檢閱](#)

COST10-BP02 定期檢閱和分析此工作負載

現有的工作負載會根據每個定義的程序定期接受審查，以確認是否可採用新服務、是否可取代現有服務、或是否可重新建構工作負載。

未建立此最佳實務時的曝險等級：中

實作指引

AWS 不斷新增新功能，因此您可以使用最新的技術更快地進行實驗和創新。[AWS 新功能](#)詳細說明 AWS 了如何執行此操作，並在發佈 AWS 服務、功能和區域擴展公告時提供快速概觀。您可以深入探討已公告推出的項目，並將其用來審查和分析現有的工作負載。為了實現新 AWS 服務和功能的優點，您可以檢閱工作負載並視需要實作新服務和功能。這表示您可能需要取代用於工作負載的現有服務，或將工作負載現代化，以採用這些新 AWS 服務。例如，您可以檢閱工作負載，並使用 Amazon Simple Email Service 取代簡訊元件。這消除了營運和維護執行個體叢集的成本，同時以較低的成本提供所有功能。

若要分析工作負載並凸顯潛在機會，您不僅應考慮使用新服務，也應使用新方法來建置解決方案。檢閱上的[這是我的架構](#)影片 AWS，以了解其他客戶的架構設計、挑戰及其解決方案。檢查 [All-In 系列](#)，了解 AWS 服務和客戶案例的真實應用程式。也可以觀看[回歸基礎](#)影片系列，其中說明、檢查和細分基本的雲端架構模式最佳實務。另一個來源是[如何建置此](#)影片，其設計旨在協助人們提出有關如何使用 AWS 服務實現最小可行產品（MVP）的重大想法。對於來自世界各地的建置者來說，這是一種從經驗豐富的 AWS 解決方案架構師取得架構指導的強烈想法。最後，可以檢閱[入門](#)資源材料，其中包含逐步教學課程。

在開始審查程序之前，請遵循您的企業在工作負載、安全和資料隱私權等方面的要求，以期在執行您同意的審查程序時，能夠採用特定的服務或區域和效能要求。

實作步驟

- 定期審查工作負載：使用您定義的程序，以指定的頻率執行審查。確認您在每個元件上付出正確的工作量。此程序與您選取服務來進行成本優化的初始設計程序類似。分析服務以及服務會帶來的效益，此時需考慮變更成本，而不僅僅是長期效益。
- 實作新服務：如果分析結果是要實作變更，請先執行工作負載的基準，以了解每個輸出的目前成本。實作變更，然後執行分析以確認每個輸出的新成本。

資源

相關文件：

- [AWS 新聞部落格](#)
- [AWS最新消息](#)
- [AWS 文件](#)
- [AWS 入門](#)
- [AWS 一般資源](#)

相關影片：

- [AWS - 這是我的架構](#)
- [AWS - 返回基本作業要點](#)
- [AWS - All-In 系列](#)
- [如何建置此方法](#)

COST 11. 如何評估工作的成本？

最佳實務

- [COST11-BP01 執行操作自動化](#)

COST11-BP01 執行操作自動化

評估雲端上的營運成本，著重於量化管理任務、部署中節省的時間和精力，並透過自動化降低人為錯誤、法規遵循和其他操作的風險。評估營運工作所需的時間和相關成本，並實作管理任務的自動化，以盡可能地減少手動工作量。

未建立此最佳實務時的曝險等級：低

實作指引

將操作自動化可減少人工作業的頻率、提升效率，且客戶可在部署、管理或操作工作負載時享有一致而穩定的體驗。您可以將基礎設施資源從手動操作任務中解放出來，並將其用於價值更高的任務與創新，這可提升商業價值。企業需要以經過實證和測試的方式來管理其雲端中的工作負載。該解決方案必須安全、快速且具有成本效益，並具有最低的風險和最大的可靠性。

首先，考慮整體營運成本，根據所需的工作量確定操作活動的優先順序。例如，在雲端中部署新資源、對現有資源進行優化變更，或實作所需的組態，分別需要多久的時間？透過考慮運營和管理成本來確定人為行動的總成本。排定管理任務的自動化優先順序，以減少人力。

審查工作量應反映潛在的效益。例如，檢查手動執行任務所花費的時間 (對照自動執行)。優先考慮自動化重複、高價值、耗時且複雜的活動。具有高價值或高人為錯誤風險的活動通常是開始自動化的起點，因為這類風險通常會產生不必要的額外營運成本 (例如營運團隊的加班費)。

使用 AWS Systems Manager 或 AWS Config 等自動化工具，簡化操作、合規、監控、生命週期和終止流程。使用 AWS 服務、工具或第三方產品時，可以自訂您實作的自動化，以滿足特定需求。下表顯示您可以透過 AWS 服務取得哪些核心操作功能與能力，以自動執行管理與操作：

- [AWS Audit Manager](#)：持續稽核 AWS 使用情況，以簡化風險與合規性評估
- [AWS Backup](#)：集中管理和自動化資料保護。
- [AWS Config](#)：配置計算資源，評定、審核、評估組態和資源清單。
- [AWS CloudFormation](#)：使用基礎設施即程式碼啟動高可用性資源。
- [AWS CloudTrail](#)：IT 變更管理、合規性和控制。
- [Amazon EventBridge](#) 會排程事件並觸發 AWS Lambda 以採取行動。
- [AWS Lambda](#)：透過事件觸發重複性流程，或使用 AWS EventBridge 按固定排程執行它們，從而進行自動化。
- [AWS Systems Manager](#)：啟動和停止工作負載、修補作業系統、自動化組態，以及持續管理。
- [AWS Step Functions](#)：排程工作並自動化工作流程。
- [AWS Service Catalog](#)：範本使用，具有合規性和控制的基礎設施即程式碼。

如果您希望藉由使用 AWS 產品和服務立即採用自動化，但您的組織沒有這類技能，請聯絡 [AWS Managed Services \(AMS\)](#)、[AWS 專業服務](#) 或 [AWS 合作夥伴](#)，以提高自動化的採用，並改善雲端的營運卓越性。

AWS Managed Services (AMS) 是代表企業客戶和合作夥伴營運 AWS 基礎設施的服務。它提供安全且合規的環境，您可以將工作負載部署至其中。AMS 使用企業雲端營運模型與自動化，讓您符合組織需求、更快速地遷移至雲端，以及降低持續管理成本。

AWS Professional Services 也可以幫助您實現所需的業務成果並使用 AWS 來自動化操作。它們可協助客戶部署已針對雲端進行優化的自動化、穩健而靈活的 IT 營運及管控能力。如需詳細的監控範例和建議的最佳實務，請參閱《卓越營運支柱》白皮書。

實作步驟

- 建置一次並多次部署：使用 CloudFormation、AWS SDK 或 AWS CLI 等基礎設施即程式碼，進行一次部署，然後多次用於類似的環境或災難復原案例。在部署時加上標籤以追蹤您的使用量，如其他最佳實務所定義。使用 [AWS Launch Wizard](#) 可減少部署許多熱門企業工作負載的時間。AWS Launch Wizard 會遵循 AWS 最佳實務，引導您完成企業工作負載的調整、設定和部署。也可以使用 [Service Catalog](#)，協助您建立和管理基礎設施即程式碼核准的範本，以便在 AWS 上使用，這樣任何人都能發現經核准的自助式雲端資源。
- 自動化持續合規性：考慮根據預先定義的標準，自動化記錄的組態的評估和修復。當您將 AWS Organizations 與 AWS Config 和 [AWS CloudFormation](#) 的功能結合使用時，可以有效管理和自動化數百個成員帳戶的大規模組態合規性。您可審核 AWS 資源之間組態和關係的變更，並深入了解資源組態的歷史。
- 自動化監控任務 AWS 提供各種可用來監控服務的工具。您可以設定這些工具來自動執行監控工作。建立和實作監控計畫來收集工作負載的全面監控資料，以便在出現多點故障時能更輕鬆地偵錯。例如，您可以使用自動化監控工具來監控 Amazon EC2，並在系統狀態檢查、執行個體狀態檢查和 Amazon CloudWatch 警示發生錯誤時向您回報。
- 自動化維護和操作：自動執行例行操作，無須人為介入。使用 AWS 服務和工具，可以選擇實作哪些 AWS 自動化，並根據您的特定需求進行定制。例如，使用 [EC2 Image Builder](#) 建置、測試和部署虛擬機器和容器映像，以便在 AWS 或內部部署中使用，或使用 AWS SSM 修補 EC2 執行個體。如果無法使用 AWS 服務完成所需操作，或者您需要透過篩選資源進行更複雜的操作，請使用 [AWS Command Line Interface](#) (AWS CLI) 或 AWS SDK 工具來自動化操作。AWS CLI 可使用指令碼來自動化控制和管理 AWS 服務的整個流程，而無須使用 AWS 管理主控台。選取您慣用的 AWS SDK 與 AWS 服務互動。如需其他程式碼範例，請參閱 AWS SDK 程式碼[範例儲存庫](#)。
- 使用自動化建立持續的生命週期：建立並保留成熟的生命週期政策非常重要，這不僅適用於法規或備援，還適用於成本最佳化。可使用 AWS Backup 集中管理和自動化資料存放區 (例如儲存貯體、磁碟區、資料庫和檔案系統) 的資料保護。也可以使用 Amazon Data Lifecycle Manager 來自動建立、保留和刪除 EBS 快照和 EBS 後端 AMI。
- 刪除不必要的資源：在沙箱或開發 AWS 帳戶中累積未使用的資源很常見。開發人員會在正常開發週期中建立並試驗各種服務和資源，然後在不再需要這些資源時不會刪除它們。未使用的資源可能會為組織帶來不必要的、有時甚至很高的成本。刪除這些資源可以降低操作這些環境的成本。如果不確定，請確保不再需要資料或已備份。可以使用 AWS CloudFormation 來清理已部署的堆疊，這會自動刪除範本中定義的大部分資源。或者，可以使用 [aws-nuke](#) 之類的工具，建立自動化功能，以刪除 AWS 資源。

資源

相關文件：

- [在 AWS 雲端中將營運現代化](#)
- [用於自動化的 AWS Services](#)
- [基礎設施和自動化](#)
- [AWS Systems Manager Automation](#)
- [自動和手動監控](#)
- [SAP 管理和操作的 AWS 自動化](#)
- [AWS Managed Services](#)
- [AWS 專業服務](#)

相關影片：

- [在 AWS 中大規模自動化持續合規性](#)
- [AWS Backup 示範：跨帳戶和跨區域備份](#)
- [修補您的 Amazon EC2 執行個體](#)

相關範例：

- [重塑自動化操作 \(第一部分\)](#)
- [重塑自動化操作 \(第二部分\)](#)
- [使用 aws-nuke 自動刪除 AWS 資源](#)
- [使用 AWS Config 和 AWS SSM 刪除未使用的 Amazon EBS 磁碟區](#)
- [在 AWS 中大規模自動化持續合規性](#)
- [使用 AWS Lambda 進行 IT 自動化](#)

永續性

在建置雲端工作負載時，永續性實務是關於了解所使用服務的影響、量化整體工作負載生命週期的影響，以及應用設計原則和最佳實務來減少這些影響。可以在[永續性支柱白皮書](#)中找到實作指引。

最佳實務領域

- [區域選擇](#)
- [因應需求](#)
- [軟體和架構](#)
- [資料](#)
- [硬體和服務](#)
- [程序和文化](#)

區域選擇

問題

- [SUS 1 如何為您的工作負載選取區域？](#)

SUS 1 如何為您的工作負載選取區域？

您為工作負載選擇的區域會極大程度地影響其 KPI，包括效能、成本和碳足跡。若要有效改進這些 KPI，請根據您的業務要求和永續性目標，選擇工作負載的區域。

最佳實務

- [SUS01-BP01 根據業務要求和永續性目標選擇區域](#)

SUS01-BP01 根據業務要求和永續性目標選擇區域

根據您的業務要求和永續性發展目標選擇工作負載的區域，以最佳化 KPI，包括效能、成本和碳足跡。

常見的反模式：

- 您可以根據自身所在位置選取工作負載的區域。
- 您可以將所有工作負載資源合併到單一地理位置。

建立此最佳實務的優勢：將工作負載放在 Amazon 可再生能源專案附近或所公佈的碳強度較低的區域附近，有助於降低雲端工作負載的碳足跡。

未建立此最佳實務時的曝險等級：中

實作指引

AWS 雲端 是會持續擴張的區域和連接點 (POP) 網路，並透過全球網路基礎設施彼此連結起來。您為工作負載選擇的區域會極大程度地影響其 KPI，包括效能、成本和碳足跡。若要有有效改進這些 KPI，請根據您的業務要求和永續性發展目標，選擇工作負載的區域。

實作步驟

- 列入候選清單的可能區域：遵循以下步驟，根據您的業務要求 (包括合規、可用功能、成本和延遲) 評估工作負載的可能區域，並將這些區域列入候選清單：
 - 根據必須遵守的當地法規，確認這些區域符合規範 (例如資料主權)。
 - 使用 [AWS 區域服務清單](#) 來檢查區域是否有您執行工作負載時所需的服務和功能。
 - 使用 [AWS 定價計算工具](#) 計算工作負載在每個區域的成本。
 - 測試終端使用者所在位置和每個 AWS 區域 之間的網路延遲。
- 選擇區域：選擇 Amazon 可再生能源專案附近的區域，以及電網公佈的碳強度低於其他位置 (或區域) 的區域。
 - 識別相關的永續性指引，根據[溫室氣體協定](#) (市場型和位置型方法) 來追蹤和比較逐年的碳排放。
 - 根據您用來追蹤碳排放的方法來選擇區域。如需根據永續性指引來選擇區域的詳細資訊，請參閱 [How to select a Region for your workload based on sustainability goals](#)。

資源

相關文件：

- [了解您的碳排放估算](#)
- [全球 Amazon](#)
- [可再生能源方法](#)
- [為工作負載選取區域時應考慮的事項](#)

相關影片：

- [AWS re:Invent 2023 - AWS 全球基礎設施的永續創新](#)
- [AWS re:Invent 2023 - 永續架構：過去、現在和未來](#)
- [AWS re:Invent 2022 - 提供可持續、高效能的架構](#)
- [AWS re:Invent 2022 - 永續架構並降低您的 AWS 碳足跡](#)

- [AWS re:Invent 2022 - AWS 全球基礎設施的永續性](#)

因應需求

問題

- [SUS 2 如何根據您的需求取得適當的雲端資源？](#)

SUS 2 如何根據您的需求取得適當的雲端資源？

使用者和應用程式使用工作負載和其他資源的方式，可協助您找到改善的機會，以達成永續性目標。擴展基礎架構以持續符合需求，並確認您僅使用支援使用者所需的最低資源。讓服務層級符合客戶需求。妥善放置資源，以限制使用者和應用程式使用資源所需的網路。移除未使用的資產。為團隊成員提供滿足其需求的裝置，同時將對永續性的影響降至最低。

最佳實務

- [SUS02-BP01 動態擴展工作負載基礎設施](#)
- [SUS02-BP02 SLAs與永續發展目標保持一致](#)
- [SUS02-BP03 停止建立和維護未使用的資產](#)
- [SUS02-BP04 根據其聯網需求最佳化工作負載的地理定位](#)
- [SUS02-BP05 為執行的活動最佳化團隊成員資源](#)
- [SUS02-BP06 實作緩衝或調節，以扁平需求曲線](#)

SUS02-BP01 動態擴展工作負載基礎設施

使用雲端的彈性並動態擴展您的基礎設施，以達到雲端資源的供需平衡，避免工作負載出現過度佈建的容量。

常見的反模式：

- 您不隨著使用者負載擴展基礎設施。
- 您一律手動擴展基礎設施。
- 您在擴展事件之後維持增加容量，而不是縮減規模。

建立此最佳實務的優勢：設定並測試工作負載彈性有助於有效達到雲端資源的供需平衡，並避免過度佈建的容量。您可以利用雲端中的彈性，在需求尖峰期間或之後自動擴展容量，以確保您使用的資源數量正好足以滿足業務所需。

未建立此最佳實務時的曝險等級：中

實作指引

雲端提供的彈性可透過各種機制來動態擴展或減少資源，以滿足需求的變化。平衡供需關係可將工作負載受到的影響降到最低。

需求可為固定或可變，需要指標和自動化以確保該項管理不致成為繁重的工作。應用程式可藉由修改執行個體大小進行垂直調整 (縱向擴展或縮減規模)、藉由修改執行個體數目進行水平調整 (縮減或橫向擴展)，或進行兩者的合併調整。

您可以使用多種不同的方法達到資源的供需平衡。

- 目標追蹤法：監控您的擴展指標，並視需要自動增加或減少容量。
- 預測擴展：縮減每日和每週趨勢的預期。
- 排程法：根據可預測的負載變化設定您自己的擴展排程。
- 服務擴展：挑選按設計原本就會擴展的服務 (例如無伺服器)，或提供自動擴展功能。

辨別使用率低或無使用率的時期，並調整資源規模以移除過剩容量、提高效率。

實作步驟

- 彈性會比對您擁有的資源供應與這些資源的需求。執行個體、容器和函數提供了彈性機制，可與自動擴展功能結合使用，或是作為服務功能提供。AWS 提供了多種自動擴展機制，以確保工作負載可在使用者負載較低時迅速輕易地縮減規模。以下是自動擴展機制的幾個範例：

自動擴展機制	在哪裡使用
Amazon EC2 Auto Scaling	用來確認您擁有正確數量的 Amazon EC2 執行個體可處理應用程式的使用者負載。
Application Auto Scaling	用來自動將個別 AWS 服務的資源擴展到 Amazon EC2 以外，例如 Lambda 函數或 Amazon Elastic Container Service (Amazon ECS) 服務。

自動擴展機制	在哪裡使用
Kubernetes Cluster Autoscaler - 我們常將擴展與 Amazon EC2 執行個體或 AWS Lambda 函數等運算服務一起討論。請考慮設定非運算服務 (例如 Amazon DynamoDB 讀取和寫入容量單位或 Amazon Kinesis Data Streams 碎片) 以符合需求。 - 確認會對要部署的工作負載類型驗證擴充或縮減規模的指標。如果您要部署影片轉碼應用程式，則預期為 100% CPU 使用率，且不應做為您的主要指標。您可以將 自訂指標 (例如記憶體使用率) 用於擴展政策 (如有必要)。若要選擇正確的指標，請考量 Amazon EC2 的下列指引： - 指標應為有效的使用率指標，並說明執行個體的忙碌程度。 - 指標值必須與 Auto Scaling 群組中的執行個體數成比例增加或減少。 - 對於 Auto Scaling 群組請使用 動態擴展，而非 手動擴展。我們也建議您在動態擴展中使用 目標追蹤擴展政策。 - 確認工作負載部署可處理橫向擴展和縮減事件。建立縮減事件的測試案例，以確認工作負載的行為符合預期，且不會對使用者體驗造成影響 (例如失去黏性工作階段)。您可以使用 活動歷史記錄 來驗證 Auto Scaling 群組的擴展活動。 - 評估工作負載以取得可預測模式，並在預計發生預測中的變化和隨需規劃變化時主動擴展。透過預測性擴展，可以消除過度佈建容量的需求。如需詳細資訊，請參閱 Predictive Scaling with Amazon EC2 Auto Scaling。	用來自動擴展 AWS 上的 Kubernetes 叢集。

資源

相關文件：

- [Amazon EC2 Auto Scaling 入門](#)
- [EC2 的預測擴展，採用機器學習技術](#)
- [使用 Amazon OpenSearch Service、Amazon Data Firehose 和 Kibana 分析使用者行為](#)
- [什麼是 Amazon CloudWatch？](#)
- [在 Amazon RDS 上使用 Performance Insights 監控資料庫負載](#)
- [介紹對於 Amazon EC2 Auto Scaling 預測擴展的原生支援](#)
- [介紹 Karpenter - 一個開放原始碼的高效能 Kubernetes Cluster Autoscaler](#)
- [深入探討 Amazon ECS 叢集自動擴展](#)

相關影片：

- [AWS re:Invent 2023 - 針對前 1000 萬個使用者在 AWS 上進行擴展](#)
- [AWS re:Invent 2023 - 永續架構：過去、現在和未來](#)
- [AWS re:Invent 2022 - 建置具成本、能源和資源效率的運算環境](#)
- [AWS re:Invent 2022 - 將容器從一個使用者擴展到數百萬名使用者](#)
- [AWS re:Invent 2023 - 使用 Amazon SageMaker AI 將 FM 推論擴展到數百種模型](#)
- [AWS re:Invent 2023 - 利用 Karpenter 的力量來擴展、最佳化和升級 Kubernetes](#)

相關範例：

- [自動擴展](#)

SUS02-BP02 SLAs與永續發展目標保持一致

根據您的永續性目標檢閱和最佳化工作負載服務層級協議（SLA），以盡可能減少支援工作負載所需的資源，同時繼續滿足業務需求。

常見的反模式：

- 工作負載SLAs未知或不明確。
- 您可以為可用性和效能定義您的 SLA。
- 您對所有工作負載使用相同的設計模式 (例如多可用區域架構)。

建立此最佳實務的好處：SLAs與永續發展目標保持一致，可在滿足業務需求的同時獲得最佳資源使用量。

未建立此最佳實務時的曝險等級：低

實作指引

SLAs 定義從雲端工作負載預期的服務層級，例如回應時間、可用性和資料保留。其影響範圍涵蓋雲端工作負載的架構、資源用量和環境影響。在定期檢視SLAs和進行權衡，以顯著降低資源使用量，以換取可接受的服務水準降低。

實作步驟

- 了解永續性目標：識別組織中的永續性目標，例如減碳或提高資源使用率。

- 檢閱 SLAs：評估您的 SLAs 以評估它們是否支援您的業務需求。如果您超過 SLAs，請執行進一步檢閱。
- 了解權衡取捨：了解工作負載複雜度 (例如大量並行使用者)、效能 (例如延遲) 和永續性影響 (例如所需資源) 之間的衡量取捨。通常情況下，優先考慮其中兩個因素會以犧牲第三個因素為代價。
- 調整 SLAs：SLAs 透過進行權衡來大幅降低永續性影響，以換取可接受的服務水準降低來調整您的。
 - 永續性和可靠性：高可用性的工作負載往往會耗用較多資源。
 - 永續性和效能：使用較多資源以提升效能，可能會對環境造成較大的影響。
 - 永續性和安全性：保護過度的工作負載可能會對環境造成較大的影響。
- 如果可能 SLAs，請定義永續性：包含工作負載 SLAs 的永續性。例如，將最低使用率層級定義為運算執行個體 SLA 的永續性。
- 使用有效的設計模式：在 上使用微服務等設計模式 AWS，以排定業務關鍵函數的優先順序，並允許非關鍵函數較低的服務層級 (例如回應時間或復原時間目標)。
- 溝通和建立責任：SLAs 與所有相關利益相關者共用，包括您的開發團隊和客戶。使用報告來追蹤和監控 SLAs。指派問責，以符合的永續性目標 SLAs。
- 使用激勵和獎勵：使用激勵和獎勵來實現或超越 SLAs 永續發展目標。
- 檢閱和迭代：定期檢閱和調整您的 SLAs，以確保其符合不斷變化的永續性和效能目標。

資源

相關文件：

- [了解恢復模式和權衡取捨以便在雲端中高效進行架構](#)
- [服務水準協議對 SaaS 供應商的重要性](#)

相關影片：

- [AWS re：Invent 2023 - 容量、可用性、成本效益：挑選三個](#)
- [AWS re：Invent 2023 - 永續架構：過去、現在和未來](#)
- [AWS re：Invent 2023 - 鬆散耦合系統的進階整合模式和權衡](#)
- [AWS re：Invent 2022 - 提供永續、高效能的架構](#)
- [AWS re：Invent 2022 - 建置具成本、能源和資源效益的運算環境](#)

SUS02-BP03 停止建立和維護未使用的資產

將您工作負載中未使用的資產除役，以降低支援您個人需求所需的雲端資源數量，並盡可能減少浪費。

常見的反模式：

- 您未分析應用程式是否有冗餘或不再需要的資產。
- 您未移除冗餘或不再需要的資產。

建立此最佳實務的優勢：移除未使用的資產可釋出資源，並改善工作負載的整體效率。

未建立此最佳實務時的曝險等級：低

實作指引

未使用的資產會耗用儲存空間和運算能力等資源。識別這些資產並將其消除可以釋出這類資源，進而提升雲端架構的效能。定期分析應用程式資產 (例如預先編譯的報告、資料集、靜態影像和資產存取模式)，以識別冗餘、未充分利用和可以除役的目標。移除這類冗餘資產，避免工作負載中的資源浪費。

實作步驟

- 執行清查：進行全面清查，以識別工作負載內的所有資產。
- 分析用量：使用持續監控功能識別不再需要的靜態資產。
- 移除未使用的資產：制定計畫來移除不再需要的資產。
 - 移除任何資產之前，均應先評估該移除對架構的影響。
 - 合併重疊產生的資產以消除冗餘處理。
 - 更新您的應用程式，使其不再產生及儲存不需要的資產。
- 與第三方通訊：指示第三方停止生產和儲存代表您管理但不再需要的資產。請求合併冗餘資產。
- 使用生命週期政策：使用生命週期政策來自動刪除未使用的資產。
 - 您可以使用 [Amazon S3 生命週期](#)，以在物件的整個生命週期中管理物件。
 - 您可以使用 [Amazon Data Lifecycle Manager](#) 自動建立、保留和刪除 Amazon EBS 快照與 Amazon EBS 支援的 AMI。
- 審查和最佳化：定期審查工作負載，以識別並移除任何未使用的資產。

資源

相關文件：

- [最佳化您的 AWS 永續性基礎設施，第 II 部分：儲存](#)
- [如何終止我的 AWS 帳戶 帳戶上不再需要的作用中資源？](#)

相關影片：

- [AWS re:Invent 2023 - 永續架構：過去、現在和未來](#)
- [AWS re:Invent 2022 - 使用 Amazon S3 保護數位媒體資產並使其價值最大化](#)
- [AWS re:Invent 2023 - 最佳化多帳戶環境中的成本](#)

SUS02-BP04 根據其聯網需求最佳化工作負載的地理定位

為您的工作負載選取可減少網路流量傳輸距離的區域和服務，並減少支援工作負載所需的整體網路資源。

常見的反模式：

- 您可以根據自身所在位置選取工作負載的區域。
- 您可以將所有工作負載資源合併到單一地理位置。
- 所有流量都流經現有資料中心。

建立此最佳實務的優勢：將工作負載分配到使用者附近的位置，可提供最低的延遲，同時減少網路間的資料移動，並降低環境影響。

未建立此最佳實務時的曝險等級：中

實作指引

AWS 雲端 基礎設施是以區域、可用區域、置放群組和邊緣位置等位置選項為基礎，例如 [AWS Outposts](#) 和 [AWS Local Zones](#)。這些位置選項負責維護應用程式元件、雲端服務、邊緣網路和內部部署資料中心之間的連線。

分析工作負載中的網路存取模式，以識別如何使用這些雲端位置選項，以及減少網路流量必須輸送的距離。

實作步驟

- 分析您工作負載中的網路存取模式，以識別使用者如何使用您的應用程式。
 - 使用監控工具，例如 [Amazon CloudWatch](#) 和 [AWS CloudTrail](#)，收集網路活動的資料。

- 分析資料以識別網路存取模式。
- 根據下列關鍵元素，為您的工作負載部署選取區域：
 - 您的永續性目標：相關說明請見[區域選擇](#)。
 - 資料所在位置：對於資料密集型應用程式 (例如大數據和機器學習)，應用程式碼執行時應盡可能接近資料。
 - 使用者所在位置：對於面向使用者的應用程式，請選擇接近工作負載使用者的一或多個區域。
 - 其他限制：考慮諸如成本和合規性之類的限制，如[為工作負載選取區域時應考慮的事項](#)中所述。
- 針對常用資產，使用本機快取或 [AWS 快取解決方案](#) 以提升效能、減少資料移動以及降低環境影響。

服務	使用情況
Amazon CloudFront	使用 快取靜態內容，例如影像、指令碼和影片，以及動態內容，例如API回應或 Web 應用程式。
Amazon ElastiCache	用來快取 Web 應用程式的內容。
DynamoDB Accelerator	用來將記憶體內加速新增至 DynamoDB 資料表。

- 使用可協助您在更接近工作負載使用者的位置執行程式碼的服務：

服務	使用情況
Lambda@Edge	用於在物件未經快取時起始的大量運算作業。
Amazon CloudFront Functions	用於可由短期函數啟動的簡單使用案例，例如 HTTP (s) 請求或回應操作。
AWS IoT Greengrass	用來為連線的裝置執行本機運算、傳訊和資料快取。

- 使用連線集區來支援連線重複使用，減少所需資源。
- 使用不仰賴持續連線和同步更新的分散式資料存放區來實現一致性，以服務區域的人口。
- 以共用動態容量取代預先佈建的靜態網路容量，與其他訂閱者分攤網路容量的永續性影響。

資源

相關文件：

- [最佳化 AWS 基礎設施以實現永續性，第 部分III：聯網](#)
- [Amazon ElastiCache 文件](#)
- [什麼是 Amazon CloudFront？](#)
- [Amazon CloudFront 金鑰功能](#)
- [AWS 全球基礎設施](#)
- [AWS 本機區域 和 AWS Outposts，為您的邊緣工作負載選擇正確的技術](#)
- [置放群組](#)
- [AWS 本機區域](#)
- [AWS Outposts](#)

相關影片：

- [在 上解密資料傳輸 AWS](#)
- [在新一代 Amazon EC2執行個體上擴展網路效能](#)
- [AWS Local Zones 解說器影片](#)
- [AWS Outposts：概觀和運作方式](#)
- [AWS re：Invent 2023 - 邊緣和內部部署工作負載的遷移策略](#)
- [AWS re：Invent 2021 - AWS Outposts：將 AWS 體驗帶到內部部署](#)
- [AWS re：Invent 2020 - AWS Wavelength：在 5G 邊緣以超低延遲執行應用程式](#)
- [AWS re：Invent 2022 - AWS Local Zones：為分散式邊緣建置應用程式](#)
- [AWS re：Invent 2021 - 使用 Amazon 建置低延遲網站 CloudFront](#)
- [AWS re：Invent 2022 - 使用 改善效能和可用性 AWS Global Accelerator](#)
- [AWS re：Invent 2022 - 使用 建置您的全球廣域網路 AWS](#)
- [AWS re：Invent 2020：使用 Amazon Route 53 進行全域流量管理](#)

相關範例：

- [AWS 網路研討會](#)
- [永續性架構 - 盡可能減少跨網路的資料移動](#)

SUS02-BP05 為執行的活動最佳化團隊成員資源

最佳化提供給團隊成員的資源，以盡量減少對環境永續性的影響，同時支援他們的需求。

常見的反模式：

- 您忽略團隊成員所使用的裝置對雲端應用程式的整體效率產生的影響。
- 您手動管理及更新團隊成員所使用的資源。

建立此最佳實務的優勢：最佳化團隊成員資源，可為啟用雲端的應用程式改善整體效率。

未建立此最佳實務時的曝險等級：低

實作指引

了解團隊成員用來使用您的服務的資源、其預期生命週期，以及財務和永續性的影響。實作將這些資源最佳化的策略。例如，在使用率高的可擴展基礎設施上執行複雜的操作 (例如轉譯和編譯)，而不是在使用率低的高功率單一使用者系統上執行。

實作步驟

- 使用節能工作站：為團隊成員提供節能的工作站和周邊裝置。在這些裝置中使用高效電源管理功能 (例如低功耗模式)，以減少其能源用量。
- 使用虛擬化：使用虛擬桌面和應用程式串流來限制升級與裝置要求。
- 鼓勵遠端協作：鼓勵團隊成員使用遠端協作工具 (如 [Amazon Chime](#) 或 [AWS Wickr](#)) 以減少出差需求和關聯的碳排放。
- 使用節能軟體：移除或關閉不必要的功能和流程，為團隊成員提供節能軟體。
- 管理生命週期：評估程序和系統對裝置生命週期的影響，並選取在滿足業務需求的同時可將裝置更換需求降至最低的解決方案。定期維護和更新工作站或軟體，以維護和提高效率。
- 遠端裝置管理：為裝置實作遠端管理，以減少必要商務差旅時間。
 - [AWS Systems Manager Fleet Manager](#) 是一種整合式使用者介面 (UI) 體驗，可協助您從遠端管理在 AWS 或內部部署環境執行的節點。

資源

相關文件：

- [什麼是 Amazon WorkSpaces？](#)

- [Amazon WorkSpaces 的成本優化器](#)
- [Amazon AppStream 2.0 文件](#)
- [NICE DCV](#)

相關影片：

- [在 AWS 上管理 Amazon WorkSpaces 的成本](#)

SUS02-BP06 實作緩衝或調節，以扁平需求曲線

緩衝和限流可讓需求曲線趨於扁平化，並減少您的工作負載所需的已佈建容量。

常見的反模式：

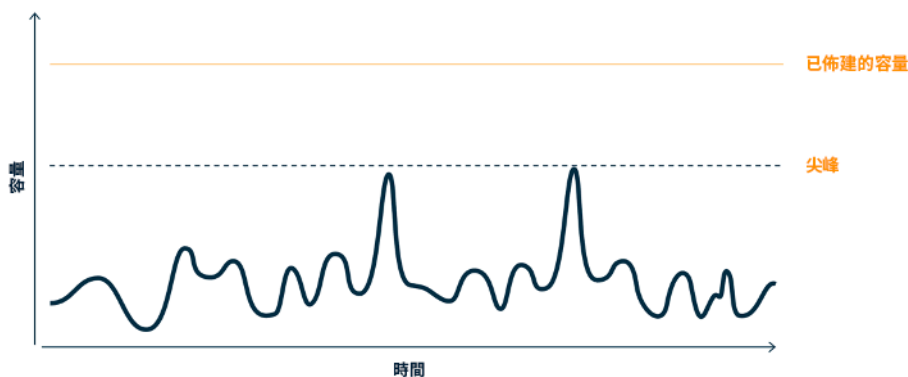
- 您非必要地立即處理用戶端請求。
- 您未分析用戶端要求的需求。

建立此最佳實務的優勢：讓需求曲線趨於扁平化，可減少工作負載所需的已佈建容量。減少已佈建的容量意味著較低的能源耗用量和環境影響。

未建立此最佳實務時的曝險等級：低

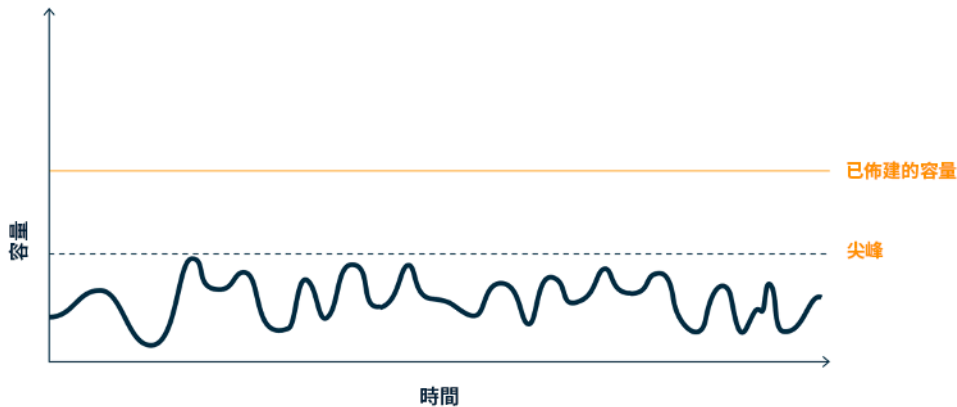
實作指引

使工作負載需求曲線扁平化，有助於減少工作負載所需的已佈建容量，以及降低對環境造成的影響。假設某个工作負載的需求曲線如下圖所示。此工作負載有兩個尖峰，為了處理這些尖峰，已佈建了資源容量 (以橙色線顯示)。用於此工作負載的資源和能源並非由需求曲線底下的區域表示，而是已佈建的容量底下的區域，因為這兩個尖峰必須用已佈建的容量處理。



需求曲線圖，內含兩個需要大量佈建容量的相異尖峰。

您可以使用緩衝或限流來修改需求曲線，並使尖峰趨緩，意即減少佈建容量和耗用的能源。在用戶端可以執行重試時實作限流。實作緩衝機制以儲存請求，並將處理的時間往後延遲。



限流對需求曲線和佈建容量的影響。

實作步驟

- 分析用戶端請求以確定如何予以回應。要考慮的問題包括：
 - 此請求是否可進行非同步處理？
 - 用戶端是否有重試能力？
- 如果用戶端有重試功能，您可以實作限流，以告知來源若目前無法處理請求，則應稍後再試。
 - 您可以使用 [Amazon API Gateway](#) 實作限流。
- 針對無法執行重試的用戶端，需要實作緩衝區使需求曲線扁平化。緩衝會延遲請求處理，讓以不同速率執行的應用程式能夠有效地通訊。緩衝型方法使用佇列或串流來接受生產者傳出的訊息。消費者可讀取訊息並進行處理，允許以符合取用者業務要求的速度運作訊息。
 - [Amazon Simple Queue Service \(Amazon SQS\)](#) 是一種受管服務，提供佇列，允許單一取用者讀取個別訊息。
 - [Amazon Kinesis](#) 可提供串流，允許許多取用者讀取相同訊息。
- 分析整體需求、變更率及所需的回應時間，以適當調整所需的調節或緩衝區大小。

資源

相關文件：

- [Amazon 入門 SQS](#)
- [使用佇列與訊息進行應用程式整合](#)
- [管理和監控工作負載中的API限流](#)
- [使用 API Gateway REST API 大規模限流分層多租戶](#)
- [使用佇列與訊息進行應用程式整合](#)

相關影片：

- [AWS re : Invent 2022 - 微服務的應用程式整合模式](#)
- [AWS re : Invent 2023 - 智慧節省：Amazon EC2成本最佳化策略](#)
- [AWS re : Invent 2023 - 鬆散耦合系統的進階整合模式和權衡](#)

軟體和架構

問題

- [SUS 3 如何利用軟體和架構模式，來支援您的永續性發展目標？](#)

SUS 3 如何利用軟體和架構模式，來支援您的永續性發展目標？

實施可執行負載順暢並保持已部署資源一致高使用率的模式，將資源消耗降至最低。由於使用者行為隨時間改變，元件可能會因缺乏使用而閒置。修改模式和架構來整合未充分利用的元件，提高整體使用率。淘汰不再需要的元件。了解工作負載元件的效能，並最佳化消耗最多資源的元件。留意客戶用來存取服務的裝置，並實施盡量減少裝置升級需求的模式。

最佳實務

- [SUS03-BP01 最佳化非同步與排程任務的軟體和架構](#)
- [SUS03-BP02 移除或重構使用量低或完全未使用的工作負載元件](#)
- [SUS03-BP03 最佳化程式碼中耗用最多時間或資源的區域](#)
- [SUS03-BP04 最佳化對裝置和設備的影響](#)
- [SUS03-BP05 使用最能支援資料存取和儲存模式的軟體模式和架構](#)

SUS03-BP01 最佳化非同步與排程任務的軟體和架構

使用有效率的軟體和架構模式 (例如佇列驅動)，讓所部署的資源一直保持高使用率。

常見的反模式：

- 在雲端工作負載中過度佈建資源以滿足未預料到的突增需求。
- 您的架構未透過傳訊元件將非同步訊息的傳送者與接受者分離。

建立此最佳實務的優勢：

- 高效率的軟體和架構模式可盡量減少工作負載中的未使用資源，並改善整體效率。
- 您可以將非同步訊息的處理與接收分開擴展。
- 透過傳訊元件，可用性要求會比較寬鬆，不用太多資源即可滿足。

未建立此最佳實務時的曝險等級：中

實作指引

使用有效率的架構模式 (例如[事件驅動架構](#))，以便能平均地使用元件，並盡量避免工作負載過度佈建。使用高效率的架構模式可盡量地讓閒置資源不會因為需求隨時間發生變化而有乏人問津的情形。

了解工作負載元件的要求，並採用能夠提升整體資源使用率的架構模式。淘汰不再需要的元件。

實作步驟

- 分析工作負載需求以確定如何回應。
- 如果請求或任務不需要同步回應，請使用佇列驅動的架構和自動擴展工作節點，以將使用率最大化。以下是您可能會考慮使用佇列驅動架構的一些範例：

佇列機制	Description
AWS Batch 任務佇列	AWS Batch 任務會提交至任務佇列並停留其中，直到能夠排定在運算環境中執行為止。
Amazon Simple Queue Service 和 Amazon EC2 Spot 執行個體	搭配使用 Amazon SQS 與 Spot 執行個體即可建置能容錯且有效率的架構。

- 對於可以隨時處理的佇列或任務，請使用排程機制來批次處理任務，以提升效率。以下是 AWS 上排程機制的幾個範例：

排程機制	Description
Amazon EventBridge 排程器	Amazon EventBridge 的一項功能，可讓您大規模地建立、執行和管理已排程任務。
AWS Glue 時間型排程	在 AWS Glue 中為編目程式和任務定義時間型排程。
Amazon Elastic Container Service (Amazon ECS) 已排程任務	Amazon ECS 支援建立排程任務。排程任務使用 Amazon EventBridge 規則來安排程或在回應 EventBridge 事件時執行任務。
Instance Scheduler	設定 Amazon EC2 和 Amazon Relational Database Service 執行個體的開始和停止排程。

- 如果您的架構中使用輪詢和 Webhook 機制，請將其取代為事件。使用[事件驅動的架構](#)來建置高效率的工作負載。
- 利用 [AWS 上的無伺服器](#)來消除過度佈建的基礎設施。
- 將架構的個別元件調整為適當大小，避免閒置資源等待輸入。
 - 您可以使用 [AWS Cost Explorer 中的適當調整大小建議](#)或 [AWS Compute Optimizer](#) 識別適當調整大小的機會。
 - 如需詳細資訊，請參閱[適當調整大小：佈建執行個體以符合工作負載](#)。

資源

相關文件：

- [什麼是 Amazon Simple Queue Service？](#)
- [什麼是 Amazon MQ？](#)
- [根據 Amazon SQS 進行擴展](#)
- [什麼是 AWS Step Functions？](#)
- [什麼是 AWS Lambda？](#)
- [搭配使用 AWS Lambda 與 Amazon SQS](#)
- [什麼是 Amazon EventBridge？](#)

- [使用 REST API 管理非同步工作流程](#)

相關影片：

- [AWS re:Invent 2023 - 將旅程導向無伺服器事件驅動架構](#)
- [AWS re:Invent 2023 - 針對事件驅動架構和網域驅動設計使用無伺服器](#)
- [AWS re:Invent 2023 - 使用 Amazon EventBridge 提升事件驅動型模式](#)
- [AWS re:Invent 2023 - 永續架構：過去、現在和未來](#)
- [非同步訊息模式 | AWS 事件](#)

相關範例：

- [具有 AWS Graviton 處理器和 Amazon EC2 Spot 執行個體的事件驅動架構](#)

SUS03-BP02 移除或重構使用量低或完全未使用的工作負載元件

移除未使用且不再需要的元件，並重構使用率低的元件，以盡可能避免工作負載中的浪費。

常見的反模式：

- 您未定期檢查個別工作負載元件的使用率水準。
- 您未查看並分析 AWS 適當調整大小的工具 (例如 [AWS Compute Optimizer](#)) 所提供的建議。

建立此最佳實務的優勢：移除未使用的元件可盡量避免浪費，並改善雲端工作負載的整體效率。

未建立此最佳實務時的曝險等級：中

實作指引

雲端工作負載中未使用或未充分利用的元件，會耗用不必要的運算、儲存或網路資源。移除或重構這些元件，可直接減少浪費，並提升雲端工作負載的整體效率。有一個迭代改進程序可由需求的變更或新雲端服務的發行來啟動。例如，[AWS Lambda](#) 函數執行時間的大幅下降可能表示必須縮減記憶體大小。此外，隨著 AWS 發行新的服務和功能，工作負載的最佳服務與架構可能會改變。

持續監控工作負載活動，並找機會改善個別元件的使用率水準。藉由移除閒置元件和執行適當調整大小的活動，您將可用最少的雲端資源達到業務要求。

實作步驟

- 清查您的 AWS 資源：建立 AWS 資源的庫存清單。在 AWS 中，您可以開啟 [AWS 資源總管](#) 探索和組織您的 AWS 資源。如需詳細資訊，請參閱 [AWS re:Invent 2022 - 在 AWS 上如何大規模管理資源和應用程式](#)。
- 監控使用率：監控及擷取工作負載關鍵元件的使用率指標 (例如 [Amazon CloudWatch 指標](#) 中的 CPU 使用率、記憶體使用率或網路輸送量)。
- 識別未使用的元件：識別架構中完全未使用或使用率不足的元件。
 - 對於穩定的工作負載，請定期檢查 AWS 適當調整大小的工具 (例如 [AWS Compute Optimizer](#))，以識別閒置、未使用或未充分利用的元件。
 - 對於暫時性工作負載，請評估使用率指標以識別閒置、未使用或未充分利用的元件。
- 移除未使用的元件：淘汰不再需要的元件和關聯的資產 (例如 Amazon ECR 映像)。
 - [自動清理 Amazon ECR 中未曾使用的映像](#)
 - [使用 AWS Config 和 AWS Systems Manager 刪除未曾使用的 Amazon Elastic Block Store \(Amazon EBS\) 磁碟區](#)
- 重構未充分利用的元件：重構或整合未充分利用的元件與其他資源，以提高利用效率。例如，您可將多個小資料庫佈建至單一 [Amazon RDS](#) 資料庫執行個體，而不要在未充分利用的個別執行個體上執行資料庫。
- 評估改進：了解[您的工作負載佈建以完成工作單位的資源](#)。使用此資訊來評估移除或重構元件所達成的改進。
 - [使用永續性代理指標測量和追蹤雲端效率，第 I 部分：什麼是代理指標？](#)
 - [使用永續性代理指標來測量和追蹤雲端效率，第 II 部分：建立指標管道](#)

資源

相關文件：

- [AWS Trusted Advisor](#)
- [什麼是 Amazon CloudWatch？](#)
- [適當調整大小：佈建執行個體以符合工作負載](#)
- [使用精簡化建議將您的成本最佳化](#)

相關影片：

- [AWS re:Invent 2023 - 容量、可用性、成本效率：挑選三項](#)

SUS03-BP03 最佳化程式碼中耗用最多時間或資源的區域

最佳化您的架構不同元件中執行的程式碼，將資源用量降至最低，同時發揮最大效能。

常見的反模式：

- 您略過資源用量的程式碼最佳化。
- 您通常藉由增加資源來回應效能問題。
- 您的程式碼審查和開發程序未追蹤效能變更。

建立此最佳實務的優勢：使用有效率的程式碼可將資源用量壓到最低，並改善效能。

未建立此最佳實務時的曝險等級：中

實作指引

請務必檢查各個功能領域 (包括雲端架構應用程式的程式碼)，以最佳化其資源用量和效能。持續監控您的工作負載在建置環境和生產環境中的效能，並找機會改進資源用量特別高的程式碼片段。採用定期審查程序，在您的程式碼內識別低效使用資源的錯誤或反模式。使用簡單有效的演算法為您的使用案例產生相同結果。

實作步驟

- 使用高效率的程式設計語言：使用適用於工作負載的高效率作業系統和程式設計語言。如需有關高性能程式設計語言 (包括 Rust) 的詳細資訊，請參閱 [Rust 的永續性](#)。
- 使用 AI 編碼配套：考慮使用 AI 編碼配套 (例如 [Amazon Q Developer](#)) 以有效率地撰寫程式碼。
- 自動執行程式碼檢閱：在擬定工作負載時採用自動化程式碼檢閱程序，以改善品質並識別錯誤和反模式。
 - [使用 Amazon CodeGuru Reviewer 自動進行程式碼審核](#)
 - [使用 Amazon CodeGuru 偵測並行錯誤](#)
 - [使用 Amazon CodeGuru 提升 Python 應用程式的程式碼品質](#)
- 使用程式碼分析工具：使用程式碼分析工具來識別程式碼中使用最多時間或資源的區域，作為最佳化目標。
 - [透過 Amazon CodeGuru Profiler 降低組織的碳足跡](#)
 - [透過 Amazon CodeGuru Profiler 了解 Java 應用程式中的記憶體用量](#)

- [透過 Amazon CodeGuru Profiler 改善客戶體驗並降低成本](#)
- 監控和最佳化：使用持續監控資源來識別資源需求高或組態不夠好的元件。
 - 將需要大量運算資源的演算法取代為會產生相同結果、但更簡單有效率的版本。
 - 移除不必要程式碼，例如排序和格式化。
- 使用程式碼重構或轉換：探索用於應用程式維護和升級的 [Amazon Q 程式碼轉換](#) 的可能性。
 - [使用 Amazon Q 程式碼轉換升級語言版本](#)
 - [AWS re:Invent 2023 - 使用 Amazon Q 程式碼轉換自動執行應用程式升級與維護](#)

資源

相關文件：

- [什麼是 Amazon CodeGuru Profiler？](#)
- [FPGA 執行個體](#)
- [在 AWS 上建置的工具中的 AWS SDK](#)

相關影片：

- [使用 Amazon CodeGuru Profiler 改善程式碼效率](#)
- [使用 Amazon CodeGuru 自動執行程式碼檢閱和應用程式效能建議](#)

SUS03-BP04 最佳化對裝置和設備的影響

了解您的架構中使用的裝置和設備，並使用策略降低其用量。這樣可以盡量減輕對雲端工作負載的整體環境影響。

常見的反模式：

- 您忽略了客戶使用的裝置所受到的環境影響。
- 您手動管理及更新客戶所使用的資源。

建立此最佳實務的優勢：實作為客戶裝置最佳化的軟體模式和功能，可降低雲端工作負載的整體環境影響。

未建立此最佳實務時的曝險等級：中

實作指引

實作為客戶裝置最佳化的軟體模式和功能，可透過數種方式降低環境影響：

- 實作具回溯相容性的新功能，可減少硬體更換的數量。
- 最佳化應用程式以在裝置上有效執行，有助於降低能源耗用量及延長電池使用壽命 (若是由電池供電)。
- 最佳化裝置的應用程式也可減少網路上的資料傳輸。

了解您的架構中使用的裝置和設備、其預期生命週期，以及更換這些元件的影響。實作適當的軟體模式和功能，以盡可能減少裝置能源耗用量，以及客戶更換裝置和手動加以升級的需求。

實作步驟

- 執行清查：清查您的架構中使用的裝置。裝置可以是行動裝置、平板裝置、IOT 裝置、智慧電燈，或是工廠的智慧裝置。
- 使用節能裝置：考慮在您的架構中使用節能裝置。在不使用時，使用裝置上的電源管理組態進入低功耗模式。
- 執行高效應用程式：最佳化裝置上執行的應用程式：
 - 採用在背景執行任務之類的策略來降低能源耗用量。
 - 在建置承載時考慮網路頻寬和延遲，並實施可協助應用程式在低頻寬、高延遲連結上良好運作的功能。
 - 將承載和檔案轉換為裝置所需的最佳化格式。例如，您可以使用 [Amazon Elastic Transcoder](#) 或 [AWS Elemental MediaConvert](#) 將大型高品質數位媒體檔案轉換為使用者可以在行動裝置、平板電腦、Web 瀏覽器和外接電視機上播放的格式。
 - 在伺服器端執行需要大量運算的活動 (例如影像渲染)，或使用應用程式串流來改善舊裝置的使用者體驗。
 - 對輸出進行分段和分頁，特別是對於互動式工作階段，以管理承載並限制本機儲存要求。
- 鼓勵供應商參與其中：與使用永續性材料並提供供應鏈和環境認證透明度的裝置供應商合作。
- 使用空中 (OTA) 更新：使用自動化空中 (OTA) 機制將更新部署至一或多個裝置。
 - 您可以使用 [CI/CD 管道](#) 更新行動應用程式。
 - 您可以使用 [AWS IoT Device Management](#) 從遠端大規模管理連網裝置。
- 使用受管 Device Farm：若要測試新功能和更新，請使用具有代表性硬體集的受管 Device Farm，並迭代開發以最大化支援的裝置。如需詳細資訊，請參閱[SUS06-BP05 使用受管 Device Farm 進行測試](#)。

- 繼續監控和改善：追蹤裝置的能源使用情況，以識別需要改善的區域。使用新技術或最佳實務來增強這些裝置對環境的影響。

資源

相關文件：

- [什麼是 AWS Device Farm？](#)
- [WorkSpaces 應用程式文件](#)
- [NICE DCV](#)
- [在執行 FreeRTOS 的裝置上更新韌體的 OTA 教學](#)
- [最佳化您的 IoT 裝置以實現環境永續性](#)

相關影片：

- [AWS re:Invent 2023 - 使用 AWS Device Farm 改善行動裝置和 Web 應用程式品質](#)

SUS03-BP05 使用最能支援資料存取和儲存模式的軟體模式和架構

了解資料在工作負載中的使用方式、使用者的使用方式、傳輸方式以及儲存方式。使用最能支援資料存取和儲存的軟體模式與架構，以盡可能減少支援工作負載所需的運算、聯網和儲存資源。

常見的反模式：

- 您假設所有工作負載具有類似的資料儲存和存取模式。
- 您只使用一個存儲層 – 假設所有工作負載都適合該層。
- 您假設資料存取模式不會隨著時間改變。
- 您的架構支援潛在的高資料存取爆量，這會導致資源在大部分的時間處於閒置狀態。

建立此最佳實務的優勢：根據資料存取和儲存模式選取及最佳化您的架構，有助於降低開發複雜性並提升整體使用率。了解何時使用全域表、資料分割和快取將協助您降低營運負擔，並根據您的工作負載需求進行擴展。

未建立此最佳實務時的曝險等級：中

實作指引

為了改善長期工作負載永續性，請使用支援工作負載的資料存取和儲存特性的架構模式。這些模式可協助您有效率地擷取和處理資料。例如，搭配針對獨特分析使用案例而最佳化的專用服務來使用 [AWS 上的現代資料架構](#)。這些架構模式有利於高效率資料處理並降低資源用量。

實作步驟

- 了解資料特性：分析您的資料特性和存取模式，以識別雲端資源的正確組態。應考量的重要特性包括：
 - 資料類型：結構化、半結構化、非結構化
 - 資料成長：有界限、無界限
 - 資料耐用性：持續性、暫時性、臨時
 - 存取模式：讀取或寫入、更新頻率、尖峰或一致
- 使用最佳架構模式：使用最能支援資料存取和儲存模式的架構模式。
 - [啟用資料持續性的模式](#)
 - [開始建構吧！現代資料架構](#)
 - [AWS 上的資料庫：使用正確的工具完成任務](#)
- 使用專用服務：使用符合用途的技術。
 - 利用可原生處理壓縮資料的技術。
 - [Athena 壓縮支援檔案格式](#)
 - [在 AWS Glue 中的 ETL 輸入與輸出格式選項](#)
 - [使用 Amazon Redshift 從 Amazon S3 載入壓縮資料檔案](#)
 - 使用專用[分析服務](#)進行架構中的資料處理。如需有關 AWS 專用分析服務的詳細資訊，請參閱 [AWS re:Invent 2022 - 在 AWS 上建置現代資料架構](#)。
 - 使用最能支援您主導查詢模式的資料庫引擎。管理您的資料庫索引以確保高效率查詢。如需進一步詳細資訊，請參閱 [AWS 資料庫](#)和 [AWS re:Invent 2022 - 使用專用資料庫將應用程式現代化](#)。
- 資料傳輸最小化：選取可減少架構中消耗的網路容量的網路協定。

資源

相關文件：

- [使用 Amazon Redshift 從單欄資料格式複製](#)
- [在 Firehose 中轉換您的輸入記錄格式](#)

- [轉換為單欄格式，提高 Amazon Athena 的查詢效能](#)
- [在 Amazon Aurora 上使用 Performance Insights 監控資料庫負載](#)
- [在 Amazon RDS 上使用 Performance Insights 監控資料庫負載](#)
- [Amazon S3 Intelligent-Tiering 儲存類別](#)
- [使用 Amazon DynamoDB 建置 CQRS 事件存放區](#)

相關影片：

- [AWS re:Invent 2022 - 在 AWS 上建置資料網格架構](#)
- [AWS re:Invent 2023 - 深入探索 Amazon Aurora 及其創新](#)
- [AWS re:Invent 2023 - 提高 Amazon EBS 效率並更具成本效益](#)
- [AWS re:Invent 2023 - 使用 Amazon S3 最佳化儲存價格和效能](#)
- [AWS re:Invent 2023 - 在 Amazon S3 上建置和最佳化資料湖](#)
- [AWS re:Invent 2023 - 使用 Amazon EventBridge 提升事件驅動型模式](#)

相關範例：

- [AWS 專用資料庫研討會](#)
- [AWS 現代資料架構 Immersion Day](#)
- [在 AWS 上建置資料網格](#)

資料

問題

- [SUS 4 如何利用資料管理政策和模式，來支援您的永續性目標？](#)

SUS 4 如何利用資料管理政策和模式，來支援您的永續性目標？

實作資料管理實務來減少支援工作負載所需的佈建儲存，以及減少為了使用它所需的資源。了解您的資料，並使用更有效支援資料業務價值及其使用方式的儲存技術和組態。當需求減少時，將資料循環到效率較高、效能較低的儲存，並刪除不再需要的資料。

最佳實務

- [SUS04-BP01 實作資料分類政策](#)

- [SUS04-BP02 使用支援資料存取和儲存模式的技術](#)
- [SUS04-BP03 使用政策來管理資料集的生命週期](#)
- [SUS04-BP04 使用彈性和自動化來擴展區塊儲存或檔案系統](#)
- [SUS04-BP05 移除不需要或多餘的資料](#)
- [SUS04-BP06 使用共用檔案系統或儲存體存取通用資料](#)
- [SUS04-BP07 盡可能減少跨網路的資料移動](#)
- [SUS04-BP08 僅在難以重新建立時才備份資料](#)

SUS04-BP01 實作資料分類政策

將資料分類以了解其對業務成果的關鍵性，並選擇適當的節能儲存層來儲存資料。

常見的反模式：

- 您未以正在處理或已儲存的類似特性來識別資料資產 (例如敏感性、業務關鍵性或法規要求)。
- 您未實作資料目錄以清查資料資產。

建立此最佳實務的優勢：實作資料分類政策，可讓您確認最節能的資料儲存層。

未建立此最佳實務時的曝險等級：中

實作指引

資料分類涉及識別組織擁有或營運的資訊系統中正在處理和儲存的資料類型。此外涉及確認資料的關鍵性，以及資料損毀、遺失或誤用可能造成的影響。

若要實作資料分類政策，請從資料的情境使用採取逆向思維，並建立適當的分類機制，將指定資料集的關鍵性程度納入組織操作的考量中。

實作步驟

- 執行資料清查：對您工作負載現有的各種資料類型執行清查。
- 將資料分組：根據組織面臨的風險，確定資料的關鍵性、機密性、完整性和可用性。使用這些要求，將資料分組為您採用的其中一個資料分類層。範例請見[分類資料及保護新創公司的四個簡單步驟](#)。
- 定義資料分類層級和政策：針對每個資料群組，定義資料分類層級 (例如公開或機密) 和處理政策。相應地標記資料。如需有關資料分類類別的詳細資訊，請參閱《資料分類》白皮書。

- 定期審查：定期審查與稽核您的環境，以尋找未標記及未分類的資料。使用自動化功能來識別這些資料，並適當地分類和標記資料。範例請見 [AWS Glue中的資料型錄和編目程式](#)。
- 建立資料型錄：建立提供稽核及管控能力的資料型錄。
- 文件：記錄每個資料類別的資料分類政策和處理程序。

資源

相關文件：

- [利用 AWS 雲端 以支援資料分類](#)
- [來自 的標籤政策 AWS Organizations](#)

相關影片：

- [AWS re : Invent 2022 - 在 上啟用資料治理的靈活性 AWS](#)
- [AWS re : Invent 2023 - 資料保護和 AWS 儲存的彈性](#)

SUS04-BP02 使用支援資料存取和儲存模式的技術

使用最能支援您的資料存取和儲存方式的儲存技術，以在支援工作負載的同時，也將佈建的資源降至最低。

常見的反模式：

- 您假設所有工作負載具有類似的資料儲存和存取模式。
- 您只使用一個存儲層 – 假設所有工作負載都適合該層。
- 您假設資料存取模式不會隨著時間改變。

建立此最佳實務的優勢：根據資料存取和儲存模式來選取及最佳化您的儲存技術，可協助您降低達成商業需求所需的雲端資源，並改善雲端工作負載的整體效率。

未建立此最佳實務時的曝險等級：低

實作指引

選取最適合您的存取模式的儲存解決方案，或者考慮變更存取模式，以符合儲存解決方案，從而達到最大的效能效率。

實作步驟

- 評估資料和存取特性：評估您的資料特性和存取模式，以收集儲存需求的重要特性。應考量的重要特性包括：
 - 資料類型：結構化、半結構化、非結構化
 - 資料成長：有界限、無界限
 - 資料耐用性：持續性、暫時性、臨時
 - 存取模式：讀取或寫入、頻率、尖峰或一致
- 選擇適當的儲存技術：將資料遷移至支援您的資料特性和存取模式的適當儲存技術。以下提供 AWS 儲存技術及其重要特性的一些範例：

Type	技術	重要特性
物件儲存	Amazon S3	一項物件儲存服務，具有不受限的可擴展性、高可用性，以及多個可存取性選項。對 Amazon S3 輸入和存取物件時，可以使用 Transfer Acceleration 或 Access Points 之類的服務來支援您的位置、安全需求和存取模式。
封存儲存	Amazon Glacier	針對資料封存而建置的 Amazon S3 儲存類別。
共用檔案系統	Amazon Elastic File System (Amazon EFS)	可供多種類型的運算解決方案存取的可掛載檔案系統。Amazon EFS 會自動增長及縮減儲存體，藉以進行效能最佳化而提供一致的低延遲。
共用檔案系統	Amazon FSx	建置於最新的 AWS 運算解決方案之上，用以支援四個常用的檔案系統：NetApp ONTAP、OpenZFS、Windows File Server 和 Lustre。Amazon FSx 的 延

Type	技術	重要特性
		<u>遲、輸送量和 IOPS</u> 會隨著檔案系統而不同，當您為工作負載需求選取適當的檔案系統時，應予以考量。
區塊儲存	Amazon Elastic Block Store (Amazon EBS)	專為 Amazon Elastic Compute Cloud (Amazon EC2) 設計的可擴展、高效能區塊儲存服務。Amazon EBS 包含支援 SSD 的儲存，適用於交易型、IOPS 密集型工作負載，以及適用於輸送量密集型工作負載的支援 HDD 的儲存。
關聯式資料庫	Amazon Aurora 、 Amazon RDS 、 Amazon Redshift	旨在支援 ACID (單元性、一致性、隔離行為、持續性) 交易，並維護參考完整性和強大的資料一致性。許多傳統應用程式、企業資源規劃 (ERP)、客戶關係管理 (CRM) 和電子商務系統都使用關聯式資料庫來儲存資料。
鍵值資料庫	Amazon DynamoDB	已針對常見的存取模式進行最佳化，通常用於儲存和擷取大量資料。高流量 Web 應用程式、電子商務系統和遊戲應用程式是鍵值資料庫的典型使用案例。

- 自動執行儲存空間配置：對於固定大小的儲存系統 (例如 Amazon EBS 或 Amazon FSx)，請監控可用儲存空間，並且在接近閾值時自動執行儲存空間配置。您可以利用 Amazon CloudWatch 來收集及分析 [Amazon EBS](#) 和 [Amazon FSx](#) 的不同指標。
- 選擇適當的儲存類別：選擇適當的資料儲存類別。
 - Amazon S3 儲存類別可以在物件層級設定。單一儲存貯體可以包含儲存於所有儲存類別的物件。

- 您可以使用 [Amazon S3 生命週期政策](#) 在儲存類別之間自動轉換物件或移除資料，而無須進行任何應用程式變更。在考量這些儲存機制時，您通常需要在資源效率、存取延遲與可靠性之間做出取捨。

資源

相關文件：

- [Amazon EBS 磁碟區類型](#)
- [Amazon EC2 執行個體儲存體](#)
- [Amazon S3 Intelligent-Tiering](#)
- [Amazon EBS I/O 特性](#)
- [使用 Amazon S3 儲存類別](#)
- [什麼是 Amazon Glacier？](#)

相關影片：

- [AWS re:Invent 2023 - 提高 Amazon EBS 效率並更具成本效益](#)
- [AWS re:Invent 2023 - 使用 Amazon S3 最佳化儲存價格和效能](#)
- [AWS re:Invent 2023 - 在 Amazon S3 上建置和最佳化資料湖](#)
- [AWS re:Invent 2022 - 在 AWS 上建置現代資料架構](#)
- [AWS re:Invent 2022 - 使用專用資料庫將應用程式現代化](#)
- [AWS re:Invent 2022 - 在 AWS 上建置資料網格架構](#)
- [AWS re:Invent 2023 - 深入探索 Amazon Aurora 及其創新](#)
- [AWS re:Invent 2023 - 使用 Amazon DynamoDB 的進階資料建模](#)

相關範例：

- [Amazon S3 範例](#)
- [AWS 專用資料庫研討會](#)
- [專為開發人員打造的資料庫](#)
- [AWS 現代資料架構 Immersion Day](#)

- [在 AWS 上建置資料網格](#)

SUS04-BP03 使用政策來管理資料集的生命週期

管理所有資料的生命週期並自動執行刪除，將工作負載所需的儲存總量降至最低。

常見的反模式：

- 您手動刪除資料。
- 您未刪除任何工作負載資料。
- 您未根據資料的保留和存取要求，將資料移至更節能的儲存層。

建立此最佳實務的優勢：使用資料生命週期政策可確保工作負載中的資料會以有效率的方式存取和保留。

未建立此最佳實務時的曝險等級：中

實作指引

資料集在其生命週期內，通常會有不同的保留和存取要求。例如，應用程式可能需要在一段時間內頻繁存取某些資料集。這段時間過後，便不會頻繁存取這些資料集。為了隨時間提高資料儲存和運算的效率，請實施生命週期政策，這些政策主要定義如何隨時間處理資料的相關規則。

有了生命週期組態規則後，便能指示特定儲存服務將資料集轉移至更節能的儲存層、將其封存，或加以刪除。此做法可將作用中資料儲存和擷取操作降至最低程度，進而降低能耗量。此外，封存或刪除過時資料等做法，支援法規遵循和資料控管。

實作步驟

- 使用資料分類：[分類工作負載中的資料集](#)。
- 定義處理規則：定義每個資料類別的處理程序。
- 啟用自動化：設定自動生命週期政策以強制執行生命週期規則。下面幾個範例會說明如何為不同的 AWS 儲存服務設定自動化生命週期政策：

儲存服務	如何設定自動化生命週期原則
Amazon Simple Storage Service (Amazon S3)	您可以使用 Amazon S3 生命週期 ，以在物件的整個生命週期中管理物件。如果存取模

儲存服務	如何設定自動化生命週期原則
	式不明、會變化或是無法預測，則可以使用 Amazon S3 Intelligent-Tiering ，讓其監控存取模式，並自動將未存取的物件移至成本較低的存取層。可以利用 Amazon S3 Storage Lens 指標，找出生命週期管理中的最佳化機會和差距。
Amazon Elastic Block Store	您可以使用 Amazon Data Lifecycle Manager 自動建立、保留和刪除 Amazon EBS 快照與 Amazon EBS 支援的 AMI。
Amazon Elastic File System	Amazon EFS 生命週期管理 會自動管理檔案系統的檔案儲存。
Amazon Elastic Container Registry	Amazon ECR 生命週期政策 會根據存留時間長短或計數讓映像過期，藉此自動清理容器映像。
AWS Elemental MediaStore	您可以使用 物件生命週期政策 來管控物件儲存在 MediaStore 容器內的時間長度。

- 刪除未使用的資產：請刪除已超過保留期間的未使用磁碟區、快照和資料。使用原生服務功能 (例如 [Amazon DynamoDB Time To Live](#) 或 [Amazon CloudWatch 日誌保留](#)) 來執行刪除作業。
- 彙總和壓縮：根據生命週期規則，在適用的情況下彙總和壓縮資料。

資源

相關文件：

- [使用 Amazon S3 儲存類別分析最佳化 Amazon S3 生命週期規則](#)
- [使用 AWS Config 規則 評估資源](#)

相關影片：

- [AWS re:Invent 2021 - Amazon S3 最佳化儲存支出的生命週期最佳實務](#)
- [AWS re:Invent 2023 - 使用 Amazon S3 最佳化儲存價格和效能](#)

- [使用 Amazon S3 生命週期來簡化資料生命週期並最佳化儲存成本](#)
- [使用 Amazon S3 Storage Lens 降低儲存成本](#)

SUS04-BP04 使用彈性和自動化來擴展區塊儲存或檔案系統

隨著資料的增長使用彈性和自動化擴充區塊儲存或檔案系統，以盡可能縮小整體的已佈建儲存。

常見的反模式：

- 您為了日後的需求購買大型區塊儲存或檔案系統。
- 您過度佈建檔案系統的每秒輸入和輸出操作（IOPS）。
- 您未監控資料磁碟區的使用率。

建立此最佳實務的優勢：盡可能減少儲存系統的過度佈建可減少閒置資源，並改善工作負載的整體效率。

未建立此最佳實務時的曝險等級：中

實作指引

使用適合工作負載的大小分配、輸送量和延遲，建立區塊儲存和檔案系統。隨著資料的增長使用彈性和自動化擴充區塊儲存或檔案系統，而無需過度佈建這些儲存服務。

實作步驟

- 對於 [Amazon EBS](#) 等固定大小的儲存體，請確認您正在監控使用的儲存體數量與整體儲存體大小的比較，並在可能的情況下建立自動化，以在達到閾值時增加儲存體大小。
- 使用彈性磁碟區和受管區塊資料服務，以在持久性資料增長時自動分配額外的儲存空間。例如，您可以使用 [Amazon EBS Elastic Volumes](#) 來變更磁碟區大小、磁碟區類型或調整 Amazon EBS 磁碟區的效能。
- 為您的檔案系統選擇適當的儲存類別、效能模式和輸送量模式以因應商業需求 (勿過量)。
 - [Amazon EFS 效能](#)
 - [Linux 執行個體上的 Amazon EBS 磁碟區效能](#)
- 設定資料磁碟區的目標使用率水準，並調整超出預期範圍的磁碟區大小。
- 根據資料調整唯讀磁碟區的大小。
- 將資料遷移到物件存放區，避免從區塊儲存的固定磁碟區大小佈建多餘容量。
- 定期審查彈性磁碟區和檔案系統以終止閒置磁碟區，並縮減過度佈建的資源以符合目前資料大小。

資源

相關文件：

- [在調整EBS磁碟區大小後擴展檔案系統](#)
- [使用 Amazon EBS Elastic Volumes 修改磁碟區](#)
- [Amazon FSx 文件](#)
- [什麼是 Amazon Elastic File System ?](#)

相關影片：

- [深入探索 Amazon EBS Elastic Volumes](#)
- [Amazon EBS和 Snapshot 最佳化策略可提供更好的效能和節省成本](#)
- [使用最佳實務來最佳化 Amazon EFS 的成本和效能](#)

SUS04-BP05 移除不需要或多餘的資料

移除不需要或多餘的資料，以盡量降低儲存資料集時所需的儲存資源。

常見的反模式：

- 您複製可以輕鬆取得或重新建立的資料。
- 您備份所有資料，而不考慮該資料是否重要。
- 您只會不定期地刪除資料、在發生營運事件時刪除資料，或完全不刪除資料。
- 您重複儲存資料，而不理會儲存服務的耐用性。
- 您在沒有任何商務理由的情況下開啟 Amazon S3 版本控制。

建立此最佳實務的優勢：移除不需要的資料會降低工作負載所需的儲存大小，以及工作負載環境所受到的影響。

未建立此最佳實務時的曝險等級：中

實作指引

當移除不需要的備援資料集時，即可縮減儲存成本和環境耗用面積。這種做法也可讓運算更有效率，因為運算資源只會處理重要資料，而不是不需要的資料。請自動刪除不需要的資料。使用會在檔案層級和區塊層級刪除重複資料的技術。使用服務原生資料複寫和備援功能。

實作步驟

- 評估公有資料集：評估您是否可以藉由使用 [AWS Data Exchange](#) 和 [AWS 上的開放資料](#) 中現有公開提供的資料集，以避免儲存資料。
- 刪除重複資料：使用可在區塊和物件層級刪除重複資料的機制。下面幾個範例會說明如何在 AWS 上刪除重複資料：

儲存服務	重複資料刪除機制
Amazon Simple Storage Service (Amazon S3)	使用 AWS Lake Formation FindMatches ，透過新的 FindMatches ML Transform 來尋找整個資料集內的相符記錄 (包括沒有識別碼的記錄)。
Amazon FSx	在適用於 Windows 的 Amazon FSx 上使用 重複資料刪除 。
Amazon Elastic Block Store 快照	快照為遞增備份，這表示只會儲存您上次執行裝置快照後發生變更的區塊。

- 使用生命週期政策：使用生命週期政策來自動刪除不需要的資料。使用原生服務功能 (例如 [Amazon DynamoDB Time To Live](#)、[Amazon S3 Lifecycle](#) 或 [Amazon CloudWatch 日誌保留](#)) 執行刪除作業。
- 使用資料虛擬化：使用 AWS 上的資料虛擬化功能以保有其來源處的資料，並避免資料重複。
 - [AWS 上的雲端原生資料虛擬化](#)
 - [使用 Amazon Redshift 資料共用來最佳化資料模式](#)
- 使用增量備份：使用可以進行增量備份的備份技術。
- 使用原生耐久性：利用 [Amazon S3](#) 的耐久性和 [Amazon EBS 的複寫功能](#) 來滿足耐久性目標，而非利用自我管理的技術 (例如獨立硬碟冗餘陣列 (RAID))。
- 使用有效率的記錄：集中日誌和追蹤資料、刪除重複的日誌項目，並建立根據需要微調詳細程度的機制。
- 使用有效率的快取：僅在合理的情況下預先填入快取。
- 建立快取監控和自動化，據以調整快取大小。
- 移除舊版資產：推送工作負載新版本時，從物件存放區和邊緣快取移除過時的部署和資產。

資源

相關文件：

- [變更 CloudWatch Logs 中的日誌資料保留](#)
- [Amazon FSx for Windows File Server 上的重複資料刪除](#)
- [Amazon FSx for ONTAP 的功能，包括重複資料刪除](#)
- [使 Amazon CloudFront 上的檔案失效](#)
- [使用 AWS Backup 來備份和還原 Amazon EFS 檔案系統](#)
- [什麼是 Amazon CloudWatch Logs？](#)
- [在 Amazon RDS 上使用備份](#)
- [使用 AWS Lake Formation 整合及刪除重複資料集](#)

相關影片：

- [Amazon Redshift 資料共用使用案例](#)

相關範例：

- [我要如何使用 Amazon Athena 分析 Amazon S3 伺服器存取日誌？](#)

SUS04-BP06 使用共用檔案系統或儲存體存取通用資料

採用共用檔案系統或儲存體以避免資料重複，並且讓工作負載有更高效率的基礎設施。

常見的反模式：

- 您為每個用戶端佈建儲存體。
- 您未從非作用中用戶端卸離資料磁碟區。
- 您未提供跨平台和系統的儲存體存取。

建立此最佳實務的優勢：使用共用檔案系統或儲存裝置，無須複製資料即可與一或多個取用者共用資料。這有助於減少工作負載所需的儲存資源。

未建立此最佳實務時的曝險等級：中

實作指引

如果有多個使用者或應用程式在存取相同的資料集，則務必使用共用儲存技術，讓您的工作負載使用高效的基礎設施。共用儲存技術提供了集中儲存和管理資料的位置，可避免資料重複。此外也可強制執行跨不同系統的資料一致性。再者，共用儲存技術可讓您更有效率地使用運算能力，因為多個運算資源可同時平行存取及處理資料。

請在必要時才從這些共用儲存服務擷取資料，且應卸離未使用的磁碟區以釋出資源。

實作步驟

- 使用共用儲存體：當資料有多個取用者時，將資料遷移到共用儲存體。以下是 AWS 共用儲存技術的一些範例：

儲存選項	使用情況
Amazon EBS Multi-Attach	Amazon EBS Multi-Attach 可讓您將單一佈建 IOPS SSD (io1 或 io2) 磁碟區連接至在相同可用區域中的多個執行個體。
Amazon EFS	請參閱 何時應選擇 Amazon EFS 。
Amazon FSx	請參閱 選擇 Amazon FSx 檔案系統 。
Amazon S3	不需要檔案系統結構、且設計為使用物件儲存的應用程式，可使用 Amazon S3 做為可大規模擴展、耐久、低成本的物件儲存解決方案。

- 視需要擷取資料：僅在需要時才將資料複製到共用檔案系統或從中擷取資料。例如，您可以建立[Amazon S3 支援的 Amazon FSx for Lustre 檔案系統](#)，並僅將處理任務所需的資料子集載入至 Amazon FSx。
- 刪除不需要的資料：根據您的使用模式適度地刪除資料，如[SUS04-BP03 使用政策來管理資料集的生命週期](#)中所述。
- 卸離非作用中用戶端：將磁碟區與未積極使用它們的用戶端分開。

資源

相關文件：

- [將您的檔案系統連結到 Amazon S3 儲存貯體](#)

- [在無伺服器應用程式中對 AWS Lambda 使用 Amazon EFS](#)
- [Amazon EFS Intelligent-Tiering 最佳化變更存取模式的工作負載成本](#)
- [將 Amazon FSx 與內部部署資料儲存庫搭配使用](#)

相關影片：

- [使用 Amazon EFS 最佳化儲存成本](#)
- [AWS re:Invent 2023 - AWS 檔案儲存最新消息](#)
- [AWS re:Invent 2023 - Amazon Elastic File System 上適用於建構者和資料科學家的檔案儲存](#)

SUS04-BP07 盡可能減少跨網路的資料移動

使用共用檔案系統或物件儲存體存取通用資料，將支援工作負載資料移動所需的整體聯網資源降至最低。

常見的反模式：

- 無論資料使用者位於何處，您都將所有資料儲存在相同的 AWS 區域 中。
- 您未最佳化資料大小和格式，便將其移至網路。

建立此最佳實務的優勢：最佳化整個網路間的資料移動，可減少工作負載所需的整體聯網資源，並降低其環境影響。

未建立此最佳實務時的曝險等級：中

實作指引

要在您的組織移動資料，需要運算、聯網和儲存資源。使用相關技術盡可能減少資料移動，並改善工作負載的整體效率。

實作步驟

- 使用鄰近性：[選取工作負載的區域時](#)，可將區域與資料或使用者的鄰近性視為決策因素。
- 分區服務：對區域性使用的服務進行分區，以便將區域專屬的資料存放在使用它的區域內。
- 使用有效率的檔案格式：使用有效率的檔案格式 (例如 Parquet 或 ORC)，並在透過網路移動資料之前先壓縮資料。
- 盡可能減少資料移動：請勿移動未使用的資料。一些有助於避免移動未使用資料的範例：

- 減少 API 回應 (僅回應相關資料)。
- 彙總詳細的資料 (不需要記錄層級資訊)。
- 參閱 [Well-Architected 實驗室 - 使用 Amazon Redshift 資料共用來最佳化資料模式](#)。
- 考慮在 [AWS Lake Formation](#) 中使用跨帳戶資料共用。
- 使用邊緣服務：使用可協助您在更接近工作負載使用者的位置執行程式碼的服務。

服務	使用情況
Lambda@Edge	用於在物件未經快取時執行的大量運算作業。
CloudFront Functions	處理簡單的使用案例，例如可由短期函數起始的 HTTP(s) 請求/回應操作。
AWS IoT Greengrass	為連線的裝置執行本機運算、傳訊和資料快取。

資源

相關文件：

- [最佳化您的 AWS 永續性基礎架構，第 III 部分：聯網](#)
- [AWS 全球基礎設施](#)
- [Amazon CloudFront 主要功能，包括 CloudFront Global Edge Network](#)
- [在 Amazon OpenSearch Service 中壓縮 HTTP 請求](#)
- [使用 Amazon EMR 進行中間資料壓縮](#)
- [從 Amazon S3 載入壓縮資料檔案至 Amazon Redshift](#)
- [使用 Amazon CloudFront 提供壓縮檔案服務](#)

相關影片：

- [揭密 AWS 上的資料傳輸](#)

SUS04-BP08 僅在難以重新建立時才備份資料

避免備份沒有商業價值的資料，以盡可能降低工作負載的儲存資源需求。

常見的反模式：

- 您沒有資料的備份策略。
- 您備份了可輕易重新建立的資料。

建立此最佳實務的優勢：避免備份非關鍵資料可減少工作負載所需的儲存資源，並降低其環境影響。

未建立此最佳實務時的曝險等級：中

實作指引

避免備份非必要的資料，有助於降低成本和工作負載所使用的儲存資源。僅備份具有商業價值或需要滿足合規要求的資料即可。檢查備份政策，並在復原案例中排除沒有價值的暫時性儲存。

實作步驟

- 將資料分類：實作資料分類政策，如 [SUS04-BP01 實作資料分類政策](#) 所述。
- 設計備份策略：根據您的[復原時間點目標 \(RTO\)](#) 和[復原點目標 \(RPO\)](#) 使用資料分類的關鍵性，並設計備份策略。避免備份非關鍵資料。
 - 排除可輕易重新建立的資料。
 - 從備份排除暫時性資料。
 - 排除資料的本機副本，除非從共同位置還原資料所需的時間不符合服務水準協議 (SLA) 的要求。
- 使用自動備份：使用自動化解決方案或受管服務來備份業務關鍵資料。
 - [AWS Backup](#) 是一項全受管的服務，可輕鬆在雲端和內部部署環境內集中管理並自動化各 AWS 服務間的資料保護。如需有關如何使用 AWS Backup 建立自動化備份的實作指引，請參閱 [Well-Architected 實驗室 - 測試備份並還原資料](#)。
 - [使用 AWS Backup 自動進行 Amazon EFS 的備份及最佳化備份成本](#)。

資源

相關的最佳實務：

- [REL09-BP01 識別並備份所有需要備份的資料，或從來源複製資料](#)
- [REL09-BP03 自動執行資料備份](#)
- [REL13-BP02 使用定義的復原策略來滿足復原目標](#)

相關文件：

- [使用 AWS Backup 來備份和還原 Amazon EFS 檔案系統](#)
- [Amazon EBS 快照](#)
- [在 Amazon Relational Database Service 上使用備份](#)
- [APN 合作夥伴：可以協助備份的合作夥伴](#)
- [AWS Marketplace：可用於備份的產品](#)
- [備份 Amazon EFS](#)
- [備份 Amazon FSx for Windows File Server](#)
- [Amazon ElastiCache \(Redis OSS\) 的備份和還原](#)

相關影片：

- [AWS re:Invent 2023 - 提升恢復能力的備份與災難復原策略](#)
- [AWS re:Invent 2023 - AWS Backup 最新消息](#)
- [AWS re:Invent 2021 - 使用 AWS 進行備份、災難復原和勒索軟體保護](#)

硬體和服務

問題

- [SUS 5 如何選取雲端硬體和服務並在您的架構中使用，以支援您的永續性目標？](#)

SUS 5 如何選取雲端硬體和服務並在您的架構中使用，以支援您的永續性目標？

透過變更硬體管理實務，尋求降低工作負載永續性影響的機會。將佈建和部署所需的硬體量降至最低，並為個別工作負載選取最高效率的硬體和服務。

最佳實務

- [SUS05-BP01 使用最低數量的硬體來滿足需求](#)
- [SUS05-BP02 使用影響最小的執行個體類型](#)
- [SUS05-BP03 使用受管服務](#)
- [SUS05-BP04 將硬體型運算加速器的使用方式最佳化](#)

SUS05-BP01 使用最低數量的硬體來滿足需求

使用最低數量的硬體讓您的工作負載有效達成商業需求。

常見的反模式：

- 您未監控資源使用率。
- 您的架構中有資源處於低使用率水準。
- 您未審查靜態硬體的使用率以確定是否應調整其大小。
- 您未根據業務 KPI 設定運算基礎設施的硬體使用率目標。

建立此最佳實務的優勢：適當調整雲端資源大小有助於降低工作負載的環境影響、節省金錢並維護效能基準。

未建立此最佳實務時的曝險等級：中

實作指引

建議選取您的工作負載所需的硬體總數，以改善其整體效率。AWS 雲端 提供的彈性可透過各種機制 (例如 [AWS Auto Scaling](#)) 動態擴充或減少資源數量，以因應需求的變化。此外也提供 [API 和 SDK](#)，讓修改資源變得非常輕鬆。使用這些功能可以頻繁變更工作負載實作。此外，使用 AWS 工具提供的適當調整大小指引可讓您有效操作雲端資源，而達到您的商業需求。

實作步驟

- 選擇執行個體類型：選擇最適合您的需求的執行個體類型。若要了解如何選擇 Amazon Elastic Compute Cloud 執行個體以及使用屬性型執行個體選擇等機制，請參閱下列內容：
 - [如何為工作負載選擇適當的 Amazon EC2 執行個體類型？](#)
 - [Amazon EC2 Fleet 的屬性型執行個體類型選取。](#)
 - [使用屬性型執行個體類型選取來建立 Auto Scaling 群組。](#)
- 擴展：使用小增量來擴展變數工作負載。
- 使用多種運算購買選項：使用多種運算購買選項平衡執行個體彈性、可擴展性和成本節省。
 - [Amazon EC2 隨需執行個體](#)最適用於新的有狀態尖峰工作負載 (其執行個體類型、位置或時間不具彈性)。
 - [Amazon EC2 Spot 執行個體](#)很適合用來補強具有容錯能力和彈性的應用程式適用的其他選項。
 - 對於狀態穩定、允許隨著您的需求變更保有彈性 (例如 AZ、區域、執行個體系列或執行個體類型) 的工作負載，請使用 [Compute Savings Plans](#)。
- 使用執行個體和可用區域多樣性：利用多樣化執行個體和可用區域來最大化應用程式可用性，並善用多餘容量。

- 適當調整執行個體大小：使用 AWS 工具提供的適當調整大小建議，調整您的工作負載。如需詳細資訊，請參閱[利用精簡化建議最佳化您的成本](#)和[適當調整大小：佈建執行個體以符合工作負載](#)
- 使用 AWS Cost Explorer 或 [AWS Compute Optimizer](#) 中的適當調整大小建議，識別適當調整大小的機會。
- 協商服務水準協議 (SLA)：協商允許暫時減少容量的 SLA，同時自動化部署替換資源。

資源

相關文件：

- [最佳化您的 AWS 永續性基礎設施，第 I 部分：運算](#)
- [Amazon EC2 Fleet 的 Auto Scaling 屬性型執行個體類型選取](#)
- [AWS Compute Optimizer 文件](#)
- [操作 Lambda：效能最佳化](#)
- [Auto Scaling 文件](#)

相關影片：

- [AWS re:Invent 2023 - Amazon EC2 最新消息](#)
- [AWS re:Invent 2023 - 智慧型節約：Amazon Elastic Compute Cloud 成本最佳化策略](#)
- [AWS re:Invent 2022 - 在 AWS 上針對效能和成本最佳化 Amazon Elastic Kubernetes Service](#)
- [AWS re:Invent 2023 - 永續性運算：使用 AWS 降低成本和碳排放量](#)

SUS05-BP02 使用影響最小的執行個體類型

持續監控並使用新的執行個體類型，讓能源效率方面的改進充分發揮效用。

常見的反模式：

- 您僅使用一個執行個體系列。
- 您僅使用 x86 執行個體。
- 您在 Amazon EC2 Auto Scaling 組態中指定了一個執行個體類型。
- 您以不符合設計宗旨的方式使用 AWS 執行個體 (例如，您將運算最佳化執行個體用於記憶體密集型工作負載)。
- 您未定期評估新的執行個體類型。

- 您未查看 AWS 適當調整大小的工具 (例如 [AWS Compute Optimizer](#)) 所提供的建議。

建立此最佳實務的優勢：藉由使用節能且適當調整大小的執行個體，將可大幅降低環境受到的影響以及工作負載成本。

未建立此最佳實務時的曝險等級：中

實作指引

在雲端工作負載中使用高效執行個體，是降低資源用量和提高成本效益的關鍵。持續關注新執行個體類型的發佈，並運用能源效率改進，包括旨在支援特定工作負載 (例如機器學習訓練和推論以及影片轉碼) 的執行個體類型。

實作步驟

- 了解並探索執行個體類型：尋找可降低工作負載對環境之影響的執行個體類型。
 - 訂閱 [AWS 最新消息](#)，隨時掌握最新的 AWS 技術和執行個體。
 - 了解不同的 AWS 執行個體類型。
 - 觀看下列資源，了解 AWS Graviton 型執行個體如何在 Amazon EC2 中的能源使用提供最佳效能功耗比：[re:Invent 2020 - 深入探討搭載 AWS Graviton2 處理器的 Amazon EC2 執行個體](#)與[深入探討 AWS Graviton3 和 Amazon EC2 C7g 執行個體](#)。
- 使用影響最小的執行個體類型：進行相關規劃，將工作負載轉移至影響程度最低的執行個體類型。
 - 定義一個程序來評估工作負載的新功能和執行個體。利用雲端的靈活性快速測試新功能類型對您的工作負載環境永續性有何改善。使用代理指標，測量您需要多少資源才能完成一個工作單位。
 - 如果可行，請修改工作負載，以使用不同數量的 CPU 和不同數量的記憶體，從而最大化您選擇執行個體類型的空間。
 - 考慮將您的工作負載轉移至 Graviton 型執行個體，以改善工作負載的效能效率。如需有關將工作負載移至 AWS Graviton 的詳細資訊，請參閱 [AWS Graviton Fast Start](#) 和 [將工作負載轉移至基於 AWS Graviton 的 Amazon Elastic Compute Cloud 執行個體時的注意事項](#)。
 - 考慮在您使用的 [AWS 受管服務](#) 中選取 AWS Graviton 選項。
 - 將工作負載遷移至有執行個體對永續性影響最小，且仍符合業務要求的區域。
 - 針對機器學習工作負載，請利用專供工作負載使用的專用硬體，例如 [AWS Trainium](#)、[AWS Inferentia](#) 和 [Amazon EC2 DL1](#)。AWS Inferentia 執行個體 (例如 Inf2 執行個體) 所提供的效能功耗比最多會比同類 Amazon EC2 執行個體高出 50%。
 - 使用 [Amazon SageMaker AI Inference Recommender](#) 適當調整 ML 推論端點的大小。
 - 對於尖峰工作負載 (不常需要額外容量的工作負載)，請使用 [爆量效能執行個體](#)。

- 對於無狀態和容錯工作負載，請使用 [Amazon EC2 Spot 執行個體](#)，以提高雲端整體使用率，並減少未使用資源的永續性影響。
- 操作和最佳化：操作並最佳化您的工作負載執行個體。
 - 對於暫時性工作負載，請評估[執行個體 Amazon CloudWatch 指標](#) (例如 CPUUtilization)，以識別執行個體是否閒置或未充分利用。
 - 對於穩定的工作負載，請定期檢查 AWS 適當調整大小的工具 (例如 [AWS Compute Optimizer](#))，以識別對執行個體進行最佳化和適當調整大小的機會。如需更多範例和建議，請參閱下列實驗室：
 - [Well-Architected 實驗室 - 適當調整大小的建議](#)
 - [Well-Architected 實驗室：使用 Compute Optimizer 適當調整大小](#)
 - [Well-Architected 實驗室 - 最佳化硬體模式和觀察永續性 KPI](#)

資源

相關文件：

- [最佳化您的 AWS 永續性基礎設施，第 I 部分：運算](#)
- [AWS Graviton](#)
- [Amazon EC2 DL1](#)
- [Amazon EC2 容量保留機群](#)
- [Amazon EC2 Spot 機群](#)
- [函數：Lambda 函數組態](#)
- [Amazon EC2 Fleet 的屬性型執行個體類型選取](#)
- [在 AWS 上建置永續性、高效且成本最佳化的應用程式](#)
- [Contino 永續性儀表板如何協助客戶最佳化碳足跡](#)

相關影片：

- [AWS re:Invent 2023 - AWS Graviton：AWS 工作負載的最佳性價比](#)
- [AWS re:Invent 2023 - 在 AWS 管理主控台 中新建 Amazon Elastic Compute Cloud 生成式 AI 功能](#)
- [AWS re:Invent 2023 - Amazon Elastic Compute Cloud 的最新消息](#)
- [AWS re:Invent 2023 - 智慧型節約：Amazon Elastic Compute Cloud 成本最佳化策略](#)
- [AWS re:Invent 2021 - 深入探討 AWS Graviton3 和 Amazon EC2 C7g 執行個體](#)

- [AWS re:Invent 2022 - 建置具成本、能源和資源效率的運算環境](#)

相關範例：

- [解決方案：在 AWS 上最佳化深度學習工作負載以維持永續性的指引](#)

SUS05-BP03 使用受管服務

使用受管服務以提高雲端中的操作效率。

常見的反模式：

- 您可以使用低使用率的 Amazon EC2 執行個體來執行應用程式。
- 您的內部團隊僅管理工作負載，而沒有時間專注於創新或簡化。
- 您為在受管服務上可更高效執行的任務部署及維護技術。

建立此最佳實務的優勢：

- 使用 受管服務將責任轉移到 AWS，這具有數百萬個客戶之間的洞察，有助於推動新的創新和效率。
- 基於多租用戶控制平面，受管服務將服務產生的環境影響分散到眾多使用者。

未建立此最佳實務時的曝險等級：中

實作指引

受管服務將責任轉移到 AWS，以維持已部署硬體的高使用率和永續性最佳化。受管服務也免除了維護服務的營運和管理重擔，讓您的團隊有更多時間可專注於創新。

檢閱您的工作負載，以識別可以由 AWS 受管服務取代的元件。例如，[Amazon RDS](#)、[Amazon Redshift](#) 和 [Amazon ElastiCache](#) 提供受管資料庫服務。[Amazon Athena](#)[Amazon EMR](#)和 [Amazon OpenSearch Service](#) 提供受管分析服務。

實作步驟

1. 清查工作負載：清查工作負載中的服務和元件。
2. 識別候選項：評估並識別可能被受管服務取代的元件。以下舉例說明您可能會考慮使用受管服務的時機：

任務	要在 上使用什麼 AWS
託管資料庫	使用受管 Amazon Relational Database Service (Amazon RDS) 執行個體，而不是在 Amazon Elastic Compute Cloud (Amazon EC2) 上維護您自己的 Amazon RDS執行個體。
託管容器工作負載	使用 AWS Fargate ，而非實作您自己的容器基礎設施。
託管 Web 應用程式	使用 AWS Amplify 託管 ，供靜態網站和伺服器端轉譯的 Web 應用程式作為全受管 CI/CD 和託管服務。

3. 建立遷移計畫：識別相依項並建立遷移計畫。據以更新執行手冊和程序手冊。
 - [AWS Application Discovery Service](#) 會自動收集並提供應用程式相依性和利用率的詳細資訊，協助您在計劃遷移時做出更明智的決定
4. 執行測試：在遷移至受管服務之前先測試服務。
5. 取代自我託管服務：使用遷移計畫將自我託管服務取代為受管服務。
6. 監控和調整：在遷移完成後繼續監控服務，以便在必要時進行調整及最佳化服務。

資源

相關文件：

- [AWS 雲端 產品](#)
- [AWS 總擁有成本 \(TCO \) 計算器](#)
- [Amazon DocumentDB](#)
- [Amazon Elastic Kubernetes Service \(EKS \)](#)
- [Amazon Managed Streaming for Apache Kafka \(Amazon MSK \)](#)

相關影片：

- [AWS re : Invent 2021 - 大規模雲端操作 AWS Managed Services](#)

- [AWS re : Invent 2023 - 在上操作的最佳實務 AWS](#)

SUS05-BP04 將硬體型運算加速器的使用方式最佳化

將加速運算執行個體的使用方式最佳化，以降低工作負載的實體基礎設施需求。

常見的反模式：

- 未監控 GPU 使用率。
- 針對工作負載使用一般用途執行個體，但專用執行個體可以提供更高的效能、較低的成本，以及更優異的效能功耗比。
- 您使用硬體型運算加速器來執行任務，但使用 CPU 型運算加速器來執行時會更有效率。

建立此最佳實務的優勢：藉由將硬體型加速器的使用方式最佳化，您可以降低工作負載的實體基礎設施需求。

未建立此最佳實務時的曝險等級：中

實作指引

如果需要高處理能力，使用加速運算執行個體可讓您獲得好處，因為其可讓您存取硬體型運算加速器，例如圖形處理單元 (GPU) 和現場可程式化閘道陣列 (FPGA)。這些硬體加速器比基於 CPU 的替代品更有效地執行某些功能，例如圖形處理或資料模式匹配。許多加速的工作負載 (例如轉譯、轉碼和機器學習) 在資源使用方面變化很大。只在需要時執行此硬體，不需要時便將其自動除役，以將資源消耗降至最低。

實作步驟

- 探索運算加速器：確定哪些[加速運算執行個體](#)可以滿足您的需求。
- 使用專用硬體：針對機器學習工作負載，請利用專供工作負載使用的專用硬體，例如 [AWS Trainium](#)、[AWS Inferentia](#) 和 [Amazon EC2 DL1](#)。AWSInferentia 執行個體 (例如 Inf2 執行個體) 所提供的[效能功耗比最多會比同類 Amazon EC2 執行個體高出 50%](#)。
- 監控用量指標：收集加速運算執行個體的用量指標。例如，可以使用 CloudWatch 代理程式為您的 GPU 收集 utilization_gpu 和 utilization_memory 等指標，如[使用 Amazon CloudWatch 收集 NVIDIA GPU 指標](#)中所示。
- 大小調整：優化硬體加速器的程式碼、網路運作和設定，以確保系統會充分利用基礎硬體。
 - [最佳化 GPU 設定](#)
 - [深度學習 AMI 中的 GPU 監控和最佳化](#)

- [將 I/O 最佳化以針對 Amazon SageMaker AI 中的深度學習訓練進行 GPU 效能調校](#)
- 保持最新狀態：使用最新的高效能程式庫和 GPU 驅動程式。
- 發布不需要的執行個體：使用自動化來發布未使用的 GPU 執行個體。

資源

相關文件：

- [加速運算](#)
- [開始建構吧！使用自訂晶片和加速器來進行建構](#)
- [如何為工作負載選擇適當的 Amazon EC2 執行個體類型？](#)
- [Amazon EC2 VT1 執行個體](#)
- [選擇最佳 AI 加速器和模型編譯來以 Amazon SageMaker AI 推斷電腦視覺](#)

相關影片：

- [AWS re:Invent 2021 - 如何為深度學習選取 Amazon EC2 GPU 執行個體](#)
- [AWS 線上技術講座 - 部署具成本效益的深度學習推論](#)
- [AWS re:Invent 2023 - 搭配 AWS 和 NVIDIA 的尖端 AI](#)
- [AWS re:Invent 2022 - \[最新發佈！\] 介紹基於 AWS Inferentia2 的 Amazon EC2 Inf2 執行個體](#)
- [AWS re:Invent 2022 - 使用 AWS Trainium 加速深度學習和創新](#)
- [AWS re:Invent 2022 - 透過 NVIDIA 在 AWS 上進行深度學習：從訓練到部署](#)

程序和文化

問題

- [SUS 6 您的組織程序如何支援您的永續性目標？](#)

SUS 6 您的組織程序如何支援您的永續性目標？

透過變更開發、測試和部署實務來尋找降低永續性影響的機會。

最佳實務

- [SUS06-BP01 傳達和串聯您的永續發展目標](#)

- [SUS06-BP02 採用可快速導入永續性改進的方法](#)
- [SUS06-BP03 讓工作負載保持在最新狀態](#)
- [SUS06-BP04 提高建置環境的使用率](#)
- [SUS06-BP05 使用受管 Device Farm 進行測試](#)

SUS06-BP01 傳達和串聯您的永續發展目標

技術是永續發展的關鍵推動因素。IT 團隊在推動組織永續發展目標的有意義變革上，扮演著關鍵角色。這些團隊應該清楚了解公司的永續發展目標，並努力在其營運中傳達和串聯這些優先事項。

常見的反模式：

- 您不知道組織的永續發展目標，以及如何將這些目標應用在您的團隊上。
- 您對雲端工作負載對環境影響的意識和訓練不足。
- 您不確定需優先處理的特定領域。
- 您未讓員工和客戶參與您的永續發展計畫。

建立此最佳實務的優勢：從基礎設施和系統的最佳化到創新技術的使用，IT 團隊可以減少組織的碳排放，並將資源消耗量降到最低。永續發展目標的傳達為 IT 團隊提供持續改進，並適應不斷變化的永續發展挑戰的能力。此外，這些永續最佳化通常也會轉化為成本節約，進而強化商業案例。

未建立此最佳實務時的曝險等級：中

實作指引

IT 團隊的主要永續發展目標是最佳化系統和解決方案，旨在提高資源效率，並將組織的碳足跡和整體環境影響降至最低。培訓計畫和操作儀表板等共用服務和計畫，支援組織最佳化 IT 維運並建置解決方案，協助大幅減少碳足跡。雲端提供以下機會：將實體基礎設施和能源採購責任移至雲端供應商共同責任，並持續最佳化雲端服務的資源效率。

當團隊使用雲端的固有效率和共同責任模型時，可以大幅降低組織對環境造成的影響。但反過來看，這卻能為組織的整體永續發展目標做出貢獻，並在邁向更加永續的未來過程中，展現這些團隊成為策略合作夥伴的價值。

實作步驟

- 定義目標和目的：為您的 IT 計畫建立明確的目標。這包括從 IT、永續發展和財務等不同部門取得負責之利害關係人的意見。這些團隊應定義符合組織永續發展目標的可衡量目標，包括減碳和資源最佳化等領域。

- 了解企業的碳核算界限：了解溫室氣體 (GHG) 通訊協定等碳核算方法與雲端工作負載的關係 (如需詳細資訊，請參閱[雲端永續性](#))。
- 使用雲端解決方案進行碳核算：[在 AWS 上使用碳核算解決方案](#)等雲端解決方案，追蹤營運、產品組合及價值鏈中 GHG 排放的範圍一、二和三。透過這些解決方案，各組織可以簡化 GHG 排放資料擷取、簡化報告，以及衍生洞見以將此深入見解知會其氣候策略。
- 監控 IT 產品組合的碳足跡：追蹤和報告 IT 系統的碳排放量。使用 [AWS 客戶碳足跡工具](#) 追蹤、測量、檢閱和預測因使用 AWS 而產生的碳排放。
- 透過代理指標向您的團隊傳達資源用量：[透過代理指標追蹤和報告資源用量](#)。在雲端的隨需定價模型中，資源用量與成本息息相關，這是合乎情理的正常指標。至少，將成本用作代理指標，以傳達每個團隊的資源用量和進步情況。
 - 在 Cost Explorer 中啟用每小時精細程度，並建立[成本和用量報告 \(CUR\)](#)：CUR 為所有 AWS 服務提供每日或每小時用量精細程度、費率、成本和用量屬性。使用[雲端智慧儀表板](#)及其永續性代理指標儀表板，將其做為處理和視覺化成本和用量型資料的起點。如需詳細資訊，請參閱下列內容：
 - [使用永續性代理指標測量和追蹤雲端效率，第 I 部分：什麼是代理指標？](#)
 - [使用永續性代理指標來測量和追蹤雲端效率，第 II 部分：建立指標管道](#)
- 持續最佳化和評估：使用[改進程序](#)持續最佳化您的 IT 系統 (包括雲端工作負載) 以實現高效率和永續性。在實作最佳化策略之前和之後監控碳足跡。藉由減少碳足跡評估有效性。
- 培養永續發展文化：利用培訓計畫 (例如 [AWS Skill Builder](#)) 教育您的員工有關永續發展的資訊。讓他們參與永續發展計畫。分享並慶祝他們的成功案例。採用獎勵措施，在達成永續發展目標時不吝給予獎勵。

資源

相關文件：

- [了解您的碳排放估算](#)

相關影片：

- [AWS re:Invent 2023 - 使用 AWS 加速資料導向的循環經濟倡議](#)
- [AWS re:Invent 2023 - AWS 全球基礎設施的永續創新](#)
- [AWS re:Invent 2023 - 永續架構：過去、現在和未來](#)
- [AWS re:Invent 2022 - 提供永續且高效能的架構](#)
- [AWS re:Invent 2022 - 永續架構並降低您的 AWS 碳足跡](#)

- [AWS re:Invent 2022 - AWS 全球基礎設施的永續性](#)

相關範例：

- [Well-Architected 實驗室 - 將成本和用量報告轉換為效率報告](#)

相關培訓：

- [AWS 上的永續性轉型](#)
- [SimuLearn - 永續發展報告](#)
- [使用 AWS 去碳化](#)

SUS06-BP02 採用可快速導入永續性改進的方法

採用相關方法和程序來驗證潛在改善、盡可能降低測試成本，以及提供小幅改善。

常見的反模式：

- 審查應用程式的永續性是僅需在專案開始時執行一次的任務。
- 您的工作負載已過時，因為發行程序太繁瑣而無法導入資源效率的小幅變更。
- 您沒有改善工作負載以維持永續性的機制。

建立此最佳實務的優勢：建立導入和追蹤永續性改善的程序後，您將可持續採用新的特性和功能、消除問題，並改善工作負載效率。

未建立此最佳實務時的曝險等級：中

實作指引

在將潛在永續性改善部署到生產環境之前，先加以測試和驗證。在計算改善所帶來的未來潛在利益時，應考慮測試成本。開發低成本測試方法以提供小幅改善。

實作步驟

- 了解並傳達組織永續發展目標：了解您的組織的永續發展目標，例如減碳或水資源管理。將這些目標轉化為雲端工作負載的永續需求。將這些需求傳達給主要利害關係人。
- 將永續發展需求加入到待辦事項：在開發待辦事項中新增永續改善需求。

- 迭代和改善：使用[迭代改善程序](#)對這些改善進行識別、評估、優先順序設定、測試及部署。
- 使用最簡可行產品 (MVP) 進行測試：使用最簡可行的代表元件開發並測試可能的改善，以降低測試成本和對環境的衝擊。
- 簡化流程：持續改進並簡化您的開發流程。例如，使用持續整合與持續交付 (CI/CD) 管道自動執行軟體交付流程，以測試及部署可能的改善，進而減少工作量和手動流程導致的錯誤。
- 培訓和認知：為您的團隊成員執行培訓計畫，帶他們了解永續發展及其活動如何影響組織的永續發展目標。
- 評估和調整：持續評估改善的影響，並視需要進行調整。

資源

相關文件：

- [AWS 啟用永續性解決方案](#)

相關影片：

- [AWS re:Invent 2023 - 永續架構：過去、現在和未來](#)
- [AWS re:Invent 2022 - 提供永續且高效能的架構](#)
- [AWS re:Invent 2022 - 永續架構並降低您的 AWS 碳足跡](#)
- [AWS re:Invent 2022 - AWS 全球基礎設施的永續性](#)
- [AWS re:Invent 2023 - AWS 可觀測性和營運最新消息](#)

SUS06-BP03 讓工作負載保持在最新狀態

將工作負載保持在最新狀態，以採用高效功能、去除問題，以及改善工作負載的整體效率。

常見的反模式：

- 假設您目前的架構為靜態，且不會隨著時間的推移而更新。
- 您沒有任何系統或定期規律可評估更新的軟體與套件是否與您的工作負載相容。

建立此最佳實務的優勢：建立讓工作負載保持在最新狀態的程序後，您將可採用新的特性和功能、解決問題，並改善工作負載效率。

未建立此最佳實務時的曝險等級：低

實作指引

最新的作業系統、執行時期、中介軟體、程式庫和應用程式可改善工作負載效率，讓您更輕鬆地採用更有效率的技術。隨著供應商提供符合自身永續性目標的功能，最新軟體也可能包含更準確測量工作負載對永續性影響的功能。定期以最新的功能和版本將工作負載保持在最新狀態。

實作步驟

- **定義程序：**定義相關程序和排程來評估工作負載的新功能和執行個體。利用雲端的靈活性快速測試新功能對您的工作負載有何改善，藉以：
 - 降低永續性的影響。
 - 獲得效能效率。
 - 消除已計劃改善的障礙。
 - 提升測量和管理工作負載影響的能力。
- **執行清查：**清查工作負載軟體和架構，並識別需要更新的元件。
 - 您可以使用 [AWS Systems Manager 庫存](#)，從您的 Amazon EC2 執行個體收集作業系統 (OS)、應用程式和執行個體中繼資料，並快速了解哪些執行個體正在執行您的軟體政策所需的軟體與組態，以及哪些執行個體需要更新。
- **了解更新程序：**了解如何更新工作負載的元件。

工作負載元件	如何更新
機器映像	使用 EC2 Image Builder 管理適用於 Linux 或 Windows Server 映像的 Amazon Machine Images (AMI) 的更新。
容器映像	搭配使用 Amazon Elastic Container Registry (Amazon ECR) 與您現有的管道來 管理 Amazon Elastic Container Service (Amazon ECS) 映像 。
AWS Lambda	AWS Lambda 包含 版本管理功能 。

- **使用自動化：**自動化更新，以減少部署新功能的工作量，並避免手動程序引起的錯誤。
 - 您可以使用 [CI/CD](#) 自動更新 AMI、容器映像，以及其他與您的雲端應用程式相關的成品。

- 您可以使用 [AWS Systems Manager Patch Manager](#) 之類的工具自動執行系統更新的程序，並使用 [AWS Systems Manager Maintenance Windows](#) 排定活動時程。

資源

相關文件：

- [AWS 架構中心](#)
- [AWS 最新消息](#)
- [AWS 開發人員工具](#)

相關影片：

- [AWS re:Invent 2022 - 使用最佳實務指引最佳化您的 AWS 工作負載](#)
- [所有事項修補程式：AWS Systems Manager](#)

SUS06-BP04 提高建置環境的使用率

提高資源的使用率以開發、測試及建置您的工作負載。

常見的反模式：

- 您以手動方式佈建或終止您的建置環境。
- 您讓建置環境在測試、建置或發行活動以外執行 (例如，在開發團隊成員的非上班時間執行環境)。
- 您為建置環境過度佈建資源。

建立此最佳實務的優勢：藉由提高建置環境的使用率，您將可改善雲端工作負載的整體效率，同時為建置者配置有效開發、測試和建置所需的資源。

未建立此最佳實務時的曝險等級：低

實作指引

使用自動化和基礎設施即程式碼，在需要時啟動建置環境，並在不使用時將其關閉。常見的模式是排程可用性時間，使之與開發團隊成員的工作時間一致。您的測試環境應該會與生產組態近似。不過，請找機會使用具有高載容量的執行個體類型、Amazon EC2 Spot 執行個體、自動調整規模資料庫服務、容器和無伺服器技術，以根據使用量調整開發和測試容量。將資料量限定為剛好達到測試要求。如果在測試中使用生產資料，請尋求從生產環境共用資料的可能性，而不要移動資料。

實作步驟

- 使用基礎設施即程式碼：使用基礎設施即程式碼來佈建您的建置環境。
- 使用自動化：使用自動化來管理開發和測試環境的生命週期，並且讓建置資源發揮最大效益。
- 最大化使用率：使用策略讓開發和測試環境達到最大的使用率。
 - 使用最低可行的代表環境來開發和測試潛在改善。
 - 在情況允許時使用無伺服器技術。
 - 使用隨需執行個體補充開發人員裝置。
 - 使用具有高載容量的執行個體類型、Spot 執行個體和其他技術，以根據使用量調整建置容量。
 - 採用原生雲端服務來獲得安全的執行個體 Shell 存取，而非部署堡壘主機機群。
 - 根據您的建置任務自動調整建置資源規模。

資源

相關文件：

- [AWS Systems Manager Session Manager](#)
- [Amazon EC2 爆量效能執行個體](#)
- [什麼是 AWS CloudFormation？](#)
- [什麼是 AWS CodeBuild？](#)
- [Instance Scheduler on AWS](#)

相關影片：

- [AWS re:Invent 2023 - AWS 的持續整合與交付](#)

SUS06-BP05 使用受管 Device Farm 進行測試

使用受管 Device Farm 有效測試代表性硬體集上的新功能。

常見的反模式：

- 您在個別實體裝置上手動測試及部署應用程式。
- 您未在真正的實體裝置上使用應用程式測試服務來測試及操作應用程式 (例如 Android、iOS 和 Web 應用程式)。

建立此最佳實務的優勢：使用受管 Device Farm 來測試啟用雲端功能的應用程式有許多優勢：

- 將有更多高效率功能可用來測試各種裝置上的應用程式。
- 無需再以內部基礎設施進行測試。
- 提供多種裝置類型 (包括較舊且較不熱門的硬體)，因而無需再進行不必要的裝置升級。

未建立此最佳實務時的曝險等級：低

實作指引

使用受管 Device Farm 有助於簡化對代表性硬體集上的新功能進行測試的程序。受管 Device Farm 提供多種裝置類型 (包括較舊且較不熱門的硬體)，並避免不必要的裝置升級對客戶的永續性造成影響。

實作步驟

- 定義測試要求：定義您的測試要求和計畫 (例如測試類型、作業系統和測試排程)。
 - 您可以使用 [Amazon CloudWatch RUM](#) 來收集和分析用戶端資料，並研擬您的測試計畫。
- 選取受管 Device Farm：選取可支援測試要求的受管 Device Farm。例如，您可以使用 [AWS Device Farm](#) 來測試和了解您的變更對代表性硬體集有何影響。
- 使用自動化：使用自動化和持續整合/持續部署 (CI/CD) 來排程和執行測試。
 - [整合 AWS Device Farm 與您的 CI/CD 管道以執行跨瀏覽器 Selenium 測試](#)
 - [使用 AWS DevOps 和行動服務建置及測試 iOS 和 iPadOS 應用程式](#)
- 審查與調整：持續審查測試結果並進行必要的改進。

資源

相關文件：

- [AWS Device Farm 裝置清單](#)
- [檢視 CloudWatch RUM 儀表板](#)

相關影片：

- [AWS re:Invent 2023 - 使用 AWS Device Farm 改善行動裝置和 Web 應用程式品質](#)
- [AWS re:Invent 2021 - 透過最終使用者洞察與 Amazon CloudWatch RUM 最佳化應用程式](#)

相關範例：

- [Android 的 AWS Device Farm 範例應用程式](#)
- [iOS 的 AWS Device Farm 範例應用程式](#)
- [AWS Device Farm 的 Appium Web 測試](#)

注意

客戶有責任對本文件中的資訊進行自己的獨立評定。本文件：(a) 僅供參考之用；(b) 代表目前 AWS 產品供應與實務，如有變更恕不另行通知；以及 (c) 不構成 AWS 及其附屬公司、供應商或授權人的任何承諾或保證。AWS 產品或服務均以「原樣」提供，不作任何形式的明示或暗示的保證、陳述或條件。AWS 對其客戶的責任與義務應由 AWS 協議管轄，本文並非 AWS 與其客戶之間的任何協議的一部分，也並非上述協議的修改。

Copyright © 2024 Amazon Web Services, Inc. 或其附屬公司。

AWS 詞彙表

如需最新的 AWS 術語，請參閱AWS 詞彙表 參考 中的[AWS 詞彙表](#)。