

IT@英特尔

为未来而准备的英特尔 数据中心网络架构

将网络架构升级，实现高于 10 GbE 的速度，这对于优化数据中心基础架构，以便更快速响应业务需求，同时显著改进 IT 为整个企业提供的服务和价值来说至关重要。

Sanjay Rungta
英特尔 IT 高级首席工程师

Matt Ammann
英特尔 IT 高级网络工程师

Mohammad Ali
英特尔 IT 高级网络工程师

Kevin Connell
英特尔 IT 高级网络工程师

要点概述

由于数据中心计算和传输的增长，对英特尔网络有更高的要求，英特尔 IT 将数据中心网络架构升级为 10 GB 与 40 GB 连接组合。现有的 100 Mbps 和 1 GbE 连接已经无法支持英特尔业务增长。全新 10/40 GbE 数据中心结构设计可为当下用量提供支持，并可满足未来的网络使用所需；我们预计最终将升级至 100 GbE 结构。

英特尔 IT 正在制定一种垂直化 (verticalization) 的战略来优化数据中心资源，以满足不同计算领域的特定业务需求。由于数据中心在以下方面的趋势，促使我们决定升级：

- 办公、企业和服务¹ 计算环境中的服务器虚拟化和整合
- 设计计算应用及其性能需求的增加
- 互联网连接以每年 40% 的速度快速增长

在设计全新的数据中心网络时，我们对多款 10/40 GbE 连接产品进行了测试，并最终选择了能够提供最佳性能、最稳定和总体拥有成本最低的解决方案。

新的数据中心网络设计提供了诸多优势：

- **降低数据中心的复杂程度。**随着虚拟化使用的增多，10/40 GbE 网络有助于减少所需要的物理服务器和交换机数量。

¹ 英特尔将其数据中心基础架构环境划分为五个不同的垂直业务，它们分别代表主要的商业计算解决方案领域 (称为 DOMES)，包括设计 (Design)、办公 (Office)、制造 (Manufacturing)、企业 (Enterprise) 和服务 (Services)。

目录

- 1 要点概述
- 2 背景
- 3 解决方案
 - 简化办公、企业和服务应用的虚拟化
 - 带宽需求：主要驱动因素、预测和规划
 - 目前可用的物理线缆技术概述
- 6 英特尔使用案例需要特定的解决方案
- 7 为办公、企业和服务 I/O 与存储整合的未来计划
- 8 结论

缩略语

CFP	C 型可插拔
MMF	多模光纤
MPO	多纤连接器 (Multi-fiber Push-On)
MPT	多路直通 (multiplex pass-through)
OM	光多模
OM3	光多模 3
OM4	光多模 4
QSFP+	四通道小型可插拔
SFP+	小型可插拔
SMF	单模光纤
Tbps	每秒兆兆位
µm	微米

- **降低虚拟环境的总体拥有成本。** 10/40 GbE 网络能够将虚拟环境的网络成本降低 18% 到 25%，这主要是因为 LAN 和线缆基础架构得到了简化。而且，新系统需要的数据中心空间、电源和散热资源也更少。
- **增加吞吐率。** 借助更快的连接和缩短的网络延迟，设计工程师能够更快完成工作负载，并显著提高工作效率。
- **提高灵活度。** 能够根据不断变化的业务轻松调整，以满足未来的传输需求，如添加额外的存储容量。

如要优化数据中心基础架构，以便在提升 IT 为企业所提供的服务和价值之际，同时加快对业务需求的响应速度，将网络架构升级，使其支持大于 10 GbE 的速度十分必要。

背景

英特尔 IT 运营在 64 个数据中心内若 55,000 台服务器，为 104,000 位员工提供服务。² 在过去的十多年里，英特尔 IT 一直致力于在办公、企业和服务数据中心环境中实现服务器虚拟化。此外，每周还负责为 4,500 万个计算密集型设计工作负载提供高性能计算。

英特尔快速的业务增长对数据中心资源提出了更高的要求。英特尔 IT 正在部署一种垂直化的战略，以便对每个应用领域进行评估，同时提供相应的技术解决方案来满足特定的业务需求。我们开发了一个办公、企业和服务私有云环境，并准备扩展云计算，以支持制造环境。

这些战略以及计算应用领域中的下列重要趋势，迫使我们现有的 1 GbE 网络基础架构重新评估，以确定其是否可满足网络基础架构的需求。

- 办公、企业和服务计算环境中的大规模虚拟化。
- 提高设计计算领域的计算密度。

此外，高性能的英特尔® 处理器和集群技术可提高文件服务器的性能。也就是说，限制提高吞吐量的因素是网络，而不是文件服务器。我们的互联网连接使用正以每年 40% 的速度增加，所需的连接速度必要比 1 GbE 数据中心网络所能提供的更快捷。

² 请参阅“数据中心战略助力业务转型”白皮书。

解决方案

在 2010 年，我们决定将数据中心网络架构从 100 Mbps 和 1 GbE 连接升级至 10 GbE 连接，以满足这些计算需求。现在又将网络升级为 10 GbE 与 40 GbE 连接的组合。借助此次升级，既能够满足当前的传输所需，又能够整合新技术以满足未来网络要求，例如最终升级至 100 GbE。

简化办公、企业和服务应用的虚拟化

英特尔在办公、企业和服务领域的数据中心战略需要借助虚拟化和整合技术，以便降低数据中心成本和功耗，同时缩短供应服务器的时间。我们目前的整合率为 20:1。借助基于英特尔® 至强® 处理器的新型双路服务器，目标是实现 30:1 的整合率。

通过将网络升级为支持 10 GbE 及更高速度的设计，可以简化服务器连接。例如，将 LAN 端口的数量从数十个 1 GbE 连接升级为两个 10 GbE 连接。这可以显著减少线缆数量并降低基础架构的复杂程度。随着我们迁移至密度更高的 10 GbE 服务器而使用更多交换机互连，这将需要更高的速度，如 40 GbE 和 100 GbE。

此外，存储也要求使用更高的网络速度。目前，并非所有存储工作负载都运行在 IP 网络上；SAN 目前使用光纤通道连接和 HBA。但是，在下一个虚拟化阶段把 SAN 迁移至 IP 网络，这将为以太网带来更多负载。

除了能够简化物理基础架构以外，与 1 GbE 结构相比，10/40 GbE 网络结构可将每台服务器的 LAN 组件的总体拥有成本降低 18.5%。图 1 显示了 10 GbE 结构可节约的成本。预计更高的速度（如 40 GbE）会有类似的节约。

带宽需求：主要驱动因素、预测和规划

如图 2 所示，英特尔 IT 在高速网络技术推出时便第一时间采用，例如，在 2010 年和 2013 年便分别引入 10 GbE 和 40 GbE。促使带宽需求增长的主要因素不尽相同，具体取决于计算环境。在设计计算环境中，存储同比增长约 35%，计算容量同比增长约 30-40%，促使了带宽需求的增加。相比之下，在办公、企业和服务计算环境中，主要驱动因素是：每台主机上虚拟机的数量不断添加，以及虚拟机总体数量增长了 15%。

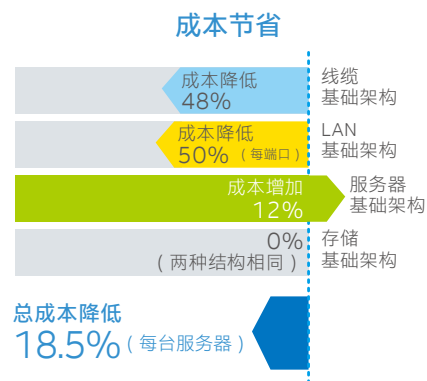


图 1. 基于 10 GbE 结构而可节约的成本。

与 1 GbE 结构相比
总成本降低
18.5%
(每台服务器)



根据当前及预测的使用增长，我们预计到 2015 年数据中心将需要 100 GbE，但是并不准备将所有数据中心端口都转换为高速端口，而是计划在未来三年内，仍然将 1 GbE 和 10 GbE 作为数据中心的主要端口，如表 1 所示。但是，随着向更高的速度迁移，我们必须开发一款物理基础架构，使其将来能够迁移至更高的速度，如 1 Tbps 和 10 Tbps。

目前可用的物理线缆技术概述

大于 10 GbE 的传输速度使用两种方法：并行光纤（parallel optics）或多股铜线。传输性能标准在 2010 年由 IEEE 制定，其中包括 802.3ba-2010 标准。第五号标准描述了基于并行光纤和铜线的配置及其相关性能的要求。

光纤解决方案普遍适用于整个行业且速度更快，因此较之于基于铜线的解决方案更被广泛接受采用。虽然铜线的价格较为低廉，但是由于距离的限制，其应用范围仅限于机柜间布线（intercabinet）连接。IEEE 标准规定 40 GbE 和 100 GbE 传输的最大距离为 7 米。使用并行光纤的光纤解决方案支持更远的距离和更灵活的设计。并行光纤主干由多条包含 12 股线的单模光纤（Single-Mode Fiber，简称 SMF）或多模光纤（Multimode Fiber，简称 MMF）（从 12 股至 192 股不等）组成。

OM3 与 OM4 MMF 比较

MMF 的分类基于其模式带宽性能，又称光纤多模（Optical Multimode，简称 OM）评级。ISO 标准中最新添加的评级是 OM4，其于 2009 年获得认可。40 GbE 和 100 GbE 传输的最低要求是使用 OM3 MMF。OM3 和 OM4 MMF 的性能比较如图 3 所示。

英特尔目前一般使用 OM3 MMF 50 pm，如图 4 所示。英特尔在 SMF 8.2 pm 实施方面始终坚持行业标准。但是，如果要实现 40 GbE 和 100 GbE 最远距离的部署，我们更倾向于使用 OM4。如欲获得更高的性能，40 GbE 和更高的速度需要 OM4 并行光纤。

上市日期和英特尔的采用日期

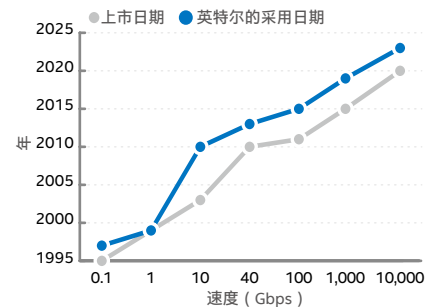


图 2. 每当有更高速度的网络技术上市时，英特尔 IT 马上就采用。

表 1. 预计每年端口速度分布

	2013	2014	2015	2016
100 Mbps 和 1 GbE	77%	71%	62%	58%
10 GbE	23%	28%	32%	31%
40 GbE	0.07%	0.71%	5.40%	10.49%
100 GbE	0.00%	0.00%	0.07%	0.33%

性能对比

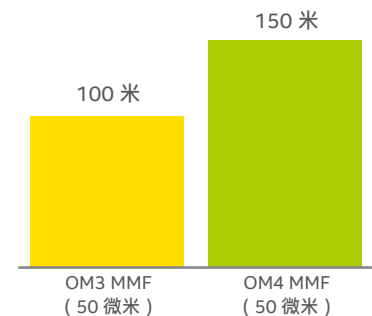


图 3. OM3 和 OM4 MMF 并行光纤性能比较

多模光纤线缆技术：行业标准和英特尔采用情况

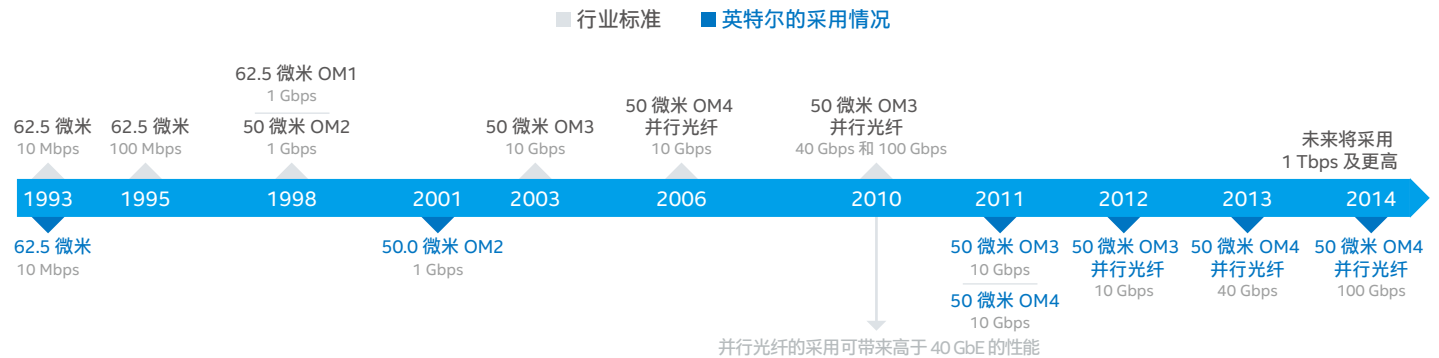


图 4. 英特尔采用线缆技术的情况与行业标准之比较。40 GbE 及更高的速度需要 OM4 并行光纤。

40 GbE 终端

在有源设备（active equipment）中的 40 GbE 连接在终端使用四通道小型可插拔（Quad Small-Form-Factor Pluggable，简称 QSFP+）收发器，以便接收多纤连接器/多路直通（Multifiber Push-On/Multiplex Pass-Through，简称 MPO/MPT）主干。针对近距离的 QSFP+ 收发器使用多模 MPO 主干。当使用多股 MMF 实施 40 GbE 交换机 — 交换机互联时，还需要考虑其极性（polarity）。要正常使用链接，这些连接的极性必须为 Method B。

远距离 QSFP+ 收发器也可在 SMF 上运行。该收发器中包含额外的电子设备，可多路传输/多路解码（mux/de-mux）并行信号 — 即将信号作为一对 SMF 来接收。这些链接使用 Little Connector 端接，最长支持 10 km，可用于 40 GbE 连接互联构建。

QSFP+ 接收器还可用于 MMF 40 GbE 到 4 个 10 GbE 的分区应用。连接的一端使用 MPO/MPT 配置端接，另一端包括四对 QSFP+ 收发器的单独通道，分别使用 Little Connector 端接。该配置可通过 OM4 MMF 支持长达 150 米的距离。这种连接可用于支持信道化 10 GbE 上行链接的分配，支持其向上传输至核心层交换机，并可尽量提高在访问层中可添加的 10 GbE 客户端的数量。

100 GbE 终端

100 GbE 连接使用 C 型可插拔（C Form-Factor Pluggable，简称 CFP）收发器。该收发器家族支持 100 GbE，相当于 QSFP+ 收发器的 40 GbE。目前有多种光纤接口选项可用，但是都还没有成为行业标准。目前业内主要使用两种 CFP：CFP2 和 CFP4。它们的主要区别在于物理密度和传输/接收通道配置。CFP2 支持 100GBASE-SR10、100BASE-LR4 和 100GBASE-ER4 光纤接口。CFP4 在线卡上的端口密度是 CFP2 的两倍，且支持 100GBASE-SR4、100GBASE-LR4 和 100GBASE-ER4 光纤接口。

英特尔使用案例需要特定的解决方案

除了 10 GbE 以外，英特尔还积极部署 40 GbE 连接；一般使用交换机 — 交换机互联和 40 GbE 端口分区来支持服务器连接。我们在特定的情形（主要是机柜间交换机 — 交换机连接）下使用了铜线，但大部分 40 GbE 连接使用了光纤。英特尔目前没有部署 100 GbE，但是正在对组件进行评估。从电缆厂（cable plant）的角度来看，持续向使用并行光纤的光纤技术投资，将会促使物理基础架构最终向 100 GbE 及更高速度迁移。

交换机 — 交换机互联

40 GbE 交换机 — 交换机互联使用以下三种方法中的一种。根据具体案例的不同，英特尔 IT 团队使用了所有三种方法。

- **QSFP+ 收发器和 MMF MPO 主干。**该配置必须使用 Method B 极性的 MPO 主干。
- **远距离 QSFP+ 收发器和标准的双股 SMF 连接。**在英特尔，这种配置用于跨数据中心或园区内不同建筑物之间的交换机 — 交换机互联。
- **有源光缆（active optical cable）。**这是一个预端接并行光缆解决方案，其包括一束 12 股 MMF，每端连接一个 QSFP+ 收发器。这种线缆可在长达 100 米的标准长度中使用。在英特尔，这种配置用于 40 GbE 连接，涵盖数据中心内的各行设备。

40 GbE 端口分区

英特尔部署的这种交换机可将 40 GbE 端口逻辑分区为 4 个 10 GbE。这有助于英特尔在最少的占地空间内尽可能提高 10 GbE 端口密度。这种连接需要扇出线缆，其由一束 12 股的 MMF 组成，一端由 MPO 接头端接，分出 4 个 2 股的连接。该线缆可在长达 100 米的标准长度中使用。

10 GbE 线缆技术

更高的传输速度要求我们实施新的线缆技术来优化 10 GbE 基础架构：

- **10GBASE-T。**该连接（无论使用非屏蔽还是屏蔽双绞线）使用 Category 6a 线缆端接时支持超过 100 米（330 英尺）的距离；使用 Category 6 线缆端接时支持 55 米；使用 Category 5e 端接时支持 45 米。我们将范围有限的 10GBASE-T 用于支持机架内的高密度连接。10GBASE-T 拥有一些成本优势，但是其功耗也高于光纤技术。

“我们用了 5 至 6 年的时间
从 1 GbE 迁移至 10 GbE。
从 10 GbE 向多 GbE 的
迁移速度将会更快。”

- Sanjay Rungta
英特尔 IT 高级首席工程师

- **小型可插拔 (Small-Form-Factor Pluggable, 简称 SFP+) 直连线缆。**这些双轴线缆可支持短距离 (最长 7 米) 的 10 GbE 连接。某些供应商生产的效率最高支持 15 米的传输距离。
- **带接头的线缆。**使用该技术能简化布线工作并降低安装成本，因为它支持 SFP+ 端口。使用一根中继线缆便可在最长 90 米的范围内支持 10 GbE，并且提供了 6 个单独的连接。在支持相同密度的情况下，这有助于将所需要的空间减少 66%。中继支持 (terminate on) 不同的选件，有助于打造灵活的系统。此外，我们还使用了 MPO 线缆 — 一种包括多股中继束与中继盒 (multi-strand trunk bundles and cassettes) 的带接头的光纤技术。该技术支持 1 GbE 和 10 GbE 连接，而且只需更换中继盒，即可轻松升级以便支持 40 和 100 GbE 并行光纤连接。在 OM3 MMF 和 SMF 下，10 GbE 架构目前的覆盖范围分别是 300 米和 10 公里。

为了实现 10 GbE 和未来 40 GbE/100 GbE 可支持最远的距离，我们更改了英特尔的光纤标准 — 以便尽可能反映最低的 OM3 MMF 和 OM4 (见图 4) — 并使用了节能的 SFP+ 端口。

在以太网上实现 I/O 融合后，多种流量类型 (局域网、存储以及进程间通信) 可被整合到一个易于使用的网络架构中。

为办公、企业和服务 I/O 与存储整合的未来计划

过去，由于以太网的带宽有限，不能用于某些应用领域，如 I/O、存储和进程间通信。因此，我们使用其它类型的结构来满足用户的高带宽、低延迟与数据包零丢失的需求，如光纤通道。10 GbE 技术可让我们把全部网络的使用都聚合在一个灵活的基础架构上。

数据中心内 I/O 的增加取决于多个因素。首先，向数据中心添加更多的服务器会增加 IOPS，从而相应地增加网络的传输量。另外，随着每一代新的处理器变得更加复杂，与芯片设计相关的数据量也显著增加，这同样会促使对网络的需求提升。最后，采用高达 1 TB 内存的系统日益普遍，这些系统同样需要高速网络来读取、写入和备份大量的数据。

通过升级到 10 GbE 以上的产品，将能够整合办公、企业和服务应用的存储，同时降低 10 GbE 的每端口成本。在以太网上实现 I/O 融合后，多种流量类型 (局域网、存储以及 IPC) 可被整合到一个易于使用的网络架构中。我们执行了多阶段的测试，在不远的将来，这些 10 GbE 端口将能够支持多种流量类型。

结论

支持 10 GbE 及更高网络速度的高性能数据中心基础架构能够简化办公、企业和服务应用虚拟化并降低每台服务器的总体拥有成本。

从分析显示，对于虚拟环境来说，10 GbE 基础架构可将我们的网络总体拥有成本最高降低 18% 到 25%。对于需要低延迟的设计应用而言，10 GbE 无需购买昂贵的低延迟技术即可提高吞吐量。我们计划到 2015 年为英特尔的数据中心部署 100 GbE，但是在未来的三年内，大部分端口仍将使用 1 GbE 和 10 GbE。物理基础架构在提升性能方面起到至关重要的作用。新的结构可降低数据中心的复杂程度并令网络更为灵活，从而满足英特尔对数据中心不断增长的需求。

如欲获得更多有关英特尔 IT 最佳实践的信息，请访问：

intel.com/cn/IT

IT@Intel

英特尔 IT 解决了一些当今最为严苛和复杂的技术问题，希望在开放式的对等论坛中与同行 IT 专业人员直接分享这些经验。

我们的目标很简单 — 提高整个组织的效率，增强 IT 投资的业务价值。

如欲了解更多信息，请访问 intel.com/cn/IT 或联系您当地的英特尔代表。

相关内容

请访问 intel.com/cn/IT，查看相关主题的内容：

- 英特尔 IT 白皮书：数据中心战略助力业务转型
- 英特尔 IT 数据中心解决方案：效率提升战略

本文旨在提供一般的信息，并非特定指南。推荐（包括潜在成本节省）全部基于英特尔的体验，仅为预估。英特尔不确认或担保他人会得出类似结果。

本文件中包含关于英特尔产品的信息。本文件不代表英特尔公司或其它机构向任何人明确或隐含地授予任何知识产权。除英特尔产品销售的条款和条件规定的责任外，英特尔不承担任何其他责任。英特尔在此作出免责声明：本文件不构成英特尔关于其产品的使用和/或销售的任何明示或暗示的保证，包括不就其产品的（i）对某一特定用途的适用性、（ii）适销性以及（iii）对任何专利、版权或其他知识产权的侵害的承担任何责任或作出任何担保。

Intel、英特尔、Intel 标识、Xeon、至强、Look Inside、Look Inside 标识是英特尔在美国和/或其他国家的商标。

* 其他的名称和品牌可能是其他所有者的资产。

英特尔公司 © 2014 年版权所有。所有权保留。

♻️ 请注意环保

0714/WWES/KC/PDF

