



超大用户如何改变高科技 供应商市场



Hari Kannan 和 Christopher Thomas (唐睿思)

据估计，过去三年超大用户对数据中心的投资达到了 1850 亿美元，仅 2017 年就约有 750 亿美元。亚马逊、微软、谷歌、苹果和脸书占到这一巨大资本支出的近 70% (约 500 亿美元) [1]。超大用户所占据的网络流量达到整个数据中心网络流量的近 1/3。思科认为，未来三年，该比例将跃升至 55% 以上，而这些用户所控制的数据中心服务器将达到总数的 50% 以上[2]。

2017 年，数据中心的相关支出增加了 20%。整个供应商市场和 IT 价值链均已感受到了这一爆炸式增长所带来的影响。这对数据中心的存储、服务器、网络、电力、冷却和外围设备等元器件供应商来说将极为有利可图，但也并非一帆风顺。供应商在处理许多超大客户提出的集中且不定期的新需求时，遇到了很多重大难题。为满足他们的需求，成功的供应商已极大程度地改变了他们的研发、制造和销售方式。

这些成功的供应商学会了如何在新旧客户之间达到微妙的平衡——在满足现有合作伙伴，即原始设备制造商 (OEM) 的需求的同时，建立全新的长期超大用户客户群。我们将在本文对此进行探讨，深入理解超大用户、超大用户对供应商的影响，以及企业赢得超大用户所需的能力。

近几年，超大用户公有云和数据中心的收入取得了前所未有的增长。例如，微软云 (Microsoft Azure) 2018 财年第一季度的收入同比飙升约 90% [3]，亚马逊网络服务 (Amazon Web Services) 2018 年的收入同比

攀升 40%以上[4]。IDC 认为，公有云服务的相关支出仍将以每年 20%以上的速度增长，2021 年将达到 2770 亿美元[5]，这意味着，超大用户的数据中心基础设施将持续扩张。

如前所述，麦肯锡估计排名前五的超大用户 2017 年的资本支出超过了 500 亿美元（图 1）[6]，年均增速 20%以上，而且这一显著增长势头未来几年仍将持续。虽然数据中心的支出并不公开，但据我们估算，这些资本支出大部分（特别是增加的部分）都用在了数据中心及相关基础设施上，包括服务器、存储、网络、内存、专用集成电路（ASIC）和现场可编程门阵列（FPGA）等加速器、设施、电力和散热。

图1 超大用户正大力投入资本，尤其是对数据中心的投入

2017年超大用户的资本支出，
10亿美元



¹复合年增长率

来源：年报；新闻稿；Synergy Research Group

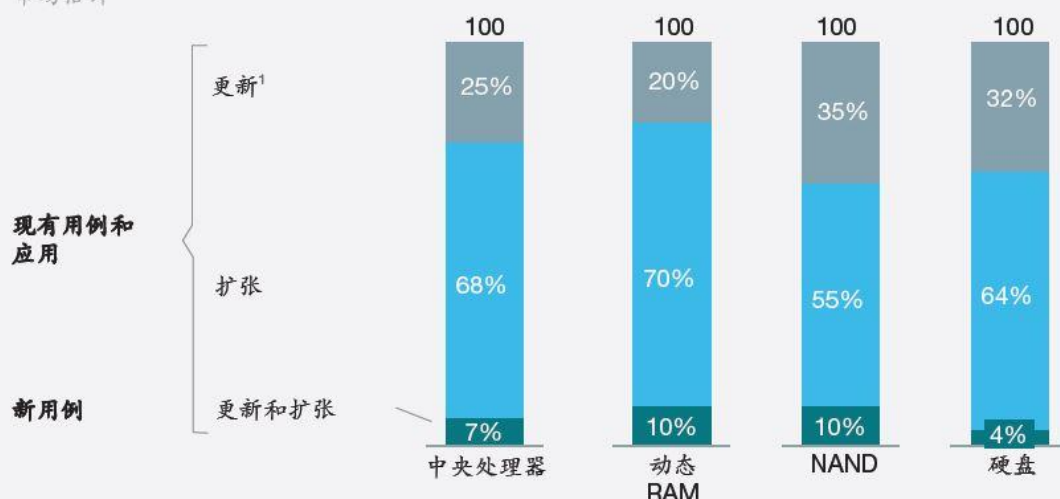
公有云技术栈的数据中心基础设施很容易被视为终端客户服务的一部分。但超大用户的核心业务（比如搜索、社交、电子商务，以及语音、视

觉、增强和虚拟现实等新拓展的平台能力) 也对数据中心的持续扩张提出了要求, 尤其是对处理能力的投资, 比如图形处理器 (GPU)、中央处理器 (CPU)、DRAM 和 NAND 等芯片以及包括磁存储和闪存在内的存储驱动器。

规划和部署这些数据中心是一项复杂且繁重的任务, 超大用户也不例外 [7]。极端地说, 超大用户的需求反映了现有设备的更新需求以及地域和能力的扩张需求。我们估计, 现有用例和应用占这一需求的绝大部分(图 2)。

图2 超大用户的绝大部分需求来自现有用例和应用

2017年超大用户需求, 按更新和扩张、新老用例细分
市场估计



¹包括亚马逊、Alphabet、脸书、微软、阿里巴巴、腾讯和百度

虽然不同的超大用户、不同地区、不同的元器件和不同类型的终端应用、设备和基础设施的具体占比存在差异, 但与现有设备更新相关的采购需求占总需求的比例不足五分之一。其余的需求都来自扩张。外部因素如新 CPU 更新周期等的的能力逐步变化、3D NAND 等的的能力变化、现有用例

的渗透水平、当年新用例的推出数量等等，会导致每年的更新和扩张率出现波动，进一步增加复杂性，因此需求规划者必须对其可能产生的影响了然于心。

这些变量也会对基础设施的选择带来很大的不确定性：超大用户必须做出许多与部署密切相关的决定，比如选择 CPU 还是 GPU。估算不同类型元器件的需求还必须了解不同工作量和用例的增长速度，各类型服务器的需求亦是如此（图 3）。新用例的出现和市场渗透率的巨大起伏可带来强大的产能，鉴于此，采用基于场景概率的方法预测需求就变得势在必行。一般而言，长期预测和近期需求之间存在很大差异。

图3 各种类型的服务器令超大用户的处理能力规划变得愈加复杂

关键服务器类型和样本用例

类型	一般用途	高性能计算	存储优化	记忆优化	GPU驱动
描述	• 使用最广；通常是公有云的起点	• 处理能力最大	• 经优化，可快速移动大量数据	• 经优化，可用于大量内存速度访问的读写操作	• 配有高性能GPU和配套硬件
优秀用例	• 开发和测试工作负载	• 数值模拟	• 高吞吐量内容	• 内存分析 • 内存数据库 • 大规模索引和缓存	• 强力渲染，仿真 • 机器学习、人工智能和深度学习 • 增强和虚拟现实
典型配置	• 中等性能CPU • 每核约3-4 gb • 中高存储，混合固态和硬盘驱动器	• 高性能CPU • 最大RAM • 特定应用的集成电路，用于特定的工作负载	• 中等性能CPU • 高动态RAM • 兼具快速和慢速存储，取决于工作负载	• 高性能CPU • 最大记忆 • 高性能存储	• 高性能CPU和GPU，最大限度作用于核 • 最大动态RAM • 高性能固态硬盘
能耗和散热强度	• 中	• 高	• 低	• 中	• 非常高
网络通速率需求	• 中	• 低	• 高	• 高	• 高

第二个复杂因素来自自上而下和自下而上的需求驱动因素的相互作用。虽然收益构成是各生产线经理自上而下估计的，但配套数据中心的增速却一定是自下而上或基于局部产生的。即便是扩大数据中心或更新设备的需求规划和部署也基于局部。将各个数据中心的需求汇聚到一起，才能进行集中采购，统一管理众多不同的全球供应商。供应链系统应能确保世界各地的数据中心都能获得合适的设备和部件。但是，因自下而上和自上而下的估算经常出现不匹配，因此额外的需求-预测改进周期或许也十分必要。

供应商对订单的反应时间因此变得更少了。

最后一个复杂因素是创新周期。超大用户持续为客户开发新的用例并不断提升现有服务效率。他们通常也是最早使用新技术的人，而且很多情况下，他们会引入定制元器件来保证用例实施。因此，他们能满足一定比例的内部需求，但同时也加剧了数据中心规划、部署和运营的复杂程度。

供应商习惯于与 OEM 供应商合作，再服务于 IT 客户。事实上，终端客户要求和需求预测通常由 OEM 供应商负责。但超大用户的大部分需求都并没有涉及 OEM，而是直接由超大客户与产品供应商联系。超大用户正快速成长为利润丰厚的终端客户，并快速扩张（图 4），在这种情况下，成为他们的供应商有助于各供应商企业获得巨大增长和成功。但是要在这一新的客户环境中顺利前行，供应商必须适应新的市场态势，投资正确的能力。以下是他们面临的一些共同挑战。

预览已结束，完整报告链接和二维码如下：

https://www.yunbaogao.cn/report/index/report?reportId=1_33781

