探索跨多个业务领域开发边缘云原生应用的挑战

Pawissanutt Lertpongrujikorn*、Hai Duc Nguyen†、Juahn Kwon*、Mohsen Amini Salehi*
*北德克萨斯大学高性能云计算（HPCC）实验室，美国 电子邮箱：{pawissanutt.lertpongrujikorn、juahn.kwon、mohsen.aminisalehi}@unt.edu
†阿贡国家实验室与芝加哥大学，美国 电子邮箱：hai.nguyen@anl.gov

arXiv:2603.03738v1 [cs.DC] 4 Mar 2026

摘要—随着云计算与先进网络的融合持续重塑现代软件开发，边缘云原生范式已成为实现可扩展、弹性且敏捷的数字服务的关键，这些服务依赖高性能、低时延和可靠通信。本研究通过对来自多个领域的专业人士进行深度访谈，探讨开发、部署和维护边缘云原生应用所面临的实践挑战，受访者涵盖信息技术、金融、医疗保健、教育和工业界。尽管云技术取得了显著进展，实践者——尤其是非技术背景人士——仍持续面临相当大的复杂性，原因包括工具链碎片化、学习曲线陡峭，以及管理分布式网络与计算、确保混合环境中的性能一致，并应对云网边界处陡峭的学习曲线所带来的运维开销。各行业的参与者一致认为，生产力、服务质量和易用性比成本或迁移等传统关切更为重要。这些发现凸显了对运维简化、具备服务等级协议感知能力且对开发者友好的平台的需求，以简化完整的应用生命周期。本研究从实践出发，为边缘云原生系统与现代企业的现实情况及需求相契合提供了视角，并为推动无缝云网融合提供了重要见解。

索引词—**边缘云原生、云计算、多领域、访谈、实践挑战**

I. 引言

云计算与高速网络的融合，通过实现前所未有的可扩展性、敏捷性和韧性，从根本上重塑了现代商业战略 [1]。这一演进正推动计算向数据源头靠近，催生了边缘计算和雾计算等分布式架构。因此，市场对相关模式的需求日益增长：这些模式不仅支持计算，还能确保低时延响应能力、本地化决策，并支持在混合云到边缘的连续体系中灵活部署 [2], [3]。底层网络结构不再只是简单的传输层，而是服务交付中至关重要、可编程的组成部分。这些能力对多个业务领域都至关重要，例如医疗保健、物联网、金融和机器人；在这些领域，快速、灵活且具备网络感知能力的软件开发与部署不可或缺。

尽管云原生技术取得了重大进展，学术研究中认识到的挑战与从业者在现实业务领域中所面临的挑战之间仍存在长期未解的差距 [4]。学术研究往往忽视了一些关键限制，例如陡峭的学习曲线、支离破碎的工具链以及有限的技术能力——这些问题在小型企业和特定领域的组织中尤为普遍。这些疏漏部分源于学术界依赖理想化假设，包括低估成本影响、简化系统模型，以及很少考虑组织限制。云平台的快速演进，以及学术研究人员与行业利益相关者之间缺乏持续合作，进一步扩大了这一脱节。因此，许多提出的云原生框架尽管在概念上前景可观，但在运营环境中应用时仍无法达到预期。

为弥合研究与实践鸿沟，本文介绍了对直接参与云原生系统设计、开发和采用的领域专业人员进行访谈后获得的初步实证发现。这些访谈旨在揭示学术讨论中往往未得到充分呈现的问题——具体而言，从业者面临的实际痛点、他们对新兴技术的期望，以及影响云采用和迁移决策的业务因素。通过聚焦这些现实视角，本研究力求揭示阻碍学术创新在实践中落地的潜在挑战。为此，我们倡导制定更切合实际且不断迭代的研究议程：不仅优先考虑技术新颖性，也兼顾现代软件开发生态系统在运维、组织和经济方面的现实情况。

本研究由以下假设驱动：

假设。 当前现实中的云原生应用开发、部署和维护过程复杂，需要投入大量时间和专业知识，从而增加总体成本。

这一假设源于对现有云原生系统和行业报告的全面审视。通过

我们的初步调查，我们发现了一个反复出现的模式：领域专业人员常常难以应对碎片化的工具链、复杂的部署流程，以及与分布式计算相关的陡峭学习曲线。

为验证这一假设，我们对多个业务领域的21名开发人员和行业专业人士进行了访谈，提出了一系列问题，重点关注他们面临的痛点、解决方案或对解决方案的期望，以及实施这些方案的障碍。

本文的主要贡献如下：

1）通过访谈（或调查）识别工业界在开发云原生应用时面临的真实需求与挑战。2）分析访谈结果，提炼与实践中应用开发相关的关键洞见。3）基于数据驱动的发现与实际需求，提出未来研究方向。

本文其余部分结构如下。在第II节，我们阐述了本研究的动机。第III 节说明了所开展访谈的策略。第IV、V以及VI 节分别提供了源自访谈的详细数据——痛点、期望以及迁移中的顾虑。在第VII节，我们讨论了主要分析。第VIII节总结了本工作，并对未来工作提出了简要建议。

II. 动机

云计算于 2000 年代初兴起，与虚拟化和网络托管技术的发展同步[5]。VMware 等虚拟化工具 [6] 实现了硬件抽象，使多台虚拟机（VM）能够在单台物理服务器上运行，从而大幅提高资源利用率。这种抽象为现代云平台奠定了基础。通过隐藏基础设施的复杂性，云有望减轻基础设施管理负担，并赋能组织，尤其是技术能力有限的组织，更轻松地开发和扩展数字服务。这一愿景促使容器编排系统 [7] 和无服务器计算 [8], [9]等云原生技术得到广泛采用。

尽管现代云平台潜力巨大，但它们已演变成一个高度复杂且支离破碎的生态环境。事实上，主要的商业云服务提供商（例如，亚马逊 AWS [5], Google Cloud [10], 和 Microsoft Azure [11]）自**2006**[12]以来累计推出了超过**500**项新服务。其中许多服务需要深厚的技术专长才能有效地配置、集成和运维。Kubernetes [7], 函数即服务（FaaS） [8], [9],CI/CD 流水线 [13], 和微服务架构等技术叠加在一起，带来了陡峭的学习曲线，并要求开发人员掌握数量庞大且不断增长的工具和最佳实践。因此，部署应用程序或配置其性能等基本日常任务，如今都需要具备较高水平的云原生原则知识和特定平台知识 [14]。

这种不断升级的技术与概念复杂性，直接与云关于精简的根本承诺相矛盾

信息技术和普遍可访问性。对于没有专职技术团队的初创企业、小型企业以及聚焦特定领域的组织而言，采用云原生解决方案可能复杂到难以承受。这类组织常常面临人才短缺、开发周期延长、调试与集成开销，以及预算受限等挑战。此外，领域专家与技术团队之间的协作可能因沟通鸿沟和心智模型不匹配而受阻。将特定领域需求转化为稳健的云原生实现往往并非易事，从而进一步提高了入门门槛。

为更好地了解这些挑战的范围和性质，我们提议对参与云原生应用设计、开发、部署和维护的专业人士开展结构化访谈，包括开发人员、解决方案架构师、领域专家和组织决策者。这些访谈旨在评估现实世界中的云实践是否符合该假设。

III. 访谈方法

为验证这一假设，我们与来自不同领域的技术专业人员进行了**21** 次独立访谈。这些访谈旨在揭示实际挑战、评估工具使用情况，并了解影响云原生应用开发与部署成本和复杂性的决策流程。访谈主要通过 Zoom 在线进行，时间为2025年2月至3月。每场访谈通常持续30分钟，先围绕一组统一的核心问题展开（见下文），随后进行开放式讨论，以深入了解每位参与者的独特经历和专业知识。为确保透明度、一致性和可追溯性，每场访谈期间均由两至三名访谈员详细记录，包括保存访谈录音和访谈笔记。

领域为确保广泛的代表性，我们从农业、教育、金融、医疗保健、工业、物联网、社交网络、硬件制造、网络安全和技术顾问这十个不同领域的组织中选取了参与者。这些领域涵盖了各种系统架构、运营约束和特定领域的需求。为简化分析，我们将组织分为两组：

- 技术（占访谈的 38%）：提供数字解决方案的组织，例如网络安全服务、技术咨询公司和硬件供应商。
- 非技术（占 62%）：核心业务并非以技术为中心、但仍依赖软件平台和计算基础设施的其他组织（例如农业和医疗保健领域的组织）。

参与者及其角色：我们还涵盖了不同角色的参与者，包括软件开发人员、DevOps工程师、测试工程师、初创企业家和领域专家。这种组合使我们能够收集一线技术视角和高层战略见解，了解

云原生技术如何在其组织中得到采用、维护和发展。

参与者人口统计特征与技术格局：为全面呈现我们的样本，表I总结了 21 名参与者的人口统计特征和技术采用模式。

人口统计分布呈现出几个显著模式。首先，我们的样本涵盖了广泛的领域，涉及十个不同的业务领域。非技术型组织与技术型组织占比为 62%/38%，这符合我们的目标：既收集可能缺乏深厚技术专长的领域导向型企业的观点，也纳入技术导向型公司的观点。

在基础设施采用方面，传统的基础设施即服务仍最为普遍（62%），反映出组织仍持续依赖虚拟机并直接控制基础设施。不过，托管服务正逐渐普及，52% 的组织使用平台即服务，48% 采用无服务器／函数即服务架构。这表明组织正逐步转向更高层次的抽象，以减少运维开销。

值得注意的是，38% 的参与者部署了边缘计算，这凸显出分布式架构对于延迟敏感型和网络受限型应用日益重要。在这些参与者中，29% 采用了边缘云混合配置，这表明纯边缘或纯云解决方案往往无法满足现实需求。多云策略（19%）和私有云部署（14%）的存在，进一步凸显了现代云原生格局的异质性和复杂性。

组织规模分布显示，三分之一的参与者来自大型企业（>500 名员工），而近40%来自中小型组织。这种构成有助于了解资源约束和组织能力如何影响云采用策略及痛点。

在确定样本的人口构成后，我们接下来转向访谈实施。半结构化访谈采用标准化协议进行，以确保一致性，同时允许对新出现的主题进行开放式探索。

核心访谈问题每次访谈都从一组一致的引导性问题开始，以收集可比数据，同时仍为领域特定阐述保留灵活性。重点领域包括：

- 技术。 目前正在使用哪些技术？正在考虑和评估哪些替代方案（如有）？
- 期望的功能：参与者希望底层技术具备哪些能力，以便更好地支持其需求？
- 痛点： 他们在开发、部署和维护方面面临的最明显的技术或组织挑战是什么？
- 解决痛点这些痛点是否已得到解决？参与者是否会考虑将切换到新平台作为解决方案的一部分？

哪些技术、成本或组织障碍会阻碍此类转变？

在以下章节中，我们将根据参与者对这些问题的回答呈现研究发现。不过，在继续之前，需要指出，我们在呈现数据时对主题的编排，并不意味着这些主题在不同领域中的相对重要性。例如，某些组织明确讨论了成本，而其他组织没有，这并不意味着成本无关紧要或被忽视——它可能只是在那些访谈中没有成为主要主题。此外，由于各种原因（主要是个人或业务相关原因），参与者无需回答所有问题。因此，任何未作答的回答均标记为未知，并排除在分析之外。

IV. 当前实践的痛点

在访谈中，参与者描述了采用和扩缩容基于云的解决方案时遇到的各种挑战。这些挑战涉及性能瓶颈、DevOps 局限性、基础设施权衡（尤其是在混合云和边缘云环境中）。根据参与者的回答，我们将报告的问题归为六个不同类别。表 II 总结了这些痛点类别及其代表性示例、明确描述各项问题的参与者总数和按领域划分的百分比，以及这些问题在访谈时是否已解决。下文重点介绍主要发现和趋势。

大多数痛点源于与技术相关的类别：运维可维护性 是报告频率最高的痛点，影响了57%的参与者。这个问题通常与缺乏专职技术人员、云服务提供商支持有限，以及配置和管理云服务的复杂性有关。因此，排名第二和第三的挑战性痛点分别是开发复杂性 （48%）和响应能力 （43%），这也就不足为奇了；要解决这些问题，同样需要扎实的技术专长。另一方面，成本 似乎不太受关注，这可能看起来有些违背直觉。这或许是因为部分受访者，尤其是担任技术或运维岗位的受访者，由于很少参与预算编制或财务决策，可能没有特别强调这一问题。

非技术参与者面临的挑战明显多于技术参与者。在所有领域中，遭遇痛点的非技术参与者比例普遍高于技术参与者。值得注意的是，在开发复杂性 和安全与隐私方面，差距分别为 4× 和 3× 。这一趋势表明，缺乏技术背景会严重阻碍云服务的有效采用和使用。

大多数痛点仍未解决 如表 II所示，在访谈时，所报告的痛点中已完全解决的不到一半（不包括因商业敏感性而未予答复的情况）。尤其是，可靠性与可用性 方面的关切大多由缺乏专职 DevOps 或 SRE 团队的组织报告。

特征	数量	百分比
组织类型		
技术型组织	8	38%
非科技组织	13	62%
业务领域		
教育	3	14%
金融	3	14%
医疗保健	2	10%
工业（石油与天然气、制造业）	3	14%
IoT	3	14%
农业	2	10%
技术咨询	2	10%
网络安全	1	5%
芯片/硬件	2	10%
社交网络	1	5%
在用的基础设施类型（可多选）		
基础设施即服务（IaaS）	13	62%
平台即服务（PaaS）	11	52%
函数即服务（FaaS／无服务器）	10	48%
边缘计算	8	38%
混合云（边缘 +云）	6	29%
多云	4	19%
私有云	3	14%
人工智能专用云	2	10%
组织规模		
初创企业/小型企业（<50 名员工）	4	19%
中型企业（50-500 名员工）	4	19%
大型企业（>500 名员工）	7	33%
未指定/未知	6	29%

表 I：参与者人口统计信息与技术采用（N=21）

类别	痛点	业务领域	已解决
Cost	云服务使用成本高，且 为优化资源利用率付出巨大努力	非技术(40%) 技术(60%) 总计(24%)	部分解决(20%) 未知(80%)
运维 可维护性	缺少专职技术人员， 云服务提供商支持延迟，缺少云 对遗留程序的支持，且 基于云的业务服务配置复杂	非技术(67%) 技术(33%) 总计(57%)	已解决(33%) 部分解决(33%) 未知(33%)
开发 复杂性	缺乏专职技术人员，分布式 边缘云系统，以及混合数据流 （从本地部署到云）	非技术人员（80%） 技术(20%) 总计(48%)	已解决(10%) 部分解决(30%) 未解决(20%) 未知(40%)
响应能力	数据库瓶颈以及缺乏架构 支持数据局部性	非技术(56%) 技术(44%) 总计(43%)	已解决(33%) 未知(67%)
可靠性与 可用性	没有站点可靠性工程人员或全职 DevOps，边缘系统中的网络连接不可靠，且系统 连接，且系统 配置错误	非技术(57%) 技术(43%) 总计(33%)	已解决(0%) 部分解决(57%) 未知(43%)
安全与隐私	在不可靠网络下进行边缘身份验证， 处理敏感数据以及政府法规 （例如，PDPA）	非技术(75%) 技术(25%) 总计(38%)	已解决(37.5%) 部分解决(25%) 未知(37.5%)

表 II：跨基础设施与业务领域的痛点

这些组织均未报告该问题已完全解决，不过有些组织已实施了部分缓解措施。

A. 访谈示例

为具体说明表格中所识别的痛点II，我们根据访谈数据提炼出具有代表性的案例。这些案例展示了多项挑战如何在真实部署中交织出现。

案例 1：医疗保健初创公司——运维与可靠性挑战

一家小型医疗保健技术公司（员工不足50人）运营着一个结合无服务器和传统基础设施的云平台。该组织报告称，在不同移动平台上的通知交付可靠性方面存在持续挑战，计划外的数据库维护窗口导致服务中断，以及频繁的网络连接问题。由于没有专职DevOps 人员，事件响应需要数小时的手动日志分析才能识别受影响的服务。尽管该组织通过数据库冗余和自动化监控解决了部分问题，但缺乏全职运维人员这一根本性资源限制仍未解决。

此案例体现了可靠性与可用性（运维人员配置不足）、运维可维护性（依赖云服务提供商支持的响应能力

不足），以及响应能力（数据库可用性缺口）之间的交集。这些挑战是资源受限的小型企业在没有专门的站点可靠性工程团队的情况下尝试维护生产系统时的典型特征。

案例 2：工业领域——边缘云混合架构复杂性

一家工业自动化公司部署了混合边缘云架构，用于远程设备监控，但在网络连接不稳定的分布式环境中面临着特有的挑战。该公司需要能够在长时间网络中断期间正常运行的身份验证机制、即使网络连接受限也能立即撤销凭证的能力，以及防范远程地点的物理设备遭入侵。此外，每台设备每秒产生约 1 MB 的海量遥测数据，因此在发生网络分区时，数据同步和重复数据删除都面临挑战。

这则案例凸显了安全与隐私方面的顾虑（离线身份验证和物理安全）、可靠性 方面的问题（网络连接时断时续），以及开发复杂性（无冲突数据同步）。标准云平台缺乏原生支持，无法提供能够适应长时间断连的身份验证系统，也无法满足这些应用场景所需的复杂数据同步机制。

案例 3：教育技术——无服务器可扩展性缺口

某教育平台利用无服务器计算提供个性化学生交互，但当函数级自动扩缩未能扩展至数据库层时，便遇到了架构限制。尽管计算资源会根据需求自动扩缩容，但在使用高峰期，数据库吞吐量成了瓶颈。该平台实施了自定义同步技术和序列化策略来绕过这些限制，大幅增加了开发工作量。无服务器环境缺乏内存处理支持，也进一步加剧了状态管理需求的复杂性。

此案例说明了计算层与数据层扩缩容能力不匹配所带来的响应能力挑战。采用无服务器架构的组织往往必须设计定制解决方案，以弥补云平台抽象无法提供端到端可扩展性保证的不足。

案例 4：金融机构——混合数据驻留要求

一家大型金融服务机构实行严格的数据驻留政策，所有敏感数据均保留在本地，仅以加密形式传输至公有云环境进行处理。尽管这种混合架构满足了监管要求，但由于容器编排平台和无服务器函数环境中的配置参数不断增多，带来了巨大的开发复杂性。该机构的组织结构进一步加剧了运维难题，因为开发团队承担了 DevOps职责，却没有专门的运维支持人员

。

尽管该机构的技术相当先进，混合云架构带来的运维开销仍高于纯云部署。这表明，监管合规要求可能会迫使组织采用增加而非降低系统复杂性的架构模式。

案例 5：政府教育——政策与遗留约束

一家使用私有云基础设施的政府教育机构，因数据隐私法规要求进行繁复的认证流程，而在采用公有云方面面临障碍。除了合规认证成本高昂之外，组织结构也带来了额外挑战：基础设施团队与应用开发团队各自独立运作，沟通有限，反映出遗留组织边界的影响。此外，合同采购周期限制了技术迁移决策，因为只有在硬件合同到期后才能进行平台变更，无论其技术价值如何。

此案例凸显了成本 方面的顾虑（监管合规带来的额外负担）、组织约束 （团队孤岛）以及运维可维护性 方面的挑战（基础设施与应用集成存在缺口）。僵化的采购周期表明，在公共部门情境中，非技术因素可能主导平台采用周期。

这些案例表明，痛点很少孤立出现。组织通常会面临多个类别中相互交叠的挑战，而资源约束、监管要求和架构决策会进一步加剧云原生采用的复杂性。

总结（痛点）： 报告中的大多数痛点仍未解决，尤其是缺乏技术专长和专职支持团队的非科技组织。

V. 对云提供商/平台的期望

表 III 总结了参与者认为云平台或其他底层技术必须满足的关键期望，以解决当前痛点，并在未来更好地支持应用开发、部署和管理。我们将这些期望归为五个功能组，每组代表云平台应提供的一项能力（见期望 列）。例如，可用性和可靠性 体现了平台在发生故障或灾难（例如大规模中断）时维持应用平稳运行的能力。对于每个类别，我们报告了表达此类支持需求的非技术、技术及所有参与者的百分比。主要发现如下。

技术参与者优先考虑服务质量支持。大多数技术参与者强调了性能相关保障的重要性，包括可用性和可靠性（67%）、安全与隐私（67%）以及服务质量保障 （57%）。这反映出，他们的解决方案通常是数字服务。这些服务高度依赖底层技术的性能（例如，正常运行时间、吞吐量和延迟），以提供高质量服务

类别	期望	非技术	Tech	总计
可用性与可靠性	高可用性（最低 99.9%）、快速联系支持团队，以及完善明确的灾难恢复机制	67%	33%	43%
安全与隐私	强大且有韧性的安全系统，并采用数据加密	67%	33%	29%
服务质量保障	提供高服务等级协议保障，并具备高级且丰富的可配置性，以支持性能优化、系统可扩展性以及相应能力	57%	43%	33%
统一可维护性	支持混合云，并提供自动化技术支持（DevOps）和日志聚合	43%	57%	33%
可编程性	开发便捷性和用户友好型界面	50%	50%	38%

表 III：受访者期望与基础设施和业务领域的对应关系

并保持良好的用户体验，这对用户留存和业务成功至关重要。

非技术参与者优先考虑简化性和易用性。相比之下，非技术参与者更经常强调简化平台复杂性的需求。如前几节所述，缺乏技术专长是这些用户面临的主要障碍。因此，能够降低运维复杂性或支持无代码/最少工作量开发的功能，例如统一可维护性（57%）和可编程性（50%），在这一群体中被提及的频率最高。

各业务领域都需要广泛的支持。汇总技术参与者和非技术参与者的回答后，我们发现五个类别的期望分布相对均衡。这表明人们强烈需要全面且统一的平台支持。这说明，未来的云解决方案不仅应提供高性能，还应确保可访问性和易用性，以服务广泛的应用和用户。

要点（期望）：未来的云平台必须提供服务质量和简便性，以有效支持技术用户和非技术用户。

VI. 迁移准备度

接下来，我们考察参与者是否愿意将其解决方案迁移到新的底层技术或云平台（例如，从基于容器的服务 [15]，[16] 转向函数即服务），以彻底解决其痛点或提升解决方案的质量。这些问题源于云技术的快速发展以及服务抽象缺乏标准化。这些因素使服务迁移成为一项重大工程——它可能带来长期改进，但也可能造成短期干扰或风险。在询问迁移准备度时，参与者的反应不一。尽管总体而言，参与者对迁移持开放态度，但如表IV所示，实际采用仍受到五项关键障碍的制约。

许多人都愿意进行迁移。在所有报告的问题中，只有少数参与者反对迁移。最常被提出的担忧是技术难度，有38%的参与者提到这一点。其他所有问题的提及率均低于30%。值得注意的是，内部阻力（例如担心造成 disruption 或裁员）很少被认定为主要阻碍，仅有5%的参与者提及。这些发现反映出，人们普遍认识到迁移能够带来明确的长期收益，例如性能提升、成本优化以及获得成熟服务，这些收益往往超过短期成本和投入。

非技术参与者对迁移更加犹豫。例如，非技术参与者提出成本相关顾虑的比例是技术参与者的两倍。内部阻力的情况更为突出，所有反对意见都仅来自非技术参与者。这一趋势与此前的发现一致：缺乏深厚技术背景的非技术参与者，在面对可能带来颠覆性变更的情况时往往更厌恶风险。相比之下，技术参与者通常更有信心评估利弊，因此在长期收益足以证明其合理性时，也更愿意推进迁移。

要点（迁移准备度）：大多数参与者愿意考虑迁移，但非技术用户仍持谨慎态度，因为他们担心成本、复杂性以及迁移可能造成的中断。

VII. 讨论

A. 全面洞察

我们在第 IV、V 和第 VI 节对数据表进行了详细解读并总结了要点，涵盖痛点、期望和迁移方面的顾虑。基于这些证据，我们提出两项重要结论，以完善并拓展原有假设（第 II 节）。

技术进步背景下的复杂性尽管云技术不断发展，访谈清楚地表明，云用户——尤其是非技术领域的用户——在实际应用中仍难以应对复杂性。许多现代云范式非但没有简化开发和部署流程，反而引入了新的抽象层，需要更多专业知识并带来更大的运维开销。这些发现印证了原有假设：当前的云原生实践仍然复杂；但结论还不止于此：复杂性并非仅仅源于技术尚不成熟，它往往根植于现代云平台本身的设计之中。这一发现凸显了迫切需要一种统一、简化的云基础设施方案，以减少碎片化并降低采用技术的门槛。

将生产力和服务质量置于成本和迁移之上与人们普遍认为从业者在采用新的业务赋能技术时主要考虑成本和迁移问题不同，参与者揭示了另一番情况。在不同的角色和领域中，成本都不是首要障碍

类别	迁移问题	非技术	Tech	总计
Cost	难以估算迁移成本且前期成本高 迁移成本	67%	33%	14%
技术 难度	对当前云平台的软件依赖，以及有限的 迁移技术熟练度，以及迁移期间不间断的 SLA 合规性	75%	25%	38%
组织 约束	与现有平台的合同义务相关的约束，以及 当前平台以及与数据政策相关的法律约束	67%	33%	14%
开发人员投入	大量手动迁移工作，耗时的 迁移过程，以及可能需要重写代码	60%	40%	24%
内部 阻力	对现有平台更加熟悉，以及员工 对采用新解决方案感到不安	100%	0%	5%

表 IV：各业务领域的迁移障碍

迁移也不是首要障碍。相反，他们关注的是服务质量（QoS）、简化的可编程性和更高的可访问性等期望。事实上，许多参与者表示，如果平台能够显著改善服务质量、运维简便性和开发速度，他们愿意承受短期的迁移挑战或更高的前期成本。这一发现完善了我们最初的假设：高成本虽然是解决方案生命周期复杂所带来的结果，但其根本原因更为深层，源于当今云生态系统中的生产力差距和架构复杂性。

B. 特定领域的技术采用模式

对访谈数据的分析揭示了不同业务领域各异的技术采用模式，反映出行业特定需求和约束如何影响基础设施选择。

受监管行业青睐混合云和私有云架构

在受到严格监管的领域，尤其是金融和医疗保健领域，组织明显更倾向于采用混合云或私有云部署。接受访谈的金融行业参与者共 3 位，三者均采用混合架构，将敏感数据保留在本地部署环境中，同时在公有云环境中进行计算。案例 4（金融行业）生动地体现了这一模式：参与者明确指出，本地数据存储与加密云处理之间存在严格的隔离。同样，政府附属教育机构（案例 5）指出，隐私法规（例如，符合个人数据保护法（PDPA））是采用公有云的重大阻碍。

这一模式反映出对数据主权、数据驻留和合规要求的严格规定，而标准公有云服务难以轻易满足这些要求。因此，组织必须在混合云边界内管理数据迁移、加密流水线和合规验证，进而加重开发复杂性和运维可维护性负担。

边缘计算集中于物联网和工业应用

边缘计算的采用主要集中在这样的领域：与数据源的物理距离和网络不可靠性是关键考量因素。在物联网参与者中（共 3 家），三家都部署了边缘基础设施。同样，工业领域的参与者（石油与天然气、制造业：共 3 家）均采用边缘部署，且通常采用边缘云混合配置。

代表性用例包括：农业中的牲畜监测系统，由于农场地处偏远、连接不便，需要进行本地数据采集；石油钻井平台监测会产生每秒 1 MB 的遥测数据，但蜂窝网络不可靠（案例 2）；以及物联网应用中的可穿戴空气质量传感器，这类应用优先考虑实时响应能力和数据隐私。这些场景面临着一些共同挑战：网络连接不可靠、对延迟敏感的处理要求，以及云连接不可用时实现平稳降级的需求。

值得注意的是，75%（8 个中的 6 个）边缘部署采用的是边缘云混合架构，而非纯边缘解决方案，这表明边缘计算通常是对云基础设施的补充，而非替代。混合方式兼顾了本地响应能力与基于云的分析、存储和管理能力。

无服务器技术采用偏向技术型组织

在技术/非技术分类中，函数即服务和无服务器技术的采用情况并不均衡。在技术型组织中（共 8 个），62.5%（5 位参与者）采用了无服务器计算。相比之下，非科技组织中只有 38%（13 个中的 5 个）采用了函数即服务。

这种差异很可能源于无服务器复杂性：尽管函数即服务承诺实现基础设施抽象，但也带来了诸多概念层面的挑战，包括无状态函数设计、冷启动延迟、编排以及分布式工作流的调试。技术能力较强的组织能够应对这些权衡，而非技术组织往往缺乏有效利用无服务器范式所需的专业知识。

有趣的是，尽管面临监管限制，金融行业的无服务器技术采用率仍然很高，3 位金融行业受访者中有 2 位使用函数即服务处理特定工作负载（例如实时广告生成、事务处理）。这表明，只要架构设计得当，无服务器技术就能与严格的合规要求并存。

多云策略仍属小众

尽管业界一直在讨论供应商锁定问题，但只有

19%（21 位参与者中有 4 位）采用了多云策略。这些采用者主要集中在资源充足、能够管理多云部署所带来的运维复杂性的金融和工业领域大型企业中。

多云采用率有限，这表明其理论优势——供应商谈判优势、冗余以及选择最优服务——不及实际问题重要：运维开销增加、技能组合碎片化，以及难以在不同平台间保持安全、监控和部署工作流的一致性。

所有领域中的 IaaS 持续应用

传统 IaaS 在所有领域仍然普遍存在（总体采用率为 62 %），其中也包括那些原本可能更青睐高层抽象的技术型组织。这种持续应用反映了多个因素：与旧版应用的兼容性需求、对底层基础设施控制的需求，以及 IaaS 工具和运维实践的成熟度。

值得注意的是，几位参与者描述了混合云 IaaS/PaaS 或 IaaS/FaaS 架构，其中不同的服务层级适用于不同的工作负载特征。例如，对性能可预测性要求高的应用仍运行在 IaaS 上，而无状态微服务则利用无服务器架构实现自动伸缩。这种模式表明，平台选择正日益取决于具体工作负载，而非整个组织统一决定。

C. 相关工作

多项研究致力于阐明各种业务场景下云原生软件开发所面临的挑战。例如，Oyeniran 等人 [17] 全面评述了云原生技术如何提升软件开发中的可扩展性和弹性——这两项问题在工业界和学术讨论中都备受关注。Alka 等人 [18]研究了创业背景下云原生技术的特征、影响与挑战。他们的研究强调，云原生范式的兴起改变了创业环境并带来了新的机遇，同时也指出了这一领域特有的九项主要挑战。Deng 等人 [19] 系统分析了云原生应用开发中的四个核心研究领域——构建、编排、运维和维护，并指出了贯穿云原生生命周期的关键面向服务的挑战。

尽管这些研究为理解云原生软件开发在各业务领域面临的挑战提供了宝贵见解，但它们主要依赖文献综述和学术假设，缺乏基于从业者经验的实证验证。因此，这些研究忽视了组织准备度、技能短缺、集成复杂性以及特定领域的期望等现实约束，而这些因素往往对实际采用场景起着决定性作用。相比之下，我们的研究基于对技术领域和非技术领域专业人士的直接访谈，从而更贴近实际、更结合具体情境地理解组织面临的痛点、期望和迁移方面的顾虑。

D. 潜在解决方案

为了满足未来云平台的需求——这些平台必须同时提供服务质量和简便性，以支持广泛的用户群体——近期研究日益聚焦于具备 SLA 感知能力的抽象和用户友好的部署模型。

当代边缘云平台通过支持服务等级协议（SLA）的规范来强调服务质量（QoS），这些协议通常以延迟、可用性和成本为衡量指标。这些服务等级协议指导函数即服务函数在异构基础设施上的自动化部署和资源管理 [20], [21]。例如，SLA-ORECS [22] 可动态分配资源，以满足不同的服务级别协议约束，同时抽象掉基础设施决策的复杂性。大量研究通过将服务级别协议驱动的部署和调度构建为优化问题，进一步推动了这一方向，以平衡用户定义的服务质量与成本效益 [23]–[25]。

许多无服务器平台将函数和状态数据整合为统一的部署单元，例如“参与者” [26]或“进程” [27]，，从而简化函数及其相关状态的部署。OaaS 及相关研究 [14], [28]–[30], 处于这一趋势的前沿，利用面向对象原则，在统一的云部署单元中封装函数、状态和非功能性需求。

VIII. 总结

本研究基于对技术和非技术领域专业人士开展的 21 次深度访谈，证实复杂性仍是云原生应用开发的一大障碍。这种复杂性往往根植于现代平台的架构之中。参与者一致强调，提升生产力、服务质量（QoS）和开发者可访问性十分必要，而成本或迁移问题并非主要驱动因素。这些见解挑战了传统假设，并揭示了当前云原生实践与用户实际需求之间的脱节。

总而言之，研究结果凸显了重新思考云基础设施设计方式的迫切需要，重点应放在降低抽象开销并适应不同用户的能力上。未来研究应探索简化运维、感知服务等级协议（SLA）且易于使用的抽象，在服务保证与简单性之间取得平衡。这包括设计端到端模型，以简化整个应用生命周期——从开发和部署到监控与扩缩容。

参考文献

[1] B. M. Harve, D. M. Bidkar, M. S. Krishnappa, G. Pandy, V. Jayaram, P. K. Veerapaneni 和 G. Mehta, 《云原生革命：云驱动世界中的微服务》，载于*2024*年智能控制论技术与应用国际会议 *(ICICyTA)*，2024 年，第 1043–1048 页。[2] Z. Fatima、M. H. Tanveer、Waseemullah、S. Zardari、L. F. Naz、H. Khadim、N. Ahmed 和 M. Tahir, 《借助物联网和工业5.0实现生产工厂与仓库自动化》，应用科学，第 12 卷，第 4 期，第 2053 页，2022 年。

[3] O. T. Modupe、A. A. Otitoola、O. J. Oladapo、O. O. Abiona、O. C. Oyeniran、A. O. Adewusi、A. M. Komolafe 和 A. Obijuru，《回顾边缘计算对实时数据处理与分析的变革性影响》，Computer Science & IT Research Journal，第 5 卷，第 3 期，第 603–702 页，2024 年。[4] A. Esashi、P. Lertpongrujikorn、S. Kato 和 M. A. Salehi，《Action engine：FaaS 中的自动工作流生成》，Future Generation Computer Systems (FGCS)，第 174 卷，第 107947 页，2026 年。[5] 亚马逊云科技，《AWS 免费套餐》，2025 年，访问日期：2025-04-21。[在线]。网址：https://aws.amazon.com/free/ [6] VMware, Inc.，《VMware》，https://www.vmware.com/，访问日期：2022 年 7 月 23 日。[7] Cloud Native Foundation，《Kubernetes》，https://kubernetes.io/，在线；访问日期：2023 年 11 月 12 日。[8] 亚马逊，《AWS Lambda |无服务器计算 - 亚马逊云科技》，https://aws.amazon.com/lambda/，在线；访问日期：2025 年 11 月 12 日。[9] G. Cloud，《Cloud Functions | Google Cloud》，https://cloud.google.com/ functions/，在线；访问日期：2023 年 11 月 12 日。[10] 谷歌，《Google Cloud Platform》，https://cloud.google.com/，访问日期：2025-04-21。[11] 微软，《持久实体 - Azure Functions》，https://docs.microsoft.com/en-us/azure/azure-functions/durable/durable-functions-entities，在线；访问日期：2025 年 11 月 12 日。[12] P. Borra，《领先云服务提供商（AWS、Azure 和 GCP）的比较与分析》，International Journal of Advanced Research in Engineering and Technology (IJARET) Volume，第 15 卷，278，2024 年 266 页。[13] GitHub，《GitHub Actions》，https://github.com/features/actions，访问日期：2025 年 4 月 21 日。[14] P. Lertpongrujikorn、H. D. Nguyen 和 M. A. Salehi，《通过稳健封装简化云原生应用的开发与部署》，载于2024 年 ACM 云计算研讨会论文集，2024 年，第 847–865 页。[15] T. Chanikaphon 和 M. A. Salehi，《UMS：在自主计算系统之间实时迁移容器化服务》，载于GLOBECOM 2023 - 2023 IEEE 全球通信大会，2023 年，第 467–472 页。[16] S. Manatura、T. Chanikaphon、C. Chantrapornchai 和 M. A. Salehi，《FastMig：利用 FastFreeze 在云 2.0 中实现稳健的服务流动性》，载于第 17 届 IEEE 云计算国际会议 (CLOUD) 论文集，2024 年，第 81–90 页。[17] O. C. Oyeniran、O. T. Modupe、A. A. Otitoola、O. O. Abiona、A. O. Adewusi、O. J. Oladapo 等，《利用云原生技术提升软件开发可扩展性与弹性的综合综述》，International Journal of Science and Research Archive，第 11 卷，第 2 期，第 330–337 页，2024 年。

[18] T. Alka、A. Sreenivasan和M. Suresh，“创业策略助力可持续增长：深入探讨云原生技术及其应用”，*Future Business Journal*，第11卷，第1期，第14页，2025。[19] S. Deng、H. Zhao、B. Huang、C. Zhang、F. Chen、Y. Deng、J. Yin、S. Dustdar和A. Y. Zomaya，“云原生计算：从服务视角的综述”，*Proceedings of the IEEE*，第112卷，第1期，第12–46页，2024。[20] M. Goudarzi、M. Palaniswami和R. Buyya，“边缘与雾计算环境中的物联网应用调度：分类体系与未来方向”，*ACM Computing Surveys*，第55卷，第7期，第1–41页，2022。[21] H. D. Nguyen和A. A. Chien，“Storm-rts：面向多云与云边的性能稳定流处理”，收录于*2023*年*IEEE*第*16*届云计算国际会议（*CLOUD*）。IEEE，2023，第45–57页。[22] S. Lan、Z. Duan、S. Lu、B. Tan、S. Chen、Y. Liang和S. Chen，“SLA-ORECS：一种面向服务级别协议的框架，用于边缘云系统中的资源再分配”，*Journal of Cloud Computing*，第13卷，第1期，第18页，2024。[23] G. R. Russo、D. Ferrarelli、D. Pasquali、V. Cardellini和F. L. Presti，“云到边缘连续体中无服务器函数的感知服务质量的卸载策略”，*Future Generation Computer Systems*，第156卷，第1–15页，2024。[24] L. Baresi、D. Y. X. Hu、G. Quattrocchi和L. Terracciano，“Neptune：一种在边缘管理无服务器函数的综合框架”，*ACM Transactions on Autonomous and Adaptive Systems*，第19卷，第1期，第1–32页，2024。[25] J. Jarachanthan、L. Chen、F. Xu和B. Li，“Astrea：面向成本效益与服务质量感知的自动无服务器分析”，*IEEETransactions on Parallel and Distributed Systems*，第33卷，第12期，第3833–3849页，2022。[26] J. Spenger、P. Carbone和P. Haller，“面向无服务器计算的类Actor编程模型综述”，收录于*Active Object Languages: Current Research Trends*。Springer，2024，第123–146页。[27] M. Copik、A. Calotoiu、G. Rethy、R. B¨ohringer、R. Bruno和T. Hoefler，“进程即服务：以无服务器进程统一弹性云与有状态云”，收录于*2024* 年 *ACM* 云计算研讨会论文集，2024，第223–242页。[28] P. Lertpongrujikorn和M. A. Salehi，“对象即服务（OaaS）：在无服务器云中实现对象抽象”，收录于*2023*年*IEEE*第*16*届云计算国际会议（*CLOUD*）。IEEE，2023，第238–248页。[29] P. Lertpongrujikorn和M. Amini Salehi，“对象即服务：通过无服务器对象抽象简化云原生开发”，*IEEE Transactions on Computers*，第75卷，第2期，第423–434页，2026。[30] P. Lertpongrujikorn和M. A. Salehi，“教程：对象即服务（OaaS）无服务器云计算范式”，收录于第*44*届分布式计算系统国际会议研讨会论文集（*ICDCSW*）。IEEE，2024，第5–8页。