

技术简报

架顶式网络的业务价值

日期：2010年6月 作者：首席分析师 Jon Oltsik

摘要：随着企业数据中心规模和重要性的增长，部署的最小单元已从单一的服务器变为刀片架，现在已演变至整个机架。为支持这种向机架级部署的转变，预配置机架和横向扩展的网络架构已成为必需。架顶式网络设计具有标准、灵活以及成本低的优势，可以逐步地对容量进行横向扩展以及实现配置和变更管理的自动化。IBM BNT的架顶式网络交换机能够按需进行网络的横向扩展，满足公有云及私有云网络的快速部署需求。

概述

在 10Gb 融合增强型以太网（CEE）出台之前，标准的企业架构采用的都是多个独立（通常为弹性的）的网络。这些网络通过安装于机架上的接线板和导光管，以线缆连接至数据中心的中央接线区。但是，如重新配置接线则须手动操作，并且线缆有识别标签，缺失或错位连接均会导致系统停机。因为网络流量是按照由边缘至核心来设计的，所有通过一组核心网络交换机（具有一定的扇出比）的所有流量均经过精心计算，以避免拥塞和帧丢失。

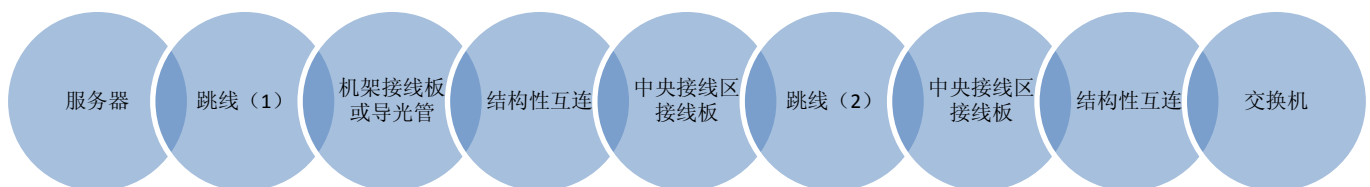
目前，网络架构使用架顶式、一次布线的配置来构建单一的弹性网络。虚拟网络取代了物理网络，并且成对的 MAC 或 WWN 地址能够识别连接而非线缆标签。甚至是基本的布线也得以显著简化—ESG 估计，仅在布线这一项上每台服务器就可能节省 2600 美元。网络管理员使用管理控制台而非跳线和网络布置图来重配网络连接，架顶式交换机取代了接线板和导光管，机架配置成了非常轻松的事情。随着我们逐步向横向扩展的存储和服务器配置过渡，网络流量的主要形式也变为边缘至边缘，而非边缘至核心。因此，核心交换机不仅体积更小，而且更具有经济性。

以服务器为中心—传统网络的物理设计

上个世纪，网络设计一般包括：一个或多个数据网络、一个存储网络、一个管理网络和一个包括键盘、视频、鼠标（KVM）的网络。

每个服务器网络端口通过机架和中央接线区内的一组接线板或光纤导光管连接至交换机。每个服务器网络端口需要连接至正确的交换机才能实现其服务目的。这包括增加至少 2 条跳线：第一条位于机架内，连接服务器至机架接线板（1）；而另一条位于接线区（2）与对应的交换机相连接。

图 1. 传统的服务器至交换机的连接过程



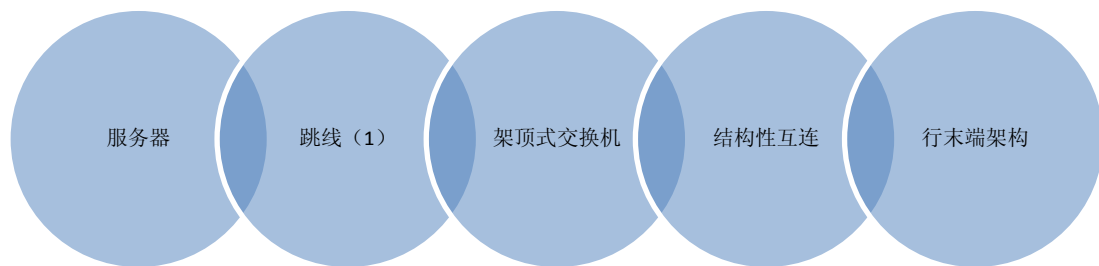
来源：Enterprise Strategy Group, 2010

通常，一台服务器可能需连接六个或更多的网络端口。如果从一个网络切换到另一个网络，则会涉及对数据中心物理访问及跳线的更换。这不仅昂贵、灵活性差、复杂且易出错。此外，故障诊断需要手动进行且过程单调。

以机架为中心—架顶式网络的物理设计

在架顶式设计中，交换机网络安装在每列的末端。在配置机架后，2台交换机安装于每个新机架的顶部，并将线路回接至机架架构的末端（使用10Gb或40Gb结构性互连上行链路）。在机架内安装了服务器或刀片架后，它们通过跳线（1）直接连接至架顶式交换机。

图2. 架顶式服务器至交换机的连接过程



来源：Enterprise Strategy Group, 2010

通常，一台服务器需要2个网络端口进行互连。变更可以通过中央控制台或使用自动系统程序、脚本和模板来进行。由于无人访问数据中心，风险和成本降低，且变更周期也缩至最短。故障诊断变得既简单又准确。高度虚拟化的环境使用hypervisor或其周围的支持系统的功能，可以实现这些变更的全自动化，并将其与负载均衡、自动故障隔离和修复功能进行集成。

什么是融合增强型以太网（CEE）？

融合这种网络方式利用了以太网技术的大量改善功能，这些改善统称为融合增强型以太网（CEE），也叫做数据中心桥接（DCB），包括：

- **基于优先级的流量控制：**旨在确保拥塞网络中某些特定服务级别数据流的零丢失，这是支持以太网光纤通道（FCoE）的必要前提。
- **高性能：**10Gb或40Gb—这是传输大量虚拟化网络流量的关键。
- **拥塞通知：**将流量控制智能移出至边缘网络，并限制传输率，由此增强了类似TCP协议的性能。

所带来的唯一影响就是，CEE可在单一网络上传输所有流量，包括对数据丢失和延迟敏感的光纤通道存储数据流。图3显示的是单一CEE网络上以太网和FCoE的混合架构。

图3. 支持混合架构的 CEE



来源：Enterprise Strategy Group, 2010

网络设计人员可选择在同一线缆上混用基于 IP 的 LAN 流量和 FCoE 或 iSCSI，他们知道 CEE 的高级功能会创建一个低延迟的无损平台，从而确保高可用性、可预测的性能和可靠性。融合设计只需变更软件配置即可允许应用灵活地从 FCoE 迁移至 iSCSI 块存储，或迁移到文件及目标结构性存储，无需移动线缆，也无需用 CEE 网络适配器更换光纤通道 HBA。

运营效率

服务器配置

架顶式设计是“一次布线，即可高枕无忧”的设计。LAN 和 SAN 的配置和路由可以集中进行—或通过控制台手动操作，或更有可能作为整合式虚拟化产品（如 VMware、Microsoft Hyper-V 或 Citrix Xen）的一部分进行操作。实际上，网络越具备“虚拟感知”的功能，服务器配置任务实现自动化的程度就越高。例如，IBM BNT 的 VMready 可以察觉到服务器之间漂移的虚拟机；出现此情况时，虚拟化管理和网络可协同工作以保持诸如 VLAN 或 802.1p 标签等属性。有了这种智能性，大型企业就可利用服务器虚拟化的灵活性和移动性功能，而不需要破坏网络架构或打开安全漏洞。

数据中心的横向扩展

架顶式融合的设计理念是能够让企业数据中心的服务器机架负载和存储实现横向扩展。在每个新机架到位之后，小规模的一组上行链路（通常是以 5:1 的扇出比，即每个上行链路端口有 5 个架顶式端口）连接至每行交换机的末端。而使用 IBM BNT 的基础设施可以彻底改善数据中心网络的经济性。怎么做呢？机架网络的作用是访问和聚合交换机，这就减少了数据中心对更为昂贵的外核/聚合交换机的需求。因为多个网络层在机架内连接，而不是通过数据中心地板下的接线板连接。并且，这种连接方式还有助于削减布线成本。

当机架到位时，交换机也应能及时采购到位，这提高了资本的效率，并减少了未利用端口产生的能耗。

传统的三层交换机的设计，要求在数据中心初建之时就部署整个网络，这极大地影响了资本效率并浪费了更多的能源和热量。由于过早地设计和部署了所有的设备，那些较新的高性能（40GbE 和 100GbE）部件就不能整合到传统的设计之中。

成本分析

众所周知，成本波动的频繁性使得成本比较相当困难。因此，在某种程度上，表 1 中的许多项目（适配器和交换机）会随时间发生变化，并且也会根据区域和市场条件的变化而变动。实际上，在相似的部件之间布线成本的差异相当大—架顶式设计只需要相对来说非常小的一部分。另外，大型企业如果使用更多的架顶式交换机（每个交换机的端口更少），则能够实现显著的容量节省，而无需投资于容量/端口数过量的昂贵的核心交换机。

表 1. 成本模型—（100 台服务器、5 个机架、常规的 4 个 LAN、2 个 SAN、2 个架顶式 CEE）

资金成本	传统设计	架顶式融合	节省
适配器	\$219,000	\$142,900	35%
交换机成本	\$203,960	\$124,000	39%
接线板和线缆成本	\$283,600	\$24,000	92%
总计	\$706,560	\$291,000	59%

来源：Enterprise Strategy Group, 2010

重要事实

如果企业在考虑建造新的数据中心时未仔细评估架顶式设计，将不仅耗费额外的成本，而且还会降低运营的灵活性。数据中心内接线板的安装、配置和管理都是相当复杂的任务，需要投入大量的精力并考虑诸多细节问题。架顶式网络设计不仅减少了人为错误或电力故障的风险，亦能更快地进行变更，而且还降低了成本。

当选择架顶式交换机时，不仅要评估硬件的“速度和投入”，软件功能的评估也是相当重要的。最好的交换机具有易于使用的软件，以便用于配置管理、故障诊断和跨多个物理交换机的集中式管理和控制。并且，软件的功能性可延伸至基本的以太网之外，朝着统一的架构进行扩展。这意味着它将与服务器虚拟化管理和以太网存储进行紧密的集成。

最后，架顶式交换机须适合数据中心的整体架构。既然服务器虚拟化的迅速普及让更多用户继续使用刀片式服务器，基于机架的网络架构就更具意义了。明智的 IT 管理人员会认识到这种趋势，并考虑他们要如何将 IBM BNT 的产品与他们的数据中心战略规划相结合。