

## 面向企业的万兆以太网交换技术概述和应用

技术上的重要进展，以及价格、性能的大幅改进，使得万兆以太网不仅被部署到数据中心，还开始被应用于园区网络。随着带宽需求的增长，以及企业应用总数的增加(详见本文所要介绍的示例)，万兆以太网的部署范围正在加速扩大。

### 简介

自从 IEEE 802.3ae 标准于 2002 年中获得批准以来，万兆以太网端口的售销量已经从每季度几百个端口增加到了每季度几万个端口。万兆以太网的部署量之所以会出现如此快速增长，主要源自于下列因素：

- **万兆以太网每端口价格的大幅降低**——目前，万兆以太网的价格还不到 2002 年中时的五分之一。因此，在智能模块化交换机中，万兆以太网目前的性价比（包括光纤成本）已经与基于光纤的千兆以太网相差无几。

- **新型光纤扩大了万兆以太网的部署范围**——目前，新型光纤的出现让万兆以太网可以部署到从数据中心到配线间的任何环境之中，而且可以沿用现有的光纤布线。

- **带宽持续增长**——首先，千兆以太网到桌面的部署量已经在 2004 年底之前达到了每季度数百万个端口。如此广泛的部署大大提高了网络其他部分的超额使用率。万兆以太网有助于将超额使用率降低到了网络设计最佳实践所要求的水平。其次，服务器适配器和 PIC 总线技术的发展使得服务器能够生成超过 7Gbps 的流量，这提高了为服务器使用万兆以太网连接的需求。最后，新型应用在企业园区、数据中心内部和数据中心之间，都催生了对于万兆以太网性能的要求。以下章节将会详细介绍这些应用。

以上因素预计将会继续推动万兆以太网市场的发展。根据 Dell' Oro Group 的预测，这个市场将会从 2004 年的 3.85 亿美元，增长到 2009 年的 29 亿美元。

本文将介绍万兆以太网与企业网络相关的技术和应用。针对电信运营商的万兆以太网技术和应用（例如 WAN PHY 和接入点 [POP] 内部互联等）不属于本文的讨论范围。

### 技术概述

#### MAC 层属性

因为万兆以太网仍然属于以太网，所以它可以充分利用经过多年发展的以太网技术，简化向这种更高速技术的迁移过程。与此前的快速以太网和千兆以太网一样，万兆以太网采用了 IEEE 802.3 以太网 MAC 协议、以太网帧格式和帧尺寸。它支持标准的以太网服务，例如 802.3ad 链路汇聚，最多可以将 8 个万兆以太网链路汇聚到一个虚拟的 80Gbps 连接上。因为万兆以太网也是全双工的点对点技

术，它可以在不导致数据包冲突的情况下，同时支持来自于链路两端的流量。因此，它不存在固有的距离限制。最大链路距离取决于传输机制和传输介质光纤，而不取决于以太网冲突域的范围大小。

## 物理层属性

### 通用接口的命名规范和工作范围

在某种新型以太网技术面世时，人们首先提出的问题之一就是：“它能传输多远的距离？”与以前的以太网技术一样，万兆以太网的传输距离取决于用户所使用的物理接口类型。由于目前存在多种可供选择的万兆以太网接口，所以需要一种统一的命名规范来区分不同的光纤接口、光纤类型和传输距离（参见表 1）。

万兆以太网的物理层接口通常使用下列命名规范：

- 前缀 = “10GBASE-” = 10Gbps 基带通信
- 首个后缀 = 介质类型或者波长（如果介质类型是光纤的话）
- 第二个后缀 = PHY 编码类型
- 第三个后缀 = 宽波分复用（WDM）波长或者 XAUI 通道个数

表 1. 万兆以太网的常用物理接口命名规范

表 1. 万兆以太网的常用物理接口命名规范			
前缀	首个后缀 = 介质类型或者波长	第二个后缀 = PHY 编码类型	第三个后缀 = WDM 波长或者 XAUI 通道个数
10GBASE -	示例： C = 铜缆 (双轴) S = 短 (850 nm) L = 长 (1310 nm) E = 加长 (1550 nm) Z = 超远加长 (1550 nm)	示例： R = LAN PHY (64B/66B) X = LAN PHY (8B/10B) W = WAN PHY (64B/66B)	示例： 如果省略，缺省值 = 1 (串行) 4 = 4 个 WDM 波长或者 4 个 XAUI 通道

例如，10GBASE-LX4 光传输模块使用一个 1310 纳米 (nm) 的激光束，LAN PHY (8B/10B) 编码，4 个 WDM 波长。10GBASE-SR 光传输模块使用一个串行 850nm 的激光束，LAN PHY (64B/66B) 编码，1 个波长。IEEE 802.3an 任务组计划在 2006 年的稍晚些时候，确定基于双绞线铜缆的万兆以太网 (10GBASE-T) 的标准。

表 2 汇总了可在企业环境中使用的万兆以太网接口所支持的传输范围和介质类型。

表 2. 万兆以太网的传输范围

10GE物理接口	典型部署	基于下列介质的传输范围			
		62.5微米多模光纤 (FDDI等级)	50微米多模光纤 (MMF)	10微米单模光纤 (SMF)	双轴铜缆
10GBASE-CX4	数据中心	—	—	—	15米
10GBASE-SR	数据中心	26米—33米	66米—300米	—	—
10GBASE-LX4	园区或者数据中心	300米	240米—300米	—	—
10GBASE-LR	园区或者城域	—	—	10公里	—
10GBASE-ER	城域	—	—	40公里	—
10GBASE-ZR	城域或者长距离	—	—	80公里	—
DWDM	城域或者长距离	—	—	80公里—32个基于单束SMF的波长	—

在现有的从园区分发层到配线间的光纤布线中，有超过 75%都是 FDDI 级别的（62.5 微米）多模光纤（MMF）。距离要求通常超过 100 米（m）。因此，要在现有的 FDDI 级别 MMF 上为配线间部署万兆以太网，通常需要使用 10GBASE-LX4 光传输模块。

## 外型

万兆以太网可插拔接口具有多种外型，例如 XENPAK、X2 和 XFP。从部署的角度而言，这些外型之间的主要区别在于：1) 某个指定外型所支持的万兆以太网物理接口的宽度；2) 物理尺寸。例如，由于空间限制，XFP 外型目前并不支持 10GBASE-LX4 和 10BASE-CX4 光传输模块。只要链路两端的万兆以太网物理接口类型（例如 10GBASE-LX4 或者 10GBASE-SR）相同，不同类型的接口就能在光传输方面进行交互操作。

## 万兆以太网相对于汇总多条千兆以太网链路的优势

很多网络管理人员都在考虑，是应当汇聚多条千兆以太网链路，还是部署一个万兆以太网链路。与过去一样，这两种方式都各有利弊，需要根据实际情况选择合适的方式。但是，与汇聚多条千兆以太网链路相比，万兆以太网可以提供一些重要的优势：

- **减少光纤使用量**——一条万兆以太网链路所使用的光纤束少于千兆以太网汇聚方式。后者需要为每条千兆以太网链路使用一个光纤束。万兆以太网的这种优势可以降低数据中心的布线复杂度。对于那些可能因为成本原因而无法铺设更多光纤的园区环境，万兆以太网能够更加有效地使用现有的光纤布线。

- **为大型数据流提供更加有力的支持**——由于终端设备的数据包排序要求，汇聚的千兆以太网链路所能支持的链路可能只限于 1Gbps 的数据流。相比之下，由于单个万兆以太网链路具有更高的容量，所以万兆以太网能够更加有效地支持那些会产生多 Gb 数据流的应用。

- **更久的部署使用寿命**——万兆以太网能够提供高于多个千兆以太网链路的可扩展性，延长部署的使用寿命。最多可以将 8 个万兆以太网链路汇总到一个虚拟的 80Gbps 连接。

## 万兆以太网的企业应用场景

如前所述，万兆以太网现在可以部署在从数据中心到配线间上行链路的已有光纤布线之上（如图 1 所示）。随着终端设备连接带宽的增加，万兆以太网部署还可以继续拓展到网络核心之外，从而提升网络的可扩展性。例如，千兆以太网到桌面的部署量已经在 2004 年底之前达到了每季度数百万个端口。如此广泛的部署大大提高了配线间上行链路的超额使用率，尤其是考虑到超过 90% 的配线间流量都会以由北往南的方式流经上行链路。

图 2 显示了一个典型的高密度园区配线间的发展历程。在 20 世纪 90 年代后期，常见的做法是为桌面部署 10/100 以太网，再配以冗余千兆以太网上行链路。如果每台交换机连接了 192 个用户，那么超额使用率就约为 19:1。根据标准网络设计最佳实践的要求，配线间带宽超额使用率应该介于 15:1 到 20:1 之间，因此这种做法没有超出规定的范围。但是，随着千兆以太网到桌面近些年来的日益普及，这些超额使用率迅速地攀升到了 48:1，甚至 96:1。即使配线间上行链路已经增加到两个或者四个千兆以太网通道，情况也没有得到改观。通过为当前的交换解决方案部署万兆以太网上行链路，有助于将配线间的超额使用率恢复到网络设计最佳实践所要求的范围之内，并可以根据未来的需要进一步扩展带宽容量。

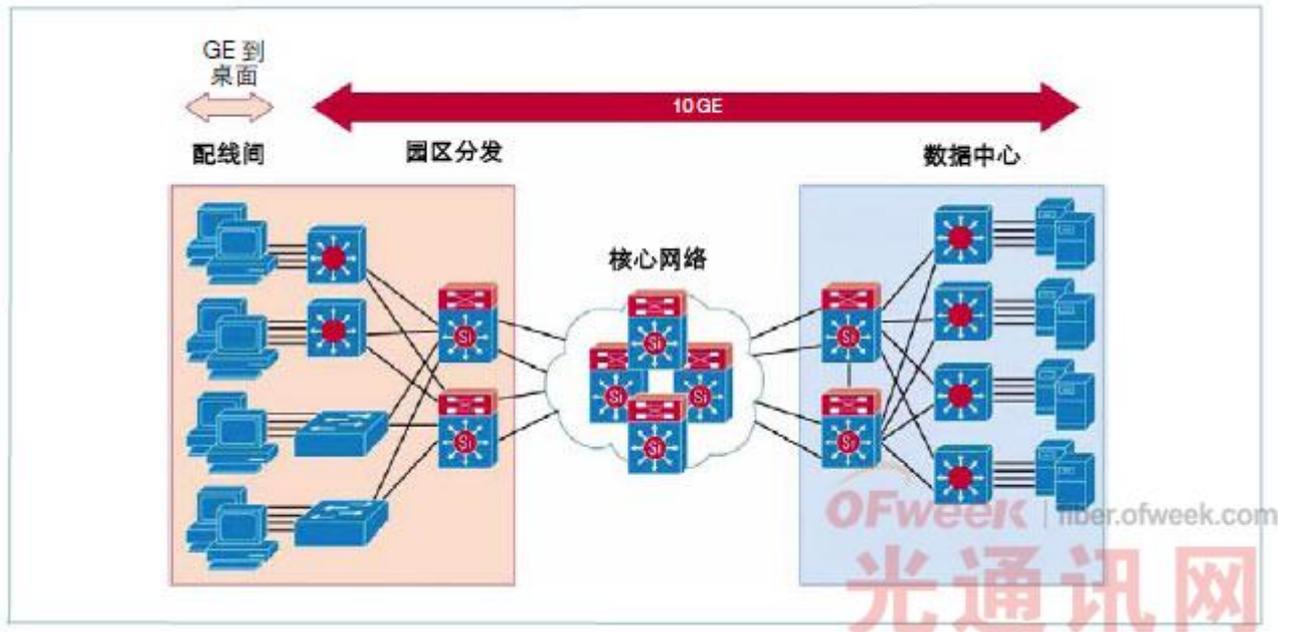


图 1. 整个企业的万兆以太网部署

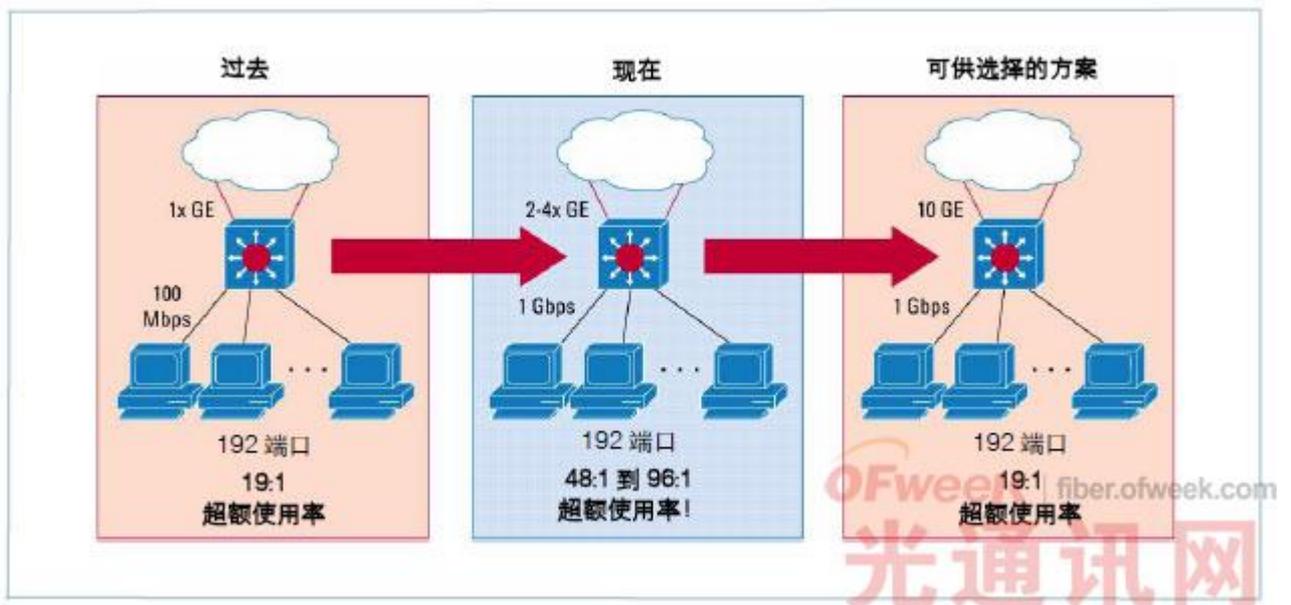


图 2. 通过万兆以太网扩展配线间的上行链路

### 桌面应用

覆盖整个企业的万兆以太网部署可以支持不断增多的桌面应用。这些应用大大提升了企业对于带宽的需求，其中包括：

- **总桌面数据负荷**——由于桌面负荷的不断增长（如图 3 所示）和新型应用对于更高带宽的要求，每个桌面的总带宽需求也在持续增加。例如，由于越来越多的员工依赖于最新的 PC 数据，PC 备份应用显得尤为重要。通过自动一一而

不是由用户一一执行备份任务，可以减少数据丢失，提高备份频率。对一个企业中的所有桌面进行频繁的 PC 备份会给网络带来沉重的负担，尤其是考虑到文件尺寸不断增长（例如 Microsoft Outlook 数据文件和 PowerPoint 演示）所带来的影响。另外，企业正在从传统的客户端/服务器应用（即在每个桌面上使用臃肿的、专用的客户端）转向基于 Web 的应用（即在每个桌面上使用简便的标准浏览器），以利用 Web 技术在运营、开发方面所具备的成本节约优势。但是，与专用客户端相比，浏览器需要更多地与服务器进行通信，才能获得信息并执行处理任务，进而导致了更高的带宽使用量。



图 3. 不断增长的桌面工作负载

来源：Intel 工作负载预测

- IP 视频应用**——很多企业都在部署高带宽的 IP 视频应用，以提高工作效率和降低运营成本。例如，在线学习让员工能够以较低的成本全天候地访问重要的培训信息，获得及时的销售培训、关于如何提供某项服务的快速培训以及课程、技能和法规培训等，从而提高员工的工作效率。企业和管理层 IP 视频通信有助于加强企业员工对于业务目标的共识，提升员工士气。这也是在跨国企业内部促进交流的一种极为有效的方法。IP 视频监控解决方案则被用于提高安全可见度，加快存档资料的检索和分析速度。对于那些需要面对面交流，但是没有时间前往指定地点的员工而言，IP 视频会议可以在他们之间实现有效的协作。所有这些 IP 视频应用都能够产生多个高达数 Gb 的 IP 视频数据流，具体取决于企业希望获得的视频质量。这无疑会占用大量的网络带宽。

- **针对特定行业的应用**——很多行业都有一些需要大量带宽容量和高性能的定制应用。无论这些应用采用的是群集方式，还是基于客户端—服务器模式，万兆以太网都能够迅速地提升网络的性能。例如在医疗行业，数字成像应用（例如图像存档系统[PACS]）通常被用于降低成本，减少获取、分析医疗图像（例如X光、MRI和CAT扫描）的延时，提高医护人员的工作效率。在媒体和广告行业，数字视频应用可以帮助企业有效地制作视频片段，并在分散的团队之间编辑、评审这些视频。在制造行业，越来越多的大型CAD和CAM设计文件需要在身处不同地点的团队之间分享。在金融行业，对于更有价值的、实时的金融信息的持续需求进一步提升了企业对于网络性能的需求。

这些应用示例和其他一些桌面应用的不断发展，都推动了在企业网络中使用万兆以太网的强烈需求。

## 存储网络

在客户服务、消息传递、电子商务、在线多媒体和目录内容等应用的推动下，企业对于存储容量的需求一直在不断增长。这样的“信息爆炸”要求IT经理们找出经济有效地访问、管理和保护这些数据的方法。

要从以服务器为中心的直连式存储转向以网络为中心的共享式存储是实现这些目标的重要策略。通过在数据中心、城域网络和企业内部共享网络化存储，可以带来下列优势：

- 在更大范围内，以共享方式最大限度地使用存储和信息资源
- 简化存储环境的管理
- 最大限度地降低存储的总体拥有成本（TCO）
- 提高数据的可用性和完整性

利用万兆以太网，IT经理现在可以将他们的网络存储环境提升到新的水平，并利用基于以太网的网络支持要求最严格的存储解决方案，例如：

- **通过数据中心备份和灾难恢复提高业务永续性**——为了满足严格的业务要求，企业一直面临着如何开发出经济高效、安全、可扩展的业务连续性和灾难恢复策略的挑战。企业之所以要采用城域存储网络，一个重要的因素是他们需要在远程地点建立备份和远程镜像，以拓展已经达到容量极限的数据中心，或者集中位于多个园区或者地点的数据中心资源。万兆以太网的远程传输能力让企业可以在相距80公里的两地之间提供高速的网络连接。通过使用光学放大器和散射补偿器，传输距离还能进一步延长。因此，企业能够支持在此范围内的多个园区，实现存储到服务器和存储到存储的数据传输。利用万兆以太网和智能交换所提供的高带宽、低延时和安全性，企业可以更加轻松地、在企业存储系统的某些距离遥远的组件之间，实现无缝的数据传输。图4显示了一个可以支持所有基于IP存储的城域解决方案和技术的万兆以太网基础设施，包括网络附加存储（NAS）、

互联网小型计算机系统接口（iSCSI）、基于 IP 的光纤通道（FCIP）和网络数据管理协议（NDMP）。

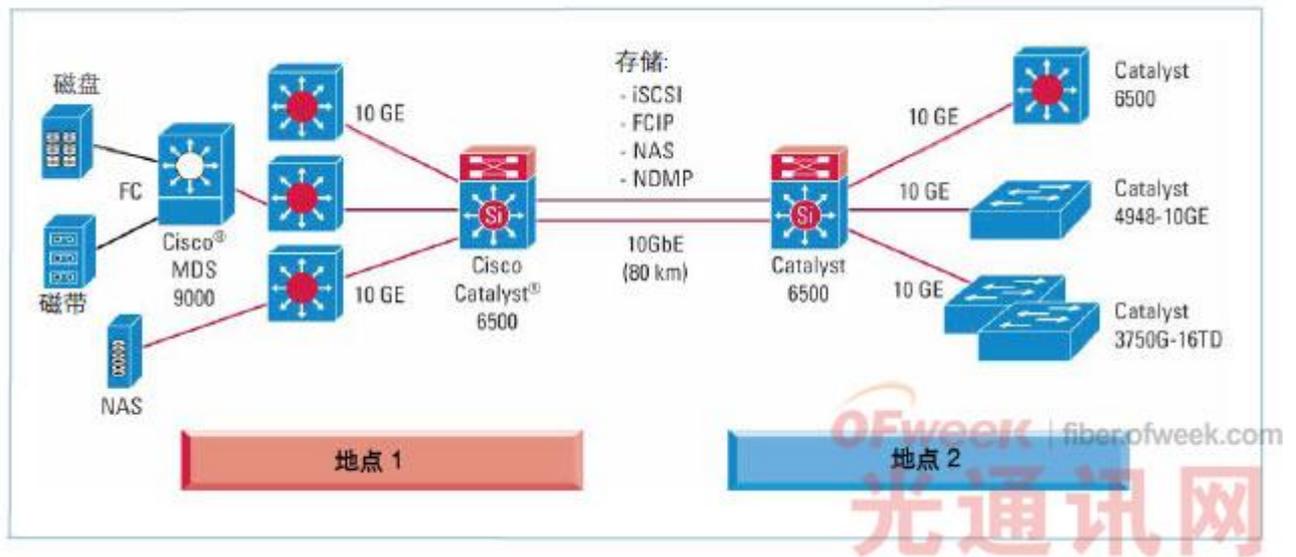


图 4. 用于备份和数据恢复的万兆以太网

对于需要更高带宽的汇聚、更远的传输距离、低延时，以及支持非 IP 技术（例如光纤通道或者 IBM 的企业系统连接[ESCON]协议）的部署，密集波分复用（DWDM）能够在城域网（MAN）中，提供高容量、独立于协议的存储访问和传输功能。这种基于光纤的城域网的关键性存储应用包括：备份、远程镜像、灾难恢复、群集和存储外包。同步镜像需要极低的延时和很高的带宽，而万兆以太网可以提供这些要素的理想组合，满足这种关键任务型的业务需求。

- 用于高性能数据共享和存储整合的网络附加存储（NAS）——NAS 已经成为基于 IP 的存储整合和文件共享的主流部署方式。NAS 在很多环境中得到了广泛的应用，包括协作式工作组开发、工程设计、电子邮件、Web 服务和一般性文件服务。由于 NAS 操作系统的可定制性，NAS 文件服务器能够极为有效地处理 I/O。这使得它很容易就会以线速占满多个千兆以太网通道。因此，企业迫切需要为 NAS 存档整合采用万兆以太网，如图 5 所示。另外，越来越多的企业需要为 NAS 文件服务器建立直接的万兆以太网连接，以支持那些单个数据流超过 1Gbps 的高性能应用。相比之下，802.3ad 链路汇聚无法支持这样的应用。

除了提供对共享文件的高性能访问以外，万兆以太网基础设施还能利用多种协议，例如网络数据管理协议（NDMP），实现服务器之间的复制和备份到磁盘等功能。

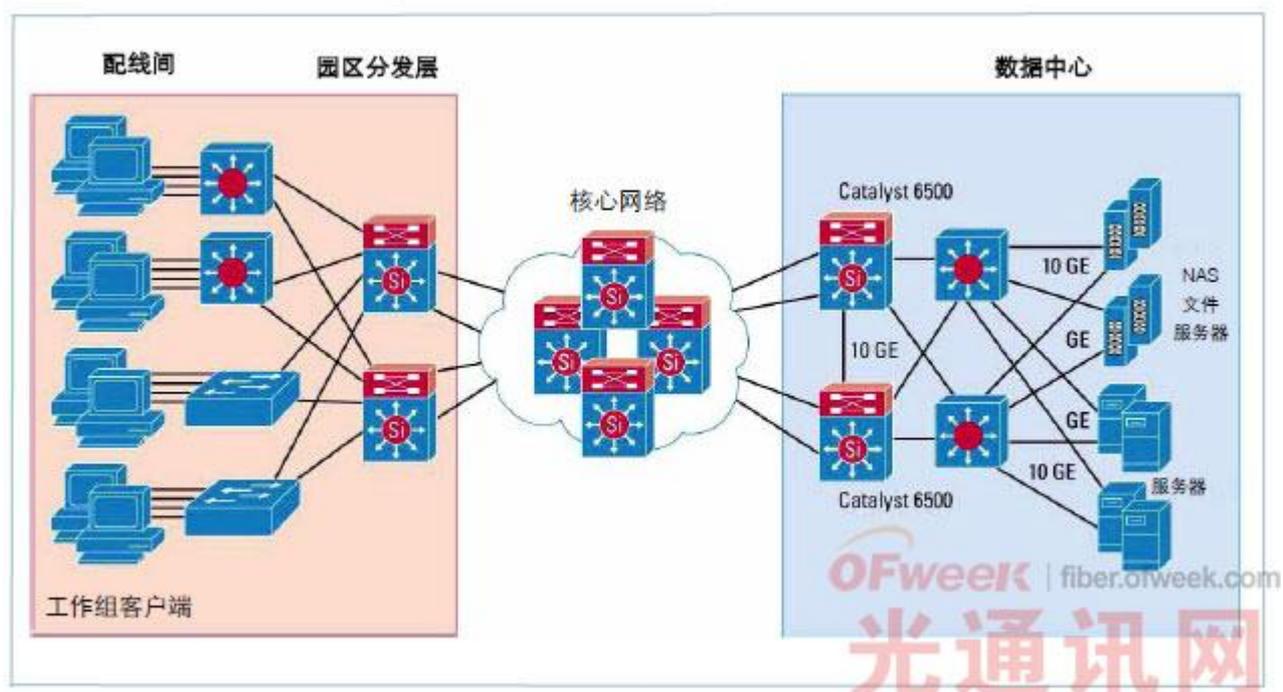


图 5. 用于 NAS 数据共享和整合的万兆以太网

**日益增长的共享存储扇出需求**——随着直连式存储管理成本的不断提高，以及存储子系统容量的日益增长（达到数百 TB 级别），企业迫切需要整合一些过去不被视为存储网络组成部分的系统。要做到这一点，企业面临的挑战主要围绕着将存储网(SAN)扩展到数量有限的高性能节点以外所需要的成本和可扩展性。实践证明，利用经济高效的 iSCSI 协议，在 IP 网络的基础上实现整个企业对存储的访问，是将存储扇出到成百上千台过去不属于存储网络的服务器的有效途径。在图 6 中，园区中支持 iSCSI 的服务器可以通过万兆以太网基础设施和 Cisco MDS 9500 访问数据中心的光纤通道 SAN。Cisco MDS 9500 在这里可以充当一个指向光纤通道存储的 iSCSI 网关。万兆以太网提供了必要的网络可扩展性，以帮助越来越多的分散式设备访问整个企业中的共享存储。

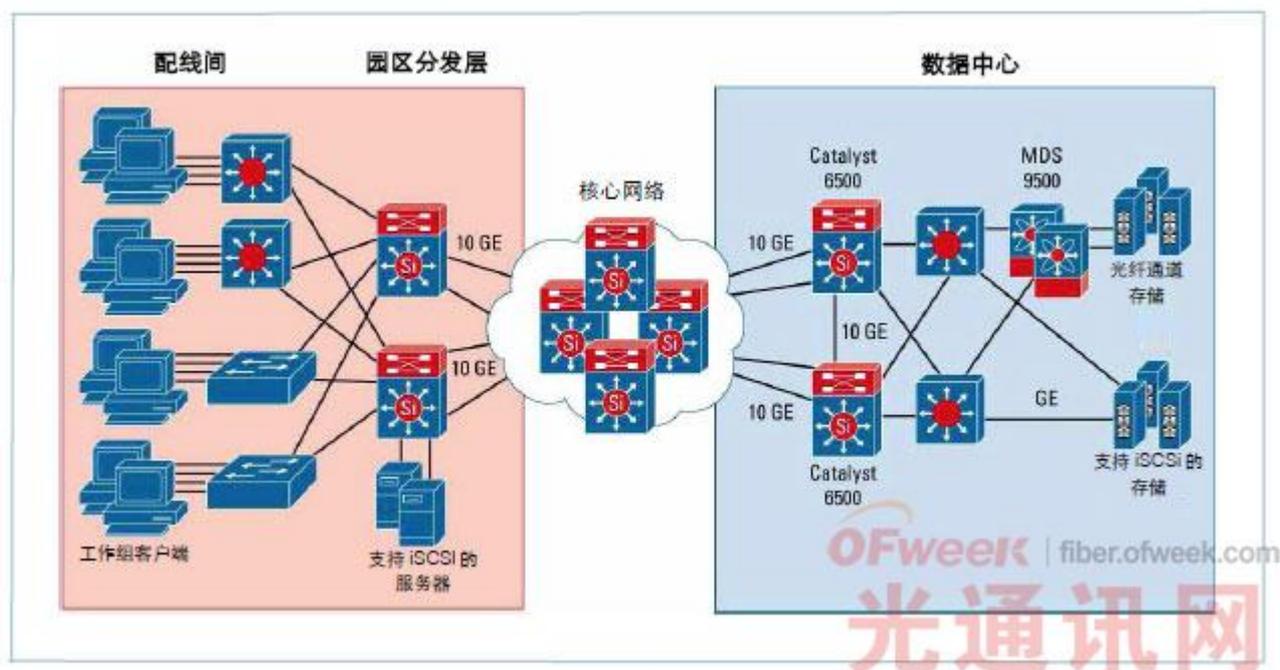


图 6. 用于提高存储整合扇出能力的万兆以太网

### 群集和网络计算

群集和网络计算旨在满足那些需要进行大量 CPU 计算、任务处理和 I/O 传输的应用的要求。这些应用需要多台服务器才能有效地完成工作负载。群集为将计算需求扩展到多台服务器提供了一种经济高效的方法，可以让多个计算节点作为一个庞大的、虚拟的计算节点协同工作。群集应用可能对计算节点之间的互联性能极为敏感，因而对于连接这些节点的网络基础设施提出了很高的要求。因此，通过借助万兆以太网的低延时特性最大限度地提高网络性能，群集应用可以获得有力的支持。为了最大限度地减少服务器延时和 CPU 负荷，企业开始采用一些新颖的服务器端技术，例如系统级 I/O 加速，TCP/IP 卸载引擎（TOE），以及远程直接内存访问（RDMA）。这些在网络和服务器性能方面取得的重要进展，也能够受益于广泛部署的以太网和 IP 技术在互操作性、管理和投资保护方面提供的优势。

虽然群集计算部署主要供科研机构使用，但是越来越多的商业机构也开始采用这项技术。数据库和应用服务器供应商已经在其产品中加入了群集计算的支持。群集计算还被广泛地应用到了其他一些高性能计算（HPC）应用之中，例如金融分析和建模，石油天然气勘测分析，以及工程建模等。图 7 显示了万兆以太网在一个 HPC 群集内部，以及两个分布式 HPC 群集之间，在 DWDM 的基础上提供高性能的网络连接。

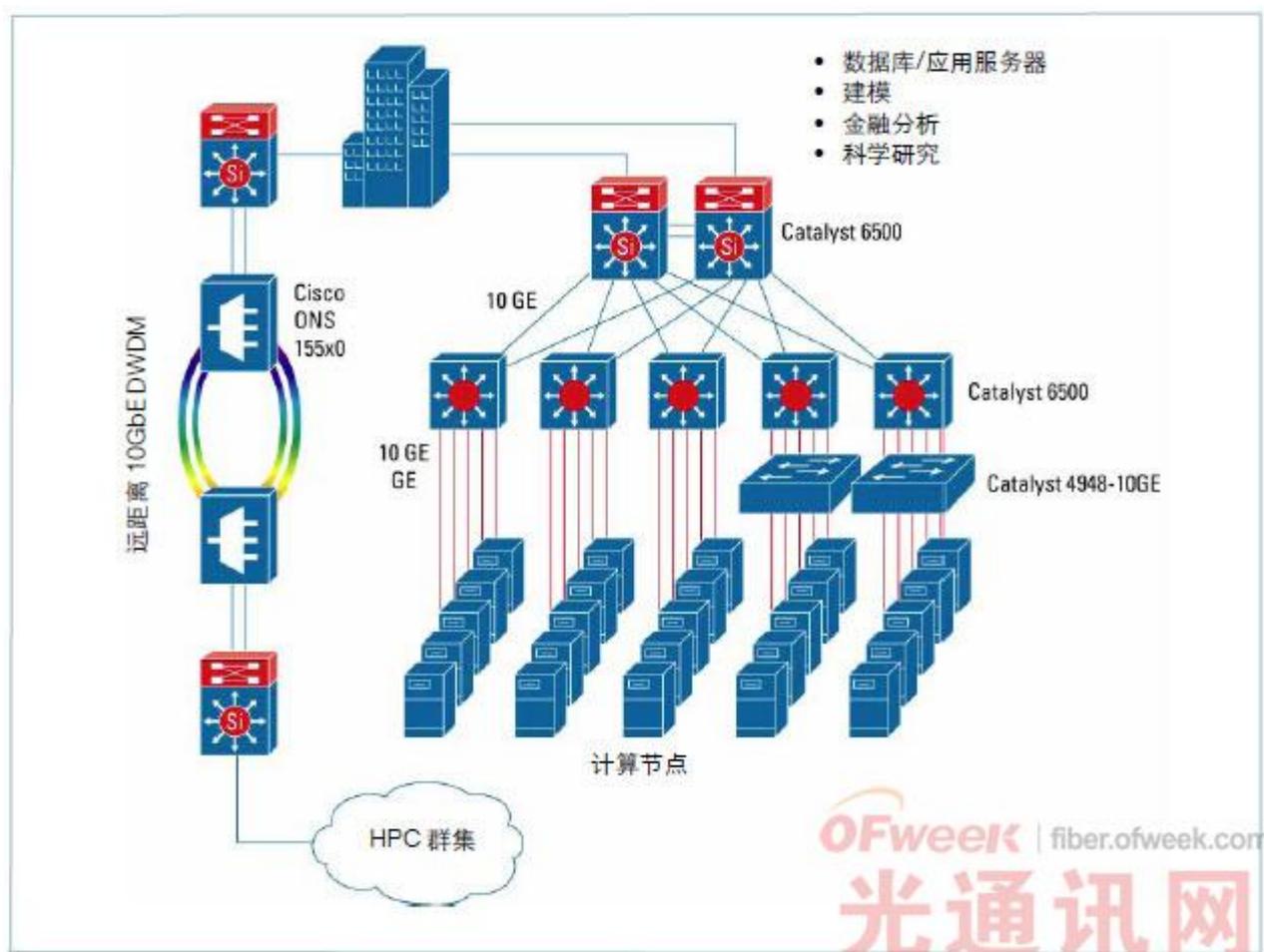


图 7. 用于群集和网络计算的万兆以太网

## 总结

由于价格、性能目标的实现，新型光纤接口对于更广泛部署的支持，以及日益增加的新型应用不断提高对带宽的要求，万兆以太网的部署量正在迅速增长。但是，万兆以太网仅仅是一个范围更广泛的交换解决方案的网络接口。成功的万兆以太网部署还会结合一些领先的智能交换服务，例如集成化安全、高可用性、交付优化和增强的可管理性等，从而为新型应用提供必要的支持。另外，为了最大限度地降低成本，企业在转向万兆以太网的过程中，应当充分地利用对模块、机箱和其他组件的已有交换投资。