

iSCSI 方案设计

简介

随着越来越多的企业从 DAS 向 SAN 环境移植，以及企业对于整合企业存储资源的需求的不断增长，人们迫切希望将整合工作的范围拓展到中端和低端应用服务器。此外，将整合 SAN 的范围拓展到城域网和广域网也成了一种必须达到的要求。思科的 MDS 9000 系列多层控制器 (Director) 和光纤通道交换机可以帮助企业建立大规模的光纤通道 SAN 并将这些 SAN 拓展到中端服务器和城域、广域网络。利用 FCIP 和 iSCSI 协议，企业现在可以利用以太网和 IP 技术进一步拓展他们的存储环境，并不断地通过存储整合节约成本。利用 16 端口无阻塞光纤通道 (FC) 交换模块或者 32 端口共享带宽光纤通道交换模块，企业可以增加它们的存储设备、磁带库和主机总线适配器 (HBA)，在一个交换机中设置多达 224 个端口。由于增加的 IP 存储服务模块可以为 iSCSI 和 FCIP 服务提供 8 个千兆以太网端口，企业可以利用 iSCSI 将 SAN 拓展到其他中低端服务器，或者通过 FCIP 协议，在 IP 网络的基础上连接 SAN。利用 Cisco MDS 9000 系列交换机提供的所有选项，大规模的、高端口密度的 SAN 将成为现实。客户可以利用他们现有的 IP 基础设施和他们自己的 IP 专业技术，优化企业的存储整合。利用 Cisco MDS 9000 系列广泛的多协议管理功能，企业 SAN 的管理也会变得更加简便。

本设计指南将着重介绍如何利用 Cisco MDS 9000 IP 存储服务模块中的 iSCSI 协议拓展 SAN。本文将介绍设计时需要考虑的因素和典型部署，以帮助终端用户了解怎样利用思科的 MDS 9000 IP 存储服务模块部署一个 iSCSI 解决方案。本文将不讨论关于 iSCSI 的 MDS 部署的应用服务器配置。如需查看关于 iSCSI 的 MDS 部署的技术说明，请访问思科连接在线网站，网址是：
<http://www.cisco.com/go/storagenetworking>。

iSCSI 简介

iSCSI 协议的设计目的是使用 TCP/IP 传输 SCSI 协议。从理论上说，iSCSI+TCP+IP 可以为串行光纤通道协议 (FCP) 提供一种类似的传输模式，FCP 也可以传输 SCSI。iSCSI 的基本想法是利用用户对现有 IP 网络的投资，建设和拓展存储局域网 (SAN)。这是通过利用 TCP/IP 协议，在主机或者发起设备 (Initiator) 和存储设备或者目标设备 (Target) (例如存储子系统和磁带设备) 之间传输 SCSI 命令、数据和状态信息。

过去，SAN 需要一个专用的基础设施来互联主机和存储系统。这种互联所采用的最主要的传输协议是光纤通道（FC）。光纤通道网络主要可以为 SCSI 协议提供一个串行传输通道。此外，用户还需要建设 IP 数据传输网络，以支持 IP 应用服务器及其相关存储的前端和后端。

与 IP 不同，光纤通道本身并不能方便地在带宽较低的远程 WAN 网络上传输，因而需要特殊的网关硬件和协议。iSCSI 在 IP 网络上的使用并不一定会取代 FC 网络，但是可以为支持 IP 的主机提供一种传输方式，以访问基于光纤通道的目标(Target)。

IP 网络基础设施可以为服务器和面向数据块的存储设备的互联提供显著的优势。最主要的是，IP 存储网络可以提供明显的成本优势，因为以太网及其相关设备的价格要远远低于基于光纤通道的同类设备。此外，IP 网络可以在一个传统的光纤通道网络上提供增强的安全性、可扩展性、互操作性和网络管理。

IP 网络的优势包括：

- 用于管理、安全和服务质量（QoS）的网络协议和中间件的通用性
- 将在 IP 网络的设计和管理中获得的经验用于 IP 存储局域网。经过培训的、经验丰富的 IP 网络管理人员可以安装和管理这些网络。
- 通过在整个企业中使用一套标准的 IP 基础设施、产品和服务而获得经济效益
- iSCSI 与现有的 IP LAN 和 WAN 基础设施兼容
- 传输距离只受应用性能要求的限制，而不受 IP 协议的限制

iSCSI 的价值

由于 iSCSI 建立在现有的 IP 网络上，所以用户无需添加主机适配器(HBA)就可以将主机连接到存储设备。此外，iSCSI SAN 可以提高存储网络资源的利用率，消除用户对于单独的并行 WAN 和 MAN 基础设施的需要。因为 iSCSI 利用 TCP/IP 作为传输 SCSI 的手段，数据可以在以太网中通过现有的基于 IP 的主机连接传输。由于它能够提高现有的 FC 后端存储资源的利用率，所以能够提供更多的价值。因为主机可以用它们现有的 IP/以太网连接来访问存储组件，存储整合能够以相对较低的成本拓展到中端服务器，同时提高现有存储设备的利用率和可扩展性。

iSCSI 标准的发展

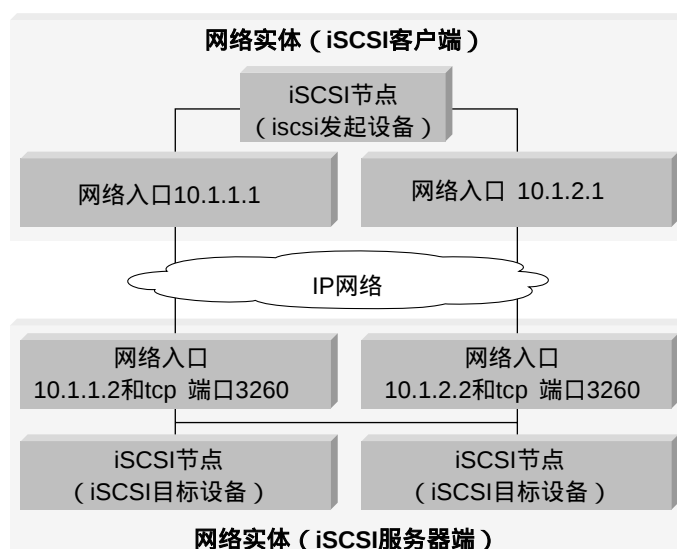
iSCSI 标准是 IETF 的 IP 存储（IPS）工作组制定和发布的一系列协议中的一个。IP 存储工作组一直致力于研究新型服务，包括经过改进的安全服务、目录服务和无需磁盘的客户端启动服务。此外，因为 iSCSI 主要使用以太网，传输协议的互操作性已经在网络行业打下了很好的基础。这种情况消除了光纤通道目前仍然面临的一个重要障碍。

iSCSI 术语和协议

iSCSI 标准使用了网络实体(Network Entity)的概念,它表示附加到某个 IP 网络上的设备或者网关。这种网络实体必须包含一个或者多个能够提供对 IP 网络的实际连接的网络端口。网络实体中包含的一个 iSCSI 节点可以利用任何一个网络入口(Network Portal)访问 IP 网络。iSCSI 节点可能是一个 iSCSI 发起设备或者目标设备,它在一个网络实体中通过它的 iSCSI 名称标明自己的属性。对于 iSCSI 而言,SCSI 设备是 iSCSI 节点中的一个可以提供 SCSI 功能的组件。每个 iSCSI 节点中只有一个 SCSI 设备。

网络入口(Network Portal)是网络实体中负责部署 TCP/IP 协议栈的核心组件。相对于发起设备,网络入口只需要通过它的 IP 地址标明属性。对于一个 iSCSI 目标设备而言,它的 IP 地址和它的 TCP 监听端口可以识别网络入口。对于 iSCSI 通信,一个发起设备网络入口和一个目标设备网络入口之间会建立一个连接。一个发起设备 iSCSI 节点和一个目标设备 iSCSI 节点之间的一组 TCP 连接可以构成一个 iSCSI 进程(Session)。这类似于但不等同于 SCSI I_T 的关系。

图 1 iSCSI 客户端/服务器架构



iSCSI 协议的实质是将 SCSI 发起设备和目标设备 (远程程序调用、参考 SCSI 架构模式、SAM) 模式映射到 TCP/IP 协议。iSCSI 可以提供它自己的、独立于它所传输的 SCSI CDB 信息的概念层。通过这种方式, SCSI 命令将被 iSCSI 请求传输,而 SCSI 应答和状态信息由 iSCSI 响应处理。而且, iSCSI 协议任务由同一个 iSCSI 请求和响应机制执行。

图 2 iSCSI 协议模式



与 SCSI 协议一样，iSCSI 采用了发起设备、目标设备的概念和名为协议数据单元（PDU）的通信消息。同样，iSCSI 传输方向是单独针对发起设备定义的。作为一种提高性能的方法，iSCSI 支持“phase-collapse”，让一个 SCSI 命令或者响应及其相关的数据可以通过一个单一的 iSCSI PDU 发送。

iSCSI 的 Cisco MDS 9000 系列 IPS 部署

iSCSI 命名和寻址

一个 iSCSI 节点的名称是独立于地址的，因为它并不会为某个 iSCSI 发起设备或者目标设备提供一个 IP 地址、一个在整个环境中独一无二的地址，或者一个永久性标识符。这使得它可以通过多种网络接口或者网络入口连接。目前有两种基于 iSCSI 标准的命名惯例：iSCSI 合格名称（iqn）和 EUI 格式。带有 IP 存储服务模块的 Cisco MDS 9000 系列可以支持这两种命名格式。但是，最常用的命名方式是 iqn 命名格式。

EUI 名称包含一个 eui（扩展独特标识符）和一个独一无二的、包含 64 个字符的字符串。这个字符串的名称与光纤通道全局名称（WWN）使用的名称一样。这种格式的一个例子是：eui.02004567A425678D。

IQN 名称包含一个 iqn 关键词和一个合格域名称。这种格式的一个例子是：

iqn.5886.com.acm.diskarrays-sn-a8675309。

管理或支持工具使用 iSCSI 地址格式来识别某个 iSCSI 节点。iSCSI 地址可以将该节点的名称和它所使用的网络地址结合到一起。iSCSI 地址的一个例子是：

iSCSI://172.16.1.1:3260/eui.02004567A425678D 或者

iSCSI://172.16.1.1:3260/iqn.com.acme.diskarrays.jbod1

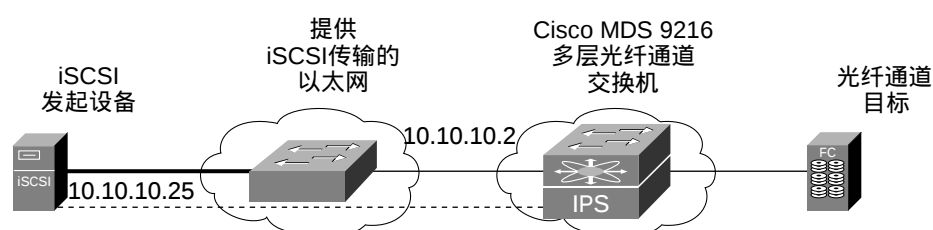
VLAN

MDS IPS 模块支持虚拟 LAN (VLAN)。虚拟 LAN (VLAN) 可以在一个单一的物理 LAN 上创建多个虚拟第二层网络。VLAN 可以提供流量隔离、安全和广播控制功能。每个千兆位以太网端口可以被配置为一个汇聚端口(Trunking)，并使用 IEEE 802.1Q 标准标签协议进行 VLAN 封装。

iSCSI 访问方法

Cisco MDS 9000 iSCSI 部署所采用的 iSCSI 访问方法是让 iSCSI 发起设备与光纤通道目标通信。这是第一种已经部署的模式。这种模式的相反方式在将来作为一项软件功能提供。

图 3 iSCSI 访问方法



为了便于理解这种访问方法，需要引入一个 FV_Port 的概念。FV_Port 是 IP 存储服务设备为了在千兆以太网和光纤通道设备之间转发数据帧而创建的一种逻辑端口。因为 Cisco MDS 9000 系列上的每个物理 FC 端口都需要通过协商成为一个 F_Port、FL_Port、E_Port 或者 TE_Port，并能够根据该端口所获得的硬件索引转发 FC 帧，IP 存储服务模块上的每个以太网端口都需要一个类似的索引。

iSCSI 发起设备访问 FC 目标

要想让一个 iSCSI 发起设备能够通过 MDS 9000 系列交换机访问 FC 目标，必须执行 4 个基本的步骤。附录 A 中提供了一个分阶段配置的实例。

1. 在 MDS 9000 IP 存储服务模块上设置 iSCSI 访问功能。
2. 设置 iSCSI 发起设备节点名称或者 IP 地址，并将其加入一个有效的 VLAN。
3. 创建 iSCSI 目标，并将它们映射到 FC 目标。
4. 设置一个包含 iSCSI 发起设备和 FC 目标的 FC 区。

为 iSCSI 配置 MDS 9000 IP 存储服务模块

第一步是设置供 iSCSI 客户端访问的 IP 地址。用户可以用不同的参数设置千兆位以太网端口，例如 MTU 大小、身份认证模式等。一旦千兆以太网端口设置完毕，用户将需要将每个需要使用的端口设为 iSCSI 端口。因此在 MDS 9000 IP 存储服务模块中，千兆位以太网端口可以同时支持 iSCSI 和 FCIP，所以必须让每个必要的千兆以太网端口专门运行 iSCSI。

设置一个 iSCSI 发起设备、IP 地址和 VSAN

用户可以通过 iSCSI 驱动程序，设置一个独一无二的 iSCSI 发起设备节点名称。如果用户不能设置一个静态名称，驱动器可以自动创建一个独特的 iSCSI 节点名称。如果节点名称是动态创建的，iSCSI 发起设备必须至少在 MDS 9000 IP 存储服务模块上登陆一次，以便让 MDS 9000 IP 存储服务模块认可这个自动创建的节点名称。要把 iSCSI 发起设备加入到正确的 VSAN 并进行相应的分区，必须要使用这个节点名称。在 MDS 9000 IP 存储部署中，iSCSI 发起设备能够跨越多个 VSAN，因而能够访问任何 VSAN 上的任何 FC 目标。

MDS 9000 IP 存储服务模块的 iSCSI 部署也允许按照 IP 地址进行分区。在配置任何分区之前，必须将发起设备的 IP 地址添加到特定的 VSAN。与 iSCSI 发起设备一样，IP 地址也可以跨越多个 VSAN。

创建 iSCSI 目标和 FC 目标

iSCSI 发起设备并不能直接附加到光纤通道目标。iSCSI 发起设备只能连接到用于代表一个或者多个光纤通道目标的 iSCSI 虚拟目标(Virtual Target)。为了实现这种功能，MDS IP 存储服务模块必须通过用 IQN 格式将所有可用的光纤通道目标的名称通知 iSCSI 发起设备，将光纤通道目标转化为 iSCSI 目标。IP 存储设备是通过用所需要的 iqn 字符串代替光纤通道 WWN 而做到这一点的。IP 存储服务模块可以通过一个基本的光纤通道名称服务器查询，获知目标的光纤通道 WWN。当 MDS 9000 IP 存储服务模块从某个 iSCSI 发起设备获得一个 SendTargets iSCSI 命令之后，该 iSCSI 发起设备就可以连接这些 iSCSI 目标了。

目前可以通过两种方法创建能够被转变为 iSCSI 目标的光纤通道目标。iSCSI 目标的首选创建方式是动态创建，或者也可以通过创建某个虚拟 iSCSI 目标进行静态设置。从本质上说，虚拟目标(Virtual Target)是通过从光纤通道到 iSCSI 的目标和 LUN 映射人工定义的。通过创建虚拟目标，发起设备将获得一个明确的目标名称，它们可以用该名称访问特定的光纤通道目标和特定的 LUN。

对于 SAN 中的 SAN 目标设备，IP 存储服务模块可以利用它自己的 FC_ID（由 SAN 分配）和一个相关的 pWWN，将 iSCSI 发起设备表现为 SAN 中的一个 N_Port 设备。

为了用 iSCSI 方式表示 FC 目标，每个 IP 存储模块千兆以太网端口都将利用它自己的端口群组标签（Port Group Tag, PGT），将一个 iSCSI 目标的名称设为“iqn.xxx”。该群组标签在该端口所在的物理交换机中是唯一的。

iSCSI 发起设备或者 IP 地址在 FC 目标中的分区

利用交换阵列中的分区(Zoning)功能，iSCSI 发起设备节点名称和/或 IP 地址可以象任何一个连接到光纤通道矩阵的光纤通道实体一样，被加入某个分区。这种部署可以提供极高的灵活性，尤其是在多路径环境中。光纤通道标准可以实现对某个象征性的节点名称（代表 iSCSI 发起设备或者 IP 地址）的分区。与任何一个光纤通道发起设备一样，iSCSI 发起设备可以跨越多个分区。

访问控制

传统的光纤通道 SAN 中的访问控制是通过部署分区服务实现的。在 Cisco MDS 9000 系列中推出了 VSAN 功能之后，VSAN 和分区都可以被用于访问控制。VSAN 可以用于将物理光纤通道 SAN 分为若干个逻辑交换阵列。这种功能非常类似于 VLAN 在以太网环境中所扮演的角色。分区服务可以限制 VSAN 中不同终端之间的通信。每个 VSAN 都拥有它自己的一组分区服务。

光纤通道或者 iSCSI 发起设备只能访问同一个分区和同一个 VSAN 中的光纤通道或者 iSCSI 目标。利用 iSCSI 的 MDS 部署，一个 iSCSI 发起设备不仅限于某个特定的 VSAN。相反，iSCSI 发起设备可以通过配置加入到任何一个所选择的 VSAN 中。这种灵活性使 iSCSI 发起设备可以通过配置，访问网络的任何一个 VSAN 上的任何一个光纤通道设备。

除了普通的访问控制以外，iSCSI 还可以部署基于 IP 的身份认证机制，以限制对于任何目标的访问权限。身份认证步骤发生在 iSCSI 的登录阶段。Cisco MDS 9000 系列交换机所部署的身份认证算法是通用的握手验证协议（CHAP）。如果必要的话还可以禁用身份认证功能，但是我们建议不要采用这种做法。iSCSI 还支持其他的身份认证算法（例如 SRP、公共密钥方法（SPKM-1 或 2）），并将在未来的软件版本中部署。

iSCSI LUN 映射

iSCSI 的 Cisco MDS 9000 部署可以支持先进的 LUN 映射功能，以提高物理磁盘的可用性和提供高等级的灵活性。下面列出了现有的 LUN 映射方法：

- 将不同的 FC 目标的 LUN 映射到一个 iSCSI 虚拟目标（在未来的软件版本中支持）
- 将一个 FC 目标的 LUN 子集映射到多个 iSCSI 虚拟目标

很多存储阵列可以让一个光纤通道目标端口能够看到多个 LUN。通过将一个光纤通道目标 LUN 屏蔽/映射到多个逻辑 iSCSI 虚拟目标，IT 管理员可以获得很大的灵活性。这种灵活性让管理员可以对容量庞大的、价格昂贵的磁盘阵列资源进行逻辑划分，将其分配到多个 iSCSI 目标，而这些 iSCSI 目标可以被不同的 iSCSI 用户群组使用。以前，这只能通过在磁盘阵列控制器上进行 LUN 屏蔽和映射而实现。但是，利用思科的 MDS 9000 IP 存储服务模块，这种功能可以在网络中实现。这种功能还可以在访问控制方面提供更高的安全性。如果身份认证流程发现某个 iSCSI 主机没有获准访问逻辑 iSCSI LUN，它的访问请求就会被拒绝。

iSCSI 高可用性

Cisco MDS 9000 iSCSI 部署支持 iSCSI 冗余功能，可以提供很高的可用性。这些冗余功能包括以太网通道和虚拟路由器冗余协议（VRRP）。以太网通道功能使用户可以将多个物理以太网连接捆绑成一个单一的更高带宽的逻辑连接。在最初的版本中，以太网通道在一个以太网通道捆绑包中只支持两个连续的连接端口，

而且它们必须位于同一个 IP 存储服务模块上。对 802.3ad 端口汇聚标准的全面支持将在未来的软件版本中提供。VRRP 让用户可以创建一个虚拟 IP 地址（第三层）和一个虚拟 MAC 地址（第二层）对，并在多个以太网网关端口上共享这些地址。Cisco MDS 9000 系列 iSCSI 部署可以在同一个或者不同的 MDS 9000 交换机或 IP 存储服务模块的多个端口上支持 VRRP。如果 VRRP 功能因为某个网关发生故障而被调用，TCP 进程信息将会失去同步，这要求 iSCSI 发起设备重新与备用交换机或者网关端口建立一个连接。

安全地将一个 iSCSI 主机集成到一个光纤通道 SAN 中

Cisco MDS 9000 系列交换机及其业界领先的可用性、可扩展性、安全性和高性能架构可以利用 IP 存储服务模块将 SAN 扩展到 IP 网络。连接到某个基于 MDS 9000 系列的矩阵的光纤通道存储可以利用 iSCSI 协议，扩展到没有光纤通道主机总线适配器（HBA）的中端服务器。具有一个 10/100M 或者千兆以太网卡的服务器，或者为了进一步提高性能而采用了一个 TCP 卸载引擎（TOE）NIC 卡的服务器，现在可以访问光纤通道存储。通过结合 IP 存储服务模块对于 FCIP 的支持，Cisco MDS 9000 系列成为一种真正业界领先的集成化多协议交换平台。

MDS 9000 系列中内含的光纤通道安全机制（例如 VSAN 和分区）可以通过 iSCSI 及其相关服务提供的安全功能而得到进一步加强。iSCSI 所添加的安全服务（例如通过 CHAP 对 iSCSI 发起设备进行身份认证）可以拓展 SAN 安全措施，安全地集成 iSCSI 主机。创建 iSCSI “虚拟目标”的灵活性可以在分配光纤通道存储方面为 iSCSI 发起设备提供基于 LUN 级别的精确度。这种功能在很多具有较低的 I/O 需求的 iSCSI 发起设备需要通过一个光纤通道存储阵列接口访问存储时尤为有用。

通过将 iSCSI 协议作为一个面向数据块的 SCSI 协议的传输方式，很多中低端服务器现在可以被集成到 SAN 中，进行集中的管理。目前，很多这样的服务器使用的是直接附加存储（DAS），因而很难进行扩展，无法充分地利用它们的存储资源。例如，服务器 A 和服务器 B 可能都拥有 100GB 的直接附加存储。但是，服务器 A 只使用了 30% 的存储容量，而服务器 B 使用了 90%。对于 DAS，用户无法在必要的情况下，方便地将服务器 A 上未被利用的存储资源移植到服务器 B。光纤通道 SAN 是实现存储资源共享的理想解决方案，但是很多企业并不愿意采用 SAN，因为昂贵的端口成本通常会超出他们对于这些低中高端服务器制定的预算。而且，通常这些服务器的 I/O 需求也比较低——大约介于 5Mbps 到 30Mbps 之间，因而不能证明移植到光纤通道网络的必要性。现在利用 iSCSI 协议和思科的 MDS 9000 iSCSI 部署，用户能够以一种更加经济的方法，方便地将这些服务器加入 SAN。利用千兆以太网连接所提供的带宽和 iSCSI 服务器通常较低的 I/O 要求，用户能够将多个 iSCSI 服务器连接到一个单一的千兆以太网端口。利用 IP 存储服务模块提供的 8 个千兆以太网端口，用户可以更加方便地扩展 iSCSI 客户端。利用服务器网络接口卡（NIC）——无论是 10/100M 还是千兆以太网，以及思科和微软为 Windows 平台提供的 iSCSI 驱动程序，这种服务器可以充分地享有 SAN 所带来的优势。

通过将 iSCSI 添加到 iSCSI 发起设备中的 IP 堆栈，iSCSI 客户端的 CPU 将需要执行额外的处理，以发送和接收 iSCSI 数据包和保持 iSCSI 进程。因此，iSCSI 可能会提高系统的整体 CPU 利用率。为了帮助系统进行这些额外的处理任务，一些传统的 HBA 和网络供应商开发了 iSCSI 主机总线适配器，也被称为 TCP 卸载引擎（TCP Offload Engine, TOE 卡）。大部分供应商在不同的平台上为它们的 TOE 卡提供了它们自己的 iSCSI 驱动程序。有些供应商可以通过主机 CPU 完全分担 iSCSI 堆栈处理功能。有些厂商则只提供了分担 TCP 堆栈处理的功能。

iSCSI 性能测试

我们利用著名的测试工具 IOmeter 对 Cisco MDS 9000 系列的 IP 存储服务模块的性能进行了测试。本小节的目的是说明不同的 I/O 模式对于 IP 存储服务模块上的 iSCSI 性能的影响。不同的测试采用了不同的 I/O 模式、不同大小的数据块和不同的读写比例。

测试配置

下面概括地介绍了用于获得本文中介绍的测试结果的测试配置：

服务器：Windows Dell 1650，内置 GE NIC，1.13 GHz CPU，2GB RAM，Windows 2000 Server SP3。

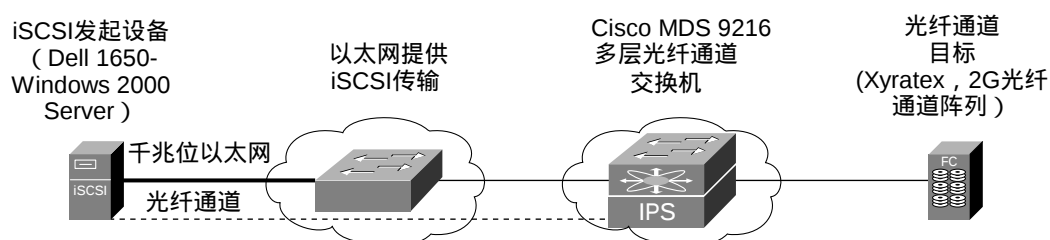
服务器：思科 iSCSI 驱动程序 3.1.1 版，一个 QLogic 2300 光纤通道主机总线适配器被用于基准测试。测试中使用了一个能够分担 TCP 处理任务但不能分担全部 iSCSI 处理任务的第三方 TOE 卡。

存储：Xyratex 2Gig RAID 控制器存储，具有 8 个 73GB 的、10K RPM 的存储驱动器。

交换机：Cisco MDS9216，具有一个 IP 存储服务模块，版本为 1.1.(1)。

Xyratex 存储阵列被连接到 MDS 9000 系列交换机，服务器通过一个 QLogic 2300 主机总线适配器（设置为 1Gbps）连接到 MDS 9000 系列交换机。Xyratex 上的 LUN 被创建为 RAID 0 LUN，分布在 8 个独立的磁盘上。本测试所采用的磁盘采用了基于 Windows 的 NTFS 文件系统。

图 4 iSCSI 测试方案

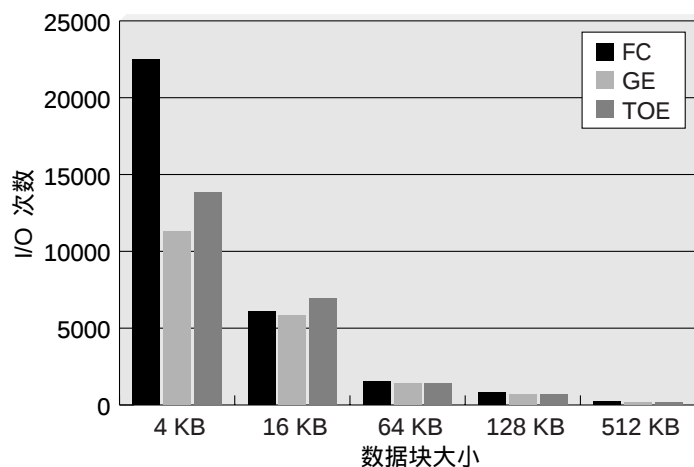


I/O 线程的大小：4KB、16KB、64KB、128KB、512KB

测试结果

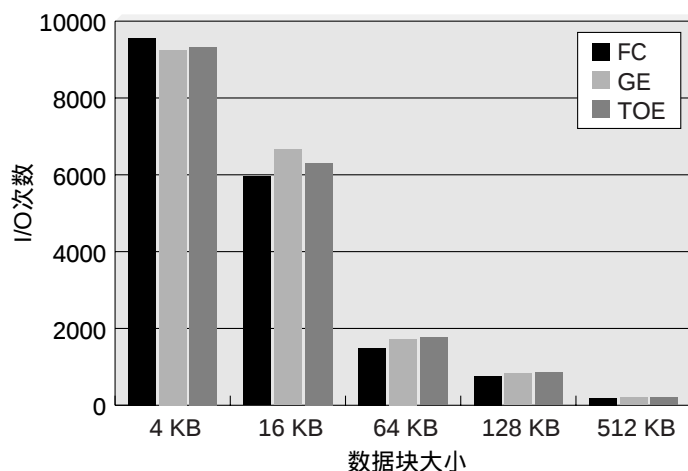
详细的测试结果见附录 B。

图 5 IOP 对比——100%读取——100%连续



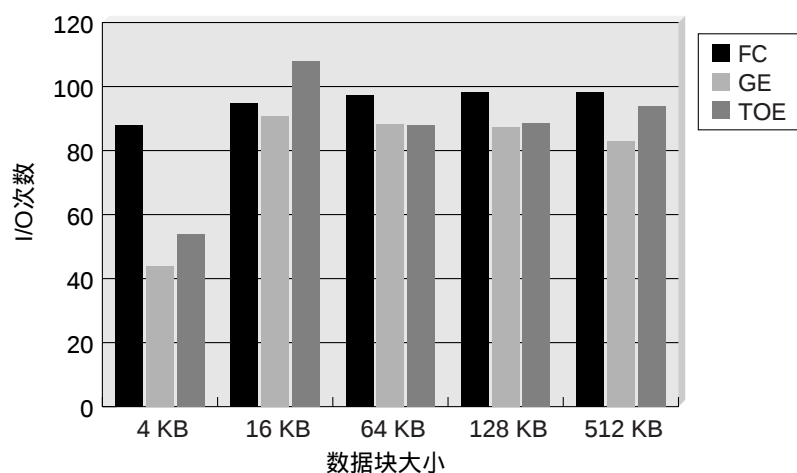
不同测试中的每秒 I/O 次数表明，随着数据块的增大，各种测试中的 I/O 次数之间的差距会缩小。因为 iSCSI 会给 CPU 带来额外的负担，所以数据块越小，所需要的 CPU 资源就越多，这可以解释 FC 和 iSCSI 之间的 I/O 差距。

图 6 IOP 对比——100%写入——100%连续



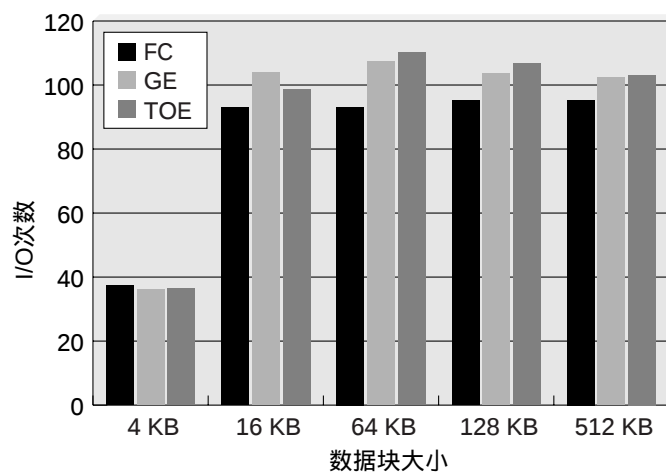
上图所显示的写入性能表明，三种测试方案的测试结果非常接近。应当指出的是，由于本测试中所使用的驱动器的数量较少，CPU 不能为 FC HBA 或者 iSCSI TOE 卡提供足够的空间。更多的驱动器空间可以支持更高的 I/O，使用更多未被利用的 CPU 资源。

图 7 吞吐量对比——100%读取——100%连续



从上图可以看出，iSCSI 在读取较大的区块时的性能与光纤通道差不多（如果不是更好的话）。在不同的测试中，iSCSI 在传输较小的数据块时的吞吐量有所降低，这是因为在这种情况下 iSCSI 对 CPU 的占用率较高。

图 8 吞吐量对比——100%写入——100%连续



在上图显示的写入吞吐量中，iSCSI 的性能与光纤通道差不多（如果不是更好的话）。对于较少的数据块，iSCSI 的吞吐量会受到影响，这主要是因为少量的驱动器和它们内置的 I/O 处理功能。如果在后端添加更多的驱动器，性能将会进一步提升。

图 9 CPU 对比——100%读取——100%连续

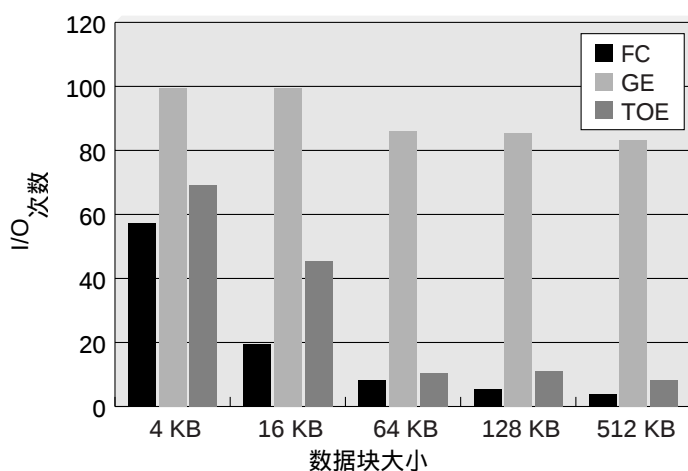
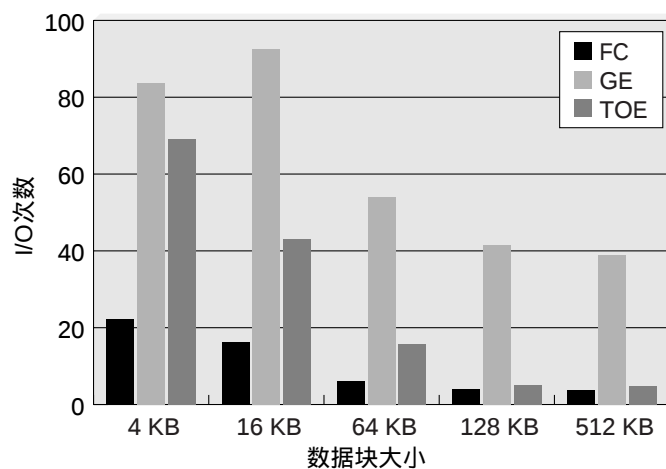


图 10 CPU 对比——100%写入——100%连续



在上面的两幅图中，因为 iSCSI 会增加 CPU 的负荷，所以图中显示了不同测试的 CPU 利用率的区别。通过用 TCP 卸载引擎降低 CPU 的利用率，可以有效地减少 CPU 的负荷。如果 TOC 卡供应商可以完全分担 iSCSI 的处理任务，CPU 的利用率可以进一步降低。

结论

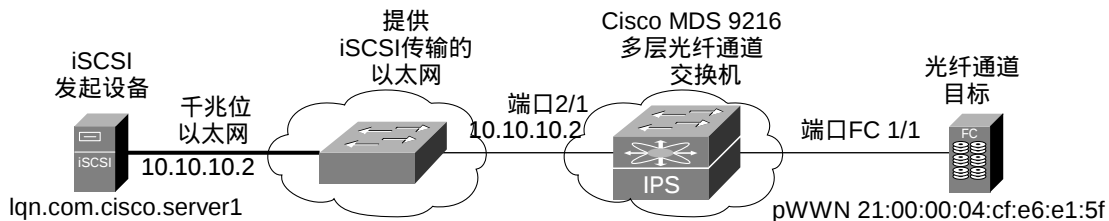
企业环境现在可以利用 MDS 9000 系列创建大规模的光纤通道 SAN。而且，利用 MDS 9000 系列 IP 存储服务模块，企业可以部署高度可用的、可扩展的、支持 FCIP 和 iSCSI 的多协议 SAN 基础架构。Cisco MDS 9000 系列可以提供一种多协议 SAN 企业级解决方案，为企业提供很高的可用性、可扩展性和更加简便的管理。通过将 SAN 拓展到中低端服务器，存储管理人员现在可以在他们的整个应用环境和所有应用服务器

中充分发挥 SAN 的优势。通过让企业利用现有的 IP 基础设施，将中低端应用服务器集成到一个集中式 SAN，思科可以为企业提供一个全面的存储解决方案，并为他们提供出色的投资回报。

附录 A

下面是一个将 iSCSI 发起设备连接到光纤通道目标的基本配置的示例。下面的示意图将介绍怎样在 MDS 9000 系列 IP 存储服务模块上配置 iSCSI。在这种基本的配置中，所有 iSCSI 发起设备和存储端口都位于缺省的 VSAN——VSAN1——之中。

图 11 iSCSI 配置示例



要让上述服务器访问光纤通道存储，必须执行下列步骤。在配置 iSCSI 之前，光纤通道存储必须被连接到模块 fc 1/1 上的 MDS，并被启用。

1. 在 VLAN 5 中配置用于 iSCSI 访问的 IP 存储服务模块千兆位以太网端口：

```
interface GigabitEthernet2/1.5
ip address 10.10.11.30 255.255.255.0
no shutdown
interface iscsi2/1
mode store-and-forward
no shutdown
```
2. 在这个步骤中，需要按照 IP 地址进行分区。因此，iSCSI 发起设备的 IP 地址必须被添加到存储所在的 VSAN1：

```
iscsi initiator name 10.10.11.230
vsan 1
```
3. 在这个步骤中，需要启用向 iSCSI 目标动态创建 FC 目标的功能。而且，还需要启动 CHAP 身份认证功能。下面是这个配置的输出结果：

```
iscsi authentication chap
iscsi import target fc
username cisco password 7 fewhg1xnkfy1sewsm1 iscsi
```

- 在上述步骤完成之后，用户需要将 iSCSI 发起设备的 IP 地址和光纤通道存储划分到同一个分区中。配置方法如下：

```
zoneset name ZS1 vsan 1
member Path1
zoneset activate name ZS1 vsan 1
zone name Path1 vsan 1
member pwwn 21:00:00:04:cf:e6:e1:5f
member symbolic-nodename 10.10.11.230
```

- 因为 iSCSI 发起设备的 IP 地址与 IP 存储服务模块千兆以太网地址不在同一个子网中，所以用户需要为发起设备创建一个静态路由，以便让它可以访问 MDS 9000 IP 存储服务模块。配置方法如下：

```
ip route 10.10.11.0 255.255.255.0 10.10.1.2
```

附录 B

下列表格提供了从我们针对测试基础设施开展的一系列测试中获得的性能结果。

100%读取 - 100%连续

IOPS	FC	GE	TOE
4KB	22517.75	11275.21	13815.29
16KB	6076.81	5809.4	6900.96
64KB	1555.13	1410.71	1407.68
128KB	784.87	699.31	709.1
512KB	196.33	165.58	187.49

100%写入 - 100%连续

IOPS	FC	GE	TOE
4KB	9568.51	9253.31	9332.11
16KB	5954.32	6655.51	6304.96
64KB	1490.47	1718.75	1763.09
128KB	760.38	828.39	853.25
512KB	190.69	204.66	206.27

100%读取 - 100%连续

吞吐量	FC	GE	TOE
4KB	87.96	44.04	53.97
16KB	94.95	90.77	107.83
64KB	972	88.17	87.98
128KB	98.11	8741	88.64
512KB	98.15	82.79	93.74

100%写入 - 100%连续

吞吐量	FC	GE	TOE
4KB	3735	36.15	36.45
16KB	93.02	103.99	98.52
64KB	93.15	107.42	110.19
128KB	95.05	103.55	106.66
512KB	95.33	102.33	103.13

100%读取 - 100%连续

CPU	FC	GE	TOE
4KB	57.32	99.56	69.28
16KB	19.55	99.39	45.53
64KB	8.21	86.17	10.41
128KB	5.54	85.32	11.12
512KB	3.88	83.28	8.23

100%写入 - 100%连续

CPU	FC	GE	TOE
4KB	22.21	83.71	68.99
16KB	16.32	92.43	43.09
64KB	6.13	53.95	15.78
128KB	4.13	41.64	5.16
512KB	3.77	39.05	4.74

