制定教育领域的单云、混合云和多云策略

AWS 规范性指导

AWS 规范性指导: 制定教育领域的单云、混合云和多云策略

Table of Contents

简介 1
概述 1
云部署策略 3
单云 3
混合云 3
多云 3
建议 4
选择主要战略云提供商 4
建立一个 CCo E 5
区分 SaaS 应用程序和基础云服务 7
为每个云服务提供商制定安全和治理要求 9
在切实可行的情况下，尽可能采用云原生托管服务 10
在现有本地投资激励继续使用时，实施混合架构 13
仅为无法通过单一云提供商满足其技术或业务要求的工作负载保留多云 15
使用案例示例 17
虚拟计算机实验室 17
预测学生成功 19
身份联合验证和单点登录 21
研究计算的云爆发 22
后续步骤 24
贡献者 25
延伸阅读 26
文档历史记录 27
术语表 28
..... 28
A 28
B 31
C 32
D 35
E 38
F 40
G 41
H 42
我 43

L 45
M 46
O 50
P 52
Q 54
R 55
S 57
T 60
U 61
V 62
W 62
Z 63
..... lxiv

制定教育领域的单云、混合云和多云策略

Amazon Web Services（贡献者）

2023 年 9 月（文档历史记录）

教育机构正在寻求利用云计算提供的敏捷性、成本节省、安全性和韧性来支持远程学习、研究、学生体验、数据洞察和管理等功能。作为数字化转型的一部分，许多组织都在评测混合云和多云部署。

本文为评估其云方案的教育机构高管和决策者提供制定单云、混合云和多云技术和治理策略的规范指引。该指引基于我们在 AWS 与全球 14000 多所不同规模的教育机构（从小学、中学到高等教育）合作的经验。

概述

随着教育机构数字化转型，为学生、家长、教职员工和社区提供差异化的服务和体验，其面临着大量的技术决策。许多组织都已经决定采用云来提高敏捷性、弹性、韧性、安全性和成本节省。根据各团队现有的关系和投资，大多数组织都使用本地数据中心、主机托管设施和云提供商的某种组合。鉴于提供多种云方案，教育机构必须经常从单云、混合云和多云部署模式（在云部署策略一节中定义）中进行选择。

多云（即使用至少两家云服务提供商的服务）如今在许多机构中并不罕见。您的 IT 团队可能首选一家云提供商，而其他团队、部门或个人用户可能选择或已在使用其他提供商。如果教育机构没有明确的策略指导其采用适当的云部署模式，就会面临诸多挑战。这些挑战包括不必要的复杂性、不断增加的员工需求、不一致的治理以及最低共同点的方法，将其限制为各提供商通用的基本功能子集。每一项挑战都会扼杀创新，减缓数字化转型的步伐。

相反，如果您有指导使用单云、混合云和多云的云策略，则可以满足您的教育使命要求，同时以运营上可持续的方式实现云优势，从而取得长期成功。为制定此策略，我们建议采取以下措施：

- 选择主要战略云提供商。
- 建立云卓越中心 (CCoE)。
- 区分软件即服务（SaaS）应用程序和基础云服务。
- 为每个云服务提供商制定安全和治理要求。
- 在切实可行的情况下，尽可能采用云原生托管解决方案。
- 在现有本地投资激励继续使用时，实施混合架构。
- 仅为无法通过单一云提供商满足技术或业务要求的工作负载保留多云。

本文的建议一节详细讨论这些最佳实践。每项建议都很重要，但机构的优先事项取决于其云采用阶段。例如，如果您刚刚开始采用云，请专注于选择主要的战略云提供商，建立 CCo企业并采用云原生托管解决方案。如果您已经使用单一云提供商，请重点关注制定核心安全和治理要求，并在现有数据中心投资激励继续使用时考虑混合架构。如果您的组织已在使用多个云提供商，请重点关注区分 SaaS 应用程序，并保留多云部署用于真正需要的少数工作负载。

内容

- 云部署策略
- 建议
- 用例示例
- 后续步骤
- 贡献者
- 延伸阅读
- 文档历史记录

云部署策略

AWS 将云计算定义为通过互联网按需提供 IT 资源，采用即用即付定价模式。您无需购买、拥有和维护物理数据中心和服务器，而是可以根据需要从云提供商处获得计算能力、存储和数据库等技术服务。云计算可以让教育机构避免硬件采购、维护和容量规划等无差异化的繁重工作。在采用和部署云解决方案时，您可以从多种模式进行选择：单云、混合云和多云。

单云

此模式仅使用单一云服务提供商。单云应用程序和工作负载可以直接在云实施，也可以之前托管在其他环境中并迁移到云。这些工作负载可使用其云提供商提供的下层基础设施服务，也可以利用上层托管服务。无论如何，此模式采用单一云提供商，并且仅使用该提供商的云服务。

混合云

混合云模式在组织自己的本地数据中心和至少一家云服务提供商之间分配资源。通常，这种模式的目的是将组织的基础设施扩展到云，同时保持与位于本地现有内部系统的私有连接。

多云

多云模式在至少两家云服务提供商之间分配资源，并使用其提供的服务。组织可能会选择多云，但更常见的情况是，各团队、部门或员工因其对不同云提供商的偏好无意中形成的结果。

建议

既然您已经对单云、混合云和多云有基本的了解，那么本节将提供有关选择模式的详细建议。

- 选择主要战略云提供商
- 建立一个 CCo E
- 区分 SaaS 应用程序和基础云服务
- 为每个云服务提供商制定安全和治理要求
- 在切实可行的情况下，尽可能采用云原生托管服务
- 在现有本地投资激励继续使用时，实施混合架构
- 仅为无法通过单一云提供商满足其技术或业务要求的工作负载保留多云

选择主要战略云提供商

云采用可以带来对于 IT 现代化、成本效益和创新至关重要的诸多优势。然而，除了有限的 SaaS 应用程序之外，采用云技术也可能带来一些挑战，教育机构必须谨慎规划，以避免不必要的成本和复杂性。在云实施工作负载涉及技术和业务变革，需要对员工进行赋能，以及调整核心基础设施，包括网络、安全、治理和运营。

有效应对这些挑战的最佳方法，是选择一家主要战略云提供商来支持您的大部分工作负载，尤其是在您的组织处于云之旅的早期阶段时。从以该提供商为中心的重点采用开始，这样便可简化和加快实现云优势。选择主要云提供商并非排他性的、不可逆转的决定。它使您的组织能够以迭代方式演进云采用。您可以先专注于少数几项服务，然后根据需要扩展到其他云服务，而不会延迟云的整体优势。此方法使组织能够最大限度地利用供应商的能力、集中发展员工技能和第三方合作伙伴关系，以及简化供应商管理。

我们看到过一些客户通过尝试同时采用多个云提供商来踏上云之旅，但后来却对这一决定及其引入的复杂性懊悔不已。Gartner 在其文章 6 Steps for Planning a Cloud Strategy 中分享了这一见解，其中步骤 2 是“Prioritize a primary provider in multicloud architectures”。

每家云提供商都引入了不同的运营和支持模式、身份和访问管理、网络、运营、合规性功能等。最好一次掌握一家云提供商的运营模式。然后，您可以在合理化的情况下，以迭代和增量方式整合其他云服务。许多因素都可能会影响您采用主要云提供商的决定，但请使用以下关键问题来指导您的选择。

- 提供商所提供服务的广度和深度如何？

不同的云提供商提供不同的服务。至少确保主要提供商具备支持您所有功能需求以及安全、治理和自动化等跨领域运营需求必要的能力。选择能够提供这些能力、在创新和卓越运营方面拥有实绩的提供商。不仅要考虑您的应用程序，还要考虑您的数据。考虑未来的数据集成和传输模式，以限制在提供商之间移动大量数据带来的成本、延迟和复杂性。选择服务广度和深度尽可能大的提供商，既能满足您当前的应用程序和数据需求，又能解锁新使用案例以适应随时间变化的机构需求。

- 提供商是否能支持您的所有安全与合规需求？

  在教育领域，安全与合规对任何技术部署都至关重要。选择能够满足您所有安全与合规需求的云提供商。通过提供按需访问安全与合规报告的中心资源，AWS Artifact 等工具可以帮助您评估提供商。不仅要考虑云提供商自身基础设施和服务的安全与合规，还要考虑使用这些服务构建安全、合规解决方案的难易程度。优先选择提供预构建解决方案、快速入门和规范指引的某种组合的提供商，以加快云的安全采用。

- 提供商是否拥有强大的合作伙伴网络？

  任何组织都无法独自完成云转型。为了加快采用速度，您应该利用云提供商及其合作伙伴网络的服务和专业知识。该网络包括技术合作伙伴（提供在云运行、与云技术集成或支持云技术的软件），以及咨询合作伙伴（可以帮助您在云设计、构建、运行和管理自己的应用程序）。您会发现，许多已经与您合作的教育技术提供商、独立软件供应商 (ISVs)、顾问和经销商都是云提供商合作伙伴网络的成员。优先选择拥有最强大的合作伙伴网络且其合作伙伴具备经过验证能力的云提供商。拥有具备成熟行业和技术专业知识的合作伙伴至关重要。

- 提供商提供哪些支持和赋能？

  要成功采用任何新技术，您需要请求培训和帮助的机制，包括最佳实践建议、配置指导和故障修复解决方案。选择能够提供强大支持和培训方案的云提供商，为您的成功奠定基础。探索提供商的官方支持模式和资源，以及任何可用的第三方或基于社区的资源，例如博客、论坛、视频和操作指南。不仅要考虑提供商的技术支持计划，还要考虑侧重于业务和文化转型的计划。例如，AWS 云采用框架 (AWS CAF) 通过关注包括业务流程和人员在内的视角，而不仅仅是技术，来帮助组织实现数字化转型。优先选择能够提供广泛培训选项以及成熟、可靠支持模式和社区的云提供商。

建立一个 CCo E

考虑通过转型办公室或卓越云中心 (CCoE) 来发展您的云领导职能。 CCoE 开发并宣传了一种在整个组织中大规模实施云技术的方法。要成功采用云技术，请设计您的 CCo E，使其包括可以代表相关团队和部门发言的代表。从小处着手，逐步发展 CCo E 以满足您在转型之旅中的需求。您的主要云提供商代表（例如您的 AWS 客户经理和解决方案架构师）可以提供资源来指导您完成创建 CCo E。 CCoE

可以提高您建立主题专业知识、获得认同、赢得整个组织的信任以及制定满足任务要求的有效指导方针的能力。不存在适用于每个机构的单一组织结构，但是以下问题将帮助您设计自己的 CCo E.

- 你应该在你的 CCo E中加入谁？

  一开始， CCoE 可能只包括少数早期采用者和云端拥护者。 CCoE 可能仍然很小，但它应该演变为包括能够代表受云采用影响的业务职能和技术职能的拥护者。业务职能包括变更管理、利益相关者要求、治理、培训、采购和沟通。这些职能通常由您所在机构的行政和教学团队成员代表。技术职能包括基础设施、自动化、运营工具、安全、性能和可用性。这些职能通常由您所在机构的 IT 团队成员代表。如有必要， CCoE还应设法让供应商和合作伙伴参与进来，以提供主题方面的专业知识。CCoE 是一个活生生的组织。其成员资格、形式和职能可能会随着时间的推移而发生变化，甚至可能在未来某个成熟阶段解散。

- 欧盟如何与 CCo其利益相关者互动？

  CCoE 为其他团队服务，仅用于告知和实现云的成功采用。看看将 CCo E 的某些部分嵌入到各个部门、学校和职能中。这样便可实现对更广泛资源的访问以及更快速的内部反馈。重点在于尽早在利益相关者之间建立合作关系和开放的沟通渠道，以在机构内部建立信任并打破组织壁垒。 CCoE 应明确与利益相关者沟通、收集反馈和培训用户的机制。 CCoE的成功指标应反映这种合作和沟通。如果仅以构建技术来衡量一个团队，那么将会构建更多的技术，但其使用和成果却会被忽略。相反，你的指标应该衡量一些因素，例如通过 CCo E 的工作实现自给自足的团队数量、 CCo E 走上关键举措道路的次数、举办的培训活动数量或 CCo E 产出的采用广度。精心构建、值得信赖 CCo的企业可以成为建立在信任基础上的更大规模组织转型的垫脚石。

- 你应该如何建立 CCo E？

  大多数组织会从具体、有针对性的试点项目开始其云采用之旅。在这些项目中建立 CCo E。良好的开端对于整个旅程的成败至关重要。

  - 从业务问题入手。为了技术而技术是一种糟糕的策略。如果您正在尝试云技术，请确定一个有说服力的业务使用案例，无论它看起来多么微不足道。然后，根据该使用案例进行反向推导，设定技术如何提供帮助的明确目标。不要在孤立环境中实施解决方案。在项目实施之前和期间，不断听取业务利益相关者的意见。所有成功的云项目都依赖于与将使用该技术的机构部门的密切合作。
  - 从小处着手。选择一个提供“双向门”的低风险项目。这意味着该项目是可逆的，并且可以快速纠正任何错误。试点项目旨在进行实验。避免大规模、高风险的项目可以让您更好地控制实施和结果。它有助于针对具体的、可定义的问题，而不是设定宽泛的目标。例如，如果最终目标是自动化，那么就应该着眼于自动化特定任务，而不是整个工作。
  - 定义并衡量成果。设定明确的指标以评测每个项目的进度和表现。提前明确定义所需的最终状态，以避免利益相关者的期望不匹配。与业务利益相关者和组织内的其他领导者密切合作，以确定期望

和可衡量的收益。将结果转化为非技术语言也至关重要。从机构目标的角度进行阐述，例如项目如何提高客户留存率和降低客户流失率，如何降低成本和加快交付速度等。

- 从舒适区开始。在机构熟悉的领域中选择一个项目。通过这种方式，您可以确保该项目具备有意义、易理解且具有实际影响的目标。这样的项目将建立信心，为您的组织带来更好的长期结果。例如，如果您已经具备数据分析方面的专业知识，则可以从分析项目入手，在利用现有技能集的同时，开启您的云之旅。每个机构都有不同的专长，需要找到自身独特的组成要素，才能制定成功的数字化转型战略。

区分 SaaS 应用程序和基础云服务

大多数教育机构已经采用软件即服务（SaaS）应用程序。SaaS 为您的机构提供由服务提供商运行和管理的完整解决方案。常见的 SaaS 应用程序包括文字处理和电子邮件等生产力应用程序，但许多任务关键型工作负载中也存在 SaaS 选项，例如企业资源规划（ERP）、学生信息系统（SIS）和学习管理系统（LMS）。当您的机构采用 SaaS 产品时，您的 IT 团队不必考虑如何维护服务或如何管理基础设施，您的用户只需使用该服务即可。这种交付模式减轻了 IT 员工的管理负担。许多机构选择在其 IT 策略中采用“SaaS 优先”方法，特别是其 IT 团队没有足够的时间、资源或技能来充分自行托管相同的应用程序时。即使您有可以自行托管的资源，采用 SaaS 解决方案并将资源投入到其他项目中，可能仍然更具成本效益。

当您使用 SaaS 应用程序时，您的 IT 团队不必管理底层基础设施，因此供应商托管应用程序的位置（本地数据中心、您的主要云提供商或备用云提供商）变得不那么重要。选择主要战略云提供商后，您可以选择使用托管在其他云提供商或本地供应商数据中心内的 SaaS 产品。相反，即使您的 SaaS 应用程序托管在一个云提供商中，您也可以根据该提供商在非 SaaS 工作负载方面的优势选择不同的主要战略云提供商。对于 SaaS 来说，托管环境之间的区别不如自托管应用程序那么重要。但是，在评估 SaaS 如何作为您的 IT 策略的一部分融入云时，仍应考虑以下关键问题。

- SaaS 应用程序是否具有高可用性和可扩展性？

  许多供应商已经决定在其 SaaS 产品中采用云。这样供应商就能享受到云带来的提高可用性和可扩展性的优势。此外，由于供应商可以采用云责任共担模式，而不是管理和维护物理基础设施，因此其可以投入更多的时间和资源来提供新功能。由于这些优势，您应该更喜欢云优先并提供云托管解决方案的提供商。

- SaaS 应用程序能否满足您的安全要求？

  在评估 SaaS 时，务必了解应用程序存储了哪些数据、如何使用这些数据以及采取了哪些安全控制措施来保护这些数据。尽管您可能无法像在自有的自托管环境中那样直接控制数据存储，但应确保供应商拥有相应的机制和控制措施来妥善处理您的数据。注意哪些安全功能是 SaaS 解决方案内置的，哪

些功能需要额外配置。云使 SaaS 提供商能够构建更多可用和可扩展的解决方案，并且由于责任共担模式，他们还可以构建更安全的解决方案。您应该更喜欢利用云安全工具和服务作为其解决方案一部分的提供商。

- 哪些人员拥有 SaaS 应用程序数据？您如何访问这些数据？

  使用 SaaS 时，您信任提供商能够妥善处理您所在机构的数据。请务必查看 SaaS 应用程序的服务条款和服务水平协议，以了解数据所有权、可用性和耐久性等影响因素。评估备份或导出数据的机制；如果您决定更换提供商或提供商停止服务，这些机制尤其重要。

- 您的其他服务和自托管应用程序是否能与 SaaS 应用程序集成，而无论环境如何？

  在采用 SaaS 解决方案时，很容易假设共享相同托管环境的服务和应用程序（即使用同一云提供商或同一供应商数据中心的应用程序）将实现更加无缝的集成。但是，当今大多数 SaaS 解决方案都广泛支持 API 和第三方集成，因此不必将自身局限于在同一环境中托管的解决方案。如果存在必要的集成，这些解决方案不必共享相同的底层环境。例如，假设你正在使用 SaaS 解决方案，例如 Google 云端硬盘或 Microsoft 来存储基于云 OneDrive 的学生文件。要向学生提供虚拟桌面和应用程序流，您可能会认为 Amazon Appl WorkSpaces ic ations最适合您的要求。尽管这些服务在不同的环境中运行，但 WorkSpaces 应用程序已与 Google 云端硬盘和 Microsoft 进行了原生集成 OneDrive，因此您的学生可以继续使用其现有存储空间。

- SaaS 应用程序是否支持集中式身份管理？

  为防止您的 IT 团队不得不管理分散的身份存储和避免用户必须记住多组凭证，请确保您的 SaaS 解决方案支持与您现有的身份管理或单点登录解决方案集成。分散的身份管理会降低工作效率，并可能导致权限蔓延和弱密码等不良安全实践。如果所需的 SaaS 解决方案不支持单点登录或您现有的身份存储，请评估采用该解决方案的业务价值是否超过给用户和员工增加的负担。

- 如何保护与 SaaS 应用程序的网络通信？

  在某些情况下，您可能需要自托管应用程序才能与 SaS 应用程序通信。通常，这种通信将通过 APIs 适当的身份验证和授权机制进行保护。但是，根据这两个应用程序的托管环境，可能需要备选或额外的机制来简化或保护该通信。例如，如果您向云提供商自托管应用程序，并且需要将其与在同一云提供商托管的 SaaS 应用程序集成，则该供应商可能会提供多种连接选项。你可以使用云特定的对等连接 APIs、私有接口或私有接口，例如AWS PrivateLink防止该通信通过公共互联网。同样，如果您的本地应用程序通过 AWS Direct Connect 等服务与云提供商建立专用网络连接，则您可以使用相同的连接与在同一云提供商托管的 SaaS 应用程序进行通信。

为每个云服务提供商制定安全和治理要求

教育机构必须实现各种合规、治理和网络安全目标。未能实现这些目标的风险可能包括机构声誉损失、罚款、勒索、敏感数据泄露、知识产权剽窃，以及关键任务功能降级或完全丧失。由于责任共担模式，采用云服务的机构可以通过将基础设施安全的部分责任分载给云服务提供商来减轻管理负担。此外，您可以受益于专门构建的云原生安全服务，这些服务提供的功能在本地部署中通常不可用、难以管理或成本高昂。示例包括 Web 应用程序保护、分布式拒绝服务防护和AWS Shield GuardDuty用于威胁检测的 Amazon 等服务。AWS WAFDDo成功的云安全和治理策略使 IT 和安全团队能够专注于构建“通过设计确保安全”的系统，帮助机构快速适应不断演进的任务要求，并为教职人员和研究人员提供安全的环境，以支持突破性的学习与创新。要评估您的安全和治理要求，请考虑以下关键问题。

- 您的工作负载必须与哪些合规框架保持一致？

  由于教育机构支持的利益相关者和工作量众多，因此必须遵循许多合规框架。这些合规框架包括《家庭教育权利和隐私法案》（FERPA）、《健康保险流通与责任法案》（HIPAA）、联邦风险与授权管理计划（FedRAMP）、网络安全成熟度模型认证（CMMC）、《美国国际武器贸易条例》（ITAR）、刑事司法信息服务（CJIS）和支付卡行业数据安全标准（PCI DSS）。在某些情况下，例如 CMMC，研究经费只有在相关工作负载获得合规认证后才会发放。每个框架都有其独特性，可能仅适用于部分工作负载。请确保您了解哪些工作负载必须遵循哪些要求，并且能够在每个工作负载的环境中实现这些要求。在云环境中，请务必厘清自身责任与云服务提供商责任之间的划分。您应该具备实现和保持合规所需的知识、资源和技能集。

- 您采用了哪些机制，能够在不阻碍创新的前提下，在多个云提供商之间强制执行合规？

  如果您的学术机构不熟悉云，我们建议您选择一家主要战略云服务提供商，并重点了解如何架构、设计和运营“通过设计确保安全”的云环境。理想情况下，自动嵌入自助服务系统中的安全控制措施允许用户在最少 IT 团队干预的情况下快速部署安全的云环境。专注于单一提供商会限制您为确保安全与合规而必须投入的资源和时间。最成功的机构会选择能够支持大多数合规要求、拥有强大合作伙伴网络、提供预先构建的合规解决方案并提供安全自助服务自动化功能的云服务提供商。如果必须确保多个云提供商的安全与合规，则需要额外投资来构建管理每个环境合规性所需的技能集和资源。如果每个云提供商都使用不同的基础环境或登录区，则需要了解每个登录区可以支持哪些合规标准和要求，这可能会决定某些工作负载是否可以在该提供商托管。您可以单独管理每个提供商的合规性，也可以使用自定义构建的解决方案或合作伙伴解决方案，从而能够跨提供商集中管理。AWS Marketplace 提供可满足合规要求的交钥匙解决方案。

- 如何评测和控制跨多个云提供商的成本和使用情况？

  如果您的学术机构不熟悉云，我们建议您建立成本可见性和控制机制，以深入了解正在使用哪些云服务、云资源的归属、这些云资源的用途以及通过优化使用可以实现哪些潜在的成本节省。机构可以通

过与云服务提供商合作迁移与现代化任务关键型系统来获得可观的投资回报，因为他们可以协商企业级协议，享受批量定价优惠，并利用云服务提供商的专业知识。如果您必须控制跨多个提供商的成本和使用情况，请考虑如何汇总和分析每个提供商的成本和使用情况，可以通过内部流程和工具，也可以使用合作伙伴解决方案。许多组织开始将云财务运营 (FinOps) 确定为一项关键职能，并投入资源来宣传和实施云成本管理和优化的能力。

- 您是否已建立能够随着时间的推移轻松管理用户权限的机制？

  我们建议学术机构在首次使用云时了解核心利益相关者的需求。机构系统的用户包括学生、教职人员、研究人员、IT 人员、行政人员、安保人员、公众以及第三方协作者。您应该明确这些用户的核心需求，并确保落实适当的机制以授予其访问云服务的权限。不同类型的用户需要不同类型的云服务访问权限。例如，学生、教职人员和公众需要访问应用程序；IT 人员、行政人员和安保人员需要访问云基础设施；研究人员及其第三方协作者需要访问安全的研究环境；教职人员需要访问安全的教学环境，甚至可能希望为学生提供访问云技术的实践操作。您应该准备好工具，以自动方式集中管理这些身份，并使用既定流程在角色和职责随着时间的推移而发生变化时识别、授予和撤消权限。

- 您是否已建立将新系统与您的身份管理解决方案妥善集成的机制？

  我们建议学术机构简化新系统与其身份管理系统的集成。通过允许利益相关者购买和构建可轻松集成到身份管理系统的系统，使机构能够灵活地支持各种任务关键型职能。通过简化集成流程，利益相关者将不太可能使用自己的访问控制措施，因为这些措施可能无法强制执行单点登录、通行密钥和多重身份验证（MFA）等安全最佳实践。确保您的身份管理系统能够通过原生集成或行业标准协议与必要的系统互操作。

- 您是否已建立启用有效事件检测和响应的机制？

  教育机构经常成为网络攻击和勒索软件的目标。为了帮助有效检测和应对此类事件，我们建议采用双轨方法：

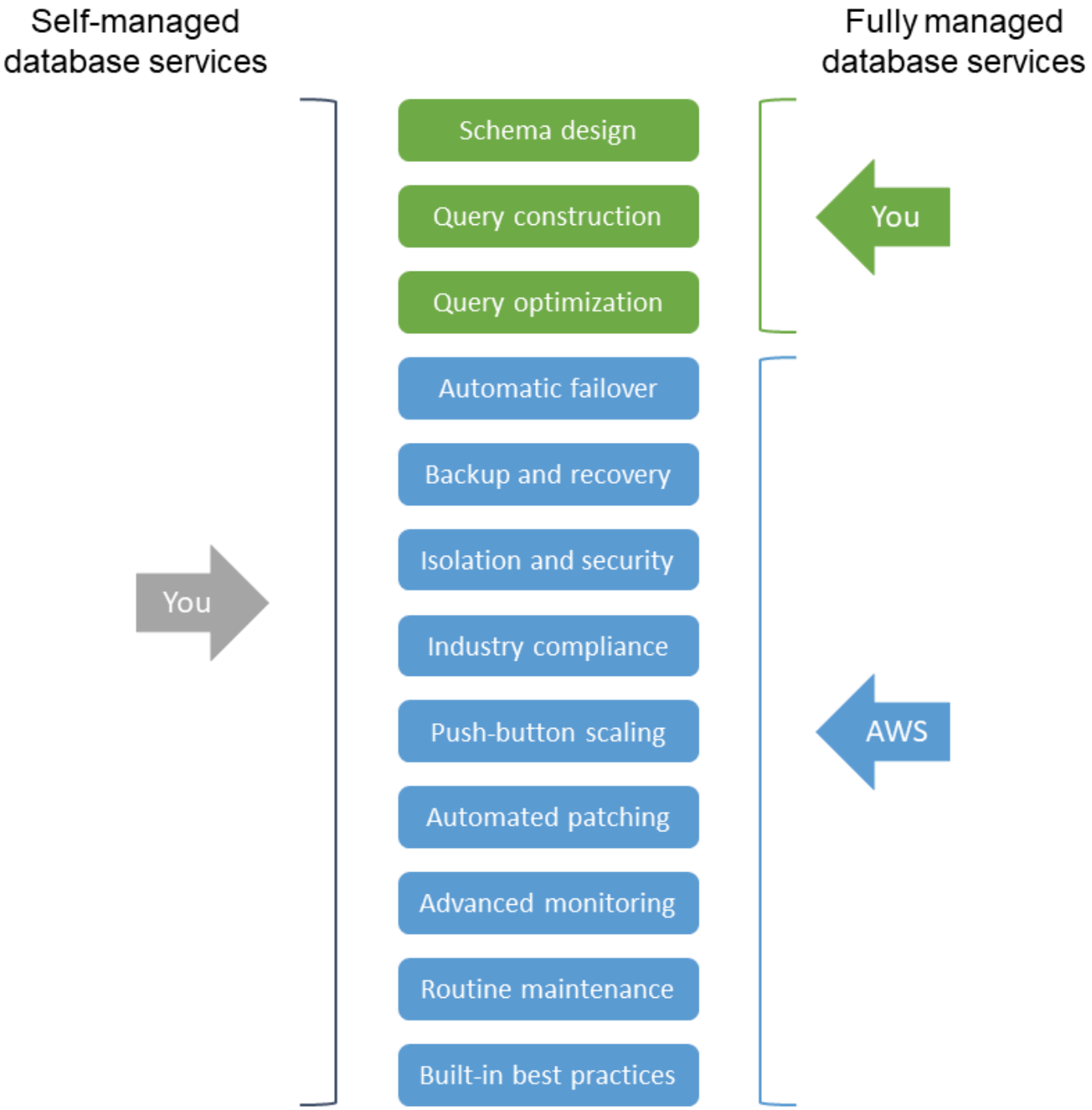
  - 将精力集中在预防性措施，即自动嵌入云环境中的安全控制措施。
  - 实施检测能力，帮助网络事件响应者及时检测、遏制和缓解安全漏洞。

与合规一样，您必须确保具备在每个环境中检测、预防和响应事件所需的资源、技能集和工具。通过专注于单一主要云提供商，您可以限制所需的资源。没有成熟安全运营团队的学术机构应向独立软件供应商、托管检测和响应提供商以及网络安全顾问寻求这些领域的帮助。

在切实可行的情况下，尽可能采用云原生托管服务

当您最初考虑如何利用云服务时，使用团队熟悉的基础设施服务和开发工具似乎是最佳前进路径。但是，选择云原生托管服务，尤其是无服务器选项，可以显著降低成本、工作量和复杂性。

云原生托管服务消除了许多无差异化 IT 任务，这些任务需要员工付出时间和精力，使其无法更好地专注于以任务为中心的活动。此外，随着提供商提升其服务能力，您的解决方案会自然继承效率、安全性、韧性、性能和其他特征方面的渐进改进。例如，完全托管的数据库服务是一个功能丰富的关系数据库管理系统，但您不必预调配和管理运行数据库的底层服务器和操作系统。这样可以省去您在自己的数据中心或在云中预调配的自托管虚拟服务器上维护关系数据库时通常需要执行的管理任务。下图展示了这种差异。



将任何云原生托管服务与类似的自托管方法进行比较时，取消基础设施管理的优势便显而易见。因此，每当您需要部署将在其上运行已购买或自定义开发应用程序的组件时，都应使用云原生托管服务，以减少所需的时间和精力。

当您的团队负责在云构建、部署或管理解决方案时，请使用云原生托管服务，以充分利用云提供商的差异化功能和创新。该策略使您能够选择、集成和部署云服务，从而减少这些项目所需的时间和精力，同时提高其韧性和安全性。要成功制定云策略，请在将自定义解决方案迁移到云、在云开发新解决方案或在云部署许可软件时考虑采用这些云原生构建数据块。在评估云原生托管服务的选项时，请考虑以下关键问题。

- 您是否需要将更多员工时间和精力集中投入到教育使命的核心功能上？

  管理服务器（即使是虚拟服务器）也需要时间和精力，以确保其与系统软件升级和补丁保持同步。使用为您处理这些任务的托管服务，可以让 IT 员工将时间用于更直接符合机构使命的活动。例如，如果您需要部署容器，可以考虑使用 AWS Fargate 等无服务器托管服务，这样就无需配置和维护服务器了。通过消除采购、预调配和管理底层基础设施的需求，您可以专注于提供新功能、优化性能和改善用户体验。在根据自托管选项评估托管服务时，请考虑这一优势。

- 您的团队需要完成哪些工作才能采用云原生托管服务？

  使用云原生托管服务设计和实施解决方案可能会存在学习曲线，但是这些投入将通过在解决方案的生命周期内降低成本、缩短时间和减少复杂性获得回报。由于云计算的按需 pay-as-you-go性质，云原生服务使您能够以更敏捷的方式快速迭代和试验，同时避免前期投资。这将增加创新，并缩短项目周期。但是，要有效实现这些好处，请考虑采用和使用该服务可能需要什么，例如对员工进行有关最佳使用模式的培训，以及为适应特定服务而进行的代码重构。 APIs即使该服务使用行业标准或开源 APIs，您也可能需要重构或配置应用程序以处理功能差异或版本不匹配问题。

- 您目前如何部署和管理基础设施？ 是否需要保持该控制水平？

  在云托管和管理基础设施的方式多种多样，包括使用裸机主机、虚拟机、托管容器服务和无服务器产品。即使您目前在本地环境中使用虚拟机或容器等类似的基础设施，也要考虑是否有其他方法适合某些工作负载。例如，与其在虚拟机上运行所有应用程序，不如考虑容器化应用程序，并利用 Amazon Elastic Container Service（Amazon ECS）等托管容器服务。这可能需要重构，但您可以使用 AWS App2Container 等工具来简化和协助容器化。更进一步，与其为所有组件部署服务器或容器，不如考虑完全无服务器的选项。无服务器技术具有自动扩展、内置高可用性和 pay-for-use计费模式，可提高敏捷性并优化成本。同时，其无需管理服务器和规划容量。AWS Lambda 等无服务器计算服务是无服务器架构的核心。Lambda 支持常见的编程语言，允许开发人员专注于应用程序代码，而不是管理基础设施。探索每种工作负载的这些选项，并考虑学习曲线、管理开销、成本和许可等因素。

- 是否必须为任何许可软件部署和管理基础设施？

在部署和管理来自独立软件供应商 (ISVs) 的许可软件时，使用云基础设施模仿本地部署似乎是合乎逻辑的。例如，您可考虑将本地虚拟机替换为云托管的虚拟机。尽管这是一个可行的选择，但请考虑是否可以将架构的任何组件替换为云原生托管服务。例如，您可以将自托管数据库服务器替换为完全托管的数据库服务，从而在运行相同的数据库引擎时减轻管理负担。许多人 ISVs已经在使用利用托管服务的云架构，甚至可能提供预先构建的模板来简化部署。在可能的情况下，您应该更倾向 ISVs 于为云部署提供规范性指导和支持。在将许可软件部署到云之前，请务必咨询您的 ISV，以了解云环境许可与本地许可有何不同。

- 您是否担心使用托管服务可能会导致供应商锁定？

  许多云原生托管服务都是为支持常见的行业标准而构建的， APIs. 例如，诸如AWS Glue和 Amazon EMR 之类的分析服务建立在 Apache Spark 和 Apache Parquet 等行业标准处理和存储框架之上。 AWS Lambda原生支持 Java、Go、微软 PowerShell、Node.js、C#、Python 和 Ruby 代码。Amazon Relational Database Service（Amazon RDS）支持多种版本的常用数据库引擎，包括 SQL Server、Oracle、PostgreSQL 和 MySQL。当服务具有专有解决方案时 APIs，可以使用与云无关的通用协议与 APIs之交互的原生解决方案或合作伙伴解决方案。例如，Amazon Simple Storage Service（Amazon S3）拥有直接集成的服务特定 API，但您也可以在使用 AWS Storage Gateway 时通过网络文件系统（NFS）、服务器消息块（SMB）和互联网小型计算机系统接口（iSCSI）等标准存储协议与其交互。您仍应专注于选择最符合自身需求的云原生托管服务，同时最大限度地降低运营开销，但您可能更愿意使用或提供通用行业标准和协议的服务。

在现有本地投资激励继续使用时，实施混合架构

大多数教育机构都投资了不同规模的本地数据中心，以托管企业应用程序、数据存储解决方案、终端用户计算（EUC）环境和共享计算资源。这些数据中心中的所有资源都要遵循不同的刷新周期，在此期间，您必须考虑未来的增长，并预调配足够的容量以容纳峰值规模，而这种规模每年可能只需要几次。因此，资源经常处于闲置状态，直到下一个刷新周期。规划、预算、采购和部署新硬件可能需要数周、甚至数月或更长时间。这种漫长的过程扼杀了创新，并可能延误学习和研究。

云计算解决了其中许多挑战。云提供按需 pay-as-you-go的 IT 资源，因此您可以更紧密地将当前容量与实际需求相匹配，而无需进行大量的前期规划和投资。但是，如果您已经对本地硬件和资源进行了大量投资，则应寻求高效利用这些资源，并根据需要在混合模式下使用云技术对其进行扩充。

成功的混合云策略可以利用现有投资，同时提供比仅靠这些投资能够支持的更高的敏捷性、可扩展性和可靠性。以下考虑因素可帮助您开始使用。

- 当您必须托管新的工作负载时，是否会先考虑云？

如何同时使用公有和私有云基础设施，定义了您的混合云策略。云优先方法并不意味着云是所有工作负载的更好选择。但是，在规划新工作负载时，应将云作为第一选择进行评估，特别是对于需要新技术或超出本地可用存储和计算容量的工作负载。具有瞬态、不一致使用模式、需要快速结果、易于移植或需要最新硬件的工作负载是实现云可扩展性和弹性的理想选择。此外，还要考虑工作负载是否可从任何本地不可用的云原生托管服务受益，即使您确实有可用容量。

- 您是否了解本地环境的 TCO，并在进行新投资时与 CFO 合作？

  我们建议您了解维护自己本地数据中心的真实总拥有成本（TCO）。在本地拥有和运营基础设施有许多隐性成本，不仅包括硬件、软件和支持，还包括设施、公用事业、保险和员工工时。这些成本会对员工的工作效率、运营韧性和业务灵活性产生负面影响。同时评估您当前的许可结构及其续订和维护期限。与您的首席财务官（CFO）合作，可以帮助您在计划进行新投资时明确所有隐性成本。有些许可证可能在云提供自带许可（BYOL）选项，或者其可能或多或少有利于云服务。了解当前基础设施的真实 TCO 有助于您优先将组织总 TCO 影响最大的工作负载采用云。您的 AWS 客户团队拥有随时可用的工具，可帮助您更好地了解本地总拥有成本。

- 您需要哪些基础设施来支持混合部署？

  要成功采用混合模式，您将需要基础网络、安全和基础设施工具。确保能够与云提供商保持足够的网络连接。这可以通过将现有的互联网连接、虚拟专用网络 (VPNs)、专用连接（例如 Direct Connect 第三方连接提供商）或 Internet2 以及区域研究和教育网络结合起来实现。确保您的本地和云环境中拥有统一的身份和访问管理。建立工具和流程，以强制实施一致的安全、成本和使用护栏措施。

- 您的 IT 员工是否已准备好运营混合部署？

  云服务可能需要您的团队不具备的特定技能集。为减少提升 IT 人员有效云采用技能所需的培训和赋能，请考虑云提供商是否提供任何在本地和云可重复使用并利用现有技能集构建的服务。例如，如果您使用并熟悉 Kubernetes，则可以考虑使用 Amazon Elastic Kubernetes Service（Amazon EKS）或 Amazon EKS Anywhere。如果您使用并熟悉 NetApp，则可以考虑使用 Amazon FSx for NetApp ONTAP。同样，还要考虑您使用的任何现有合作伙伴解决方案是否有原生集成或支持云环境。

- 您是否能将长期存储或低使用率计算从本地分载到云？

  云存储为长期数据存储提供了几种经济实惠的选择。例如，Amazon Simple Storage Service（Amazon S3）提供针对不同使用案例优化的各种存储层。如果您的机构需要长时间保留某些数据，请考虑使用 Amazon Glacier 等冷存储解决方案。将这些数据分载到云存储可以腾出宝贵的高性能本地存储。AWS Storage Gateway 等服务使本地应用程序可以轻松地通过 SMB、NFS 和 iSCSI 等标准协议访问云存储层。同样，可以考虑分载所有不常用或使用率较低的计算任务。如果您有专用于此类任务的本地服务器，则可以改用可扩展的云计算服务，即按需预调配资源并只需为实际

用量付费。这些低成本、长期存储和低使用率的计算选项也使云成为备份和灾难恢复的理想选择。您可以在云使用安全、耐用、可扩展的存储和计算来保护您的数据，并在发生灾难时快速恢复，而无需自行维护必要的存储和计算基础设施。

- 您是否有足够的本地容量进行实验和创新？

  在固定规模的本地环境中，缺乏弹性和敏捷性可能会限制用户可用的服务和技术。如果您有严格的刷新周期，则新工作负载可能需要等到下一个周期才能实施。这种运营模式可能会限制实验并减缓创新。当您需要测试新的或创新的工作负载时，可以考虑使用可扩展的弹性云服务。云资源可以按需预调配和取消预调配，您只需为实际使用的资源付费，因此您可以快速实验和实施快速失效机制，同时最大限度地降低组织风险。

- 是否有独特的合规或性能要求，迫使您将数据保存在本地？

  具有严格数据驻留或延迟要求的工作负载可能会要求您将数据保留在本地或尽可能靠近用户的地方。对于这些使用案例，您可以优先使用现有的本地资源。但是，请考虑云提供商是否提供在本地使用基于云技术的边缘服务或机制。边缘服务在离您自己的端点更近的地方提供数据处理、分析和存储，并使您能够在标准云提供商数据中心之外部署工具。例如，AWS 提供 AWS Local Zones 和 AWS Wavelength 等服务，以在离最终用户更近的特定位置部署应用程序。您还可以使用 AWS Outposts、AWS Storage Gateway、Amazon ECS Anywhere 和 Amazon EKS Anywhere 等服务将云服务和功能引入现有的数据中心。

仅为无法通过单一云提供商满足其技术或业务要求的工作负载保留多云

多云是指使用多个（两个或更多）云服务提供商的云服务。制定多云策略可以提供某些优势，例如可以选择解锁多个云提供商的差异化功能，或者能够满足单一云提供商可能无法满足的数据主权要求。但是，对于您使用的每个提供商，请确保您拥有适当的人员、技能、培训和工具集，以便有效地使用该提供商。此外，如果您想对特定工作负载使用多云策略，则需要额外的资源来集成每个云提供商的必要服务并进行互操作。我们建议您仅在收益大于增加的投资时才考虑使用多云环境。要确定是否应选择多云策略，请考虑以下关键问题。

- 您是否有足够的资源和技能集驾驭不同云提供商提供的服务？

  多个云提供商提供各种产品和服务时，您的员工需要具备基本技能才能驾驭每个提供商的能力。仅使用一家云提供商的服务可能需要提升员工的技能和进行培训，具体取决于您使用的服务和功能。如果您正在考虑多云策略，请评估您的现有资源，以确定要有效使用来自多个云提供商的服务需要具备哪些额外的技能集。除了单一云提供商所需之外，您可能需要增加员工或投入更多时间和金钱来提升

技能和进行培训。如果您已经有单独的团队或用户在使用不同的云提供商，请考虑将他们 case-by-case逐一整合到主云提供商对组织带来的好处。

- 特定的多云架构会带来哪些额外开销？

  多云的一个常见驱动因素是希望使用一家提供商的特定托管服务，该服务具备与其他云提供商服务不同的能力。例如，您可能希望使用一家云提供商满足您的基础设施需求，而使用另一家提供商的托管服务提供域和目录服务。但是，即使该单一托管服务减轻了管理负担并简化了该架构组件的管理，仍可能会给代码重构、私有连接需求或手动集成工作等其他工作负载带来额外的开销。预先确定此额外开销，并确保其不会抵消或掩盖您的团队将从差异化服务中获得的收益。

- 如何实现跨云供应商的集中监控与管理？

  在开始使用来自不同云提供商的资源来部署应用程序和功能时，请考虑如何标记、监控和管理此类资源。每个提供商都有自己的工具，您可以将其扩展到其他环境。例如，您可以使用 Amazon CloudWatch 监控关键指标和日志，创建警报，并可视化单云、混合云和多云环境中的应用程序和基础架构。您还可以使用 AWS Systems Manager 来提升资源可见性和控制能力，快速诊断和修复运营问题，并自动化跨环境更新和修补虚拟机等流程。如果您有提供商工具无法支持的要求，则可以探索合作伙伴解决方案，但这可能会增加额外的成本或集成工作。

- 使用不同的云提供商时，如何通过自动化管理基础设施即代码？

  当您在云运行资源时，自动化预调配和管理资源可帮助您高效地管理各种环境。 APIs 和原生自动化工具因云提供商而异。如果可能，请考虑使用一组可容纳不同云提供商资源的通用编排和部署工具。这提供了更大的灵活性并简化了跨多个云的运营。但是，单独使用每个提供商的原生自动化并建立组织流程以确保适当使用可能更简单。

- 您是否有每个云提供商都必须满足的合规与监管要求？

  您可能有监管方面的考虑，这些因素决定应如何存储和处理数据。重点是标准化策略（例如网络流量、存储和安全），使其能够自动应用到跨云提供商的每个云环境。考虑您的应用程序将如何与其数据通信，并将其托管到同一个提供商。如果您的应用程序及其数据分散在各个提供商之间，则很难确保您满足合规与监管要求。通常最好让应用程序尽可能靠近数据，以最大限度地减少网络延迟、最大化数据吞吐量并限制数据传出，同时简化安全和访问控制。

- 跨云提供商部署应用程序时，是否能够最大限度地降低 TCO 并最大限度地提高定价折扣？

  考虑多云时，务必将总拥有成本（TCO）考虑在内。跨多个云提供商运行应用程序可能会增加维护和管理每个环境中资源的运营成本和管理开销。此外，将使用量分散到多个提供商会使您难以利用特定提供商的批量定价折扣或企业协议。在判断多云带来的收益是否值得增加 TCO 时，请一并考虑这些因素。

使用案例示例

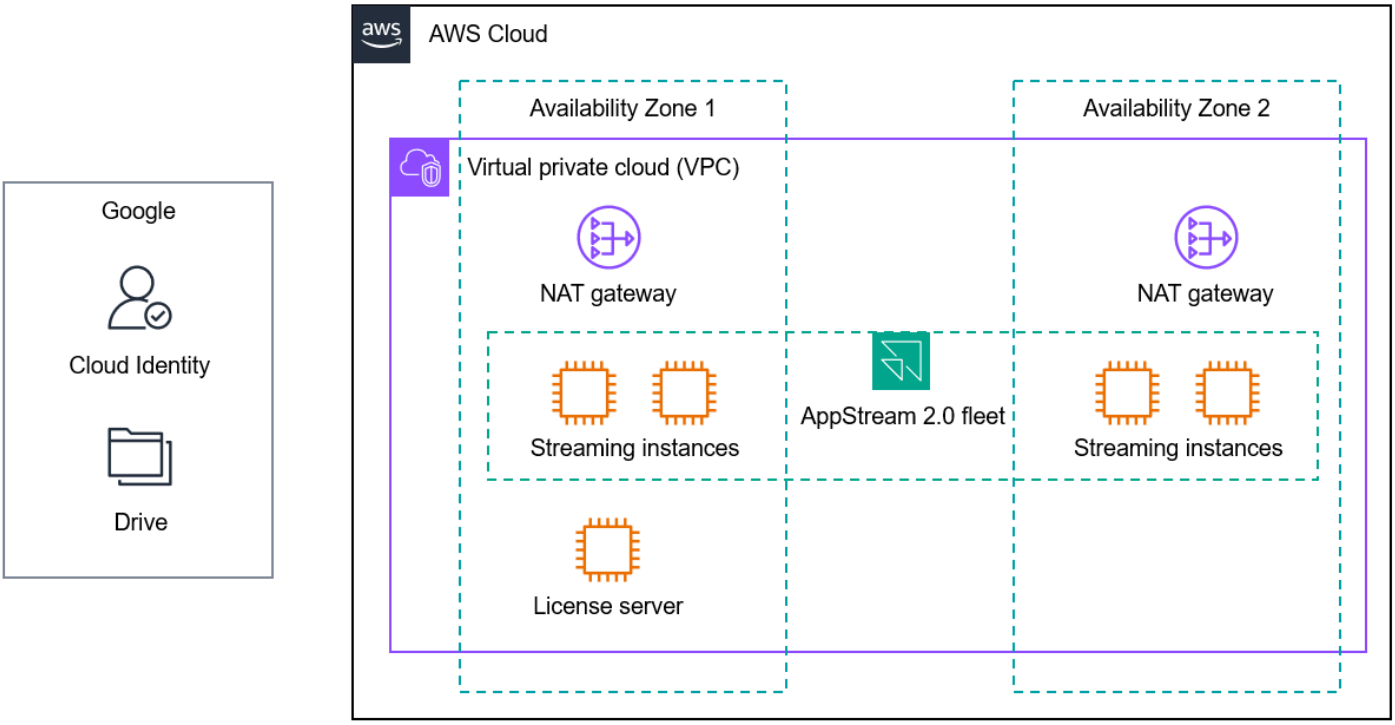
为了更好地理解这些原则在不同场景中的应用，让我们来讨论一些示例使用案例。这些使用案例基于现实世界中教育机构采用云服务的方式。

- 虚拟计算机实验室
- 预测学生成功
- 身份联合验证和单点登录
- 研究计算的云爆发

虚拟计算机实验室

尽管基于 Web 的学习工具很受欢迎，而笔记本电脑、Chromebook 和平板电脑等用户设备也日益普及，但大多数教育机构仍保留了实体计算机实验室，用于资源密集型或旧版应用程序。这些计算机实验室通常是科学、技术、工程和数学（STEM）、职业和技术教育（CTE）、媒体和艺术、工程以及类似课程的必备设施。学校可以使用基于云的虚拟桌面或应用程序流服务来增强或取代实体计算机实验室，以确保所有学生都可以随时随地在任何设备上访问所需的应用程序。这可提升数字公平性，实现远程学习，确保一致的用户体验，并在降低成本的同时保护远程访问。

在小学和中学（K12）教育中，许多美国学校使用完全托管的桌面和应用程序流媒体服务亚马逊 WorkSpaces 应用程序来提供虚拟计算机实验室，以提供对Adobe Creative Cloud、Autodesk软件、STEM和CTE课程（例如PLTW）等的访问权限。许多 K12 组织已经通过 SaaS 应用程序 Google Workspace 和 Google Drive 管理学生单点登录和文件存储。这些机构可以通过 SAML 2.0 联盟在 Google Workspace 和 WorkSpaces 应用程序之间设置单点登录。他们还可以在 WorkSpaces 应用和 Google 云端硬盘之间配置原生集成，以便学生可以使用现有存储空间。下图说明了此用例的 WorkSpaces 应用程序部署。



此架构遵循以下建议：

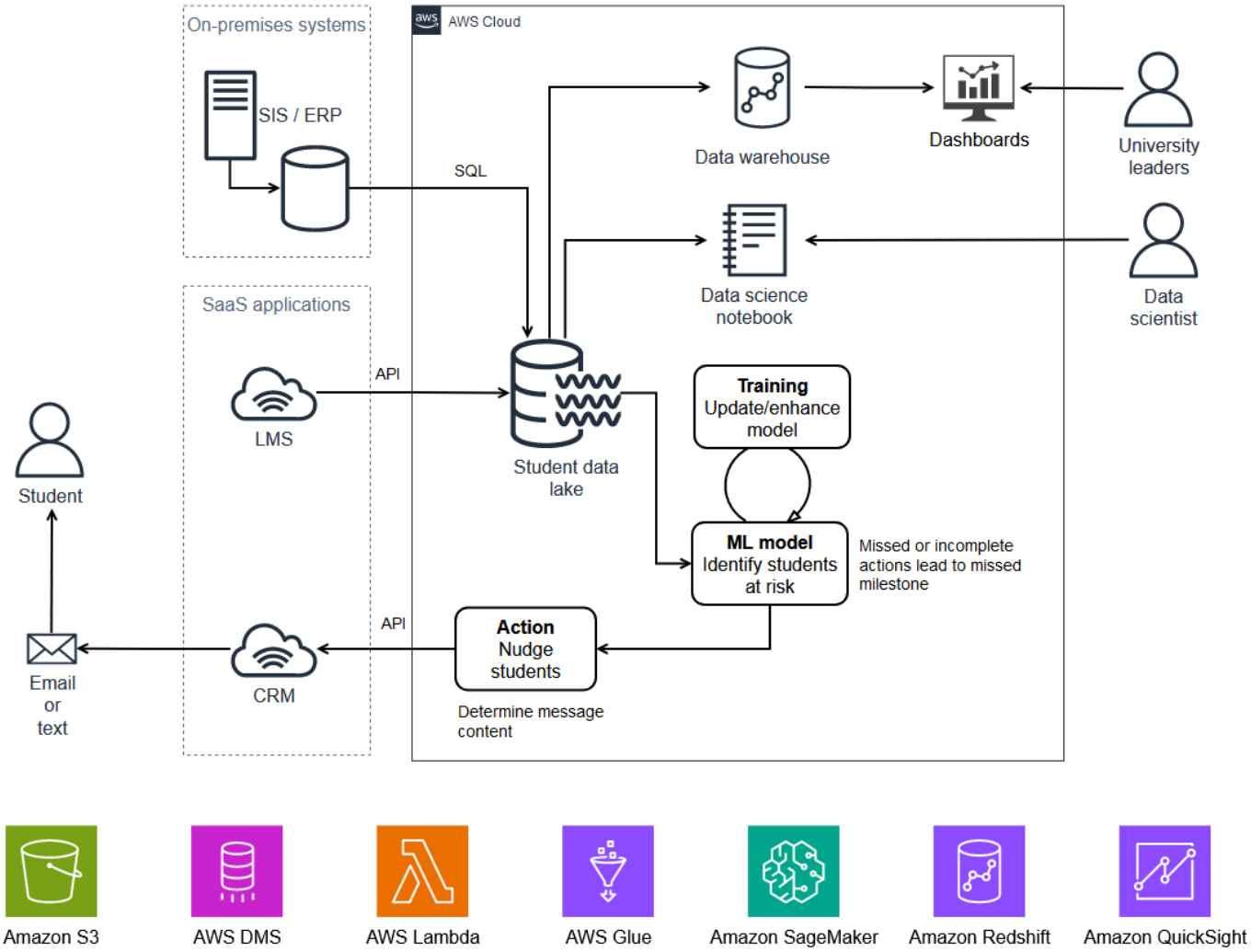
- 选择主要战略云提供商。此架构使用来自一个主要云提供商的云服务。尽管其包含与不在同一提供商上托管的 SaaS 应用程序的集成，但这些集成只需简单的配置即可完成。只有部署和管理主要云提供商的服务时，才需要云专业知识和技能集。
- 区分 SaaS 应用程序和基础云服务。Google Workspace 和 Google 云端硬盘与 AppStream 2.0 托管在同一个云提供商上，但这是可以接受的，因为这种部署提供了必要的集成。单点登录可实现集中式身份管理，并通过 SAML 2.0 进行安全配置。要为学生启用永久云存储，只需在 Google 云端硬盘和 WorkSpaces 应用程序中进行简单的配置更改。
- 为每个云服务提供商制定安全和治理要求。此架构中使用的服务和集成可帮助满足机构的安全和治理要求。流式传输流量已加密。通过 Google Workspace 进行联合身份验证以实现集中式身份管理。Amazon Virtual Private Cloud（Amazon VPC）等网络服务支持子网、路由和防火墙的配置。您可以使用 DNS 配置、代理、虚拟设备或 Amazon Route 53 Resolver DNS 防火墙等托管服务来筛选内容。您可以使用诸如的服务AWS Control Tower来帮助确保托管 WorkSpaces 应用程序的 AWS 账户遵守标准的组织护栏和控制措施。
- 尽可能采用云原生托管解决方案。 WorkSpaces 应用程序是一项用于桌面和应用程序流的托管服务。您可以流式传输桌面和应用程序，而无需预调配、扩展或维护服务器。您只需安装应用程序，连接相应的身份、网络和存储解决方案，然后集中管理这些应用程序并将其流式传输给您的用户。这省去了管理自有虚拟桌面流式传输解决方案所需的大量无差异化繁重工作。

预测学生成功

美国中西部的一所大学发现，对于新入学的一年级学生来说，一些关键活动能够高度预测其是否能成功完成第一学期的课程以及最终获得学位。该大学希望实施一个系统，监控这些活动的完成情况，并在关键截止日期临近或已过时，鼓励学生完成这些步骤。

SaaS 学习管理系统（LMS）数据是该解决方案的关键输入，但事实证明，使用大学 IT 团队的数据仓库工具难以访问和处理其数据。此外，发送给学生的消息必须通过学校基于云的客户关系管理（CRM）系统发送。为了构建功能性解决方案并评测给学生的提示的有效性，该大学必须通过 CRM 启动消息并从中收集数据。

该大学在单云环境中开发并部署解决方案。该解决方案混合了云原生托管服务、预调配云服务器以及与本地系统和基于云的 SaaS 应用程序的集成。如下图所示，该解决方案将来自学生信息系统（SIS）、LMS 和 CRM 的数据摄取到数据湖中。它使用这些数据来识别有可能错过关键活动的学生，通过 CRM 向其发送消息，并为大学领导层提供控制面板。



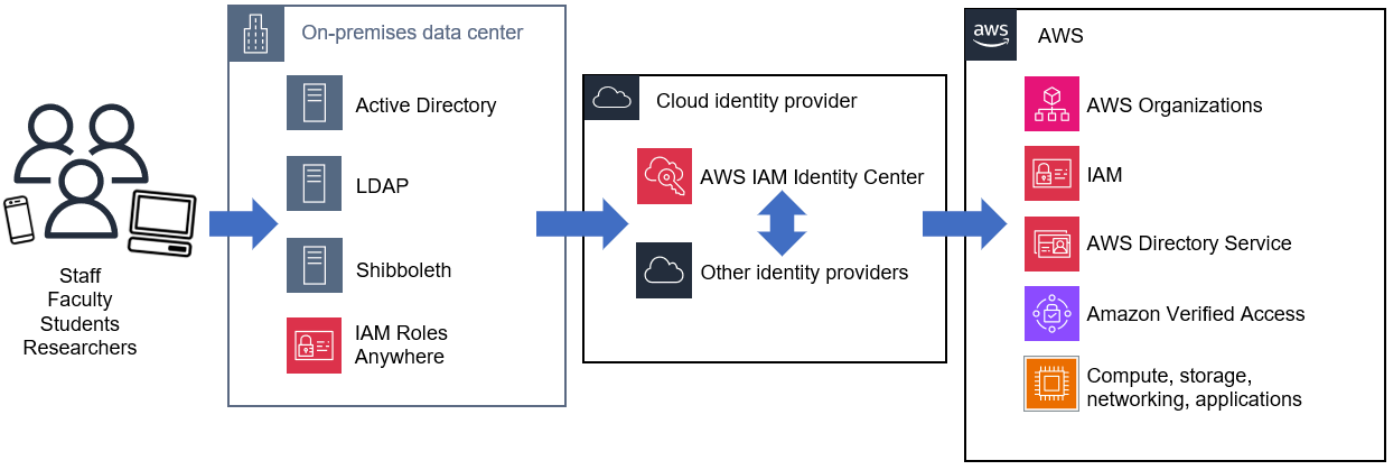
此架构遵循以下建议：

- 选择主要战略云提供商。该大学的战略云提供商容纳已部署的全部解决方案。这使得 IT 和业务人员能够专注于提升单一集成云能力集的技能。
- 区分 SaaS 应用程序和基础云服务。该大学会区分 SaaS 应用程序与核心云分析服务，并使用与 SaaS 应用程序的集成来收集数据并启动相应的通信。
- 为每个云服务提供商制定安全和治理要求。该大学通过实施护栏和控制措施（包括传输中和静态加密）来确保架构的所有组件安全，从而妥善处理学生数据。
- 在切实可行的情况下，尽可能采用云原生托管解决方案。云原生托管服务用于数据摄取、存储、数据库以及提取、转换和加载 (ETL) 功能，从而缩短了开发数据处理工作流程的 end-to-end时间。

身份联合验证和单点登录

确保核心系统之间一致的身份管理是成功且安全地采用任何技术的关键。教育机构越来越多地采用基于云的身份和单点登录解决方案，例如 AWS IAM Identity CenterMicrosoft Entra ID（前身为Azure Active Directory）JumpCloud、Okta OneLogin、、、、Ping Identity，以及 CyberArk 为了简化身份管理、减轻运营负担和集中实施最佳实践，例如多因素身份验证和最低权限访问。

其中许多机构仍在为其本地环境维护 Active Directory 和 Shibboleth 等身份管理和目录服务。其可以与基于云的解决方案集成，为您的学生和教职员工提供集中式身份管理和单点登录。云解决方案提供商应拥有强大的 easy-to-integrate身份管理平台，允许您通过云身份提供商将身份与现有应用程序、SaaS 解决方案和云服务联合起来。下图展示一个示例架构。



此架构遵循以下建议：

- 选择主要战略云提供商。此架构 AWS 用作主要的云提供商。通过与云身份提供者以及本地现有身份管理和目录服务集成，此架构支持自动预调配和管理对主要云提供商服务以及其他应用程序和 SaaS 解决方案的访问权限。随着越来越多的应用程序和服务添加到该机构的技术产品组合，这可确保以一致、易于管理的方式满足安全和治理要求。
- 区分 SaaS 应用程序和基础云服务。该架构集成了多种类型的基于云的、SaaS 和本地身份系统，以提供对 AWS 云 服务和其他应用程序的访问。许多基于云的身份提供者和单点登录解决方案也是 SaaS 应用程序，它们可以使用原生集成和 SAML 等标准协议来跨环境工作。
- 为每个云服务提供商制定安全和治理要求。此架构符合众多安全框架发布的身份和访问管理指引，包括美国国家标准与技术研究所（NIST）网络安全框架（CSF）、NIST 800-171 和 NIST 800-53。与 AWS Organizations、AWS Identity and Access Management （IAM） 以及其他 AWS 安全、身份与合规服务的集成可帮助根据组权限提供安全、精细的访问控制。

- 在切实可行的情况下，尽可能采用云原生托管服务。此架构使用基于云的托管服务进行身份管理和单点登录。这可减少在基础设施管理上花费的时间和精力，使维护这些关键系统变得更加容易。
- 在现有本地投资激励继续使用时，实施混合架构。此架构集成了用于托管 Active Directory、轻型目录访问控制（LDAP）和 Shibboleth 工作负载的基础设施现有的本地投资，并提供了最终将核心身份服务迁移到基于云的基础设施的路径。此外，如果您的本地工作负载需要基于证书的 AWS 资源访问权限，则可以使用AWS Identity and Access Management Roles Anywhere。

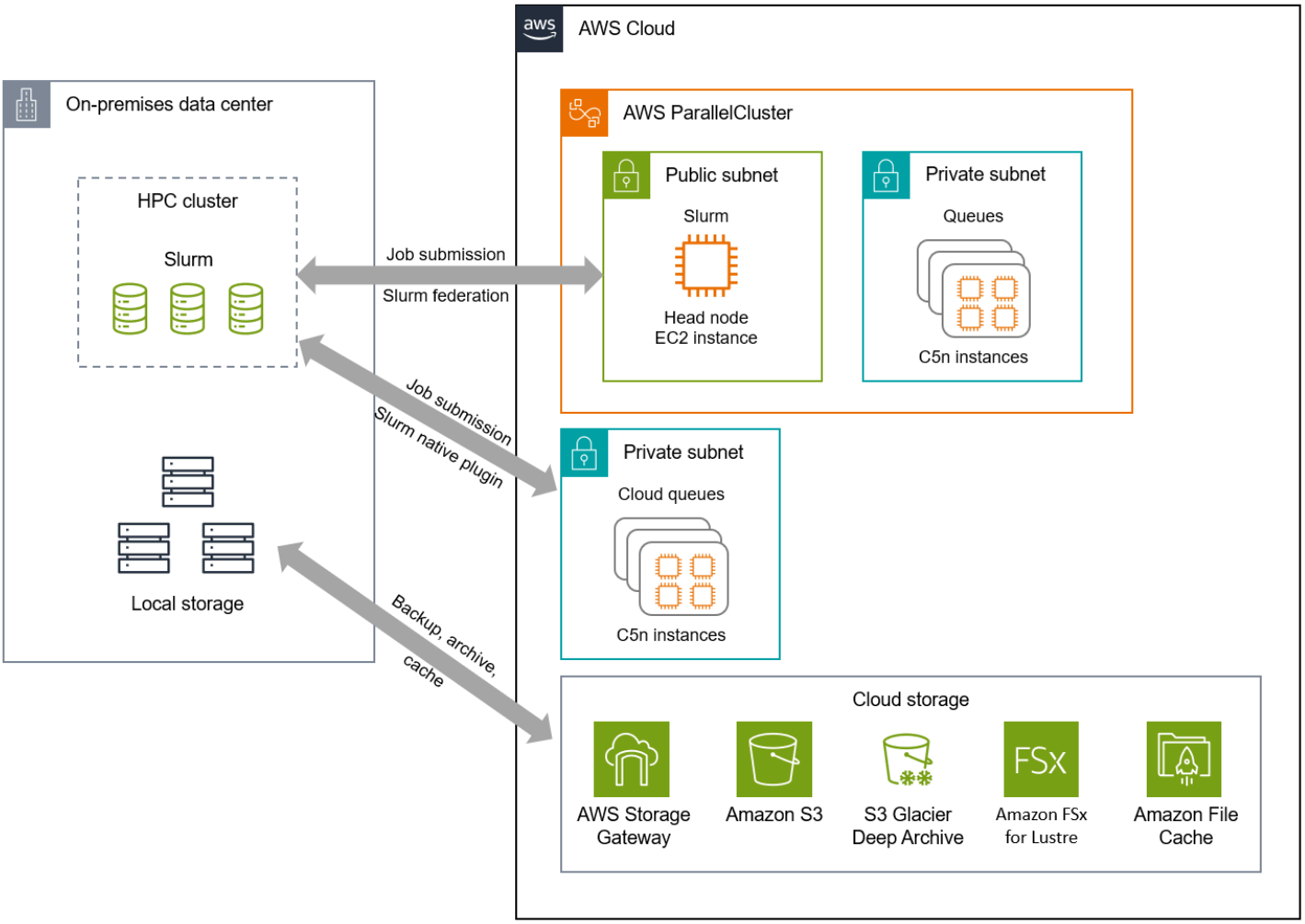
研究计算的云爆发

美国一所 R1 级（博士类大学：研究活动非常高）研究机构的研究计算团队多年来一直使用 Slurm 调度器运行本地高性能计算（HPC）集群。除了几周的计划维护外，集群的利用率为 80-95%，且大部分队列均为满载。

该机构的研究活动不断增长，这给容量和能力都带来了挑战。一些备受瞩目的研究人员总是在某些队列中执行长时间模拟，这增加了其他用户的等待时间。新聘用的教师需要进行大量的天气模拟，以构建一种新颖的人工智能和机器学习（AI/ML）模型进行天气预报，但其所需容量超出了可用容量。研究计算小组还收到了更多关于提供用于训练机器学习模型的最新图形处理单元 (GPUs) 的请求。即使有了购买新机架的资金 GPUs，该团队也需要等待几个月才能获得扩大数据中心机架空间的批准。

许多研究人员不愿删除旧数据，因此本地存储容量也面临挑战。我们需要一种更具可扩展性的长期存储方案，以腾出宝贵的本地高性能存储。

云通过混合计算和存储解决方案来应对这些挑战，当本地容量不足时，您可以将研究计算爆发到云。以下架构图展示使用 AWS ParallelCluster 和 AWS Storage Gateway 等工具的几种计算和存储爆发方法。



此架构遵循以下建议：

- 选择主要战略云提供商。此架构使用一个主要云提供商，以避免受到最小共同点方法的限制。这样，该机构就可以利用主要云提供商提供的创新以及原生计算和存储服务。研究计算团队可以专注于优化主要云提供商所提供环境中的工作负载，而不是如何在不同的云环境中工作。
- 为每个云服务提供商制定安全和治理要求。此架构中使用的每种服务和工具均可进行配置，以满足研究计算团队的安全和治理要求，包括私有连接、传输中和静态数据加密、活动日志记录等。
- 在切实可行的情况下，尽可能采用云原生托管服务。此架构能够使用托管存储和计算服务以及工具来简化集群管理。这样，研究计算团队无需自行管理集群或底层基础设施，这些工作往往复杂且耗时。
- 在现有本地投资激励继续使用时，实施混合架构。此架构使机构能够继续使用其本地资源，并利用云来增加容量和按需扩展计算能力。借助云，机构可以调整计算类型的适合大小，以实现最大性价比；并使用最新技术来促进创新，而无需对额外的本地硬件进行大量前期投资。

后续步骤

为云工作负载选择正确的部署模式需要审慎考虑。使用本文中概述的建议来指导您的决策，以避免常见的陷阱，例如不必要的复杂性、不断增加的员工需求、不一致的治理以及最低共同点方法。通过遵循这些最佳实践，您可以加快云采用的速度，从而更有效地实现和超越您的机构目标。

记得选择一家主要的战略云提供商，并建立卓越云中心 (CCoE)，以帮助提高组织成熟度，从而确保您的长期成功。区分 SaaS 应用程序与基础云服务，并明确各自的核心安全和治理要求。尽可能采用云原生托管服务，并在现有数据中心投资激励继续使用时实施混合架构。最后，仅为真正需要多云的工作负载保留多云。

AWS 完全有能力帮助您管理单云、混合云和多云环境。无论您的环境如何，您的机构都可以使用 AWS 管理和可观察性解决方案（例如AWS Systems ManagerAWS Config、和 Amazon CloudWatch）来简化和集中对基础设施和应用程序的管理和监控。借助 Amazon Athena、AWS Glue 和 AWS DataSync 等数据和分析服务，无论数据存储在何处，您都可以从所有数据中获得洞察。混合解决方案（例如AWS OutpostsAWS Wavelength、和 ）可AWS Snow Family让您将 AWS 基础设施和服务带到任何需要的地方。诸如 Amazon EKS Distro 之类的工具可帮助您在云上 AWS、本地或其他云上构建自我管理的 Kubernetes 集群。

在定义云策略时，请考虑以下后续步骤：

1. 查看AWS 云采用框架 (AWS CAF)，确定转型机会并确定其优先级，评估和改善您的云就绪性，并迭代发展您的转型路线图。
2. 确定一个用于云实施的系统，作为概念验证的起点。这将帮助您定义云基础或框架以验证任何假设，还将支持未来的云实施。
3. 与您的AWS 客户团队接触，讨论您的云实施目标。 AWS 客户团队可以帮助提供澄清、建议方法、确定依赖关系，还可以与您的团队合作规划从最初的概念到实施的旅程。

贡献者

本指南的贡献者包括：

- Kevin Arand，AWS 教育解决方案架构高级经理
- Kevin McCandless，AWS K-12 教育高级解决方案架构师
- Craig Jordan，AWS 教育首席解决方案架构师
- Jesse Roberts，AWS SLG 和 K-12 教育首席解决方案架构师
- Jianjun Xu，AWS 教育首席解决方案架构师
- Josh Badal，AWS 教育高级解决方案架构师
- Raj Chary，AWS 教育高级解决方案架构师

延伸阅读

如需了解其他信息，请参阅：

- AWS 架构中心
- Public Sector Cloud Transformation
- AWS Cloud Adoption Framework（AWS CAF）
- 适用于混合云和多云的 AWS 解决方案

文档历史记录

下表介绍了本指南的一些重要更改。如果您希望收到有关未来更新的通知，可以订阅 RSS 源。

变更	说明	日期
初次发布	—	2023 年 9 月 15 日

AWS 规范性指导词汇表

以下是 AWS 规范性指导提供的策略、指南和模式中的常用术语。若要推荐词条，请使用术语表末尾的提供反馈链接。

数字

7 R

将应用程序迁移到云中的 7 种常见迁移策略。这些策略以 Gartner 于 2011 年确定的 5 R 为基础，包括以下内容：

- 重构/重新架构：充分利用云原生功能来提高敏捷性、性能和可扩展性，以迁移应用程序并修改其架构。这通常涉及到移植操作系统和数据库。示例：将本地 Oracle 数据库迁移到 Amazon Aurora PostgreSQL 兼容版。
- 更换平台：将应用程序迁移到云中，并进行一定程度的优化，以利用云功能。示例：将本地 Oracle 数据库迁移到 AWS 云中的 Amazon Relational Database Service（Amazon RDS）for Oracle。
- 重新购买：转换到其他产品，通常是从传统许可转向 SaaS 模式。示例：将客户关系管理（CRM）系统迁移到 Salesforce.com。
- 重新托管（直接迁移）：将应用程序迁移到云中，无需进行任何更改即可利用云功能。示例：将本地 Oracle 数据库迁移到 AWS 云中 EC2 实例上的 Oracle。
- 重新放置（虚拟机监控器级直接迁移）：将基础设施迁移到云中，无需购买新硬件、重写应用程序或修改现有操作。您将服务器从本地平台迁移到同一平台的云服务中。示例：将Microsoft Hyper-V应用程序迁移到 AWS。
- 保留（重访）：将应用程序保留在源环境中。其中可能包括需要进行重大重构的应用程序，并且您希望将工作推迟到以后，以及您希望保留的遗留应用程序，因为迁移它们没有商业上的理由。
- 停用：停用或删除源环境中不再需要的应用程序。

A

ABAC

请参阅基于属性的访问控制。

抽象服务

请参阅托管服务。

ACID

请参阅原子性、一致性、隔离性、持久性。

主动-主动迁移

一种数据库迁移方法，在这种方法中，源数据库和目标数据库保持同步（通过使用双向复制工具或双写操作），两个数据库都在迁移期间处理来自连接应用程序的事务。这种方法支持小批量、可控的迁移，而不需要一次性割接。它比主动-被动迁移更灵活，但工作量更大。

主动-被动迁移

一种数据库迁移方法，在这种方法中，源数据库和目标数据库保持同步，但在将数据复制到目标数据库时，只有源数据库处理来自连接应用程序的事务。目标数据库在迁移期间不接受任何事务。

聚合函数

一种 SQL 函数，它对一组行进行操作并计算该组的单个返回值。聚合函数的示例包括 SUM 和 MAX。

AI

请参阅人工智能。

AIOps

请参阅人工智能运营。

匿名化

永久删除数据集中个人信息的过程。匿名化可以帮助保护个人隐私。匿名化数据不再被视为个人数据。

反模式

一种用于解决反复出现的问题的常用解决方案，而在这类问题中，此解决方案适得其反、无效或不如替代方案有效。

应用程序控制

一种安全方法，仅允许使用经批准的应用程序，以帮助保护系统免受恶意软件的侵害。

应用程序组合

有关组织使用的每个应用程序的详细信息的集合，包括构建和维护该应用程序的成本及其业务价值。这些信息是产品组合发现和分析过程的关键，有助于识别需要进行迁移、现代化和优化的应用程序并确定其优先级。

人工智能（AI）

计算机科学领域致力于使用计算技术执行通常与人类相关的认知功能，例如学习、解决问题和识别模式。有关更多信息，请参阅什么是人工智能？

人工智能操作 (AIOps)

使用机器学习技术解决运营问题、减少运营事故和人为干预以及提高服务质量的过程。有关如何在 AIOps AWS 迁移策略中使用的更多信息，请参阅操作集成指南。

非对称加密

一种加密算法，使用一对密钥，一个公钥用于加密，一个私钥用于解密。您可以共享公钥，因为它不用于解密，但对私钥的访问应受到严格限制。

原子性、一致性、隔离性、持久性（ACID）

一组软件属性，即使在出现错误、电源故障或其他问题的情况下，也能保证数据库的数据有效性和操作可靠性。

基于属性的访问权限控制（ABAC）

根据用户属性（如部门、工作角色和团队名称）创建精细访问权限的做法。有关更多信息，请参阅 AWS Identity and Access Management (I AM) 文档 AWS中的 AB AC。

权威数据来源

存储主要数据版本的位置，被认为是最可靠的信息源。您可以将数据从权威数据来源复制到其他位置，以便处理或修改数据，例如对数据进行匿名化、编辑或假名化。

可用区

中的一个不同位置 AWS 区域 ，不受其他可用区域故障的影响，并向同一区域中的其他可用区提供低成本、低延迟的网络连接。

AWS 云采用框架 (AWS CAF)

该框架包含指导方针和最佳实践 AWS ，可帮助组织制定高效且有效的计划，以成功迁移到云端。AWS CAF将指导分为六个重点领域，称为视角：业务、人员、治理、平台、安全和运营。业务、人员和治理角度侧重于业务技能和流程；平台、安全和运营角度侧重于技术技能和流程。例如，人

员角度针对的是负责人力资源（HR）、人员配置职能和人员管理的利益相关者。从这个角度来看，AWS CAF 为人员发展、培训和沟通提供了指导，以帮助组织为成功采用云做好准备。有关更多信息，请参阅 AWS CAF 网站和 AWS CAF 白皮书。

AWS 工作负载资格框架 (AWS WQF)

一种评估数据库迁移工作负载、推荐迁移策略和提供工作估算的工具。 AWS WQF 包含在 AWS Schema Conversion Tool (AWS SCT) 中。它用来分析数据库架构和代码对象、应用程序代码、依赖关系和性能特征，并提供评测报告。

B

恶意机器人

一种旨在扰乱或伤害个人或组织的机器人。

BCP

请参阅业务连续性计划。

行为图

一段时间内资源行为和交互的统一交互式视图。您可以使用 Amazon Detective 的行为图来检查失败的登录尝试、可疑的 API 调用和类似的操作。有关更多信息，请参阅 Detective 文档中的行为图中的数据。

大端序系统

一个先存储最高有效字节的系统。另请参阅字节顺序。

二进制分类

一种预测二进制结果（两个可能的类别之一）的过程。例如，您的 ML 模型可能需要预测诸如“该电子邮件是否为垃圾邮件？”或“这个产品是书还是汽车？”之类的问题

bloom 筛选条件

一种概率性、内存高效的数据结构，用于测试元素是否为集合的成员。

蓝/绿部署

一种部署策略，您可以创建两个独立但完全相同的环境。在一个环境中运行当前应用程序版本（蓝色），在另一个环境中运行新应用程序版本（绿色）。此策略可帮助您在影响最小的情况下快速回滚。

自动程序

一种通过互联网运行自动任务并模拟人类活动或交互的软件应用程序。有些机器人是有用或有益的，例如在互联网上索引信息的 Web 爬网程序。还有一些被称为恶意机器人的机器人，其目的是扰乱或伤害个人或组织。

僵尸网络

被恶意软件感染并受单方（称为僵尸网络控制者或僵尸网络操作者）控制的僵尸网络。僵尸网络是最著名的扩展机器人及其影响力的机制。

分支

代码存储库的一个包含区域。在存储库中创建的第一个分支是主分支。您可以从现有分支创建新分支，然后在新分支中开发功能或修复错误。为构建功能而创建的分支通常称为功能分支。当功能可以发布时，将功能分支合并回主分支。有关更多信息，请参阅关于分支（GitHub 文档）。

紧急（break-glass）访问

在特殊情况下，通过批准的流程，用户 AWS 账户 可以快速访问他们通常没有访问权限的内容。有关更多信息，请参阅 AWS Well-Architected Guidance 中的 Implement break-glass procedures 指示器。

棕地策略

您环境中的现有基础设施。在为系统架构采用棕地策略时，您需要围绕当前系统和基础设施的限制来设计架构。如果您正在扩展现有基础设施，则可以将棕地策略和全新策略混合。

缓冲区缓存

存储最常访问的数据的内存区域。

业务能力

企业如何创造价值（例如，销售、客户服务或营销）。微服务架构和开发决策可以由业务能力驱动。有关更多信息，请参阅在 AWS上运行容器化微服务白皮书中的围绕业务能力进行组织部分。

业务连续性计划（BCP）

一项计划，旨在应对大规模迁移等破坏性事件对运营的潜在影响，并使企业能够快速恢复运营。

C

CAF

请参阅 AWS 云采用框架。

金丝雀部署

缓慢而渐进地向最终用户发布版本。当您确信无误后，即可部署新版本，并完全替换当前版本。

CCoE

请参阅云卓越中心。

CDC

请参阅更改数据捕获。

更改数据捕获（CDC）

跟踪数据来源（如数据库表）的更改并记录有关更改的元数据的过程。您可以将 CDC 用于各种目的，例如审计或复制目标系统中的更改以保持同步。

混沌工程

故意引入故障或破坏性事件来测试系统的韧性。您可以使用 AWS Fault Injection Service (AWS FIS) 来执行实验，对您的 AWS 工作负载施加压力并评估其响应。

CI/CD

请参阅持续集成和持续交付。

分类

一种有助于生成预测的分类流程。分类问题的 ML 模型预测离散值。离散值始终彼此不同。例如，一个模型可能需要评估图像中是否有汽车。

客户端加密

在目标 AWS 服务 收到数据之前，对数据进行本地加密。

云卓越中心 (CCoE)

一个多学科团队，负责推动整个组织的云采用工作，包括开发云最佳实践、调动资源、制定迁移时间表、领导组织完成大规模转型。有关更多信息，请参阅 AWS 云 企业战略博客上的 CCoE 帖子。

云计算

通常用于远程数据存储和 IoT 设备管理的云技术。云计算通常连接到边缘计算技术。

云运营模型

在 IT 组织中，一种用于构建、完善和优化一个或多个云环境的运营模型。有关更多信息，请参阅构建您的云运营模型。

云采用阶段

组织迁移到 AWS 云中时通常会经历四个阶段：

- 项目 - 出于概念验证和学习目的，开展一些与云相关的项目
- 基础 — 进行基础投资以扩大云采用率（例如，创建着陆区、定义 CCo E、建立运营模型）
- 迁移 - 迁移单个应用程序
- 重塑 - 优化产品和服务，在云中创新

Stephen Orban在 AWS 云 企业战略博客的博客文章《云优先之旅和采用阶段》中定义了这些阶段。有关它们与 AWS 迁移策略的关系的信息，请参阅迁移准备指南。

CMDB

请参阅配置管理数据库。

代码存储库

通过版本控制过程存储和更新源代码和其他资产（如文档、示例和脚本）的位置。常见的云存储库包括 GitHub 或 Bitbucket Cloud。每个版本的代码都称为一个分支。在微服务结构中，每个存储库都专门用于一个功能。单个 CI/CD 管线可以使用多个存储库。

冷缓存

一种空的、填充不足或包含过时或不相关数据的缓冲区缓存。这会影响性能，因为数据库实例必须从主内存或磁盘读取，这比从缓冲区缓存读取要慢。

冷数据

很少访问的数据，且通常是历史数据。查询此类数据时，通常可以接受慢速查询。将这些数据转移到性能较低且成本更低的存储层或类别可以降低成本。

计算机视觉（CV）

一种 AI 领域，它使用机器学习来分析和提取数字图像和视频等视觉格式中的信息。例如，Amazon SageMaker AI 为 CV 提供了图像处理算法。

配置偏移

对于工作负载而言，一种偏离预期状态的配置更改。这可能会导致工作负载变得不合规，且通常是渐进的，不是故意的。

配置管理数据库（CMDB）

一种存储库，用于存储和管理有关数据库及其 IT 环境的信息，包括硬件和软件组件及其配置。您通常在迁移的产品组合发现和分析阶段使用来自 CMDB 的数据。

合规性包

一系列 AWS Config 规则和补救措施，您可以汇编这些规则和补救措施，以自定义您的合规性和安全性检查。您可以使用 YAML 模板将一致性包作为单个实体部署在 AWS 账户 和区域或整个组织中。有关更多信息，请参阅 AWS Config 文档中的一致性包。

持续集成和持续交付（CI/CD）

自动执行软件发布过程的源代码、构建、测试、暂存和生产阶段的过程。 CI/CD 通常被描述为管道。 CI/CD 可以帮助您实现流程自动化、提高生产力、提高代码质量和更快地交付。有关更多信息，请参阅持续交付的优势。CD 也可以表示持续部署。有关更多信息，请参阅持续交付与持续部署。

CV

请参阅计算机视觉。

D

静态数据

网络中静止的数据，例如存储中的数据。

数据分类

根据网络中数据的关键性和敏感性对其进行识别和分类的过程。它是任何网络安全风险管理策略的关键组成部分，因为它可以帮助您确定对数据的适当保护和保留控制。数据分类是 Well-Architecte AWS d Framework 中安全支柱的一个组成部分。有关详细信息，请参阅数据分类。

数据漂移

生产数据与用来训练机器学习模型的数据之间的有意义差异，或者输入数据随时间推移的有意义变化。数据漂移可能降低机器学习模型预测的整体质量、准确性和公平性。

传输中数据

在网络中主动移动的数据，例如在网络资源之间移动的数据。

数据网格

一种架构框架，可提供分布式、去中心化的数据所有权以及集中式管理和治理。

数据最少化

仅收集并处理绝对必要数据的原则。在中进行数据最小化 AWS 云 可以降低隐私风险、成本和分析碳足迹。

数据边界

AWS 环境中的一组预防性防护措施，可帮助确保只有可信身份才能访问来自预期网络的可信资源。有关更多信息，请参阅在上构建数据边界。 AWS

数据预处理

将原始数据转换为 ML 模型易于解析的格式。预处理数据可能意味着删除某些列或行，并处理缺失、不一致或重复的值。

数据溯源

在数据的整个生命周期跟踪其来源和历史的过程，例如数据如何生成、传输和存储。

数据主体

正在收集和处理其数据的个人。

数据仓库

一种支持商业智能（例如分析）的数据管理系统。数据仓库通常包含大量历史数据，通常用于查询和分析。

数据库定义语言（DDL）

在数据库中创建或修改表和对象结构的语句或命令。

数据库操作语言（DML）

在数据库中修改（插入、更新和删除）信息的语句或命令。

DDL

请参阅数据库定义语言。

深度融合

组合多个深度学习模型进行预测。您可以使用深度融合来获得更准确的预测或估算预测中的不确定性。

深度学习

一个 ML 子字段使用多层人工神经网络来识别输入数据和感兴趣的目标变量之间的映射。

defense-in-depth

一种信息安全方法，经过深思熟虑，在整个计算机网络中分层实施一系列安全机制和控制措施，以保护网络及其中数据的机密性、完整性和可用性。当你采用这种策略时 AWS，你会在 AWS

Organizations 结构的不同层面添加多个控件来帮助保护资源。例如，一种 defense-in-depth方法可以结合多因素身份验证、网络分段和加密。

委派管理员

在中 AWS Organizations，兼容的服务可以注册 AWS 成员帐户来管理组织的帐户并管理该服务的权限。此账户被称为该服务的委托管理员。有关更多信息和兼容服务列表，请参阅 AWS Organizations 文档中使用 AWS Organizations的服务。

部署

使应用程序、新功能或代码修复在目标环境中可用的过程。部署涉及在代码库中实现更改，然后在应用程序的环境中构建和运行该代码库。

开发环境

请参阅环境。

侦测性控制

一种安全控制，在事件发生后进行检测、记录日志和发出提醒。这些控制是第二道防线，提醒您注意绕过现有预防性控制的安全事件。有关更多信息，请参阅在 AWS上实施安全控制中的侦测性控制。

开发价值流映射（DVSM）

用于识别对软件开发生命周期中的速度和质量产生不利影响的限制因素并确定其优先级的流程。DVSM 扩展了最初为精益生产实践设计的价值流映射流程。其重点关注在软件开发过程中创造和转移价值所需的步骤和团队。

数字孪生

真实世界系统的虚拟再现，如建筑物、工厂、工业设备或生产线。数字孪生支持预测性维护、远程监控和生产优化。

维度表

星型架构中的一种较小的表，其中包含事实表中定量数据的数据属性。维度表属性通常是文本字段或行为类似于文本的离散数字。这些属性通常用于查询约束、筛选和结果集标注。

灾难

阻止工作负载或系统在其主要部署位置实现其业务目标的事件。这些事件可能是自然灾害、技术故障或人为操作的结果，例如无意的配置错误或恶意软件攻击。

灾难恢复（DR）

您用来最大程度地减少由灾难造成的停机时间和数据丢失的策略和流程。有关更多信息，请参阅 Well-Architected Fr ame AWS work 中的 “工作负载灾难恢复：云端 AWS 恢复”。

DML

请参阅数据库操作语言。

领域驱动设计

一种开发复杂软件系统的方法，通过将其组件连接到每个组件所服务的不断发展的领域或核心业务目标。Eric Evans 在其著作领域驱动设计：软件核心复杂性应对之道（Boston: Addison-Wesley Professional, 2003）中介绍了这一概念。有关如何将领域驱动设计与 strangler fig 模式结合使用的信息，请参阅使用容器和 Amazon API Gateway 逐步将原有的 Microsoft ASP.NET（ASMX）Web 服务现代化。

DR

请参阅灾难恢复。

偏差检测

跟踪与基准配置的偏差。例如，您可以使用 AWS CloudFormation 来检测系统资源中的偏差，也可以使用 AWS Control Tower 来检测着陆区中可能影响监管要求合规性的变化。

DVSM

请参阅开发价值流映射。

E

EDA

请参阅探索性数据分析。

EDI

请参阅电子数据交换。

边缘计算

该技术可提高位于 IoT 网络边缘的智能设备的计算能力。与云计算比较时，边缘计算可以减少通信延迟并缩短响应时间。

电子数据交换（EDI）

组织之间业务文件的自动交换。有关更多信息，请参阅什么是电子数据交换。

加密

一种将人类可读的纯文本数据转换为加密文字的计算流程。

加密密钥

由加密算法生成的随机位的加密字符串。密钥的长度可能有所不同，而且每个密钥都设计为不可预测且唯一。

字节顺序

字节在计算机内存中的存储顺序。大端序系统先存储最高有效字节。小端序系统先存储最低有效字节。

端点

请参阅服务端点。

端点服务

一种可以在虚拟私有云（VPC）中托管，与其他用户共享的服务。您可以使用其他 AWS 账户 或 AWS Identity and Access Management (IAM) 委托人创建终端节点服务， AWS PrivateLink 并向其授予权限。这些账户或主体可通过创建接口 VPC 端点来私密地连接到您的端点服务。有关更多信息，请参阅 Amazon Virtual Private Cloud（Amazon VPC）文档中的创建端点服务。

企业资源规划（ERP）

一种自动化和管理企业关键业务流程（例如会计、MES 和项目管理）的系统。

信封加密

用另一个加密密钥对加密密钥进行加密的过程。有关更多信息，请参阅 AWS Key Management Service (AWS KMS) 文档中的信封加密。

环境

正在运行的应用程序的实例。以下是云计算中常见的环境类型：

- 开发环境 — 正在运行的应用程序的实例，只有负责维护应用程序的核心团队才能使用。开发环境用于测试更改，然后再将其提升到上层环境。这类环境有时称为测试环境。
- 下层环境 — 应用程序的所有开发环境，比如用于初始构建和测试的环境。

- 生产环境 — 最终用户可以访问的正在运行的应用程序的实例。在 CI/CD 管道中，生产环境是最后一个部署环境。
- 上层环境 — 除核心开发团队以外的用户可以访问的所有环境。这可能包括生产环境、预生产环境和用户验收测试环境。

epic

在敏捷方法学中，有助于组织工作和确定优先级的功能类别。epics 提供了对需求和实施任务的总体描述。例如， AWS CAF 安全史诗包括身份和访问管理、侦探控制、基础设施安全、数据保护和事件响应。有关 AWS 迁移策略中 epics 的更多信息，请参阅计划实施指南。

ERP

请参阅企业资源规划。

探索性数据分析（EDA）

分析数据集以了解其主要特征的过程。您收集或汇总数据，并进行初步调查，以发现模式、检测异常并检查假定情况。EDA 通过计算汇总统计数据和创建数据可视化得以执行。

F

事实表

星型架构中的中心表。它存储有关业务运营的定量数据。通常，事实表包含两种类型的列：包含度量的列和包含维度表外键的列。

快速失效机制

一种使用频繁且增量式的测试来缩短开发生命周期的理念。这是敏捷方法的关键部分。

故障隔离边界

在中 AWS 云，诸如可用区 AWS 区域、控制平面或数据平面之类的边界，它限制了故障的影响并有助于提高工作负载的弹性。有关更多信息，请参阅 AWS 故障隔离边界。

功能分支

请参阅分支。

特征

您用来进行预测的输入数据。例如，在制造环境中，特征可能是定期从生产线捕获的图像。

特征重要性

特征对于模型预测的重要性。这通常表示为数值分数，可以通过各种技术进行计算，例如 Shapley 加法解释（SHAP）和积分梯度。有关更多信息，请参阅使用机器学习模型的可解释性 AWS。

功能转换

为 ML 流程优化数据，包括使用其他来源丰富数据、扩展值或从单个数据字段中提取多组信息。这使得 ML 模型能从数据中获益。例如，如果您将“2021-05-27 00:15:37”日期分解为“2021”、“五月”、“星期四”和“15”，则可以帮助学习与不同数据成分相关的算法学习精细模式。

少样本提示

在要求 LLM 执行类似任务之前，先向其提供少量示例，以演示任务和预期输出。此技术是上下文内学习的一种应用，其中模型可以从提示中嵌入的示例（样本）中学习。对于需要特定格式、推理或领域知识的任务，少样本提示可能非常有效。另请参阅零样本提示。

FGAC

请参阅精细访问控制。

精细访问控制（FGAC）

使用多个条件允许或拒绝访问请求。

快闪迁移

一种数据库迁移方法，通过更改数据捕获使用连续数据复制，在极短的时间内迁移数据，而非使用分阶段方法。目标是将停机时间降至最低。

FM

请参阅基础模型。

基础模型（FM）

一个大型深度学习神经网络，一直在广义和未标记数据的大量数据集上进行训练。 FMs 能够执行各种各样的一般任务，例如理解语言、生成文本和图像以及用自然语言进行对话。有关更多信息，请参阅什么是基础模型。

G

生成式人工智能

AI 模型的一个子集，这些模型已经过大量数据训练，可以使用简单的文本提示来创建新的内容和构件，例如图像、视频、文本和音频。有关更多信息，请参阅什么是生成式人工智能。

地理阻止

请参阅地理限制。

地理限制（地理阻止）

在 Amazon 中 CloudFront，一种阻止特定国家/地区的用户访问内容分发的选项。您可以使用允许列表或阻止列表来指定已批准和已禁止的国家/地区。有关更多信息，请参阅 CloudFront 文档中的限制内容的地理分布。

GitFlow 工作流程

一种方法，在这种方法中，下层和上层环境在源代码存储库中使用不同的分支。Gitflow 工作流程被认为是传统的工作流程，而基于中继的工作流程则是现代的、首选的方法。

黄金映像

系统或软件的快照，用作部署该系统或软件的新实例的模板。例如，在制造业中，黄金映像可用于在多个设备上预调配软件，并有助于提高设备制造操作的速度、可扩展性和生产效率。

全新策略

在新环境中缺少现有基础设施。在对系统架构采用全新策略时，您可以选择所有新技术，而不受对现有基础设施（也称为棕地）兼容性的限制。如果您正在扩展现有基础设施，则可以将棕地策略和全新策略混合。

防护机制

帮助管理各组织单位的资源、策略和合规性的高级规则 (OUs)。预防性防护机制会执行策略以确保符合合规性标准。它们是使用服务控制策略和 IAM 权限边界实现的。侦测性护栏会检测策略违规和合规性问题，并生成提醒以进行修复。它们通过使用 AWS Config、、Amazon、 AWS Security Hub CSPM GuardDuty AWS Trusted Advisor、Amazon Inspector 和自定义 AWS Lambda 支票来实现。

H

HA

请参阅高可用性。

异构数据库迁移

将源数据库迁移到使用不同数据库引擎的目标数据库（例如，从 Oracle 迁移到 Amazon Aurora）。异构迁移通常是重新架构工作的一部分，而转换架构可能是一项复杂的任务。AWS 提供了 AWS SCT 来帮助实现架构转换。

高可用性（HA）

在遇到挑战或灾难时，工作负载无需干预即可连续运行的能力。HA 系统旨在自动进行故障转移、持续提供良好性能，并以最小的性能影响处理不同负载和故障。

历史数据库现代化

一种用于实现运营技术（OT）系统现代化和升级以更好满足制造业需求的方法。历史数据库是一种用于收集和存储工厂中各种来源数据的数据库。

保留数据

从用于训练机器学习模型的数据集中保留的一部分标注的历史数据。通过将模型预测与保留数据进行比较，您可以使用保留数据来评估模型性能。

同构数据库迁移

将源数据库迁移到共享同一数据库引擎的目标数据库（例如，从 Microsoft SQL Server 迁移到 Amazon RDS for SQL Server）。同构迁移通常是更换主机或更换平台工作的一部分。您可以使用本机数据库实用程序来迁移架构。

热数据

经常访问的数据，例如实时数据或近期的转化数据。这些数据通常需要高性能存储层或存储类别才能提供快速的查询响应。

修补程序

针对生产环境中关键问题的紧急修复。由于其紧迫性，修补程序通常是在典型的 DevOps 发布工作流程之外进行的。

hypercare 周期

割接之后，迁移团队立即管理和监控云中迁移的应用程序以解决任何问题的时间段。通常，这个周期持续 1-4 天。在 hypercare 周期结束时，迁移团队通常会将应用程序的责任移交给云运营团队。

我

IaC

请参阅基础设施即代码。

基于身份的策略

附加到一个或多个 IAM 委托人的策略，用于定义他们在 AWS 云 环境中的权限。

空闲应用程序

90 天内平均 CPU 和内存使用率在 5% 到 20% 之间的应用程序。在迁移项目中，通常会停用这些应用程序或将其保留在本地。

IIoT

请参阅工业物联网。

不可变基础设施

一种模型，可为生产工作负载部署新的基础设施，而不是更新、修补或修改现有基础设施。不可变基础设施本质上比可变基础设施更一致、更可靠、更可预测。有关更多信息，请参阅 AWS Well-Architected Framework 中的使用不可变基础设施进行部署最佳实践。

入站（入口）VPC

在 AWS 多账户架构中，一种接受、检查和路由来自应用程序外部的网络连接的 VPC。AWS 安全参考架构建议设置您的网络帐户，包括入站、出站和检查， VPCs 以保护您的应用程序与更广泛的互联网之间的双向接口。

增量迁移

一种割接策略，在这种策略中，您可以将应用程序分成小部分进行迁移，而不是一次性完整割接。例如，您最初可能只将几个微服务或用户迁移到新系统。在确认一切正常后，您可以逐步迁移其他微服务或用户，直到停用遗留系统。这种策略降低了大规模迁移带来的风险。

工业 4.0

该术语由 Klaus Schwab 在 2016 年提出，指的是通过连接、实时数据、自动化、分析和 AI/ML 的进步来实现制造流程的现代化。

基础设施

应用程序环境中包含的所有资源和资产。

基础设施即代码（IaC）

通过一组配置文件预调配和管理应用程序基础设施的过程。IaC 旨在帮助您集中管理基础设施、实现资源标准化和快速扩展，使新环境具有可重复性、可靠性和一致性。

工业物联网 (IIoT)

在工业领域使用联网的传感器和设备，例如制造业、能源、汽车、医疗保健、生命科学和农业。有关更多信息，请参阅制定工业物联网 (IIoT) 数字化转型战略。

检查 VPC

在 AWS 多账户架构中，一种集中式 VPC，用于管理对 VPCs（相同或不同 AWS 区域）、互联网和本地网络之间的网络流量的检查。AWS 安全参考架构建议设置您的网络帐户，包括入站、出站和检查， VPCs 以保护您的应用程序与更广泛的互联网之间的双向接口。

物联网（IoT）

由带有嵌入式传感器或处理器的连接物理对象组成的网络，这些传感器或处理器通过互联网或本地通信网络与其他设备和系统进行通信。有关更多信息，请参阅什么是 IoT？

可解释性

它是机器学习模型的一种特征，描述了人类可以理解模型的预测如何取决于其输入的程度。有关更多信息，请参阅使用机器学习模型的可解释性 AWS。

物联网

请参阅物联网。

IT 信息库（ITIL）

提供 IT 服务并使这些服务符合业务要求的一套最佳实践。ITIL 是 ITSM 的基础。

IT 服务管理（ITSM）

为组织设计、实施、管理和支持 IT 服务的相关活动。有关将云运营与 ITSM 工具集成的信息，请参阅运营集成指南。

ITIL

请参阅 IT 信息库。

ITSM

请参阅 IT 服务管理。

L

基于标签的访问控制（LBAC）

强制访问控制（MAC）的一种实施方式，其中明确为用户和数据本身分配了安全标签值。用户安全标签和数据安全标签之间的交集决定了用户可以看到哪些行和列。

登录区

landing zone 是一个架构精良的多账户 AWS 环境，具有可扩展性和安全性。这是一个起点，您的组织可以从这里放心地在安全和基础设施环境中快速启动和部署工作负载和应用程序。有关登录区的更多信息，请参阅设置安全且可扩展的多账户 AWS 环境。

大语言模型（LLM）

一种基于大量数据进行预训练的深度学习 AI 模型。LLM 可以执行多项任务，例如回答问题、总结文档、将文本翻译成其他语言以及完成句子。有关更多信息，请参阅什么是 LLMs。

大规模迁移

迁移 300 台或更多服务器。

LBAC

请参阅基于标签的访问控制。

最低权限

授予执行任务所需的最低权限的最佳安全实践。有关更多信息，请参阅 IAM 文档中的应用最低权限许可。

直接迁移

请参阅 7 R。

小端序系统

一个先存储最低有效字节的系统。另请参阅字节顺序。

LLM

请参阅大型语言模型。

下层环境

请参阅环境。

M

机器学习（ML）

一种使用算法和技术进行模式识别和学习的人工智能。ML 对记录的数据（例如物联网（IoT）数据）进行分析和学习，以生成基于模式的统计模型。有关更多信息，请参阅机器学习。

主分支

请参阅分支。

恶意软件

旨在危害计算机安全或隐私的软件。恶意软件可能会破坏计算机系统、泄露敏感信息或获得未经授权的访问权限。恶意软件的示例包括病毒、蠕虫、勒索软件、木马、间谍软件和键盘记录器。

托管式服务

AWS 服务 它 AWS 运行基础设施层、操作系统和平台，您可以访问端点来存储和检索数据。Amazon Simple Storage Service（Amazon S3）和 Amazon DynamoDB 就是托管服务的示例。这些服务也称为抽象服务。

制造执行系统（MES）

一种软件系统，用于跟踪、监控、记录和控制在车间将原材料转化为成品的生产过程。

MAP

请参阅迁移加速计划。

机制

一个完整的过程，您可以在其中创建工具，推动工具的采用，然后检查结果以进行调整。机制是一种在运作过程中自我强化和改善的循环。有关更多信息，请参阅在 Well-Architect AWS ed 框架中构建机制。

成员账户

AWS 账户 除属于组织中的管理账户之外的所有账户 AWS Organizations。一个账户一次只能是一个组织的成员。

MES

请参阅制造执行系统。

消息队列遥测传输（MQTT）

一种基于发布/订阅模式的轻量级 machine-to-machine (M2M) 通信协议，适用于资源受限的物联网设备。

微服务

一种小型的独立服务，通过明确的定义进行通信 APIs ，通常由小型的独立团队拥有。例如，保险系统可能包括映射到业务能力（如销售或营销）或子域（如购买、理赔或分析）的微服务。微服务

的好处包括敏捷、灵活扩展、易于部署、可重复使用的代码和恢复能力。有关更多信息，请参阅使用 AWS 无服务器服务集成微服务。

微服务架构

一种使用独立组件构建应用程序的方法，这些组件将每个应用程序进程作为微服务运行。这些微服务使用轻量级通过定义明确的接口进行通信。 APIs该架构中的每个微服务都可以更新、部署和扩展，以满足对应用程序特定功能的需求。有关更多信息，请参阅在上实现微服务。 AWS

迁移加速计划（MAP）

AWS 该计划提供咨询支持、培训和服务，以帮助组织为迁移到云奠定坚实的运营基础，并帮助抵消迁移的初始成本。MAP 提供了一种以系统的方式执行遗留迁移的迁移方法，以及一套用于自动执行和加速常见迁移场景的工具。

大规模迁移

将大部分应用程序组合分波迁移到云中的过程，在每一波中以更快的速度迁移更多应用程序。本阶段使用从早期阶段获得的最佳实践和经验教训，实施由团队、工具和流程组成的迁移工厂，通过自动化和敏捷交付简化工作负载的迁移。这是 AWS 迁移策略的第三阶段。

迁移工厂

跨职能团队，通过自动化、敏捷的方法简化工作负载迁移。迁移工厂团队通常包括运营、业务分析师和所有者、迁移工程师、开发人员和冲刺 DevOps 领域的专业人员。20% 到 50% 的企业应用程序组合由可通过工厂方法优化的重复模式组成。有关更多信息，请参阅本内容集中有关迁移工厂的讨论和云迁移工厂指南。

迁移元数据

有关完成迁移所需的应用程序和服务器的信息。每种迁移模式都需要一套不同的迁移元数据。迁移元数据的示例包括目标子网、安全组和 AWS 账户。

迁移模式

一种可重复的迁移任务，详细列出了迁移策略、迁移目标以及所使用的迁移应用程序或服务。示例：使用 AWS 应用程序迁移服务重新托管向 Amazon EC2 的迁移。

迁移组合评测（MPA）

一种在线工具，提供了用于验证迁移到 AWS 云的业务案例的信息。MPA 提供了详细的组合评测（服务器规模调整、定价、TCO 比较、迁移成本分析）以及迁移计划（应用程序数据分析和数据收集、应用程序分组、迁移优先级排序和波次规划）。所有 AWS 顾问和 APN 合作伙伴顾问均可免费使用 MPA 工具（需要登录）。

迁移准备情况评测（MRA）

使用 AWS CAF 深入了解组织的云就绪状态、确定优势和劣势以及制定行动计划以缩小已发现差距的过程。有关更多信息，请参阅迁移准备指南。MRA 是 AWS 迁移策略的第一阶段。

迁移策略

将工作负载迁移到 AWS 云的方法。有关更多信息，请参见术语表中的 7 R 词条，以及动员您的组织以加快大规模迁移。

ML

请参阅机器学习。

现代化

将过时的（原有的或单体）应用程序及其基础设施转变为云中敏捷、弹性和高度可用的系统，以降低成本、提高效率和利用创新。有关更多信息，请参阅在 AWS 云中实现应用程序现代化的策略。

现代化准备情况评估

一种评估方式，有助于确定组织应用程序的现代化准备情况；确定收益、风险和依赖关系；确定组织能够在多大程度上支持这些应用程序的未来状态。评估结果是目标架构的蓝图、详细说明现代化进程发展阶段和里程碑的路线图以及解决已发现差距的行动计划。有关更多信息，请参阅在 AWS 云中评估应用程序的现代化准备情况。

单体应用程序（单体式）

作为具有紧密耦合进程的单个服务运行的应用程序。单体应用程序有几个缺点。如果某个应用程序功能的需求激增，则必须扩展整个架构。随着代码库的增长，添加或改进单体应用程序的功能也会变得更加复杂。若要解决这些问题，可以使用微服务架构。有关更多信息，请参阅将单体分解为微服务。

MPA

请参阅迁移组合评测。

MQTT

请参阅消息队列遥测传输。

多分类器

一种帮助为多个类别生成预测（预测两个以上结果之一）的过程。例如，ML 模型可能会询问“这个产品是书、汽车还是手机？”或“此客户最感兴趣什么类别的产品？”

可变基础设施

一种用于更新和修改生产工作负载的现有基础设施的模型。为了提高一致性、可靠性和可预测性，Well-Architect AWS ed Framework 建议使用不可变基础设施作为最佳实践。

O

OAC

请参阅来源访问控制。

OAI

请参阅来源访问身份。

OCM

请参阅组织变革管理。

离线迁移

一种迁移方法，在这种方法中，源工作负载会在迁移过程中停止运行。这种方法会延长停机时间，通常用于小型非关键工作负载。

OI

请参阅运营集成。

OLA

请参阅运营级别协议。

在线迁移

一种迁移方法，在这种方法中，源工作负载无需离线即可复制到目标系统。在迁移过程中，连接工作负载的应用程序可以继续运行。这种方法的停机时间为零或最短，通常用于关键生产工作负载。

OPC-UA

请参阅开放流程通信 – 统一架构。

开放流程通信 – 统一架构（OPC-UA）

一种用于工业自动化的 machine-to-machine（M2M）通信协议。OPC-UA 提供了一个包含数据加密、身份验证和授权方案的互操作性标准。

运营级别协议（OLA）

一项协议，阐明了 IT 职能部门承诺相互交付的内容，以支持服务水平协议（SLA）。

运营准备情况审查（ORR）

一份问题核对清单和关联的最佳实践，可帮助您了解、评估、预防或缩小事件和可能的故障的范围。有关更多信息，请参阅 AWS Well-Architected Framework 中的运营准备情况审查（ORR）。

运营技术（OT）

与物理环境配合使用以控制工业运营、设备和基础设施的硬件和软件系统。在制造业中，OT 和信息技术（IT）系统的集成是工业 4.0 转型的关键重点。

运营整合（OI）

在云中实现运营现代化的过程，包括就绪计划、自动化和集成。有关更多信息，请参阅运营整合指南。

组织跟踪

由 AWS CloudTrail 此创建的跟踪记录组织 AWS 账户 中所有人的所有事件 AWS Organizations。该跟踪是在每个 AWS 账户 中创建的，属于组织的一部分，并跟踪每个账户的活动。有关更多信息，请参阅 CloudTrail文档中的为组织创建跟踪。

组织变革管理（OCM）

一个从人员、文化和领导力角度管理重大、颠覆性业务转型的框架。OCM 通过加快变革采用、解决过渡问题以及推动文化和组织变革，帮助组织为新系统和战略做好准备和过渡。在 AWS 迁移策略中，该框架被称为人员加速，因为云采用项目需要变更的速度。有关更多信息，请参阅 OCM 指南。

来源访问控制（OAC）

在中 CloudFront，一个增强的选项，用于限制访问以保护您的亚马逊简单存储服务 (Amazon S3) 内容。OAC 全部支持所有 S3 存储桶 AWS 区域、使用 AWS KMS (SSE-KMS) 进行服务器端加密，以及对 S3 存储桶的动态PUT和DELETE请求。

来源访问身份（OAI）

在中 CloudFront，一个用于限制访问权限以保护您的 Amazon S3 内容的选项。当您使用 OAI 时，CloudFront 会创建一个 Amazon S3 可以对其进行身份验证的委托人。经过身份验证的委托人只能通过特定 CloudFront 分配访问 S3 存储桶中的内容。另请参阅 OAC，其中提供了更精细和增强的访问控制。

ORR

请参阅运营准备情况审查。

OT

请参阅运营技术。

出站（出口）VPC

在 AWS 多账户架构中，一种处理从应用程序内部启动的网络连接的 VPC。AWS 安全参考架构建议设置您的网络帐户，包括入站、出站和检查， VPCs 以保护您的应用程序与更广泛的互联网之间的双向接口。

P

权限边界

附加到 IAM 主体的 IAM 管理策略，用于设置用户或角色可以拥有的最大权限。有关更多信息，请参阅 IAM 文档中的权限边界。

个人身份信息（PII）

直接查看其他相关数据或与之配对时可用于合理推断个人身份的信息。PII 的示例包括姓名、地址和联系信息。

PII

请参阅个人身份信息。

playbook

一套预定义的步骤，用于捕获与迁移相关的工作，例如在云中交付核心运营功能。playbook 可以采用脚本、自动化运行手册的形式，也可以是操作现代化环境所需的流程或步骤的摘要。

PLC

请参阅可编程逻辑控制器。

PLM

请参阅产品生命周期管理。

policy

一个对象，可以定义权限（请参阅基于身份的策略）、指定访问条件（请参阅基于资源的策略）或定义 AWS Organizations 的组织中所有账户的最大权限（请参阅服务控制策略）。

多语言持久性

根据数据访问模式和其他要求，独立选择微服务的数据存储技术。如果您的微服务采用相同的数据存储技术，它们可能会遇到实现难题或性能不佳。如果微服务使用最适合其需求的数据存储，则可以更轻松地实现微服务，并获得更好的性能和可扩展性。

组合评测

一个发现、分析和确定应用程序组合优先级以规划迁移的过程。有关更多信息，请参阅评估迁移准备情况。

谓词

返回 `true` 或 `false` 的查询条件，通常位于 `WHERE` 子句中。

谓词下推

一种数据库查询优化技术，可在传输之前筛选查询中的数据。这将减少从关系数据库检索和处理的数据量，并提高查询性能。

预防性控制

一种安全控制，旨在防止事件发生。这些控制是第一道防线，帮助防止未经授权的访问或对网络的意外更改。有关更多信息，请参阅在 AWS上实施安全控制中的预防性控制。

主体

中 AWS 可以执行操作和访问资源的实体。此实体通常是 IAM 角色的根用户或用户。 AWS 账户有关更多信息，请参阅 IAM 文档中角色术语和概念中的主体。

隐私设计

一种在整个开发过程中都考虑隐私的系统工程方法。

私有托管区

一个容器，其中包含有关您希望 Amazon Route 53 如何响应针对一个或多个 VPCs域名及其子域名的 DNS 查询的信息。有关更多信息，请参阅 Route 53 文档中的私有托管区的使用。

主动控制

一种安全控制，旨在防止部署不合规资源。这些控制会在资源预置之前对其进行扫描。如果资源与控制不兼容，则不会预置它。有关更多信息，请参阅 AWS Control Tower 文档中的控制参考指南，并参见在上实施安全控制中的主动控制 AWS。

产品生命周期管理（PLM）

对产品在其整个生命周期内的数据和流程的管理，从设计、开发和发布，到增长和成熟，再到衰退和淘汰。

生产环境

请参阅环境。

可编程逻辑控制器（PLC）

在制造业中，一种高度可靠、适应性强的计算机，用于监控机器并实现制造过程自动化。

提示串接

使用一个 LLM 提示的输出作为下一个提示的输入，以生成更好的响应。该技术用于将复杂的任务分解为子任务，或者迭代地完善或扩展初步响应。它有助于提高模型响应的准确性和相关性，并允许获得更精细的个性化结果。

假名化

用占位符值替换数据集中个人标识符的过程。假名化可以帮助保护个人隐私。假名化数据仍被视为个人数据。

publish/subscribe (pub/sub)

一种支持微服务间异步通信的模式，可提高可扩展性和响应能力。例如，在基于微服务的 MES 中，微服务可以将事件消息发布到其他微服务可以订阅的频道。系统可以在不更改发布服务的情况下添加新的微服务。

Q

查询计划

一系列用于访问 SQL 关系数据库系统中的数据的步骤，类似于指令。

查询计划回归

当数据库服务优化程序选择的最佳计划不如数据库环境发生特定变化之前时。这可能是由统计数据、约束、环境设置、查询参数绑定更改和数据库引擎更新造成的。

R

RACI 矩阵

请参阅责任、问责、咨询和知情（RACI）。

RAG

请参阅检索增强生成。

勒索软件

一种恶意软件，旨在阻止对计算机系统或数据的访问，直到付款为止。

RASCI 矩阵

请参阅责任、问责、咨询和知情（RACI）。

RCAC

请参阅行列访问控制。

只读副本

用于只读目的的数据库副本。您可以将查询路由到只读副本，以减轻主数据库的负载。

重新架构

请参阅 7 R。

恢复点目标（RPO）

自上一个数据恢复点以来可接受的最长时间。这决定了从上一个恢复点到服务中断之间可接受的数据丢失情况。

恢复时间目标（RTO）

服务中断和服务恢复之间可接受的最大延迟。

重构

请参阅 7 R。

Region

地理区域内的 AWS 资源集合。每一个 AWS 区域 都相互隔离，彼此独立，以提供容错、稳定性和弹性。有关更多信息，请参阅指定您的账户可以使用的 AWS 区域。

回归

一种预测数值的 ML 技术。例如，要解决“这套房子的售价是多少？”的问题 ML 模型可以使用线性回归模型，根据房屋的已知事实（如建筑面积）来预测房屋的销售价格。

重新托管

请参阅 7 R。

版本

在部署过程中，推动生产环境变更的行为。

重新放置

请参阅 7 R。

更换平台

请参阅 7 R。

重新购买

请参阅 7 R。

韧性

应用程序抵御中断或从中断中恢复的能力。在 AWS 云中规划韧性时，高可用性和灾难恢复是常见的考虑因素。有关更多信息，请参阅 AWS 云 韧性。

基于资源的策略

一种附加到资源的策略，例如 AmazonS3 存储桶、端点或加密密钥。此类策略指定了允许哪些主体访问、支持的操作以及必须满足的任何其他条件。

责任、问责、咨询和知情（RACI）矩阵

定义参与迁移活动和云运营的所有各方的角色和责任的矩阵。矩阵名称源自矩阵中定义的责任类型：负责（R）、问责（A）、咨询（C）和知情（I）。支持（S）类型是可选的。如果包括支持，则该矩阵称为 RASCI 矩阵，如果将其排除在外，则称为 RACI 矩阵。

响应性控制

一种安全控制，旨在推动对不良事件或偏离安全基线的情况进行修复。有关更多信息，请参阅在AWS上实施安全控制中的响应性控制。

保留

请参阅 7 R。

停用

请参阅 7 R。

检索增强生成（RAG）

一种生成式人工智能技术，其中 LLM 在生成响应之前引用其训练数据来源之外的权威数据来源。例如，RAG 模型可以对组织的知识库或自定义数据执行语义搜索。有关更多信息，请参阅什么是 RAG。

轮换

定期更新密钥以使攻击者更难访问凭证的过程。

行列访问控制（RCAC）

使用已定义访问规则的基本、灵活的 SQL 表达式。RCAC 由行权限和列掩码组成。

RPO

请参阅恢复点目标。

RTO

请参阅恢复时间目标。

运行手册

执行特定任务所需的一套手动或自动程序。它们通常是为了简化重复性操作或高错误率的程序而设计的。

S

SAML 2.0

许多身份提供商 (IdPs) 使用的开放标准。此功能支持联合单点登录 (SSO)，因此用户无需在 IAM 中为组织中的所有人创建用户即可登录 AWS 管理控制台 或调用 AWS API 操作。有关基于 SAML 2.0 的联合身份验证的更多信息，请参阅 IAM 文档中的关于基于 SAML 2.0 的联合身份验证。

SCADA

请参阅监督控制和数据采集。

SCP

请参阅服务控制策略。

机密密钥

在中 AWS Secrets Manager，您以加密形式存储的机密或受限信息，例如密码或用户凭证。它由密钥值及其元数据组成。密钥值可以是二进制、单个字符串或多个字符串。有关更多信息，请参阅 Secrets Manager 文档中的什么是 Amazon Secrets Manager 密钥？。

安全设计

一种在整个开发过程中都考虑安全的系统工程方法。

安全控制

一种技术或管理防护机制，可防止、检测或降低威胁行为体利用安全漏洞的能力。安全控制有以下四种类型：预防性、检测性、响应性和主动性。

安全固化

缩小攻击面，使其更能抵御攻击的过程。这可能包括删除不再需要的资源、实施授予最低权限的最佳安全实践或停用配置文件中不必要的功能等操作。

安全信息和事件管理（SIEM）系统

结合了安全信息管理（SIM）和安全事件管理（SEM）系统的工具和服务。SIEM 系统会收集、监控和分析来自服务器、网络、设备和其他来源的数据，以检测威胁和安全漏洞，并生成警报。

安全响应自动化

一种预定义的程序化操作，旨在自动响应或修复安全事件。这些自动化可作为侦探或响应式安全控制措施，帮助您实施 AWS 安全最佳实践。自动响应操作的示例包括修改 VPC 安全组、修补 Amazon EC2 实例或轮换凭证。

服务器端加密

由接收数据的人在目的地对数据 AWS 服务 进行加密。

服务控制策略（SCP）

一种策略，用于集中控制组织中所有账户的权限 AWS Organizations。 SCPs 定义防护措施或限制管理员可以委托给用户或角色的操作。您可以使用 SCPs 允许列表或拒绝列表来指定允许或禁止哪些服务或操作。有关更多信息，请参阅 AWS Organizations 文档中的服务控制策略。

服务端点

的入口点的 URL AWS 服务。您可以使用端点，通过编程方式连接到目标服务。有关更多信息，请参阅 AWS 一般参考 中的 AWS 服务 端点。

服务水平协议（SLA）

一份协议，阐明了 IT 团队承诺向客户交付的内容，比如服务正常运行时间和性能。

服务水平指示器（SLI）

对服务性能方面的衡量，例如错误率、可用性或吞吐量。

服务水平目标（SLO）

代表服务运行状况的目标指标，由服务水平指示器衡量。

责任共担模式

描述您在云安全与合规方面共同承担 AWS 的责任的模型。 AWS 负责云的安全，而您则负责云中的安全。有关更多信息，请参阅责任共担模式。

SIEM

请参阅安全信息和事件管理系统。

单点故障（SPOF）

应用程序的单个关键组件出现故障，可能会中断系统。

SLA

请参阅服务水平协议。

SLI

请参阅服务水平指示器。

SLO

请参阅服务水平目标。

split-and-seed 模型

一种扩展和加速现代化项目的模式。随着新功能和产品发布的定义，核心团队会拆分以创建新的产品团队。这有助于扩展组织的能力和服务，提高开发人员的工作效率，支持快速创新。有关更多信息，请参阅在 AWS 云中实现应用程序现代化的分阶段方法。

SPOF

请参阅单点故障。

星型架构

一种数据库组织结构，它使用一个大型事实表来存储事务数据或测量数据，并使用一个或多个较小的维度表来存储数据属性。此结构专为在数据仓库中使用或用于商业智能目的而设计。

strangler fig 模式

一种通过逐步重写和替换系统功能直至可以停用原有的系统来实现单体系统现代化的方法。这种模式用无花果藤作为类比，这种藤蔓成长为一棵树，最终战胜并取代了宿主。该模式是由 Martin Fowler 提出的，作为重写单体系统时管理风险的一种方法。有关如何应用此模式的示例，请参阅使用容器和 Amazon API Gateway 逐步将原有的 Microsoft ASP.NET（ASMX）Web 服务现代化。

子网

您的 VPC 内的一个 IP 地址范围。子网必须位于单个可用区中。

监督控制和数据采集（SCADA）

在制造业中，一种使用硬件和软件来监控实物资产和生产操作的系统。

对称加密

一种加密算法，它使用相同的密钥来加密和解密数据。

综合测试

以模拟用户交互的方式测试系统，以检测潜在问题或监控性能。你可以使用 Amazon S CloudWatch ynthetics 来创建这些测试。

系统提示

一种为 LLM 提供上下文、说明或准则以指导其行为的技术。系统提示有助于设置上下文并制定与用户交互的规则。

T

标签

键值对，用作组织资源的元数据。 AWS 标签有助于您管理、识别、组织、搜索和筛选 资源。有关更多信息，请参阅标记您的 AWS 资源。

目标变量

您在监督式 ML 中尝试预测的值。这也被称为结果变量。例如，在制造环境中，目标变量可能是产品缺陷。

任务列表

一种通过运行手册用于跟踪进度的工具。任务列表包含运行手册的概述和要完成的常规任务列表。对于每项常规任务，它包括预计所需时间、所有者和进度。

测试环境

请参阅环境。

训练

为您的 ML 模型提供学习数据。训练数据必须包含正确答案。学习算法在训练数据中查找将输入数据属性映射到目标（您希望预测的答案）的模式。然后输出捕获这些模式的 ML 模型。然后，您可以使用 ML 模型对不知道目标的新数据进行预测。

中转网关

一个网络传输中心，可用于将您的网络 VPCs 和本地网络互连。有关更多信息，请参阅 AWS Transit Gateway 文档中的什么是公交网关。

基于中继的工作流程

一种方法，开发人员在功能分支中本地构建和测试功能，然后将这些更改合并到主分支中。然后，按顺序将主分支构建到开发、预生产和生产环境。

可信访问权限

向您指定的服务授予权限，该服务可代表您在其账户中执行任务。 AWS Organizations 当需要服务相关的角色时，受信任的服务会在每个账户中创建一个角色，为您执行管理任务。有关更多信息，请参阅 AWS Organizations 文档中的AWS Organizations 与其他 AWS 服务一起使用。

优化

更改训练过程的各个方面，以提高 ML 模型的准确性。例如，您可以通过生成标签集、添加标签，并在不同的设置下多次重复这些步骤来优化模型，从而训练 ML 模型。

双披萨团队

一个小 DevOps 团队，你可以用两个披萨来喂食。双披萨团队的规模可确保在软件开发过程中充分协作。

U

不确定性

这一概念指的是不精确、不完整或未知的信息，这些信息可能会破坏预测式 ML 模型的可靠性。不确定性有两种类型：认知不确定性是由有限的、不完整的数据造成的，而偶然不确定性是由数据中固有的噪声和随机性导致的。

无差别任务

也称为繁重工作，即创建和运行应用程序所必需的工作，但不能为最终用户提供直接价值或竞争优势。无差别任务的示例包括采购、维护和容量规划。

上层环境

请参阅环境。

V

vacuum 操作

一种数据库维护操作，包括在增量更新后进行清理，以回收存储空间并提高性能。

版本控制

跟踪更改的过程和工具，例如存储库中源代码的更改。

VPC 对等连接

两者之间的连接 VPCs ，允许您使用私有 IP 地址路由流量。有关更多信息，请参阅 Amazon VPC 文档中的什么是 VPC 对等连接。

漏洞

损害系统安全的软件缺陷或硬件缺陷。

W

热缓存

一种包含经常访问的当前相关数据的缓冲区缓存。数据库实例可以从缓冲区缓存读取，这比从主内存或磁盘读取要快。

暖数据

不常访问的数据。查询此类数据时，通常可以接受中速查询。

窗口函数

一种对与当前记录有某种关联的一组行执行计算的 SQL 函数。窗口函数对于处理任务很有用，例如计算移动平均值或根据当前行的相对位置访问行的值。

工作负载

一系列资源和代码，它们可以提供商业价值，如面向客户的应用程序或后端过程。

工作流

迁移项目中负责一组特定任务的职能小组。每个工作流都是独立的，但支持项目中的其他工作流。例如，组合工作流负责确定应用程序的优先级、波次规划和收集迁移元数据。组合工作流将这些资产交付给迁移工作流，然后迁移服务器和应用程序。

WORM

请参阅一次写入多次读取。

WQF

请参阅 AWS 工作负载资格鉴定框架。

一次写入多次读取（WORM）

一种存储模型，可一次写入数据并防止数据被删除或修改。授权用户可以根据需要多次读取数据，但无法对其进行更改。此数据存储基础设施被认为不可变。

Z

零日漏洞利用

一种利用零日漏洞的攻击，通常为恶意软件。

零日漏洞

生产系统中不可避免的缺陷或漏洞。威胁主体可能利用这种类型的漏洞攻击系统。开发人员经常因攻击而意识到该漏洞。

零样本提示

为 LLM 提供执行任务的说明，但没有可以帮助指导的示例（样本）。LLM 必须使用预先训练的知识来处理任务。零样本提示的有效性取决于任务的复杂性和提示的质量。另请参阅少样本提示。

僵尸应用程序

平均 CPU 和内存使用率低于 5% 的应用程序。在迁移项目中，通常会停用这些应用程序。

本文属于机器翻译版本。若本译文内容与英语原文存在差异，则一律以英文原文为准。