



# 10GBASE-T

**10GBase-T** 推广了数据中心的万兆应用

作者：  
Carl G. Hansen, Intel  
Carrie Higbie, Siemon



## 10 GbE 万兆以太网：应用驱动

数据中心中为了减少IT成本而不断增长的虚拟化应用使得许多目前使用千兆以太网(1GbE)的管理员都在认真地考虑采用万兆以太网 (10GbE)来减少复杂性。与虚拟相关的服务器合并对网络的I/O输入输出有着很大的影响力，因为此时把几个机器和其他背景服务如现场直播迁移、以太网网络联合到一个机器上了。

随着统一网络的趋势，可以使用单个以太网就能同时支持数据和存储，这就增加了对I/O的要求使得1GbE网络在数据中心中会成为瓶颈而增加复杂性。因此需要重新考虑数据中心网络。由于1GbE 连接可以支持单个流量，在高峰时需要支持多个流量就显得带宽不足。这就需要多个1GbE的连接。

采用万兆以太网(10GbE)就可以解决这些网络问题，它能提供更多的带宽，并通过汇总多个千兆端口到一个万兆连接来简化网络基础设施。数据中心管理员有多种10GbE 接口的选择，包括CX4，SFP+ 光纤，SFP+ DAC直连铜缆，和 10GBASE-T。如今，绝大部分选择了 10GbE 光纤或者SFP+ DAC。然而，这些接口的局限性阻止了它们在数据中心中被广泛采用。

光纤链路的成本太高，SFP+ DAC 又受7米的限制，因此需要一个完整的基础设施的升级。CX4是个较早的技术，不能支持高密度。对于10GBASE-T，以往大家的看法是要求的功耗太大，因此无法被广泛采用。而这些问题已经得到了解决，最新的制造工艺可以极大地减少10GBASE-T的功耗和成本。

只有一个有成本效益的解决方案才能被广泛采用，这个解决方案必须能向下兼容，而且具有极大的灵活性可支持数据中心中绝大部分的交换机和服务器。这本白皮书着重于那些促使选择10GbE的驱动性因素，以及10GBASE-T被广泛采用，包括在服务器底版中的采用的趋势。这本白皮书同时还概述了在数据中心中采用10GBASE-T的优势，包括增加带宽，提高灵活性，简化基础设施，易于升级，和成本节省。

## 10GbE 万兆以太网的需求

多种技术的改进和趋势驱动着数据中心对10GbE 需求的增加。例如，在服务器中多核处理器和多插座平台的广泛采用。这种性能使得用户可以采用单个服务器运行多个应用，这就导致多个应用竞相争用服务器上有限的输入输出资源。同时用户还在使用虚拟化来把多个服务器合并到一个物理上的服务器，来减少设备和电力的支出成本。采用最新的Intel® Xeon® 处理器技术的服务器可以把合并服务器的比例提高到15:1。

不管怎样，服务器的合并和虚拟化使得以前几台服务器的I/O需求，现在只通过单个服务器的网络资源来实现，所以对服务器的网络带宽提出了很高的要求。为了配合这个网络需求，IT已经把每台服务器的千兆以太网的连接数量扩大到了双倍、三倍，甚至四倍。这个模式导致了网络的复杂性增加，因为它要求额外的以太网适配器，网络线缆和交换机端口。

向统一网络的过渡增加了对网络高带宽的需求。IT部门正在通过把iSCSI, NAS, 和 FCoE等局域网和存储区域网络的流量汇聚到单个的以太网数据中心协议的方式，使得在迁移到统一网络时，能够帮助简化网络基础设施。这种把多种流量类型汇聚到单一以太网构造的方式的确简化了网络，但同时极大地增加了网络I/O需求。

继续使用千兆以太网的链路是不太合理的，因为要满足现在和未来的I/O 需求，千兆链路会急剧增加网络的成本和复杂性，如增加的千兆以太网适配器，电力消耗等。若采用10GbE万兆以太网则可以用单口或双口的10G连接来替代多个的千兆链路，这就在增加了带宽的同时又极大地简化了网络，降低了电力消耗。

1 Source: Results have been estimated based on internal Intel analysis and are provided for informational purposes only. Any difference in system hardware or software design or configuration may affect actual performance.

## 10GbE的媒介选择

虽然业界一致认同朝着10GbE万兆以太网的迁移，但是广泛采用10GbE还是受到了一些因素的限制。要理解这个动态的状况，需要调查目前对10GbE媒介选择的赞成和反对的理由。

### 10GbE 的媒介选择

| Technology  | Pros  | Cons  | Cable  |
|---|---|---|--|
| <b>10GBASE-SR</b><br>(SFP+ Fiber)                 | <ul style="list-style-type: none"> <li>• 3.2w / port</li> <li>• Latency: &lt; 1 <math>\mu</math>s</li> <li>• 300 meter reach (SR)</li> <li>• 10km reach (LR)</li> <li>• &lt; \$699 / port</li> </ul>  | <ul style="list-style-type: none"> <li>• Requires new SFP+ switches</li> <li>• High cost of purchase and deployment</li> <li>• Not compatible with installed 1000BASE-T switches</li> <li>• Due to the higher purchase price ongoing maintenance costs are higher</li> </ul>  | <ul style="list-style-type: none"> <li>• Multimode fiber</li> </ul>                  |
| <b>10GBASE-DAC</b><br>(SFP+ Direct-Attach Copper) | <ul style="list-style-type: none"> <li>• 2.9w / port</li> <li>• Latency: &lt; 1 <math>\mu</math>s</li> <li>• &lt; \$399 / port</li> </ul>   | <ul style="list-style-type: none"> <li>• Limited to 7 meters</li> <li>• Not compatible with existing 1000BASE-T switches</li> <li>• Requires new SFP+ switches</li> <li>• Can lead to a significant oversubscription of ports due to limited reach</li> </ul>   | <ul style="list-style-type: none"> <li>• Twinax copper</li> </ul>                    |
| <b>10GBASE-CX4</b>                                | <ul style="list-style-type: none"> <li>• 2.25w / port</li> </ul>  | <ul style="list-style-type: none"> <li>• Limited to 15 meters</li> <li>• Large connector form factor limits high density</li> <li>• Requires 10GBASE-CX4 switches, with few new products available</li> <li>• Not compatible with existing 1000BASE-T</li> <li>• Larger diameter cabling is difficult to manage in legacy cabinets</li> </ul> | <ul style="list-style-type: none"> <li>• Twinax copper</li> </ul>                    |
| <b>10GBASE-T</b>                                  | <ul style="list-style-type: none"> <li>• &lt;\$399 / port</li> <li>• 9.2w / port (&lt;80m)</li> <li>• 10.1w / port (&gt;80m)</li> <li>• CAT6 55m</li> <li>• CAT6a 100m</li> <li>• CAT7 100m</li> <li>• Backward compatibility with 1000BASE-T</li> <li>• Uses existing high performance structured cabling</li> </ul> | <ul style="list-style-type: none"> <li>• Higher power than SFP+</li> <li>• Latency: 2 - 4 <math>\mu</math>s</li> </ul>  | <ul style="list-style-type: none"> <li>• UTP copper</li> <li>• FTP copper</li> </ul> |

目前 IT 经理面对10GbE的挑战是每种选择都有些缺点，要么是成本、电力消耗，或者是支持的距离。

## 10GBASE-CX4

10GBASE-CX4 是个较早的 10GbE 部署方案，然而由于其笨重而昂贵的线缆，以及只有15米的支持距离，其采用受到了限制。CX4 接头的尺寸限制了交换机的高密度，而这在大型项目中是非常需要的。固定长度的大直径线缆导致了更多松驰线缆的管理问题。管道和空间都不够来处理这么大的线缆。

## SFP+

SFP+可以同时支持光纤和DAC直连铜缆使得它比CX4有更好的灵活性。SFP+ 如今在慢慢增长，但有些局限阻止了它被所有的服务器采用。

### 10GBASE-SR (SFP+ 光纤)

光纤在延时和支持距离（超过300米）上是非常棒的，但它非常昂贵。光纤提供了较低的电力消耗，但是其昂贵的电子设备使得在所有数据中心中部署光纤网络是高得买不起的。光电子设备通常是同类铜设备的4-5倍，这也意味着将来的维护费用（通常是基于设备采购的价格），也会贵得多。当采用光纤时，不仅意味着需要光纤的交换机端口，而且也意味着需要光纤的服务器网络接口。

### 10GBASE-SFP+ DAC

DAC直连铜缆相对于光纤来说是个较低成本的选择，但是它只能支持 7 米，而且不能向下与现有的千兆以太网交换机兼容。DAC需要购买一个适配器卡，而且需要一台新的柜顶（ToR）交换机。所需线缆比起结构化布线的铜缆信道更加昂贵，而且不能在现场端接。这使得 DAC 比起 10GBASE-T来说更加昂贵。在底版的局域网接口中采用DAC的可能性很小，因为它没有10GBASE-T的灵活性和支持距离。

## 10GBASE-T

10GBASE-T 提供了最大的灵活性，是个成本最低的媒介类型，并与现有的千兆以太网向下兼容。

### 支持的距离

像所有的以太网实施一样，10GBASE-T 能支持长达100米的链路，在连接数据中心的设备时，给了IT 经理极大的灵活性。有了距离的灵活性，10GBASE-T 可以适应柜顶，列中，或者列末的网络拓扑结构。由于可通过现有的结构化布线系统来工作，10GBASE-T给了 IT 经理最大的灵活性来放置服务器。

采用更高等级的布线产品（6A 类或以上），10GBASE-T 可以在30米之内的信道上运行低功耗模式（这也同样适用于数据中心模式）。这相当于比在更长的100米模式下每个端口有进一步的电力节省。数据中心可以采用 any-to-all 的配线区域来确保信道少于30米从而实现这些节省。

## 向下兼容性

由于 10GBASE-T 与 1000BASE-T 向下兼容，因此可以在数据中心中采用6类、6A类或更高等级的布线系统暂时支持现有的 1GbE 交换机设备，使得 IT 既可保持较低成本，也可在未来有很方便的升级到 10GbE 的路径。

## 功率

10GBASE-T 面临的挑战是早先的物理层接口芯片 (PHYs) 的功耗实在太而不能被广泛采用。同样的事情在千兆以太网发布时也曾经经历过。最初的千兆芯片的功耗是大约 6.5 瓦/每端口。随着处理器的改进，芯片从一代到下一代都得到了改善，以致于现在的千兆端口的功耗已经低于 1瓦/每端口。同样的改进已经发生在 10GBASE-T。针对10GBASE-T 的好消息是这些物理层接口芯片 (PHYs) 已经极大地从最新的制造流程中得益。PHYs 的改进与著名的摩尔定律很相符，更新的处理器技术将持续不断地降低新的10GBASE-T PHYs的功耗和成本。

当 10GBASE-T 适配器在 2008年第一次被采用时，它们需要每端口 25瓦的功率。功率已经在连着几代的更新更小的处理器技术中降低了。新近的 10GBASE-T 适配器只需要每端口 10瓦的功率。进一步的改进将减少更多的功率。在2011之前，功率将降低到每端口 5瓦以下，使得 10GBASE-T 适合于底版集成和高密度的交换机。

## 延时

根据包的大小，1000BASE-T 的延时从亚微秒到12 微妙之间。10GBASE-T 的延时在2微秒到4微秒之间，更窄的延时范围。对于以太网中包的大小在512B 或者更大时，10GBASE-T的整体吞吐量相比1000BASE-T将有很大的优势。在较大的传输包中，10GBASE-T 的延时比起1000BASE-T低过3倍多。只有在对延时非常敏感的应用如 HPC 或者高频交易系统中会注意到一点延时的问题。

10GBASE-T 中增加的2微秒的延时对绝大多数用户来说是不重要的。对绝大多数的企业应用以前能运行在1000BASE-T的延时环境，现在的10GBASE-T 延时只会更好。许多 LAN 产品还特意加了很小的延时以减少电力消耗或者 CPU 的开销。一个常用的缺省时就被激活的 LAN 特色是中断调节，这个特色通常会增加大约100微秒的延时，这就允许中断被合并并极大地减少CPU 负担。多很多用户来说，这个协议总体上带来了一些积极的好处。

## 成本

由于最近三代产品的功耗已经急剧地下降了，所以成本也有一个类似的下降曲线。第一代的 10GBASE-T 适配器的价格是每端口 \$1000。如今第三代的双端口10GBASE-T 适配器的价格每端口已经低于 \$400了。在2011年，10GBASE-T将被设计成底版中的局域网接口 (LOM) 而且会被包括在服务器的价格中。通过利用LOM模块中内含的10GBASE-T，用户将会看到一个显著的节省，不再需要购买更贵的SFP+ DAC 和光纤适配器，而且可以释放一些服务器中的I/O 插槽。



## 数据中心万兆以太网的网络结构选择

如下的表格列出了一些典型的适用于10GbE的数据中心网络结构。  
这个表格清楚地展示了10GBASE-T技术比起另外两个类似的铜缆选项提供更大的设计灵活性。

| Technology            | Data Center Network Architectures | Connectivity  |
|-----------------------|-----------------------------------|---|
| 10GBASE-SR SFP+ Fiber | • Top of Rack (ToR)               | Uplinks from ToR switches to aggregation layer switches |
|                       | • Middle of Row (MoR)             | Inter-cabinet connectivity from servers to MoR switches |
|                       | • End of Row (EoR)                | Inter-cabinet connectivity from servers to EoR switches |
|                       | • Core network                    | Backbone  |
| 10GBASE-SFP+ DAC      | • Top of Rack                     | Intra-cabinet connectivity from servers to ToR switches |
| 10GBASE-CX4           | • Top of Rack                     | Intra-cabinet connectivity from servers to ToR switches |
| 10GBASE-T             | • Top of Rack (ToR)               | Intra-cabinet connectivity from servers to ToR switches |
|                       | • Middle of Row (MoR)             | Inter-cabinet connectivity from servers to MoR switches |
|                       | • End of Row (EoR)                | Inter-cabinet connectivity from servers to EoR switches |

### 10GBASE-T的未来

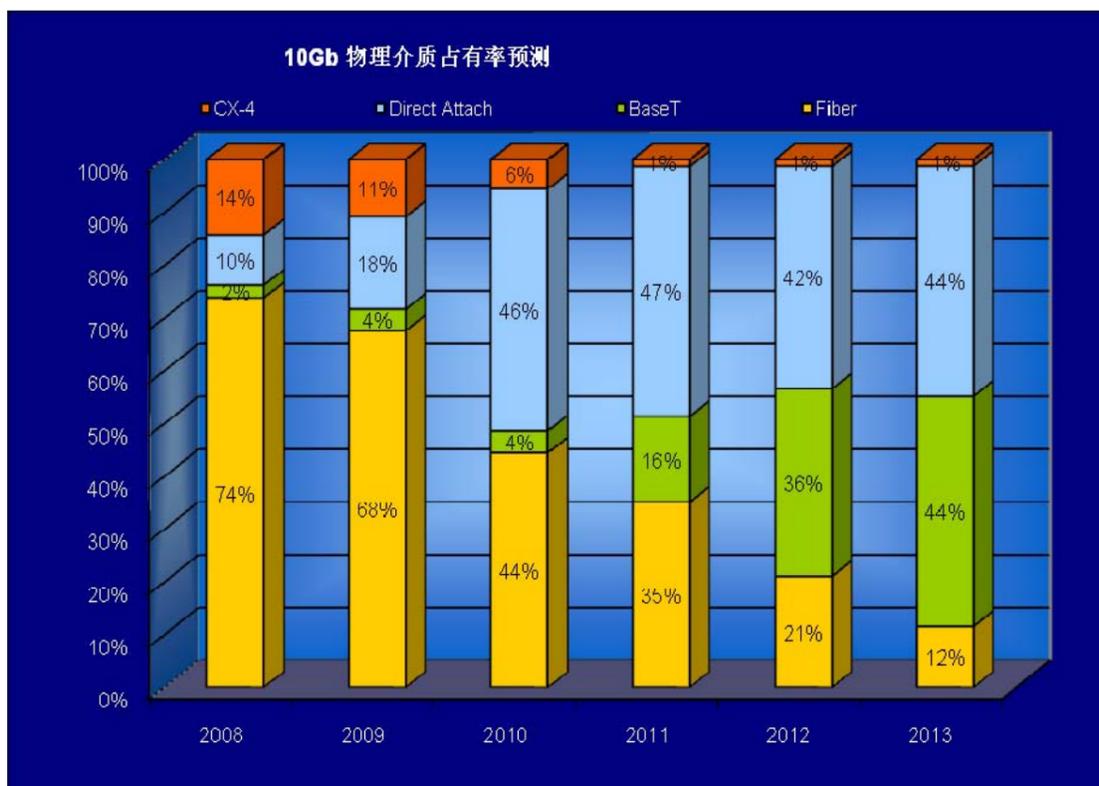
Intel 看见了在 10GbE 中 10GBASE-T 的广泛采用。在 2010 年数据中心中采用光纤来支持 10GbE 的比例是 44%，到了 2013 年，这个比例将会降低到大约 12%。随着大量 IP 数据中心和高性能计算的部署，DAC 直连铜缆连接将会在接下来几年中得到增长，到了 2013 年将会达到 44%。10GBASE-T 将会从 2010 年仅有的 4% 增长到 2013 年的 44%，最终将成为主要的媒介选项。

### 10GBASE-T 作为 LOM

服务器的 OEMs 将会把 10G BASE-T 作为机架式或者塔式服务器中广泛部署 10GbE 的标准化媒介。10GBASE-T 提供了性能和支持距离上的最大灵活性。OEMs 可以创建单个的底版设计来支持在 100 米之内的千兆、万兆以太网或其他的应用。1GBASE-T 是现今数据中心中当之无愧被用得最多的，10GBASE-T 自然将是下一步的首选。

### 结论

10GBASE-T 的广泛部署将简化数据中心的基础设施，在提供高带宽来支持大量虚拟化和 I/O 密集型的应用的同 时，使得服务器连接的管理也更容易。当销量上去时，价格将持续走低，而且新的芯片处理器已经降低了功耗和 热量。这些优势使得 10GBASE-T 非常适合集成在服务器底版中。这个层次的集成，如大家所知的底版中的局域网 接口（LOM）将使得 10GbE 成为数据中心的所有的服务器类型的主流接口。



Source: Intel Market Forecast