制定多云战略的久经考验的实践

AWS 规范性指导

AWS 规范性指导: 制定多云战略的久经考验的实践

Table of Contents

制定多云战略的久经考验的实践

Amazon Web Services 的 Tom Godden 和 Ellie Tamari

2025 年 9 月（文档历史记录）

如今，关于采用多云的组织，面临着相互矛盾的信息。有些人建议完全不要这样做，而另一些人则声称每个人都在切换到多云环境。现实介于这些极端之间：支持和反对多云战略的正当理由都存在，成功取决于在潜在的业务价值与固有的复杂性和风险之间取得平衡。

在 AWS，我们对互操作性的承诺是许多客户选择我们平台的关键原因。我们坚信，无论您的工作负载在哪里，您都可以自由地进行创新，并使您能够选择最适合自己需求的技术。在 AWS，我们一直处于开发解决方案的最前沿，这些解决方案使您能够在任何环境中构建和部署应用程序。这种以客户为中心的方法是受到全球数百万客户的信赖的基础。 AWS Cloud

我们知道，客户需要能够与现有工具和未来技术选择无缝协作的云平台。添加来自其他提供商的功能时，您不必重新构建所有内容。您的云应该可以帮助您跨环境连接、保护和管理工作负载，而不必强迫您成为每个平台的专家。 AWS 无论您的策略是 AWS 专门使用还是遵循选择性的多云方法，都可以直接在其服务中建立连接点，以帮助您高效运营。

我们认识到，每个组织都有独特的业务需求来推动其云战略决策。无论您是主要在上运行工作负载 AWS、跨多个云运行工作负载，还是将其 AWS 用作更广泛的多云架构的一部分，我们都致力于帮助您取得成功。 AWS 提供深度和广度的工具和功能，无论您的工作负载位于何处，都能帮助您更轻松、更快速地构建、迁移和操作。 AWS 工具可简化提供商之间的管理，同时最大限度地提高云投资的性能和价值。

这篇论文重点介绍成功实施多云战略的久经考验的原则，包括多云方法何时何地有意义，以及如何 AWS 帮助企业成功实施多云战略。它提供了规范性指导，帮助高管做出与采用多云相关的明智战略和决策选择。本 paper 没有对多云实施进行技术性的深入讨论。要获得技术实施支持和帮助以应对您的特定挑战，我们建议您与您的 AWS 解决方案架构师合作。

本 paper 根据我们与 AWS 企业客户打交道的经验，提出了九项久经考验的多云成功原则。每项原则都涉及多云战略的一个关键方面，从调整业务目标到安全实施。通过应用这些原则，组织可以自信地应对多云的复杂性。

原则 1. 使多云目标与您的战略保持一致

Gartner的研究和行业趋势表明，越来越多的组织采用多云方法来满足特定的业务需求。以下场景演示了多云基础架构何时具有战略优势。

合并和收购

并购 (M&A) 可以立即制定有关云战略的决策。尽管运营多个云可能会增加成本和复杂性，但快速整合可能会延迟集成价值并中断业务运营。您的云决策成为实现并购收益的核心。

整合规划应考虑整个技术格局。每项工作负载都需要在您的集成时间表和业务优先级的背景下进行评估。

我们的指导：

- 制定以业务为导向的整合战略，在即时集成需求和长期运营效率之间取得平衡。在仓促整合可能会中断关键业务运营或延迟并购价值实现的情况下，首先要维护多个云。
- 创建与您的集成时间表一致的明确工作负载放置标准。优先考虑创收应用程序和核心业务流程，同时考虑技术依赖关系和运营需求。

希望利用另一家 CSP 的长期差异化能力

对错过机会的恐惧促使一些公司想要每种云都占有一席之地。工作负载配置决策会影响整个组织——从工程团队到财务再到安全运营。

因此，Organizations需要研究他们追求多云的理由。有人认为，每个工作负载都应由最能满足其需求的云服务提供商 (CSP) 负载。但是，个人工作量优化必须与更广泛的组织影响相平衡。每增加一个云提供商，就有可能增加运营复杂性，产生新的人才需求，并引入影响整个技术组织的安全注意事项。

我们的指导：

- 遵循 80/20 方法：为大多数工作负载选择主要提供商，仅针对特定的高价值用例考虑其他提供商。该策略最大限度地提高了效率和人才留存率，同时降低了复杂性。
- 考虑跨云运营的总成本。在分析中包括安全工具、治理产品、财务管理系统和运营开销。
- 评估每个工作负载的依赖关系和相互作用。工作负载很少孤立运行；它们共享数据、安全控制和操作流程。
- 对供应商进行全面的性价比分析。不仅要比较直接成本，还要比较管理多个环境的开销。

控股公司采用多云技术，运营公司或业务线为主云

私募股权公司和控股公司面临着独特的云战略考虑。他们的投资组合公司通常保持独立的云战略，这通常源于过去的并购活动。这种结构降低了通常与多云运营相关的复杂性，因为每个业务部门都独立运营。但是，这种独立性可能会限制利用企业范围内的批量折扣和购买激励措施的机会。

云战略在控股公司层面的有效性取决于投资组合公司的自主权及其各自的技术需求。尽管整合可能会产生购买杠杆，但它可能与控股公司和私募股权投资组合中典型的独立运营模式相冲突。

我们的指导：

- 了解 CSP 批量 discount 结构。每个提供商都提供了在企业协议中添加或删除子公司以及将业务部门分拆为单独实体的机制。这些代表了双向决策。
- 仔细计划云购买承诺。尽早与您的云解决方案提供商的客户团队联系，或者联系 AWS Partner 具有AWS 云运营能力的人寻求帮助。
- 在独立性与效率之间取得平衡。考虑共享服务或购买协议，使投资组合公司受益，同时又不限制其运营。
- 首先关注业务目标。制定支持您的运营模式的技术战略，而不是仅仅为了多云战略而推行多云战略。
- 从投资组合管理的角度评估云策略。考虑一下云选择如何影响潜在的资产剥离或未来的收购。

宗旨 2. 注意多云的误解

在制定多云策略时，请避免以下各节中讨论的常见误解。

每个人都在采用多云策略

咨询公司和媒体公司描绘了采用多云的复杂画面。研究表明，人们对多云方法普遍感兴趣，但支出模式往往会讲述不同的故事。实际上，许多企业要么维护单一云环境，要么维护明确的 primary/secondary CSP 关系。这种脱节凸显了将目光投向头条新闻之外，转而关注组织的特定需求的重要性。

我们的指导方针：

- 根据您的特定业务需求做出云决策，而不是关注行业趋势。重点关注组织可衡量的成本和风险。
- 在您的行业背景下研究多云用例。适用于消费科技公司的云策略可能无法转化为金融服务、制造业或游戏环境。
- 将数据引力视为工作负载放置决策的主要因素。数据的位置和移动通常决定了最有效的云架构。
- 除了采用率统计数据之外，还可以了解支出模式。报告的高多云采用率通常掩盖了实际的支出模式。
- 在承诺使用多云环境之前，请评估技术限制。当某些工作负载的组件保留在单个云环境中时，其性能最佳。

多云降低了供应商锁定的风险

在制定云战略时，供应商灵活性是一个合理的考虑因素。Organizations 重视能够随着业务需求的变化调整其技术选择。这种担忧反映了以前在传统 IT 投资方面的经验，这些投资产生了具有约束力的长期承诺。 云服务围绕提供商的灵活性提供了不同的动态。 AWS 提供兼容开源的服务和数据可移植性选项，可减少迁移的技术障碍。但是，灵活性和运营效率之间的权衡仍然很重要。组织必须权衡维护提供商选项的商业价值和与主要提供商提供的专业服务深度集成的技术优势。

一些客户试图通过设计使用容器的与云无关的解决方案来避免锁定。这种方法通常将他们限制在基本的计算和存储服务上，而忽略了高级云功能的优势。我们的经验表明，与使用原生服务相比，由于需要更多的开发时间和资源，这种策略增加了相当大的复杂性。

我们的指导方针：

- 考虑与云无关的架构的全部成本。额外的工程开销可能不足以证明便携性优势是合理的。

- 使用云原生功能实现最大价值。仅靠基本的计算和存储服务往往会牺牲在安全性、可扩展性和创新方面的显著优势。
- 根据业务需求规划云战略。当多云实施增加明显的价值（例如能够在多个平台上为用户提供服务）时，额外的工程投资就会变得物有所值。
- 评估现实的退出情景和成本。将更换提供商的可能性和费用与使用整套提供商的好处进行比较 AWS 服务。
- 在的开源基础上构建 AWS。 AWS 诸如 Amazon Relational Database Service (Amazon RDS) 之类的托管服务可为您提供灵活性和卓越的运营，并支持您当今使用的数据库引擎。
- 利用提供的全面迁移工具 AWS。我们可以帮助您将工作负载向任何方向移动，并在您离开时使用其他提供商 AWS 时提供免费的数据流量。如需了解更多信息，请参阅 AWS 博客文章《迁出时免费将数据传输到互联网》 AWS。

多云可提高可用性和弹性

在中断期间，人们相信云提供商之间可以无缝切换工作负载，这促使一些组织转向多云战略。这种思维方式使人们对云基础设施弹性的看法过于简单，忽略了基本的技术现实。

根据多年来与多云客户合作的经验 AWS，我们发现，在提供商之间保持完全的工作负载可移植性通常会带来极大的复杂性，而无法带来所有预期的收益。由于数据引力的限制，数据密集型应用程序面临着难以克服的挑战。实际上，在我们看来，组织几乎不可能成功地为数据密集型工作负载实施真正无缝的多云故障转移。

Gartner杰出副总裁分析师Lydia Leong在社交媒体帖子中强化了这一观点：“多云故障转移既复杂又昂贵，几乎总是不切实际的，也不是解决云弹性风险的特别有效的方法。” 提供商在网络、存储、数据库、机器学习和安全性方面的固有差异使得真正的可移植性几乎是不可能的。在提供商之间分散工作负载可能会增加风险，因为任何一个环境中的故障都可能触发所有环境的中断。

我们的指导方针：

- 专注于掌握单个工作负载的 AWS 功能，而不是追求复杂的多云架构。
- 通过 AWS 区域 和可用区建立弹性，而不是尝试跨提供商故障转移。要深入了解 AWS 如何在物理数据中心之间自动进行工作负载故障转移，请参阅 AWS 博客文章 Zonal autoshift — 当我们检测到潜在问题时，自动将您的流量从可用区域转移出去。
- 策略性地将工作负载迁移到一个应用程序 AWS，一次只能专注于一个应用程序，从而最大限度地提高成功率。

多云提供更优惠的定价

对于多云环境来说，价格竞争力可能是最弱的论据。各组织在复杂、昂贵的软件或数据中心合同上签订多年协议的经历使他们在采购IT服务时谨慎行事。传统的采购方法尚未适应采 pay-as-you-go购、批量折扣或云端价格竞争的现实。（截至2025年1月，自成立以来 AWS 已降价151次。）

成本降低的最大驱动因素是管理良好、经过优化的云环境。一家公司通过主要与服务具有性价比优势（例如基于 AWS Gravit on等定制设计芯片的计算实例）并拥有卓越的云财务管理解决方案的提供商合作，可以实现更好的成本优化。根据 2022年Hackett集团对1,000多家组织进行的一项研究，与多云组织相比， AWS 客户的基础设施支出占IT总支出的百分比降低了20%。

我们的经验表明，各公司不会预料到在多云环境中运营会增加成本和复杂性，也不会适当地权衡这一成本与 head-to-head采购参与的预期收益。

我们的指导方针：

- 在 Well-Architecte AWS d Framework 成本优化支柱上制定成本优化策略。有五个设计原则：
  - 实施云财务管理：要在云端取得财务成功并加速实现业务价值，您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。与您的安全或运营能力一样，您需要通过知识建设、计划、资源和流程来提高能力，以帮助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在工作周内每天仅使用八个小时。您可以在不使用这些资源时将其停用，这样可以节省 75% 的成本（40 小时比 168 小时）。
  - 衡量整体效率：衡量工作负载的业务产出以及与交付相关的成本。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。
  - 停止将钱花在无差别的繁重工作上： CSPs 完成数据中心运营的繁重工作，例如机架、堆叠和为服务器供电。它们还通过使用托管服务消除了管理操作系统和应用程序的运营负担。这使您可以专注于客户和业务项目，而不是 IT 基础架构。
  - 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。
- 考虑到在不同提供商之间运营的财务开销，我们引导客户在自动化和成本优化工具上进行大量投资。每个 CSP 都提供了该领域的大量原生工具，例如。AWS 成本优化中心大多数原生工具都为云环境中的客户提供了出色的功能。但是，要了解多项支出 CSPs，您可以从一组丰富的 ISV 和软件即服务 (SaaS) 产品中进行选择，这些产品扩展了这些功能，以提供单一的成本优化体验。

- 通过支出权益策略稀释购买力并不能产生商业价值。它可能会破坏潜在的批量折扣，并可能破坏技术设计。使用云服务的最有效方法是在您的大部分运营中使用主要提供商，而 CSPs 仅在可以增加业务价值的地方使用其他提供商。

宗旨 3. 制定明确的战略和治理来支持它

决定推行多云战略是不够的；您必须制定实现目标的策略，包括明确管理哪些工作负载将流向何处以及为什么。应使用评估标准来优化工作负载及其依赖关系。如果 CSPs 将评估留给个人，那么不协调的蔓延可能会侵蚀多云战略的价值。我们建议您定期评估 CSP 工作负载性能，并将评估作为对 CSP 选择、标准和 future 使用情况的关键输入。

有效的治理策略需要了解整个企业中使用的服务、应用程序和组件的总数。其中不可或缺的是一种强大的标记策略，该策略涵盖 CSPs 并明确了所有已部署资源的所有权、使用情况和环境（例如开发、QA、暂存和生产）。所有内容都应标记为所有者；如果未标记或无法识别所有者，则应将其删除。我们与一家大型金融服务组织密切合作，该组织会自动查找和删除任何未标记的资源，无论它给开发团队带来什么不便，我们都认为这是一种最佳实践。这种标记方法编纂了治理规则并实现了执法自动化，而不是阻碍进展（也就是说，它实施的是护栏，而不是大门）。成本、运营和安全必须以同样的方式进行跟踪、监控和采取行动，同时保持同样的数据深度和透明度 CSPs。

在实施多云战略时，跨云提供商建立清晰一致的客户结构对于维护运营控制和安全至关重要。我们建议采用一种 hub-and-spoke模型，即 AWS 账户 为不同的业务部门单独创建。这些账户由两个关键的中央账户支撑：一个用于整合合规和安全监控的 security/audit 账户，以及一个用于管理互联的中央网络账户。（这种方法已编入的AWS Control Tower设计中。 但是，最低特权和职责分离的原则同样适用于其他云层。 Well-Ar chitected Framework 详细讨论了这些概念，强烈建议技术AWS 受众使用。）这种基础方法应在云提供商之间进行反映，以保持治理和运营的一致性。工作负载帐户应按环境（开发、暂存、生产）或职能进行组织，并为帐户的创建和删除制定明确的流程。

我们的指导方针：

- 实施全面的标签策略，在所有云资源中保持明确的所有权和使用模式。通过一致的标签策略跟踪环境、成本中心、应用程序和业务部门。移除缺少适当标签的资源，以强制执行治理标准并保持环境清晰度。
- 建立统一的合规框架，在您的多云环境中映射监管要求。保留清晰的文档，说明每个云提供商的控制措施和认证如何支持您的合规义务。
- 通过自动化而不是使用手动批准流程自动执行治理。将您的治理规则编码到自动化系统中，以便在违反政策的行为发生之前加以防范。这样可以消除人为错误，同时保持开发速度。
- 在具有集中安全和网络控制的 hub-and-spoke模型中构建帐户。创建用于安全审计和网络管理的专用帐户，以集中管理关键功能。此基础可在整个组织中实现一致的安全策略和网络连接。
- 要保持运营边界，请为不同的环境和功能创建单独的账户、订阅或项目（取决于您的 CSP 的命名法）。按开发、暂存和生产环境划分工作负载。这种分离可以防止安全事件的蔓延，并保持清晰的操作域。

- 通过整个环境中一致的指标来监控成本、运营和安全性。对资源利用率、安全事件和支出模式实施统一监控。使用这些数据来优化工作负载放置和资源分配决策。
- 通过组织策略和自动控制来防止未经授权的云使用。为账户创建和资源配置定义明确的流程。实施服务控制策略 (SCPs)，强制所有账户都遵守组织标准。
- 建立侦查和预防性控制措施，防止影子 IT 通过未经授权的提供商帐户出现。通过费用报告和网络流量监控未经授权的云使用情况。阻止未经授权的提供商访问，同时保持经批准的创新路径。

宗旨 4. 不要将连续的工作负载分散到云中

将连续的工作负载分散到多个云提供商会带来不必要的复杂性、风险和成本。当同时处理和分析数据的工作负载跨越多个提供商时，组织将面临数据移动、同步和一致性方面的挑战。团队必须为每个提供商浏览不同的 APIs管理界面、安全模型和操作流程，这会增加出错的可能性并增加运营开销。这种复杂性增加了出错和运营开销的机会，并可能阻碍敏捷性和可扩展性。

但是，在某些实际场景中，由于特定的业务或技术要求，组织可能需要在云之间分配连续的工作负载。在这些情况下，我们建议您制定明确的标准和指导原则来评估权衡取舍，并确保该方法与组织的整体多云战略保持一致。

当组织选择在多个云中分配工作负载时，采用以消息传递和松散耦合为中心的架构可以缓解许多相关的挑战。这是在云之间区分关注点的最佳方法，也是缩小提供商受损时的影响范围的最佳方法。最具时限的业务，例如金融交易，最好保持在单一环境中。切勿允许一个环境的中断危及另一个环境中的工作负载。

我们的指导方针：

- 设计云工作负载以实现运营独立性，从而最大限度地减少提供商之间的实时依赖性。当需要分配工作负载时，应实施高效的批量数据传输机制，而不是保持持续的跨云连接。
- 根据明确的业务标准评估每项建议的分布式工作负载。既要考虑分发带来的战略好处，又要考虑运营复杂性。

宗旨 5. 制定长期整合策略

在不同云中的应用程序之间移动大量数据时要小心，特别是如果您的计算资源和应用程序部署在一个 CSP 中，而您的数据存储资源部署在另一个 CSP 中。这种情况可能会增加复杂性和延迟，从而抵消感知到的好处。我们与许多客户进行了交谈，这些客户在一个云上拥有数据湖，但希望使用另一个 CSP 的工具进行机器学习 (ML) 或分析。决定在多云环境中放置工作负载的位置是组织面临的最关键、通常也是最具挑战性的决策之一。我们建议您通过三个关键维度来评估每项工作负载安置决策：技术要求、业务需求和提供商优势。

通过映射每个工作负载的基本特征（计算能力、数据操作、响应时间需求和增长要求）来开始技术评估。当应用程序位于数据附近时，它们自然会表现最好。将应用程序从其数据源中移开会造成不必要的技术障碍并降低性能。

业务决策必须考虑提供商的定价、数据驻留要求和供应商合同。每个工作负载的放置都会影响整个组织的运营、安全性和生产力。孤立地看待工作负载会导致做出次优决策。

我们的指导方针：

- 在云之间实现批量数据传输，而不是实时访问。使用高效的批量操作来安排定期数据刷新，而不是在云之间使用持续的 API 调用。这种方法可以降低成本、提高可靠性并保持稳定的性能。例如，导出汇总的每日销售数据，而不是跨云查询单个交易。
- 在设计工作负载布局时，请考虑数据引力。让应用程序靠近其主数据源，以保持性能并降低成本。机器学习模型、分析引擎和交易处理系统都受益于直接访问其数据。将这些工作负载从其数据中移开会造成不必要的网络延迟和复杂性。
- 在完整的云战略背景下评估工作负载决策，而不是孤立地对其进行审查。考虑每种安置选择如何影响整个组织的运营流程、安全控制和团队能力。从整体上看，似乎最适合单个工作负载的决策可能会使监控复杂化或增加安全风险。
- 定义明确的数据所有权和治理政策，指定不同类型数据的存放位置。创建数据分类框架，推动云提供商之间就数据放置做出一致的决策。

宗旨 6. 策略性地使用容器

容器在支持多云战略方面可以发挥宝贵的作用，但也必须认识到其局限性。对于任何现代的云原生应用程序来说，使用容器通常都是一个好主意，因为它们为不同环境的可移植性和一致性提供了好处。容器不受平台限制，这意味着它们可以在任何支持容器化技术的云平台或基础设施（例如 Kubernetes）上运行。使用容器的组织只需开发和打包一次应用程序，即可在多个云提供商或本地环境中一致地部署它们，而无需进行重大修改。通过将应用程序代码、依赖关系和运行时环境封装在容器中，您可以实现高度的可移植性，这使您能够在云提供商之间或云和本地数据中心之间无缝移动工作负载。

但是，容器可能无法解决所有用例，也无法消除组织在采用多云策略时可能面临的所有挑战。容器最适合基于微服务的现代架构，但它们可能不太适合大型单片应用程序。此外，尽管容器可以解决可移植性的某些方面，例如应用程序运行时，但它们并不能自动解决有关数据管理、安全策略和其他跨云依赖关系的问题。组织仍然需要仔细规划和架构其多云解决方案，以确保一致的数据管理、统一的安全控制以及云托管组件和本地组件之间的无缝集成。

我们的指导方针：

- 使用每家云提供商的原生容器管理功能，最大限度地提高业务价值并加快交付。这种方法可确保最佳性能，同时避免创建与云无关的解决方案的复杂性，而这些解决方案很少能带来有意义的回报。
- 制定容器策略，以解决整个运营问题，包括数据管理、安全性和跨云依赖关系。在做出容器架构决策时，请专注于业务成果。

宗旨 7. 有一个 CCo E，但要专攻它

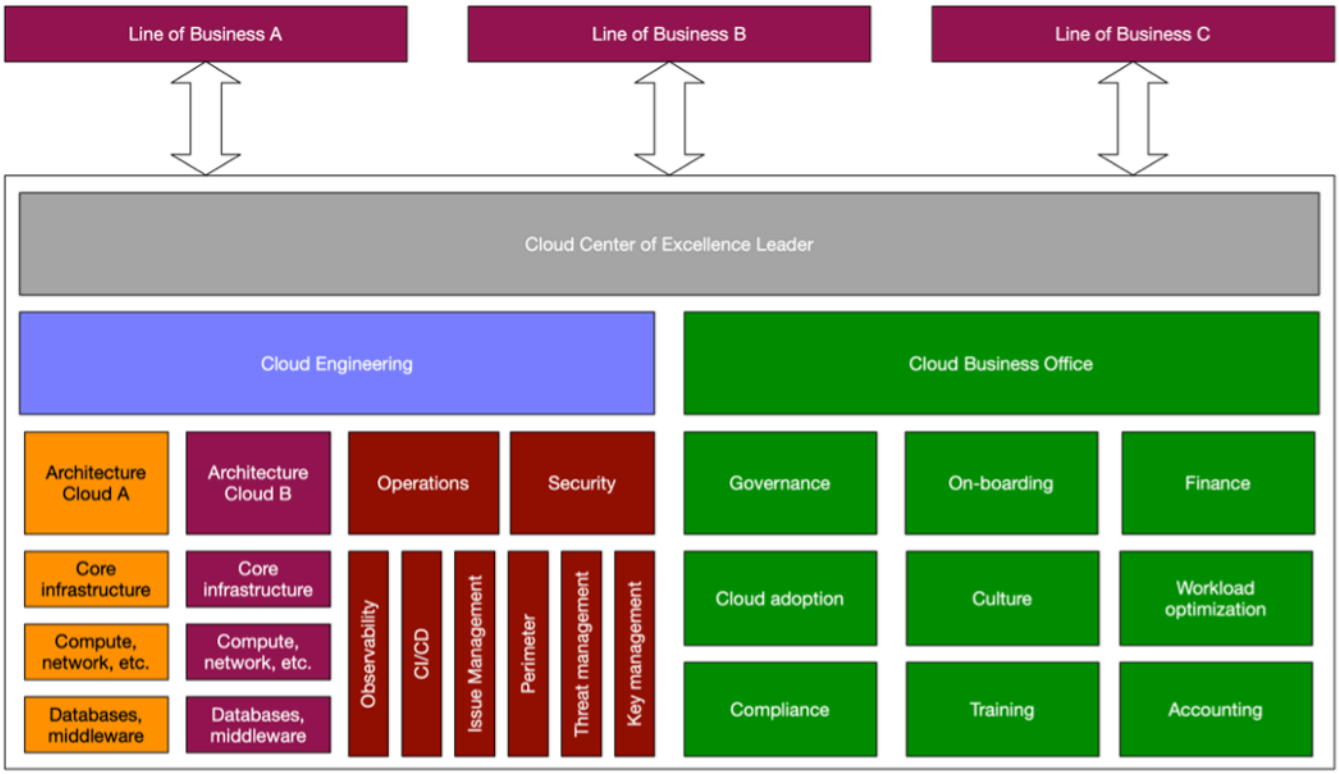
正如我们建议许多 AWS 客户的那样，您应该在组织内建立卓越云中心 (CCoE)，以提供领导力、标准化和加速您的云之旅。在多云环境方面，我们发现最成功的公司会采取平衡的方法来处理他们的E. CCo

我们建议您使用一个统一的 CCo E 来监督组织的多云战略，而不是为每个 CSP 建立单独的 CCo E。这有助于确保采用协调、一致的方法，而不是可能导致分歧、重新设计和浪费的孤立努力。确保您的单个 CCo E 中的团队拥有组织使用的每个 CSP 所必需的专业技能、工具和机制。这些专业知识使 CCo 得 E 能够有效地管理、支持和加速不同云平台的使用。

例如， CCoE 应该有深入了解 AWS Cloud、服务和最佳实践的 AWS特定专家，以及可以 CSPs 指导组织使用这些云技术的其他专家。单一 CCo E 中的这种专业知识可以帮助您的组织从集中式方法的协调和标准化中受益，同时确保每个云平台得到最佳利用。

单一 CCo E 应作为中央管理机构，负责为组织的多云战略制定标准、政策和最佳实践。云工作负载和项目的实际实施可以分配给专业团队或业务部门，而 CCOE 则提供监督、支持和协调。这种平衡的方法有助于确保统一的多云战略，同时在组织内部提供必要程度的灵活性和自主权。

下图说明了 CCo E 如何为多个业务线 (LOBs)、云工程团队和云业务办公室 (CBO) 团队提供集中式方法和治理。



我们的指导方针：

- 组织您的 CCo E 以保持战略监督，同时为每个云提供商嵌入专业知识。专注于招聘个人云平台方面的深厚专业知识，而不是寻找稀有的多云专家，并促进内部知识共享以建立组织能力。
- 让您的 CCo电子人员能够针对安全性和可观察性等跨领域问题制定企业级标准，同时让各个团队能够自主地使用云原生工具和服务在这些指导方针内执行这些指导方针。
- 制定全面的人才战略，在主要云平台的深厚专业知识和更广泛的架构知识之间取得平衡。专注于组建将强大的云特定技能与企业架构经验相结合的团队。

宗旨 8. 确保安全始终是重中之重

多云方法会增加未经授权访问的风险，从而更难确保安全，因为您的安全态势必须考虑到更多的攻击面。多云策略通常会迫使公司 CSPs在身份管理、网络安全、资产管理和审计日志等领域处理多种安全模式。这种复杂性有可能使透明度变得更加困难，增加安全团队的负担并增加风险。

在多云环境中，安全自动化至关重要。身份管理必须跨环境无缝运行；它必须连接现有的身份提供商，同时保持一致的访问策略。安全需要跨数据、网络和端点层的集成保护。数据分类、加密和生命周期管理构成了基础。网络安全建立在标准化设计和连接模式之上。端点保护通过一致的补丁管理和基于主机的控制来完善框架。

这些基础要素对于成功安全地采用多家云提供商至关重要，因此在任何多云战略规划中都必须尽早考虑这些要素。

我们的指导方针：

- 在您的多云环境中实施一个集成的安全框架，该框架侧重于三个核心要素：通过标准化分类和加密实现数据保护，通过一致的设计模式实现网络安全，以及通过系统控制和补丁管理实现端点保护。
- 建立统一的安全运营模式，利用每个云提供商的原生安全功能，同时通过标准化的工具和流程保持集中的可见性和控制。
- 使用 Amazon Security Lake 集中收集和分析安全数据。该平台将来自其他云提供商 AWS、SaaS 应用程序和本地系统的安全信息聚合到一个视图中。它支持开放网络安全架构框架 (OCSF)，并支持跨混合云和多云环境的标准化分析。这种集中式方法可以改善威胁检测和响应，同时简化安全操作。
- 部署每个提供商的本地安全工具，以增强您的保护能力。这些专门构建的服务可解决提供商特定的功能，同时将数据反馈给您的集中式安全平台。本机工具和集中式可视性相结合，有助于为您的整个基础架构提供全面的安全保护。
- 实施统一的可观测性策略，从头开始提供对整个云环境的全面可见性，包括运营和安全数据。采用业界领先的监控方法实现标准化，无论业务服务在何处运营，都能实现一致的跟踪。
- 建立企业范围的运营数据收集和可视化标准，从而在多云环境中快速识别和解决问题。专注于为技术和业务利益相关者提供服务的运营洞察创建单一事实来源。

宗旨 9. 采用 80/20 方法而不是等额分配

如何在提供商之间分配工作负载从根本上决定了您的多云成功。许多组织错误地追求云分布的平等，并试图在提供商之间平均分配工作负载。这种方法增加了复杂性，但不会带来成比例的收益。平等分配会削弱您的技术能力，削弱您的购买力，并造成不必要的运营开销。当团队被迫同时在多个平台上保持能力时，他们很难发展深厚的专业知识。

事实证明，80/20 方法提供的结果要好于在云间均匀分布。将80%的投资集中在一家主要提供商身上，同时有选择地使用其他提供商来实现特定功能，从而形成一种平衡的策略，可以降低成本和复杂性。这种集中的方法可以加速创新，因为您的团队可以利用主平台的高级服务积累深厚的专业知识。您的技术人员可以成为一个架构的专家，而不必在多个环境中保持表面层面的知识。当工程师掌握一个平台时，他们可以更高效地构建，更快地排除故障，并实施更复杂的解决方案。

遵循80/20方法的公司通常会报告人才留存率更高，因为他们的团队培养了宝贵的、适销对路的专业知识，而不是在多种技术上捉襟见肘。这种集中的策略还可以限制不同提供商之间不同安全模型的复杂性，从而帮助简化安全管理。主云将您在安全工具、监控解决方案和运营流程方面的大部分投资用于主云。与资源平均分配相比，这创造了更强大的安全基础。

我们的指导方针：

- 选择符合您的大部分业务和技术要求的主要云提供商。该提供商应支持您至少 80% 的工作负载，并成为您的云战略的基础。将您的培训投资、架构标准和运营流程集中在最大限度地提高该主要平台的价值上。
- 为需要在二级云上部署的工作负载制定明确的标准。这些标准应侧重于您的主要提供商无法实现的特定业务价值。不要仅仅为了维持支出公平或在提供商之间进行人为平衡而将工作负载放在二级云上。
- 组织您的企业协议以反映您的 80/20 方法。根据集中支出与您的主要提供商协商批量折扣，并在特定用例中与二级提供商保持灵活性。与平均分配支出相比，这种方法可以最大限度地提高您的购买杠杆率，并且通常可以获得更好的总体定价。
- 使您的人才战略与 80/20 方法保持一致。投资于在主要提供商的服务方面积累深厚的专业知识，同时保持对辅助平台的足够了解，以支持特定的工作负载。这种有针对性的人才战略可以提高生产力，加快交付速度，并降低出现关键技能缺口的风险。
- 定期衡量多云战略的业务成果。跟踪显示从每个提供商那里获得的价值的指标，并在必要时调整您的分布。目标不是完全避免多云，而是战略性地实施多云，让特定工作负载真正受益于其他提供商独有的功能。

结论

此 paper 概述了制定有效多云战略的九个关键原则。Organizations 通过主云方法取得最大的成功，并在特定业务需求需要时战略性地使用其他提供商。我们所描述的80/20方法在重点与灵活性之间取得平衡，使组织能够发展更深入的专业知识，保持更牢固的提供商关系，培养更有价值的人才，同时仍能满足合法的多云需求。

成功的多云实施需要对业务需求进行清晰的评估，而不是关注行业趋势。公司必须建立强大的治理，将安全性作为重中之重，避免在提供商之间分散互联的工作负载，保留应用程序及其交易数据，识别容器的限制，并维护统一但专业的卓越云中心。

云 AWS 方法从根本上建立在客户选择和互操作性之上。我们设计的工具和服务可以跨环境无缝运行，因为我们知道您的业务需求通常超出单一提供商的范围。从混合连接解决方案到跨环境的容器编排，AWS 提供的功能可帮助您在整个技术环境中高效运营。

与其强迫您成为多个平台的专家，不如通过直观的工具和一致的界面来 AWS 简化多云管理。我们专注于消除复杂性，因此您可以专注于创新。这些功能可帮助您按照自己的条件实施多云策略，无论这意味着 AWS 仅使用还是与其他环境一起使用特定的 AWS 服务 环境。

云应该为您的业务战略提供支持，而不是限制它。通过应用本 paper 中概述的原则并利用 AWS 互操作性功能，您可以构建一种云方法，该方法可以最大限度地提高价值，最大限度地减少不必要的复杂性，并使您的组织在当今不断变化的业务环境中取得长期成功。

要详细了解可帮助简化混合云和多云环境管理的 AWS 解决方案，请参阅多云AWS 解决方案。

资源

参考

- 使用云卓越中心 (CCOE) 实现整个企业的转型（AWS 博客文章）
- AWS 架构完善的框架
- 通过成本优化中心发现机会（AWS Cost Management 文档）
- 迁移到亚马逊 Web Services 的商业价值（哈克特集团，2022 年 2 月）
- 出门时可免费将数据传输到互联网 AWS（AWS 博客文章）

工具

- 区域自动切换 — 当我们检测到潜在问题时，自动将您的流量从可用区域转移出去（AWS 博客文章）
- AWS 多云解决方案

AWS 合作伙伴

- AWS Cloud 运营能力

文档历史记录

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

变更	说明	日期
初次发布	—	2025 年 9 月 3 日

AWS 规范性指导词汇表

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

数字

7 R

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

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

A

A2A () Agent-to-Agent

一种支持任务委托和状态转移的代理到代理协作的状态协议。

ABAC

请参阅基于属性的访问控制。

抽象服务

请参阅托管服务。

ACID

请参阅原子性、一致性、隔离性、持久性。

主动-主动迁移

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

主动-被动迁移

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

座席

一种能够使用工具自主推理、计划和采取行动来实现目标的人工智能系统。

特工行动

在生产环境中大规模构建、测试、部署和运行 AI 代理的操作实践。

聚合函数

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

AI

请参阅人工智能。

AIOps

请参阅人工智能运营。

匿名化

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

反模式

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

应用程序控制

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

应用程序组合

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

人工智能（AI）

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

人工智能运营（AIOps）

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

非对称加密

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

原子性、一致性、隔离性、持久性（ACID）

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

基于属性的访问权限控制（ABAC）

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

权威数据来源

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

可用区

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

AWS 云采用框架 (AWS CAF)

该框架包含指导方针和最佳实践 AWS ，可帮助组织制定高效且有效的计划，以成功迁移到云端。AWS CAF将指导分为六个重点领域，称为视角：业务、人员、治理、平台、安全和运营。业务、人员和治理角度侧重于业务技能和流程；平台、安全和运营角度侧重于技术技能和流程。例如，人员角度针对的是负责人力资源（HR）、人员配置职能和人员管理的利益相关者。从这个角度来看，AWS CAF 为人员发展、培训和沟通提供了指导，以帮助组织为成功采用云做好准备。有关更多信息，请参阅 AWS CAF 网站和 AWS CAF 白皮书。

AWS 工作负载资格框架 (AWS WQF)

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

B

恶意机器人

一种旨在扰乱或伤害个人或组织的机器人。

BCP

请参阅业务连续性计划。

行为图

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

大端序系统

一个先存储最高有效字节的系统。另请参阅字节顺序。

二进制分类

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

bloom 筛选条件

一种概率性、内存高效的数据结构，用于测试元素是否为集合的成员。

blue/green 部署

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

自动程序

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

僵尸网络

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

分支

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

紧急（break-glass）访问

在特殊情况下，通过批准的流程，用户 AWS 账户 可以快速访问他们通常没有访问权限的内容。有关更多信息，请参阅指南中的 “实施破碎玻璃程序” AWS Well-Architected 指示器。

棕地策略

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

缓冲区缓存

存储最常访问的数据的内存区域。

业务能力

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

业务连续性计划（BCP）

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

C

CAF

请参阅 AWS 云采用框架。

金丝雀部署

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

CCoE

请参阅云卓越中心。

CDC

请参阅更改数据捕获。

更改数据捕获（CDC）

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

混沌工程

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

CI/CD

请参阅持续集成和持续交付。

分类

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

公民开发者

使用无code/low代码平台创建 AI 应用程序但没有专业技术技能的企业用户。

客户端加密

在目标 AWS 服务 收到数据之前，对数据进行本地加密。

云卓越中心（CCoE）

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

云计算

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

云运营模型

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

云采用阶段

组织迁移到 AWS Cloud中时通常会经历四个阶段：

- 项目 - 出于概念验证和学习目的，开展一些与云相关的项目
- 基础 - 进行基础投资以扩大云采用率（例如，创建登录区、定义 CCoE、建立运营模型）
- 迁移 - 迁移单个应用程序
- Re-invention — 优化产品和服务，在云端进行创新

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

CMDB

请参阅配置管理数据库。

代码存储库

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

冷缓存

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

冷数据

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

计算机视觉（CV）

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

配置偏移

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

配置管理数据库（CMDB）

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

合规性包

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

持续集成和持续交付 (CI/CD)

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

CV

请参阅计算机视觉。

D

静态数据

网络中静止的数据，例如存储中的数据。

数据分类

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

数据漂移

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

传输中数据

在网络中主动移动的数据，例如在网络资源之间移动的数据。

数据网格

一种架构框架，可提供分布式、去中心化的数据所有权以及集中式管理和治理。

数据最少化

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

数据边界

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

数据预处理

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

数据溯源

在数据的整个生命周期跟踪其来源和历史的过程，例如数据如何生成、传输和存储。

数据主体

正在收集和处理其数据的个人。

数据仓库

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

数据库定义语言（DDL）

在数据库中创建或修改表和对象结构的语句或命令。

数据库操作语言（DML）

在数据库中修改（插入、更新和删除）信息的语句或命令。

DDL

请参阅数据库定义语言。

深度融合

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

深度学习

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

深度防御

一种信息安全方法，经过深思熟虑，在整个计算机网络中分层实施一系列安全机制和控制措施，以保护网络及其中数据的机密性、完整性和可用性。当你采用这种策略时 AWS，你会在 AWS Organizations 结构的不同层面添加多个控件来帮助保护资源。例如，深度防御方法可能将多因素身份验证、网络分段和加密结合起来。

委派管理员

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

部署

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

开发环境

请参阅环境。

侦测性控制

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

开发价值流映射（DVSM）

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

数字孪生

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

维度表

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

灾难

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

灾难恢复（DR）

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

DML

请参阅数据库操作语言。

领域驱动设计

一种开发复杂软件系统的方法，通过将其组件连接到每个组件所服务的不断发展的领域或核心业务目标。埃里克·埃文斯（Eric Evans）在他的《Domain-Driven 设计：解决软件核心的复杂性》（波士顿： Addison-Wesley 专业版，2003年）一书中介绍了这个概念。有关如何使用带有 strangler fig 模式的域驱动设计的信息，请参阅使用容器和 Amazon API Gateway 逐步实现传统微软 ASP.NET (ASMX) 网络服务的现代化。

DR

请参阅灾难恢复。

偏差检测

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

DVSM

请参阅开发价值流映射。

E

EDA

请参阅探索性数据分析。

EDI

请参阅电子数据交换。

边缘计算

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

电子数据交换（EDI）

组织之间业务文件的自动交换。有关更多信息，请参阅什么是电子数据交换。

加密

一种将人类可读的纯文本数据转换为加密文字的计算流程。

加密密钥

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

字节顺序

字节在计算机内存中的存储顺序。 Big-endian 系统首先存储最重要的字节。 Little-endian 系统首先存储最低有效字节。

端点

请参阅服务端点。

端点服务

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

企业资源规划（ERP）

一种自动化和管理企业关键业务流程（例如会计、MES 和项目管理）的系统。

信封加密

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

环境

正在运行的应用程序的实例。以下是云计算中常见的环境类型：

- 开发环境 — 正在运行的应用程序的实例，只有负责维护应用程序的核心团队才能使用。开发环境用于测试更改，然后再将其提升到上层环境。这类环境有时称为测试环境。
- 下层环境 — 应用程序的所有开发环境，比如用于初始构建和测试的环境。
- 生产环境 — 最终用户可以访问的正在运行的应用程序的实例。在 CI/CD 管道中，生产环境是最后一个部署环境。
- 上层环境 — 除核心开发团队以外的用户可以访问的所有环境。这可能包括生产环境、预生产环境和用户验收测试环境。

epic

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

ERP

请参阅企业资源规划。

探索性数据分析（EDA）

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

F

事实表

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

快速失效机制

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

故障隔离边界

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

功能分支

请参阅分支。

特征

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

特征重要性

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

功能转换

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

少样本提示

在要求 LLM 执行类似任务之前，先向其提供少量示例，以演示任务和预期输出。这种技术是情境学习的应用，模型可以从提示中嵌入的示例（镜头）中学习。 Few-shot 对于需要特定格式、推理或领域知识的任务，提示可能非常有效。另请参阅零样本提示。

FGAC

请参阅精细访问控制。

精细访问控制（FGAC）

使用多个条件允许或拒绝访问请求。

快闪迁移

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

FM

请参阅基础模型。

基础模型（FM）

一个大型深度学习神经网络，它已使用海量的通用和未标注数据集进行训练。FM 能够执行各种常规任务，例如理解语言、生成文本和图像以及使用自然语言进行对话。有关更多信息，请参阅什么是基础模型。

FM 网关

一种集中式中介，用于控制和规范对基础模型的访问。也称为 LLM 网关。

G

生成式人工智能

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

地理阻止

请参阅地理限制。

地理限制（地理阻止）

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

GitFlow 工作流程

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

黄金映像

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

全新策略

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

防护机制

一种高级规则，用于跨组织单位（OU）管理资源、策略和合规性。预防性防护机制会执行策略以确保符合合规性标准。它们是使用服务控制策略和 IAM 权限边界实现的。侦测性护栏会检测策略违规和合规性问题，并生成提醒以进行修复。它们通过使用 AWS Config、、Amazon、 AWS Security Hub CSPM GuardDuty AWS Trusted Advisor、Amazon Inspector 和自定义 AWS Lambda 支票来实现。

护栏 (AI)

用于过滤、验证和限制代理输入和输出的安全机制，有助于确保负责任和安全的 AI 行为。

H

HA

请参阅高可用性。

异构数据库迁移

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

高可用性（HA）

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

历史数据库现代化

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

保留数据

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

人机在圈 (HitL)

一种工作流程模式，其中代理执行在关键决策点暂停以供人工审查和批准。

同构数据库迁移

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

热数据

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

修补程序

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

hypercare 周期

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

我

IaC

请参阅基础设施即代码。

基于身份的策略

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

空闲应用程序

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

IIoT

请参阅工业物联网。

不可变基础设施

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

入站（入口）VPC

在 AWS 多账户架构中，一种接受、检查和路由来自应用程序外部的网络连接的 VPC。AWS 安全参考架构建议使用入站、出站和检查 VPC 设置网络账户，保护应用程序与广泛的互联网之间的双向接口。

增量迁移

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

工业 4.0

该术语由克劳斯·施瓦布（Klaus Schwab）在2016年推出，指的是通过连接性、实时数据、自动化、分析和的进步实现制造流程的现代化。 AI/ML

基础设施

应用程序环境中包含的所有资源和资产。

基础设施即代码（IaC）

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

工业物联网（IIoT）

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

检查 VPC

在 AWS 多账户架构中，一种集中式 VPC，用于管理 VPC（相同或不同 AWS 区域）、互联网和本地网络之间的网络流量检查。AWS 安全参考架构建议使用入站、出站和检查 VPC 设置网络账户，保护应用程序与广泛的互联网之间的双向接口。

物联网（IoT）

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

可解释性

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

物联网

请参阅物联网。

IT 信息库（ITIL）

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

IT 服务管理（ITSM）

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

ITIL

请参阅 IT 信息库。

ITSM

请参阅 IT 服务管理。

L

基于标签的访问控制（LBAC）

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

登录区

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

大语言模型（LLM）

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

大规模迁移

迁移 300 台或更多服务器。

LBAC

请参阅基于标签的访问控制。

最低权限

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

直接迁移

请参阅 7 R。

小端序系统

一个先存储最低有效字节的系统。另请参阅字节顺序。

LLM

请参阅大型语言模型。

下层环境

请参阅环境。

M

机器学习（ML）

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

主分支

请参阅分支。

恶意软件

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

托管式服务

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

制造执行系统（MES）

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

MAP

请参阅迁移加速计划。

MCP

参见模型上下文协议。

模型上下文协议（MCP）

一种用于代理与工具通信的无状态协议。

MCP 服务器

一种通过模型上下文协议公开一个或多个工具的服务。

机制

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

成员账户

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

MES

请参阅制造执行系统。

消息队列遥测传输（MQTT）

一种基于publish/subscribe模式的轻量级机器对机器 (M2M) 通信协议，适用于资源受限的物联网设备。

微服务

一种小型独立服务，通过明确定义的 API 进行通信，通常由小型独立团队拥有。例如，保险系统可能包括映射到业务能力（如销售或营销）或子域（如购买、理赔或分析）的微服务。微服务的好处包括敏捷、灵活扩展、易于部署、可重复使用的代码和恢复能力。有关更多信息，请参阅使用 AWS 无服务器服务集成微服务。

微服务架构

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

迁移加速计划（MAP）

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

大规模迁移

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

迁移工厂

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

迁移元数据

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

迁移模式

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

迁移组合评测（MPA）

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

迁移准备情况评测（MRA）

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

迁移策略

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

ML

请参阅机器学习。

现代化

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

现代化准备情况评估

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

单体应用程序（单体式）

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

MPA

请参阅迁移组合评测。

MQTT

请参阅消息队列遥测传输。

多分类器

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

可变基础设施

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

O

OAC

请参阅来源访问控制。

OAI

请参阅来源访问身份。

OCM

请参阅组织变革管理。

离线迁移

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

OI

请参阅运营集成。

OLA

请参阅运营级别协议。

在线迁移

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

OPC-UA

请参阅开放流程通信 – 统一架构。

开放流程通信-统一架构 (OPC-UA)

一种用于工业自动化的机器对机器 (M2M) 通信协议。 OPC-UA 提供了数据加密、身份验证和授权方案的互操作性标准。

运营级别协议（OLA）

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

运营准备情况审查（ORR）

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

运营技术（OT）

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

运营整合（OI）

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

组织跟踪

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

组织变革管理（OCM）

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

来源访问控制（OAC）

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

来源访问身份（OAI）

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

ORR

请参阅运营准备情况审查。

OT

请参阅运营技术。

出站（出口）VPC

在 AWS 多账户架构中，一种处理从应用程序内部启动的网络连接的 VPC。AWS 安全参考架构建议使用入站、出站和检查 VPC 设置网络账户，保护应用程序与广泛的互联网之间的双向接口。

P

权限边界

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

个人身份信息（PII）

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

PII

请参阅个人身份信息。

playbook

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

PLC

请参阅可编程逻辑控制器。

PLM

请参阅产品生命周期管理。

policy

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

多语言持久性

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

组合评测

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

谓词

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

谓词下推

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

预防性控制

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

主体

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

隐私设计

一种在整个开发过程中都考虑隐私的系统工程方法。

私有托管区

私有托管区就是一个容器，其中包含的信息说明您希望 Amazon Route 53 如何响应一个或多个 VPC 中的某个域及其子域的 DNS 查询。有关更多信息，请参阅 Route 53 文档中的私有托管区的使用。

主动控制

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

产品生命周期管理（PLM）

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

生产环境

请参阅环境。

可编程逻辑控制器（PLC）

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

提示串接

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

假名化

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

publish/subscribe (pub/sub)

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

Q

查询计划

一系列用于访问 SQL 关系数据库系统中的数据的步骤，类似于指令。

查询计划回归

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

R

RACI 矩阵

请参阅责任、问责、咨询和知情（RACI）。

RAG

请参阅检索增强生成。

勒索软件

一种恶意软件，旨在阻止对计算机系统或数据的访问，直到付款为止。

RASCI 矩阵

请参阅责任、问责、咨询和知情（RACI）。

RCAC

请参阅行列访问控制。

只读副本

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

重新架构

请参阅 7 R。

恢复点目标（RPO）

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

恢复时间目标（RTO）

服务中断和服务恢复之间可接受的最大延迟。

重构

请参阅 7 R。

Region

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

回归

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

重新托管

请参阅 7 R。

版本

在部署过程中，推动生产环境变更的行为。

重新放置

请参阅 7 R。

更换平台

请参阅 7 R。

重新购买

请参阅 7 R。

韧性

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

基于资源的策略

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

责任、问责、咨询和知情（RACI）矩阵

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

响应性控制

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

保留

请参阅 7 R。

停用

请参阅 7 R。

检索增强生成（RAG）

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

轮换

定期更新密钥以使攻击者更难访问凭证的过程。

行列访问控制（RCAC）

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

RPO

请参阅恢复点目标。

RTO

请参阅恢复时间目标。

运行手册

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

S

SAML 2.0

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

SCADA

请参阅监督控制和数据采集。

SCP

请参阅服务控制策略。

机密密钥

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

安全设计

一种在整个开发过程中都考虑安全的系统工程方法。

安全控制

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

安全固化

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

安全信息和事件管理（SIEM）系统

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

安全响应自动化

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

服务器端加密

由接收数据的人在目的地对数据 AWS 服务 进行加密。

服务控制策略（SCP）

一种策略，用于集中控制 AWS Organizations的组织中所有账户的权限。SCP 为管理员可以委托给用户或角色的操作定义了防护机制或设定了限制。您可以将 SCP 用作允许列表或拒绝列表，指定允许或禁止哪些服务或操作。有关更多信息，请参阅 AWS Organizations 文档中的服务控制策略。

服务端点

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

服务水平协议（SLA）

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

服务水平指示器（SLI）

对服务性能方面的衡量，例如错误率、可用性或吞吐量。

服务水平目标（SLO）

代表服务运行状况的目标指标，由服务水平指示器衡量。

责任共担模式

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

暗影人工智能

在组织内受管控渠道之外构建或使用的未经授权的 AI 应用程序。

SIEM

请参阅安全信息和事件管理系统。

单点故障（SPOF）

应用程序的单个关键组件出现故障，可能会中断系统。

SLA

请参阅服务水平协议。

SLI

请参阅服务水平指示器。

SLO

请参阅服务水平目标。

split-and-seed 模式

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

SPOF

请参阅单点故障。

星型架构

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

strangler fig 模式

一种通过逐步重写和替换系统功能直至可以停用原有的系统来实现单体系统现代化的方法。这种模式用无花果藤作为类比，这种藤蔓成长为一棵树，最终战胜并取代了宿主。该模式是由 Martin

Fowler 提出的，作为重写单体系统时管理风险的一种方法。有关如何应用此模式的示例，请参阅使用容器和 Amazon API Gateway 逐步实现传统微软 ASP.NET (ASMX) 网络服务的现代化。

子网

您的 VPC 内的一个 IP 地址范围。子网必须位于单个可用区中。

监督控制和数据采集（SCADA）

在制造业中，一种使用硬件和软件来监控实物资产和生产操作的系统。

对称加密

一种加密算法，它使用相同的密钥来加密和解密数据。

综合测试

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

系统提示

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

T

标签

Key-value 对充当用于组织 AWS 资源的元数据。标签有助于您管理、识别、组织、搜索和筛选 资源。有关更多信息，请参阅标记您的 AWS 资源。

目标变量

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

任务列表

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

测试环境

请参阅环境。

训练

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

工具

代理可以调用以在外部系统中执行操作的函数或 API。

中转网关

中转网关是网络中转中心，您可用它来互连 VPC 和本地网络。有关更多信息，请参阅 AWS Transit Gateway 文档中的什么是公交网关。

基于中继的工作流程

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

可信访问权限

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

优化

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

双披萨团队

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

U

不确定性

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

无差别任务

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

上层环境

请参阅环境。

V

vacuum 操作

一种数据库维护操作，包括在增量更新后进行清理，以回收存储空间并提高性能。

版本控制

跟踪更改的过程和工具，例如存储库中源代码的更改。

VPC 对等连接

两个 VPC 之间的连接，允许您使用私有 IP 地址路由流量。有关更多信息，请参阅 Amazon VPC 文档中的什么是 VPC 对等连接。

漏洞

损害系统安全的软件缺陷或硬件缺陷。

W

热缓存

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

暖数据

不常访问的数据。查询此类数据时，通常可以接受中速查询。

窗口函数

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

工作负载

一系列资源和代码，它们可以提供商业价值，如面向客户的应用程序或后端过程。

工作流

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

WORM

请参阅一次写入多次读取。

WQF

请参阅 AWS 工作负载资格鉴定框架。

一次写入多次读取（WORM）

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

Z

零日漏洞利用

一种利用零日漏洞的攻击，通常为恶意软件。

零日漏洞

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

零样本提示

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

僵尸应用程序

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

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