AWS Well-Architected 框架

成本优化支柱

Copyright © 2025 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

成本优化支柱: AWS Well-Architected 框架

Copyright © 2025 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon 的商标和商业外观不得用于任何非 Amazon 的商品或服务，也不得以任何可能引起客户混淆、贬低或诋毁 Amazon 的方式使用。所有非 Amazon 拥有的其他商标均为各自所有者的财产，这些所有者可能附属于 Amazon、与 Amazon 有关联或由 Amazon 赞助，也可能不是如此。

Table of Contents

摘要和简介 1
- 摘要 1
- 简介 1

成本优化 3
- 设计原则 3
- 定义 3

践行云财务管理 5
- COST01-BP01 确立成本优化的责任归属模式 7
  - 实施指导 7
  - 资源 9
- COST01-BP02 在财务和技术人员之间建立合作关系 9
  - 实施指导 7
  - 资源 9
- COST01-BP03 建立云预算和预测 13
  - 实施指导 7
  - 资源 9
- COST01-BP04 在组织流程中落实成本意识 16
  - 实施指导 7
  - 资源 9
- COST01-BP05 报告和通知成本优化 18
  - 实施指导 7
  - 资源 9
- COST01-BP06 主动监控成本 20
  - 实施指导 7
  - 资源 9
- COST01-BP07 及时了解新发布的服务 21
  - 实施指导 7
  - 资源 9
- COST01-BP08 建立对成本敏感的文化 22
  - 实施指导 7
  - 资源 9
- COST01-BP09 量化通过成本优化实现的业务价值 24
  - 实施指导 7
  - 资源 9

支出和使用情况意识 27
治理 27
COST02-BP01 根据组织的要求制定各种策略 27
COST02-BP02 实施方向性目标和执行性目标 31
COST02-BP03 实施账户结构 35
COST02-BP04 实施组和角色 38
COST02-BP05 实施成本控制 39
COST02-BP06 跟踪项目生命周期 41
监控成本和使用情况 43
COST03-BP01 配置详细信息源 43
COST03-BP02 在成本和使用情况中添加组织信息 46
COST03-BP03 确定成本归属类别 47
COST03-BP04 建立组织指标 50
COST03-BP05 配置账单和成本管理工具 51
COST03-BP06 根据工作负载指标分配成本 54
停用资源 55
COST04-BP01 在资源生命周期内跟踪资源 55
COST04-BP02 实施停用流程 57
COST04-BP03 停用资源 59
COST04-BP04 自动停用资源 60
COST04-BP05 执行数据留存策略 61
具有成本效益的资源 63
选择服务时评估成本 63
COST05-BP01 确定组织对成本的要求 63
COST05-BP02 分析工作负载的所有组件 65
COST05-BP03 对每个组件进行全面分析 67
COST05-BP04 选择具有成本效益许可的软件 69
COST05-BP05 选择此工作负载的组件，以便根据组织的优先事项优化成本 70
COST05-BP06 对不同时间的不同使用情况执行成本分析 72
选择正确的资源类型、规模和数量 73
COST06-BP01 执行成本建模 73
COST06-BP02 根据数据选择资源类型、规模和数量 75
COST06-BP03 根据指标自动选择资源类型、规模和数量 77
COST06-BP04 考虑使用共享资源 79
选择最佳定价模式 82
COST07-BP01 执行定价模式分析 86

COST07-BP02 根据成本选择区域 88
COST07-BP03 选择具有经济实惠条款的第三方协议 90
COST07-BP04 针对此工作负载的所有组件实施定价模式 91
COST07-BP05 在管理账户级别执行定价模式分析 93
制定数据传输计划 95
COST08-BP01 执行数据传输建模 95
COST08-BP02 选择组件以便优化数据传输成本 97
COST08-BP03 实施服务以便降低数据传输成本 99
管理需求和供应资源 101
COST09-BP01 对工作负载需求执行分析 101
实施指导 7
资源 9
COST09-BP02 实施缓冲区或节流来管理需求 103
实施指导 7
资源 9
COST09-BP03 动态供应资源 106
实施指导 7
实施步骤 8
资源 9
持续优化 112
定义审查流程并定期分析工作负载 112
COST10-BP01 制定工作负载审核流程 112
COST10-BP02 定期审核和分析此工作负载 114
自动实施运营 115
COST11-BP01 执行运营自动化 115
结论 119
贡献者 120
延伸阅读 121
文档修订 122
版权声明 124
AWS 术语表 125

成本优化支柱 – AWS Well-Architected Framework

发布日期：2024 年 6 月 27 日（文档修订）

摘要

本白皮书重点介绍亚马逊云科技（AWS）Well-Architected Framework 的成本优化支柱。文中所述指导可帮助客户在 AWS 环境的设计、交付和维护过程中应用最佳实践。

经过成本优化的工作负载能够充分利用所有资源，以尽可能低的价格实现成果，并满足您的功能要求。本白皮书提供了有关在组织内增强能力、设计工作负载、选择服务、配置和运营服务以及应用成本优化技术的深入指导。

简介

AWS Well-Architected Framework 能够帮助您理解您在 AWS 上构建工作负载时所作的决策。该框架提供了在云中设计和运行可靠、安全、高效、经济实惠且可持续的工作负载的架构最佳实践。它提供了一种方法，使您能够根据最佳实践一致地衡量架构，并确定需要改进的方面。我们相信，拥有架构完善的工作负载能够大大提高实现业务成功的可能性。

该框架基于六大支柱：

- 卓越运营
- 安全性
- 可靠性
- 性能效率
- 成本优化
- 可持续性

本白皮书重点介绍成本优化支柱以及如何以最高效的方式使用服务和资源来构建工作负载，从而以最低的价格实现业务成果。

您将了解如何在组织内应用成本优化支柱的最佳实践。对传统的本地解决方案进行成本优化可能比较困难，因为您必须要预测未来容量和业务需求，同时要掌控复杂的采购流程。采用本白皮书中的实践可帮助您的组织实现以下目标：

- 践行云财务管理
- 支出和使用情况意识
- 具有成本效益的资源
- 管理需求和供应资源
- 持续优化

本白皮书的目标读者是技术和财务岗位的人员，例如首席技术官（CTO），首席财务官（CFO）、架构师、开发人员、财务总监、财务规划师、业务分析师和运营团队成员。本白皮书不提供实施细节或架构模式，但包含了适当资源参考。

成本优化

成本优化是在工作负载的整个生命周期中不断完善和改进的过程。本白皮书中的实践有助于构建和运营对成本敏感的工作负载，这些工作负载可在帮助实现业务成果的同时，最大限度降低成本并助力组织最大限度提高其投资回报。

主题

- 设计原则
- 定义

设计原则

请考虑以下成本优化设计原则：

实施云财务管理：要获得财务上的成功并加速在云中实现业务价值，必须投资云财务管理。组织必须投入必要的时间和资源，增强自身在这个新的技术和使用管理领域中的能力。与安全或运营能力类似，组织需要通过知识积累、计划、资源和流程来增强自身能力，从而成为一个具有成本效益的组织。

采用消费模型：仅为所用的计算资源付费，并可根据业务要求增加或减少使用量。例如，开发和测试环境通常只需要在每个工作日运行八个小时。可以在不需要的时候停用这些资源，这样有可能节省 75% 的成本（40 小时对比 168 小时）。

衡量整体效率：衡量工作负载的业务产出及其交付成本。使用此数据了解通过提高产出、增加功能和降低成本获得的收益。

不再把钱花在千篇一律的繁重工作上：AWS 会帮您料理繁重的数据中心运营工作，如安装、堆叠和驱动服务器。它还消除了使用托管服务管理操作系统和应用程序的运营负担。因此，您可以集中精力处理客户和业务项目而非 IT 基础设施。

分析并划分支出属性：使用云，可以更轻松地准确了解工作负载的成本和使用情况，从而将 IT 成本透明地归属到收入来源和各个工作负载拥有者。这有助于衡量投资回报率（ROI），并让工作负载拥有者能够据此优化资源和降低成本。

定义

云中的成本优化包括五个重点方面：

- 践行云财务管理

- 支出和使用情况意识
- 具有成本效益的资源
- 管理需求和供应资源
- 持续优化

与 Well-Architected Framework 中的其他支柱类似，考虑成本优化时也需要权衡取舍。例如，是要优化上市速度还是优化成本？ 在某些情况下，最好优化上市速度以便快速上市、交付新功能或按时完成任务，而不是优化预付成本。

设计决策有时是仓促作出的，并未结合数据进行考虑，而且人们往往会过度补偿，而不是花时间进行基准测试以确定最合算的部署。过度补偿可能会导致部署过度预置且优化不足。但是，如果您必须将资源从本地环境“直接迁移”到云，然后再进行优化，这可能也算合理之举。

通过预先在成本优化策略中投入适量的精力，可以确保始终如一地遵守最佳实践，避免不必要的过度预置，从而更轻松地实现云的经济优势。以下部分介绍了一些技巧和最佳实践，可帮助您开始并持续实施工作负载的云财务管理和成本优化。

践行云财务管理

管理云财务需要改进现有财务流程，针对您的 AWS 环境建立和运营成本透明化、控制、规划和优化。

在动态云使用场景中应用传统的静态瀑布式规划，以及 IT 预算和成本评估模型可能会带来风险，导致规划不准确和可见性降低。最终会导致丧失有效优化和控制成本，以及实现长期业务价值的机会。为了规避这些陷阱，需要在整个云之旅中积极管理成本，无论是在云中原生构建应用程序，将工作负载迁移到云，还是增加对云服务的采用。

云财务管理（CFM）允许财务、产品、技术和业务组织在增加 AWS 上的使用量和进行扩缩时管理、优化和规划成本。CFM 的主要目标是让客户能够以最具成本效益的方式实现其业务成果，并加速创造经济和业务价值，同时在敏捷性和控制力之间找到适当的平衡。

CFM 解决方案通过成本透明化、控制、预测和优化帮助实现业务转型。这些解决方案还有助于树立注重成本的企业文化，在所有团队和职能部门间形成问责制。财务团队可以看到成本的来源，以最少的意外费用执行运营，规划动态云使用，并节省云费用，同时团队可在云中扩缩其采用规模。与工程团队共享此信息可以为他们的资源选择、使用和优化提供必要的财务背景。

AWS CFM 提供了一组功能，可在保持业务敏捷性的同时，对云成本进行管理、优化和规划。CFM 不仅对有效管理成本至关重要，而且还能确保投资推动实现预期业务成果。以下是云财务管理框架的四大支柱AWS Cloud：查看、节省、计划和运营。这些支柱中的每一个都有一套活动和功能。

云财务管理的四大支柱。

- 查看：您目前如何衡量、监控云支出并为其建立问责制？ 如果您对 AWS 或使用 AWS 进行规划不熟悉，是否有建立成本节省和使用情况可见性的计划？

要了解您的 AWS 成本并优化支出，您需要了解这些成本的来源。这需要对您的账户和资源进行深思熟虑的结构设计，以帮助财务组织跟踪支出流，并确保团队对他们的利润部分负责。

AWS 服务：AWS Control Tower、AWS Organizations、成本分配标记、标记策略、AWS Resource Groups、AWS Cost Categories、AWS Cost Explorer、AWS 成本和使用情况报告、RI 和 SP

资源：AWS 标记最佳实践、AWS Cost Categories

- 节省：您目前正在使用哪些成本优化方法来优化支出？ 如果不使用 AWS，是否熟悉常见的基于使用情况和基于定价模式的优化？

  在节省原则中，我们通过定价和资源建议来优化成本。优化成本首先要为新的云运营模式定义良好的策略。理想情况下，应该在您的云之旅中尽早开展此行动，并通过正确的流程和行为为注重成本的企业文化奠定基础。

  可以采用多种不同的方法来优化云成本。其中一个方法是选择正确的采购模型（RI 和 SP），或者工作负载是否不可变且已容器化，以便可以采用 Amazon EC2 竞价型实例。此外，还可以使用 Amazon EC2 Auto Scaling 组来扩展工作负载。

  AWS 服务：RI 和 SP、Amazon EC2 Auto Scaling 组、竞价型实例

  资源：预留实例、节省计划、处理 Amazon EC2 的最佳实践

- 计划：您目前如何规划未来的云使用情况和支出？ 您是否有量化生成新迁移价值的方法？ 您是否改进了当前的预算和预测流程，以采用云的可变使用情况？

  计划原则意味着通过灵活的预算和预测来改进规划。在建立可见性和成本控制后，您可能需要规划和设定云项目支出的预期。AWS 可以帮助您灵活地构建动态预测和预算制定流程，以便您可以随时了解成本是否低于或超出预算限制。

  AWS 服务：AWS Cost Explorer、AWS 成本和使用情况报告、AWS Budgets

  资源：基于使用情况的预测、AWS Budgets 报告和警报

- 运营：您目前正在使用哪些运营流程和工具来管理云支出，谁在负责这些工作？ 您是否考虑过开始使用 AWS 后从日常运营角度将如何运作？

  运营原则实际上是管理计费和成本控制。可以建立防护机制并设置治理机制，确保支出与预算保持一致。AWS 提供了多种工具来帮助您入门。

  AWS 服务：AWS 账单与成本管理 控制台、AWS Identity and Access Management、服务控制策略（SCP）、AWS Service Catalog、AWS Cost Anomaly Detection、AWS Budgets

资源：AWS Billing 控制台入门

以下是云财务管理的最佳实践：

最佳实践

- COST01-BP01 确立成本优化的责任归属模式
- COST01-BP02 在财务和技术人员之间建立合作关系
- COST01-BP03 建立云预算和预测
- COST01-BP04 在组织流程中落实成本意识
- COST01-BP05 报告和通知成本优化
- COST01-BP06 主动监控成本
- COST01-BP07 及时了解新发布的服务
- COST01-BP08 建立对成本敏感的文化
- COST01-BP09 量化通过成本优化实现的业务价值

COST01-BP01 确立成本优化的责任归属模式

创建一个团队（云业务办公室、云卓越中心或 FinOps 团队），负责在整个组织内建立并维护成本意识。成本优化的负责人可以是了解整个组织和云财务的个人或团队（需要来自财务、技术和业务团队的人员）。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

这是云业务办公室（CBO）或云卓越中心（CCoE）部门或团队的介绍，此部门或团队负责建立并维护一种对云计算成本敏感的文化。此部门可以是一个现有的个人、组织内的一个团队，也可以是一个由整个组织的关键财务、技术和组织利益相关者组成的新团队。

此部门（个人或团队）会排定成本管理和成本优化活动的优先级，并根据需要为这些活动投入一定比率的时间。相对于较大型企业中的全职部门，小型组织的这一部门在此方面花费的时间可能更少。

此部门需要采取多学科方法，并具备项目管理、数据科学、财务分析和软件或基础设施开发的能力。此部门可通过三种不同的责任归属模式来执行成本优化，以提高工作负载的效率：

- 集中式：通过 FinOps 团队、云财务管理（CFM）团队、云业务办公室（CBO）或云卓越中心（CCoE）等指定团队，客户可以设计和实施治理机制，并在全公司范围内推广最佳实践。
- 分散式：由具有影响力的技术团队执行成本优化。
- 混合式：同时利用集中式和分散式团队，两者可以合作执行成本优化。

可以对照成本优化目标（例如工作负载效率指标）来衡量此部门的执行和交付能力。

您必须为此部门获得高管支持，这是取得成功的一个关键因素。支持者负责倡导注重成本效益的云消费理念，为成本优化团队提供升级上报支持，确保按组织确定的优先级开展成本优化活动。否则，相关方面会忽视指导意见，并且不会优先考虑节省成本的机会。支持者与团队共同确保组织有效利用云资源并创造业务价值。

如果您制定了商业、企业入门或企业支持计划，并需要帮助来建立此团队或部门，请通过您的客户团队联系云财务管理（CFM）专家。

实施步骤

- 明确主要成员：组织的所有相关部门都必须关注成本管理，并做出贡献。组织中的常见团队通常包括：财务、应用程序或产品负责人、管理和技术团队（DevOps）。有些是全职工作（财务或技术），有些则是根据需要定期工作。执行 CFM 的个人或团队需要掌握以下技能组合：
  - 软件开发：在构建脚本和自动化的情况下。
  - 基础设施工程：部署脚本，自动化流程，并了解如何预置服务或资源。
  - 运营敏锐性：CFM 的宗旨是通过衡量、监控、修改、规划和扩展对云的有效使用，在云上高效地运行。
- 明确目标和指标：此部门需要以不同方式为组织创造价值。相关目标已确定，并将随着组织的发展而不断完善。常见活动包括：在整个组织内创建并执行关于成本优化的培训计划、制定涉及整个组织的标准，例如监控和报告成本优化，以及设置关于优化的工作负载目标。此部门还需要定期向组织报告其成本优化能力。

  可以定义基于价值或成本的关键绩效指标（KPI）。定义 KPI 时，可以根据效率和预期业务成果计算预期成本。基于价值的 KPI 将成本和使用情况指标与业务价值驱动因素联系起来，有助于合理调整 AWS 支出。获得基于价值的 KPI 的第一步是跨组织协作，选择并商定一组标准 KPI。
- 建立定期沟通机制：该小组（财务、技术和业务团队）应该定期开会，以审查目标和指标。典型的定期沟通机制包括审查组织的状态、当前执行的任何计划以及整体财务和优化指标。随后，更详细地报告关键工作负载。

在这些定期审查中，可以审查工作负载效率（成本）和业务成果。例如，工作负载的成本增加 20% 可能与客户使用情况的增加保持一致。在这种情况下，这 20% 的成本增加可以理解为一项投资。这些定期沟通通话有助于团队确定价值 KPI，为整个组织带来意义。

资源

相关文档：

- AWS CCOE 博客
- 创建云业务办公室
- CCOE – 云卓越中心

相关视频：

- Vanguard CCOE 成功案例

相关示例：

- 使用云卓越中心（CCoE）实现整个企业转型
- 构建 CCOE 以实现整个企业转型
- 构建 CCOE 时应避免的 7 个陷阱

COST01-BP02 在财务和技术人员之间建立合作关系

在云之旅的所有阶段，都让财务和技术团队参与成本和使用情况的讨论。团队定期开会，讨论组织目标、成本和使用情况的当前状态以及财务和会计实务等主题。

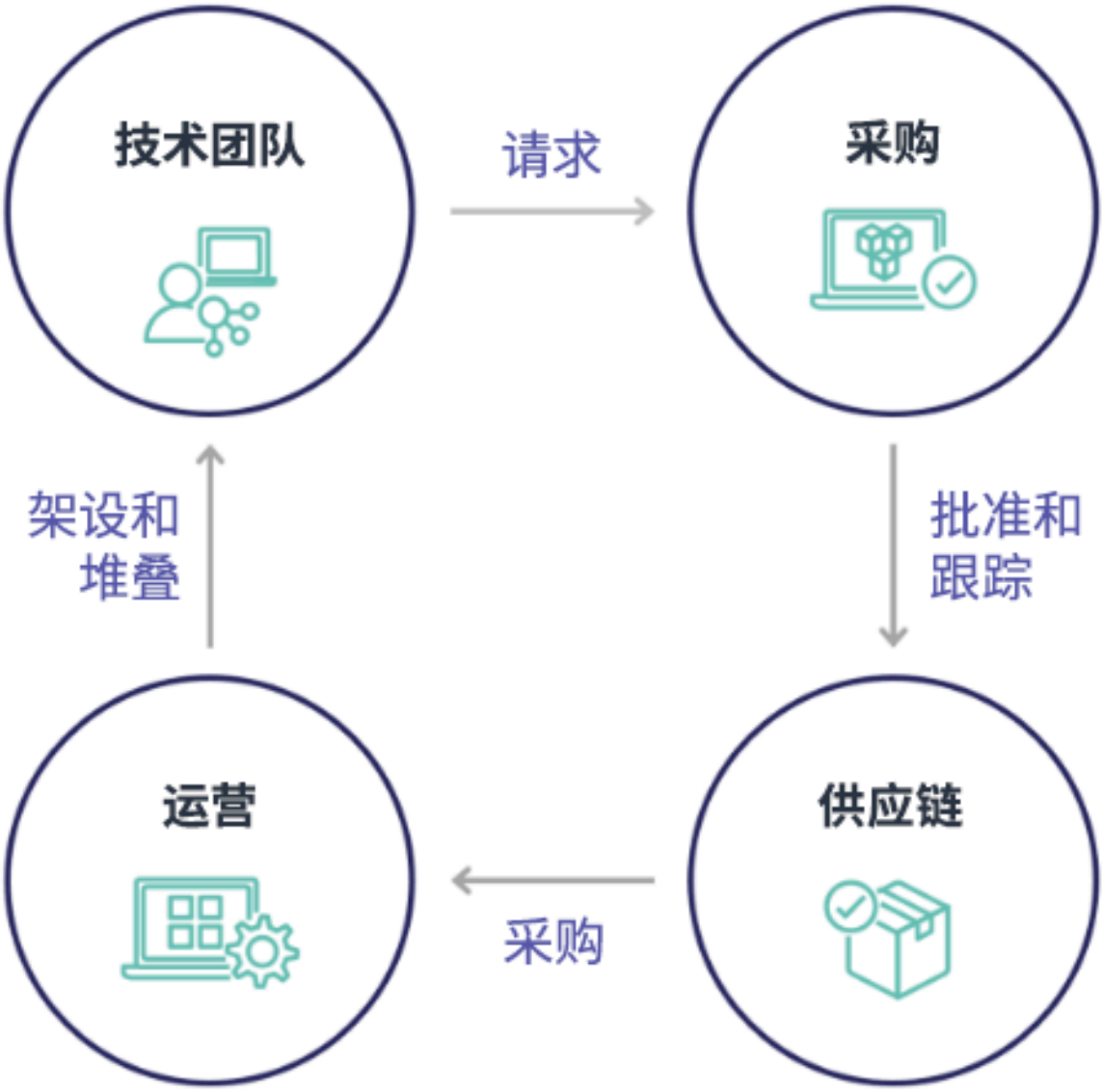
在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

由于缩短了审批、采购和基础设施部署周期，技术团队在云端的创新速度更快。这可能是对财务组织的一种调整，以前，他们习惯于运行耗时的资源密集型流程，以便在数据中心和本地环境中获取和部署资金，并且只在项目批准时进行成本分配。

从财务和采购组织的角度来看，资本预算编制、资本请求、审批、采购和安装物理基础设施这一流程是我们几十年来一直在学习和实现标准化的流程：

- 工程团队或 IT 团队通常是请求者
- 不同的财务团队充当审批者和采购者
- 运营团队架设、堆叠并移交可供使用的基础设施



随着云的采用，基础设施的采购和消费不再受制于一连串的依赖关系。在云模式下，技术和产品团队不再仅仅是构建者，还是产品的运营者和负责人，他们负责过去与财务和运营团队有关的大部分活动，包括采购和部署。

预置云资源真正需要的只是一个账户和一组正确的权限。这也有助于降低 IT 和财务风险，因为团队只需点击几下鼠标或进行几次 API 调用，就可以终止空闲或不必要的云资源。这也使技术团队能够更快创新，因为他们获得了启动实验以及之后拆除实验的敏捷性和能力。虽然从资本预算编制和预测的角度来看，云消费的可变性质可能会影响可预测性，但云为组织提供了降低过度预置成本的能力，以及降低与保守预置不足相关的机会成本的能力。

在关键的财务和技术利益相关方之间建立合作关系，让他们就组织目标达成共识，并开发在云计算的可变支出模型中获得财务成功的机制。组织内的相关团队必须在云之旅的各个阶段参与成本和使用情况讨论，包括：

- 财务领导：首席财务官、财务总监、财务规划师、业务分析师、采购、供应商开发人员和应付账款负责人必须了解消费、采购选项和每月开票流程的云模型。财务部门需要与技术团队合作创建有关 IT 价值的沟通内容并广泛传播，帮助业务团队了解技术支出与业务成果之间的联系。这样，技术支出就不会被视为成本，而会被视为投资。由于云运营（如使用情况的变化速率、即用即付定价模式、分级

定价、定价模式以及详细的计费和使用情况信息）与本地运营之间存在根本差异，财务组织必须了解云的使用对业务方面的影响，包括采购流程、激励跟踪、成本分配和财务报表。

- 技术领导：技术领导（包括产品和应用程序拥有者）必须了解财务要求（如预算约束）和业务要求（如服务水平协议），如此才能实施工作负载并实现组织的预期目标。

财务与技术人员的合作可带来以下好处：

- 财务和技术团队几乎可以实时看到成本和使用情况。
- 财务和技术团队建立了标准操作程序来处理云支出差异。
- 在如何使用资本购买承诺折扣（例如预留实例或 AWS 节省计划）以及如何使用云来发展组织方面，财务利益相关方充当战略顾问。
- 将现有的应付账款和采购流程用于云部署。
- 财务和技术团队协作预测未来的 AWS 成本和使用情况，以调整和建立组织预算。
- 通过共通的语言以及对财务概念的一致理解，更好地进行跨组织沟通。

组织中应参与成本和使用情况讨论的其他利益相关方包括：

- 业务部门负责人：业务部门负责人必须了解云业务模式，以便为业务部门和整个公司提供指导。在需要预测增长和工作负载使用情况，以及评测不同购买选项（例如预留实例或节省计划）时，这方面的云知识至关重要。
- 工程团队：在财务和技术团队之间建立合作关系对于建立对成本敏感的文化至关重要，这可以鼓励工程师在云财务管理（CFM）上采取行动。CFM 或财务运营从业者和财务团队的一个常见问题是让工程师了解云上的整个业务，遵循最佳实践，并采取建议的行动。
- 第三方：如果组织使用第三方（例如顾问或工具），请确保他们与您的财务目标一致，并且可以通过参与模式和投资回报（ROI）证明一致性。第三方通常会帮助报告和分析其管理的任何工作负载，并且提供他们设计的任何工作负载的成本分析。

要想实施 CFM 并取得成功，需要财务、技术和业务团队之间彼此协作，并在整个组织内就如何转变云支出进行沟通和评估。将工程团队包括在内，让他们在所有阶段参与这些成本和使用情况的讨论；还要鼓励他们遵循最佳实践并采取相应的商定行动。

实施步骤

- 明确主要成员：确认财务和技术团队的所有相关成员都参与合作。相关财务成员将是与云账单进行交互的人员。他们通常是首席财务官、财务总监、财务规划师、业务分析师和采购员。技术成员通常是

产品和应用程序负责人、技术经理和所有在云上执行构建的团队的代表。其他成员可能包括业务部门负责人（例如影响产品使用的营销部门）和第三方（例如顾问），以确保与您的目标和机制保持一致，并协助进行报告。

- 明确讨论主题：明确团队之间的共同主题，或者需要达成共识的主题。从生成成本之时跟踪成本，直到账单已付为止。请注意所涉及的任何成员，以及需要应用的组织流程。了解经过的每个步骤或流程以及相关信息，例如可用的定价模式、分级定价、折扣模型、预算编制和财务要求。
- 建立定期沟通机制：为了让财务和技术人员展开合作，建立定期沟通机制，以促进并保持一致。该小组需要定期开会，审查目标和指标。典型的定期沟通机制包括审查组织的状态、当前执行的任何计划以及整体财务和优化指标。然后，更详细地报告关键工作负载。

资源

相关文档：

- AWS 新闻博客

COST01-BP03 建立云预算和预测

调整现有的组织预算编制和预测流程，使之适应云成本和使用情况的易变特性。流程必须是动态的，可以使用基于趋势或基于业务驱动因素的算法，也可以将两者结合使用。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

在传统的本地 IT 设置中，客户在规划偶有变动的固定成本时，经常会面临挑战。一般情况下，客户会购买新的 IT 硬件和服务来满足高峰需求，而这就会带来成本变动挑战。相反，AWS Cloud 采用了另一种方法，使客户只需为在满足他们实际的 IT 和业务需求时使用的资源付费。在云环境中，需求会每月、每天甚至每小时发生波动。

借助云，客户可以提升效率、速度和敏捷性，最终打造高度可变的成本和使用模式。随着工作负载效率的提高，或者新工作负载和功能的部署，成本可能会降低，有时也会提高。当工作负载扩展以便为不断扩大的客户群提供服务时，云的使用情况和成本也会相应上升，这是因为资源可访问性得到了提高。云服务的这种灵活性会延伸到成本和预测，从而带来一定的弹性。

您必须将这些不断变化的业务需求和需求驱动因素紧密结合起来，并尽可能准确地规划。传统的组织预算编制流程需要作出调整来适应这种可变性。

在预测新工作负载的成本时，可以考虑使用成本建模。通过成本建模，可以基本了解预期的云成本，这有助于您执行总拥有成本（TCO）、投资回报率（ROI）和其他财务分析，与利益相关方一起设定目标和期望，并确定成本优化机会。

组织应了解成本定义和可接受的分组。做出的预测的详细程度可能会因组织的结构和内部工作流程而异。选择符合特定要求和组织机构特点的粒度。请务必了解预测是在哪个级别进行的：

- 管理账户或 AWS Organizations 级别：管理账户是用于创建 AWS Organizations 的账户。默认情况下，Organizations 有一个管理账户。
- 关联账户或成员账户：Organizations 中的账户是标准 AWS 账户，其中包含您的 AWS 资源以及可以访问这些资源的身份。
- 环境：环境是运行应用程序版本的 AWS 资源的集合。可以使用多个关联账户或成员账户来创建环境。
- 项目：项目是在固定期限内要完成的既定目标或任务的组合。在预测期间，请务必考虑项目生命周期。
- AWS 服务：组或类别，例如计算或存储服务，可以在其中对 AWS 服务进行分组以进行预测。
- 自定义分组：可以根据组织需求创建自定义组，例如业务部门、成本中心、团队、成本分配标签、成本类别、关联账户或这些元素的任意组合。

确定可能影响使用成本的业务驱动因素，并分别针对每个驱动因素进行预测，以便提前计算出预期使用量。其中一些驱动因素可能与组织中的 IT 和产品团队有关。其他业务驱动因素（如营销活动、促销、地域扩张、合并和收购）则是销售、营销和业务领导者所了解的，因此与他们合作，将所有这些需求驱动因素考虑在内也很重要。

可以根据历史支出，使用 AWS Cost Explorer 对确定的未来时间范围进行基于趋势的预测。AWS Cost Explorer 的预测引擎会根据付费类型（例如预留实例）对历史数据进行细分，并结合使用机器学习和基于规则的模型来分别预测所有付费类型的支出。

建立预测流程并构建模型后，可以使用 AWS Budgets 设置精细的自定义预算，即通过指定时间段、循环或金额（固定或可变），及添加筛选条件（如服务、AWS 区域 和标标签）来实现。预算通常针对一年进行编制，并且保持固定，这将要求所有相关人员严格遵守。相比之下，预测更为灵活，可以在全年重新调整，并提供一年、两年或三年的动态预测。在确定各技术和业务利益相关方的财务预期时，预算编制和预测都起着至关重要的作用。预测和实施要保持准确，还需要直接负责预置成本的利益相关方承担责任，这也能提高他们的整体成本意识。

为了随时了解现有预算的执行情况，可以创建 AWS Budgets 报告并安排好时间表，定期以电子邮件的形式发送给您和您的利益相关方。您还可以根据实际成本（本质上是反应式的）或预测成本（可以留出

时间缓解潜在的成本超支情况）创建 AWS Budgets 警报。您的成本或使用情况实际超出一定量，或预计将超出预算金额时，系统会向您发送警报。

使用基于趋势（将历史成本用作输入）的算法或者基于驱动因素（例如新产品发布、区域扩张或用于工作负载的新环境）的算法（适用于动态和可变支出环境），调整现有的预算编制和预测流程，使这些流程更为灵活。使用 Cost Explorer 或任何其他工具确定基于趋势的预测后，使用 AWS 定价计算器 即可估算您的 AWS 应用场景和基于预期使用情况（流量、每秒请求数或所需的 Amazon EC2 实例）的未来成本。

跟踪预测的准确性，因为预算的编制需要基于这些预测计算和估算值。监控集成云成本预测的准确性和有效性。定期针对预测审查实际支出，并根据需要作出调整来提高预测精度。跟踪预测差异，对报告的差异进行根本原因分析，以采取行动并调整预测。

如 COST01-BP02 在财务和技术人员之间建立合作关系 中所述，在 IT 部门、财务部门和其他利益相关方之间建立合作关系和沟通机制非常重要，以确保他们都使用相同的工具或流程来保持一致性。在预算可能需要更改的情况下，提高沟通频率，来更快地应对这些更改。

实施步骤

- 定义组织内的成本语言：在组织内创建一种具有多个维度和分组的通用 AWS 成本语言。确保利益相关方了解预测粒度、定价模式和成本预测级别。
- 分析基于趋势的预测：使用基于趋势的预测工具，例如 AWS Cost Explorer 和 Amazon Forecast。从服务、账户、标签和成本类别等多个维度分析使用成本。
- 分析基于驱动因素的预测：确定业务驱动因素对云使用情况的影响，并分别针对每个驱动因素进行预测，以预先计算预期的使用成本。与业务部门负责人和利益相关方密切合作，了解新驱动因素的影响，计算预期成本变化以确定准确的预算。
- 调整现有的预测和预算编制流程：根据所采用的预测方法（如基于趋势的预测方法、基于业务驱动因素的预测方法，或者结合使用两种预测方法），确定预测和预算编制流程。预算应经过计算得出且切合实际，并以预测为基础。
- 配置警报和通知：使用 AWS Budgets 警报和成本异常检测来获取警报和通知。
- 与关键利益相关方开展定期审查：例如，与 IT、财务、平台团队和其他业务领域的利益相关方就业务方向和使用情况方面的变化达成一致。

资源

相关文档：

- AWS Cost Explorer

- AWS 成本和使用情况报告
- 利用 Cost Explorer 进行预测
- QuickSight 预测
- AWS Budgets

相关视频：

- 如何使用 AWS Budgets 来跟踪我的支出和使用情况
- AWS 成本优化系列：AWS Budgets

相关示例：

- Understand and build driver-based forecasting
- How to establish and drive a forecasting culture
- How to improve your cloud cost forecasting
- Using the right tools for your cloud cost forecasting

COST01-BP04 在组织流程中落实成本意识

在影响使用情况的新流程或现有流程中落实成本意识、创建成本透明度和成本问责制，并利用现有流程提高成本意识。在员工培训中落实成本意识。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

必须在新的和现有的组织流程中落实成本意识。它是其他最佳实践的基础性先决条件之一。建议在可能的情况下重用和修改现有流程，这可以最大限度地减少对敏捷性和速度的影响。向技术团队以及业务和财务团队的决策者报告云成本，以提高成本意识，并为财务和业务利益相关方建立效率关键性能指标（KPI）。以下建议有助于在工作负载中落实成本意识：

- 确认变更管理包含成本度量，以量化变更对财务的影响。这有助于主动解决与成本相关的问题，并强调成本节省。
- 确认成本优化是运营能力的核心组成部分。例如，可以利用现有的事件管理流程来调查和确定成本及使用情况异常或成本超支的根本原因。

- 通过自动化或工具加快节省成本和实现业务价值。在考虑实施成本时，请在对话中加入投资回报（ROI）信息，以证明投入时间或资金的合理性。
- 通过对云支出（包括在基于承诺的购买选项、共享服务和市场购买上的支出）实施 对账或扣款来分配云成本，以推动极具成本意识的云消费。
- 扩展现有的培训和发展计划，在整个组织中开展成本意识培训。建议在其中加入持续的培训和认证。这样有助于建立一个能够自我管理成本和使用情况的组织。
- 利用免费的 AWS 原生工具，例如 AWS Cost Anomaly Detection、AWS Budgets 和 AWS Budgets Reports。

当组织持续采用云财务管理（CFM）实践时，这些行为就会在工作和决策过程中落地生根。最终带来的是在从构建新的云原生应用程序的开发人员，到分析这些新的云投资的 ROI 的财务经理之间建立起更具有成本意识的企业文化。

实施步骤

- 确定相关的组织流程：每个组织单位审核自己的流程，并确定影响成本和使用情况的流程。任何导致资源创建或终止的流程都需要进行审核。查找能够在企业中支持成本意识的流程，例如事件管理和培训。
- 建立可自我维持的对成本敏感的文化：确保所有相关的利益相关方都认同变更原因和影响是一种成本，这样他们就能理解云成本。这将使贵组织能够在创新方面建立一种自我维持的对成本敏感的文化。
- 将成本意识融入流程：对每个流程进行修改以提高成本意识。流程可能需要执行额外的预检查，例如评估成本的影响，或执行后期检查，以验证成本和使用情况是否发生了预期变化。可以扩展支持流程（如培训和事件管理），以包括成本和使用情况项目。

要获得帮助，请通过您的客户团队与 CFM 专家联系，或浏览以下资源和相关文档。

资源

相关文档：

- AWS 云财务管理

相关示例：

- 高效云成本管理战略

- 成本控制博客系列 3：如何应对成本冲击
- AWS Cost Management 初学者指南

COST01-BP05 报告和通知成本优化

设置云预算并配置用于检测异常使用情况的机制。配置相关工具，根据预定义目标发出成本和使用情况警报，并在任何使用情况超过这些目标时接收通知。定期召开会议，分析工作负载的成本效益，提高成本意识。

在未建立这种最佳实践的情况下暴露的风险等级：低

实施指导

必须定期报告组织内成本和使用情况的优化情况。可以安排专门的会议来讨论成本绩效，或者将成本优化纳入工作负载的常规运营报告周期。利用服务和工具定期监控成本绩效，并抓住机会实施成本节省措施。

使用 AWS Cost Explorer 按多种筛选条件和粒度查看成本和使用情况，此工具提供控制面板和报告，例如按服务或按账户统计的成本、每日成本或市场成本。通过 AWS Budgets Reports 跟踪成本和使用情况在所配置预算中的消耗进度。

使用 AWS Budgets 设置自定义预算以跟踪成本和使用情况，并在超过阈值时快速响应从电子邮件或 Amazon Simple Notification Service（Amazon SNS）通知收到的警报。将首选预算期设置为每日、每月、每季度或每年，并创建具体的预算限制，以随时了解实际或预测成本和使用情况相对于预算阈值的情况。还可以根据这些警报设置警报和操作，以便自动运行，或在超出预算目标时通过审批流程运行。

启用关于成本和使用情况的通知，确保在成本和使用情况发生意外变化时您能够迅速采取行动。AWS Cost Anomaly Detection 让您能够减少意外成本，加强控制，同时不放慢创新速度。AWS Cost Anomaly Detection 可识别异常支出并找出根本原因，这有助于降低出现账单意外的风险。只需简单三步，即可创建自己的情境化监控器，在检测到任何异常支出时接收警报。

还可以将 QuickSight 与 AWS 成本和使用情况报告（CUR）数据结合使用，以提供包含更精细数据的高度定制的报告。利用 QuickSight，可以安排报告，并定期接收关于历史成本和使用情况或成本节省机会的成本报告电子邮件。查看我们的 Cost Intelligence Dashboard（CID）解决方案，该解决方案基于 QuickSight 构建，可为您提供高级可见性。

使用 AWS Trusted Advisor，该工具可提供指导，以验证预置的资源是否符合 AWS 的成本优化最佳实践。

通过可视化图表，对照精细的成本和使用情况数据检查节省计划建议。每小时图表显示按需支出和建议的节省计划承诺，让用户深入了解估计的节省额、节省计划覆盖率和节省计划使用率。这有助于组织了解节省计划如何应用于每小时的支出，而不必投入时间和资源来构建支出分析模型。

定期创建报告，其中包含来自 AWS Cost Explorer 的节省计划、预留实例和 Amazon EC2 合理调整大小建议的提要，以开始降低与稳定状态工作负载、空闲和未充分利用的资源相关的成本。确定并收回与已部署资源的云浪费有关的支出。当创建的资源大小不正确，或者观察到不同于预期的使用模式时，就会发生云浪费。遵循 AWS 最佳实践以减少浪费，或者请客户团队和合作伙伴协助优化并节省您的云成本。

定期生成报告，为您的资源提供更好的购买选项，以降低工作负载的单位成本。诸如节省计划、预留实例或 Amazon EC2 竞价型实例等购买选项可为容错工作负载节省大量成本，并让利益相关方（业务负责人、财务和技术团队）能够参与有关这些承诺的讨论。

分享包含可能有助于降低云总拥有成本（TCO）的机会或新发布公告的报告。采用新的服务、区域、功能、解决方案或新的方式来进一步削减成本。

实施步骤

- 配置 AWS Budgets：在所有账户中为工作负载配置 AWS Budgets。通过使用标签设置账户总支出预算和工作负载预算。
  - Well-Architected Lab：成本和使用情况治理
- 报告成本优化：定期讨论和分析工作负载的效率。使用已确立的指标，报告实现的指标和实现成本。找出任何负面趋势，加以解决，并确定可以在整个组织中推广的正面趋势。报告的参与者应该包括应用程序团队和负责人、财务团队以及云支出方面的关键决策者的代表。

资源

相关文档：

- AWS Cost Explorer
- AWS Trusted Advisor
- AWS Budgets
- AWS 成本和使用情况报告
- AWS Budgets 最佳实践
- Amazon S3 分析

相关示例：

- 开始优化 AWS 云成本的关键方法

COST01-BP06 主动监控成本

利用工具和控制面板主动监控工作负载的成本。定期用已配置的工具或开箱即用的工具审核成本，不要只在收到通知时才查看成本和类别。主动监控和分析成本有助于确定正面趋势，让您能够在整个组织中推广这些趋势。

在未建立这种最佳实践的情况下暴露的风险等级：中

实施指导

建议在组织内部主动监控成本和使用情况，而不仅仅是在出现异常或意外时才监控。在整个办公室或工作环境中，一目了然的控制面板能确保关键人员可以访问所需信息，并凸显组织对成本优化的重视程度。通过可见的控制面板，可以积极推动成功的结果，并在整个组织中加以实施。

创建一个每日或经常性的例程，以使用 AWS Cost Explorer 或任何其他控制面板（如 Amazon QuickSight）来查看成本并主动分析。通过分组和筛选，在 AWS 账户级、工作负载级或特定 AWS 服务级分析 AWS 服务的使用情况和成本，并验证它们是否符合预期。使用小时级和资源级粒度和标签来筛选和识别主要资源所产生的成本。还可以使用 Cost Intelligence Dashboard（一种 Amazon QuickSight 解决方案，由 AWS 解决方案架构师构建）构建自己的报告，并将预算与实际成本和使用情况进行比较。

实施步骤

- 报告成本优化：定期讨论和分析工作负载的效率。使用已确立的指标，报告实现的指标和实现成本。找出任何负面趋势，加以解决，并确定正面趋势，以便在整个组织中推广。报告应该包括应用程序团队和负责人、财务团队和管理团队的代表。
- 为成本和使用情况创建并激活每日粒度 AWS Budgets，以便及时采取行动来防止任何潜在的成本超支情况：可以使用 AWS Budgets 配置警报通知，以便在任何预算类型超出预配置阈值时，都会收到通知。利用 AWS Budgets 的最佳方式是将预期成本和使用情况设置为您的限值，这样一来，任何超出预算的情况均视为超支。
- 创建 AWS Cost Anomaly Detection 用于成本监控：AWS Cost Anomaly Detection 使用先进的机器学习技术来识别异常支出和根本原因，以便您快速采取行动。可以利用它来配置成本监控，以定义您想要评估的支出部分（例如，各项 AWS 服务、成员账户、成本分配标签和成本类别），并设置何

时、何地以及如何接收警报通知。对于每个监控器，为业务负责人和技术团队附加多个警报订阅，包括每个订阅的名称、成本影响阈值和警报频率（单个警报、每日总结、每周总结）。

- 使用 AWS Cost Explorer 或将 AWS 成本和使用情况报告（CUR）数据与 Amazon QuickSight 控制面板集成，使组织的成本可视化：AWS Cost Explorer 有一个易于使用的界面，让您能够可视化、理解和管理随时间变化的 AWS 成本和使用情况。Cost Intelligence Dashboard 是一个可定制且可访问的控制面板，可帮助创建您自己的成本管理和优化工具的基础。

资源

相关文档：

- AWS Budgets
- AWS Cost Explorer
- 每日成本和使用情况预算
- AWS Cost Anomaly Detection

相关示例：

- AWS Cost Anomaly Detection 警报与 Slack 集成

COST01-BP07 及时了解新发布的服务

定期咨询专家或 AWS 合作伙伴，以便确定哪些服务和功能的成本更低。查看 AWS Blog 和其他信息源。

在未建立这种最佳实践的情况下暴露的风险等级：中

实施指导

AWS 不断添加新功能，以便您可以利用前沿技术来更快地进行试验和创新。您或许可以实施新的 AWS 服务和功能，以提高工作负载的成本效率。请定期查阅 AWS 成本管理、AWS News 博客、AWS 成本管理博客和 AWS 的新功能，了解有关新服务和功能发布的信息。“新增功能”博客文章简要概述了 AWS 发布的所有服务、功能和区域扩展公告。

实施步骤

- 订阅博客：转到 AWS Blog 页面，订阅“新增功能”博客和其他相关博客。您可以在通信偏好页面上用您的电子邮件地址进行注册。
- 订阅 AWS News：请定期查阅 AWS News 博客和 AWS 的新功能，以了解有关新服务和功能发布的信息。订阅 RSS 源，或通过电子邮件关注公告和发布的信息。
- 关注 AWS 降价：我们会定期对各项服务进行降价，这是 AWS 将我们的规模所带来的经济效益传递给客户的一种标准方式。截至 2023 年 9 月 20 日，AWS 自 2006 年以来已降价 134 次。如果您有任何因价格问题而悬而未决的业务决策，可在降价和新服务整合后再次考虑这些决策。可以了解以前的降价情况，包括 Amazon Elastic Compute Cloud（Amazon EC2）实例（在 AWS News 博客的降价类别中了解这些情况）。
- AWS 活动和交流会：参加当地的 AWS 峰会，以及与当地其他组织的任何当地交流会。如果您不能亲自参加，请尝试参加虚拟活动，从 AWS 专家和其他客户的业务案例中了解更多信息。
- 与客户团队交流：定期与您的客户团队交流，讨论行业趋势和 AWS 服务。与您的客户经理、解决方案架构师和支持团队沟通。

资源

相关文档：

- AWS 成本管理
- AWS 新增功能
- AWS 新闻博客

相关示例：

- Amazon EC2 – 15 年来为您优化和节省 IT 成本
- AWS News 博客 – 降价

COST01-BP08 建立对成本敏感的文化

在整个组织中实施更改或计划，以建立对成本敏感的文化。建议先从小范围着手，然后随着能力的增强和组织对云的使用的增加，在更广泛的范围实施更大型的计划。

在未建立这种最佳实践的情况下暴露的风险等级：低

实施指导

在对成本敏感的文化中，可以在整个组织中以有机和分散的方式执行最佳实践，从而扩展成本优化和云财务管理（财务运营、云卓越中心、云运营团队等）。与严格的自上而下的集中式方法相比，只需要很少的工作，成本意识就能让您在整个组织中培养起较高的能力水平。

建立云计算成本意识（尤其对于云计算中的主要成本驱动因素）可以让团队了解成本角度的任何更改的预期结果。访问云环境的团队应了解定价模式，以及传统的本地数据中心与云计算之间的区别。

对成本敏感的文化的主要好处是，技术团队会主动且持续地优化成本（例如，在构建新工作负载或更改现有工作负载时，优化成本会被视为一项非功能性需求），而不是根据需要被动地进行成本优化。

文化上的细微改变可对您当前和将来的工作负载的效率产生重大影响。这种情况的示例包括：

- 在工程团队中提供可见性并提高意识，以了解他们所做的事情，以及他们在成本方面的影响。
- 以游戏方式展示整个组织的成本和使用情况。这可以通过一个公开可见的控制面板或比较各个团队的规范化成本和使用情况的报告（例如，每个工作负载的成本和每笔交易的成本）来完成。
- 认可成本效率。公开或私下奖励自愿或主动实现的成本优化成果，并从错误中吸取教训，以免日后重蹈覆辙。
- 创建自上而下的组织要求，确保工作负载按预定义的预算运行。
- 询问变更的业务要求，以及所请求的架构基础设施或工作负载配置变更的成本影响，以确保您只需为所需的内容付费。
- 确保变更规划者了解对成本有影响的预期变更，并确保这些变更得到利益相关方的确认，以经济高效的方式实现业务成果。

实施步骤

- 向技术团队报告云成本：提高成本意识，为财务和业务利益相关方建立效率 KPI。
- 将计划变更告知利益相关方或团队成员：创建一个议程项目，在每周变更会议期间讨论计划变更及其对工作负载的成本效益影响。
- 与客户团队交流：建立与客户团队定期沟通的机制，讨论行业趋势和 AWS 服务。与您的客户经理、架构师和支持团队沟通。
- 分享成功案例：分享关于任何工作负载、AWS 账户 或组织的成本降低的成功案例，以便围绕成本优化树立积极的态度并给予鼓励。
- 培训：确保技术团队或团队成员接受了关于 AWS Cloud 资源成本意识方面的培训。
- AWS 活动和交流会：参加当地的 AWS 峰会，以及与当地其他组织的任何当地交流会。

- 订阅博客：转到 AWS Blog 页面，订阅新增功能博客和其他相关博客，以关注 AWS 分享的新发布、实施、示例和更改。

资源

相关文档：

- AWS 博客
- AWS 成本管理
- AWS 新闻博客

相关示例：

- AWS 云财务管理

COST01-BP09 量化通过成本优化实现的业务价值

通过量化成本优化带来的业务价值，可以了解组织获得的全部益处。由于成本优化是一项必要的投资，因此量化业务价值之后，您就可以向利益相关方说明投资回报。如果能够量化业务价值，在未来的成本优化投资中，就可以从利益相关方那里得到更多支持，并获得一个框架来衡量组织成本优化活动的成果。

在未建立这种最佳实践的情况下暴露的风险等级：中

实施指导

量化业务价值是指衡量企业从他们所采取的行动和作出的决策中获得的益处。业务价值可以是有形的（例如减少支出或增加利润），也可以是无形的（例如改善品牌声誉或提高客户满意度）。

量化成本优化带来的业务价值，意味着要确定您从提高支出效率的努力中获得了多少价值或益处。例如，如果一家公司花费 10 万美元在 AWS 上部署工作负载，然后对其进行优化，则在不牺牲质量或产出的情况下，新成本仅为 8 万美元。在这种情况下，成本优化带来的量化业务价值是节省 2 万美元。但是，除了节约成本外，企业还可以从缩短交付时间、提高客户满意度或成本优化工作带来的其他指标等方面量化价值。利益相关方需要就成本优化的潜在价值、优化工作负载的成本和回报价值作出决定。

除了报告通过成本优化节省的费用外，建议量化其创造的附加价值。成本优化的益处通常以每项业务成果降低的成本来量化。例如，当您购买节省计划时，可以量化 Amazon Elastic Compute

Cloud（Amazon EC2）节省的成本，这可以降低成本并维持工作负载输出水平。当闲置的 Amazon EC2 实例移除，或者未连接的 Amazon Elastic Block Store（Amazon EBS）卷删除时，可以量化削减的 AWS 支出成本。

然而，成本优化带来的益处绝不仅仅在于降低或规避成本。考虑捕获其他数据来衡量效率提升值和业务价值。

实施步骤

- 评估业务效益：这是分析和调整 AWS Cloud 成本的过程，其方式可以最大限度地提高从花费的每一美元中获得的效益。与其专注于降低成本而忽略业务价值，不如考虑业务效益和成本优化的投资回报，这样可能会让您所花的钱产生更大价值。这就是要明智地花钱，在能产生最大回报的领域进行投资和支出。
- 分析预测 AWS 成本：预测可帮助财务利益相关方与其他内部和外部组织利益相关方设定预期，并可以提高组织的财务可预测性。AWS Cost Explorer 可用于预测成本和使用情况。

资源

相关文档：

- AWS Cloud 经济性
- AWS 博客
- AWS 成本管理
- AWS 新闻博客
- Well-Architected 可靠性支柱白皮书
- AWS Cost Explorer

相关视频：

- 在 AWS 上使用 Windows 挖掘业务价值

相关示例：

- 衡量和最大限度提升 Customer 360 的业务价值
- 采用亚马逊云科技托管式数据库的业务价值
- 亚马逊云科技为独立软件供应商带来的业务价值

- 云现代化的业务价值
- 向亚马逊云科技迁移的业务价值

支出和使用情况意识

了解组织的成本和驱动因素，对于有效管理成本和使用情况以及找到降低成本的机会至关重要。在组织中，通常会有多个团队运行多个工作负载。这些团队可能在不同的组织单位，每个单位都有自己的收入来源。将资源成本分摊到工作负载、各个组织或产品拥有者可以推动更高效的资源使用模式，减少浪费。准确的成本和使用情况监控能够帮助您了解各部门和产品如何盈利，并让您能够针对组织内的资源分配做出更明智的决策。组织中各层级的人员都了解使用情况是推动变化的关键，因为使用情况变化会导致成本变化。

考虑采用多元方法来了解使用情况和支出情况。团队必须收集数据、进行分析，然后报告。要考虑的关键因素包括：

主题

- 治理
- 监控成本和使用情况
- 停用资源

治理

为了管理云中的成本，您必须通过以下治理领域来管理使用情况：

最佳实践

- COST02-BP01 根据组织的要求制定各种策略
- COST02-BP02 实施方向性目标和执行性目标
- COST02-BP03 实施账户结构
- COST02-BP04 实施组和角色
- COST02-BP05 实施成本控制
- COST02-BP06 跟踪项目生命周期

COST02-BP01 根据组织的要求制定各种策略

制定策略，确定组织应该如何管理资源，并定期执行检查。策略应该涵盖资源和工作负载的成本，包括在资源生命周期内的创建、修改和停用。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

了解组织的成本和驱动因素，对于有效管理成本和使用情况以及找到降低成本的机会至关重要。在组织中，通常会有多个团队运行多个工作负载。这些团队可能在不同的组织单位，每个单位都有自己的收入来源。将资源成本分摊到工作负载、各个组织或产品负责人，这样既推动更高效的资源使用模式，又能减少浪费。准确的成本和使用情况监控有助于您了解如何优化工作负载，以及各部门和产品的盈利能力。利用这些知识，可以就在组织内的何处分配资源作出更明智的决策。组织中各层级的人员都了解使用情况是推动变化的关键，因为使用情况变化会导致成本变化。考虑采用多元方法来了解使用情况和支出情况。

执行治理的第一步是按照组织要求来针对云的使用制定策略。这些策略定义组织如何使用云以及如何管理资源。策略应涵盖与成本或使用情况有关的资源和工作负载的所有方面，包括资源生命周期内的创建、修改和停用。确认云环境中的任何更改都是遵循相应策略和程序实施的。在 IT 变更管理会议期间，提出问题以了解计划变更的成本影响（无论是增加还是减少）、业务理由以及预期结果。

策略应该简单易懂，能够在整个组织中有效实施。策略还需要易于遵守和解释（以便落实），并且要具体（不会在各团队之间造成误解）。此外，需要定期审查这些策略（例如我们的机制），并在客户业务状况或业务优先级发生变化时及时更新策略，以免导致策略过时。

从广泛的、高层级的策略开始，例如使用哪个地理区域，或者一天中应该运行资源的时间。逐步为各组织单位和工作负载细化策略。常见策略包括可以使用哪些服务和功能（例如，测试和开发环境中性能较低的存储），哪些类型的资源可供不同团队使用（例如，开发账户中最大规模的资源为中等大小），以及这些资源将使用多长时间（可能是临时使用、短期使用或在特定时间段内使用）。

策略示例

以下是侧重于成本优化的策略示例，可以参考该策略来创建自己的云治理策略。请确保根据组织和利益相关方的要求调整策略。

- 策略名称：定义清晰的策略名称，例如“资源优化和成本削减策略”。
- 目的：解释为什么应使用此策略以及预期结果如何。此策略的目标是确认，在为了满足业务要求而部署和运行所需工作负载方面，有最低的成本要求。
- 范围：明确定义谁应使用此策略以及何时应使用此策略，例如 DevOps X 团队，为美国东部的客户在 X 环境（生产或非生产）中使用此策略。

策略声明

1. 根据工作负载的环境和业务要求（开发、用户验收测试、预生产或生产），选择 us-east-1 或多个美国东部区域。
2. 安排 Amazon EC2 和 Amazon RDS 实例在早上 6 点到晚上 8 点 [美国东部标准时间（EST）] 之间运行。
3. 在处于不活动状态 8 小时之后，停止所有未使用的 Amazon EC2 实例，在处于不活动状态 24 小时之后，停止未使用的 Amazon RDS 实例。
4. 在非生产环境中处于不活动状态 24 小时之后，终止所有未使用的 Amazon EC2 实例。提醒 Amazon EC2 实例拥有者（根据标签）查看其生产环境中已停止的 Amazon EC2 实例，并告知他们，如果其 Amazon EC2 实例在 72 小时内未使用，将会被终止。
5. 使用通用实例系列和大小（如 m5.large），然后使用 AWS Compute Optimizer，根据 CPU 和内存利用率调整实例大小。
6. 优先使用自动扩缩，根据流量动态调整运行的实例数量。
7. 为非关键工作负载使用竞价型实例。
8. 查看容量需求，为可预测的工作负载使用节省计划或预留实例，并通知云财务管理团队。
9. 使用 Amazon S3 生命周期策略，将不经常访问的数据移动到更便宜的存储层。如果未定义保留策略，请使用 Amazon S3 Intelligent Tiering，自动将对象移动到存档层。
10. 使用 Amazon CloudWatch 监控资源利用率并设置警报来触发扩展事件。
11. 对于每个 AWS 账户，使用 AWS Budgets，根据成本中心和业务单位为您的账户设置成本和使用情况预算。
12. 使用 AWS Budgets 为账户设置成本和使用情况预算，有助于您控制支出和避免意外账单，从而更好地控制成本。

程序：提供实施此策略的详细过程，或参考介绍如何实施各个策略声明的其他文档。此部分应提供关于如何执行策略要求的分步说明。

要实施此策略，可以使用各种第三方工具或 AWS Config 规则来检查是否符合策略声明，并使用 AWS Lambda 函数触发自动修复操作。您也可以使用 AWS Organizations 来强制执行策略。此外，您应定期查看资源使用情况，并在必要时调整策略，以确认策略继续满足业务需求。

实施步骤

- 与利益相关方交流：要制定策略，让组织内的利益相关方（云业务办公室、工程师或实施策略的职能部门决策者）详细说明他们的要求并记录下来。采用迭代方法，首先大致进行，然后在每一步中不断细化到最小单元。团队成员包括与工作负载切身相关的人员（例如组织单位或应用程序负责人）以及支持小组（例如安全和财务团队）。

- 获得确认：确保团队就哪些人可以访问和部署到 AWS Cloud 的策略达成一致。确保这些人遵守组织的策略，并确认其资源创建符合商定的策略和程序。
- 创建入职培训课程：要求新组织成员完成入职培训课程，以建立成本意识并了解组织要求。他们可能会采取不同于以往经验的策略，或者根本不考虑这些策略。
- 定义运行工作负载的位置：定义工作负载的运行位置，包括国家/地区以及国家/地区中的区域。此信息用于映射到 AWS 区域 和可用区。
- 定义服务和资源并对其进行分组：定义工作负载所需的服务。对于每项服务，指定类型、大小和所需资源数量。按职能定义资源组，如应用程序服务器或数据库存储。资源可属于多个组。
- 定义用户并按职能对其进行分组：定义与工作负载交互的用户，侧重于用户的工作范畴及其使用工作负载的方式，而不是侧重于他们的身份或其在组织中的职位。将类似用户或职能分组在一起。可以使用 AWS 托管式策略作为指南。
- 定义操作：使用前面确定的位置、资源和用户，定义每项在其生命周期（开发、运行和停用）内实现工作负载成果所需的操作。根据每个位置的组（而不是组中的个别元素）确定操作。首先广泛读写，然后细化到每项服务的具体操作。
- 定义审核期：工作负载和组织要求可能会随着时间的推移而发生变化。定义工作负载审核计划，确保其与组织优先事项保持一致。
- 记录策略：确认已定义的策略是否可按组织的要求访问。这些策略用于实施、维护和审核对环境的访问。

资源

相关文档：

- 云中的变更管理
- 针对工作职能的 AWS 托管式策略
- AWS 多账户计费策略
- AWS 服务的操作、资源和条件键
- AWS 管理和治理
- 使用 IAM 策略控制对 AWS 区域 的访问
- 全球基础设施区域和可用区

相关视频：

- 大规模的 AWS 管理和治理

COST02-BP02 实施方向性目标和执行性目标

实施工作负载的成本和使用情况方向性目标和执行性目标。方向性目标为组织在预期结果方面指明了方向，而执行性目标则提供了要为工作负载实现的具体可衡量结果。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

为组织制定成本和使用情况方向性目标和执行性目标。作为一个在 AWS 上不断发展壮大的组织，针对成本优化设定方向性目标并进行跟踪，这一点非常重要。这些目标或关键绩效指标（KPI）可以是按需支出百分比等内容，也可以是采用某些优化服务（比如 AWS Graviton 实例或 gp3 EBS 卷类型）。设定可衡量且可实现的方向性目标，有助于您衡量效率的提高情况，这对于业务运营不可或缺。方向性目标为组织提供有关预期结果的指引和方向。

执行性目标则明确要实现的具体可衡量的结果。简言之，方向性目标是指您前进的方向，而执行性目标就是在这个方向上的进展情况，以及应何时实现这个目标（使用具体、可衡量、可分配、切合实际、适时的指导原则，即 SMART 指导原则）。方向性目标的一个示例是：在略微（非线性）增加成本的情况下，显著提升平台使用量。执行性目标的一个示例是：在成本增长不到 5% 的情况下，将平台使用量提升 20%。另一个常见的方向性目标是每 6 个月提高一次工作负载的效率。相应的执行性目标则是，需要每 6 个月将各项业务指标的成本缩减 5%。使用正确的指标，合理计算来为组织设定 KPI。可以从基本 KPI 开始，以后再根据业务需求进行改进。

就成本优化而言，其方向性目标是提高工作负载效率，这对应于逐步降低每项业务成果的工作负载成本。为所有工作负载实施此方向性目标，并设定执行性目标，例如每 6 至 12 个月将效率提高 5%。在云端，可以通过培养成本优化能力以及发布新服务和功能来做到这一点。

执行性目标是您为实现方向性目标而要达到的可量化基准，这些基准将您的实际结果与执行性目标进行比较。使用 KPI 为计算服务（例如竞价型实例采用、Graviton 采用、最新实例类型和按需覆盖范围）、存储服务（例如 EBS GP3 采用、过时的 EBS 快照和 Amazon S3 Standard 存储）或数据库服务使用情况（例如 RDS 开源引擎、Graviton 采用和按需覆盖范围）的单位成本建立基准。这些基准和 KPI 有助于验证您是否通过最经济高效的方式使用 AWS 服务。

下表提供了标准 AWS 指标列表以供参考。每个组织都可以为这些 KPI 设定不同的执行性目标值。

类别	KPI	描述
计算	EC2 使用情况覆盖率	使用 SP+RI+Spot 的 EC2 实例（以成本或小时计）与 EC2 实

类别	KPI	描述
		例的总计（以成本或小时计）的比较
计算	计算 SP/RI 使用率	使用的 SP 或 RI 小时数与可用 SP 或 RI 总小时数的比较
计算	EC2/小时成本	EC2 成本除以该小时内运行的 RDS 实例数量
计算	vCPU 成本	所有实例每个 vCPU 的成本
计算	最新一代实例	Graviton（或其他现代实例类型）上的实例百分比
数据库	RDS 覆盖率	使用 RI 的 RDS 实例（以成本或小时计）与 RDS 实例的总计（以成本或小时计）的比较
数据库	RDS 使用率	使用的 RI 小时数与可用 RI 总小时数的比较
数据库	RDS 正常运行时间	RDS 成本除以该小时内运行的 RDS 实例数量
数据库	最新一代实例	Graviton（或其他现代实例类型）上的实例百分比
存储	存储使用率	优化的存储成本（例如 Glacier、Deep Archive 或不频繁访问）除以总存储成本

类别	KPI	描述
标记	未标记的资源	Cost Explorer： 1. 筛选出服务抵扣金、折扣、税款、退款、商城，并复制最新的月度成本 2. 在 Cost Explorer 中选择仅显示未标记的资源 3. 将未标记资源中的数字除以每月成本。

此表包含执行性目标值或基准值，这些值应根据您的组织目标计算得出。您需要衡量业务的某些指标，并了解该工作负载的业务成果，以便确立准确、切合实际的 KPI。在评估组织内部的绩效指标时，要区分用于不同目的的各类指标。这些指标主要衡量技术基础设施的性能和效率，而不是直接衡量总体业务影响。例如，它们可能会跟踪服务器响应时间、网络延迟或系统正常运行时间。要评测基础设施能否有效支持组织的技术运营，这些指标必不可少。但是，您无法通过它们直接了解更广泛的业务目标，例如客户满意度、收入增长或市场份额。要全面了解业务绩效，请使用与业务成果直接相关的战略业务指标，补充这些效率指标。

您需要能够近乎实时地监控 KPI 及相关的成本节省机会，并不断跟踪进度。要着手定义和跟踪 KPI 目标，我们建议使用 Cloud Intelligence Dashboard（CID）中的 KPI 控制面板。根据来自成本和使用情况报告（CUR）的数据，KPI 控制面板提供了一系列建议的成本优化 KPI，能够设定自定义的方向性目标并不断跟踪进度。

如果通过其他解决方案来设定和跟踪 KPI 方向性目标，请务必让组织中的所有云财务管理利益相关方都采用这些方法。

实施步骤

- 定义预期使用量：首先，重点关注使用水平。与应用程序负责人、营销团队和更大的业务团队交流，了解工作负载的预期使用水平。客户需求可能如何随时间而变化？会因季节性增长或营销活动发生哪些变化？
- 定义工作负载资源和成本:定义使用水平后，量化满足这些使用量所需的工作负载资源变化。您可能需要增加工作负载组件的资源大小或数量，增加数据传输，或者在特定级别将工作负载组件更改为不同的服务。明确每个关键点的成本，并预测使用情况发生变化时的成本变化。

- 定义业务目标：从预期的使用情况和成本变化中获取输出，将其与预期的技术变化或正在开展的任何计划相结合，制定工作负载目标。方向性目标必须阐明使用情况和成本，以及两者之间的关系。方向性目标必须简明扼要，并能让人们了解企业对结果的期望（例如，确保未使用的资源保持在一定的成本水平以下）。您不需要为每种未使用的资源类型定义方向性目标，也不需要定义会导致方向性目标和执行性目标损失的成本。确认制定有组织计划（例如培训和教育等能力培养计划），以防成本呈预期变化，而使用情况无变化。
- 定义执行性目标：对于定义的每个方向性目标，指定一个可衡量的执行性目标。如果方向性目标是提高工作负载的效率，则执行性目标将量化改进量（通常为每一美元支出的业务产出）及获益时间。例如，可以设定一个方向性目标，最大限度地减少因过度预置而造成的浪费。有了此方向性目标，执行性目标可以是，因第一层生产工作负载中的计算过度预置而造成的浪费不应超过层级计算成本的 10%。此外，第二个执行性目标可以是，因第二层生产工作负载中的计算过度预置而造成的浪费不应超过层级计算成本的 5%。

资源

相关文档：

- 针对工作职能的 AWS 托管式策略
- AWS 多账户计费策略
- 使用 IAM 策略控制对 AWS 区域 的访问
- S.M.A.R.T. 目标
- How to track your cost optimization KPIs with the CID KPI Dashboard

相关视频：

- Well-Architected Lab：方向性目标和执行性目标（第 100 级）

相关示例：

- 什么是单位指标？
- 选择单位指标来支持您的业务
- 实践中的单位指标 – 经验教训
- 单位指标如何有助于在业务职能之间建立一致性

COST02-BP03 实施账户结构

实施与组织对应的账户结构。这有助于在整个组织内分摊和管理成本。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

AWS Organizations 允许创建多个 AWS 账户，这有助于在扩展 AWS 上的工作负载时集中管理环境。可以通过在组织单位（OU）结构中分组 AWS 账户，并在每个 OU 下创建多个 AWS 账户，对组织层次结构进行建模。要创建账户结构，首先需要确定哪个 AWS 账户 将成为管理账户。之后，可以按照管理账户最佳实践和成员账户最佳实践，根据设计的账户结构创建新 AWS 账户 或选择现有账户作为成员账户。

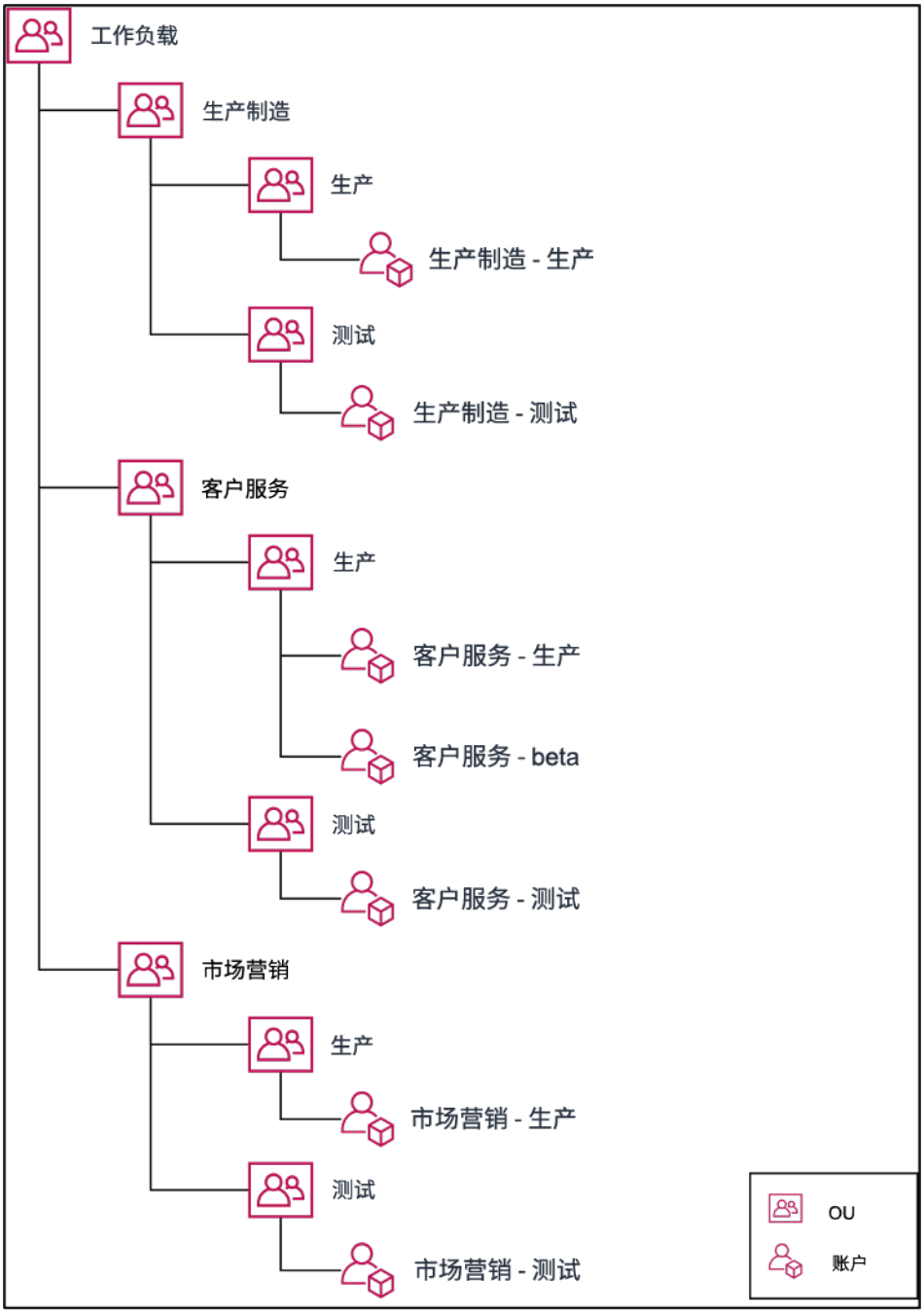
建议无论组织规模或使用情况如何，始终至少有一个管理账户和一个与之链接的成员账户。所有工作负载资源应仅驻留在成员账户中，不应在管理账户中创建任何资源。对于您应该拥有多少个 AWS 账户这一问题，没有标准答案。评测当前和未来的运营和成本模型，确保 AWS 账户 结构反映了组织的目标。有些公司出于业务原因会创建多个 AWS 账户，例如：

- 需要在组织部门、成本中心或特定工作负载之间实施管理或财务和计费隔离。
- AWS 服务限制设置为针对特定工作负载。
- 必须对工作负载和资源进行隔离和分离。

在 AWS Organizations 内，整合账单会在一个或多个成员账户与管理账户之间创建结构。通过成员账户，可以按组隔离和区分成本和使用情况。常见做法是每个组织部门（如财务、营销和销售）、每个环境生命周期（如开发、测试和生产）或每个工作负载（工作负载 a、b 和 c）具有单独的成员账户，然后使用整合账单将这些关联账户汇总在一起。

通过整合账单，可以将多个成员 AWS 账户 的付款整合至一个管理账户下，同时仍可查看每个关联账户的活动。由于成本和使用情况在管理账户中汇总，因此，可以最大限度地提高服务量折扣，并最大限度地利用承诺折扣（节省计划和预留实例）来获得最高折扣。

下图显示了如何对组织单位（OU）使用 AWS Organizations，以对多个账户进行分组，并在每个 OU 下放置多个 AWS 账户。建议将 OU 用于各种应用场景和工作负载，这提供了组织账户的模式。



在组织单位下将多个 AWS 账户 分组的示例。

AWS Control Tower 可以快速设置和配置多个 AWS 账户，确保治理符合组织要求。

实施步骤

- 定义分离要求：分离要求涉及多项因素，包括安全性、可靠性和财务结构。按顺序阐明每项因素，并详细说明工作负载或工作负载环境是否应与其他工作负载分开。安全性有助于满足访问和数据要求。可靠性管理限制，以便环境和工作负载不会影响其他项。定期查看 Well-Architected Framework 的

安全性和可靠性支柱，并遵循提供的最佳实践。财务结构创建严格的财务分离（不同的成本中心、工作负载所有权和问责制）。常见的分离示例包括在单独的账户中运行生产和测试工作负载，或使用单独的账户，以便可以将发票和账单数据提供给组织中的单个业务单位或部门，或拥有该账户的利益相关方。

- 定义分组要求：分组要求并不覆盖分离要求，而是用于协助管理。将无需分离的类似环境或工作负载分组在一起。例如，将一个或多个工作负载的多个测试或开发环境分组在一起。
- 定义账户结构：使用这些分离和分组，为每个组指定一个账户，并确保持续满足分离要求。这些账户有成员账户或关联账户。通过将这些成员账户分组到一个管理账户或付款人账户下，可以合并使用量，从而可以跨所有账户享有更大的批量折扣，这为所有账户提供一个账单。可以分离账单数据，并为每个成员账户提供其账单数据的单独视图。如果成员账户不能让任何其他账户看到自己的使用情况或账单数据，或者，如果需要 AWS 提供单独的账单，请定义多个管理账户或付款人账户。在这种情况下，每个成员账户都有自己的管理账户或付款人账户。资源应始终放置在成员账户或关联账户中。管理账户或付款人账户应只用于管理。

资源

相关文档：

- 使用成本分配标签
- 针对工作职能的 AWS 托管式策略
- AWS 多账户计费策略
- 使用 IAM 策略控制对 AWS 区域 的访问
- AWS Control Tower
- AWS Organizations
- 管理账户和成员账户的最佳实践
- 使用多个账户整理您的 AWS 环境
- 开启共享的预留实例和节省计划折扣
- 整合账单
- 整合账单

相关示例：

- 拆分 CUR 并共享访问权限

相关视频：

- AWS Organizations 简介
- 为 AWS Organizations 设置使用最佳实践的多账户 AWS 环境

相关示例：

- 为电信公司定义 AWS 多账户策略
- Best Practices for Optimizing AWS 账户
- Best Practices for Organizational Units with AWS Organizations

COST02-BP04 实施组和角色

实施与策略一致的组和角色，控制每个组中谁可以创建、修改或停用实例和资源。例如，实施开发组、测试组和生产组。这适用于 AWS 服务和第三方解决方案。

在未建立这种最佳实践的情况下暴露的风险等级：低

实施指导

用户角色和组是设计和实施安全、高效的系统的基本构件。角色和组有助于组织在控制力需求与灵活性和生产率要求之间取得平衡，最终满足组织目标和用户需求。正如 AWS Well-Architected Framework 安全性支柱的身份与访问管理部分所建议的那样，您需要强大的身份管理和权限，以便在适当的条件下为合适的人员提供对正确资源的访问权限。用户仅获得完成任务所需的访问权限。这可最大限度地降低与未经授权访问或滥用相关的风险。

制定策略后，可以在组织内创建逻辑组和用户角色。这让您可以分配权限、控制使用情况，并有助于实施强健的访问控制机制，防止未经授权访问敏感信息。从大致的人员分组开始，这通常与组织单位和岗位角色（例如，IT 部门的系统管理员、财务主管或业务分析师）对应。这些组将执行相似任务并需要相似访问权限的人员划分在一起。角色定义组必须做什么。管理组和角色的权限比管理单个用户的权限更容易。角色和组可一致且系统性地为所有用户分配权限，防止出现错误和不一致。

当用户的角色发生变化时，管理员可以在角色或组级别上调整访问权限，而不是重新配置单个用户账户。例如，IT 部门的系统管理员需要创建所有资源的权限，而分析团队成员仅需要创建分析资源。

实施步骤

- 实施组：如有必要，请使用组织策略中定义的用户组实施相应的组。有关用户、组和身份验证的最佳实践，请参阅 AWS Well-Architected Framework 的安全性支柱。

- 实施角色和策略：使用组织策略中定义的操作，创建所需的角色和访问策略。有关角色和策略的最佳实践，请参阅 AWS Well-Architected Framework 的安全性支柱。

资源

相关文档：

- 针对工作职能的 AWS 托管式策略
- AWS 多账户计费策略
- AWS Well-Architected Framework 安全性支柱
- AWS Identity and Access Management（IAM）
- AWS Identity and Access Management 策略

相关视频：

- 为何使用身份与访问管理

相关示例：

- 使用 IAM 策略控制对 AWS 区域 的访问
- 开启您的云财务管理之旅：云成本运营

COST02-BP05 实施成本控制

根据组织策略以及定义的组和角色来实施控制。这样可以确保成本只根据组织要求的规定产生，例如，控制用户对区域或资源类型的访问。

在未建立这种最佳实践的情况下暴露的风险等级：中

实施指导

实施成本控制的第一步通常是进行相关通知设置，以便在发生成本或使用情况超出策略的事件时触发通知。可以迅速采取行动，并验证是否需要采取纠正措施，而不会限制工作负载或新活动，抑或是对它们产生负面影响。了解工作负载和环境限制后，就可以强制实施治理。AWS Budgets 允许您为 AWS 成本、使用情况和承诺折扣（节省计划和预留实例）设置通知和定义每月预算。可以在总成本级别（如所有成本）创建预算，也可以在更细粒度的级别创建预算，其中只包含特定的维度，如关联的账户、服务、标签或可用区。

使用 AWS Budgets 设置预算限制后，请使用 AWS Cost Anomaly Detection 来降低意外成本。AWS Cost Anomaly Detection 是一种成本管理服务，使用机器学习持续监控您的成本和使用情况，以检测异常支出。它可以帮助您识别异常支出和根本原因，以便您可以快速采取行动。首先，在 AWS Cost Anomaly Detection 中创建成本监控，然后通过设置美元阈值来选择提醒首选项（例如，对影响大于 1000 美元的异常情况发出提醒）。收到提醒后，可以分析造成异常情况的根本原因，以及对成本的影响。您还可以在 AWS Cost Explorer 中监控并执行自己的异常分析。

在 AWS 中通过 AWS Identity and Access Management 和 AWS Organizations 服务控制策略（SCP）强制执行治理策略。借助 IAM，可以安全地管理对 AWS 服务和资源的访问。可以使用 IAM 控制谁能创建或管理 AWS 资源、可创建的资源类型，以及可在何处创建。这最大限度地降低了在定义的策略之外创建资源的可能性。使用先前创建的角色和组，并分配 IAM 策略以强制实施正确的使用量。SCP 用于集中管控组织中所有账户的最大可用权限，从而确保账户始终在访问控制准则允许的范围内。SCP 仅在启用了所有功能的组织中可用，并且可以将 SCP 配置为默认情况下拒绝或允许对成员账户执行操作。有关实施访问管理的更多详细信息，请参阅《Well-Architected 安全性支柱白皮书》。

也可以通过 AWS 服务配额管理实施治理。通过确保为服务配额设置最低开销并进行准确维护，可以最大限度地减少组织要求以外的资源创建。要实现这一点，必须了解要求的改变速度、正在进行的项目（资源的创建和停用），并考虑配额更改的实施速度。必要时，服务配额可用于增加配额。

实施步骤

- 实施支出通知：使用定义的组织策略，创建 AWS Budgets 以在支出超出策略时收到通知。配置多个成本预算，每个账户一个，以通知您账户的总支出情况。在每个账户中为该账户内的较小单元配置额外的成本预算。这些单元因账户结构而异。一些常见示例包括 AWS 区域、工作负载（使用标签）或 AWS 服务。将电子邮件通讯组列表配置为通知的收件人，而不是个人的电子邮件账户。可以为超出金额的情况配置实际预算，或者使用预测预算通知预测使用情况。还可以预配置 AWS Budgets 操作，以强制实施特定 IAM 或 SCP 策略，或停止目标 Amazon EC2 或 Amazon RDS 实例。可以自动执行预算操作，也可以要求工作流程审批。
- 对异常支出实施通知：使用 AWS Cost Anomaly Detection 降低组织中的意外成本，并分析潜在异常支出的根本原因。创建成本监控来以指定粒度识别异常支出，并在 AWS Cost Anomaly Detection 中配置通知后，它会在检测到异常支出时向您发送提醒。这将让您能够分析导致异常的根本原因，并了解对成本的影响。配置 AWS Cost Anomaly Detection 时使用 AWS 成本类别来确定哪个项目团队或业务部门团队可以分析导致意外成本的根本原因，并及时采取必要的措施。
- 实施使用情况控制：使用定义的组织策略，实施 IAM 策略和角色，指定用户可执行和无法执行的操作。一个 AWS 策略中可能包含多个组织策略。采用定义策略时所用的方式，首先大致进行，然后在每一步施加更细粒度的控制。服务限制也是一种有效的使用情况控制措施。对所有账户实施正确的服务限制。

资源

相关文档：

- 针对工作职能的 AWS 托管式策略
- AWS 多账户计费策略
- 使用 IAM 策略控制对 AWS 区域 的访问
- AWS Budgets
- AWS Cost Anomaly Detection
- 控制您的 AWS 成本

相关视频：

- 如何使用 AWS Budgets 来跟踪我的支出和使用情况

相关示例：

- IAM 访问管理策略示例
- 服务控制策略示例
- AWS Budgets 操作
- 创建 IAM 策略以使用标签控制对 Amazon EC2 资源的访问权限
- 限制 IAM Identity 对特定 Amazon EC2 资源的访问权限
- Slack integrations for Cost Anomaly Detection using Amazon Q Developer in chat applications

COST02-BP06 跟踪项目生命周期

跟踪、衡量并审核项目、团队和环境的生命周期，以避免使用不必要的资源并为此付费。

在未建立这种最佳实践的情况下暴露的风险等级：低

实施指导

通过有效地跟踪项目生命周期，组织可以加强规划、管理和资源优化，进而更好地控制成本。通过跟踪获得的见解对于作出明智的决策非常宝贵，有助于提高项目的成本效益和整体成功率。

跟踪工作负载的整个生命周期，有助于您了解何时不再需要工作负载或工作负载组件。现有的工作负载和组件可能看起来仍在使用中，但是当 AWS 发布新的服务或功能时，它们可能会被停用或采用。检查

工作负载的先前阶段。在工作负载进入生产之后，可以停用以前的环境或大幅降低其容量，直到再次需要它们为止。

可以用时间范围或提醒标记资源，以便确定工作负载的审核时间。例如，如果上一次审核开发环境是在几个月前，现在可能非常适合再次审核该环境，以探究是否能采用新服务或该环境是否在使用中。可以在 AWS 上使用 myApplications 对应用程序进行分组和标记，以管理和跟踪元数据，例如重要性、环境、上次审查时间和成本中心。您既可以跟踪工作负载的生命周期，也可以监控和管理应用程序的成本、运行状况、安全态势和性能。

AWS 提供了各种可用于实体生命周期跟踪的管理和治理服务。可以使用 AWS Config 或 AWS Systems Manager 提供一份详尽的 AWS 资源和配置清单。建议集成现有项目或资产管理系统，以跟踪组织内的活动项目和产品。将当前系统与 AWS 提供的丰富事件和指标集结合起来，您就可以构建重要生命周期事件的视图并主动管理资源，以减少不必要的成本。

与应用程序生命周期管理（ALM）类似，跟踪项目生命周期应涉及多个流程、工具、团队协同工作，例如设计和开发、测试、生产、支持及工作负载冗余。

通过仔细监控项目生命周期的每个阶段，组织可以获得重要的见解并增强掌控力，从而促进成功的项目规划、实施和完成。这种仔细的监督可以确保项目不仅符合质量标准，而且在预算范围内按时交付，从而提高总体成本效率。

有关实施实体生命周期跟踪的更多详细信息，请参阅《AWS Well-Architected 卓越运营支柱白皮书》。

实施步骤

- 建立项目生命周期监控流程：云卓越中心团队必须建立项目生命周期监控流程。建立结构化的系统化方法来监控工作负载，以改善项目的控制、可见性和性能。使监控过程透明化、协作化并注重持续改进，以最大限度地提高其有效性和价值。
- 执行工作负载审查：根据组织策略的定义，定期安排时间审核现有项目并执行工作负载审查。在审核方面投入的工作量应与组织的大致风险、价值或成本成比例。主要审核领域包括组织面临的意外事件或中断风险，或对组织所做的贡献（以收入或品牌声誉进行衡量）、工作负载的成本（以资源的总成本和运营成本进行衡量）和工作负载的使用情况（以单位时间的组织产出量进行衡量）。如果这些领域在生命周期内发生变化，则需要对工作负载进行调整，例如全部停用或部分停用。

资源

相关文档：

- Guidance for Tagging on AWS

- 什么是 ALM（应用程序生命周期管理）？
- 针对工作职能的 AWS 托管式策略

相关示例：

- 使用 IAM 策略控制对 AWS 区域 的访问

相关工具

- AWS Config
- AWS Systems Manager
- AWS Budgets
- AWS Organizations
- AWS CloudFormation

监控成本和使用情况

支持团队详细了解工作负载，从而针对成本和使用情况采取行动。要进行成本优化，首先必须详尽了解成本和使用情况明细，能够对未来的支出、使用情况和功能进行建模和预测，并实施足够的机制以使成本和使用情况符合组织的目标。以下是监控成本和使用情况时必须完成的步骤：

最佳实践

- COST03-BP01 配置详细信息源
- COST03-BP02 在成本和使用情况中添加组织信息
- COST03-BP03 确定成本归属类别
- COST03-BP04 建立组织指标
- COST03-BP05 配置账单和成本管理工具
- COST03-BP06 根据工作负载指标分配成本

COST03-BP01 配置详细信息源

设置成本管理和报告工具，进一步分析成本和使用情况数据并提高透明度。配置工作负载以创建日志条目，便于跟踪和分割成本和使用情况。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

利用成本管理工具中的详细账单信息（如每小时粒度），组织能够更详细地跟踪使用情况，并有助于组织找出成本增加的一些原因。这些数据来源最确切地反映了整个组织中的成本和使用情况。

可以使用 AWS Data Exports 创建 AWS 成本和使用情况报告（CUR）2.0 的导出。这是从 AWS 接收详细的成本和使用情况数据的新推荐方式。这种方法可以提供所有收费 AWS 服务的每日或每小时使用粒度、费率、成本和使用属性（与 CUR 的信息相同），以及一些改进。CUR 中包括所有可能的维度，例如标签、位置、资源属性和账户 ID。

根据要创建的导出的类型，有三种导出类型：标准数据导出、导出到集成 QuickSight 的成本与使用情况控制面板，或旧版数据导出。

- 标准数据导出：定期交付到 Amazon S3 的表的自定义导出。
- 成本与使用情况控制面板：导出到 QuickSight 和与之集成，用于部署预构建的成本与使用情况控制面板。
- 旧版数据导出：旧版 AWS 成本和使用情况报告（CUR）的导出。

可以使用以下自定义项创建数据导出：

- 包括资源 ID
- 拆分成本分配数据
- 每小时粒度
- 版本控制
- 压缩类型和文件格式

对于在 Amazon ECS 或 Amazon EKS 上运行容器的工作负载，启用拆分成本分配数据，以便可以根据容器工作负载使用共享计算和内存资源的方式，将容器成本分配给各个业务部门和团队。拆分成本分配数据将新的容器级资源的成本和使用情况数据引入 AWS 成本和使用情况报告。拆分成本分配数据通过计算集群上运行的各个 ECS 服务和任务的成本来计算。

成本与使用情况控制面板定期将成本与使用情况控制面板表导出到 S3 存储桶，并将预构建的成本与使用情况控制面板部署到 QuickSight。如果您想快速部署包含成本和使用情况数据的控制面板，且不需要自定义功能，请使用此选项。

如果需要，您仍然可以导出旧版 CUR，并可在其中集成其他处理服务（例如 AWS Glue）来准备数据以供分析，并使用 Amazon Athena 进行数据分析，通过 SQL 查询数据。

实施步骤

- 创建数据导出：使用所需数据创建自定义导出并控制导出的架构。使用基本 SQL 创建账单和成本管理数据导出，并通过集成 QuickSight 实现账单和成本管理数据的可视化。您还能以标准模式导出数据，以使用 Amazon Athena 等其他处理工具来分析数据。
- 配置成本和使用情况报告：使用账单控制台，配置至少一个成本和使用情况报告。配置采用每小时粒度的报告，以便包括所有标识符和资源 ID。还可以创建采用不同粒度的其他报告，以提供概括性摘要信息。
- 在 Cost Explorer 中配置每小时粒度：要以每小时粒度访问过去 14 天的成本和使用情况数据，请考虑在账单控制台中启用每小时和资源级别数据。
- 配置应用程序日志记录：确认应用程序记录所交付的每项业务成果，以便进行跟踪和衡量。确保该数据的粒度至少为每小时一次，以便与成本和使用情况数据匹配。有关日志记录和监控的更多详细信息，请参阅《Well-Architected 卓越运营支柱》。

资源

相关文档：

- AWS Data Exports
- AWS Glue
- QuickSight
- AWS 成本管理定价
- 为AWS资源添加标签
- 使用 Cost Explorer 分析成本
- 管理 AWS 成本和使用情况报告
- Well-Architected Operational Excellence Pillar

相关示例：

- AWS 账户设置
- Data Exports for AWS Billing and Cost Management

- AWS Cost Explorer 常见应用场景

COST03-BP02 在成本和使用情况中添加组织信息

根据组织、工作负载属性和成本分配类别定义标记方案，以便可以在成本管理工具中筛选和搜索资源或监控成本和使用情况。在可能的情况下，根据目的、团队、环境或与业务相关的其他标准，在所有资源中应用一致的标签。

在未建立这种最佳实践的情况下暴露的风险等级：中

实施指导

在 AWS 中应用标签以将组织信息添加到资源中，然后将该信息添加到成本和使用情况信息中。标签是键值对：键经过定义，必须在整个组织中唯一，值则对于一组资源唯一。键值对的一个示例是键为 `Environment`，值为 `Production`。生产环境中的所有资源都有这个键值对。借助标签，可以使用有意义、相关的组织信息对成本进行分类和跟踪。可以应用代表组织类别（例如成本中心、应用程序名称、项目或负责人）的标签，标识工作负载和工作负载的特征（例如测试或生产），在整个组织中对成本和使用情况进行归属。

当您将标签应用于 AWS 资源（如 Amazon Elastic Compute Cloud 实例或 Amazon Simple Storage Service 存储桶）并激活标签后，AWS 会将此信息添加到成本和使用情况报告中。可以在带标签和无标签的资源上运行报告并执行分析，以更好地遵守内部成本管理策略，并确保准确归属。

跨组织账户创建和实施 AWS 标签标准之后，您将能够一致且统一地管理和治理 AWS 环境。使用 AWS Organizations 中的标签策略定义有关如何在 AWS Organizations 账户的 AWS 资源上使用标签的规则。借助标签策略，可以采用标准化方法轻松标记 AWS 资源

AWS 标签编辑器可用于为多个资源添加、删除和管理标签。通过使用标签编辑器，您可以搜索要标记的资源，然后在搜索结果中管理这些资源的标签。

AWS Cost Categories 可用于向成本分配组织含义，而无需在资源上添加标签。可以将成本和使用情况信息映射到唯一的内部组织结构。可以定义类别规则，以使用账单维度（例如账户和标签）对成本进行映射和分类。除了标记之外，这还提供了另一个级别的管理能力。您还可以将特定账户和标签映射到多个项目。

实施步骤

- 定义标记方案：召集整个企业的所有利益相关方来定义方案。这通常包括担任技术、财务和管理角色的人员。定义所有资源必须具有的标签列表，以及资源应该具有的标签列表。确认标签名称和值在整个组织中是否一致。

- 标签资源：使用定义的成本归属类别，根据类别在工作负载中的所有资源上放置标签。使用 CLI、标签编辑器或 AWS Systems Manager 等工具提高效率。
- 实施 AWS Cost Categories：无需应用标签即可创建成本类别。Cost Categories 使用现有的成本和使用情况维度。根据方案创建类别规则，并在 Cost Categories 中加以实施。
- 实现标记自动化：为确保对所有资源保持高水平的标记，请自动进行标记，以便在创建资源时自动对其进行标记。使用诸如 AWS CloudFormation 之类的服务来验证资源在创建时是否已标记。还可以创建自定义解决方案以便使用 Lambda 函数实现自动标记，或使用微服务定期扫描工作负载并删除没有标记的任何资源，此方法非常适合测试和开发环境。
- 监控和报告标记：为确保在整个组织中保持高水平的标记，请跨工作负载报告和监控标签。可以使用 AWS Cost Explorer 查看带标签和无标签的资源的成本，也可以使用标签编辑器等服务。定期审核无标签的资源的数量，并执行操作添加标签，直到达到所需的标记级别。

资源

相关文档：

- 标记最佳实践
- AWS CloudFormation 资源标签
- AWS Cost Categories
- 为AWS资源添加标签
- 使用 AWS Budgets 分析成本
- 使用 Cost Explorer 分析成本
- 管理 AWS 成本和使用情况报告

相关视频：

- 如何标记 AWS 资源，以便按成本中心或项目划分账单
- 标记 AWS 资源

COST03-BP03 确定成本归属类别

确定组织类别，例如业务单位、部门或项目，这些类别可用于在组织中按照内部使用实体分配成本。利用这些类别来执行支出问责制，树立成本意识，并推动高效的使用行为。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

成本分类过程在预算编制、会计、财务报告、决策、基准测试和项目管理中至关重要。通过对费用进行分级和分类，团队可以更好地了解自己整个云之旅中会产生的成本类型，从而有助于团队作出明智的决策并有效地管理预算。

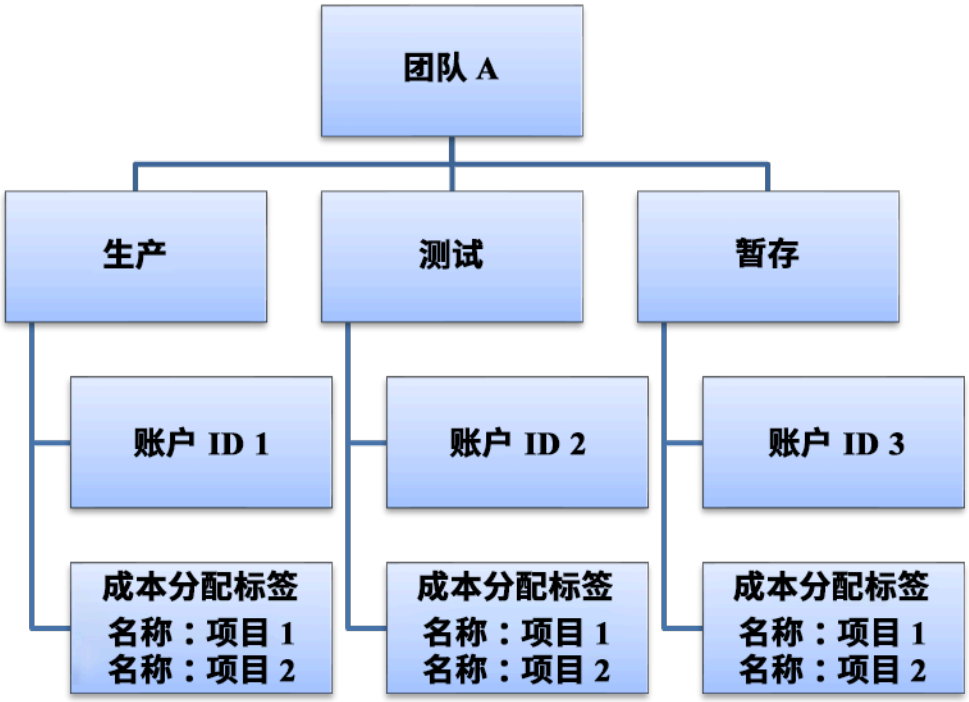
云支出问责制为严明的需求和成本管理提供了强有力的激励措施。最终，组织在将大部分云支出分配到使用这些资源的业务单位或团队后，可以显著地节省云成本。此外，通过分配云支出，有助于组织采用更多集中式云治理的最佳实践。

与财务团队和其他相关利益相关方合作，在定期沟通通话期间，了解必须如何在组织内部分配成本的要求。必须将工作负载成本分配至整个生命周期，包括开发、测试、生产和停用。了解组织如何对学习、员工培养和创意构思进行成本归类。这有助于将用于此目的的账户正确分配给培训和开发预算，而不是一般的 IT 成本预算。

与组织中的利益相关方一起定义成本归属类别后，使用 AWS Cost Categories 将成本和使用情况信息分组为 AWS Cloud 中有意义的类别，例如特定项目的成本、部门或业务部门的 AWS 账户。可以创建自定义类别，并根据使用各种维度（如账户、标签、服务或费用类型）定义的规则，将成本和使用情况信息映射到这些类别中。设置成本类别后，可以按这些类别查看成本和使用情况信息，从而让组织能够作出更好的战略和采购决策。这些类别也可在 AWS Cost Explorer、AWS Budgets 和 AWS 成本和使用情况报告 中查看。

例如，为业务部门（DevOps 团队）创建成本类别，并在每个类别下创建多个规则（每个子类别的规则），这些规则具有基于所定义分组的多个维度（AWS 账户、成本分配标签、服务或费用类型）。通过 Cost Categories，您可以使用基于规则的引擎来组织您的成本。您配置的规则按类别组织您的成本。在这些规则中，可以使用每个类别的多个维度进行筛选，例如特定 AWS 账户、AWS 服务或费用类型。可以在 AWS 账单与成本管理 和成本管理控制台中跨多个产品使用这些类别。这包括 AWS Cost Explorer、AWS Budgets、AWS 成本和使用情况报告 和 AWS Cost Anomaly Detection。

下图举例说明了如何通过多个团队（成本类别）、多个环境（规则）以及每个环境具有多个资源或资产（维度），对组织中的成本和使用情况信息进行分组。



成本和使用情况组织结构图

您也可以使用成本类别创建成本分组。创建成本类别后（为使用情况记录创建成本类别之后，允许在 24 小时内更新相应的值），这些类别将显示在 AWS Cost Explorer、AWS Budgets、AWS 成本和使用情况报告 和 AWS Cost Anomaly Detection 中。在 AWS Cost Explorer 和 AWS Budgets 中，成本类别显示为额外的计费维度。您可以使用此项筛选特定成本类别值，或者按成本类别分组。

实施步骤

- 定义组织类别：与内部利益相关方和业务部门会面，确定反映组织结构和要求的类别。这些类别应直接对应于现有财务类别的结构，例如业务部门、预算、成本中心或部门。了解云为您带来的业务成果（例如培训或教育），因为这些也是组织类别。
- 定义职能类别：与内部利益相关方和业务部门会面，确定反映业务所含职能的类别。这可以是工作负载名称或应用程序名称以及环境类型（例如生产、测试或开发）。
- 定义 AWS Cost Categories：使用 AWS Cost Categories 创建成本类别以整理成本和使用情况信息，并将 AWS 成本和使用情况映射到有意义的类别。可以将多个类别分配给一个资源，并且一个资源可以位于多个不同的类别中，因此可以根据需要定义任意数量的类别，以便可以使用 AWS Cost Categories 在分类的结构中管理成本。

资源

相关文档：

- 为AWS资源添加标签
- 使用成本分配标签
- 使用 AWS Budgets 分析成本
- 使用 Cost Explorer 分析成本
- 管理 AWS 成本和使用情况报告
- AWS Cost Categories
- 使用 AWS Cost Categories 管理成本
- 创建成本类别
- 标记成本类别
- 在成本类别中拆分费用
- AWS Cost Categories 的功能

相关示例：

- 使用 AWS Cost Categories 整理成本和使用情况数据
- 使用 AWS Cost Categories 管理成本

COST03-BP04 建立组织指标

建立此工作负载需要的组织指标。生成的客户报告或提供给客户的 Web 页面都属于工作负载指标。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

了解如何根据业务成功来衡量工作负载的输出。每个工作负载通常有一组表示性能的主要输出。如果您的工作负载复杂且包含许多组件，则可以对列表进行优先级排序，或者为每个组件定义和跟踪指标。与团队合作，了解要使用哪些指标。此部分将用于了解工作负载的效率，或每项业务产出的成本。

实施步骤

- 定义工作负载成果：与业务利益相关方召开会议，定义工作负载成果。这些主要用于衡量客户使用情况，因此必须是业务指标，而不是技术指标。每个工作负载应该有少量的概要指标（少于 5 个）。如果工作负载针对不同的应用场景产生多个结果，请将其分组为一个指标。
- 定义工作负载组件指标：如果工作负载大而复杂，或者可以轻松地将工作负载分为输入和输出定义明确的多个组件（例如微服务），则可以选择为每个组件定义指标。这项工作应反映组件的价值和成本。按照从大到小的顺序，从最大的组件开始，逐步处理较小的组件。

资源

相关文档：

- 为AWS资源添加标签
- 使用 AWS Budgets 分析成本
- 使用 Cost Explorer 分析成本
- 管理 AWS 成本和使用情况报告

COST03-BP05 配置账单和成本管理工具

根据组织的策略配置成本管理工具，以管理和优化云支出。这包括用于整理和跟踪成本和使用情况数据的服务、工具和资源，通过整合账单和访问权限增强控制，通过预算编制和预测改进规划，接收通知或提醒，并通过资源和定价优化降低成本。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

为了建立强有力的问责制，首先将您的账户策略视为成本分配策略的一部分。做好这一点，可能会为您省下许多工作。否则就会出现意识不足的情况，造成更多痛点。

为了鼓励针对云支出建立问责制，应向用户授予权限，允许他们使用可查看其成本和使用情况的工具。AWS 建议您出于以下目的配置所有工作负载和团队：

- 整理：使用您自己的标记策略和分类方法，建立成本分配和治理基准。使用 AWS Control Tower 或 AWS Organization 等工具创建多个 AWS 账户。标记支持的 AWS 资源，并根据组织结构（业务部门、部门或项目）对资源进行有目的性的分类。为特定成本中心标记账户名称，并将其与 AWS Cost Categories 对应起来，以便为业务部门就其成本中心进行账户分组，这样业务部门负责人就可以在一个位置查看多个账户的使用情况。

- 访问权限：在整合账单中跟踪组织范围的账单信息。确认相应的利益相关方和业务负责人是否拥有访问权限。
- 控制：使用服务控制策略（SCP）、标签策略、IAM 策略和预算警报时，使用适当的防护机制建立有效的治理机制，以防止出现意外情况。例如，可以使用有效的控制机制，仅允许团队在首选区域内创建特定资源，并防止没有特定标签（例如成本中心）的资源创建。
- 当前状态：配置显示当前成本和使用情况水平的控制面板。该控制面板应位于工作环境中的显眼位置，类似于操作控制面板。可以导出数据，并使用 AWS 成本优化中心中的成本与使用情况控制面板或任何支持的产品来实现此可见性。您可能需要为不同的角色创建不同的控制面板。例如，管理器控制面板可能不同于工程控制面板。
- 通知：当成本或使用情况超过定义的限制并且 AWS Budgets 或 AWS 异常检测出现异常时，提供通知。
- 报告：汇总所有成本和使用情况信息。通过详细的、可归因的成本数据，提高对云支出的认识，加强问责制。创建报告（这些报告与其使用团队相关）并包含建议。
- 跟踪：对照配置的方向性目标或执行性目标显示当前的成本和使用情况。
- 分析：允许团队成员使用不同的筛选条件（资源、账户、标签等）执行自定义和深入分析，精确到每小时、每天或每月的粒度。
- 检查：随时了解最新的资源部署和成本优化机会。使用 Amazon CloudWatch、Amazon SNS 或 Amazon SES 获取组织级别的资源部署的通知。使用 AWS Trusted Advisor 或 AWS Compute Optimizer 审核成本优化建议。
- 趋势报告：以所需的粒度显示所需期限内成本和使用情况的变化。
- 预测：使用您创建的预测控制面板，显示估计的未来成本、估计资源使用情况和支出。

可以使用 AWS 成本优化中心来了解从集中位置整合的潜在成本节约机会，并创建数据导出以与 Amazon Athena 集成。还可以使用 AWS 成本优化中心来部署成本与使用情况控制面板，该控制面板利用 QuickSight 进行交互式成本分析和安全的成本洞察共享。

如果组织不具备必备技能或带宽，则可以与 AWS ProServ、AWS Managed Services（AMS）或 AWS 合作伙伴合作。您也可以使用第三方工具，但请务必验证价值主张。

实施步骤

- 允许对工具进行基于团队的访问：配置账户并创建组，这些组可以访问所需的成本和使用情况报告来了解其使用情况，并使用 AWS Identity and Access Management 控制对 AWS Cost Explorer 等工具的访问权限。这些组必须包括负责或管理应用程序的所有团队的代表。这证明每个团队都可以访问他们的成本和使用情况信息，以跟踪使用情况。

- 整理成本标签和类别：跨团队、业务部门、应用程序、环境和项目整理成本。使用资源标签来按成本分配标签整理成本。使用标签、账户、服务等，根据维度创建成本类别，以映射成本。
- 配置 AWS Budgets：在所有账户中为工作负载配置 AWS Budgets。使用标签和成本类别，设置账户总支出预算和工作负载预算。在 AWS Budgets 中配置通知，以便在您超出预算金额或估计成本超出预算时收到警报。
- 配置 AWS 成本异常检测：针对您的账户、核心服务或成本类别使用 AWS 成本异常检测，以监控成本和使用情况，并检测异常支出。可以在汇总的报告中单独收到警报，以及在电子邮件或 Amazon SNS 主题中收到警报，以便您分析和确定异常的根本原因，并确定导致成本增加的原因。
- 使用成本分析工具：为工作负载和账户配置 AWS Cost Explorer，实现成本数据可视化，以便进一步分析。为工作负载创建一个控制面板，用于跟踪总体支出、工作负载的关键使用指标，以及基于历史成本数据的未来成本预测。
- 使用成本节省分析工具：通过 AWS 成本优化中心利用量身定制的建议（包括删除未使用的资源、合理调整大小、节省计划和预留）和 Compute Optimizer 建议来发现可节省成本的机会。
- 配置高级工具：可以选择创建视觉对象以促进交互式分析和成本洞察的分享。借助 AWS 成本优化中心上的数据导出功能，可以为组织创建由 QuickSight 提供支持的成本与使用情况控制面板，从而提供更多详细信息和粒度。还可以通过使用 Amazon Athena 中的数据导出来实现高级分析功能以进行高级查询，并在 QuickSight 上创建控制面板。与 AWS 合作伙伴合作，将云管理解决方案用于整合的云账单监控和优化。

资源

相关文档：

- What is AWS 账单与成本管理 and Cost Management?
- 建立您的最佳实践 AWS 环境
- 标记 AWS 资源的最佳实践
- 标记您的 AWS 资源
- AWS Cost Categories
- 使用 AWS Budgets 分析成本
- 使用 AWS Cost Explorer 分析成本
- What is AWS Data Exports?

相关视频：

- Deploying Cloud Intelligence Dashboards
- Get Alerts on any FinOps or Cost Optimization Metric or KPI

相关示例：

- Cost and Usage Dashboard powered by QuickSight
- AWS Cost and Usage Governance 讲习会

COST03-BP06 根据工作负载指标分配成本

根据使用情况指标或业务成果分配工作负载的成本，以便衡量工作负载的成本效率。实施一个流程，使用分析服务来分析成本和使用情况数据，以便深入了解成本因素和退款功能。

在未建立这种最佳实践的情况下暴露的风险等级：低

实施指导

成本优化是指以最低的价格实现业务成果，这只能通过按工作负载指标分配工作负载成本（按工作负载效率衡量）来实现。通过日志文件或其他应用程序监控来监控定义的工作负载指标。将此数据与工作负载的成本（可通过查看具有特定标签值或账户 ID 的成本获得）相结合。每小时进行一次此分析。如果有静态成本要素（例如，持续运行的后端数据库）且请求率不同（例如，使用量高峰在上午 9 点至下午 5 点，晚间的请求数量很少），则效率通常会变化。了解静态成本和可变成本之间的关系，有助于您专注于优化活动。

与 Amazon Elastic Container Service（Amazon ECS）和 Amazon API Gateway 上的容器化应用程序等资源相比，为共享资源创建工作负载指标可能并非易事。但是，您可以通过某些方法来分类使用情况并跟踪成本。如果您需要跟踪 Amazon ECS 和 AWS Batch 共享资源，则可以在 AWS Cost Explorer 中启用拆分成本分配数据。通过拆分成本分配数据，您可以了解并优化容器化应用程序的成本和使用情况，并根据共享计算和内存资源的使用情况，将应用程序成本分配给各个业务实体。

实施步骤

- 将成本分配到工作负载指标：使用定义的指标和配置的标签，创建结合工作负载输出和工作负载成本的指标。使用 Amazon Athena 和 Amazon QuickSight 等分析服务，为整个工作负载和任何组件创建效率控制面板。

资源

相关文档：

- 为AWS资源添加标签
- 使用 AWS Budgets 分析成本
- 使用 Cost Explorer 分析成本
- 管理 AWS 成本和使用情况报告

相关示例：

- 利用 AWS 拆分成本分配数据提高 Amazon ECS 和 AWS Batch 的成本可见性

停用资源

当您管理项目、员工和技术资源的列表之后，随着时间的推移，您将能够识别不再使用的资源或不再具备拥有者的项目。

最佳实践

- COST04-BP01 在资源生命周期内跟踪资源
- COST04-BP02 实施停用流程
- COST04-BP03 停用资源
- COST04-BP04 自动停用资源
- COST04-BP05 执行数据留存策略

COST04-BP01 在资源生命周期内跟踪资源

制定和实施一种方法，在资源生命周期内跟踪资源及其与系统的关联。您可以使用标签来标识资源的工作负载或功能。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

停用不再需要的工作负载资源。一个常见的示例是用于测试的资源：在测试完成后，可以将资源删除。使用标签跟踪资源（并对这些标签运行报告），可以帮助您识别要停用的资产，因为将不会使用这些资

产，或者这些资产的许可证将过期。使用标签是跟踪资源的一种有效方法，它通过标记资源的功能或资源的已知可停用日期来跟踪资源。然后，可以对这些标签运行报告。功能标签的示例值是 `feature-X testing`，用于根据工作负载生命周期标识资源的用途。另一个例子是对资源使用 `LifeSpan` 或 `TTL`，例如待删除的标签键名称和值，用于定义停用的时间段或具体时间。

实施步骤

- 实施标记方案：实施标记方案，标识资源所属的工作负载，从而确保相应地标记工作负载中的所有资源。标签可以帮助您按用途、团队、环境或与业务相关的其他条件对资源进行分类。有关标签应用场景、策略和技术的更多详细信息，请参阅 AWS Tagging Best Practices。
- 实施工作负载吞吐量或输出监控：实施工作负载吞吐量监控或警报，在输入请求或输出完成时启动。将其配置为在工作负载请求或输出下降到零时发出通知，指示不再使用工作负载资源。如果在正常情况下，工作负载周期性地下降到零，则加入时间因素。有关未使用或未充分利用的资源的更多详细信息，请参阅 AWS Trusted Advisor Cost Optimization checks。
- 为 AWS 资源分组：为 AWS 资源创建组。可以使用 AWS Resource Groups 来整理和管理同一个 AWS 区域 中的 AWS 资源。可以将标签添加到大多数资源中，帮助标识资源并对组织内的资源进行排序。使用标签编辑器向支持的资源批量添加标签。考虑使用 AWS Service Catalog 创建、管理和向最终用户分发已批准的产品组合，并管理产品生命周期。

资源

相关文档：

- AWS Auto Scaling
- AWS Trusted Advisor
- AWS Trusted Advisor Cost Optimization Checks
- 为AWS资源添加标签
- 发布自定义指标

相关视频：

- How to optimize costs using AWS Trusted Advisor

相关示例：

- 组织 AWS 资源

- 使用 AWS Trusted Advisor 优化成本

COST04-BP02 实施停用流程

实施一个流程来确定和停用未使用的资源。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

在整个组织中实施标准化流程，以确定和删除未使用的资源。该流程应该定义执行搜索的频率以及删除资源的流程，以确认满足所有组织要求。

实施步骤

- 创建并实施停用流程：与工作负载开发人员和负责人合作，为工作负载及其资源构建停用流程。该流程应涵盖一种方法，验证工作负载是否正在使用以及每个工作负载资源是否正在使用。详细说明停用资源所需的步骤，将资源从服务中删除，同时确保符合所有法规要求。应包含任何关联的资源，例如许可证或附加的存储。向工作负载负责人发送已开启停用流程的通知。

  使用以下停用步骤来指导您确定在流程中应检查的内容：

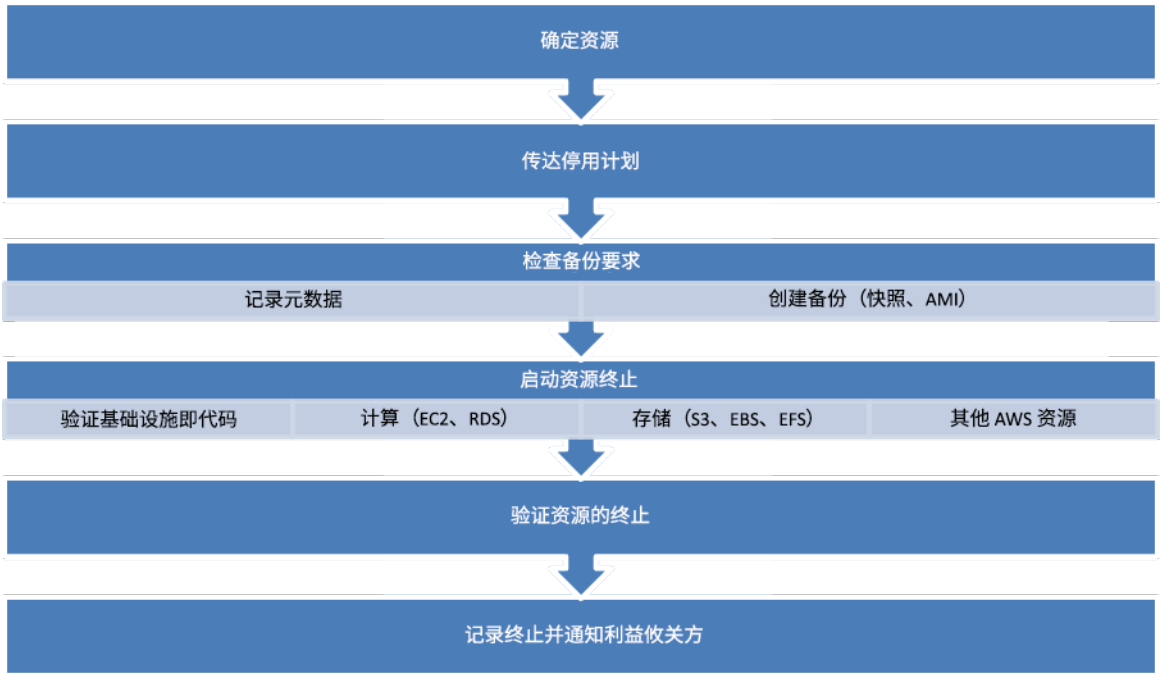
  - 确定要停用的资源：确定 AWS Cloud 中符合停用条件的资源。记录所有必要的信息并安排停用。在您的时间表中，请务必考虑在此过程中是否（以及何时）会出现意外问题。
  - 协调和沟通：与工作负载负责人合作确认要停用的资源
  - 记录元数据并创建备份：如果生产环境中的资源需要或它们是关键资源，请记录元数据（例如公有 IP、区域、可用区、VPC、子网和安全组），并创建备份（例如 Amazon Elastic Block Store 快照或取得 AMI、密钥导出和证书导出）。
  - 验证基础设施即代码：确定资源是使用 AWS CloudFormation、Terraform、AWS Cloud Development Kit (AWS CDK) 还是任何其他基础设施即代码部署工具部署的，以便在必要时可以重新部署它们。
  - 阻止访问：在一段时间内应用限制性控制，以防止在确定是否需要资源时使用资源。验证资源环境是否可以在需要时恢复到其原始状态。
  - 遵循内部停用流程：遵循组织的管理任务和停用流程，例如从组织域中删除资源、删除 DNS 记录以及从配置管理工具、监控工具、自动化工具和安全工具中删除资源。

  如果资源是 Amazon EC2 实例，请参阅以下列表。有关更多详细信息，请参阅《如何删除或终止我的 Amazon EC2 资源？》

- 停止或终止您的所有 Amazon EC2 实例和负载均衡器。Amazon EC2 实例终止后可短时间内在控制台中看到。您不需要为任何未处于运行状态的实例付费
- 删除 Auto Scaling 基础设施。
- 释放所有专属主机。
- 删除所有 Amazon EBS 卷和 Amazon EBS 快照。
- 释放所有弹性 IP 地址。
- 取消注册所有亚马逊机器映像（AMI）。
- 终止所有 AWS Elastic Beanstalk 环境。

如果资源是 Amazon S3 Glacier 存储中的对象，并且您在满足最短存储持续时间之前删除了存档，将按比例收取提前删除费用。Amazon S3 Glacier 的最短存储持续时间取决于所使用的存储类。有关每种存储类的最短存储持续时间摘要，请参阅跨 S3 存储类的性能。有关如何计算提前删除费用的详细信息，请参阅 Amazon S3 定价。

下面的简单停用过程流程图概述了停用步骤。在停用资源之前，请确认组织未使用已确定要停用的资源。



资源停用流程。

资源

相关文档：

- AWS Auto Scaling
- AWS Trusted Advisor
- AWS CloudTrail

相关视频：

- 删除 CloudFormation 堆栈，但保留某些资源
- Find out which user launched Amazon EC2 instance

相关示例：

- 删除或终止 Amazon EC2 资源
- 了解哪个用户启动了 Amazon EC2 实例

COST04-BP03 停用资源

停用由定期审核或使用情况发生变化等事件启动的资源。停用通常定期执行，可以手动停用，也可以自动停用。

在未建立这种最佳实践的情况下暴露的风险等级：中

实施指导

搜索未使用资源的频率和工作量应反映潜在的节省额，因此，与成本较高的账户相比，对成本较低的账户进行分析的频率应该更低。搜索和停用事件可由工作负载中的状态更改启动，比如产品生命周期结束或产品被更换。搜索和停用事件也可由外部事件启动，如市场条件发生变化或产品终止。

实施步骤

- 停用资源：这是不再需要的 AWS 资源的折旧阶段或许可协议的终止。在进入处置阶段并停用资源之前完成所有最终检查，以防止任何意外中断，例如拍摄快照或备份。使用停用流程，停用每项被确定为未使用的资源。

资源

相关文档：

- AWS Auto Scaling

- AWS Trusted Advisor

COST04-BP04 自动停用资源

设计您的工作负载，使其在您确定并停用非关键资源、不需要的资源或使用率低的资源时妥善处理资源的终止。

在未建立这种最佳实践的情况下暴露的风险等级：低

实施指导

使用自动化技术可以减少或消除停用流程中的相关成本。将工作负载设计为执行自动化停用，将减少工作负载在其整个生命周期内的总成本。您可以使用 Amazon EC2 Auto Scaling 或 Application Auto Scaling 来执行停用流程。还可以使用 API 或 SDK 来实施自定义代码，以自动停用工作负载资源。

现代应用程序是以“无服务器为先”理念构建的，这种策略优先考虑采用无服务器服务。AWS 为堆栈的所有三个层开发了无服务器服务：计算、集成和数据存储。使用无服务器架构将允许您在低流量期间通过自动纵向扩展和缩减来节省成本。

实施步骤

- 实施 Amazon EC2 Auto Scaling 或 Application Auto Scaling：对于支持的资源，请使用 Amazon EC2 Auto Scaling 或 Application Auto Scaling 来配置它们。这些服务可以帮助您在使用 AWS 服务时优化使用率和成本效率。当需求下降时，这些服务将自动删除任何多余的资源容量，以避免超支。
- 将 CloudWatch 配置为终止实例：可以将实例配置为使用 CloudWatch 警报终止。使用停用流程的指标，实施包含 Amazon Elastic Compute Cloud 操作的警报。在推出之前，在非生产环境中验证操作。
- 在工作负载中实现代码：可以使用 AWS SDK 或 AWS CLI 停用工作负载资源。在与 AWS 集成的应用程序中实现代码，并终止或删除不再使用的资源。
- 使用无服务器服务：优先在 AWS 上构建无服务器架构和事件驱动型架构，以构建和运行应用程序。AWS 提供多种无服务器技术服务，这些服务本质上可以自动优化资源利用率且具有自动停用功能（横向缩减和横向扩展）。通过无服务器应用程序，可自动优化资源利用率，您无需为过度预置付费。

资源

相关文档：

- Amazon EC2 Auto Scaling
- Getting Started with Amazon EC2 Auto Scaling
- Application Auto Scaling
- AWS Trusted Advisor
- AWS 无服务器
- 创建警报以停止、终止、重启或恢复实例
- 在 Amazon CloudWatch 警报中添加终止操作

相关示例：

- Scheduling automatic deletion of AWS CloudFormation stacks

COST04-BP05 执行数据留存策略

在支持的资源上定义数据留存策略，以根据组织要求处理对象的删除事宜。确定并删除非必要的或不再需要的孤立资源和对象。

在未建立这种最佳实践的情况下暴露的风险等级：中

使用数据留存策略和生命周期策略，降低停用过程的相关成本以及已确定资源的存储成本。定义数据留存策略和生命周期策略来执行自动存储类迁移和删除，将会降低其生命周期内的总体存储成本。可以使用 Amazon Data Lifecycle Manager 来自动创建和删除 Amazon Elastic Block Store 快照及基于 Amazon EBS 的亚马逊机器映像（AMI），并使用 Amazon S3 Intelligent-Tiering 或 Amazon S3 生命周期配置来管理 Amazon S3 对象的生命周期。还可以使用 API 或 SDK 实现自定义代码，为要自动删除的对象创建生命周期策略和策略规则。

实施步骤

- 使用 Amazon Data Lifecycle Manager：使用 Amazon Data Lifecycle Manager 上的生命周期策略，自动删除 Amazon EBS 快照和 Amazon EBS-backed AMI。
- 在存储桶上设置生命周期配置：根据业务要求，使用存储桶上的 Amazon S3 生命周期配置，定义 Amazon S3 在对象生命周期中要采取的操作以及对象生命周期结束时的删除操作。

资源

相关文档：

- AWS Trusted Advisor
- Amazon Data Lifecycle Manager
- 如何在 Amazon S3 存储桶上设置生命周期配置

相关视频：

- Automate Amazon EBS Snapshots with Amazon Data Lifecycle Manager
- 使用生命周期配置规则清空 Amazon S3 存储桶

相关示例：

- 使用生命周期配置规则清空 Amazon S3 存储桶

具有成本效益的资源

为工作负载使用合适的服务、资源和配置是节省成本的关键。创建具有成本效益的资源时，请考虑以下几点：

可以使用 AWS 解决方案架构师、AWS 解决方案、AWS 参考架构和 APN 合作伙伴来帮助您根据所学知识选择架构。

主题

- 选择服务时评估成本
- 选择正确的资源类型、规模和数量
- 选择最佳定价模式
- 制定数据传输计划

选择服务时评估成本

最佳实践

- COST05-BP01 确定组织对成本的要求
- COST05-BP02 分析工作负载的所有组件
- COST05-BP03 对每个组件进行全面分析
- COST05-BP04 选择具有成本效益许可的软件
- COST05-BP05 选择此工作负载的组件，以便根据组织的优先事项优化成本
- COST05-BP06 对不同时间的不同使用情况执行成本分析

COST05-BP01 确定组织对成本的要求

与团队成员合作，为此工作负载确定成本优化与其他支柱（例如性能和可靠性）之间的平衡。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

在大多数组织中，信息技术（IT）部门由多个小型团队组成，每个小团队都有自己的议程和重点领域，这反映了其团队成员的专长和技能。您需要了解组织的总体目标、优先事项、具体目标以及每个部门或

项目如何为这些总体目标做出贡献。对所有基本资源（包括人员、设备、技术、材料和外部服务）进行分类，是实现组织目标和全面预算规划的重要一环。采用这种系统化的方法来确定和了解成本，对于组织制定切合实际且稳健的成本计划至关重要。

在为工作负载选择服务时，了解组织的优先要务至关重要。在成本优化和 AWS Well-Architected Framework 的其他支柱（例如性能和可靠性）之间取得平衡。这一过程应系统性地定期进行，以反映组织目标、市场条件和运营动态的变化。完全成本优化的工作负载是最符合组织要求的解决方案，但不一定是成本最低的。与组织内的所有团队（例如产品、业务、技术和财务）会面以收集信息。在有冲突的利益或替代方法之间做出权衡并评估其影响，以便在确定工作重心或选择行动方案时作出明智的决策。

例如，加快新功能上市的速度可能会比成本优化更重要，或者您可以为非关系数据选择关系数据库，以简化迁移系统的工作，而不是迁移到针对数据类型优化的数据库和更新应用程序。

实施步骤

- 确定组织对成本的要求：与组织中的团队成员会面，这些成员包括产品管理、应用程序负责人、开发和运营团队、管理和财务人员。对此工作负载及其组件的 Well-Architected 支柱进行优先级排序，输出应该是一个按顺序排列的支柱列表。您还可以为每个支柱添加一个权重，指示该支柱体现的额外关注程度，或者两个支柱之间的关注点的相似程度。
- 解决技术债务并记录下来：在工作负载审核期间，解决技术债务。记录积压工作项以便将来重新审视工作负载，目标是重构或重新构建以进一步优化工作负载。必须向其他利益相关方明确说明所做的权衡取舍。

资源

相关最佳实践：

- REL11-BP07 构建产品以满足可用性目标和正常运行时间服务水平协议（SLA）
- OPS01-BP06 评估权衡

相关文档：

- AWS 总拥有成本（TCO）计算器
- Amazon S3 存储类
- 云产品

COST05-BP02 分析工作负载的所有组件

确认已分析工作负载的每个组件，无论当前大小或当前成本如何。审核工作应该体现可能带来的好处，例如当前成本和预期成本。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

旨在为组织提供业务价值的工作负载组件可能包含各种服务。对于每个组件，组织可以选择特定的 AWS Cloud 服务来满足业务需求。对这些服务的熟悉程度，或以往的使用经验等因素，可能会影响这种选择。

按照 COST05-BP01 确定组织对成本的要求中的说明确定组织的要求后，对工作负载中的所有组件进行全面分析。根据当前和预计的成本和规模，分析每个组件。结合工作负载在其生命周期内可能实现的节省来考虑分析成本。分析此工作负载的所有组件所花费的精力，应与优化该特定组件预期能产生的潜在节省或改进相匹配。例如，如果拟议资源的成本为每月 10 美元，在预测的负载下不会超过每月 15 美元，那么为了将成本降低 50%（每月 5 美元）而耗费的一天时间，其价值可能已经超过系统使用寿命内的潜在收益。使用更快、更高效的基于数据的预估，可为该组件带来出色的总体结果。

工作负载可能会随时间变化，如果工作负载架构或使用情况发生变化，原本合适的服务集可能不再是最优之选。为甄选服务进行分析时，必须考虑工作负载当前和未来的状态以及使用水平。为将来的工作负载状态或使用情况实施服务，可以减少或消除未来进行更改所需的工作量，从而降低总体成本。例如，最初使用 EMR Serverless 可能是合适的选择。但是，随着该服务使用情况的增加，过渡到 EC2 上的 EMR 可以降低该工作负载组件的成本。

AWS Cost Explorer 和 AWS 成本和使用情况报告（CUR）可以分析概念验证（PoC）或运行环境的成本。您还可以使用 AWS 定价计算器 估算工作负载成本。

编写一个工作流程以供技术团队用来审核其工作负载。此工作流程不仅要简单，还要涵盖所有必要步骤，以确保团队了解工作负载的每个组件及其定价。之后，组织可以遵循此工作流程，并根据每个团队的具体需求自定义此工作流程。

1. 列出用于工作负载的每项服务：这是一个很好的起点。确定当前使用的所有服务以及成本的来源。
2. 了解这些服务的定价方式：了解每项服务的定价模式。不同的 AWS 服务具有不同的定价模式，具体取决于使用量、数据传输和功能特定的定价等因素。
3. 重点关注那些工作负载成本不符合预期且与您的预期使用情况和业务结果不一致的服务：找出成本与使用 AWS Cost Explorer 或 AWS 成本和使用情况报告 的价值或使用情况不成比例时的异常值或服务。将成本与业务成果关联起来以确定优化工作的优先顺序，这一点很重要。

4. 使用 AWS Cost Explorer、CloudWatch Logs、VPC 流日志和 Amazon S3 Storage Lens 存储统计管理工具了解这些高成本的根本原因：这些工具有助于诊断高成本。每项服务都提供一种不同的分析功能来查看和分析使用情况和成本。例如，Cost Explorer 有助于确定总体成本趋势，CloudWatch Logs 可提供运营见解，VPC 流日志可显示 IP 流量，Amazon S3 Storage Lens 存储统计管理工具有助于存储分析。
5. 使用 AWS Budgets 为服务或账户设置特定金额的预算：设置预算是一种主动管理成本的方法。使用 AWS Budgets 可以设置自定义预算阈值，并在成本超过这些阈值时接收警报。
6. 配置 Amazon CloudWatch 警报以发送账单和使用情况警报：设置成本和使用情况指标的监控和警报。CloudWatch 警报可在达到特定阈值时通知您，从而缩短干预响应时间。

对所有工作负载组件（无论其当前属性如何）进行战略审核，从而推动实现显著的改进和财务节省。应仔细考量审核过程中投入的精力，同时认真考虑可能实现的潜在优势。

实施步骤

- 列出工作负载组件：构建工作负载组件的列表。使用此列表来验证是否已分析每个组件。应基于工作负载对企业优先事项的重要程度，来投入相应的工作量。按功能将资源分组可以提高效率（例如，将多个数据库合并为生产数据库存储）。
- 确定组件列表的优先级：选择组件列表并按工作量顺序对其进行优先级排序。通常按照组件的成本从最昂贵到最便宜的顺序排列，或者按照组件对组织优先事项的重要程度排列。
- 执行分析：对于列表中的每个组件，检查可用的选项和服务，然后选择最符合组织优先事项的选项。

资源

相关文档：

- AWS 定价计算器
- AWS Cost Explorer
- Amazon S3 存储类
- AWS Cloud 产品

相关视频：

- AWS 成本优化系列：CloudWatch

COST05-BP03 对每个组件进行全面分析

分析组织为每个组件付出的总体成本。通过考虑运营和管理成本（尤其是在使用云提供商提供的托管服务时）来计算总拥有成本。审核工作量应该体现可能带来的好处（例如分析时间与组件成本成正比）。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

利用节省下来的时间，您的团队将能够专注于解决技术债务、创新和增值功能，并打造企业的竞争优势。例如，您可能需要尽快将数据库从本地环境直接迁移（也称为重新托管）到云，然后再进行优化。值得探索的是，通过使用 AWS 上可消除或减少许可成本的托管服务，您可以节省多少成本。AWS 上的托管服务消除了维护服务的运营和管理负担（例如修补或更新操作系统），让您可以专注于创新和业务。

由于托管服务在云级运行，它们可以提供更低的单位事务或服务成本。您可以在不改变应用程序核心架构的情况下进行潜在优化，以获得一些切实的优势。例如，您可能希望通过迁移到 Amazon Relational Database Service（Amazon RDS）等数据库即服务平台或将应用程序迁移到完全托管式平台（例如 AWS Elastic Beanstalk），来减少管理数据库实例所花费的时间。

通常，可以设置托管服务的部分属性，以确保容量足够。您必须设置和监控这些属性，以便最大限度地减少多余容量，并最大限度地提高性能。您可以使用 AWS Management Console 或 AWS API 和 SDK 来修改 AWS Managed Services 的属性，使资源需求匹配不断变化的要求。例如，可以增加或减少 Amazon EMR 集群（或 Amazon Redshift 集群）上的节点数量，以横向扩展或缩减集群。

您还可以在 AWS 资源上打包多个实例，实现更高密度的使用情况。例如，您可以在单个 Amazon Relational Database Service（Amazon RDS）数据库实例上预置多个小型数据库。随着使用量的增长，您可以使用快照和还原过程，将其中一个数据库迁移到专用 Amazon RDS 数据库实例。

在托管服务上预置工作负载时，必须了解调整服务容量的要求。这些要求通常是时间、工作量和对正常工作负载运营的任何影响。预置的资源必须留出时间来进行任何更改，并预置所需开销以允许这样做。通过使用与系统和监控工具（如 Amazon CloudWatch）集成的 API 和软件开发工具包，可以将修改服务所需的持续工作量减少至接近零。

Amazon RDS、Amazon Redshift 和 Amazon ElastiCache 提供托管式数据库服务。Amazon Athena、Amazon EMR 和 Amazon OpenSearch Service 提供托管式分析服务。

AMS 是代表企业客户和合作伙伴运营 AWS 基础设施的服务。该服务提供了一个安全且合规的环境，您可以将工作负载部署到其中。AMS 使用具有自动化功能的企业云运营模型，让您可以满足组织要求，更快地迁移到云中并降低持续的管理成本。

实施步骤

- 执行全面的分析：使用组件列表，从最高优先级到最低优先级遍历每个组件。对于优先级较高且成本较高的组件，执行额外分析并评测所有可用选项及其长期影响。对于优先级较低的组件，评测使用情况的变化是否会更改组件的优先级，然后以适当的工作量进行分析。
- 比较托管资源和非托管资源：考虑您管理的资源的运营成本，并将其与 AWS 托管资源进行比较。例如，审查在 Amazon EC2 实例上运行的数据库，并与 Amazon RDS 选项（AWS 托管服务）进行比较，或将 Amazon EMR 与在 Amazon EC2 上运行 Apache Spark 进行比较。当从自我管理的工作负载迁移到 AWS 完全托管的工作负载时，请仔细研究您的选择。需要考虑的三个最重要的因素是您要使用的托管服务的类型、要用于迁移数据的流程以及了解 AWS 责任共担模式。

资源

相关文档：

- AWS 总拥有成本（TCO）计算器
- Amazon S3 存储类
- AWS Cloud 产品
- AWS 责任共担模式

相关视频：

- 为什么要迁移到托管数据库？
- 什么是 Amazon EMR，如何才能使用它来处理数据？

相关示例：

- 为什么要迁移到托管数据库？
- 使用 AWS DMS 将相同 SQL Server 数据库中的数据整合到单个 Amazon RDS for SQL Server 数据库中
- 将数据大规模传输到 Amazon Managed Streaming for Apache Kafka（Amazon MSK）
- 将 ASP.NET Web 应用程序迁移到 AWS Elastic Beanstalk

COST05-BP04 选择具有成本效益许可的软件

开源软件无需软件许可成本，从而大大节省了工作负载的成本。如果需要许可软件，应避免使用绑定到任意属性（如 CPU）的许可，而应使用绑定到输出或结果的许可。这些许可的成本与所提供的效益更为相当。

在未建立这种最佳实践的情况下暴露的风险等级：低

实施指导

开源起源于软件开发，表示软件符合某些自由发布的标准。任何人都可以检查、修改和增强开源软件的源代码。根据业务要求、工程师技能、预测的使用情况或其他技术依赖关系，组织可以考虑在 AWS 上使用开源软件来最大限度地降低许可成本。换言之，使用开源软件可以降低软件许可成本。随着工作负载规模的扩展，这可能对工作负载成本产生重大影响。

将许可软件能够带来的好处与总成本进行比较，以优化工作负载。对许可的任何更改以及更改对工作负载成本的影响建模。如果供应商更改了数据库许可的成本，请调查这会如何影响工作负载的整体效率。考虑供应商的历史定价公告，了解其产品中的许可更改趋势。许可成本也可以独立于吞吐量或使用情况进行扩缩，例如按硬件扩缩的许可（CPU 绑定许可）。应避免使用这些许可，因为成本会迅速增加，而且无法取得相应的结果。

例如，与在 Windows 上运行 Amazon EC2 实例相比，使用 Linux 操作系统在 us-east-1 中运行另一个 Amazon EC2 实例可以将成本降低大约 45%。

AWS 定价计算器 提供了一种全面的方法来比较具有不同许可选项的各种资源的成本，例如 Amazon RDS 实例和不同的数据库引擎。此外，AWS Cost Explorer 为现有工作负载（尤其是附带不同许可的工作负载）的成本提供了宝贵的视角。对于许可管理，AWS License Manager 提供了一种简化的方法来监督和处理软件许可。客户可以在 AWS Cloud 中部署和运行他们首选的开源软件。

实施步骤

- 分析许可选项：查看可用软件的许可条款。查看具有所需功能的开源版本，以及许可软件提供的效益是否大于成本。优惠条款可确保软件成本与所提供的效益相符。
- 分析软件提供商：查看供应商的任何历史定价或许可更改。了解与成果不符的任何更改，例如在特定供应商硬件或平台上运行的惩罚性条款。此外，还要了解他们如何执行审核和处罚。

资源

相关文档：

- Open Source at AWS
- AWS 总拥有成本（TCO）计算器
- Amazon S3 存储类
- 云产品

相关示例：

- 开源博客
- AWS 开源博客
- 优化与许可评测

COST05-BP05 选择此工作负载的组件，以便根据组织的优先事项优化成本

为工作负载选择所有组件时考虑成本因素。这包括使用应用程序级服务和托管服务或无服务器、容器或事件驱动型架构，以降低总体成本。使用开源软件、没有许可费用的软件或替代方案来尽可能减少许可成本，从而减少开支。

在未建立这种最佳实践的情况下暴露的风险等级：中

实施指导

在选择所有组件时考虑服务和选项的成本。这包括使用应用程序级服务和托管服务（例如 Amazon Relational Database Service（Amazon RDS）、Amazon DynamoDB、Amazon Simple Notification Service（Amazon SNS）和 Amazon Simple Email Service（Amazon SES）来降低总体组织成本。

使用无服务器和容器进行计算，例如将 AWS Lambda 和 Amazon Simple Storage Service（Amazon S3）用于静态网站。尽可能将应用程序容器化，并使用 AWS 托管容器服务，例如 Amazon Elastic Container Service（Amazon ECS）和 Amazon Elastic Kubernetes Service（Amazon EKS）。

使用开源软件或没有许可费用的软件，尽可能减少许可成本：例如，对计算工作负载使用 Amazon Linux 或将数据库迁移到 Amazon Aurora。

可以使用无服务器服务或应用程序级服务，例如 Lambda、Amazon Simple Queue Service（Amazon SQS）、Amazon SNS 和 Amazon SES。这些服务消除了管理资源的需要，并提供代码执行、服务排队和消息传递功能。另一个好处是，这些服务可以根据使用情况横向缩减性能和成本，从而实现有效的成本分配和归属。

对于无服务器服务，也可以使用事件驱动型架构。事件驱动型架构是推送式的，所以当事件在路由器中出现时，一切都按需发生。这样，您就不会为检查事件的连续轮询而付费。这意味着更少的网络带宽消耗、更低的 CPU 利用率、更少的闲置实例集容量，以及更少的 SSL/TLS 握手次数。

有关无服务器的更多信息，请参阅《Well-Architected 无服务器应用程序剖析白皮书》。

实施步骤

- 选择每项服务以优化成本：使用经过优先级排序的列表和分析，选择最符合组织优先事项的每个选项。与其增加容量来满足需求，不如考虑其他可以提供更高性价比的选项。例如，您需要审查数据库在 AWS 上的预期流量，并考虑增加实例大小或使用 Amazon ElastiCache 服务（Redis 或 Memcached），为数据库提供缓存机制。
- 评估事件驱动型架构：通过使用无服务器架构，您还可以为基于微服务的分布式应用程序构建事件驱动型架构，这有助于您构建可扩展、有韧性、敏捷且具有成本效益的解决方案。

资源

相关文档：

- AWS 总拥有成本（TCO）计算器
- AWS 无服务器
- 什么是事件驱动型架构
- Amazon S3 存储类
- 云产品
- Amazon ElastiCache（Redis OSS）

相关示例：

- Getting started with event-driven architecture
- 事件驱动型架构
- How Statsig runs 100x more cost-effectively using Amazon ElastiCache (Redis OSS)
- Best practices for working with AWS Lambda functions

COST05-BP06 对不同时间的不同使用情况执行成本分析

工作负载可能会随时间而变化。某些服务或功能在不同的使用水平下更具成本效益。通过根据预期使用情况对一段时间内的每个组件执行分析，工作负载可在其生命周期内保持成本效益。

在未建立这种最佳实践的情况下暴露的风险等级：中

实施指导

随着 AWS 发布新的服务和功能，适用于工作负载的最佳服务可能会发生变化。所需的工作量应体现可能带来的好处。工作负载审核频率取决于组织要求。如果工作负载的成本很高，则尽早实施新服务可最大限度地节省成本，因此提高审核频率可能是有利的。审核的另一个启动因素是使用模式发生变化。使用情况发生重大变化可能表明备用服务更加理想。

如果您需要将数据移入 AWS Cloud，可以选择 AWS 提供的任何种类的服务和合作伙伴工具，帮助您迁移数据集，无论它们是文件、数据库、机器映像、数据块卷，还是磁带备份。例如，要将大量数据移入和移出 AWS，或者要在边缘处理数据，可以使用 AWS 专用设备之一，经济高效地离线移动 PB 级数据。再例如，对于更高的数据传输率，直接连接服务可能比 VPN 更便宜，而 VPN 可以为业务提供所需的稳定一致的连接。

根据一段时间内针对不同使用情况的成本分析，审查您的扩缩活动。分析结果，看看是否可以调整扩缩策略，以增加具有多种实例类型和购买选项的实例。审查设置，看看是否可以减小最小值，用较小的实例集大小满足用户请求，并增加更多资源来满足预期的高需求。

通过与组织中的利益相关方讨论，对不同使用情况执行成本分析，并使用 AWS Cost Explorer 预测功能来预测服务变更的潜在影响。使用 AWS Budgets、CloudWatch 计费警报和 AWS Cost Anomaly Detection 监控使用水平启动因素，以尽早确定和实施最具成本效益的服务。

实施步骤

- 定义预计使用模式：与组织中的相关人员（例如营销和产品负责人）合作，记录哪些预期和预计使用模式适用于工作负载。与业务利益相关方讨论历史和预测的成本及使用量的增加，并确保增加幅度与业务要求一致。确定预计会有更多用户使用您的 AWS 资源的日历日、周或月，这表明您应该增加现有资源的容量，或采用额外的服务来降低成本和提高性能。
- 对预测的使用情况执行成本分析：使用定义的使用模式，在每个点执行分析。分析工作量应该体现潜在的成果，例如，如果使用情况变化很大，应执行彻底分析，以验证任何成本和变化。换句话说，当成本增加时，企业的使用量也应该增加。

资源

相关文档：

- AWS 总拥有成本（TCO）计算器
- Amazon S3 存储类
- 云产品
- Amazon EC2 Auto Scaling
- 云数据迁移
- AWS Snow Family

相关视频：

- AWS OpsHub for Snow Family

选择正确的资源类型、规模和数量

通过选择最佳的资源类型、规模和资源数量，可以以最低成本的资源满足技术要求。合理调整大小活动需要考虑工作负载的所有资源、各项资源的所有属性以及合理调整大小操作中涉及的工作量。合理调整大小可以是一个由使用模式变化和外部因素（如 AWS 价格下降或新 AWS 资源类型）启动的迭代过程。如果合理调整大小所需的成本超过工作负载整个生命周期中可能节省的成本，则合理调整大小也可以是一次性的。

AWS 中有许多不同的方法：

最佳实践

- COST06-BP01 执行成本建模
- COST06-BP02 根据数据选择资源类型、规模和数量
- COST06-BP03 根据指标自动选择资源类型、规模和数量
- COST06-BP04 考虑使用共享资源

COST06-BP01 执行成本建模

确定组织要求（如业务需求和现有承诺），并对工作负载及其每个组件进行成本建模（总体成本）。对不同预计负载下的工作负载执行基准测试活动，并比较成本。建模工作应该体现可能带来的好处，例如花费的时间与组件成本成正比。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

对工作负载及其每个组件执行成本建模，以了解资源之间的平衡，并在给定的具体性能水平下，确定工作负载中各项资源的适当规模。在评估计划的工作负载部署的价值实现结果时，了解成本考虑因素可为组织的业务案例和决策过程提供依据。

对不同预计负载下的工作负载执行基准测试活动，并比较成本。建模工作量应该体现可能带来的好处，例如花费的时间与组件成本或预计可节省的成本成正比。有关最佳实践，请参阅《AWS Well-Architected Framework 的性能效率支柱》的“审核”部分。

例如，要为由计算资源组成的工作负载创建成本模型，AWS Compute Optimizer 可以帮助实现对正在运行的工作负载的成本建模。它根据历史使用情况为计算资源提供合理调整大小的建议。确保将 CloudWatch 代理部署到 Amazon EC2 实例，以收集内存指标，这会通过 AWS Compute Optimizer 内更准确的建议为您提供帮助。这是计算资源的理想数据来源，因为它是一项免费服务，并且使用机器学习根据风险等级提出多项建议。

您可以使用多项服务将自定义日志作为数据来源，用于合理调整其他服务和工作负载组件（例如 AWS Trusted Advisor、Amazon CloudWatch 和 Amazon CloudWatch Logs）的操作的大小。AWS Trusted Advisor 会检查资源并标记利用率低的资源，这可以帮助您合理调整资源大小并创建成本模型。

以下是成本建模数据和指标的建议：

- 监控必须准确反映用户体验。为时间段选择正确的粒度，并仔细选择最大值或第 99 个百分位值而不是平均值。
- 为覆盖任何工作负载周期所需的分析时间段选择正确的粒度。例如，如果执行为期两周的分析，您可能会忽略高利用率的月度周期，这可能导致预置不足。
- 通过考虑现有的承诺、为其他工作负载选择的定价模式，以及更快地创新和专注于核心业务价值的能力，为计划的工作负载选择合适的 AWS 服务。

实施步骤

- 执行资源的成本建模：将工作负载或概念验证部署到具有特定资源类型和规模的单独账户，然后执行测试。使用测试数据运行工作负载，并记录输出结果以及测试运行时段的成本数据。然后，重新部署工作负载或更改资源类型和规模，并再次运行测试。在创建成本模型时，考虑您可能与这些资源一起使用的任何产品的许可费用，以及用于部署和管理这些资源的估计运营（劳动力或工程师）成本。考虑对一个时间段（每小时、每天、每月、每年或三年）进行成本建模。

资源

相关文档：

- AWS Auto Scaling
- 确定合理调整大小的机会
- Amazon CloudWatch 功能
- 成本优化：合理调整 Amazon EC2 的大小
- AWS Compute Optimizer
- AWS 定价计算器

相关示例：

- 执行数据驱动型成本建模
- 估算计划的 AWS 资源配置的成本
- 选择合适的 AWS 工具

COST06-BP02 根据数据选择资源类型、规模和数量

根据工作负载和资源特征的相关数据选择资源规模或类型，例如计算、内存、吞吐量或写入密集型资源。选择的依据通常是使用工作负载的上一个版本（本地版本）、文档或关于工作负载的其他信息源。

在未建立这种最佳实践的情况下暴露的风险等级：中

实施指导

Amazon EC2 提供许多实例类型，它们具有不同的 CPU、内存、存储和联网容量级别，适合不同的应用场景。这些实例类型具有不同的 CPU、内存、存储和联网容量组合，让您在为项目选择正确的资源组合时有更多的备选方案。每种实例类型都有多种大小，因此您可以根据工作负载的需求调整资源。要确定您需要哪种实例类型，请根据您计划在实例上运行的应用程序或软件，收集相关的系统要求详细信息。这些详细信息应包括以下各项：

- 操作系统
- CPU 核心数量
- GPU 核心数
- 系统内存（RAM）容量

- 存储类型和空间
- 网络带宽要求

确定计算要求的目的以及需要哪个实例，然后探索各种 Amazon EC2 实例系列。Amazon 提供以下实例类型系列：

- 通用
- 计算优化
- 内存优化
- 存储优化
- 加速计算
- HPC 优化型

要更深入地了解特定 Amazon EC2 实例系列可以实现的具体目的和应用场景，请参阅 AWS 实例类型。

想要选择最能满足需求的特定实例系列和实例类型，系统要求的收集至关重要。实例类型名称由系列名称和实例大小组成。例如，t2.micro 实例属于 T2 系列，而且是微型实例。

根据工作负载和资源特征（例如计算、内存、吞吐量或写入密集型）选择资源规模或类型。选择的依据通常是使用成本建模、工作负载的上一个版本（例如本地版本）、文档或关于工作负载的其他信息源（白皮书或发布的解决方案）。使用 AWS 定价计算器或成本管理工具有助于就实例类型、规模和配置作出明智的决策。

实施步骤

- 根据数据选择资源：使用成本建模数据选择预期的工作负载使用水平，然后选择指定的资源类型和规模。根据成本建模数据，确定虚拟 CPU 的数量、总内存（GiB）、本地实例存储卷（GB）、Amazon EBS 卷和网络性能级别，同时考虑实例所需的数据传输速率。务必根据详细的分析和准确的数据做出选择，这样不仅能够优化性能，还能有效管理成本。

资源

相关文档：

- AWS 实例类型

- AWS Auto Scaling
- Amazon CloudWatch 功能
- 成本优化：合理调整 EC2 的大小

相关视频：

- Selecting the right Amazon EC2 instance for your workloads
- Right size your service

相关示例：

- It just got easier to discover and compare Amazon EC2 instance types

COST06-BP03 根据指标自动选择资源类型、规模和数量

使用当前运行的工作负载的指标选择正确的规模和类型，从而优化成本。为计算、存储、数据和网络服务适当地预置吞吐量、规模和存储。这可以通过自动扩缩等反馈环路进行，也可以在工作负载中使用自定义代码来实现。

在未建立这种最佳实践的情况下暴露的风险等级：低

实施指导

在工作负载中创建一个反馈环路，此循环使用正在运行的工作负载中的活动指标来对该工作负载进行更改。您可以使用托管服务（例如 AWS Auto Scaling），您配置该服务来为您执行合理调整大小的操作。AWS 还提供 API、SDK 以及各类功能，让您可以毫不费力地修改资源。您可以对工作负载进行编程以停止和启动 Amazon EC2 实例，从而允许更改实例大小或实例类型。这带来双重好处：既合理调整了大小，又几乎消除了进行更改所需的全部运营成本。

某些 AWS 服务内置了自动类型或大小选项，如 Amazon Simple Storage Service Intelligent-Tiering。Amazon S3 Intelligent-Tiering 会根据使用模式，自动在两个访问层之间移动数据：频繁访问和不频繁访问。

实施步骤

- 通过配置工作负载指标来提高可观测性：捕获工作负载的关键指标。这些指标指明了客户体验（例如工作负载输出），并适应资源类型和规模之间的差异（例如 CPU 和内存使用情况）。对于计算

资源，请分析性能数据以合理调整 Amazon EC2 实例大小。确定空闲实例和未充分利用的实例。要查找的关键指标是 CPU 利用率和内存利用率（例如，在 90% 的时间内为 40% 的 CPU 利用率，如 Rightsizing with AWS Compute Optimizer and Memory Utilization Enabled 中所述）。确定四周内最大 CPU 使用率和内存利用率低于 40% 的实例。这些是大小适当能够降低成本的实例。对于诸如 Amazon S3 之类的存储资源，可以使用 Amazon S3 Storage Lens 存储统计管理工具，它允许您在存储桶级别查看不同类别的 28 个指标，并在控制面板中默认查看 14 天的历史数据。可以按摘要和成本优化或事件筛选 Amazon S3 Storage Lens 存储统计管理工具控制面板，以分析特定指标。

- 查看合理调整大小建议：使用 AWS Compute Optimizer 中的合理调整大小建议和成本管理控制台中的 Amazon EC2 合理调整大小工具，或者查看 AWS Trusted Advisor 如何合理调整资源大小以调整工作负载。无论是 Amazon EC2 实例、AWS 存储类还是 Amazon RDS 实例类型，都必须使用正确的工具来调整不同资源的大小，并遵循合理调整大小指南。对于存储资源，可以使用 Amazon S3 Storage Lens 存储统计管理工具，它可以让您了解对象存储使用情况、活动趋势，并提出可行的建议，以优化成本并应用数据保护最佳实践。使用 Amazon S3 Storage Lens 存储统计管理工具通过对整个组织指标的分析得出的上下文建议，您可以立即采取措施来优化存储。
- 根据指标自动选择资源类型和大小：使用工作负载指标，手动或自动选择工作负载资源。对于计算资源，配置 AWS Auto Scaling 或在应用程序中实现代码，可以减少频繁更改所需的工作量，而且实施更改的速度可能比手动操作更快。您可以启动并自动扩展单个 Auto Scaling 组中的一组按需实例和竞价型实例。除了享受使用竞价型实例的折扣外，您还可以使用预留实例或 Savings Plan 获得常规按需实例定价的折扣费率。将所有这些因素相结合，可帮助您优化 Amazon EC2 实例的成本节约，并确定应用程序所需的规模和性能。还可以在自动扩缩组（ASG）中使用基于属性的实例类型选择（ABS）策略，该策略允许您将实例要求表示为一组属性，例如 vCPU、内存和存储。您可以在发布新一代实例类型时自动使用它们，并通过 Amazon EC2 竞价型实例获得更广泛的容量。Amazon EC2 Fleet 和 Amazon EC2 Auto Scaling 会选择并启动符合指定属性的实例，无需手动选择实例类型。对于存储资源，可以使用 Amazon S3 Intelligent Tiering 和 Amazon EFS 不频繁访问功能，它们允许您自动选择存储类，以便在数据访问模式变更时自动节省存储成本，而不会影响性能或运营开销。

资源

相关文档：

- AWS Auto Scaling
- AWS 合理调整大小
- AWS Compute Optimizer
- Amazon CloudWatch 功能
- CloudWatch Getting Set Up

- CloudWatch Publishing Custom Metrics
- Getting Started with Amazon EC2 Auto Scaling
- Amazon S3 Storage Lens 存储统计管理工具
- Amazon S3 Intelligent-Tiering
- Amazon EFS 不频繁访问
- Launch an Amazon EC2 Instance Using the SDK

相关视频：

- Right Size Your Services

相关示例：

- 基于属性选择实例类型用于 Amazon EC2 Fleet 的自动扩缩
- Optimizing Amazon Elastic Container Service for cost using scheduled scaling
- Predictive scaling with Amazon EC2 Auto Scaling
- 利用 Amazon S3 Storage Lens 存储统计管理工具优化成本并深入了解使用情况

COST06-BP04 考虑使用共享资源

对于已经在组织级别为多个业务部门部署的服务，可以考虑使用共享资源来提高利用率，并降低总拥有成本（TCO）。使用共享资源是一种经济实惠的选择，可以通过使用现有解决方案和/或共享组件，实现集中管理和成本优化。在账户边界内或专用账户中管理监控、备份和连接等常见功能。还可以通过实施标准化、减少重复和降低复杂性来降低成本。

在未建立这种最佳实践的情况下暴露的风险等级：中

实施指导

如果多个工作负载产生相同的功能，则使用现有解决方案和共享组件，改善管理和优化成本。考虑使用现有资源（尤其是共享资源），例如非生产数据库服务器或目录服务，通过遵循安全最佳实践和组织法规来降低云成本。为了更好地实现价值和提高效率，将成本分配（使用对账和扣款）给产生使用情况的相关业务领域至关重要。

对账是指将云成本分为可归属类别的报告，例如消费者、业务部门、总分类账账户或其他责任实体。对账的目标是向团队、业务部门或个人展示其使用的云资源的成本。

扣款是指根据适合特定财务管理流程的策略，将中央服务支出分配给成本单位。对于客户而言，扣款会将一个共享服务账户产生的成本，分别计入适用于客户报告流程的不同财务成本类别。通过建立扣款机制，您可以报告不同业务部门、产品和团队产生的成本。

工作负载可以分类为关键和非关键工作负载。根据此分类，使用常规配置的共享资源来处理不太关键的工作负载。为了进一步优化成本，请仅为关键工作负载预留专用服务器。跨多个账户共享或预置资源，对资源进行高效管理。即使在不同的开发、测试和生产环境中，安全共享也是可行的，不会破坏组织结构。

为了加强您的了解，并优化容器化应用程序的成本和使用情况，请使用拆分成本分配数据，采用这种做法，有助于您根据应用程序对共享计算和内存资源的使用情况，将成本分配给各个业务实体。拆分成本分配数据可帮助您在容器工作负载中实现任务级对账和扣款，这些工作负载在 Amazon Elastic Container Service（Amazon ECS）或 Amazon Elastic Kubernetes Service（Amazon EKS）上运行。

对于分布式架构，请构建共享服务 VPC，集中访问每个 VPC 中的工作负载所需的共享服务。这些共享服务可以包括目录服务或 VPC 端点等资源。为了减少管理开销和成本，请从中心位置共享资源，而不是在每个 VPC 中构建资源。

使用共享资源时，可以节省运营成本、大幅提高资源利用率并提高一致性。在多账户设计中，可以集中托管一些 AWS 服务，并在中心使用多个应用程序和账户访问这些服务，此举有助于节省成本。可以使用 AWS Resource Access Manager（AWS RAM）共享其它公用资源，如 VPC subnets and AWS Transit Gateway attachments、AWS Network Firewall 或 Amazon SageMaker AI pipelines。在多账户环境中，使用 AWS RAM 一次性创建资源并与其他账户共享。

组织应有效地标记共享成本，并确认其没有大量成本未标记或未分配。如果您未有效地分配共享成本，也没有人对共享成本管理负责，则共享云成本可能会急剧增加。您应当知道自己在资源、工作负载、团队或组织层面产生了哪些成本，因为这些信息有助于您进一步了解相较于实现的业务成果，在适当层面所实现的价值。最终，通过共享云基础设施，组织将实现成本节省。建议对共享云资源进行成本分配，以优化云支出。

实施步骤

- 评估现有资源：审核对您的工作负载使用类似服务的现有工作负载。根据工作负载的组件，如果业务逻辑或技术要求允许，考虑现有平台。
- 在 AWS RAM 中使用资源共享并进行相应的限制：使用 AWS RAM 与组织内的其他 AWS 账户共享资源。共享资源时，无需在多个账户中重复构建资源，这样可以大幅减少资源维护的运营负担。此过程还有助于您与账户中的角色和用户，以及与其他 AWS 账户 安全地共享已创建的资源。

- 标记资源：标记比较适合成本报告的资源，并将其在成本类别中归类。激活这些与成本相关的资源标签来进行成本分配，以实现 AWS 资源使用情况的可见性。专注于在成本和使用情况可见性方面创建适当的粒度级别，并通过成本分配报告和 KPI 跟踪来影响云使用行为。

资源

相关最佳实践：

- SEC03-BP08 在组织内安全地共享资源

相关文档：

- 什么是 AWS Resource Access Manager？
- AWS services that you can use with AWS Organizations
- Shareable AWS resources
- AWS Cost and Usage (CUR) Queries

相关视频：

- AWS Resource Access Manager - granular access control with managed permissions
- How to design your AWS cost allocation strategy
- AWS Cost Categories

相关示例：

- How-to chargeback shared services: An AWS Transit Gateway example
- How to build a chargeback/showback model for Savings Plans using the CUR
- Using VPC Sharing for a Cost-Effective Multi-Account Microservice Architecture
- Improve cost visibility of Amazon EKS with AWS Split Cost Allocation Data
- 利用 AWS 拆分成本分配数据提高 Amazon ECS 和 AWS Batch 的成本可见性

选择最佳定价模式

执行工作负载成本建模：考虑工作负载组件的要求并了解潜在的定价模式。定义组件的可用性要求。确定工作负载中是否存在执行功能的多个独立资源，以及工作负载随着时间推移的要求。使用默认的按需定价模式和其他适用模式比较资源成本。考虑资源或工作负载组件的任何潜在更改。

定期执行账户级别的分析：定期执行成本建模可确保能够跨多个工作负载进行优化。例如，如果总体上按需使用多个工作负载，则变更的风险较低，并且实施基于承诺的折扣可降低总体成本。建议每两周到一个月定期执行一次分析。此分析让您可以进行少量调整性采购，因此定价模式的覆盖范围会随着工作负载及其组件的变化而不断变化。

使用 AWS Cost Explorer 建议工具寻找享受承诺折扣的机会。

要为竞价型实例工作负载寻找机会，请查看总体使用情况的小时视图，并确定使用情况或弹性的定期变化周期。

定价模式：AWS 有多种定价模式，支持以符合组织需求、最具成本效益的方式支付资源费用。以下部分介绍了各种采购模式：

- 按需型实例
- 竞价型实例
- 承诺折扣 – 节省计划
- 承诺折扣 – 预留实例/容量
- 地理选择
- 第三方协议和定价

按需型实例：这是默认的即用即付定价模式。当您使用资源（如按需使用 EC2 实例或 DynamoDB 等服务）时，可以按小时支付固定费用，并且无需长期承诺。可以根据应用程序的需求增加或减少资源或服务的容量。按需模式有小时费率，但是根据服务的不同，可以以 1 秒为单位计费（例如 Amazon RDS 或 Linux EC2 实例）。推荐以下应用程序使用按需模式：具有定期出现峰值的短期工作负载（例如为期四个月的项目），或无法中断的不可预测工作负载。按需模式还适用于要求运行时不间断，但运行时间不够长、无法享受承诺折扣（节省计划或预留实例）的工作负载，例如预生产环境。

竞价型实例：竞价型实例是备用的 Amazon EC2 计算容量，可提供按需价格高达 90% 的折扣，无需长期承诺。借助竞价型实例，可以在相同的预算下显著降低运行应用程序的成本，或扩展应用程序的计算容量。不同于按需模式，如果 Amazon EC2 需要恢复容量，或者竞价型实例的价格超出配置的价格，可以中断竞价型实例并触发 2 分钟的警告。平均而言，竞价型实例的中断时间在 5% 以下。

如果有队列或缓冲区，或者有多个资源独立处理请求（例如 Hadoop 数据处理），竞价型实例是理想之选。通常，这些工作负载无状态且灵活，具备容错能力，例如批处理、大数据和分析、容器化环境和高性能计算（HPC）。非关键工作负载（例如测试和开发环境）也适合使用竞价型实例。

竞价型实例也已集成到多项 AWS 服务中，例如 Amazon EC2 Auto Scaling 组、Amazon EMR、Amazon Elastic Container Service（Amazon ECS）和 AWS Batch。

如果需要回收竞价型实例，Amazon EC2 会通过 CloudWatch Events 传递的竞价型实例中断通知以及实例元数据，发送一段两分钟警告。在这两分钟内，应用程序可以利用这段时间保存其状态、耗尽运行的容器、上传最终日志文件，或从负载均衡器中删除自己。两分钟结束时，可以选择休眠、停止或终止竞价型实例。

在工作负载中采用竞价型实例时，请考虑以下最佳实践：

- 在尽可能多的实例类型之间保持灵活性：在实例类型的系列和规模方面都要灵活，以提高达到目标容量要求的可能性、获得可能的最低成本，并最大限度减小中断的影响。
- 灵活安排工作负载的运行位置：可用容量可能因可用区而异。这样做可提高实现目标容量的可能性（因为会利用多个备用容量池），并最大限度降低成本。
- 连续性设计：针对无状态和容错性设计工作负载，这样即使一些 EC2 容量被中断，也不会影响工作负载的可用性或性能。
- 我们建议将竞价型实例与按需型和节省计划/预留实例相结合，以最大限度优化工作负载成本并提高性能。

承诺折扣 – 节省计划：AWS 通过保留或承诺使用一定数量的资源并为您的资源收取折扣价，为您提供多种降低成本的方法。通过节省计划，可以在一年或三年的时间里保证每小时的花费，从而享受所有资源的折扣价。节省计划提供 Amazon EC2、AWS Fargate 和 AWS Lambda 等 AWS Compute 服务的折扣。当您做出承诺时，您每小时支付该承诺金额，并且其会以折扣率从您的按需使用费中扣除。例如，您承诺每小时 50 美元，按需使用费每小时 150 美元。考虑到节省计划定价，特定使用费都可享受 50% 的折扣。因此，50 美元的承诺额涵盖 100 美元的按需使用费。您将支付 50 美元（承诺）和剩余的 50 美元按需使用费。

计算类节省计划是最灵活的方案，提供高达 66% 的折扣。它们自动跨可用区、实例大小、实例系列、操作系统、租约、区域和计算服务应用。

实例节省计划不太灵活，但提供更高的折扣率（高达 72%）。它们自动跨可用区、实例大小、操作系统和租约应用。

有三种付款选项：

- 无预先付款：没有预先付款；然后，每个月按逐渐减少的每小时费率为当月的总小时数支付费用。
- 部分预先付款：相比无预先付款，折扣率更高。为部分使用量预先支付费用；然后，每个月按逐渐减少的每小时费率为当月的总小时数支付费用。
- 全额预先付款：预先支付整个期间使用量的费用，在承诺涵盖的剩余使用期限内不会产生任何其他费用。

可以跨工作负载任意组合使用这三种购买选项。

节省计划先应用到购入它的账户中的使用量（从最高折扣率到最低折扣率），然后应用到所有其他账户的合并使用量（从最高折扣率到最低折扣率）。

建议在没有使用量或资源的账户（例如管理账户）中购买所有节省计划。这可以确保节省计划适用于所有使用情况下的最高折扣率，最大限度提高折扣金额。

工作负载和使用量通常会随时间而改变。建议随着时间的推移持续购买少量的节省计划承诺使用量。这样可确保您保持较高的覆盖率，以最大限度提高折扣，并且计划始终与工作负载和组织要求紧密匹配。

由于折扣可能发生变化，请不要在账户中设置目标覆盖率。覆盖率低并不一定意味着潜在节省费用高。您的账户的覆盖率可能很低，但如果使用量由小型实例组成，并且使用许可的操作系统，则节省的费用可能仅为几个百分点。相反，在节省计划建议工具中跟踪和监控可能节省的费用。经常查看 Cost Explorer 中的节省计划建议（执行定期分析），并继续购买承诺使用量，直到估计的节省额低于组织所需的折扣为止。例如，跟踪并监控潜在折扣是否保持在 20% 以下（如果高于该值，则必须购买）。

监控利用率和覆盖率，但仅用于检测更改。不要以特定的利用率或覆盖率作为目标，因为这并不一定与节省额成比例。确保购买节省计划可增加覆盖率，如果覆盖率或使用率降低，则确保对其进行量化并了解相关情况。例如，您将工作负载资源迁移到较新的实例类型，这会减少现有计划的利用率，但是性能效益胜于节省的成本。

承诺折扣 – 预留实例/承诺：与节省计划类似，预留实例（RI）为做出最低资源运行量承诺的用户提供高达 72% 的折扣。预留实例可用于 Amazon RDS、Amazon OpenSearch Service、Amazon ElastiCache、Amazon Redshift 和 DynamoDB。Amazon CloudFront 和 AWS Elemental MediaConvert 也会在您做出最低使用量承诺时提供折扣。预留实例当前可用于 Amazon EC2，但是，节省计划提供相同的折扣级别，同时具有更大的灵活性，且没有管理开销。

预留实例提供相同的定价选项：无预先付款、部分预先付款和全额预先付款，并且期限相同：一年或三年。

可以在某个区域或特定可用区购买预留实例。在可用区中购买时，它们提供容量预留。

虽然 Amazon EC2 具有可转换的 RI，但应该对所有 EC2 实例使用节省计划，因其增加了灵活性并降低了运营成本。

应使用相同的流程和指标来跟踪和购买预留实例。建议不要跨账户跟踪 RI 的覆盖率。同时，建议不要监控或跟踪利用率百分比，而应在 Cost Explorer 中查看利用率报告，并使用表中的净节省额列。如果净节省额是很大的负数，则必须采取措施补救未使用的 RI。

EC2 实例集：EC2 实例集可用于定义目标计算容量，然后为实例集指定实例类型以及按需型和竞价型实例的余额。EC2 实例集将自动启动最低价格的资源组合，以满足定义的容量。

地理选择：在构建解决方案时，最佳实践是设法将计算资源放在更接近用户的位置，以提供更低的延迟和强大的数据主权。对于全球用户，您应该使用多个位置来满足这些需求。您应该选择尽可能降低成本的地理位置。

AWS Cloud 基础架构围绕区域和可用区构建。区域是指全球范围内的某个物理位置，每个区域由多个可用区组成。可用区由一个或多个分散的数据中心组成，每个都拥有独立的配套设施，其中包括冗余电源、联网和连接。

每个 AWS 区域 都在当地市场条件下运营，每个区域的资源定价可能不同。选择特定区域来运行解决方案组件或整个解决方案，以便可以在全球范围内以尽可能低的价格运行。可以使用 AWS Simple Monthly Calculator 来估算各区域中工作负载的成本。

第三方协议和定价：当您在云中使用第三方解决方案或服务时，确保定价结构与成本优化结果保持一致非常重要。定价应与其带来的结果和价值成比例。这方面的一个例子是可带来一定百分比节省额的软件，节省额（结果）越高，其价格也就越高。除非您能提供特定账单每一部分的结果，否则与账单成比例的协议通常不会与成本优化保持一致。例如，如果您使用的其他服务没有带来任何益处，提供 Amazon EC2 相关建议并按整体账单一定比例收取费用的解决方案将会变得更加昂贵。另一个示例是根据所托管资源的成本按一定百分比收费的托管服务。实例越大并不一定意味着需要更多的管理工作，但会收取更多费用。确保这些服务定价安排包括成本优化计划或服务中的功能，以提高效率。

最佳实践

- COST07-BP01 执行定价模式分析
- COST07-BP02 根据成本选择区域
- COST07-BP03 选择具有经济实惠条款的第三方协议
- COST07-BP04 针对此工作负载的所有组件实施定价模式
- COST07-BP05 在管理账户级别执行定价模式分析

COST07-BP01 执行定价模式分析

分析工作负载的每个组件。确定组件和资源是长时间运行（享受承诺折扣），还是短时间动态运行（采用竞价型或按需型实例定价）。使用成本管理工具中的建议对工作负载执行分析，并对这些建议应用业务规则以实现高回报。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

AWS 有多种定价模式，您可以根据产品，以符合组织需求且最具成本效益的方式支付资源费用。与团队合作确定最适合的定价模式。定价模式通常包含多种选项的组合，根据供应情况来决定

按需型实例允许您只需要按小时或按秒（最低 60 秒）支付计算或数据库容量费用，具体取决于您运行的实例，无需做出长期承诺或预先付款。

节省计划是灵活的定价模式，支持以较低的 Amazon EC2、Lambda 和 AWS Fargate 使用价格换取一年或三年期的持续使用量承诺（以美元/小时为单位）。

竞价型实例是一种 Amazon EC2 定价机制，允许您以折扣小时费率（最高比按需价格低 90%）申请备用计算容量，无需预先承诺。

预留实例通过预付容量费用，允许您享受高达 75% 的折扣。有关更多详细信息，请参阅使用预留优化成本。

可以选择为与生产、质量和开发环境关联的资源纳入节省计划。另外，由于沙盒资源仅在需要时才会启用，因此您可以为该环境中的资源选择按需模式。使用 Amazon 竞价型实例降低 Amazon EC2 成本，或者使用计算类节省计划降低 Amazon EC2、Fargate 和 Lambda 成本。AWS Cost Explorer 建议工具提供通过节省计划获得承诺折扣的机会。

如果您过去曾购买过 Amazon EC2 的预留实例，或者组织内部已经建立成本分配实践，则可以暂时继续使用 Amazon EC2 预留实例。但是，我们建议制定一项战略，以便在未来使用节省计划，它是一种更灵活的成本节省机制。您可以在 AWS Cost Management 中刷新节省计划（SP）建议，以便随时生成新的节省计划建议。使用预留实例（RI）降低 Amazon RDS、Amazon Redshift、Amazon ElastiCache 和 Amazon OpenSearch Service 成本。节省计划和预留实例提供三个选项：全额预付、部分预付和无预付款。使用 AWS Cost Explorer RI 和 SP 购买建议中提供的建议。

要为竞价型实例工作负载寻找机会，请查看总体使用情况的小时视图，并确定使用情况或弹性的定期变化周期。您可以为各种容错和灵活的应用程序使用 Spot 实例。示例包括无状态 Web 服务器、API 端点、大数据和分析应用程序、容器化工作负载、CI/CD 及其他灵活工作负载。

分析 Amazon EC2 和 Amazon RDS 实例，确定在不使用时（下班后和周末）是否可以将其关闭。与 24/7 全天候使用这些实例相比，您可以使用这种方法将成本降低 70% 或更多。如果您的 Amazon Redshift 集群只需要在特定时间内可用，则可以暂停集群并在稍后恢复它。停止 Amazon Redshift 集群或 Amazon EC2 和 Amazon RDS 实例时，计算计费会停止，仅收取存储费用。

请注意，按需容量预留（ODCR）不是定价折扣。无论您是否在预留容量中运行实例，都按等同的按需费率为您计算容量预留费用。当您需要为计划运行的资源提供足够的容量时，应该考虑使用这种模式。ODCR 不必做出长期承诺，因为当您不再需要时可以取消它们，但它们也可以享受节省计划或预留实例提供的折扣。

实施步骤

- 分析工作负载弹性：在 Cost Explorer 中使用每小时粒度或者使用自定义控制面板分析工作负载弹性。确定正在运行的实例数量的规律性变化。短期实例比较适合采用竞价型实例或竞价型实例集。
  - Well-Architected Lab：Cost Explorer
  - Well-Architected Lab：成本可视化
- 查看现有定价合同：查看当前合同或对长期需求的承诺。分析您目前拥有的承诺以及这些承诺中有多少正在使用。利用已有的合同折扣或企业协议。企业协议让客户可以选择定制最适合其需求的协议。对于长期承诺，请考虑预留定价折扣、特定实例类型的预留实例或节省计划、实例系列、AWS 区域和可用区。
- 执行承诺折扣分析：使用账户中的 Cost Explorer 查看节省计划和预留实例建议。要验证您是否实施了具有所需折扣和风险的正确建议，请按照 Well-Architected Lab 的说明操作。

资源

相关文档：

- Accessing Reserved Instance recommendations
- 实例购买选项
- AWS Enterprise

相关视频：

- Save up to 90% and run production workloads on Spot

相关示例：

- Well-Architected Lab：Cost Explorer
- Well-Architected Lab：成本可视化
- Well-Architected Lab：定价模式

COST07-BP02 根据成本选择区域

资源定价在每个区域中可能各不相同。确定区域成本差异，仅当需要满足延迟、数据驻留和数据主权要求时，才在成本较高的区域部署。考虑区域成本有助于您为此工作负载支付最低的总体费用。

在未建立这种最佳实践的情况下暴露的风险等级：中

实施指导

AWS Cloud 基础设施是全球性的，托管在世界各地的多个地点，围绕 AWS 区域、可用区、Local Zones、AWS Outposts 和 Wavelength Zones 构建。区域是世界上的一个物理位置，每个区域都是一个独立的地理位置，而 AWS 有多个可用区。可用区是每个区域内多个相互隔离的位置，由一个或多个独立的数据中心组成，每个数据中心都拥有冗余的电力、联网和连接。

每个 AWS 区域 都在当地市场条件下运作，由于土地、光纤、电力和税收等成本的不同，每个区域的资源定价也不同。选择特定区域来运行解决方案组件或整个解决方案，以便可以在全球范围内以尽可能低的价格运行。使用 AWS Calculator 按位置类型（区域、Wavelength Zone 和 Local Zone）和区域搜索服务，估算工作负载在不同区域的成本。

在构建解决方案时，最佳实践是设法将计算资源放在更接近用户的位置，以提供更低的延迟和强大的数据主权。根据您的业务、数据隐私、性能和安全要求来选择地理位置。对于具有全球终端用户的应用程序，请使用多个位置。

如果您不存在数据隐私、安全和业务要求方面的义务，请使用 AWS 服务价格较低的区域来部署工作负载。例如，如果您的默认区域是亚太地区（悉尼）（`ap-southwest-2`），并且不存在使用其他区域的限制（例如数据隐私、安全），则在美国东部（弗吉尼亚州北部）（`us-east-1`）部署非关键（开发和测试）Amazon EC2 实例成本更少。

	合规性	延迟	成本	服务/功能
区域 1	✓	15 ms	$$	✓
区域 2	✓	20 ms	$$$	X
区域 3	✓	80 ms	$	✓
区域 4	✓	15 ms	$$	✓
区域 5	✓	20 ms	$$$	X
区域 6	✓	15 ms	$	✓
区域 7	✓	80 ms	$	✓
区域 8	✓	15 ms	$	X

区域特征矩阵表

前面的矩阵表显示，区域 6 是这种给定场景的最佳选择，因为与其他区域相比，其延迟较低，可以使用相应服务，而且是最便宜的区域。

实施步骤

- 查看 AWS 区域 定价：分析当前区域的工作负载成本。首先使用按服务和使用类型划分的最高成本，计算其他可用区域的成本。如果预测的节省额超过迁移组件或工作负载的成本，则迁移到新区域。
- 审查多区域部署的要求：分析您的业务要求和义务（数据隐私、安全或性能），确认是否存在任何会阻止您使用多个区域的限制。如果不存在要求您只能使用单个区域的限制，则使用多个区域。
- 分析所需的数据传输：在选择区域时考虑数据传输成本。让您的数据靠近客户和资源。在数据流动和数据传输非常少时，选择成本较低的 AWS 区域。根据数据传输业务要求，您可以使用 Amazon CloudFront、AWS PrivateLink、AWS Direct Connect 和 AWS Virtual Private Network 来降低联网成本、提高性能和增强安全性。

资源

相关文档：

- Accessing Reserved Instance recommendations
- Amazon EC2 定价
- 实例购买选项
- 区域表

相关视频：

- Save up to 90% and run production workloads on Spot

相关示例：

- Overview of Data Transfer Costs for Common Architectures
- Cost Considerations for Global Deployments
- What to Consider when Selecting a Region for your Workloads

COST07-BP03 选择具有经济实惠条款的第三方协议

经济实惠的协议和条款可确保这些服务的成本与所提供的效益相称。选择与可为组织带来额外效益相称的协议和定价。

在未建立这种最佳实践的情况下暴露的风险等级：中

实施指导

市场上有多种产品可用于管理云环境中的成本。根据客户需求，它们在功能方面可能有一些差异，例如有些侧重于成本治理或成本可见性，而有些则侧重于成本优化。想要得到有效的成本优化和治理，一个关键因素是使用具有必要功能的正确工具和正确的定价模式。这些产品有不同的定价模式。有些产品按每月账单的一定比例收费，有些则按所实现节省额的一定比例收费。理想情况下，您应该只为需要的资源付费。

当您在云中使用第三方解决方案或服务时，确保定价结构契合期望结果非常重要。定价应与其带来的结果和价值成比例。例如，可带来一定百分比节省额的软件中，节省额（结果）越高，其价格也就越高。许可协议规定，随着开支的增加，您需要支付更多的费用，这可能并不总是有利于您优化成本。但是，如果供应商为您的账单的所有部分都提供明显的效益，那么这种按比例收费可能是合理的。

例如，如果您使用的其他服务没有带来任何效益，提供 Amazon EC2 相关建议并按整体账单一定比例收取费用的解决方案将会变得更加昂贵。另一个示例是根据托管资源的成本，按一定百分比收费的托管

服务。实例越大并不一定意味着需要更多的管理工作，但会收取更多费用。验证这些服务定价安排是否在其服务中包含了成本优化计划或功能，以提高效率。

客户可能会发现市场上的这些产品更先进或更易于使用。您需要考虑这些产品的成本，并考虑长期潜在的成本优化结果。

实施步骤

- 分析第三方协议和条款：查看第三方协议中的定价。基于不同的使用水平执行建模，并将新成本（例如使用新服务，或当前服务由于工作负载增长导致使用量增加）纳入考量。确定额外成本能否为业务提供所需效益。

资源

相关文档：

- Accessing Reserved Instance recommendations
- 实例购买选项

相关视频：

- Save up to 90% and run production workloads on Spot

COST07-BP04 针对此工作负载的所有组件实施定价模式

永久运行的资源应利用预留容量，如节省计划或预留实例。短期容量配置为使用竞价型实例或竞价型实例集。按需型实例仅用于无法中断并且运行时长不足以使用预留容量的短期工作负载（根据具体的资源类型，时长介于 25% 到 75% 之间）。

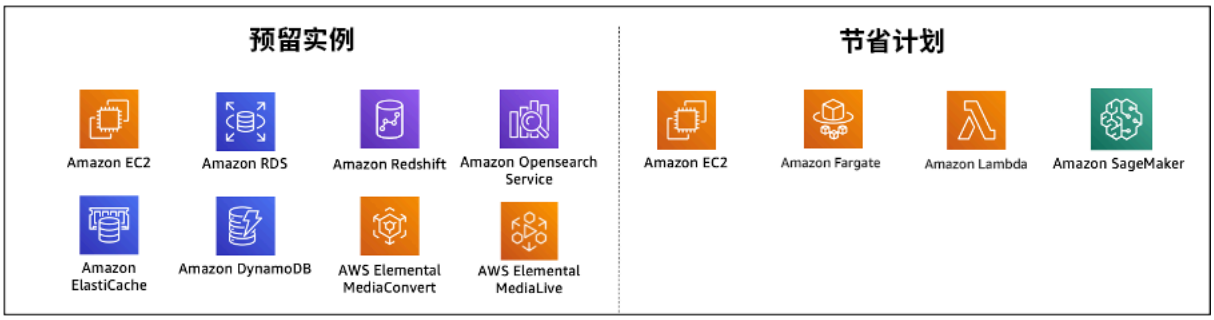
在未建立这种最佳实践的情况下暴露的风险等级：低

实施指导

为了提高成本效率，AWS 根据过去的使用情况提供多项承诺建议。您可以使用这些建议，了解可以节省的金额以及如何使用承诺。使用这些服务的方式可以是按需型实例、竞价型实例，或在一定时间内承诺使用量，通过预留实例（RI）和节省计划（SP）降低按需成本。您不仅需要了解每个工作负载组件和多项 AWS 服务，还需要了解这些服务的承诺折扣、购买选项和竞价型实例，以优化工作负载。

考虑工作负载组件的要求，并了解这些服务的不同定价模式。定义这些组件的可用性要求。确定工作负载中是否存在执行功能的多个独立资源，以及工作负载随着时间推移的要求。使用默认的按需定价模式和其他适用模式比较资源成本。考虑资源或工作负载组件的任何潜在更改。

例如，让我们看一下 AWS 上的这个 Web 应用程序架构。此示例工作负载由多项 AWS 服务组成，例如 Amazon Route 53、AWS WAF、Amazon CloudFront、Amazon EC2 实例、Amazon RDS 实例、负载均衡器、Amazon S3 存储和 Amazon Elastic File System（Amazon EFS）。您需要审查每项服务，并确定使用不同定价模式带来的潜在成本节省机会。其中一些服务可能有资格使用 RI 或 SP，而另一些可能只能使用按需方案。如下图所示，一些 AWS 服务可以使用 RI 或 SP 来承诺。



使用预留实例和节省计划承诺的 AWS 服务

实施步骤

- 实施定价模式：使用分析结果，购买节省计划、预留实例或实施竞价型实例。如果是首次承诺购买，请选择列表中的前 5 项或 10 项建议，然后在接下来的一到两个月内监控并分析结果。AWS Cost Management Console 会指导您完成整个过程。查看控制台中的 RI 或 SP 建议，自定义建议（类型、付款和期限），查看每小时承诺（例如每小时 20 美元），然后添加到购物车。折扣将自动应用于符合条件的使用情况。定期购买少量承诺折扣（例如每 2 周或每月一次）。对可以中断或者无状态的工作负载实施竞价型实例。最后，选择按需型 Amazon EC2 实例并为其余需求分配资源。
- 工作负载审核周期：实施工作负载审核周期，用于专门分析定价模式覆盖范围。工作负载达到所需覆盖范围后，可部分（每几个月）购买额外的承诺折扣，或者随着组织的使用情况变化进行购买。

资源

相关文档：

- Understanding your Savings Plans recommendations
- Accessing Reserved Instance recommendations
- 如何购买预留实例
- 实例购买选项

- 竞价型实例
- 其他 AWS 服务的预留模型
- Savings Plans Supported Services

相关视频：

- Save up to 90% and run production workloads on Spot

相关示例：

- What should you consider before purchasing Savings Plans?
- How can I use Cost Explorer to analyze my spending and usage?

COST07-BP05 在管理账户级别执行定价模式分析

检查账单和成本管理工具，并查看承诺和预留的建议折扣，以便在管理账户级别执行定期分析。

在未建立这种最佳实践的情况下暴露的风险等级：低

实施指导

定期执行成本建模可帮助您跨多个工作负载进行优化。例如，如果总体上多个工作负载使用按需型实例，则变更的风险较低，并且实施基于承诺的折扣可降低总体成本。建议每两周到一个月定期执行一次分析。这样您可以进行少量调整性采购，因此定价模式的覆盖范围会随着工作负载及其组件的变化而不断变化。

使用 AWS Cost Explorer 建议工具，寻找在管理账户中享受承诺折扣的机会。在计算管理账户级别的建议时，考虑 AWS 组织中具有预留实例（RI）或节省计划（SP）的所有账户的使用情况。它们也是在启用折扣共享的情况下计算的，以便推荐可在不同账户中最大限度实现节省的承诺。

虽然在许多情况下，在管理账户级别购买可以进行优化以最大限度地节省开支，但在某些情况下，您可能会考虑在关联账户级别购买 SP，例如当您希望首先对该特定关联账户中的资源使用情况应用折扣时。成员账户建议在个人账户级别计算，以最大限度地节省每个独立账户的成本。如果您的账户同时拥有 RI 和 SP 承诺，则它们将按以下顺序应用：

1. 区域 RI
2. 标准 RI

3. 可转换 RI
4. 实例节省计划
5. 计算类节省计划

如果您在管理账户级别购买 SP，则将根据从最高到最低的折扣百分比来应用节省额。管理账户级别的 SP 会查看所有关联账户，并在任何可以实现最高折扣的地方应用节省额。如果您希望限制应用节省额的位置，则可以在关联账户级别购买节省计划，而只要该账户在运行符合条件的计算服务，就会首先在其上应用折扣。当账户未运行符合条件的计算服务时，将在同一管理账户下的其他关联账户之间共享折扣。默认情况下，折扣共享处于启用状态，不过您可以根据需要关闭。

在整合账单系列中，节省计划首先应用于所有者账户的使用量，然后应用于其他账户的使用量。只有在启用共享时才会发生这种情况。您的节省计划将首先应用于您的最高节省百分比。如果有多个节省百分比相等的使用量，则节省计划将应用于节省计划费率最低的第一个使用量。节省计划将继续适用，直到没有剩余的使用情况或承诺用量用完为止。任何剩余的使用量按照按需费率收费。您可以在 AWS Cost Management 中刷新节省计划建议，以便随时生成新的节省计划建议。

分析实例的灵活性之后，可以按照建议进行承诺。创建成本模型：使用可能的不同资源选项分析工作负载的短期成本、分析 AWS 定价模式并使其与业务要求保持一致，以发现总拥有成本和成本优化机会。

实施步骤

执行承诺折扣分析：使用账户中的 Cost Explorer 查看节省计划和预留实例建议。确保了解节省计划建议，估算每月支出以及每月节省额。查看管理账户级别的建议，这些建议根据 AWS 组织中启用了 RI 或节省计划折扣共享的所有成员账户的使用情况计算得出，以便各账户实现最大限度的节省。您可以按照 Well-Architected Lab 的说明操作，确认您实施了具有所需折扣和风险的正确建议。

资源

相关文档：

- AWS 如何定价？
- 实例购买选项
- Saving Plan Overview
- Saving Plan recommendations
- Accessing Reserved Instance recommendations
- Understanding your Saving Plans recommendation
- How Savings Plans apply to your AWS usage

- 节省计划与整合账单
- 开启共享的预留实例和节省计划折扣

相关视频：

- Save up to 90% and run production workloads on Spot

相关示例：

- 在购买节省计划之前我应该考虑什么？
- How can I use rolling Savings Plans to reduce commitment risk?
- 何时使用竞价型实例

制定数据传输计划

云的优势之一在于它是一种托管的网络服务。不再需要管理和操作一组交换机、路由器和其他相关的网络设备。云中网络资源的耗用和付费方式与您为 CPU 和存储器付费的方式相同，即只需为使用的资源付费。要在云中优化成本，必须高效利用网络资源。

最佳实践

- COST08-BP01 执行数据传输建模
- COST08-BP02 选择组件以便优化数据传输成本
- COST08-BP03 实施服务以便降低数据传输成本

COST08-BP01 执行数据传输建模

收集组织要求，并对工作负载及其每个组件执行数据传输建模。这样可以确定满足当前数据传输要求的最低成本点。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

在云端设计解决方案时，由于习惯于设计使用本地数据中心的架构或者缺乏相应的知识，数据传输费用常常会被忽视。AWS 中的数据传输费用由流量的来源、目的地和数量决定。在设计阶段将这些费用考虑在内可以节省成本。了解工作负载中的哪些环节需要进行数据传输、传输成本及其相关好处，对于准

确估算总拥有成本（TCO）非常重要。因此，您可以作出明智的决定来修改或接受架构决策。例如，您可能有一个多可用区配置，可以在可用区之间复制数据。

您可以对在工作负载中传输数据的服务组件进行建模，并确定这是可接受的成本（类似于在两个可用区中支付计算和存储费用），以实现所需的可靠性和韧性。对不同使用水平的成本进行建模。工作负载的使用情况可能随时间而变化，不同的服务可能在不同的水平上更具有成本效益。

在对数据传输进行建模时，需考虑接收了多少数据以及这些数据来自哪里。此外，还需考虑处理了多少数据以及需要多少存储或计算容量。在建模期间，应遵循工作负载架构的联网最佳实践，以优化潜在的数据传输成本。

可利用 AWS 定价计算器 查看特定 AWS 服务和预期数据传输的估计成本。如果您已有工作负载在运行（用于测试目的或在预生产环境中），请使用 AWS Cost Explorer 或 AWS 成本和使用情况报告（CUR）来了解数据传输成本并对其建模。配置概念验证（PoC）或测试您的工作负载，并在实际的模拟负载下运行测试。可以根据不同的工作负载需求对成本进行建模。

实施步骤

- 确定要求：计划在来源和目的地之间传输数据的主要目标和业务要求是什么？ 预期的最终业务成果是什么？ 收集业务要求并定义预期结果。
- 确定来源和目的地：数据传输的数据来源和目的地是什么，例如在 AWS 区域 内部、传输到 AWS 服务或传出到互联网？
  - AWS 区域 内部的数据传输
  - AWS 区域 之间的数据传输
  - 将数据传输到互联网
- 确定数据分类：此数据传输的数据分类是什么？ 这是什么样的数据？ 数据量有多大？ 必须多久传输一次数据？ 数据是否敏感？
- 确定要使用的 AWS 服务或工具：此数据传输使用哪些 AWS 服务？ 是否可以将已经预置的服务用于其他工作负载？
- 计算数据传输成本：结合使用 AWS 定价和您之前创建的数据传输建模，计算工作负载的数据传输成本。计算不同使用水平的数据传输成本，包括工作负载使用情况的增加和减少这两种情况。如果工作负载架构有多个传输选项，则计算每个选项的成本以便进行比较。
- 将成本与成果关联：对于产生的每项数据传输成本，指定其实现的工作负载成果。如果在组件之间传输，可能是为了实现解耦；如果在可用区之间传输，则可能是为了实现冗余。
- 创建数据传输建模：收集所有信息后，为多个应用场景和不同的工作负载创建概念性基础数据传输建模。

资源

相关文档：

- AWS 缓存解决方案
- AWS 定价
- Amazon EC2 定价
- Amazon VPC 定价
- Understanding data transfer charges

相关视频：

- Monitoring and Optimizing Your Data Transfer Costs
- S3 Transfer Acceleration

相关示例：

- Overview of Data Transfer Costs for Common Architectures
- AWS Prescriptive Guidance for Networking

COST08-BP02 选择组件以便优化数据传输成本

选择所有组件然后设计架构，以便降低数据传输成本。其中包括使用广域网（WAN）优化和多可用区（AZ）配置等组件

在未建立这种最佳实践的情况下暴露的风险等级：中

实施指导

针对数据传输进行构建，可最大限度地降低数据传输成本。这可能涉及使用内容分发网络来定位更靠近用户的数据，或者使用从您的本地设施到 AWS 的专用网络链接。您还可以使用 WAN 优化和应用程序优化来减少组件之间传输的数据量。

向 AWS Cloud 传输数据或在其中传输数据时，必须根据不同的应用场景、数据的性质和可用的网络资源来了解目的地，以便选择正确的 AWS 服务来优化数据传输。AWS 提供一系列针对不同数据迁移要求量身定制的数据传输服务。根据组织内的业务需求，选择合适的数据存储和数据传输选项。

在规划或审查您的工作负载架构时，请考虑以下几点：

- 在 AWS 中使用 VPC 端点：VPC 端点允许您在 VPC 与支持的 AWS 服务之间建立私有连接。这让您可以避免使用公共互联网，而使用公共互联网会产生数据传输成本。
- 使用 NAT 网关：使用 NAT 网关，以便私有子网中的实例可以连接到互联网或 VPC 以外的服务。检查发送流量最多的 NAT 网关后面的资源是否与 NAT 网关位于同一可用区。如果未在同一可用区，则在与资源相同的可用区中创建新的 NAT 网关，从而降低跨可用区传输数据的费用。
- 使用 AWS Direct ConnectAWS Direct Connect 绕过公共互联网，并在您的本地网络和 AWS 之间建立直接的私有连接。与通过互联网传输大量数据相比，这可能更具成本效益且更加一致。
- 避免跨区域边界传输数据：AWS 区域 之间（从一个区域到另一个区域）的数据传输通常会产生费用。采用多区域方法应是一项深思熟虑的决定。有关更多详细信息，请参阅多区域场景。
- 监控数据传输：使用 Amazon CloudWatch 和 VPC 流日志，捕获有关数据传输和网络使用情况的详细信息。分析 VPC 中捕获的网络流量信息，例如进出网络接口的 IP 地址或范围。
- 分析网络使用情况：使用 AWS Cost Explorer、CUDOS 控制面板或 CloudWatch 等计量和报告工具来了解工作负载的数据传输成本。

实施步骤

- 选择用于数据传输的组件：使用 COST08-BP01 执行数据传输建模 中介绍的数据传输建模，关注产生最多数据传输成本之处，或者工作负载使用情况发生变化时产生最多数据传输成本之处。寻找替代架构或其他组件，以消除或减少数据传输的需要（或降低其成本）。

资源

相关最佳实践：

- COST08-BP01 执行数据传输建模
- COST08-BP03 实施服务以便降低数据传输成本

相关文档：

- 云数据迁移
- AWS 缓存解决方案
- 使用 Amazon CloudFront 更快地交付内容

相关示例：

- Overview of Data Transfer Costs for Common Architectures
- AWS Network Optimization Tips
- Optimize performance and reduce costs for network analytics with VPC Flow Logs in Apache Parquet format

COST08-BP03 实施服务以便降低数据传输成本

实施服务以减少数据传输。例如，使用边缘站点或内容分发网络（CDN）向终端用户交付内容，在应用程序服务器或数据库前构建缓存层，并使用专用网络连接而不是 VPN 连接到云端。

在未建立这种最佳实践的情况下暴露的风险等级：中

实施指导

可以利用多种 AWS 服务来优化网络数据传输使用情况。根据工作负载组件、类型和云架构，这些服务有助于您在云端压缩、缓存、共享和分发流量。

- Amazon CloudFront 是一个全球内容分发网络，可提供低延迟、高传输速度的数据。它在世界各地的边缘站点缓存数据，从而减少资源负担。通过使用 CloudFront，您可以减少向全球大量用户分发内容的管理工作，同时将延迟降到最低。如果您计划随着时间的推移增加使用量，安全防御节省套餐有助于您节省高达 30% 的 CloudFront 使用量。
- AWS Direct Connect 允许您建立到 AWS 的专用网络连接。与基于互联网的连接相比，这可以降低网络成本、增加带宽并提供更一致的网络体验。
- AWS VPN 允许您在专用网络和 AWS 全球网络之间建立安全的专用连接。此网络是小型办公室或业务合作伙伴的理想之选，因其提供简化的连接，并且是完全托管的弹性服务。
- VPC 端点允许通过专用网络在 AWS 服务之间建立连接，可用于减少公共数据传输和 NAT 网关成本。网关 VPC 端点不按小时收费，支持 Amazon S3 和 Amazon DynamoDB。接口 VPC 端点由 AWS PrivateLink 提供，有小时费和每 GB 使用成本。
- 与独立的 NAT 实例相比，NAT 网关提供内置的扩展和管理功能，可降低成本。将 NAT 网关放置在与高流量实例所在的同一可用区中，并考虑对需要访问 Amazon DynamoDB 或 Amazon S3 的实例使用 VPC 端点，以降低数据传输和处理成本。
- 使用具有计算资源的 AWS Snow Family 设备在边缘采集和处理数据。AWS Snow Family 设备（Snowball Edge、Snowball Edge 和 Snowmobile）可让您经济高效地将 PB 级数据离线迁移到 AWS Cloud。

实施步骤

- 实施服务：使用数据传输建模并查看 VPC 流日志，根据服务工作负载类型选择适用的 AWS 网络服务。了解产生最多成本和最多数据流之处。了解 AWS 服务并评测是否存在一种服务，可以减少或消除传输，特别是网络和内容分发。如有需要重复访问数据或存在大量数据的情况，则应了解缓存服务。

资源

相关文档：

- AWS Direct Connect
- AWS 探索我们的产品
- AWS 缓存解决方案
- Amazon CloudFront
- AWS Snow Family
- Amazon CloudFront 安全防御节省套餐

相关视频：

- Monitoring and Optimizing Your Data Transfer Costs
- AWS Cost Optimization Series: CloudFront
- How can I reduce data transfer charges for my NAT gateway?

相关示例：

- How-to chargeback shared services: An AWS Transit Gateway example
- Understand AWS data transfer details in depth from cost and usage report using Athena query and QuickSight
- Overview of Data Transfer Costs for Common Architectures
- Using AWS Cost Explorer to analyze data transfer costs
- Cost-Optimizing your AWS architectures by utilizing Amazon CloudFront features
- How can I reduce data transfer charges for my NAT gateway?

管理需求和供应资源

在您迁移到云时，您仅为所需内容付费。可以在需要时供应与工作负载需求匹配的资源，这样就无需进行昂贵且浪费的过度预置。还可以通过节流、缓冲区或队列来修改需求，满足需求并以更少的资源达成目标。

您应该在即时供应的经济优势和预置需求之间取得平衡，以协调资源故障、高可用性和预置时间。根据需求是固定的还是可变的，可以计划创建指标和自动化，这样即使在扩展时也能够确保最小化环境管理需要。修改需求时，必须知道工作负载可以允许的可接受最大延迟。

在 AWS 中，可以采用多种不同的方法来管理需求和供应资源。以下最佳实践介绍了如何使用这些方法。

最佳实践

- COST09-BP01 对工作负载需求执行分析
- COST09-BP02 实施缓冲区或节流来管理需求
- COST09-BP03 动态供应资源

COST09-BP01 对工作负载需求执行分析

分析工作负载需求随时间的变化。确认分析涵盖季节性趋势，并准确反映整个工作负载生命周期内的运行条件。分析工作应该体现可能带来的好处，例如花费的时间与工作负载成本成正比。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

分析云计算的工作负载需求涉及了解在云环境中启动的计算任务的模式和特征。这种分析有助于用户优化资源分配、管理成本并验证性能是否达到要求的水平。

了解工作负载的要求。组织要求应指出工作负载对于请求的响应时间。响应时间可用于确定需求是否得到了管理，或者是否应改变资源供应以满足需求。

分析应包括需求的可预测性和可重复性、需求的变化速率以及需求的变化量。对足够长的时间执行分析，以纳入任何季节性变化，例如月末处理或假期高峰。

分析工作应该体现实施扩缩可能带来的好处。查看组件的预期总成本，以及在工作负载生命周期内使用情况和成本的任何增加和减少。

以下是执行云计算的工作负载需求分析时，需要考虑的一些关键方面：

1. 资源利用率和性能指标：分析 AWS 资源在一段时间内的使用情况。确定高峰期和非高峰期使用模式，以优化资源分配和扩缩策略。监控性能指标，如响应时间、延迟、吞吐量和错误率。这些指标有助于评测云基础设施的整体运行状况和效率。
2. 用户和应用程序扩缩行为：了解用户行为及其对工作负载需求的影响。检查用户流量模式有助于增强内容的交付和应用程序的响应能力。分析工作负载如何随着需求的增长而扩展。确定自动扩缩参数的配置是否正确有效，以处理负载波动。
3. 工作负载类型：识别云中运行的不同类型的工作负载，如批处理、实时数据处理、Web 应用程序、数据库或机器学习。每种类型的工作负载都可能有不同的资源需求和性能特征。
4. 服务水平协议（SLA）：将实际绩效与 SLA 进行比较，确保合规性并确定需要改进的方面。

可以使用 Amazon CloudWatch 来收集和跟踪各项指标、监控日志文件、设置警报以及自动应对 AWS 资源的更改。还可以通过使用 Amazon CloudWatch 全面地了解资源利用率、应用程序性能和运行状况。

借助 AWS Trusted Advisor，您可以按照最佳实践预置资源，提高系统性能和可靠性，增强安全性，并寻找节省资金的机会。您也可以关闭非生产实例，并使用 Amazon CloudWatch 和 Auto Scaling 来匹配需求的增加或减少。

最后，可以将 AWS Cost Explorer或 QuickSight 与 AWS 成本和使用情况报告（CUR）文件或应用程序日志一起使用，来对工作负载需求执行高级分析。

总之，通过全面的工作负载需求分析，组织可以就资源预置、扩缩和优化作出明智的决策，从而提高性能、成本效益和用户满意度。

实施步骤

- 分析现有工作负载数据：分析现有工作负载中的数据、以前工作负载版本中的数据或预测使用模式中的数据。使用 Amazon CloudWatch、日志文件和监控数据来深入了解工作负载的使用情况。分析整个工作负载周期，并收集任何季节性变化数据，如月末或年末活动。分析中反映的工作应该体现出工作负载特征。应将最多的精力放在需求变化最大的高价值工作负载上。应将最少的精力放在需求变化最小的低价值工作负载上。
- 预测外部影响：与组织中会影响或更改工作负载需求的团队成员会面。通常涉及的团队包括销售、营销或业务拓展团队。与他们合作，了解其运作周期，以及是否有改变工作负载需求的任何活动。使用这些数据预测工作负载需求。

资源

相关文档：

- Amazon CloudWatch
- AWS Trusted Advisor
- AWS X-Ray
- AWS Auto Scaling
- AWS 实例调度器
- Getting started with Amazon SQS
- AWS Cost Explorer
- QuickSight

相关示例：

- 监控、跟踪和分析以实现成本优化
- Searching and analyzing logs in CloudWatch

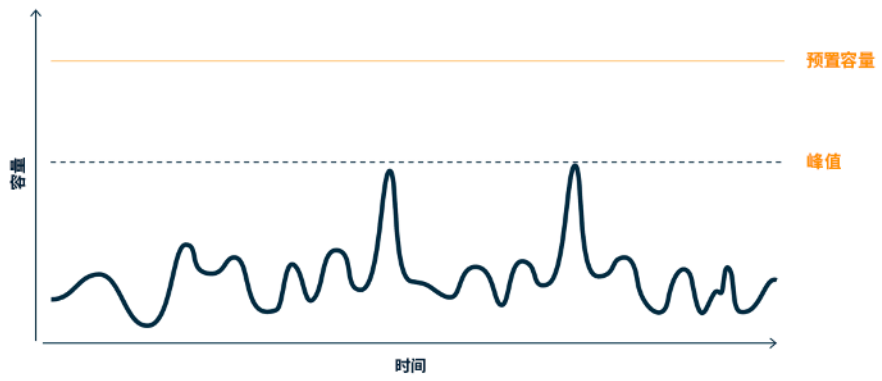
COST09-BP02 实施缓冲区或节流来管理需求

缓冲区和节流可修改工作负载需求，从而避免出现任何峰值情形。在客户端执行重试时实施节流。实施缓冲以存储请求并将处理任务往后推迟一段时间。确认设计节流和缓冲区时客户端能够在所需的时间内收到响应。

在未建立这种最佳实践的情况下暴露的风险等级：中

实施指导

在云计算领域，实施缓冲区或节流对于管理需求和减少工作负载所需的预置容量至关重要。为了获得最佳性能，必须衡量总需求，包括峰值、请求变化的速度和必需的响应时间。当客户端能够重新发送请求时，应用节流比较切实可行。相反，对于缺乏重试功能的客户端，理想的方法是实施缓冲区解决方案。此类缓冲区可舒缓请求的涌入，并优化具有不同操作速度的应用程序之间的交互。



需求曲线，有两个不同的峰值，需要高预置容量

假设工作负载的需求曲线如上图所示。此工作负载有两个峰值，为了处理这些峰值，如橙色线所示预置资源容量。因为需要预置容量来处理这两个峰值，所以此工作负载所使用的资源和能源不是由需求曲线下的区域表示，而是由预置容量线下面的区域表示。展平工作负载需求曲线有助于降低工作负载的预置容量和减少对环境的影响。为了平滑峰值，可以考虑实施节流或缓冲解决方案。

为了更好地理解它们，让我们探讨一下节流和缓冲。

节流：如果需求源具有重试功能，可以实施节流。节流会告诉需求源，如果当前无法处理请求，则应稍后再试。源将等待一段时间，然后重试请求。实施节流的优势是可限制最大资源量和工作负载成本。在 AWS 中，可以使用 Amazon API Gateway 实施节流。

基于缓冲：基于缓冲的方法使用产生器（向队列发送消息的组件）、使用器（从队列接收消息的组件）和队列（保存消息）来存储消息。然后消息将由使用器读取并处理，这样消息就能够以满足使用器业务要求的速率运行。通过使用以缓冲区为中心的方法，产生器发出的消息被存储在队列或流中，随时可供使用器以符合其运营需求的速度进行访问。

在 AWS 中，您可以从多项服务中进行选择，以便实施缓冲方法。Amazon Simple Queue Service（Amazon SQS）是一项托管服务，提供允许单个使用器读取单个消息的队列。Amazon Kinesis 提供允许众多使用器读取相同消息的流。

缓冲和节流可以通过修改工作负载需求来平滑任何峰值。在客户端重试操作时使用节流，并使用缓冲来保存请求以供以后处理。使用基于缓冲区的方法时，请构建工作负载以在所需时间内处理请求，并验证您是否能够处理重复的工作请求。分析总体需求、变化率和所需的响应时间，以使所需节流或缓冲的大小适宜。

实施步骤

- 分析客户端要求：分析客户端请求，确定它们是否能够执行重试。对于无法执行重试的客户端，需要实施缓冲区。分析总体需求、变化率和所需的响应时间，以确定所需的节流或缓冲区大小。
- 实施缓冲或限制：在工作负载中实施缓冲或限制。Amazon Simple Queue Service（Amazon SQS）之类的队列可以为工作负载组件提供缓冲区。Amazon API Gateway 可以为工作负载组件提供节流。

资源

相关最佳实践：

- SUS02-BP06 实施缓冲和节流以展平需求曲线
- REL05-BP02 限制请求

相关文档：

- AWS Auto Scaling
- AWS 实例调度器
- Amazon API Gateway
- Amazon Simple Queue Service
- Getting started with Amazon SQS
- Amazon Kinesis

相关视频：

- Choosing the Right Messaging Service for Your Distributed App

相关示例：

- Managing and monitoring API throttling in your workloads
- Throttling a tiered, multi-tenant REST API at scale using API Gateway
- Enabling Tiering and Throttling in a Multi-Tenant Amazon EKS SaaS Solution Using Amazon API Gateway
- Application integration Using Queues and Messages

COST09-BP03 动态供应资源

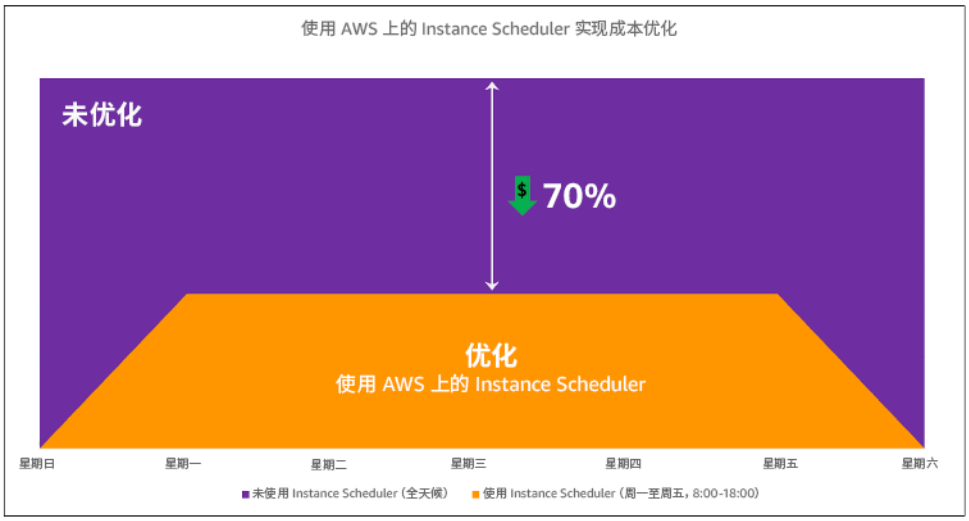
资源按计划预置。这种预置可以基于需求（例如通过自动扩缩来实现），也可以基于时间（需求可以预测，基于时间提供资源）。这些方法可以尽可能减少过度预置或预置不足的情况。

在未建立这种最佳实践的情况下暴露的风险等级：低

实施指导

AWS 客户可以通过多种方式增加可供应用程序使用的资源，并提供资源以满足需求。其中一个选项是使用 AWS 实例调度器，它可以自动启动和停止 Amazon Elastic Compute Cloud（Amazon EC2）及 Amazon Relational Database Service（Amazon RDS）实例。另一个选项是使用 AWS Auto Scaling，该服务让您可以根据应用程序或服务的需求，自动扩缩计算资源。根据需求提供资源，这样您就只需要为使用的资源付费，并且仅在有需要时启动资源，在不需要时终止资源，从而降低成本。

AWS 实例调度器让您可以将 Amazon EC2 和 Amazon RDS 实例配置为在指定的时间停止和启动，这样您就可以通过一致的时间模式满足对相同资源的需求，例如用户在每天早上八点访问 Amazon EC2 实例，晚上六点后就不再需要访问。该解决方案可停止不使用的资源，并在需要时启动它们，帮助降低运营成本。



使用 AWS 实例调度器优化成本。

您还可以通过简单的用户界面（UI）使用 AWS Systems Manager 快速设置功能，轻松地跨账户和区域来为 Amazon EC2 实例配置计划。您可以使用 AWS 实例调度器来计划 Amazon EC2 或 Amazon RDS 实例，也可以停止和启动现有实例。但是，您无法停止和启动属于自动扩缩组（ASG）或管理 Amazon Redshift 或 Amazon OpenSearch Service 等服务的实例。自动扩缩组对组中的实例和何时创建这些实例有自己的计划。

AWS Auto Scaling 有助于您调整容量以维持稳定、可预测的性能，并确保成本最低，以满足不断变化的需求。这是一项用来扩缩应用程序容量的完全托管式免费服务，与 Amazon EC2 实例和竞价型实例集、Amazon ECS、Amazon DynamoDB 与 Amazon Aurora 集成。自动扩缩提供自动资源发现功能，帮助在工作负载中找到可以配置的资源，它具有内置的扩缩策略来优化性能、成本或者在两者之间取得平衡，并提供预测性扩展来协助应对定期出现的峰值。

可以通过多种扩缩选项来扩缩自动扩缩组：

- 始终保持当前实例级别
- 手动缩放
- 按计划扩展
- 根据需求进行扩展
- 使用预测式扩展

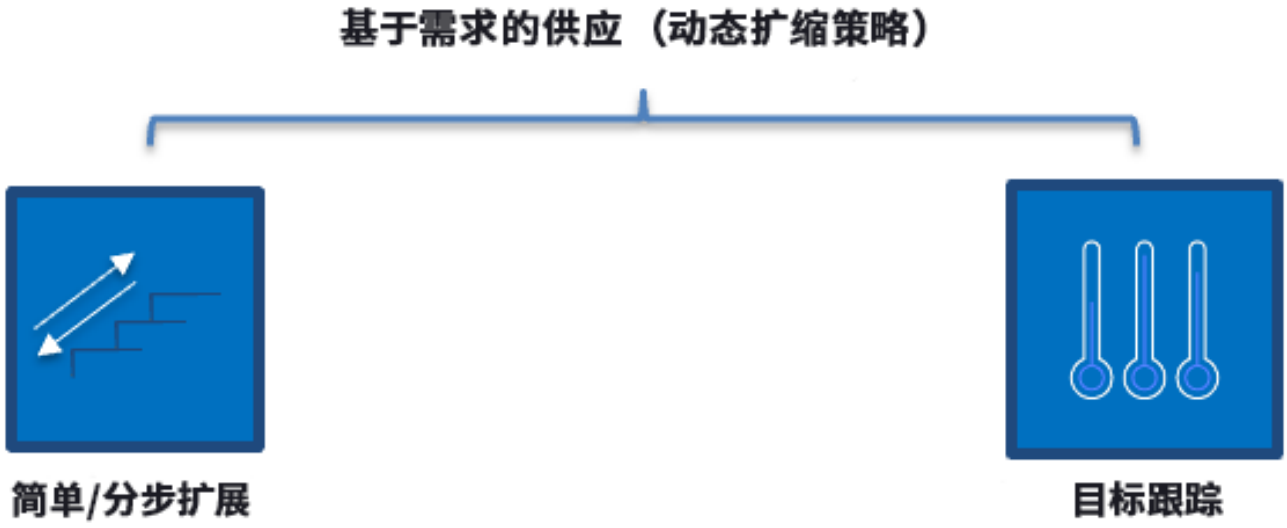
自动扩缩策略各不相同，可以分为动态扩缩策略和计划扩缩策略。动态策略为手动扩缩或动态扩缩，这可以是计划扩缩或者预测性扩缩。您可以针对动态、计划和预测性扩缩使用扩缩策略。还可以使用 Amazon CloudWatch 中的指标和警报来触发工作负载的扩缩事件。我们建议您使用启动模板，确保可以访问最新功能和改进。当您使用启动配置时，并非所有的自动扩缩功能都可用。例如，您无法创建 Auto Scaling 组来同时启动竞价型实例和按需型实例或者指定多个实例类型。您必须使用启动模板来配置这些功能。使用启动模板时，建议您对每个模板进行版本控制。利用启动模板的版本控制，您可以创建全套参数的子集。然后，您可以重复使用它来创建同一启动模板的其他版本。

可以使用 AWS Auto Scaling 或通过 AWS API 或 SDK 在代码中加入扩缩。这样省去了手动更改环境的操作成本，因而工作负载的总体成本得以降低，而且可以更快地执行更改。这还使您的工作负载资源配置随时与您的需求相匹配。为了遵循这一最佳实践，并为组织动态供应资源，您应该了解 AWS Cloud 中的水平和垂直扩缩，以及 Amazon EC2 实例上运行的应用程序的性质。您的云财务管理团队最好与技术团队合作，以遵循这一最佳实践。

弹性负载均衡（ELB）通过在多种资源之间分配需求来帮助您扩缩规模。通过使用 ASG 和弹性负载均衡，可以按照最优方式路由流量来管理传入的请求，这样就可以避免自动扩缩组中某个实例负载过高的情况。请求将以轮询方式，在目标组的所有目标上分配，而不考虑容量或利用率。

典型的指标可以是标准 Amazon 指标，例如 CPU 利用率、网络吞吐量以及弹性负载均衡观察到的请求和响应延迟。如果可能，应该使用指示客户体验的指标，通常是来自工作负载中的应用程序代码的自定义指标。在本文档中，为了详细说明如何动态满足需求，我们将自动扩缩划分为两类：基于需求的供应模型和基于时间的供应模型，并分别深入探讨这两种模型。

基于需求的供应：利用云的弹性来提供资源，根据近实时的需求状态来满足不断变化的需求。对于基于需求的供应，请使用 API 或服务功能，以编程方式改变架构中云资源的数量。这让您能够在架构中扩缩组件，并在需求高峰期间增加资源数量以保持性能，也可以在需求量降低时减少容量以降低成本。

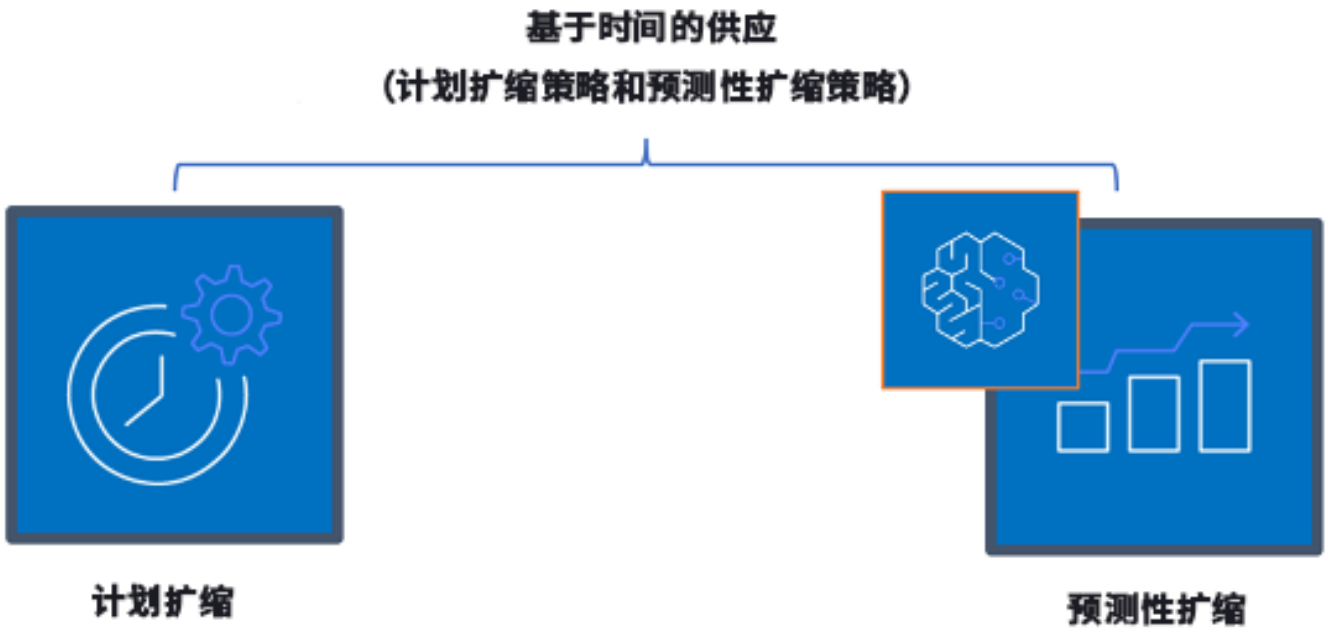


基于需求的动态扩缩策略

- 简单/步骤扩缩：监控指标，按照客户定义的步骤手动添加/删除实例。
- 目标跟踪：类似恒温器的控制机制，可自动添加或删除实例，以维护客户定义的目标指标。

当采用基于需求的方法进行构建时，请注意两个重要事项。首先，了解您必须以多快的速度预置新资源。其次，了解供应和需求之间的差额将发生变化。您必须准备好应对需求变化的速度，并准备好应对资源故障。

基于时间的供应：基于时间的方法可以协调资源容量，以满足可预测或按时间明确定义的需求。此方法通常不依赖资源的利用水平。基于时间的方法可以确保资源在需要的特定时间可用，并且提供时不会因启动程序和系统或一致性检查而发生延迟。使用基于时间的方法，您可以在繁忙时段提供额外的资源或增加容量。



基于时间的扩缩策略

您可以使用计划性或预测性自动扩缩，实施基于时间的方法。工作负载可以在规定的时间按计划横向扩展或缩减（例如办公时间开始时），从而确保用户就位或需求增加时资源可用。预测性扩缩使用相关模式进行横向扩展，而计划扩缩在预先规定的时间进行横向扩展。您还可以在自动扩缩组中使用基于属性的实例类型选择（ABS）策略，该策略允许您将实例要求表示为一组属性，例如 vCPU、内存和存储。这还允许您在新一代实例类型发布时自动使用它们，并通过 Amazon EC2 竞价型实例访问更广泛的容量。Amazon EC2 Fleet 和 Amazon EC2 Auto Scaling 会选择并启动符合指定属性的实例，无需手动选择实例类型。

可以利用 AWS API 和 SDK 以及 AWS CloudFormation，在需要时自动预置和停用整个环境。此方法非常适合仅在规定的办公时间或时间段运行的开发或测试环境。您可以使用 API 来扩缩环境中的资源大小（垂直扩缩）。例如，可以通过更改实例大小或类纵向扩展生产工作负载。这可以通过停止和启动实例，以及选择不同的实例大小或类来实现。这种技巧也可以应用于其他资源，如 Amazon EBS 弹性卷，您可以在使用时对其进行修改以增加大小、调整性能（IOPS）或更改卷类型。

当采用基于时间的方法进行构建时，请注意两个重要事项。首先，使用模式的一致性如何？ 其次，如果模式发生更改会产生什么影响？ 您可以通过两种方式提高预测的准确性：监控工作负载和使用商业智能。如果您发现使用模式发生重大更改，可以调整时间，以确保提供覆盖范围。

实施步骤

- 配置计划扩缩：对于可预测的需求变化，基于时间的扩缩可以及时提供正确的资源数量。如果资源创建和配置的速度不够快，无法响应需求变化，也可使用这种方法。根据工作负载分析，使用 AWS Auto Scaling 配置计划扩缩。要配置基于时间的计划，可以使用预测性扩缩而不是计划扩缩，根据预期或可预测的负载变化，提前增加自动扩缩组中的 Amazon EC2 实例数量。
- 配置预测性扩缩：使用预测性扩缩，可在流量流的每日和每周模式之前增加自动扩缩组中的 Amazon EC2 实例数量。如果您有定期的流量高峰和需要很长时间才能启动的应用程序，则应该考虑使用预测性扩缩。与本质上属于被动应对的单纯动态扩缩相比，预测性扩缩可帮助您在预计负载到来之前初始化容量，从而更快地扩缩。例如，如果用户在开始上班时开始使用工作负载，而在下班后不再使用，预测性扩缩可以在上班之前增加容量，这就消除了动态扩缩对流量变化作出反应的延迟。
- 配置动态自动扩缩：要根据活动工作负载指标配置扩缩，请使用自动扩缩。使用分析并配置自动扩缩以在正确的资源级别启动，并确认工作负载在所需的时间内横向缩减。您可以启动并自动扩展单个 Auto Scaling 组中的一组按需实例和竞价型实例。除了享受使用竞价型实例的折扣外，您还可以使用预留实例或 Savings Plan 获得常规按需实例定价的折扣费率。以上所有因素的综合作用是帮助您进一步节约 Amazon EC2 实例成本，同时帮助您获得应用程序所需的规模和性能。

资源

相关文档：

- AWS Auto Scaling
- AWS 实例调度器
- 扩展 Auto Scaling 组的大小
- Getting Started with Amazon EC2 Auto Scaling
- Getting started with Amazon SQS
- Scheduled Scaling for Amazon EC2 Auto Scaling
- Predictive scaling for Amazon EC2 Auto Scaling

相关视频：

- Target Tracking Scaling Policies for Auto Scaling
- AWS 实例调度器

相关示例：

- 基于属性选择实例类型用于 Amazon EC2 Fleet 的自动扩缩
- Optimizing Amazon Elastic Container Service for cost using scheduled scaling
- Predictive Scaling with Amazon EC2 Auto Scaling
- 如何将实例调度器和 AWS CloudFormation 配合使用来计划 Amazon EC2 实例？

持续优化

在 AWS 中，可以检查新服务并将其实施到工作负载中，以随着时间的推移持续优化。

AWS 会不断发布新服务和功能，建议您经常检查现有的架构决策，确保它们始终具有成本效益。随着需求的变化，一定要主动停用不再需要的资源、组件和工作负载。请考虑以下最佳实践来帮助您随着时间的推移持续优化。

在随着时间的推移持续优化工作负载并改善组织中的 CFM 文化的同时，评估云端运营的工作量成本，审查耗时的云运营，并通过采用相关 AWS 服务、第三方产品或自定义工具（如 AWS CLI 或 AWS SDK）实现自动化，以减少人力和成本。

主题

- 定义审查流程并定期分析工作负载
- 自动实施运营

定义审查流程并定期分析工作负载

最佳实践

- COST10-BP01 制定工作负载审核流程
- COST10-BP02 定期审核和分析此工作负载

COST10-BP01 制定工作负载审核流程

制定流程，定义工作负载的审核标准和流程。审核工作量应该体现可能带来的好处，例如，核心工作负载或费用占比超过 10% 的工作负载每季度或每六个月审核一次，而费用占比低于 10% 的工作负载每年审核一次。

在未建立这种最佳实践的情况下暴露的风险等级：高

实施指导

为了让工作负载始终最具成本效益，必须定期对工作负载进行审核，了解是否有机会实施新的服务、功能和组件。为了降低总体成本，审核工作量必须与潜在的节省额成比例。例如，与占总支出 5% 的工作负载相比，应更经常、更彻底地审核占总支出 50% 的工作负载。考虑任何外部因素或波动。如果工作负载服务于特定的地理位置或市场领域，并且您已预测出该领域会出现的变化，则提高审核频率可能会

节省成本。审核时要考虑的另一个因素是实施更改的工作量。如果测试和验证变更的成本很高，则审核的频率应该降低。

考虑维护过时和旧式组件及资源的长期成本，以及无法在其中实施新功能的事实。当前的测试和验证成本可能会超过预计的效益。但是，随着时间的推移，工作负载和当前技术之间的差距会增大，进行更改的成本可能会大幅增加，导致成本升高。例如，迁移到新的编程语言的成本当前可能不具成本效益。然而，五年之后，熟练使用该语言的人员的成本可能会增加，并且由于工作负载的扩展，迁移到新语言的系统规模更大，这其中涉及的工作量甚至高于以前。

将工作负载分解成多个组件，分配组件的成本（估算即可），然后在每个组件旁边列出因素（例如工作量和外部市场）。使用这些指示信息来确定每个工作负载的审核频率。例如，您可能觉得 Web 服务器的成本高、变更的工作量小、外部因素多，因而审核频率很高。而中央数据库的成本可能中等、变更的工作量很大、外部因素较少，因而审核频率也为中等。

制定相应流程，在新的服务、设计模式、资源类型和配置推出后，对它们进行评估，以优化工作负载成本。与绩效支柱审核和可靠性支柱审核流程类似，确定、验证、优化和改进活动并确定其优先顺序，然后提出补救措施并将其纳入积压工作中。

实施步骤

- 定义审核频率：定义工作负载及其组件的审核频率。分配时间和资源，以持续改进并保持审核频率，优化工作负载并提高工作负载的效率。应考虑多种因素，且这些因素可能会因组织中的工作负载以及工作负载中的组件而异。常见的因素包括：从收入或品牌角度来讲对组织的重要性、运行工作负载的总成本（包括运营和资源成本）、工作负载的复杂性、实施更改的难易程度、任何软件许可协议，以及更改是否会因惩罚性的许可而显著增加许可成本。可以在功能或技术上定义组件，例如 Web 服务器和数据库，或计算和存储资源。相应地权衡这些因素，并为工作负载及其组件制定一个周期。您可能决定每 18 个月审核一次完整的工作负载，每 6 个月审核一次 Web 服务器，每 12 个月审核一次数据库，每 6 个月审核一次计算资源和短期存储，每 12 个月审核一次长期存储。
- 定义审核的彻底性：定义在工作负载或工作负载组件的审核上投入的工作量。与审核频率类似，这也需要权衡多种因素。评估各种改进机会并确定其优先顺序，以便将精力集中在可以实现最大效益的工作上，同时估计这些活动所需的工作量。如果预期结果不符合目标，并且所需工作量会花费更多成本，则寻求其他行动方案。审核流程中应该分配专用的时间和资源，以便实现持续渐进式改进。例如，您可能决定投入一周的时间对数据库组件进行分析，投入一周的时间对计算资源进行分析，投入四小时的时间进行存储审核。

资源

相关文档：

- AWS 新闻博客
- 云计算类型
- AWS 的新功能

相关示例：

- AWS Support Proactive Services
- 定期审核 SAP 工作负载的工作负载

COST10-BP02 定期审核和分析此工作负载

根据每个明确的流程定期审核现有工作负载，以便确定是采用新服务、替换现有服务还是重新构建工作负载。

在未建立这种最佳实践的情况下暴露的风险等级：中

实施指导

AWS 在不断地增加新功能，以便您可以使用最新技术更快地进行实验和创新。AWS新增功能详细介绍了 AWS 如何做到这一点，并简要概述了 AWS 服务、功能和区域扩展公告的发布时间表。您可以深入了解已经宣布的产品发布，并使用它们来审核和分析现有的工作负载。为了享受新 AWS 服务和功能带来的优势，请对工作负载进行审核，并根据需要实施新服务和功能。这意味着您可能需要替换用于工作负载的现有服务，或对工作负载进行现代化改造，以便采用这些新的 AWS 服务。例如，可以审核工作负载，并使用 Amazon Simple Email Service 替换消息传递组件。这省去了运行和维护实例集的成本，同时能以更低的成本提供所有功能。

为了分析您的工作负载并突显潜在机会，您不仅应考虑新服务，而且还要考虑构建解决方案的新方法。观看 AWS 上的这是我的架构视频，了解其他客户的架构设计、面临的挑战和解决方案。查看 All-In series，了解 AWS 服务的真实应用和客户案例。还可以观看回归基础视频系列，该系列讲解、研究和分解基本的云架构模式最佳实践。另一个来源是如何构建此架构视频，这些视频旨在帮助有宏大想法的人们了解如何使用 AWS 服务将最简可行产品（MVP）变为现实。来自世界各地有着坚定想法的构建者可以利用这种方式从经验丰富的 AWS 解决方案架构师那里获得架构指导。最后，可以查看入门资源材料，其中包含分步教程。

在开始审核流程之前，请遵照企业对工作负载的要求、安全性和数据隐私要求，以便在遵循商定的审核流程时，运用特定的服务或区域和性能要求。

实施步骤

- 定期审核工作负载：使用定义的流程，按照指定的频率执行审核。确认在每个组件上投入适当的工作量。此流程类似于您选择服务进行成本优化的初始设计流程。分析服务及其带来的优势，这一次需考虑进行更改所产生的成本，而不仅仅是长期优势。
- 实施新服务：如果分析结果表明可以实施更改，请先执行工作负载基线，以了解每项产出的当前成本。实施更改，然后执行分析以确认每项产出的新成本。

资源

相关文档：

- AWS 新闻博客
- AWS 的新功能
- AWS 文档
- AWS 入门
- AWS 一般资源

相关视频：

- AWS – 这是我的架构
- AWS – 回归基础
- AWS - All-In series
- 如何构建此架构

自动实施运营

最佳实践

- COST11-BP01 执行运营自动化

COST11-BP01 执行运营自动化

评估云端的运营成本，专注于量化在管理任务、部署、人为错误风险缓解、合规性以及通过自动化实现的其他操作方面所节省的时间和工作量。评测运营工作所需的时间和相关成本，实现管理任务的自动化，从而最大限度地减少人工操作。

在未建立这种最佳实践的情况下暴露的风险等级：低

实施指导

通过实现操作的自动化，还可以在部署、管理或运行工作负载时提供一致且可靠的体验，从而降低手动任务的频率、提高效率以及使客户受益。您可以将基础设施资源从手动操作任务中解放出来，让他们参与更高价值的任务和创新，从而提高业务价值。企业需要行之有效、经过测试的方法来管理云中的工作负载。该解决方案必须安全、快速、经济高效，风险最低且可靠性最高。

首先着眼于总体运营成本，根据所需的工作量确定运营活动的优先级。例如，在云中部署新资源、对现有资源进行优化更改或实施必要的配置需要多长时间？ 考虑运营和管理成本，看看人工行为的总成本。优先考虑管理任务的自动化，以减少人工工作量。

审核工作应该体现可能带来的好处。例如，检查与自动执行任务相比，手动执行任务所花费的时间。优先考虑自动执行重复、高价值、耗时和复杂的活动。从那些高价值或人为错误风险高的活动开始实现自动化通常会更好，因为风险通常会带来不必要的额外运营成本（例如，运营团队加班产生的成本）。

使用 AWS Systems Manager 或 AWS Config 等自动化工具来简化运营、合规性、监控、生命周期和终止流程。借助 AWS 服务、工具和第三方产品，您可以自定义所实施的自动化来满足特定要求。下表显示了您为了自动执行管理和运营，可以通过 AWS 服务实现的一些核心运营职能和能力：

- AWS Audit Manager：持续审核 AWS 使用情况，以简化风险和合规性评测
- AWS Backup：集中管理和自动化数据保护。
- AWS Config：配置计算资源，评估、审核配置和资源清单。
- AWS CloudFormation：使用“基础设施即代码”启动高可用性资源。
- AWS CloudTrail：IT 变更管理、合规性和控制。
- Amazon EventBridge 调度事件并触发 AWS Lambda 采取行动。
- AWS Lambda：通过使用事件触发重复流程，或者使用 AWS EventBridge 按固定计划运行重复流程，实现这些流程的自动化。
- AWS Systems Manager：启动和停止工作负载、修补操作系统、自动配置和持续管理。
- AWS Step Functions：安排作业并实现工作流程自动化。
- AWS Service Catalog：模板消费，基础设施即代码，具有合规性和控制力。

如果您想在使用 AWS 产品和服务时立即采用自动化，并且组织中没有相关技能，请联系 AWS Managed Services（AMS）、AWS 专业服务团队或 AWS 合作伙伴，以提高自动化的采用率并改善您在云端的卓越运营。

AWS Managed Services（AMS）是代表企业客户和合作伙伴运营 AWS 基础设施的服务。该服务提供了一个安全且合规的环境，您可以将工作负载部署到其中。AMS 使用具有自动化功能的企业云运营模型，让您可以满足组织要求，更快地迁移到云中并降低持续的管理成本。

AWS 专业服务团队还可以帮助您实现期望的业务成果并通过 AWS 实现运营自动化。团队成员帮助客户部署自动、稳健、敏捷的 IT 运营，还提供针对云进行优化的治理能力。有关详细的监控示例和推荐的最佳实践，请参阅《卓越运营支柱白皮书》。

实施步骤

- 一次构建，多次部署：使用基础设施即代码（例如 CloudFormation、AWS SDK 或 AWS CLI）一次部署并多次用于类似环境或灾难恢复场景。在部署时进行标记，以便按照其他最佳实践中的规定跟踪使用情况。使用 AWS Launch Wizard 缩短部署许多常用企业工作负载的时间。AWS Launch Wizard 指导您按照 AWS 最佳实践完成企业工作负载的大小调整、配置和部署。还可以使用 Service Catalog，它可以帮助您创建和管理经基础设施即代码批准的模板以在 AWS 上使用，以便任何人都可以发现经批准的自助式云资源。
- 实现持续合规自动化：考虑根据预定义的标准自动评测和补救记录的配置。将 AWS Organizations 与 AWS Config 和 AWS CloudFormation 的功能相结合后，可以高效地大规模管理和自动化数百个成员账户的配置合规性。可以审核配置更改以及 AWS 资源之间的关系，并深入探究资源配置的历史记录。
- 自动执行监控任务 AWS 为您提供了各种工具用来监控服务。您可以配置这些工具来自动执行监控任务。创建并实施监控计划，以收集工作负载中所有部分的监控数据，在发生多点故障时，就能更轻松地进行调试。例如，当系统状态检查、实例状态检查和 Amazon CloudWatch 警报出现问题时，可以使用自动监控工具来观察 Amazon EC2 并向您发送报告。
- 自动执行维护和操作：自动运行例行操作，无需人工干预。使用 AWS 服务和工具，您可以选择实施哪些 AWS 自动化，并根据您的特定要求进行自定义。例如，使用 EC2 Image Builder 构建、测试和部署虚拟机和容器镜像，以在 AWS 或本地使用，或者使用 AWS SSM 修补 EC2 实例。如果无法使用 AWS 服务完成所需操作，或者需要针对筛选资源执行更复杂的操作，请使用 AWS Command Line Interface（AWS CLI）或 AWS SDK 工具自动执行操作。AWS CLI 提供了使用脚本即可自动执行控制和管理 AWS 服务的整个流程的能力，无需使用 AWS Management Console。选择首选的 AWS SDK 与 AWS 服务交互。有关其他代码示例，请参阅 AWS SDK 代码示例存储库。
- 通过自动化创建持续的生命周期：制定和保留成熟的生命周期策略非常重要，这不仅是为了法规或冗余，还是为了成本优化。可以使用 AWS Backup 集中管理和自动保护数据存储（例如，存储桶、卷、数据库和文件系统）中的数据。还可以使用 Amazon Data Lifecycle Manager 来自动创建、保留和删除 EBS 快照和 EBS 支持的 AMI。
- 删除不必要的资源：在沙盒或开发 AWS 账户 中积累未使用的资源是很常见的情况。在一般开发周期内，开发人员将创建和试验各种服务与资源，之后，当他们不再需要这些资源时，不会将其删

除。未使用的资源会给组织带来不必要的成本，这些成本有时甚至非常高。删除这些资源可降低这些环境的运营成本。确认您不需要这些数据，如果无法确定，请确认数据已备份。可以使用 AWS CloudFormation 清理已部署的堆栈，这将自动删除模板中定义的大多数资源。或者，可以使用诸如 aws-nuke 之类的工具实现 AWS 资源删除操作的自动化。

资源

相关文档：

- Modernizing operations in the AWS Cloud
- AWS Services for Automation
- 基础设施和自动化
- AWS Systems Manager Automation
- 自动和手动监控
- AWS automations for SAP administration and operations
- AWS Managed Services
- AWS Professional Services

相关视频：

- Automate Continuous Compliance at Scale in AWS
- AWS Backup Demo: Cross-Account & Cross-Region Backup
- Patching for your Amazon EC2 Instances

相关示例：

- Reinventing automated operations (Part I)
- Reinventing automated operations (Part II)
- Automate deletion of AWS resources by using aws-nuke
- Delete unused Amazon EBS volumes by using AWS Config and AWS SSM
- Automate continuous compliance at scale in AWS
- IT Automations with AWS Lambda

结论

成本优化和云财务管理是一项持续的工作。您应定期与财务和技术团队合作，审核架构方法，更新组件选项。

AWS 致力于在您构建韧性高、响应迅速的自适应部署时，帮助您最大限度地降低成本。要真正优化部署成本，请充分利用本白皮书中讨论的相关工具、技巧和最佳实践。

贡献者

本文档的贡献者包括：

- Fatih (Ben) Mergen，Amazon Web Services 的 Well-Architected 部门的成本优化支柱负责人
- Keith Jarrett，亚马逊云科技的成本优化部门业务拓展主管
- Arthur Basbaum，亚马逊云科技业务开发经理
- Jarman Hauser，亚马逊云科技商业架构师

延伸阅读

有关更多信息，请参阅：

- AWS Well-Architected Framework
- AWS 架构中心

文档修订

如需获取有关该白皮书更新的通知，请订阅 RSS 信息源。

变更	说明	日期
更新了最佳实践指南	多项最佳实践更新。新增最佳实践 COST06-BP04。	2024 年 6 月 27 日
更新了最佳实践指南	通篇对最佳实践进行了小幅更新。	2023 年 12 月 6 日
更新了最佳实践指南	根据支柱中的新指南更新了最佳实践。	2023 年 10 月 3 日
更新了最佳实践指南	根据新指南更新了最佳实践的以下方面：治理、监控成本和使用情况、选择最佳定价模式以及管理需求和供应资源。	2023 年 7 月 13 日
针对新框架进行了更新	为最佳实践更新了规范性指南并增加了新的最佳实践。增加了问题 COST 11，包括新的最佳实践 COST11-BP01。	2023 年 4 月 10 日
次要更新	定价模式部分恢复了缺失的指南。	2023 年 1 月 13 日
已更新白皮书	为最佳实践更新了新的实施指导。	2022 年 12 月 15 日
已更新白皮书	扩展了最佳实践并增加了改进计划。	2022 年 10 月 20 日
次要更新	在简介中添加了可持续性支柱。	2021 年 12 月 2 日
次要更新	更新了链接。	2023 年 4 月 25 日

次要更新	更新了链接。	2021 年 3 月 10 日
针对新框架进行了更新	已更新以加入 CFM、新服务，同时与 Well-Architected 集成。	2020 年 7 月 8 日
已更新白皮书	已更新以反映对 AWS 的更改，并加入与客户审核的过程中了解到的信息。	2018 年 7 月 1 日
已更新白皮书	已更新以反映对 AWS 的更改，并加入与客户审核的过程中了解到的信息。	2017 年 11 月 1 日
初次发布	发布了成本优化支柱 – AWS Well-Architected Framework。	2016 年 11 月 1 日

版权声明

客户有责任对本文档中的信息进行单独评测。本文档：(a) 仅供参考，(b) 代表当前的 AWS 产品和实践，如有更改，恕不另行通知，以及 (c) 不构成 AWS 及其附属公司、供应商或许可方的任何承诺或保证。AWS 产品或服务“按原样”提供，不附带任何明示或暗示的保证、陈述或条件。AWS 对其客户承担的责任和义务受 AWS 协议制约，本文档不是 AWS 与客户直接协议的一部分，也不构成对该协议的修改。

© 2023，Amazon Web Services, Inc. 或其附属公司。保留所有权利。

AWS 术语表

有关最新的 AWS 术语，请参阅 AWS 词汇表 参考中的 AWS 词汇表。