



Economia para IA agente em AWS

AWS Orientação prescritiva



AWS Orientação prescritiva: Economia para IA agente em AWS

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

As marcas comerciais e imagens de marcas da Amazon não podem ser usadas no contexto de nenhum produto ou serviço que não seja da Amazon, nem de qualquer maneira que possa gerar confusão entre os clientes ou que deprecie ou desprestige a Amazon. Todas as outras marcas comerciais que não pertencem à Amazon pertencem a seus respectivos proprietários, que podem ou não ser afiliados, patrocinados pela Amazon ou ter conexão com ela.

Table of Contents

Introdução	1
Público-alvo	1
Objetivos	2
Sobre esta série de conteúdo	2
Entendendo a economia da IA agêntica	3
Avaliação de tarefas	3
Avaliação do impacto do risco	4
Retorno sobre o investimento	6
Medindo o sucesso e o ROI	7
Use sua base	7
Estabeleça metas	7
Monitore métricas	7
Use AgentOps	7
Avaliação dos custos do processo humano	8
Custos trabalhistas	9
Custos de desempenho	9
Custos de tecnologia	10
Custos de oportunidade	10
Custos de riscos e defeitos	11
Implementação de sistemas de IA agentes	12
Incorporando feedback humano	13
aprendizagem comportamental	13
Aprendizagem contínua	14
Colaboração entre humanos e IA	14
Preços baseados em resultados	15
Modelo tradicional e inicial	15
Modelo baseado em resultados	16
Usando AWS Marketplace	17
Estudo de caso: operações de recrutamento	18
Cenário A	18
Estrutura de custo base	19
Métricas operacionais	20
Análise de custos baseada em volume	21
Análise de ROI	21

Comparação cumulativa de custos	21
Benefícios adicionais	22
Cenário B	22
Estrutura de custo base	23
Métricas operacionais	24
Análise de custos baseada em volume	24
Análise de ROI	25
Comparação cumulativa de custos	25
Comparando cenários	26
Conclusão e atributos	28
Recursos	28
Histórico do documento	30
Glossário	31
#	31
A	32
B	35
C	37
D	40
E	44
F	46
G	48
H	49
eu	51
L	53
M	54
O	59
P	61
Q	64
R	65
S	68
T	72
U	73
V	74
W	74
Z	75
.....	lxxvii

Economia para IA agêntica em AWS

Hans Schabert e Prasanta Roy, da Amazon Web Services

Janeiro de 2026 ([histórico do documento](#))

As organizações que adotam automação orientada por IA e sistemas de IA agentes precisam tomar decisões econômicas informadas entre mão de obra humana e agentes inteligentes. Isso se torna essencial para operações sustentáveis na nuvem. Este guia ajuda você a avaliar, implementar e otimizar as compensações econômicas entre a força de trabalho humana e os sistemas de IA agentes em AWS. Você pode maximizar seu retorno sobre o investimento (ROI) enquanto mantém a excelência operacional.

Nenhum sistema está 100% correto. Esse princípio fundamental impulsiona a análise econômica dos sistemas de IA humanos e agentes. As organizações devem ir além das comparações simplistas de custos para avaliar o impacto econômico total, os perfis de risco, os requisitos de qualidade das decisões e a criação de valor estratégico de longo prazo.

O comportamento do cliente está mudando drasticamente dos tradicionais investimentos iniciais em tecnologia para pay-per-outcome modelos que alinham os custos aos resultados comerciais. Essa transformação requer novas abordagens para avaliação, implementação e otimização da colaboração entre agentes humanos.

O caminho para o sucesso segue um padrão claro: comece com trabalhos adequados, meça tudo e escale o que funciona. As organizações que adotam essa abordagem obtêm uma vantagem competitiva sustentável por meio da alocação inteligente de recursos e da automação focada em resultados.

Público-alvo

Este guia se destina ao seguinte:

- Executivos (CEOs, CTOs, CFOs) que estão tomando decisões estratégicas de investimento
- Arquitetos corporativos que estão projetando estratégias de automação organizacional
- Profissionais de operações financeiras que estão otimizando o gerenciamento financeiro na nuvem
- Líderes de tecnologia que estão avaliando abordagens de implementação de IA
- Líderes de unidades de negócios que desejam entender o ROI da automação

- Profissionais de compras que estão navegando em novos modelos de preços de IA

Para entender os conceitos deste guia, recomendamos que você analise os [Fundamentos da IA agêntica em AWS](#).

Objetivos

Este guia ajuda você a entender o seguinte:

- Como avaliar o potencial de automação de uma agência de empregos
- Modelos econômicos para comparar custos de mão de obra humana com investimentos em sistemas de IA agentes
- Pay-per-outcome modelos de preços e seu impacto na economia de projetos de IA
- Técnicas de medição para demonstrar o ROI e gerenciar riscos
- Estratégias de escalabilidade que transformam custos fixos em resultados variáveis

Sobre esta série de conteúdo

Este guia faz parte de uma série sobre IA agente em AWS. Para obter mais informações e ver os outros guias desta série, consulte [Agentic AI](#) no site da AWS Prescriptive Guidance.

Entendendo a economia da IA agêntica em AWS

Um dos princípios fundamentais é determinar quando usar agentes de IA e quando usar métodos determinísticos tradicionais. As organizações devem avaliar sistematicamente quais trabalhos justificam a automação automática e quais devem usar a automação tradicional ou a operação humana contínua. Essa decisão requer a compreensão da relação entre as características da tarefa, a tolerância ao risco e a abordagem operacional.

Antes de decidir implementar a IA agente, você deve usar a estrutura de decisão para entender o impacto econômico. A estrutura de decisão inclui as seguintes três questões-chave:

1. [Avaliação de tarefas](#) — Essa tarefa é adequada para um agente de IA?
2. [Avaliação do impacto do risco](#) — Quais são os riscos envolvidos?
3. [Retorno sobre o investimento](#) — Será econômico?

Avaliação de tarefas

Tarefas com regras de decisão padronizadas e de alta complexidade podem se beneficiar de abordagens de IA agênticas. Tarefas simples e altamente padronizadas são melhor atendidas pela automação tradicional ou pela automação robótica de processos. Os sistemas de IA da Agentic se destacam no raciocínio, na compreensão do contexto ou na tomada de decisões de forma adaptativa. Eles agregam valor além do processamento baseado em regras. Implementações bem-sucedidas de IA para agências exigem sistemas capazes de aprender e se adaptar.

Considere os seguintes fatores ao avaliar uma tarefa:

- **Complexidade** — Grau de raciocínio e compreensão do contexto necessários. Tarefas que exigem compreensão contextual, interpretação diferenciada ou respostas adaptativas às mudanças nas condições favorecem as abordagens agênticas em relação à automação tradicional, enquanto tarefas puramente mecânicas ou de cálculo podem não exigir inteligência agente.
- **Padronização** — Presença de padrões e regras claras. A IA agente é recomendada se a tarefa exigir compreensão contextual. Se nenhuma adaptação ou aprendizado for necessário, considere a automação tradicional.
- **Volume** — Frequência do desempenho da tarefa. A IA agente é recomendada para atividades autônomas. A automação tradicional é recomendada para tarefas consistentes e de alto volume. No entanto, o volume por si só não determina a abordagem. Decisões de baixo volume e alto valor

podem justificar a assistência de um agente para melhorar a qualidade das decisões, em vez da redução de custos.

- Valor — impacto nos negócios por conclusão de tarefa. Considere a IA agente para obter resultados de alto valor que exijam capacidade autônoma semelhante à humana. Considere a automação tradicional para tarefas repetidas e consistentes, que podem ser feitas de forma determinística.

Avaliação do impacto do risco

Atualmente, existem quatro abordagens de implantação de IA agente: totalmente autônoma, humana no circuito, copiloto ou liderada por humanos com suporte de agentes. Cada um tem seu próprio perfil de risco e tolerância a erros, e todos envolvem humanos de alguma forma. A tabela a seguir descreve os detalhes dos riscos dessas abordagens.

Nível de autonomia	Perfil de risco	tolerância a erros	Exemplos de casos de uso	Envolvimento humano
Totalmente autônomo	Baixo risco	1-2% aceitável	<ul style="list-style-type: none"> • Categorização básica de dados • Roteamento de documentos • Geração de relatórios padrão 	<ul style="list-style-type: none"> • Supervisão mínima • Auditorias periódicas
Humano no circuito	Risco médio	Abaixo de 0,5%	<ul style="list-style-type: none"> • Esboço de respostas • Moderação de conteúdo 	<ul style="list-style-type: none"> • Revisão regular • Gerenciamento de exceções

			<ul style="list-style-type: none"> • Processamento inicial de reclamações 	<ul style="list-style-type: none"> • Garantia de qualidade
Co-piloto	Alto risco	Quase zero	<ul style="list-style-type: none"> • Entrada para o planejamento estratégico 	<ul style="list-style-type: none"> • O ser humano toma as decisões finais
			<ul style="list-style-type: none"> • Avaliações de riscos • Decisões de investimento 	<ul style="list-style-type: none"> • O agente fornece recomendações
Liderado por humanos com suporte de agentes	Risco crítico	Tolerância zero	<ul style="list-style-type: none"> • Decisões legais • Diagnóstico médico • Conformidade regulatória 	<ul style="list-style-type: none"> • Processo de condução humana • O agente fornece pesquisas ou análises e apenas informações de apoio

A tabela a seguir descreve as principais considerações ao escolher entre essas abordagens.

Consideração	Totalmente autônomo	Humano no circuito	Co-piloto	Liderado por humanos
Eficiência de custos	Mais alto	Alto	Médio	Baixo
Escalabilidade	Ilimitado	Alto	Médio	Limitado

Velocidade de processamento	Mais rápido	Fast	Médio	Lento
Gestão de riscos	Básico	Avançado	Forte	Mais forte
Tratamento da complexidade	Tarefas simples	Tarefas moderadamente complexas	Tarefas complexas	Tarefas críticas

Essa estrutura de consideração ajuda as organizações a adequar os níveis de autonomia aos perfis de risco, escalar as operações de forma adequada, equilibrar eficiência com controle, implementar a governança adequada e otimizar a alocação de recursos.

Retorno sobre o investimento

O cálculo do retorno sobre o investimento em sistemas de IA agentes começa com uma análise abrangente de custos. As organizações devem primeiro calcular seus custos humanos atuais, incluindo salários, benefícios e despesas de espaço de trabalho, juntamente com despesas específicas do processo e custos ocultos, como treinamento, cobertura e tempo de inatividade.

Para uma análise de equilíbrio, as organizações devem considerar os custos de implementação, as despesas operacionais contínuas e o volume necessário para justificar o investimento. Também é importante considerar as variações sazonais e os benefícios da curva de aprendizado que surgem à medida que os sistemas amadurecem e melhoram com o tempo.

Ao avaliar os agentes de IA, as organizações devem se lembrar de que esses sistemas geralmente têm custos iniciais mais altos, mas menores custos por transação em comparação com as operações humanas. Além disso, os agentes de IA demonstram uma melhoria no desempenho ao longo do tempo e oferecem melhor escalabilidade do que as equipes humanas. Isso os torna cada vez mais econômicos à medida que a implantação aumenta e a experiência operacional se acumula.

Medindo o sucesso e o ROI de sistemas de IA agentes

Medir o sucesso na implementação de um sistema de IA agente requer uma abordagem sistemática. Esta seção fornece uma metodologia clara para avaliação e otimização contínua que usa sua análise existente em vez de começar do zero.

Etapa 1: use sua base existente

Comece com uma avaliação abrangente de custos de acordo com as recomendações na seção [Avaliação dos custos atuais do processo](#). Isso fornece uma linha de base operacional para seus cálculos de ROI. Conforme descrito na seção [Avaliação de impacto de risco](#), escolha entre os quatro níveis de autonomia (totalmente autônomo, humano em loop, abordagem copiloto, liderada por humanos com o suporte do agente) para determinar os critérios de medição apropriados e os limites de tolerância a erros para cada processo.

Etapa 2: Estabeleça metas claras de sucesso

Estabeleça metas de arquitetura e sucesso que enfatizem sistemas capazes de aprender, conforme descrito na seção [Padrões de sucesso para a implementação de sistemas de IA agentes](#). Concentre-se na melhoria contínua em vez do desempenho estático. Defina cronogramas de ROI usando a metodologia de análise de equilíbrio demonstrada no [Estudo de caso: comparação dos custos de IA humana e agente para operações de recrutamento](#). Inclua pontos de decisão claros para demitir agentes inadimplentes.

Etapa 3: Acompanhe as principais métricas

Monitore o desempenho financeiro em relação à linha de base estabelecida e acompanhe a economia de custos e as melhorias estratégicas de valor. Avalie as métricas operacionais, incluindo taxas de erro dentro dos limites aceitáveis para o nível de autonomia escolhido, melhorias na velocidade de processamento e ganhos de consistência. Concentre-se em indicadores estratégicos que demonstrem capacidade de aprendizado e adaptação ao longo do tempo.

Etapa 4: Usar AgentOps

Aplique a estrutura de aprendizado contínuo da seção [Incorporando feedback humano em sistemas de IA agentes](#) para otimizar a tomada de decisões por meio da integração sistemática do feedback

humano. Crie sistemas de aprendizado em tempo real que incorporem insights humanos para aprimorar o desempenho. Monitore a transformação em direção a modelos de negócios baseados em resultados, conforme descrito em [Transformação econômica para preços baseados em resultados para](#) sistemas de IA agentes em. AWS

Avaliando seus custos atuais de processos humanos

Entender os verdadeiros custos do processo é fundamental para tomar decisões informadas sobre investimentos em sistemas de IA agentes. Primeiro, você deve estabelecer uma linha de base precisa do custo atual de seus processos, incluindo todas as despesas ocultas, taxas de falha e custos de oportunidade. Isso ajuda você a desenvolver cálculos precisos de ROI e tomar decisões estratégicas. Essa avaliação abrangente de custos serve como base crítica para avaliar se os sistemas de IA agentes podem oferecer valor genuíno como companheiros produtivos.

A avaliação básica dos custos é essencial pelos seguintes motivos principais:

- Precisão do ROI — Linhas de base de custo precisas suportam projeções realistas de ROI que respondem por todo o espectro das despesas operacionais atuais.
- Estratégia de implementação do agente — A compreensão abrangente dos custos ajuda as organizações a identificar os processos mais promissores para a implantação inicial do sistema de IA agente.
- Medição de desempenho — As linhas de base estabelecidas fornecem a estrutura de medição para rastrear os benefícios reais e projetados das implementações de IA agente.

As organizações devem identificar e avaliar sistematicamente todos os fatores de custo que influenciam a economia do processo antes de comparar alternativas humanas e agentes. Essa avaliação garante cálculos de linha de base precisos, contabilizando fatores de custo óbvios e ocultos. Ele dá ênfase especial aos custos de falhas, às taxas históricas de falhas e às oportunidades de negócios perdidas que representam o verdadeiro custo total dos processos atuais.

Esta seção descreve como coletar dados em cada categoria de custo para estabelecer medições de linha de base precisas para seus processos atuais. Ele discute as fontes de informação e fornece exemplos das seguintes categorias de custo:

- [Custos trabalhistas](#)
- [Custos de desempenho humano e consistência](#)
- [Custos de tecnologia e infraestrutura](#)

- [Custos de oportunidades de negócios perdidos](#)
- [Custos de riscos e defeitos](#)

Custos trabalhistas

Extraia 24 meses de dados da folha de pagamento que incluem salário base, horas extras, benefícios e custos de treinamento. Use seu sistema de informações de recursos humanos (HRIS) para monitorar as despesas de recrutamento e as taxas de rotatividade. Os sistemas de controle de tempo revelam a produtividade real versus as horas programadas. As plataformas de gerenciamento de desempenho mostram a correlação entre os níveis de habilidade e os custos de compensação. Calcule as taxas horárias totalmente carregadas que alocam para despesas gerais de gerenciamento.

Veja a seguir uma lista de exemplos de fatores de custo da mão de obra.

Gerador de custos	Impacto nos negócios
Compensação básica	\$25—150 por hora totalmente carregada
Benefícios e impostos sobre a folha de pagamento	25— 40% do salário base
Treinamento e desenvolvimento	5— 15% do custo anual de mão de obra
Sobrecarga de gerenciamento	15— 25% do custo de mão de obra direta

Custos de desempenho humano e consistência

Combine dados de sistemas de gerenciamento de projetos que mostram variações de conclusão de tarefas com sistemas de atendimento. Isso pode revelar padrões de absenteísmo e mudanças sazonais. As plataformas de atendimento ao cliente demonstram faixas de desempenho individual por meio de métricas de resolução, e os dados de gerenciamento de relacionamento com o cliente (CRM) de vendas podem mostrar variações de eficiência no fechamento de negócios. Os sistemas de gerenciamento de qualidade fornecem taxas de defeitos e dados de conformidade de processos entre equipes e locais. Os sistemas de fluxo de trabalho capturam tempos de conclusão, atrasos na aprovação e frequência de tratamento de exceções. A análise de comunicação revela a sobrecarga de coordenação por meio da frequência das reuniões e dos padrões de colaboração.

Veja a seguir uma lista de exemplos de fatores de custo para desempenho e consistência humanos.

Gerador de custos	Impacto nos negócios
Flutuações de produtividade	Faixa de desempenho de 20 a 50%
Absenteísmo e cobertura	É necessária uma capacidade adicional de 15 a 25%
Ciclos de fadiga e motivação	Variação de produtividade de 10 a 30%
Inconsistências do procedimento	10— 40% de perda de eficiência
Variações do controle de qualidade	10— 30% do custo total
Sobrecarga de coordenação	15— 25% do custo operacional

Custos de tecnologia e infraestrutura

As plataformas de gerenciamento de licenças mostram os custos e as taxas de utilização do software. O monitoramento da infraestrutura fornece dados de disponibilidade, métricas de desempenho e custos de manutenção. Os sistemas de suporte técnico monitoram a sobrecarga de suporte e os problemas técnicos recorrentes. Os sistemas de gerenciamento de fornecedores capturam os custos totais do relacionamento tecnológico, incluindo despesas de integração e desempenho do nível de serviço.

Veja a seguir uma lista de exemplos de fatores de custo para tecnologia e infraestrutura.

Gerador de custos	Impacto nos negócios
Sistemas de tecnologia	\$50—500 por usuário por mês
Espaço de trabalho e equipamento	\$200—1.000 por funcionário por mês

Custos de oportunidades de negócios perdidos

As plataformas de CRM contêm tempos de resposta de leads, taxas de conversão e documentação de oportunidades perdidas. A automação de marketing mostra os impactos do atraso no

acompanhamento nas conversões de leads. Os sistemas de suporte ao cliente revelam como os problemas operacionais afetam a satisfação e a retenção. A análise competitiva fornece requisitos de resposta do mercado e dados de ganhos ou perdas que conectam o desempenho operacional aos resultados da receita.

Veja a seguir uma lista de exemplos de fatores de custo para oportunidades de negócios perdidas.

Gerador de custos	Impacto nos negócios
Atrasos na resposta do mercado	Receita por dia de atraso
Restrições de capacidade	Oportunidades de negócios perdidas
Alocação de recursos de inovação	Custo de oportunidade do trabalho rotineiro
Atrasos na aquisição de clientes	50— 90% de perda de chumbo devido à resposta lenta

Custos de riscos e defeitos

A documentação da apólice de seguro mostra os custos de responsabilidade geral, responsabilidade profissional, compensação de trabalhadores e cobertura de responsabilidade cibernética. Os relatórios internos de avaliação de riscos identificam as vulnerabilidades operacionais e os custos de mitigação associados. Os sistemas de rastreamento de defeitos documentam falhas de produtos ou serviços, incluindo custos de detecção, despesas de substituição e reclamações de garantia. Os cronogramas de substituição de ativos mostram as taxas de falha do equipamento e os custos de substituição. Os relatórios de incidentes de segurança rastreiam acidentes de trabalho e pedidos de indenização de trabalhadores associados. Os planos de continuidade de negócios detalham os custos do sistema de backup e os investimentos em recuperação de desastres.

Veja a seguir uma lista de exemplos de fatores de custo para riscos e defeitos.

Gerador de custos	Impacto nos negócios
Custos de seguro	1— 5% do orçamento operacional
Custo dos erros	\$50—5.000 por incidente de erro
Impacto do erro humano	2— 15% do custo operacional total

Taxas de erro e retrabalho

1,5 a 4 vezes o custo original das correções

Padrões bem-sucedidos para implementar sistemas de IA agentes em AWS

O [estado da adoção da IA corporativa](#) (relatório ISG 2025) revela que a principal barreira para a implementação bem-sucedida da IA não é a capacidade técnica, mas a lacuna de aprendizado. Esse termo se refere a sistemas que não conseguem se adaptar, lembrar o contexto ou melhorar com o tempo. Organizações que implementam ferramentas estáticas de IA apresentam altas taxas de falha. A seguir estão as características comuns dos sistemas de IA agentes que alcançam o sucesso:

- Memória contextual — Sistemas que retêm o histórico de conversas e as preferências do usuário
- Integração de feedback — Capacidade de aprender com as correções e melhorar o desempenho
- Adaptação do fluxo de trabalho — ajuste automático às mudanças nos requisitos de negócios
- Melhoria contínua — Aprimoramento mensurável por meio da experiência operacional

Organizações que obtêm implementações bem-sucedidas de IA geralmente priorizam o seguinte:

- Usando ecossistemas de parceiros abrangentes em vez de criar e explorar recursos de IA de forma independente
- Sistemas capazes de aprender em vez de ferramentas estáticas
- Foco nos resultados comerciais em vez da comparação de recursos técnicos
- Integração de fluxo de trabalho em vez de ferramentas autônomas
- Adaptação contínua em vez de implementação única

[Esses padrões se alinham a muitos AWS service \(Serviço da AWS\) recursos, particularmente o acesso ao modelo básico no Amazon Bedrock, a arquitetura orientada a eventos e o monitoramento AWS Lambdaabrangente oferecido pela Amazon. CloudWatch](#) Para obter mais informações sobre como integrar feedback humano e sistemas capazes de aprender, consulte a seção [Incorporando feedback humano em sistemas de IA agentes neste guia](#).

Incorporando feedback humano em sistemas de IA agentes

Nenhum sistema é 100% bem-sucedido e a falha está fadada a acontecer. Com cada falha, há um custo de mudança associado. Human in the loop é uma abordagem de IA em que a IA executa uma tarefa, mas a intervenção ou aprovação humana é necessária. Essa abordagem deve ser usada quando o custo da falha é maior do que o custo de ter uma human-in-the-loop solução.

O sucesso dos sistemas de IA da agência depende fundamentalmente da capacidade do agente de aprender e melhorar por meio do feedback humano. O custo do esforço humano deve ser levado em consideração, dependendo do nível de esforço necessário. Ao contrário das ferramentas de automação estática que executam regras predeterminadas, as human-in-the-loop soluções têm sistemas agentes capazes de aprender que criam uma parceria dinâmica entre os agentes autônomos e o ser humano. A experiência humana aprimora continuamente o desempenho do agente, enquanto os agentes lidam com o processamento de rotina em grande escala. Essa abordagem colaborativa transforma a implementação de IA de uma implantação única em um processo contínuo de otimização. O sistema se adapta aos padrões organizacionais, internaliza os padrões de qualidade e refina suas capacidades de tomada de decisão com base na experiência operacional do mundo real. Ao capturar sistematicamente correções, aprovações e insights humanos, as organizações podem criar agentes de IA que entendam o contexto, reconheçam padrões e se alinhem cada vez mais aos objetivos de negócios ao longo do tempo.

Para soluções que não exigem intervenção ou suporte humano, não há necessidade de incluir os custos humanos específicos na economia do agente.

Aprendizagem comportamental de operadores humanos

Os operadores humanos fornecem feedback crítico que os sistemas de IA agentes podem usar para aprender, adaptar e melhorar suas respostas ao longo do tempo. Esse ciclo de feedback cria um ambiente colaborativo em que a experiência humana aprimora a capacidade do agente enquanto os agentes lidam com o processamento de rotina.

Por meio do reconhecimento de padrões de comportamento humano, os agentes aprendem com os padrões de interação humana para espelhar abordagens de comunicação bem-sucedidas. Isso os ajuda a se adaptar aos padrões de decisão organizacionais e aos níveis de tolerância ao risco. Os sistemas internalizam as expectativas de qualidade por meio de correções e aprovações humanas. Eles também podem aprender respostas apropriadas para diferentes segmentos de clientes e contextos de negócios.

Mecanismos eficazes de coleta de feedback capturam sistematicamente edições e modificações humanas nas respostas dos agentes. Eles analisam o que os revisores humanos aprovam, rejeitam ou modificam nas recomendações dos agentes. Ao entender por que certos casos exigem intervenção humana e incorporando a avaliação humana do desempenho do agente em diferentes cenários e níveis de complexidade, esses sistemas refinam continuamente suas capacidades para se alinharem mais de perto aos padrões e expectativas organizacionais.

Operações de aprendizado contínuo

A integração do aprendizado em tempo real permite que os sistemas de IA agentes incorporem feedback humano e melhorem as respostas dos agentes imediatamente por meio da atualização dinâmica do modelo. Esses sistemas usam insights humanos para identificar novos padrões e casos extremos. Isso aprimora suas capacidades de reconhecimento de padrões e, ao mesmo tempo, constrói a memória organizacional por meio de experiências de aprendizado guiadas por humanos. O refinamento contínuo com base no feedback entre humanos e operadores e nos resultados comerciais impulsiona a otimização contínua do desempenho.

O treinamento guiado por humanos captura conhecimento especializado para aprimorar as capacidades de tomada de decisão dos agentes. Ele transfere conhecimentos essenciais de operadores experientes para o sistema de IA. Por meio do aprendizado baseado em cenários, os sistemas usam exemplos criados por humanos para melhorar o tratamento de situações complexas. Eles também alinham os padrões de desempenho do agente às expectativas de qualidade humana por meio da calibração de qualidade. Essa abordagem incorpora insights humanos sobre a cultura organizacional e as expectativas dos clientes. Essa adaptação cultural ajuda os agentes a responder adequadamente em diferentes contextos.

Excelência operacional com colaboração entre humanos e IA

A otimização automatizada com reconhecimento de riscos permite a avaliação contínua das condições operacionais e da probabilidade de erro com supervisão humana para cenários de alto risco. Isso ajuda os sistemas a aprender com as avaliações de risco humanas e a melhorar a tomada de decisões futuras. [O Amazon Bedrock](#) fornece acesso a vários modelos de fundação com diferentes recursos e perfis de custo. Isso permite um roteamento inteligente que considera os perfis de custo e risco e, ao mesmo tempo, incorpora feedback humano para otimizar a seleção do modelo. O ajuste de desempenho equilibra a eficiência com a minimização da taxa de erro, incorporando feedback humano sobre padrões de qualidade e compensações aceitáveis de desempenho. As decisões automatizadas consideram o custo total de propriedade ajustado ao risco. Os operadores fornecem orientação sobre tolerância ao risco organizacional e ponderação das prioridades de

negócios. Isso ajuda você a otimizar os custos e, ao mesmo tempo, a se alinhar aos objetivos organizacionais.

Os sistemas de aprendizagem aprimorados por humanos priorizam a contribuição humana por meio do impacto do erro e das consequências comerciais. Isso cria sistemas de aprendizado que entendem a precisão técnica e o contexto comercial por meio de feedback ponderado pelo risco. A análise regular de desempenho incorpora métricas de risco e análise de custos de erros, com insights humanos fornecendo um contexto que os sistemas automatizados não conseguem capturar. O desenvolvimento de melhores práticas enfatiza o gerenciamento de riscos e a prevenção de erros, combinando o reconhecimento automatizado de padrões com a experiência e o julgamento humanos. A capacitação organizacional por meio de programas de treinamento desenvolve habilidades humanas para gerenciar sistemas de IA agentes e capacidades de agentes para apoiar a tomada de decisões humanas. Isso garante uma abordagem abrangente para a colaboração entre humanos e IA que fortalece os dois componentes da parceria.

Transformação econômica em preços baseados em resultados para sistemas de IA agentes em AWS

A mudança dos modelos tradicionais de custo fixo para preços baseados em resultados representa uma transformação fundamental na forma como as organizações estruturam suas operações econômicas e gerenciam riscos. Essa transformação serve como um caminho para a modernização constante dos processos existentes e, ao mesmo tempo, financia a transformação da IA agente. Ele permite que as organizações evoluam de operações estáticas que consomem muitos recursos para modelos de negócios dinâmicos e orientados a resultados.

Modelo tradicional e inicial

Os departamentos geralmente operam como centros de custo com custos diretos de mão de obra financiados pela alocação de custos. Normalmente, as organizações desejam reduzir essa alocação de custos. Se o processo não for modernizado, o departamento deve oferecer os mesmos resultados com uma força de trabalho menor. Isso normalmente degrada a qualidade. Os modelos de negócios tradicionais criam desafios significativos, incluindo:

- Escalonamento linear de custos com aumentos de volume — Isso exige que as organizações contratem pessoal adicional para lidar com o aumento do volume.
- Compromissos de custo fixo — Eles persistem independentemente do desempenho dos negócios e da eficiência do processo.

- Planejamento avançado — A flexibilidade limitada durante crises econômicas e restrições de capacidade requer planejamento antecipado.
- Ciclo de degradação da qualidade — Orçamentos reduzidos levam à diminuição da qualidade do serviço quando os custos são cortados sem melhorias no processo.

Modelo baseado em resultados

Modelos modernos baseados em resultados vinculam os pagamentos diretamente a resultados comerciais mensuráveis, como contratações bem-sucedidas concluídas, métricas de qualidade alcançadas, melhorias na eficiência do processo ou ganhos de produtividade obtidos. Isso basicamente transfere o risco financeiro das unidades de negócios para os prestadores de serviços, ao mesmo tempo em que cria um alinhamento natural de incentivos. A seguir estão os principais benefícios de um modelo baseado em resultados:

- Os custos aumentam diretamente com o valor comercial gerado
- Alinhamento natural entre despesas operacionais e receita
- Flexibilidade para ajustar a capacidade com base nas condições do mercado
- Pay-per-success os modelos reduzem o risco financeiro transferindo a exposição financeira do investimento inicial para o desempenho operacional contínuo
- Concentre-se em sistemas capazes de aprender que melhoram com o tempo, em vez de alternativas estáticas

Essa transformação vai muito além dos centros de custos internos para reformular fundamentalmente a forma como as organizações interagem com parceiros externos e prestadores de serviços. Ao aplicar preços baseados em resultados às colaborações de parceiros, as organizações podem promover melhorias de qualidade a longo prazo e reduzir custos, ao mesmo tempo em que enfatizam indiretamente a modernização da IA.

As organizações podem experimentar rapidamente, medir o desempenho com clareza e escalar com base no valor comercial real gerado, em vez dos compromissos tradicionais de recursos fixos. Essa abordagem permite o seguinte:

- Evolução do relacionamento com fornecedores — Os parceiros investem no sucesso do cliente em vez de apenas na prestação de serviços.
- Métricas de resultados padronizadas — Simplifique os processos de aquisição em vários fornecedores.

- Capacidade de resposta do mercado — adapte-se rapidamente às mudanças nas condições do mercado e às necessidades dos clientes.
- Vantagem competitiva — utilização superior de recursos e recursos operacionais aprimorados.
- Parcerias orientadas pela qualidade — A colaboração de longo prazo se concentra na melhoria contínua e em resultados mensuráveis.

Usando AWS Marketplace como pay-per-outcome habilitador

O principal facilitador dessa transformação é [AWS Marketplace](#), que serve como um veículo de transação para trabalho agente e preços baseados em resultados. Ele fornece acesso a centenas de agentes de IA pré-criados e soluções agentes com modelos de preços transparentes e baseados no uso. Ele pode ajudar a eliminar os custos iniciais de licenciamento, reduzir a complexidade da implementação e permitir que as organizações se concentrem em sistemas capazes de aprender que se adaptam e melhoram com o tempo, em vez de alternativas estáticas

O uso AWS Marketplace pode fornecer os seguintes benefícios:

- Experimentação rápida — teste várias soluções sem investimento significativo de capital
- Preços transparentes — custos baseados no uso com atribuição clara aos resultados comerciais
- Soluções comprovadas — Acesso a agentes testados em campo de fornecedores experientes
- Integração integrada — Conectividade perfeita com os existentes Serviços da AWS
- Mitigação de riscos — Capacidade de mudar de fornecedor com base no desempenho
- Acesso à capacidade de aprendizagem — Disponibilidade de sistemas adaptativos sem custos internos de desenvolvimento

Essa abordagem permite que as organizações comparem várias opções com base na entrega de resultados e nos recursos de aprendizado, em vez de listas de recursos. Também pode ajudá-lo a estabelecer critérios de sucesso e metodologias de medição claros e a negociar preços baseados em resultados vinculados aos resultados comerciais e à melhoria do sistema. Ao financiar a transformação agente da IA por meio de modelos baseados em resultados, as organizações podem modernizar seus processos continuamente, pagando apenas por melhorias mensuráveis e resultados bem-sucedidos.

Estudo de caso: comparação dos custos de IA humana e agente para operações de recrutamento

As operações de recrutamento fornecem um estudo de caso convincente para avaliar as compensações econômicas entre sistemas de IA humanos e agentes, mas o cálculo do ROI depende criticamente de sua linha de base operacional atual. As organizações que avaliam investimentos em IA em agências geralmente fazem uma pergunta fundamental: “E se simplesmente otimizássemos nossos processos humanos existentes?” Para abordar isso diretamente, esta análise apresenta dois cenários distintos que abrangem a faixa de eficiência operacional humana.

O [cenário A](#) modela o curriculum vitae (CV) de 45 minutos ou os tempos de triagem do currículo. O [cenário B](#) demonstra operações humanas otimizadas em 15 minutos por aplicação, o que representa uma melhoria de 66% na eficiência. Por exemplo, essa melhoria pode ser alcançada por meio de processos simplificados, recrutadores experientes ou ferramentas especializadas.

Ao comparar capacidades idênticas do sistema de agentes com essas diferentes linhas de base de desempenho humano, revelamos como a eficiência do processo existente afeta os cálculos de ROI, os cronogramas de equilíbrio e as decisões estratégicas de implementação. Essa abordagem de cenário duplo serve a vários propósitos. Isso evita que as organizações descartem a IA agente ao presumir que a otimização do processo por si só é suficiente. Também ajuda organizações com processos já eficientes a entender sua economia específica. Além disso, esses cenários destacam quando vantagens não financeiras, como disponibilidade e escalabilidade 24 horas por dia, 7 dias por semana, se tornam os principais fatores de decisão. A compreensão dessas dinâmicas econômicas em diferentes linhas de base de eficiência permite que as organizações tomem decisões informadas sobre onde e quando implantar sistemas de IA agentes para obter o máximo impacto nos negócios.

Cenário A: tempo de exibição de 45 minutos

O cenário A representa operações de recrutamento em que recrutadores humanos passam 45 minutos examinando cada currículo. Esse cenário modela um recrutador de nível médio com um custo anual total de \$112.250. Esse recrutador processa as inscrições durante o horário comercial padrão com características típicas de desempenho humano. Por outro lado, o sistema de IA agente requer um investimento inicial de \$23.000 para desenvolvimento, personalização e integração de ATS, e tem um custo operacional mensal mínimo de \$500 para a infraestrutura em nuvem. O agente processa os aplicativos em apenas 5 minutos com disponibilidade 24 horas por dia, 7 dias por

semana, atingindo uma taxa de erro de 2% e uma capacidade mensal superior a 8.600 aplicativos. Essa é uma grande lacuna de eficiência, na qual o agente opera 9 vezes mais rápido por aplicativo e 39 vezes mais capacidade mensal. Esta seção examina a análise da estrutura de custos, métricas operacionais, comparações baseadas em volume e cálculos cumulativos de ROI nos primeiros seis meses de operação.

Estrutura de custo base

A tabela a seguir mostra os custos iniciais de configuração para o cenário A.

Componente	Operações humanas	Sistema de IA agente
Desenvolvimento e personalização de agentes	N/D	\$15.000
Integração do sistema de rastreamento de candidatos (ATS)	N/D	\$5.000
Treinamento e otimização	N/D	\$3.000
Custo total de configuração inicial	\$0	\$23.000

A tabela a seguir mostra os custos fixos anuais para o cenário A.

Componente	Operações humanas	Sistema de IA agente
Salário base	\$65.000	N/D
Benefícios (30%)	\$19.500	N/D
Espaço de trabalho e equipamento	\$12.000	N/D
Supervisão da gestão (15%)	\$9.750	N/D
Treinamento e desenvolvimento	\$6.000	N/D

Custo fixo anual total	\$112.250	N/D
------------------------	-----------	-----

A tabela a seguir mostra os custos operacionais mensais do cenário A.

Componente	Operações humanas	Sistema de IA agente
Computação em nuvem	N/D	200 USD
Armazenamento	N/D	100 USD
Operações do banco de dados	N/D	100 USD
Monitoramento	N/D	100 USD
Custo fixo mensal total	\$9.354	\$500

Métricas operacionais

A tabela a seguir mostra as métricas operacionais para o cenário A.

Métrica	Operações humanas	Sistema de IA agente
Tempo de processamento por aplicativo	45 minutos	5 minutos
Capacidade horária	1.33 aplicativos	12 aplicações
Capacidade diária (24 horas)	10-11 aplicações	288 aplicações
Capacidade mensal	220 aplicações	8.640 aplicações
Custo por aplicação	\$45	\$2,50
Custo por contratação bem-sucedida	\$2.200	\$125
Taxa de erro	5%	2%
Custo de correção de erros	\$90 por erro	\$45 por escalonamento

Análise de custos baseada em volume

A tabela a seguir mostra uma análise de custo baseada em volume para o cenário A. Neste exemplo, o custo do sistema de IA agente inclui custos fixos e custos de configuração amortizados de \$1.917 por mês durante 12 meses.

Volume mensal	Custo humano	Custo do sistema de IA da Agentic	Economia mensal
100 aplicativos	\$4.500	\$750	\$3.750
500 aplicativos	\$22.500	\$2.667	\$19.833
1.000 aplicativos	\$45.000	\$4.917	\$40.083

Análise de ROI

A tabela a seguir mostra uma análise de ROI para o cenário A que se baseia no processamento de 500 aplicativos por mês.

Métrica	Valor
Custo humano mensal	\$22.500
Custo mensal do agente	\$2.667
Economia mensal	\$19.833
Economia anual	\$237.996
Período de equilíbrio	1,16 meses

Comparação cumulativa de custos

A tabela a seguir mostra uma comparação de custos cumulativos para o cenário A nos primeiros seis meses, considerando 500 aplicativos por mês.

Mês	Custo humano	Custo do sistema de IA da Agentic	Economias cumulativas
1	\$22.500	\$25.667	-\$3.167
2	\$45.000	\$28.334	\$16.666
3	\$67.500	\$31.001	\$36.499
4	\$90.000	\$33.668	\$56.332
5	\$112.500	\$36.335	\$76.165
6	\$135.000	\$39.002	\$95.998

Benefícios adicionais do sistema de IA agente

A seguir estão os benefícios adicionais fornecidos pelo sistema de IA agente no Cenário A:

- Escalabilidade — Pode lidar com picos de volume sem custo adicional
- Disponibilidade — operação 24 horas por dia, 7 dias por semana com resposta imediata
- Consistência — Aplicação uniforme de critérios de triagem
- Eficiência de tempo — Reduzida significativamente time-to-hire
- Experiência do usuário — Feedback instantâneo aos candidatos

Cenário B: tempo de exibição de 15 minutos

O cenário B modela operações de recrutamento otimizadas, nas quais recrutadores humanos simplificaram seu processo de seleção para 15 minutos por inscrição. Isso representa uma melhoria de eficiência de 66% em relação ao Cenário A. Esse cenário mantém o mesmo custo anual completo de \$112.250 para um recrutador de nível médio. No entanto, ele demonstra uma produtividade humana significativamente aprimorada, com a capacidade diária aumentando para 32 aplicações durante um turno de 8 horas e a produtividade mensal chegando a 660 aplicações. A eficiência humana aprimorada reduz o custo por aplicativo de \$45 para \$15, diminuindo a lacuna econômica com o sistema de IA agente. No entanto, o agente mantém suas vantagens estruturais: tempo de processamento de 5 minutos, disponibilidade 24 horas por dia, 7 dias por semana, permitindo

288 aplicações diárias, uma taxa de erro menor de 2% em comparação com os 5% humanos e capacidade mensal superior a 8.600 aplicações. Embora essa melhoria de eficiência estenda o período de equilíbrio de 1,16 meses para 4,76 meses e reduza a economia mensal de \$19.833 para \$4.833, a análise revela que os sistemas de agentes permanecem economicamente viáveis mesmo quando competem com operações humanas altamente otimizadas — uma visão crítica para organizações avaliarem se seus níveis atuais de eficiência de processo justificam o investimento em IA agente.

Estrutura de custo base

A tabela a seguir mostra os custos fixos anuais para o cenário B.

Componente	Operações humanas	Sistema de IA agente
Salário base	\$65.000	N/D
Benefícios (30%)	\$19.500	N/D
Espaço de trabalho e equipamento	\$12.000	N/D
Supervisão de gestão (15%)	\$9.750	N/D
Treinamento e desenvolvimento	\$6.000	N/D
Custo fixo anual total	\$112.250	N/D

A tabela a seguir mostra os custos de implementação do cenário B.

Componente	Operações humanas	Sistema de IA agente
Configuração inicial do	N/D	\$23.000
Custos fixos mensais	\$9.354	\$500

Métricas operacionais

A tabela a seguir mostra as métricas operacionais para o cenário B.

Métrica	Operações humanas	Sistema de IA agente
Tempo de processamento por aplicativo	15 minutos	5 minutos
Capacidade horária	4 aplicações	12 aplicações
Capacidade diária (turno de 8 horas)	32 aplicativos	288 aplicações
Capacidade mensal	660 aplicações	8.640 aplicações
Custo por aplicação	\$15	\$2,50
Custo por contratação bem-sucedida	\$2.200	\$125
Taxa de erro	5%	2%
Custo de correção de erros	\$30 por erro	\$45 por escalonamento

Análise de custos baseada em volume

A tabela a seguir mostra uma análise de custo baseada em volume para o cenário B. Neste exemplo, o custo do sistema de IA agente inclui custos fixos e custos de configuração amortizados de USD 1.917 por mês durante 12 meses.

Volume mensal	Custo humano	Custo do sistema de IA da Agentic	Economia mensal
100 aplicativos	\$1.500	\$750	\$750
500 aplicativos	\$7.500	\$2.667	\$4.833
1.000 aplicativos	\$15.000	\$4.917	\$10.083

Análise de ROI

A tabela a seguir mostra uma análise de ROI para o cenário B que se baseia no processamento de 500 aplicativos por mês.

Métrica	Valor
Custo humano mensal	\$7.500
Custo mensal do sistema de IA da agência	\$2.667
Economia mensal	\$4.833
Economia anual	\$57.996
Período de equilíbrio	4,76 meses

Comparação cumulativa de custos

A tabela a seguir mostra uma comparação de custos cumulativos para o cenário B nos primeiros seis meses, considerando 500 aplicativos por mês.

Mês	Custo humano	Custo do sistema de IA da Agentic	Economias cumulativas
1	\$7.500	\$25.667	-\$18.167
2	\$15.000	\$28.334	-\$13.334
3	\$22.500	\$31.001	-\$8.501
4	\$30.000	\$33.668	-\$3.668
5	\$37.500	\$36.335	\$1.165
6	\$45.000	\$39.002	\$5.998

Comparando custos e benefícios para cada cenário

Métrica	Cenário A	Cenário B	Impacto
Tempo de exibição	45 minutos	15 minutos	66% de melhoria
Capacidade diária	10 a 11 aplicativos	32 aplicativos	Aumento de 200%
Custo por aplicação	\$45	\$15	Redução de 66%
Economia mensal (500 aplicativos)	\$19.833	\$4.833	Redução de 76%
Período de equilíbrio	1,16 meses	4,76 meses	310% mais

O cenário B demonstra ganhos significativos de eficiência nas operações humanas, com melhorias no tempo de processamento que aumentam a capacidade sem número adicional de funcionários e reduzem substancialmente o custo por aplicativo. No entanto, o impacto financeiro revela uma imagem mais sutil: embora o ROI permaneça positivo, as organizações enfrentam um período de equilíbrio prolongado e economias mensais reduzidas em comparação com o Cenário A. Esses resultados destacam os fatores críticos de decisão para a implementação — o sistema de agentes permanece financeiramente viável mesmo com operações humanas otimizadas, mas as organizações devem adotar uma perspectiva de investimento de longo prazo e considerar cuidadosamente as flutuações de volume e as necessidades de escalabilidade ao avaliar os prazos de implantação e os retornos esperados.

No entanto, o sistema de IA agente ainda mantém vantagens operacionais críticas que vão além da pura economia de custos. Ele fornece disponibilidade 24 horas por dia, 7 dias por semana, para engajamento imediato do candidato, independentemente dos fusos horários ou do horário comercial. Ele oferece qualidade de peneiramento consistente aplicando critérios uniformes a cada aplicação e é dimensionado para lidar com picos de volume sem incorrer em custos adicionais. Ele oferece uma resposta imediata ao candidato que aprimora a marca do empregador e a experiência do candidato, e opera com o fator de fadiga zero, o que garante o mesmo desempenho de alta qualidade na primeira inscrição e na milésima.

Os erros humanos geralmente resultam de fadiga, distração ou lacunas de conhecimento e geralmente envolvem falhas de comunicação ou informações incorretas. Os erros do sistema de IA

da Agentic geralmente resultam de casos extremos, entradas ambíguas ou limitações de dados de treinamento. Esses erros tendem a ser de natureza mais consistente.

As métricas de qualidade e experiência revelam claras compensações entre as capacidades humanas e dos agentes:

- Satisfação do cliente — Os humanos se destacam em empatia e resolução de problemas complexos, e os agentes fornecem informações consistentes e precisas para consultas rotineiras.
- Tempo de resposta — O tempo de resposta favorece os agentes com disponibilidade imediata 24 horas por dia, 7 dias por semana. Os humanos fornecem suporte em horário comercial com possíveis atrasos nas filas.
- Consistência — os agentes fornecem respostas idênticas a consultas semelhantes. Os humanos podem variar na abordagem e na aplicação do conhecimento.
- Tratamento da escalada — questões complexas que exigem julgamento, criatividade ou inteligência emocional continuam sendo pontos fortes humanos.

Conclusão e atributos

A economia dos sistemas humanos em comparação com os sistemas de IA agentes representa mais do que uma decisão tecnológica. Ela reflete uma transformação fundamental na forma como as organizações criam valor, gerenciam riscos e alcançam vantagem competitiva. O sucesso requer avaliação sistemática das características do trabalho, medição abrangente dos resultados (incluindo fatores de risco) e escalonamento estratégico com base em resultados comprovados.

O [estado da adoção da IA corporativa](#) (relatório ISG 2025) revela que a maioria das implementações de IA falha devido a lacunas de aprendizado — sistemas que não conseguem se adaptar, lembrar o contexto ou melhorar com o tempo. Organizações que alcançam o sucesso se concentram em sistemas capazes de aprender que se integrem profundamente aos fluxos de trabalho e demonstrem melhoria contínua por meio de feedback humano e experiência operacional.

Organizações que entendem esses princípios — começando com trabalhos adequados, decompondo trabalhos em tarefas, medindo tudo, inclusive o impacto do risco, e escalando o que funciona — obterão vantagem competitiva sustentável por meio da utilização ideal de recursos e da automação focada em resultados que cresce com o sucesso dos negócios.

O futuro pertence às organizações que podem combinar de forma inteligente a experiência humana com os recursos de inteligência artificial. Isso cria modelos híbridos que oferecem resultados superiores, mantendo a flexibilidade, a capacidade de aprendizado e os benefícios colaborativos necessários para condições dinâmicas de mercado.

Recursos

Os recursos a seguir podem ajudá-lo a planejar, projetar e implementar sistemas de IA agentes em AWS:

- [Criação de arquiteturas sem servidor para IA agente ativada \(orientação prescritiva\) AWS](#) AWS
- [Operacionalizando a IA agente ativada \(orientação prescritiva AWS\)](#) AWS
- [Padrões de inteligência artificial e fluxos de trabalho em AWS](#) (orientação AWS prescritiva)
- [Inteligência artificial](#) (orientação AWS prescritiva)
- [Hub de Otimização de Custos da AWS](#) (AWS service (Serviço da AWS))
- [Documentação do Amazon Bedrock](#) (AWS service (Serviço da AWS))
- [Pilar de otimização de custos](#) (AWS Well-Architected Framework)

- [Agentes e soluções de IA](#) (AWS Marketplace)

Histórico do documento

A tabela a seguir descreve alterações significativas feitas neste guia. Se desejar receber notificações sobre futuras atualizações, inscreva-se em um [feed RSS](#).

Alteração	Descrição	Data
Publicação inicial	—	28 de janeiro de 2026

AWS Glossário de orientação prescritiva

A seguir estão os termos comumente usados em estratégias, guias e padrões fornecidos pela Orientação AWS Prescritiva. Para sugerir entradas, use o link Fornecer feedback no final do glossário.

Números

7 Rs

Sete estratégias comuns de migração para mover aplicações para a nuvem. Essas estratégias baseiam-se nos 5 Rs identificados pela Gartner em 2011 e consistem em:

- Refatorar/rearquitetar: mova uma aplicação e modifique sua arquitetura aproveitando ao máximo os recursos nativos de nuvem para melhorar a agilidade, a performance e a escalabilidade. Isso normalmente envolve a portabilidade do sistema operacional e do banco de dados. Exemplo: migrar seu banco de dados Oracle on-premises para o Amazon Aurora Edição Compatível com PostgreSQL.
- Redefinir a plataforma (mover e redefinir [mover e redefinir (lift-and-reshape)]): mova uma aplicação para a nuvem e introduza algum nível de otimização a fim de aproveitar os recursos da nuvem. Exemplo: migrar seu banco de dados Oracle on-premises para o Amazon Relational Database Service (Amazon RDS) para Oracle na Nuvem AWS.
- Recomprar (drop and shop): mude para um produto diferente, normalmente migrando de uma licença tradicional para um modelo SaaS. Exemplo: migrar seu sistema de gerenciamento de relacionamento com o cliente (CRM) para o Salesforce.com.
- Redefinir a hospedagem (mover sem alterações [lift-and-shift])mover uma aplicação para a nuvem sem fazer nenhuma alteração a fim de aproveitar os recursos da nuvem. Exemplo: migrar seu banco de dados Oracle on-premises para o Oracle em uma instância do EC2 na Nuvem AWS.
- Realocar (mover o hipervisor sem alterações [hypervisor-level lift-and-shift]): mover a infraestrutura para a nuvem sem comprar novo hardware, reescrever aplicações ou modificar suas operações existentes. Você migra servidores de uma plataforma on-premises para um serviço de nuvem para a mesma plataforma. Exemplo: Migrar um Microsoft Hyper-V aplicativo para o. AWS
- Reter (revisitar): mantenha as aplicações em seu ambiente de origem. Isso pode incluir aplicações que exigem grande refatoração, e você deseja adiar esse trabalho para um

momento posterior, e aplicações antigas que você deseja manter porque não há justificativa comercial para migrá-las.

- Retirar: desative ou remova aplicações que não são mais necessárias em seu ambiente de origem.

A

ABAC

Consulte [controle de acesso baseado em atributo](#).

serviços abstraídos

Veja [serviços gerenciados](#).

ACID

Veja [atomicidade, consistência, isolamento, durabilidade](#).

migração ativa-ativa

Um método de migração de banco de dados no qual os bancos de dados de origem e de destino são mantidos em sincronia (por meio de uma ferramenta de replicação bidirecional ou operações de gravação dupla), e ambos os bancos de dados lidam com transações de aplicações conectadas durante a migração. Esse método oferece suporte à migração em lotes pequenos e controlados, em vez de exigir uma substituição única. É mais flexível, mas exige mais trabalho do que a [migração ativa-passiva](#).

migração ativa-passiva

Um método de migração de banco de dados em que os bancos de dados de origem e de destino são mantidos em sincronia, mas somente o banco de dados de origem manipula as transações das aplicações conectadas, enquanto os dados são replicados no banco de dados de destino. O banco de dados de destino não aceita nenhuma transação durante a migração.

AGGREGATE FUNCTION

Uma função SQL que opera em um grupo de linhas e calcula um único valor de retorno para o grupo. Exemplos de funções agregadas incluem SUM e MAX.

AI

Veja [inteligência artificial](#).

AIOps

Veja [operações de inteligência artificial](#).

anonimização

O processo de excluir permanentemente informações pessoais em um conjunto de dados. A anonimização pode ajudar a proteger a privacidade pessoal. Dados anônimos não são mais considerados dados pessoais.

antipadrões

Uma solução frequentemente usada para um problema recorrente em que a solução é contraproducente, ineficaz ou menos eficaz do que uma alternativa.

controle de aplicações

Uma abordagem de segurança que permite o uso somente de aplicações aprovadas para ajudar a proteger um sistema contra malware.

portfólio de aplicações

Uma coleção de informações detalhadas sobre cada aplicação usada por uma organização, incluindo o custo para criar e manter a aplicação e seu valor comercial. Essas informações são fundamentais para [o processo de descoberta e análise de portfólio](#) e ajudam a identificar e priorizar as aplicações a serem migradas, modernizadas e otimizadas.

inteligência artificial (IA)

O campo da ciência da computação que se dedica ao uso de tecnologias de computação para desempenhar funções cognitivas normalmente associadas aos humanos, como aprender, resolver problemas e reconhecer padrões. Para obter mais informações, consulte [O que é inteligência artificial?](#)

operações de inteligência artificial (AIOps)

O processo de usar técnicas de machine learning para resolver problemas operacionais, reduzir incidentes operacionais e intervenção humana e aumentar a qualidade do serviço. Para obter mais informações sobre como AIOps é usado na estratégia de AWS migração, consulte o [guia de integração de operações](#).

criptografia assimétrica

Um algoritmo de criptografia que usa um par de chaves, uma chave pública para criptografia e uma chave privada para descryptografia. É possível compartilhar a chave pública porque ela não é usada na descryptografia, mas o acesso à chave privada deve ser altamente restrito.

atomicidade, consistência, isolamento, durabilidade (ACID)

Um conjunto de propriedades de software que garantem a validade dos dados e a confiabilidade operacional de um banco de dados, mesmo no caso de erros, falhas de energia ou outros problemas.

controle de acesso por atributo (ABAC)

A prática de criar permissões minuciosas com base nos atributos do usuário, como departamento, cargo e nome da equipe. Para obter mais informações, consulte [ABAC AWS](#) na documentação AWS Identity and Access Management (IAM).

fonte de dados autorizada

Um local onde você armazena a versão principal dos dados, que é considerada a fonte de informações mais confiável. Você pode copiar dados da fonte de dados autorizada para outros locais com o objetivo de processar ou modificar os dados, como anonimizá-los, redigi-los ou pseudonimizá-los.

Zona de disponibilidade

Um local distinto dentro de um Região da AWS que está isolado de falhas em outras zonas de disponibilidade e fornece conectividade de rede barata e de baixa latência a outras zonas de disponibilidade na mesma região.

AWS Estrutura de adoção da nuvem (AWS CAF)

Uma estrutura de diretrizes e melhores práticas AWS para ajudar as organizações a desenvolver um plano eficiente e eficaz para migrar com sucesso para a nuvem. AWS O CAF organiza a orientação em seis áreas de foco chamadas perspectivas: negócios, pessoas, governança, plataforma, segurança e operações. As perspectivas de negócios, pessoas e governança têm como foco habilidades e processos de negócios; as perspectivas de plataforma, segurança e operações concentram-se em habilidades e processos técnicos. Por exemplo, a perspectiva das pessoas tem como alvo as partes interessadas que lidam com recursos humanos (RH), funções de pessoal e gerenciamento de pessoal. Nessa perspectiva, o AWS CAF fornece orientação para desenvolvimento, treinamento e comunicação de pessoas para ajudar a preparar a organização para a adoção bem-sucedida da nuvem. Para obter mais informações, consulte o [site da AWS CAF](#) e o [whitepaper da AWS CAF](#).

AWS Estrutura de qualificação da carga de trabalho (AWS WQF)

Uma ferramenta que avalia as cargas de trabalho de migração do banco de dados, recomenda estratégias de migração e fornece estimativas de trabalho. AWS O WQF está incluído com AWS

Schema Conversion Tool (AWS SCT). Ela analisa esquemas de banco de dados e objetos de código, código de aplicações, dependências e características de performance, além de fornecer relatórios de avaliação.

B

bot malicioso

Um [bot](#) destinado a causar disrupção ou danos a indivíduos ou organizações.

BCP

Veja [planejamento de continuidade de negócios](#)

gráfico de comportamento

Uma visualização unificada e interativa do comportamento e das interações de recursos ao longo do tempo. É possível usar um gráfico de comportamento com o Amazon Detective para examinar tentativas de login malsucedidas, chamadas de API suspeitas e ações similares. Para obter mais informações, consulte [Dados em um gráfico de comportamento](#) na documentação do Detective.

sistema big-endian

Um sistema que armazena o byte mais significativo antes. Veja também [endianness](#).

classificação binária

Um processo que prevê um resultado binário (uma de duas classes possíveis). Por exemplo, seu modelo de ML pode precisar prever problemas como “Este e-mail é ou não é spam?” ou “Este produto é um livro ou um carro?”

filtro de bloom

Uma estrutura de dados probabilística e eficiente em termos de memória que é usada para testar se um elemento é membro de um conjunto.

blue/green deployment (implantação azul/verde)

Uma estratégia de implantação em que você cria dois ambientes separados, mas idênticos. Você executa a versão atual da aplicação em um ambiente (azul) e a nova versão da aplicação no outro ambiente (verde). Essa estratégia ajuda você a reverter rapidamente com o mínimo de impacto.

bot

Uma aplicação de software que executa tarefas automatizadas na internet e simula a atividade ou interação humana. Alguns bots são úteis ou benéficos, como crawlers da web que indexam informações na internet. Outros bots, conhecidos como bots maliciosos, têm como objetivo causar interrupção ou danos a indivíduos ou organizações.

botnet

Redes de [bots](#) infectadas por [malware](#) e sob o controle de uma única parte, conhecidas como bot herder ou operador de bots. Os botnets são o mecanismo mais conhecido para escalar bots e seu impacto.

ramo

Uma área contida de um repositório de código. A primeira ramificação criada em um repositório é a ramificação principal. Você pode criar uma nova ramificação a partir de uma ramificação existente e, em seguida, desenvolver recursos ou corrigir bugs na nova ramificação. Uma ramificação que você cria para gerar um recurso é comumente chamada de ramificação de recurso. Quando o recurso estiver pronto para lançamento, você mesclará a ramificação do recurso de volta com a ramificação principal. Para obter mais informações, consulte [Sobre filiais](#) (GitHub documentação).

Acesso de emergência

Em circunstâncias excepcionais e por meio de um processo aprovado, um meio rápido para um usuário obter acesso a um Conta da AWS que ele normalmente não tem permissão para acessar. Para obter mais informações, consulte o indicador [Implement break-glass procedures](#) nas orientações do AWS Well-Architected.

estratégia brownfield

A infraestrutura existente em seu ambiente. Ao adotar uma estratégia brownfield para uma arquitetura de sistema, você desenvolve a arquitetura de acordo com as restrições dos sistemas e da infraestrutura atuais. Se estiver expandindo a infraestrutura existente, poderá combinar as estratégias brownfield e [greenfield](#).

cache do buffer

A área da memória em que os dados acessados com mais frequência são armazenados.

capacidade de negócios

O que uma empresa faz para gerar valor (por exemplo, vendas, atendimento ao cliente ou marketing). As arquiteturas de microsserviços e as decisões de desenvolvimento podem

ser orientadas por recursos de negócios. Para obter mais informações, consulte a seção [Organizados de acordo com as capacidades de negócios](#) do whitepaper [Executar microsserviços containerizados na AWS](#).

planejamento de continuidade de negócios (BCP)

Um plano que aborda o impacto potencial de um evento disruptivo, como uma migração em grande escala, nas operações e permite que uma empresa retome as operações rapidamente.

C

CAF

Veja [AWS Cloud Adoption Framework](#).

implantação canário

O lançamento lento e incremental de uma versão para usuários finais. Quando estiver confiante, você implanta a nova versão e substitui a versão atual por completo.

CCoE

Veja [Centro de Excelência da Nuvem](#).

CDC

Veja [captura de dados de alteração](#).

captura de dados de alterações (CDC)

O processo de rastrear alterações em uma fonte de dados, como uma tabela de banco de dados, e registrar metadados sobre a alteração. É possível usar o CDC para várias finalidades, como auditar ou replicar alterações em um sistema de destino para manter a sincronização.

engenharia do caos

Introduzir intencionalmente falhas ou eventos disruptivos para testar a resiliência de um sistema. Você pode usar [AWS Fault Injection Service \(AWS FIS\)](#) para realizar experimentos que estressam suas AWS cargas de trabalho e avaliar sua resposta.

CI/CD

Veja [integração e entrega contínuas](#).

classificação

Um processo de categorização que ajuda a gerar previsões. Os modelos de ML para problemas de classificação predizem um valor discreto. Os valores discretos são sempre diferentes uns dos outros. Por exemplo, um modelo pode precisar avaliar se há ou não um carro em uma imagem.

criptografia no lado do cliente

Criptografia de dados localmente, antes que o alvo os AWS service (Serviço da AWS) receba.

Centro de excelência em nuvem (CCoE)

Uma equipe multidisciplinar que impulsiona os esforços de adoção da nuvem em toda a organização, incluindo o desenvolvimento de práticas recomendadas de nuvem, a mobilização de recursos, o estabelecimento de cronogramas de migração e a liderança da organização em transformações em grande escala. Para obter mais informações, consulte as [publicações CCo E](#) no blog de estratégia Nuvem AWS corporativa.

computação em nuvem

A tecnologia de nuvem normalmente usada para armazenamento de dados remoto e gerenciamento de dispositivos de IoT. A computação em nuvem é normalmente conectada à tecnologia de [computação de borda](#).

modelo operacional em nuvem

Em uma organização de TI, o modelo operacional usado para criar, amadurecer e otimizar um ou mais ambientes de nuvem. Para obter mais informações, consulte [Criar seu modelo operacional de nuvem](#).

estágios de adoção da nuvem

As quatro fases pelas quais as organizações normalmente passam ao migrar para a Nuvem AWS:

- Projeto: executar alguns projetos relacionados à nuvem para fins de prova de conceito e aprendizado
- Fundação — Fazer investimentos fundamentais para escalar sua adoção da nuvem (por exemplo, criar uma landing zone, definir um CCo E, estabelecer um modelo de operações)
- Migração: migrar aplicações individuais
- Reinvenção: otimizar produtos e serviços e inovar na nuvem

Esses estágios foram definidos por Stephen Orban na postagem do blog [The Journey Toward Cloud-First & the Stages of Adoption](#) no blog de estratégia Nuvem AWS empresarial. Para obter

informações sobre como eles se relacionam com a estratégia de AWS migração, consulte o [guia de preparação para migração](#).

CMDB

Veja [banco de dados de gerenciamento de configuração](#).

repositório de código

Um local onde o código-fonte e outros ativos, como documentação, amostras e scripts, são armazenados e atualizados por meio de processos de controle de versão. Os repositórios de nuvem comuns incluem o GitHub ou o Bitbucket Cloud. Cada versão do código é chamada de ramificação. Em uma estrutura de microsserviços, cada repositório é dedicado a uma única peça de funcionalidade. Um único pipeline de CI/CD pode usar vários repositórios.

cache frio

Um cache de buffer que está vazio, não está bem preenchido ou contém dados obsoletos ou irrelevantes. Isso afeta a performance porque a instância do banco de dados deve ler da memória principal ou do disco, um processo que é mais lento do que a leitura do cache do buffer.

dados frios

Dados que raramente são acessados e geralmente são históricos. Ao consultar esse tipo de dados, consultas lentas geralmente são aceitáveis. Mover esses dados para níveis ou classes de armazenamento de baixo desempenho e menos caros pode reduzir os custos.

visão computacional (CV)

Um campo de [IA](#) que usa machine learning para analisar e extrair informações de formatos visuais, como vídeos e imagens digitais. Por exemplo, a Amazon SageMaker AI fornece algoritmos de processamento de imagem para CV.

desvio de configuração

Em uma workload, uma alteração de configuração em relação ao estado esperado. Isso pode fazer com que a workload se torne incompatível e, normalmente, é gradual e não intencional.

banco de dados de gerenciamento de configuração (CMDB)

Um repositório que armazena e gerencia informações sobre um banco de dados e seu ambiente de TI, incluindo componentes de hardware e software e suas configurações. Normalmente, os dados de um CMDB são usados no estágio de descoberta e análise do portfólio da migração.

pacote de conformidade

Um conjunto de AWS Config regras e ações de remediação que você pode montar para personalizar suas verificações de conformidade e segurança. Você pode implantar um pacote de conformidade como uma entidade única em uma Conta da AWS região ou em uma organização usando um modelo YAML. Para obter mais informações, consulte [Pacotes de conformidade na documentação](#). AWS Config

integração contínua e entrega contínua (CI/CD)

O processo de automatizar os estágios de origem, criação, teste, preparação e produção do processo de lançamento do software. CI/CD é comumente descrito como um pipeline. CI/CD pode ajudá-lo a automatizar processos, melhorar a produtividade, melhorar a qualidade do código e entregar com mais rapidez. Para obter mais informações, consulte [Benefícios da entrega contínua](#). CD também pode significar implantação contínua. Para obter mais informações, consulte [Entrega contínua versus implantação contínua](#).

CV

Veja [visão computacional](#).

D

dados em repouso

Dados estacionários em sua rede, por exemplo, dados que estão em um armazenamento.

classificação de dados

Um processo para identificar e categorizar os dados em sua rede com base em criticalidade e confidencialidade. É um componente crítico de qualquer estratégia de gerenciamento de riscos de segurança cibernética, pois ajuda a determinar os controles adequados de proteção e retenção para os dados. A classificação de dados é um componente do pilar de segurança no AWS Well-Architected Framework. Para obter mais informações, consulte [Classificação de dados](#).

desvio de dados

Uma variação significativa entre os dados de produção e os dados usados para treinar um modelo de ML ou uma alteração significativa nos dados de entrada ao longo do tempo. O desvio de dados pode reduzir a qualidade geral, a precisão e a imparcialidade das previsões do modelo de ML.

dados em trânsito

Dados que estão se movendo ativamente pela sua rede, como entre os recursos da rede.

data mesh

Um framework de arquitetura que fornece propriedade de dados distribuída e descentralizada com gerenciamento e governança centralizados.

minimização de dados

O princípio de coletar e processar apenas os dados estritamente necessários. Praticar a minimização de dados no Nuvem AWS pode reduzir os riscos de privacidade, os custos e a pegada de carbono de sua análise.

perímetro de dados

Um conjunto de proteções preventivas em seu AWS ambiente que ajudam a garantir que somente identidades confiáveis acessem recursos confiáveis das redes esperadas. Para obter mais informações, consulte [Construindo um perímetro de dados em AWS](#)

pré-processamento de dados

A transformação de dados brutos em um formato que seja facilmente analisado por seu modelo de ML. O pré-processamento de dados pode significar a remoção de determinadas colunas ou linhas e o tratamento de valores ausentes, inconsistentes ou duplicados.

proveniência dos dados

O processo de rastrear a origem e o histórico dos dados ao longo de seu ciclo de vida, por exemplo, como os dados foram gerados, transmitidos e armazenados.

titular dos dados

Um indivíduo cujos dados estão sendo coletados e processados.

data warehouse

Um sistema de gerenciamento de dados compatível com business intelligence, como analytics. Os data warehouses geralmente contêm grandes quantidades de dados históricos e geralmente são usados para consultas e análises.

linguagem de definição de dados (DDL)

Instruções ou comandos para criar ou modificar a estrutura de tabelas e objetos em um banco de dados.

linguagem de manipulação de dados (DML)

Instruções ou comandos para modificar (inserir, atualizar e excluir) informações em um banco de dados.

DDL

Veja [linguagem de definição de banco de dados](#).

deep ensemble

A combinação de vários modelos de aprendizado profundo para gerar previsões. Os deep ensembles podem ser usados para produzir uma previsão mais precisa ou para estimar a incerteza nas previsões.

Aprendizado profundo

Um subcampo do ML que usa várias camadas de redes neurais artificiais para identificar o mapeamento entre os dados de entrada e as variáveis-alvo de interesse.

defense-in-depth

Uma abordagem de segurança da informação na qual uma série de mecanismos e controles de segurança são cuidadosamente distribuídos por toda a rede de computadores para proteger a confidencialidade, a integridade e a disponibilidade da rede e dos dados nela contidos. Ao adotar essa estratégia AWS, você adiciona vários controles em diferentes camadas da AWS Organizations estrutura para ajudar a proteger os recursos. Por exemplo, uma defense-in-depth abordagem pode combinar autenticação multifatorial, segmentação de rede e criptografia.

administrador delegado

Em AWS Organizations, um serviço compatível pode registrar uma conta de AWS membro para administrar as contas da organização e gerenciar as permissões desse serviço. Essa conta é chamada de administrador delegado para esse serviço. Para obter mais informações e uma lista de serviços compatíveis, consulte [Serviços que funcionam com o AWS Organizations](#) na documentação do AWS Organizations .

implantação

O processo de criar uma aplicação, novos recursos ou correções de código disponíveis no ambiente de destino. A implantação envolve a implementação de mudanças em uma base de código e, em seguida, a criação e execução dessa base de código nos ambientes da aplicação

ambiente de desenvolvimento

Veja [ambiente](#).

controle detectivo

Um controle de segurança projetado para detectar, registrar e alertar após a ocorrência de um evento. Esses controles são uma segunda linha de defesa, alertando você sobre eventos de segurança que contornaram os controles preventivos em vigor. Para obter mais informações, consulte [Controles detectivos](#) em Como implementar controles de segurança na AWS.

mapeamento do fluxo de valor de desenvolvimento (DVSM)

Um processo usado para identificar e priorizar restrições que afetam negativamente a velocidade e a qualidade em um ciclo de vida de desenvolvimento de software. O DVSM estende o processo de mapeamento do fluxo de valor originalmente projetado para práticas de manufatura enxuta. Ele se concentra nas etapas e equipes necessárias para criar e movimentar valor por meio do processo de desenvolvimento de software.

gêmeo digital

Uma representação virtual de um sistema real, como um prédio, fábrica, equipamento industrial ou linha de produção. Os gêmeos digitais oferecem suporte à manutenção preditiva, ao monitoramento remoto e à otimização da produção.

tabela de dimensões

Em um [esquema em estrela](#), uma tabela menor que contém atributos de dados sobre dados quantitativos em uma tabela de fatos. Os atributos da tabela de dimensões geralmente são campos de texto ou números discretos que se comportam como texto. Esses atributos normalmente são usados para restringir consultas, filtrar e rotular conjuntos de resultados.

desastre

Um evento que impede que uma workload ou sistema cumpra seus objetivos de negócios em seu local principal de implantação. Esses eventos podem ser desastres naturais, falhas técnicas ou o resultado de ações humanas, como configuração incorreta não intencional ou ataque de malware.

Recuperação de desastres (RD)

A estratégia e o processo que você usa para minimizar o tempo de inatividade e a perda de dados causados por um [desastre](#). Para obter mais informações, consulte [Recuperação de desastres de cargas de trabalho em AWS: Recuperação na nuvem no AWS Well-Architected Framework](#).

DML

Veja [linguagem de manipulação de banco de dados](#).

design orientado por domínio

Uma abordagem ao desenvolvimento de um sistema de software complexo conectando seus componentes aos domínios em evolução, ou principais metas de negócios, atendidos por cada componente. Esse conceito foi introduzido por Eric Evans em seu livro, *Design orientado por domínio: lidando com a complexidade no coração do software* (Boston: Addison-Wesley Professional, 2003). Para obter informações sobre como usar o design orientado por domínio com o padrão strangler fig, consulte [Modernizar incrementalmente os serviços web herdados do Microsoft ASP.NET \(ASMX\) usando contêineres e o Amazon API Gateway](#).

DR

Veja [recuperação de desastres](#).

Deteção da oscilação

Rastreamento de desvios de uma configuração de linha de base. Por exemplo, você pode usar AWS CloudFormation para [detectar desvios nos recursos do sistema](#) ou AWS Control Tower para [detectar mudanças em seu landing zone](#) que possam afetar a conformidade com os requisitos de governança.

DVSM

Veja [mapeamento do fluxo de valor de desenvolvimento](#).

E

EDA

Veja [análise exploratória de dados](#).

EDI

Veja [intercâmbio eletrônico de dados](#).

computação de borda

A tecnologia que aumenta o poder computacional de dispositivos inteligentes nas bordas de uma rede de IoT. Quando comparada com a [computação em nuvem](#), a computação de borda pode reduzir a latência da comunicação e melhorar o tempo de resposta.

intercâmbio eletrônico de dados (EDI)

A troca automatizada de documentos comerciais entre organizações. Para obter mais informações, consulte [O que é EDI \(Intercâmbio eletrônico de dados\)?](#).

criptografia

Um processo de computação que transforma dados de texto simples, legíveis por humanos, em texto cifrado.

chave de criptografia

Uma sequência criptográfica de bits aleatórios que é gerada por um algoritmo de criptografia. As chaves podem variar em tamanho, e cada chave foi projetada para ser imprevisível e exclusiva.

endianismo

A ordem na qual os bytes são armazenados na memória do computador. Os sistemas big-endian armazenam o byte mais significativo antes. Os sistemas little-endian armazenam o byte menos significativo antes.

endpoint

Veja [endpoint de serviço](#).

serviço de endpoint

Um serviço que pode ser hospedado em uma nuvem privada virtual (VPC) para ser compartilhado com outros usuários. Você pode criar um serviço de endpoint com AWS PrivateLink e conceder permissões a outros diretores Contas da AWS ou a AWS Identity and Access Management (IAM). Essas contas ou entidades principais podem se conectar ao serviço de endpoint de maneira privada criando endpoints da VPC de interface. Para obter mais informações, consulte [Criar um serviço de endpoint](#) na documentação do Amazon Virtual Private Cloud (Amazon VPC).

planejamento de recursos empresariais (ERP)

Um sistema que automatiza e gerencia os principais processos de negócios (como contabilidade, [MES](#) e gerenciamento de projetos) para uma empresa.

criptografia envelopada

O processo de criptografar uma chave de criptografia com outra chave de criptografia. Para obter mais informações, consulte [Criptografia de envelope](#) na documentação AWS Key Management Service (AWS KMS).

ambiente

Uma instância de uma aplicação em execução. Estes são tipos comuns de ambientes na computação em nuvem:

- ambiente de desenvolvimento: uma instância de uma aplicação em execução que está disponível somente para a equipe principal responsável pela manutenção da aplicação. Ambientes de desenvolvimento são usados para testar mudanças antes de promovê-las para ambientes superiores. Esse tipo de ambiente às vezes é chamado de ambiente de teste.
- ambientes inferiores: todos os ambientes de desenvolvimento para uma aplicação, como aqueles usados para compilações e testes iniciais.
- ambiente de produção: uma instância de uma aplicação em execução que os usuários finais podem acessar. Em um CI/CD pipeline, o ambiente de produção é o último ambiente de implantação.
- ambientes superiores: todos os ambientes que podem ser acessados por usuários que não sejam a equipe principal de desenvolvimento. Isso pode incluir um ambiente de produção, ambientes de pré-produção e ambientes para testes de aceitação do usuário.

epic

Em metodologias ágeis, categorias funcionais que ajudam a organizar e priorizar seu trabalho. Os epics fornecem uma descrição de alto nível dos requisitos e das tarefas de implementação. Por exemplo, os épicos de segurança AWS da CAF incluem gerenciamento de identidade e acesso, controles de detetive, segurança de infraestrutura, proteção de dados e resposta a incidentes. Para obter mais informações sobre epics na estratégia de migração da AWS, consulte o [guia de implementação do programa](#).

ERP

Veja [planejamento de recursos empresariais](#).

análise exploratória de dados (EDA)

O processo de analisar um conjunto de dados para entender suas principais características. Você coleta ou agrega dados e, em seguida, realiza investigações iniciais para encontrar padrões, detectar anomalias e verificar suposições. O EDA é realizado por meio do cálculo de estatísticas resumidas e da criação de visualizações de dados.

F

tabela de fatos

A tabela central em um [esquema em estrela](#). Ela armazena dados quantitativos sobre as operações comerciais. Normalmente, uma tabela de fatos contém dois tipos de colunas: as que contêm medidas e as que contêm uma chave externa para uma tabela de dimensões.

Antecipar-se à falha

Uma filosofia que usa testes frequentes e incrementais para reduzir o ciclo de vida do desenvolvimento. É uma parte essencial de uma abordagem ágil.

delimitação de isolamento contra falhas

No Nuvem AWS, um limite, como uma zona de disponibilidade, Região da AWS um plano de controle ou um plano de dados, que limita o efeito de uma falha e ajuda a melhorar a resiliência das cargas de trabalho. Para obter mais informações, consulte [AWS Fault Isolation Boundaries](#).

ramificação de recursos

Veja [ramificação](#).

recursos

Os dados de entrada usados para fazer uma previsão. Por exemplo, em um contexto de manufatura, os recursos podem ser imagens capturadas periodicamente na linha de fabricação.

importância do recurso

O quanto um recurso é importante para as previsões de um modelo. Isso geralmente é expresso como uma pontuação numérica que pode ser calculada por meio de várias técnicas, como Shapley Additive Explanations (SHAP) e gradientes integrados. Para obter mais informações, consulte [Interpretabilidade do modelo de aprendizado de máquina com AWS](#).

transformação de recursos

O processo de otimizar dados para o processo de ML, incluindo enriquecer dados com fontes adicionais, escalar valores ou extrair vários conjuntos de informações de um único campo de dados. Isso permite que o modelo de ML se beneficie dos dados. Por exemplo, se a data “2021-05-27 00:15:37” for dividida em “2021”, “maio”, “quinta” e “15”, isso poderá ajudar o algoritmo de aprendizado a aprender padrões diferenciados associados a diferentes componentes de dados.

prompt few shot

Fornecer a um [LLM](#) um pequeno número de exemplos que demonstram a tarefa e o resultado desejado antes de solicitar que ele execute uma tarefa semelhante. Essa técnica é uma aplicação do aprendizado em contexto, em que os modelos aprendem com exemplos (shots) incorporados aos prompts. Prompts few-shot podem ser eficazes para tarefas que exigem formatação, raciocínio ou conhecimento de domínio específicos. Veja também [prompts zero-shot](#).

FGAC

Veja [controle de acesso refinado](#).

Controle de acesso refinado (FGAC)

O uso de várias condições para permitir ou negar uma solicitação de acesso.

migração flash-cut

Um método de migração de banco de dados que usa replicação contínua de dados via [captura de dados de alteração](#) para migrar os dados no menor tempo possível, em vez de usar uma abordagem em fases. O objetivo é reduzir ao mínimo o tempo de inatividade.

FM

Veja [modelo de base](#).

modelo de base (FM)

Uma grande rede neural de aprendizado profundo que vem treinando em grandes conjuntos de dados generalizados e não rotulados. FMs são capazes de realizar uma ampla variedade de tarefas gerais, como entender a linguagem, gerar texto e imagens e conversar em linguagem natural. Para obter mais informações, consulte [O que são modelos de base?](#).

G

IA generativa

Um subconjunto de modelos de [IA](#) que foram treinados em grandes quantidades de dados e que podem usar um simples prompt de texto para criar novos artefatos e conteúdo, como imagens, vídeos, texto e áudio. Para obter mais informações, consulte [O que é IA generativa?](#).

bloqueio geográfico

Veja [restrições geográficas](#).

restrições geográficas (bloqueio geográfico)

Na Amazon CloudFront, uma opção para impedir que usuários em países específicos acessem distribuições de conteúdo. É possível usar uma lista de permissões ou uma lista de bloqueios para especificar países aprovados e banidos. Para obter mais informações, consulte [Restringir a distribuição geográfica do seu conteúdo](#) na CloudFront documentação.

Fluxo de trabalho do GitFlow

Uma abordagem na qual ambientes inferiores e superiores usam ramificações diferentes em um repositório de código-fonte. O fluxo de trabalho do Gitflow é considerado legado, e o [fluxo de trabalho trunk-based](#) é a abordagem moderna e preferencial.

golden image

Um snapshot de um sistema ou software usado como modelo para implantar novas instâncias desse sistema ou software. Por exemplo, na manufatura, uma golden image pode ser usada para provisionar software em vários dispositivos e ajudar a melhorar a velocidade, a escalabilidade e a produtividade nas operações de fabricação de dispositivos.

estratégia greenfield

A ausência de infraestrutura existente em um novo ambiente. Ao adotar uma estratégia greenfield para uma arquitetura de sistema, é possível selecionar todas as novas tecnologias sem a restrição da compatibilidade com a infraestrutura existente, também conhecida como [brownfield](#). Se estiver expandindo a infraestrutura existente, poderá combinar as estratégias brownfield e greenfield.

barreira de proteção

Uma regra de alto nível que ajuda a governar recursos, políticas e conformidade em todas as unidades organizacionais (OUs). Barreiras de proteção preventivas impõem políticas para garantir o alinhamento a padrões de conformidade. Elas são implementadas usando políticas de controle de serviço e limites de permissões do IAM. Barreiras de proteção detectivas detectam violações de políticas e problemas de conformidade e geram alertas para remediação. Eles são implementados usando AWS Config, AWS Security Hub CSPM, Amazon GuardDuty AWS Trusted Advisor, Amazon Inspector e verificações personalizadas AWS Lambda .

H

HA

Veja [alta disponibilidade](#).

migração heterogênea de bancos de dados

Migrar seu banco de dados de origem para um banco de dados de destino que usa um mecanismo de banco de dados diferente (por exemplo, Oracle para Amazon Aurora). A migração heterogênea geralmente faz parte de um esforço de redefinição da arquitetura, e converter

o esquema pode ser uma tarefa complexa. [O AWS fornece o AWS SCT](#) para ajudar nas conversões de esquemas.

alta disponibilidade (HA)

A capacidade de uma workload operar continuamente, sem intervenção, em caso de desafios ou desastres. Os sistemas AH são projetados para realizar o failover automático, oferecer consistentemente desempenho de alta qualidade e lidar com diferentes cargas e falhas com impacto mínimo no desempenho.

modernização de historiador

Uma abordagem usada para modernizar e atualizar os sistemas de tecnologia operacional (OT) para melhor atender às necessidades do setor de manufatura. Um historiador é um tipo de banco de dados usado para coletar e armazenar dados de várias fontes em uma fábrica.

dados de hold-out

Uma parte dos dados históricos rotulados que são retidos de um conjunto de dados usado para treinar um modelo de [machine learning](#). Você pode usar dados de hold-out para avaliar a performance do modelo comparando as previsões do modelo com os dados de retenção.

migração homogênea de bancos de dados

Migrar seu banco de dados de origem para um banco de dados de destino que compartilha o mesmo mecanismo de banco de dados (por exemplo, Microsoft SQL Server para Amazon RDS para SQL Server). A migração homogênea geralmente faz parte de um esforço de redefinição da hospedagem ou da plataforma. É possível usar utilitários de banco de dados nativos para migrar o esquema.

dados quentes

Dados acessados com frequência, como dados em tempo real ou dados translacionais recentes. Esses dados normalmente exigem uma camada ou classe de armazenamento de alto desempenho para fornecer respostas rápidas às consultas.

hotfix

Uma correção urgente para um problema crítico em um ambiente de produção. Devido à sua urgência, um hotfix geralmente é feito fora do fluxo de trabalho normal de DevOps lançamento.

período de hipercuidados

Imediatamente após a substituição, o período em que uma equipe de migração gerencia e monitora as aplicações migradas na nuvem para resolver quaisquer problemas. Normalmente,

a duração desse período é de 1 a 4 dias. No final do período de hipercuidados, a equipe de migração normalmente transfere a responsabilidade pelas aplicações para a equipe de operações de nuvem.

eu

laC

Veja [infraestrutura como código](#).

Política baseada em identidade

Uma política anexada a um ou mais diretores do IAM que define suas permissões no Nuvem AWS ambiente.

aplicação ociosa

Uma aplicação que tem um uso médio de CPU e memória entre 5 e 20% em um período de 90 dias. Em um projeto de migração, é comum retirar essas aplicações ou retê-las on-premises.

IloT

Veja [Internet das Coisas Industrial](#).

infraestrutura imutável

Um modelo que implanta uma nova infraestrutura para workloads de produção em vez de atualizar, aplicar patches ou modificar a infraestrutura existente. Infraestruturas imutáveis são inerentemente mais consistentes, confiáveis e preditivas do que [infraestruturas mutáveis](#). Para obter mais informações, consulte a prática recomendada [Implantar usando infraestrutura imutável](#) no AWS Well-Architected Framework.

VPC de entrada (admissão)

Em uma arquitetura de AWS várias contas, uma VPC que aceita, inspeciona e roteia conexões de rede de fora de um aplicativo. A [Arquitetura de Referência de AWS Segurança](#) recomenda configurar sua conta de rede com entrada, saída e inspeção VPCs para proteger a interface bidirecional entre seu aplicativo e a Internet em geral.

migração incremental

Uma estratégia de substituição na qual você migra a aplicação em pequenas partes, em vez de realizar uma única substituição completa. Por exemplo, é possível mover inicialmente

apenas alguns microsserviços ou usuários para o novo sistema. Depois de verificar se tudo está funcionando corretamente, mova os microsserviços ou usuários adicionais de forma incremental até poder descomissionar seu sistema herdado. Essa estratégia reduz os riscos associados a migrações de grande porte.

Indústria 4.0

Um termo que foi introduzido por [Klaus Schwab](#) em 2016 para se referir à modernização dos processos de manufatura por meio de avanços em conectividade, dados em tempo real, automação, analytics e IA/ML.

infraestrutura

Todos os recursos e ativos contidos no ambiente de uma aplicação.

Infraestrutura como código (IaC)

O processo de provisionamento e gerenciamento da infraestrutura de uma aplicação por meio de um conjunto de arquivos de configuração. A IaC foi projetada para ajudar você a centralizar o gerenciamento da infraestrutura, padronizar recursos e escalar rapidamente para que novos ambientes sejam reproduzíveis, confiáveis e consistentes.

Internet industrial das coisas (IIoT)

O uso de sensores e dispositivos conectados à Internet nos setores industriais, como manufatura, energia, automotivo, saúde, ciências biológicas e agricultura. Para obter mais informações, consulte [Criando uma estratégia de transformação digital industrial da Internet das Coisas \(IIoT\)](#).

VPC de inspeção

Em uma arquitetura de AWS várias contas, uma VPC centralizada que gerencia as inspeções do tráfego de rede entre VPCs (na mesma ou em diferentes Regiões da AWS) a Internet e as redes locais. A [Arquitetura de Referência de AWS Segurança](#) recomenda configurar sua conta de rede com entrada, saída e inspeção VPCs para proteger a interface bidirecional entre seu aplicativo e a Internet em geral.

Internet das coisas (IoT)

A rede de objetos físicos conectados com sensores ou processadores incorporados que se comunicam com outros dispositivos e sistemas pela Internet ou por uma rede de comunicação local. Para obter mais informações, consulte [O que é IoT?](#)

interpretabilidade

Uma característica de um modelo de machine learning que descreve o grau em que um ser humano pode entender como as previsões do modelo dependem de suas entradas. Para obter mais informações, consulte [Interpretabilidade do modelo de aprendizado de máquina com AWS](#).

IoT

Veja [Internet das Coisas](#).

Biblioteca de informações de TI (ITIL)

Um conjunto de práticas recomendadas para fornecer serviços de TI e alinhar esses serviços a requisitos de negócios. A ITIL fornece a base para o ITSM.

Gerenciamento de serviços de TI (ITSM)

Atividades associadas a design, implementação, gerenciamento e suporte de serviços de TI para uma organização. Para obter informações sobre a integração de operações em nuvem com ferramentas de ITSM, consulte o [guia de integração de operações](#).

ITIL

Veja [biblioteca de informações de TI](#).

ITSM

Veja [gerenciamento de serviços de TI](#).

L

controle de acesso baseado em etiqueta (LBAC)

Uma implementação do controle de acesso obrigatório (MAC) em que os usuários e os dados em si recebem explicitamente um valor de etiqueta de segurança. A interseção entre a etiqueta de segurança do usuário e a etiqueta de segurança dos dados determina quais linhas e colunas podem ser vistas pelo usuário.

zona de pouso

Uma landing zone é um AWS ambiente bem arquitetado, com várias contas, escalável e seguro. Um ponto a partir do qual suas organizações podem iniciar e implantar rapidamente workloads e aplicações com confiança em seu ambiente de segurança e infraestrutura. Para obter mais

informações sobre zonas de pouso, consulte [Configurar um ambiente da AWS com várias contas seguro e escalável](#).

grande modelo de linguagem (LLM)

Um modelo de [IA](#) de aprendizado profundo pré-treinado em uma grande quantidade de dados. Um LLM pode realizar várias tarefas, como responder a perguntas, resumir documentos, traduzir texto para outros idiomas e completar frases. Para obter mais informações, consulte [O que são LLMs](#).

migração de grande porte

Uma migração de 300 servidores ou mais.

LBAC

Veja [controle de acesso baseado em rótulo](#).

privilégio mínimo

A prática recomendada de segurança de conceder as permissões mínimas necessárias para executar uma tarefa. Para obter mais informações, consulte [Aplicar permissões de privilégios mínimos](#) na documentação do IAM.

mover sem alterações (lift-and-shift)

Veja [7 Rs](#).

sistema little-endian

Um sistema que armazena o byte menos significativo antes. Veja também [endianness](#).

LLM

Veja [grande modelo de linguagem](#).

ambientes inferiores

Veja [ambiente](#).

M

machine learning (ML)

Um tipo de inteligência artificial que usa algoritmos e técnicas para reconhecimento e aprendizado de padrões. O ML analisa e aprende com dados gravados, por exemplo, dados da

Internet das Coisas (IoT), para gerar um modelo estatístico baseado em padrões. Para obter mais informações, consulte [Machine learning](#).

ramificação principal

Veja [ramificação](#).

Malware

Software projetado para comprometer a segurança ou a privacidade do computador. O malware pode interromper os sistemas do computador, vaziar informações sensíveis ou obter acesso não autorizado. Exemplos de malware incluem vírus, worms, ransomware, cavalos de Troia, spyware e keyloggers.

Serviços gerenciados

Serviços da AWS para o qual AWS opera a camada de infraestrutura, o sistema operacional e as plataformas, e você acessa os endpoints para armazenar e recuperar dados. O Amazon Simple Storage Service (Amazon S3) e o Amazon DynamoDB são exemplos de serviços gerenciados. Eles também são conhecidos como serviços abstraídos.

sistema de execução de manufatura (MES)

Um sistema de software para rastrear, monitorar, documentar e controlar processos de produção que convertem matérias-primas em produtos acabados no chão de fábrica.

MAP

Veja [Programa de Aceleração da Migração](#).

mecanismo

Um processo completo em que você cria uma ferramenta, impulsiona a adoção da ferramenta e, em seguida, inspeciona os resultados para fazer ajustes. Um mecanismo é um ciclo que se reforça e se aprimora à medida que opera. Para obter mais informações, consulte [Construindo mecanismos](#) no AWS Well-Architected Framework.

conta de membro

Todos, Contas da AWS exceto a conta de gerenciamento, que fazem parte de uma organização em AWS Organizations. Uma conta só pode ser membro de uma organização de cada vez.

MES

Veja [sistema de execução de manufatura](#).

Transporte de Telemetria de Enfileiramento de Mensagens (MQTT)

[Um protocolo de comunicação leve machine-to-machine \(M2M\), baseado no padrão de publicação/assinatura, para dispositivos de IoT com recursos limitados.](#)

microsserviço

Um serviço pequeno e independente que se comunica de forma bem definida APIs e normalmente é de propriedade de equipes pequenas e independentes. Por exemplo, um sistema de seguradora pode incluir microsserviços que mapeiam as capacidades comerciais, como vendas ou marketing, ou subdomínios, como compras, reclamações ou análises. Os benefícios dos microsserviços incluem agilidade, escalabilidade flexível, fácil implantação, código reutilizável e resiliência. Para obter mais informações, consulte [Integração de microsserviços usando serviços sem AWS servidor.](#)

arquitetura de microsserviços

Uma abordagem à criação de aplicações com componentes independentes que executam cada processo de aplicação como um microsserviço. Esses microsserviços se comunicam por meio de uma interface bem definida usando leveza. APIs Cada microsserviço nessa arquitetura pode ser atualizado, implantado e escalado para atender à demanda por funções específicas de uma aplicação. Para obter mais informações, consulte [Implementação de microsserviços em. AWS](#)

Programa de Aceleração da Migração (MAP)

Um AWS programa que fornece suporte de consultoria, treinamento e serviços para ajudar as organizações a criar uma base operacional sólida para migrar para a nuvem e ajudar a compensar o custo inicial das migrações. O MAP inclui uma metodologia de migração para executar migrações legadas de forma metódica e um conjunto de ferramentas para automatizar e acelerar cenários comuns de migração.

migração em escala

O processo de mover a maior parte do portfólio de aplicações para a nuvem em ondas, com mais aplicações sendo movidas em um ritmo mais rápido a cada onda. Essa fase usa as práticas recomendadas e lições aprendidas nas fases anteriores para implementar uma fábrica de migração de equipes, ferramentas e processos para agilizar a migração de workloads por meio de automação e entrega ágeis. Esta é a terceira fase da [estratégia de migração para a AWS.](#)

fábrica de migração

Equipes multifuncionais que simplificam a migração de workloads por meio de abordagens automatizadas e ágeis. As equipes da fábrica de migração geralmente incluem operações,

analistas e proprietários de negócios, engenheiros de migração, desenvolvedores e DevOps profissionais que trabalham em sprints. Entre 20 e 50% de um portfólio de aplicações corporativas consiste em padrões repetidos que podem ser otimizados por meio de uma abordagem de fábrica. Para obter mais informações, consulte [discussão sobre fábricas de migração](#) e o [guia do Cloud Migration Factory](#) neste conjunto de conteúdo.

metadados de migração

As informações sobre a aplicação e o servidor necessárias para concluir a migração. Cada padrão de migração exige um conjunto de metadados de migração diferente. Exemplos de metadados de migração incluem a sub-rede, o grupo de segurança e AWS a conta de destino.

padrão de migração

Uma tarefa de migração repetível que detalha a estratégia de migração, o destino da migração e a aplicação ou o serviço de migração usado. Exemplo: rehoste a migração para o Amazon EC2 AWS com o Application Migration Service.

Avaliação de Portfólio para Migração (MPA)

Uma ferramenta on-line que fornece informações para validar o caso de negócios para migrar para a Nuvem AWS. O MPA fornece avaliação detalhada do portfólio (dimensionamento correto do servidor, preços, comparações de TCO, análise de custos de migração), bem como planejamento de migração (análise e coleta de dados de aplicações, agrupamento de aplicações, priorização de migração e planejamento de ondas). A [ferramenta MPA](#) (requer login) está disponível gratuitamente para todos os AWS consultores e consultores parceiros da APN.

Avaliação de Preparação para Migração (MRA)

O processo de obter insights sobre o status de prontidão de uma organização para a nuvem, identificar pontos fortes e fracos e criar um plano de ação para fechar as lacunas identificadas, usando o CAF. AWS Para mais informações, consulte o [guia de preparação para migração](#). A MRA é a primeira fase da [estratégia de migração para a AWS](#).

estratégia de migração

A abordagem usada para migrar uma workload para a Nuvem AWS. Para obter mais informações, veja a entrada [7 Rs](#) neste glossário e consulte [Mobilize sua organização para acelerar migrações em grande escala](#).

ML

Veja [machine learning](#).

modernização

Transformar uma aplicação desatualizada (herdada ou monolítica) e sua infraestrutura em um sistema ágil, elástico e altamente disponível na nuvem para reduzir custos, ganhar eficiência e aproveitar as inovações. Para obter mais informações, consulte [Strategy for modernizing applications in the Nuvem AWS](#).

avaliação de preparação para modernização

Uma avaliação que ajuda a determinar a preparação para modernização das aplicações de uma organização. Ela identifica benefícios, riscos e dependências e determina o quão bem a organização pode acomodar o estado futuro dessas aplicações. O resultado da avaliação é um esquema da arquitetura de destino, um roteiro que detalha as fases de desenvolvimento e os marcos do processo de modernização e um plano de ação para abordar as lacunas identificadas. Para obter mais informações, consulte [Evaluating modernization readiness for applications in the Nuvem AWS](#).

aplicações monolíticas (monólitos)

Aplicações que são executadas como um único serviço com processos fortemente acoplados. As aplicações monolíticas apresentam várias desvantagens. Se um recurso da aplicação apresentar um aumento na demanda, toda a arquitetura deverá ser escalada. Adicionar ou melhorar os recursos de uma aplicação monolítica também se torna mais complexo quando a base de código cresce. Para resolver esses problemas, é possível criar uma arquitetura de microsserviços. Para obter mais informações, consulte [Decompor monólitos em microsserviços](#).

MPA

Veja [Avaliação do Portfólio para Migração](#).

MQTT

Veja [Transporte de Telemetria de Enfileiramento de Mensagens](#).

classificação multiclasse

Um processo que ajuda a gerar previsões para várias classes (prevendo um ou mais de dois resultados). Por exemplo, um modelo de ML pode perguntar “Este produto é um livro, um carro ou um telefone?” ou “Qual categoria de produtos é mais interessante para este cliente?”

infraestrutura mutável

Um modelo que atualiza e modifica a infraestrutura existente para workloads de produção. Para melhorar a consistência, confiabilidade e previsibilidade, o AWS Well-Architected Framework recomenda o uso de infraestrutura [imutável](#) como uma prática recomendada.

O

OAC

Veja [controle de acesso de origem](#).

OAI

Veja [identidade de acesso de origem](#).

OCM

Veja [gerenciamento de alterações organizacionais](#).

migração offline

Um método de migração no qual a workload de origem é desativada durante o processo de migração. Esse método envolve tempo de inatividade prolongado e geralmente é usado para workloads pequenas e não críticas.

OI

Veja [integração de operações](#).

Ola

Veja [acordo de nível operacional](#).

migração online

Um método de migração no qual a workload de origem é copiada para o sistema de destino sem ser colocada offline. As aplicações conectadas à workload podem continuar funcionando durante a migração. Esse método envolve um tempo de inatividade nulo ou mínimo e normalmente é usado para workloads essenciais para a produção.

OPC-UA

Veja [Open Process Communications - Unified Architecture](#).

Open Process Communications - Unified Architecture (OPC-UA)

Um protocolo de comunicação machine-to-machine (M2M) para automação industrial. O OPC-UA fornece um padrão de interoperabilidade com esquemas de criptografia, autenticação e autorização de dados.

acordo de nível operacional (OLA)

Um acordo que esclarece o que os grupos funcionais de TI prometem oferecer uns aos outros para apoiar um acordo de serviço (SLA).

análise de prontidão operacional (ORR)

Uma lista de verificação de perguntas e práticas recomendadas associadas que ajudam você a entender, avaliar, prevenir ou reduzir o escopo de incidentes e possíveis falhas. Para obter mais informações, consulte [Operational Readiness Reviews \(ORR\)](#) no AWS Well-Architected Framework.

tecnologia operacional (TO)

Sistemas de hardware e software que trabalham com o ambiente físico para controlar operações, equipamentos e infraestrutura industriais. Na manufatura, a integração dos sistemas de tecnologia da informação (TI) e tecnologia operacional (TO) é o foco principal das transformações da [Indústria 4.0](#).

integração de operações (OI)

O processo de modernização das operações na nuvem, que envolve planejamento de preparação, automação e integração. Para obter mais informações, consulte o [guia de integração de operações](#).

trilha organizacional

Uma trilha criada por ela AWS CloudTrail registra todos os eventos de todas as Contas da AWS em uma organização em AWS Organizations. Essa trilha é criada em cada Conta da AWS que faz parte da organização e monitora a atividade em cada conta. Para obter mais informações, consulte [Criação de uma trilha para uma organização](#) na CloudTrail documentação.

gerenciamento de alterações organizacionais (OCM)

Uma estrutura para gerenciar grandes transformações de negócios disruptivas de uma perspectiva de pessoas, cultura e liderança. O OCM ajuda as organizações a se prepararem e fazerem a transição para novos sistemas e estratégias, acelerando a adoção de alterações, abordando questões de transição e promovendo mudanças culturais e organizacionais. Na estratégia de AWS migração, essa estrutura é chamada de aceleração de pessoas, devido à velocidade de mudança exigida nos projetos de adoção da nuvem. Para obter mais informações, consulte o [guia do OCM](#).

controle de acesso de origem (OAC)

Em CloudFront, uma opção aprimorada para restringir o acesso para proteger seu conteúdo do Amazon Simple Storage Service (Amazon S3). O OAC oferece suporte a todos os buckets S3 Regiões da AWS, criptografia do lado do servidor com AWS KMS (SSE-KMS) e solicitações dinâmicas ao bucket S3. PUT DELETE

Identidade do acesso de origem (OAI)

Em CloudFront, uma opção para restringir o acesso para proteger seu conteúdo do Amazon S3. Quando você usa o OAI, CloudFront cria um principal com o qual o Amazon S3 pode se autenticar. Os diretores autenticados podem acessar o conteúdo em um bucket do S3 somente por meio de uma distribuição específica. CloudFront Veja também [OAC](#), que fornece um controle de acesso mais granular e aprimorado.

ORR

Veja [análise de prontidão operacional](#).

OT

Veja [tecnologia operacional](#).

VPC de saída (egresso)

Em uma arquitetura de AWS várias contas, uma VPC que gerencia conexões de rede que são iniciadas de dentro de um aplicativo. A [Arquitetura de Referência de AWS Segurança](#) recomenda configurar sua conta de rede com entrada, saída e inspeção VPCs para proteger a interface bidirecional entre seu aplicativo e a Internet em geral.

P

limite de permissões

Uma política de gerenciamento do IAM anexada a entidades principais do IAM para definir as permissões máximas que o usuário ou perfil podem ter. Para obter mais informações, consulte [Limites de permissões](#) na documentação do IAM.

Informações de identificação pessoal (PII)

Informações que, quando visualizadas diretamente ou combinadas com outros dados relacionados, podem ser usadas para inferir razoavelmente a identidade de um indivíduo. Exemplos de PII incluem nomes, endereços e informações de contato.

PII

Veja [informações de identificação pessoal](#).

manual

Um conjunto de etapas predefinidas que capturam o trabalho associado às migrações, como a entrega das principais funções operacionais na nuvem. Um manual pode assumir a forma de scripts, runbooks automatizados ou um resumo dos processos ou etapas necessários para operar seu ambiente modernizado.

PLC

Veja [controlador lógico programável](#).

PLM

Veja [gerenciamento do ciclo de vida do produto](#).

política

Um objeto que pode definir permissões (veja [política baseada em identidade](#)), especificar condições de acesso (veja [política baseada em recurso](#)) ou definir as permissões máximas para todas as contas em uma organização no AWS Organizations (veja [política de controle de serviços](#)).

persistência poliglota

Escolher de forma independente a tecnologia de armazenamento de dados de um microsserviço com base em padrões de acesso a dados e outros requisitos. Se seus microsserviços tiverem a mesma tecnologia de armazenamento de dados, eles poderão enfrentar desafios de implementação ou apresentar baixa performance. Os microsserviços serão implementados com mais facilidade e alcançarão performance e escalabilidade melhores se usarem o armazenamento de dados mais bem adaptado às suas necessidades.

avaliação do portfólio

Um processo de descobrir, analisar e priorizar o portfólio de aplicações para planejar a migração. Para obter mais informações, consulte [Avaliar a preparação para a migração](#).

predicado

Uma condição de consulta que retorna `true` ou `false`, normalmente localizada em uma cláusula `WHERE`.

pushdown de predicados

Uma técnica de otimização de consultas de banco de dados que filtra os dados na consulta antes da transferência. Isso reduz a quantidade de dados que devem ser recuperados e processados do banco de dados relacional e melhora a performance das consultas.

controle preventivo

Um controle de segurança projetado para evitar que um evento ocorra. Esses controles são a primeira linha de defesa para ajudar a evitar acesso não autorizado ou alterações indesejadas em sua rede. Para obter mais informações, consulte [Controles preventivos](#) em Como implementar controles de segurança na AWS.

principal (entidade principal)

Uma entidade AWS que pode realizar ações e acessar recursos. Essa entidade geralmente é um usuário raiz para um Conta da AWS, uma função do IAM ou um usuário. Para obter mais informações, consulte Entidade principal em [Termos e conceitos de perfis](#) na documentação do IAM.

Privacidade por design

Uma abordagem em engenharia de sistemas que leva em consideração a privacidade em todo o processo de desenvolvimento.

zonas hospedadas privadas

Um contêiner que contém informações sobre como você deseja que o Amazon Route 53 responda às consultas de DNS para um domínio e seus subdomínios em um ou mais VPCs. Para obter mais informações, consulte [Como trabalhar com zonas hospedadas privadas](#) na documentação do Route 53.

controle proativo

Um [controle de segurança](#) desenvolvido para evitar a implantação de recursos não conformes. Esses controles verificam os recursos antes de serem provisionados. Se o recurso não estiver em conformidade com o controle, ele não será provisionado. Para obter mais informações, consulte o [guia de referência de controles](#) na AWS Control Tower documentação e consulte [Controles proativos](#) em Implementação de controles de segurança em AWS.

gerenciamento do ciclo de vida do produto (PLM)

O gerenciamento de dados e processos de um produto em todo o seu ciclo de vida, desde a concepção, o desenvolvimento e o lançamento, passando pelo crescimento e maturidade, até o declínio e a remoção.

ambiente de produção

Veja [ambiente](#).

controlador lógico programável (PLC)

Na manufatura, um computador altamente confiável e adaptável que monitora as máquinas e automatiza os processos de fabricação.

encadeamento de prompts

Uso da saída de um prompt do [LLM](#) como entrada para o próximo prompt para gerar respostas melhores. Essa técnica é usada para dividir uma tarefa complexa em subtarefas, ou para refinar ou expandir iterativamente uma resposta preliminar. Isso ajuda a melhorar a precisão e a relevância das respostas de um modelo e permite resultados mais granulares e personalizados.

pseudonimização

O processo de substituir identificadores pessoais em um conjunto de dados por valores de espaço reservado. A pseudonimização pode ajudar a proteger a privacidade pessoal. Os dados pseudonimizados ainda são considerados dados pessoais.

publish/subscribe (pub/sub)

Um padrão que permite comunicações assíncronas entre microsserviços para melhorar a escalabilidade e a capacidade de resposta. Por exemplo, em um [MES](#) baseado em microsserviços, um microsserviço pode publicar mensagens de eventos em um canal em que outros microsserviços possam assinar. O sistema pode adicionar novos microsserviços sem alterar o serviço de publicação.

Q

plano de consulta

Uma série de etapas, como instruções, usadas para acessar os dados em um sistema de banco de dados relacional SQL.

regressão de planos de consultas

Quando um otimizador de serviço de banco de dados escolhe um plano menos adequado do que escolhia antes de uma determinada alteração no ambiente de banco de dados ocorrer. Isso pode ser causado por alterações em estatísticas, restrições, configurações do ambiente, associações de parâmetros de consulta e atualizações do mecanismo de banco de dados.

R

Matriz RACI

Veja [responsável, aprovador, consultado, informado \(RACI\)](#).

RAG

Veja [geração aumentada via recuperação](#).

ransomware

Um software mal-intencionado desenvolvido para bloquear o acesso a um sistema ou dados de computador até que um pagamento seja feito.

Matriz RASCI

Veja [responsável, aprovador, consultado, informado \(RACI\)](#).

RCAC

Veja [controle de acesso por linha e coluna](#).

réplica de leitura

Uma cópia de um banco de dados usada somente para leitura. É possível encaminhar consultas para a réplica de leitura e reduzir a carga no banco de dados principal.

Redefinir arquitetura

Veja [7 Rs](#).

objetivo de ponto de recuperação (RPO).

O máximo período de tempo aceitável desde o último ponto de recuperação de dados. Isso determina o que é considerado uma perda aceitável de dados entre o último ponto de recuperação e a interrupção do serviço.

objetivo de tempo de recuperação (RTO)

O máximo atraso aceitável entre a interrupção e a restauração do serviço.

refatorar

Veja [7 Rs](#).

Região

Uma coleção de AWS recursos em uma área geográfica. Cada um Região da AWS é isolado e independente dos outros para fornecer tolerância a falhas, estabilidade e resiliência. Para obter informações, consulte [Specify which Regiões da AWS your account can use](#).

regressão

Uma técnica de ML que prevê um valor numérico. Por exemplo, para resolver o problema de “Por qual preço esta casa será vendida?” um modelo de ML pode usar um modelo de regressão linear para prever o preço de venda de uma casa com base em fatos conhecidos sobre a casa (por exemplo, a metragem quadrada).

redefinir a hospedagem

Veja [7 Rs](#).

versão

Em um processo de implantação, o ato de promover mudanças em um ambiente de produção.

realocar

Veja [7 Rs](#).

redefinir a plataforma

Veja [7 Rs](#).

recomprar

Veja [7 Rs](#).

resiliência

A capacidade de uma aplicação de resistir ou se recuperar de interrupções. [Alta disponibilidade](#) e [recuperação de desastres](#) são considerações comuns ao planejar a resiliência na Nuvem AWS. Para obter mais informações, consulte [Nuvem AWS Resilience](#).

política baseada em recurso

Uma política associada a um recurso, como um bucket do Amazon S3, um endpoint ou uma chave de criptografia. Esse tipo de política especifica quais entidades principais têm acesso permitido, ações válidas e quaisquer outras condições que devem ser atendidas.

matriz responsável, accountable, consultada, informada (RACI)

Uma matriz que define as funções e responsabilidades de todas as partes envolvidas nas atividades de migração e nas operações de nuvem. O nome da matriz é derivado dos tipos de responsabilidade definidos na matriz: responsável (R), responsabilizável (A), consultado (C) e informado (I). O tipo de suporte (S) é opcional. Se você incluir suporte, a matriz será chamada de matriz RASCI e, se excluir, será chamada de matriz RACI.

controle responsivo

Um controle de segurança desenvolvido para conduzir a remediação de eventos adversos ou desvios em relação à linha de base de segurança. Para obter mais informações, consulte [Controles responsivos](#) em Como implementar controles de segurança na AWS.

reter

Veja [7 Rs](#).

Retirada

Veja [7 Rs](#).

Geração Aumentada de Recuperação (RAG)

Uma tecnologia de [IA generativa](#) em que um [LLM](#) faz referência a uma fonte de dados autorizada que está fora de suas fontes de dados de treinamento antes de gerar uma resposta. Por exemplo, um modelo RAG pode realizar uma pesquisa semântica na base de conhecimento ou nos dados personalizados de uma organização. Para obter mais informações, consulte [O que é RAG \(geração aumentada via recuperação\)?](#).

alternância

O processo de atualizar periodicamente um [segredo](#) para dificultar o acesso de um invasor às credenciais.

controle de acesso por linha e coluna (RCAC)

O uso de expressões SQL básicas e flexíveis que tenham regras de acesso definidas. O RCAC consiste em permissões de linha e máscaras de coluna.

RPO

Veja [objetivo de ponto de recuperação](#).

RTO

Veja [objetivo de tempo de recuperação](#).

runbook

Um conjunto de procedimentos manuais ou automatizados necessários para realizar uma tarefa específica. Eles são normalmente criados para agilizar operações ou procedimentos repetitivos com altas taxas de erro.

S

SAML 2.0

Um padrão aberto que muitos provedores de identidade (IdPs) usam. Esse recurso permite o login único federado (SSO), para que os usuários possam fazer login no Console de gerenciamento da AWS ou chamar as operações da AWS API sem que você precise criar um usuário no IAM para todos em sua organização. Para obter mais informações sobre a federação baseada em SAML 2.0, consulte [Sobre a federação baseada em SAML 2.0](#) na documentação do IAM.

SCADA

Veja [controle de supervisão e aquisição de dados](#).

SCP

Veja [política de controle de serviço](#).

secret

Em AWS Secrets Manager, informações confidenciais ou restritas, como uma senha ou credenciais de usuário, que você armazena de forma criptografada. Consiste no valor secreto e em seus metadados. O valor secreto pode ser binário, uma única string ou várias strings. Para obter mais informações, consulte [What's in a Secrets Manager secret?](#) na documentação do Secrets Manager.

segurança desde a concepção

Uma abordagem em engenharia de sistemas que leva em consideração a segurança em todo o processo de desenvolvimento.

controle de segurança

Uma barreira de proteção técnica ou administrativa que impede, detecta ou reduz a capacidade de uma ameaça explorar uma vulnerabilidade de segurança. Existem quatro tipos primários de controles de segurança: [preventivos](#), [detectivos](#), [responsivos](#) e [proativos](#).

hardening da segurança

O processo de reduzir a superfície de ataque para torná-la mais resistente a ataques. Isso pode incluir ações como remover recursos que não são mais necessários, implementar a prática recomendada de segurança de conceder privilégios mínimos ou desativar recursos desnecessários em arquivos de configuração.

sistema de gerenciamento de eventos e informações de segurança (SIEM)

Ferramentas e serviços que combinam sistemas de gerenciamento de informações de segurança (SIM) e gerenciamento de eventos de segurança (SEM). Um sistema SIEM coleta, monitora e analisa dados de servidores, redes, dispositivos e outras fontes para detectar ameaças e violações de segurança e gerar alertas.

automação de resposta de segurança

Uma ação predefinida e programada projetada para responder ou remediar automaticamente um evento de segurança. Essas automações servem como controles de segurança [responsivos](#) ou [detectivos](#) que ajudam você a implementar as melhores práticas AWS de segurança. Exemplos de ações de resposta automatizada incluem a modificação de um grupo de segurança da VPC, a aplicação de patches em uma instância do Amazon EC2 ou a alternância de credenciais.

Criptografia do lado do servidor

Criptografia dos dados em seu destino, por AWS service (Serviço da AWS) quem os recebe.

política de controle de serviços (SCP)

Uma política que fornece controle centralizado sobre as permissões de todas as contas em uma organização em AWS Organizations. SCPs defina barreiras ou estabeleça limites nas ações que um administrador pode delegar a usuários ou funções. Você pode usar SCPs como listas de permissão ou listas de negação para especificar quais serviços ou ações são permitidos ou proibidos. Para obter mais informações, consulte [Políticas de controle de serviço](#) na AWS Organizations documentação.

service endpoint (endpoint de serviço)

O URL do ponto de entrada para um AWS service (Serviço da AWS). Você pode usar o endpoint para se conectar programaticamente ao serviço de destino. Para obter mais informações, consulte [Endpoints do AWS service \(Serviço da AWS\)](#) na Referência geral da AWS.

acordo de serviço (SLA)

Um acordo que esclarece o que uma equipe de TI promete fornecer aos clientes, como tempo de atividade e performance do serviço.

indicador de nível de serviço (SLI)

Uma avaliação de um aspecto de performance de um serviço, como taxa de erro, disponibilidade ou throughput.

objetivo de nível de serviço (SLO)

Uma métrica alvo que representa a integridade de um serviço, conforme avaliado por um [indicador de nível de serviço](#).

modelo de responsabilidade compartilhada

Um modelo que descreve a responsabilidade com a qual você compartilha AWS pela segurança e conformidade na nuvem. AWS é responsável pela segurança da nuvem, enquanto você é responsável pela segurança na nuvem. Para obter mais informações, consulte o [Modelo de responsabilidade compartilhada](#).

SIEM

Veja [sistema de gerenciamento de eventos e informações de segurança](#).

ponto único de falha (SPOF)

Uma falha em um único componente crítico de uma aplicação que pode interromper o sistema.

SLA

Veja [acordo de serviço](#).

SLI

Veja [indicador de nível de serviço](#).

SLO

Veja [objetivo de nível de serviço](#).

split-and-seed modelo

Um padrão para escalar e acelerar projetos de modernização. À medida que novos recursos e lançamentos de produtos são definidos, a equipe principal se divide para criar novas equipes de produtos. Isso ajuda a escalar os recursos e os serviços da sua organização, melhora a produtividade do desenvolvedor e possibilita inovações rápidas. Para obter mais informações, consulte [Phased approach to modernizing applications in the Nuvem AWS](#).

SPOF

Veja [ponto único de falha](#).

esquema em estrela

Uma estrutura organizacional de banco de dados que usa uma grande tabela de fatos para armazenar dados transacionais ou medidos e usa uma ou mais tabelas dimensionais menores para armazenar atributos de dados. Essa estrutura foi projetada para ser usada em um [data warehouse](#) ou para fins de inteligência comercial.

padrão strangler fig

Uma abordagem à modernização de sistemas monolíticos que consiste em reescrever e substituir incrementalmente a funcionalidade do sistema até que o sistema herdado possa ser desativado. Esse padrão usa a analogia de uma videira que cresce e se torna uma árvore estabelecida e, eventualmente, supera e substitui sua hospedeira. O padrão foi [apresentado por Martin Fowler](#) como forma de gerenciar riscos ao reescrever sistemas monolíticos. Para ver um exemplo de como aplicar esse padrão, consulte [Modernizar incrementalmente os serviços Web herdados do Microsoft ASP.NET \(ASMX\) usando contêineres e o Amazon API Gateway](#).

sub-rede

Um intervalo de endereços IP na VPC. Cada sub-rede fica alocada em uma única zona de disponibilidade.

controle supervisão e aquisição de dados (SCADA)

Na manufatura, um sistema que usa hardware e software para monitorar ativos físicos e operações de produção.

symmetric encryption (criptografia simétrica)

Um algoritmo de criptografia que usa a mesma chave para criptografar e descriptografar dados.

testes sintéticos

Testar um sistema de forma que simule as interações do usuário para detectar possíveis problemas ou monitorar a performance. Você pode usar o [Amazon CloudWatch Synthetics](#) para criar esses testes.

prompt do sistema

Uma técnica para fornecer contexto, instruções ou orientações a um [LLM](#) a fim de direcionar seu comportamento. Os prompts do sistema ajudam a definir o contexto e a estabelecer regras para interações com os usuários.

T

tags

Pares de valores-chave que atuam como metadados para organizar seus recursos. AWS As tags podem ajudar você a gerenciar, identificar, organizar, pesquisar e filtrar recursos da . Para obter mais informações, consulte [Marcar seus recursos do AWS](#).

variável-alvo

O valor que você está tentando prever no ML supervisionado. Ela também é conhecida como variável de resultado. Por exemplo, em uma configuração de fabricação, a variável-alvo pode ser um defeito do produto.

lista de tarefas

Uma ferramenta usada para monitorar o progresso por meio de um runbook. Uma lista de tarefas contém uma visão geral do runbook e uma lista de tarefas gerais a serem concluídas. Para cada tarefa geral, ela inclui o tempo estimado necessário, o proprietário e o progresso.

ambiente de teste

Veja [ambiente](#).

treinamento

O processo de fornecer dados para que seu modelo de ML aprenda. Os dados de treinamento devem conter a resposta correta. O algoritmo de aprendizado descobre padrões nos dados de treinamento que mapeiam os atributos dos dados de entrada no destino (a resposta que você deseja prever). Ele gera um modelo de ML que captura esses padrões. Você pode usar o modelo de ML para obter previsões de novos dados cujo destino você não conhece.

gateway de trânsito

Um hub de trânsito de rede que você pode usar para interconectar sua rede com VPCs a rede local. Para obter mais informações, consulte [O que é um gateway de trânsito](#) na AWS Transit Gateway documentação.

fluxo de trabalho baseado em troncos

Uma abordagem na qual os desenvolvedores criam e testam recursos localmente em uma ramificação de recursos e, em seguida, mesclam essas alterações na ramificação principal. A ramificação principal é então criada para os ambientes de desenvolvimento, pré-produção e produção, sequencialmente.

Acesso confiável

Conceder permissões a um serviço que você especifica para realizar tarefas em sua organização AWS Organizations e em suas contas em seu nome. O serviço confiável cria um perfil vinculado ao serviço em cada conta, quando esse perfil é necessário, para realizar tarefas de gerenciamento para você. Para obter mais informações, consulte [Usando AWS Organizations com outros AWS serviços](#) na AWS Organizations documentação.

tuning (ajustar)

Alterar aspectos do processo de treinamento para melhorar a precisão do modelo de ML. Por exemplo, você pode treinar o modelo de ML gerando um conjunto de rótulos, adicionando rótulos e repetindo essas etapas várias vezes em configurações diferentes para otimizar o modelo.

equipe de duas pizzas

Uma pequena DevOps equipe que você pode alimentar com duas pizzas. Uma equipe de duas pizzas garante a melhor oportunidade possível de colaboração no desenvolvimento de software.

U

incerteza

Um conceito que se refere a informações imprecisas, incompletas ou desconhecidas que podem minar a confiabilidade dos modelos preditivos de ML. Há dois tipos de incertezas: a incerteza epistêmica é causada por dados limitados e incompletos, enquanto a incerteza aleatória é causada pelo ruído e pela aleatoriedade inerentes aos dados.

tarefas indiferenciadas

Também conhecido como trabalho pesado, trabalho necessário para criar e operar um aplicativo, mas que não fornece valor direto ao usuário final nem oferece vantagem competitiva. Exemplos de tarefas indiferenciadas incluem aquisição, manutenção e planejamento de capacidade.

ambientes superiores

Veja [ambiente](#).

V

aspiração

Uma operação de manutenção de banco de dados que envolve limpeza após atualizações incrementais para recuperar armazenamento e melhorar a performance.

controle de versões

Processos e ferramentas que rastreiam mudanças, como alterações no código-fonte em um repositório.

emparelhamento da VPC

Uma conexão entre duas VPCs que permite rotear o tráfego usando endereços IP privados. Para ter mais informações, consulte [O que é emparelhamento de VPC?](#) na documentação da Amazon VPC.

Vulnerabilidade

Uma falha de software ou hardware que compromete a segurança do sistema.

W

cache quente

Um cache de buffer que contém dados atuais e relevantes que são acessados com frequência. A instância do banco de dados pode ler do cache do buffer, o que é mais rápido do que ler da memória principal ou do disco.

dados mornos

Dados acessados raramente. Ao consultar esse tipo de dados, consultas moderadamente lentas geralmente são aceitáveis.

função de janela

Uma função SQL que executa um cálculo em um grupo de linhas que se relacionam de alguma forma com o registro atual. As funções de janela são úteis para processar tarefas, como calcular uma média móvel ou acessar o valor das linhas com base na posição relativa da linha atual.

workload

Uma coleção de códigos e recursos que geram valor empresarial, como uma aplicação voltada para o cliente ou um processo de backend.

workstreams

Grupos funcionais em um projeto de migração que são responsáveis por um conjunto específico de tarefas. Cada workstream é independente, mas oferece suporte aos outros workstreams do projeto. Por exemplo, o workstream de portfólio é responsável por priorizar aplicações, planejar ondas e coletar metadados de migração. O workstream de portfólio entrega esses ativos ao workstream de migração, que então migra os servidores e as aplicações.

WORM

Veja [gravação única e várias leituras](#).

WQF

Veja [AWS Workload Qualification Framework](#).

gravação única e várias leituras (WORM)

Um modelo de armazenamento que grava dados uma única vez e evita que os dados sejam excluídos ou modificados. Os usuários autorizados podem ler os dados quantas vezes forem necessárias, mas não podem alterá-los. Essa infraestrutura de armazenamento de dados é considerada [imutável](#).

Z

exploração de dia zero

Um ataque, normalmente malware, que tira proveito de uma [vulnerabilidade zero-day](#).

vulnerabilidade de dia zero

Uma falha ou vulnerabilidade não mitigada em um sistema de produção. Os agentes de ameaças podem usar esse tipo de vulnerabilidade para atacar o sistema. Os desenvolvedores frequentemente ficam cientes da vulnerabilidade como resultado do ataque.

prompt zero shot

Fornecer a um [LLM](#) instruções para realizar uma tarefa, mas sem exemplos (shots) que possam ajudar a orientá-lo. O LLM deve usar seu conhecimento pré-treinado para lidar com a tarefa. A

eficácia dos prompts zero-shot depende da complexidade da tarefa e da qualidade do prompt.

Veja também [prompts few-shot](#).

aplicação zumbi

Uma aplicação que tem um uso médio de CPU e memória inferior a 5%. Em um projeto de migração, é comum retirar essas aplicações.

As traduções são geradas por tradução automática. Em caso de conflito entre o conteúdo da tradução e da versão original em inglês, a versão em inglês prevalecerá.