



第5章

SAN服务器配置与管理

学习目标

- 理解 SAN 基础知识、FC SAN、IP SAN、iSCSI 协议栈以及 FCoE 等相关理论知识。
- 掌握 SAN 服务配置与管理以及多路径链路 iSCSI 虚拟磁盘应用部署等相关知识与技能。

5.1 项目陈述

存储区域网络是一种面向网络的、以数据存储为中心的存储架构。SAN 采用可扩展的网络拓扑结构连接服务器和存储设备,并将数据的存储和管理集中在相对独立的专用网络中,向服务器提供数据存储服务。以 SAN 为核心的网络存储系统具有良好的可用性、可扩展性、可维护性,能保障存储网络业务的高效运行。本章讲解 SAN 基础知识、FC SAN、IP SAN、iSCSI 协议栈以及 FCoE 等相关理论知识,项目实践部分讲解 SAN 服务配置与管理以及多路径链路 iSCSI 虚拟磁盘应用部署等相关知识与技能。

5.2 必备知识

5.2.1 SAN 基础知识

传统 SAN 的主要支撑技术是光纤通道技术。与 NAS 完全不同,它不是把所有的存储设备集

中安装在一个服务器中,而是将这些设备单独通过光纤交换机连接起来,形成一个光纤通道存储在网络中,然后再与企业的局域网进行连接,这种技术的最大特性就是将网络、设备的通信协议与存储传输介质隔离开,因此存储数据的传输不会受网络状态的影响。

基于光纤交换机的 SAN 存储,通常会综合运用链路冗余与设备冗余的方式,同一服务器访问磁盘阵列有多条冗余路径,不论其中的部分线路或者部分光纤交换机出现故障,都不会导致服务器存储失败。这种方式部署成本较高,但对银行、数据中心等存储了大量关键数据且不允许业务中断的行业来说非常重要。

目前处于迅速成长的 IP SAN 存储,是在传统的 FC SAN 的基础上演变而来的。IP SAN 是在以太网上架构一个 SAN 存储网络,把服务器或普通工作站与存储设备连接起来的存储技术。IP SAN 在 FC SAN 的基础上更进一步,它把 SCSI 协议完全封装在 IP 协议之中。简单来说,IP SAN 就是把 FC SAN 中光纤通道解决的问题通过更为成熟的以太网实现,从逻辑上讲,它是提供区块级服务的 SAN 架构。

1. SAN 概述

传统数据存储 DAS 存在以下几方面的不足。

(1) 存储空间得不到有效的利用。DAS 存储不管是使用内置的磁盘空间,还是通过直连磁盘陈列来获取空间,都避免不了主机系统之间形成多个数据孤岛的问题。每个单独的小“岛屿”都是一个专门直接连接的存储器应用,导致存储内容无法共享,存储空间无法合理利用。例如,有些应用服务器由于应用数据少而留着很大的存储空间处于空闲状态,而有些应用服务器的数据量相当大,空闲存储空间很小;因此,存储资源得不到合理利用,浪费了很多宝贵的存储资源,增加了用户的整体拥有成本。

(2) 存储空间无法满足日益增长的数据需求。存储空间主要依赖于服务器上自带硬盘,而每台服务器能够挂接的硬盘数量是有限的,随着业务的发展,需要保存的数据量必然陡增,所需存储空间也会越来越大,仅依靠本机的存储空间已经无法满足要求。况且,每一次安装硬盘,都需要给服务器断电,进而中断业务的运行;同时数据管理变得复杂,数据放置过于分散,大大加重了管理员的负担。

(3) 存储架构无法满足日益发展的业务需求。由于采用了硬盘挂接在服务器上这种存储架构,服务器上任意部件(如内存、硬盘、硬盘连接线等)发生故障,都会使服务器运行中断,业务和应用停止服务。对用户而言,业务的中断带来了经济损失,影响经济效益,而且业务的中断对企业信誉也是一种致命的打击,造成负面影响。所以需要一种高效、安全而且可靠的存储架构来保障信息数据的安全和可用性,保证业务的持续性。

(4) 存储架构可扩展性差。可扩展性是指现有投资在受到保护的情况下,不影响业务的前提下,通过添加新设备而达到客户新的规模要求和性能要求。以扩充容量(即扩容)为例,对于传统数据存储而言,扩容不仅意味着往服务器上增加新的硬盘,还需要将一部分数据从已有硬盘迁移到新增硬盘上。这种扩容意味着现有存储架构的可扩展性不好,理想的扩容应该是在保证业务连续性的前提下,通过添加存储设备(如硬盘)从而达到客户扩容要求的方案。

针对以上问题,业界提出了存储区域网络(Storage Area Network,SAN),通常人们将 SAN 技术视为 DAS 技术的一个替代者。存储网络工业协会(Storage Networking Industry Association, SNIA)对于 SAN 的标准定义是:“A network whose primary purpose is the transfer of data

between computer systems and storage elements and among storage elements”,即 SAN 是用来在计算机系统和存储单元之间、存储单元与存储单元之间进行数据传输的网络系统。SAN 包含一个通信系统基础结构,包括物理连接、管理层、存储单元和计算机系统,以确保数据传输的安全性和稳定性。

SAN 是服务器和存储资源之间的一个高性能的专用网络体系,它提供存储装置、计算机主机及相关网络设备的管理机制,并且提供强而有力且安全的数据传输环境,它为了实现大量原始数据的传输而进行了专门的优化。SAN 通常被认为是提供数据块存取(Block I/O)服务而非档案存取服务,但这并不是 SAN 的必要条件,事实上,可以把 SAN 看成是对 SCSI 协议在长距离应用上的扩展。利用 SAN 可以构架理想的存储结构,这种理想的存储结构包括如下特征。

- (1) 具有可伸缩能力。
- (2) 可扩展到整个世界。
- (3) 非常可靠。
- (4) 提供尽可能高的传输速度。
- (5) 易于管理。

例如,华为技术有限公司,在全球有超过 100 000 名员工,假设在荷兰工作的员工,希望能够访问存储在深圳总公司存储设备上的相关数据,那么要求它是“能扩展到整个世界”的存储架构设计。一个设计良好的存储架构,可以运行很多年。当设计一个庞大的信息通信(ICT)基础设施时,需要如下设计需求列表。

- (1) 设计可以无限扩展,可以方便地增加方案中的设备数量。
- (2) 设计能够允许各个组件之间的距离没有限制或限制较小。在实践中,相距 20 000km 的设备组件是允许互联的。
- (3) 设计必须是可靠的,当发生硬件故障或者人为失误时,也不会给公司造成严重的问题。
- (4) 相互连接的组件之间能够以最快的速度进行通信。
- (5) 即使设计非常复杂,少数的技术员也可以维护和监控整个存储架构中的设备。ICT 部门不需要 100 个人来管理 100 台或者 200 台设备。低成本、高效率的管理也是一个大的设计需求。
- (6) 设计应该是灵活的。在基础设施中改变、替换和增加组件不会有任何限制,这意味着即使经过几年技术的发展,仍然可以将新技术集成到当前的基础设施中。
- (7) 设计应该支持异构。异构是指来自不同厂商的设备可以像来自同一个厂商的设备那样一起工作。支持异构是存储架构设计中的一个重要挑战,并不容易做到:一方面,大多数的客户只会购买一家公司的设备,因为客户往往只想和一个硬件供应商签订服务合同,以防止在发生技术问题时,需要去联系多个厂家的技术支持团队;另一方面,又不能过于依赖同一家公司,以免这家公司的产品出现批次问题或者其他问题时,影响本公司的正常运转。另外,如果系统支持异构,那么当从一家公司的产品切换到另外一家公司的设备时,将更容易进行迁移。

2. SAN 组网

SAN 也叫存储区域网络,它是将存储设备(如磁盘阵列、磁带库、光盘库等)与服务器连接起来的网络。结构上,SAN 允许服务器和任何存储设备相连,并直接存储所需数据,如图 5.1 所示为典型的 SAN 组网方式。

相对于传统数据存储方式,SAN 可以跨平台使用存储设备,可以对存储设备实现统一管理和

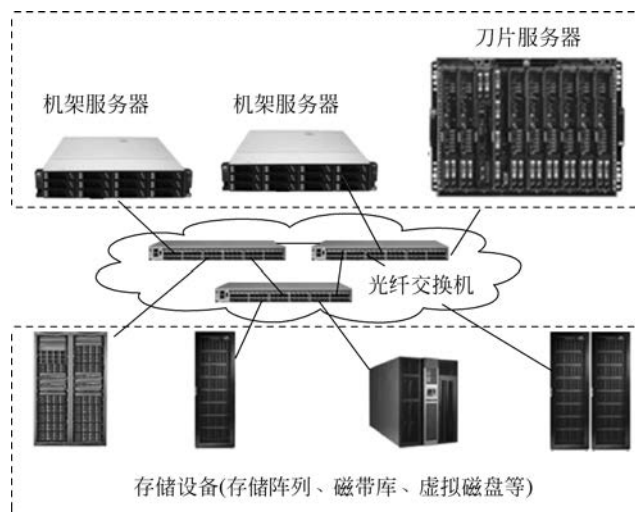


图 5.1 典型的 SAN 组网方式

容量分配,从而降低使用和维护的成本,提高存储的利用率。根据 Forrester 研究报告,使用传统独立存储方式时存储利用率为 40%~80%,平均利用率为 60%,存储通常处于低利用率状态。SAN 对存储资源进行集中管控,高效利用存储资源,有助于提高存储利用率。更高的存储利用率意味着存储设备的减少,网络中的电能能耗和制冷能耗降低,节能省电。

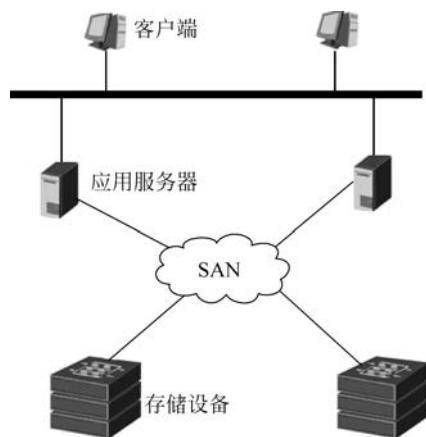


图 5.2 SAN 的网络拓扑架构示意图

此外,通过 SAN 网络主机与存储设备连通,SAN 为在其网络上的任意一台主机和存储设备之间提供专用的通信通道,同时,SAN 将存储设备从服务器中独立出来。SAN 支持通过光纤通道协议和 IP 协议组网,支持大量、大块的数据传输;同时满足吞吐量、可用性、可靠性、可扩展性和可管理性等方面的要求,如图 5.2 所示。

SAN 和 LAN 相互独立,这个特点的优势在前面已经提过,然而它会带来成本和能耗方面的一些不足。

(1) SAN 需要建立专属的网络,这就增加了网络中线缆的数量和复杂度。

(2) 应用服务器除了连接 LAN 的网卡之外,还需配备与 SAN 交换机连接的主机总线适配器(Host Bus Adapter,HBA)。

3. SAN 组件

SAN 由三个基本组件组成:服务器、网络基础设施和存储。这些组件可以进一步划分,分别是:端口、连接设备、集线器、存储阵列等。

(1) SAN 网络服务器。

在所有 SAN 解决方案中,服务器基础结构是其根本,其基础结构可以是多种服务器平台的混合体,包括 Windows、UNIX、Linux 和 macOS 等。

(2) SAN 网络存储。

光纤接口存储设备是存储基础结构核心。SAN 存储基础结构能够更好地保存和保护数据,能

够提供更好的网络可用性、数据访问性和系统管理性。SAN 为了使存储设备与服务器解耦,使其不依赖于服务器的特定总线,将存储设备直接接入网络中。从另一个角度看,存储设备做到了外置或外部化,其功能分散在整个存储系统内部。

(3) SAN 网络互连。

实现 SAN 的第一个要素是通过 FC 等通道技术实现存储和服务器组件的连通。所使用的组件是实现 LAN 和 WAN 所使用的典型组件。与 LAN 一样,SAN 通过存储接口的互连形成很多网络配置,并能够跨越很长的距离。除了线缆和连接器,还包括如下具体互连设备。

交换机:交换机是用于连接大量设备、增加带宽、减少阻塞和提供高吞吐量的一种高性能设备。

网桥:网桥的作用是使 LAN/SAN 能够与使用不同协议的其他网络通信。

集线器:通过集线器,仲裁环线路上一个逻辑环路上可以连接多达 127 个设备。

网关:网关是网络上用来连接两个或更多网络或设备的站点,是一个网络连接到另一个网络的接口,也用于两个高层协议不同的网络互连,也被称为网间连接器、协议转换器。网关产品通常用来实现 LAN 到 WAN 的访问,通过网关,SAN 可以延伸和连接。

(4) SAN 网络端口有三种常用端口。

① FC 接口使用 FC 协议,使用该种协议的 SAN 架构,称为 FC SAN。

② ETH 接口使用 iSCSI 协议,使用该种协议的 SAN 架构,称为 IP SAN。

③ FCoE 接口使用 FCoE 协议,使用该种协议的 SAN 架构,称为 FCoE SAN。

4. DAS 和 SAN 的区别

SAN 网络和 DAS 直连一样,都是以 SCSI 块的方式发送数据,将数据从存储设备传送到服务器上。当然,SAN 网络和 DAS 直连有一些显著的区别,如价格,用户购买 SAN 网络所需花销远远大于 DAS,如 DAS 缆线的连接范围在 25m 以内,而 SAN 网络连接则可以长达数百或者数千千米等。

在一个基于 SAN 网络架构的解决方案中,SAN 不只会在网上发送单独的 SCSI 协议块,而是将 SCSI 协议块封装到一个数据包或者数据帧中,利用网络将数据包传输到更远的距离。数据包就好像是一个信封,我们可以利用信封来把信传递给某人。信可以看成是用户数据,而信封就是数据包。事实上,我们不可能通过将信纸放在地上,然后让风将信纸送到收信人的地址。所以一个好的办法是将信纸装入信封,并且贴上邮票,然后写上正确的地址信息并把信塞入一个邮箱。国家邮政服务人员将信件从邮箱取出,并将它传递到收信人手中。当然,也有其他的办法可以将信送到收信人手中,一个替代办法是选择专业的快递服务公司,例如 UPS 或者 FedEx。它们有自己的投递系统,你需要将这封信放入一个特殊的信封中。然后,负责送货服务的传输系统将负责把信送到收信人手中。

现在有多种方法将 SCSI 块发送到跨 SAN 的连接中,这些方法被称为协议,每个协议都有不同的方法来描述处理 SCSI 块的传输方式。如上所述,FC、iSCSI 和 FCoE 是 SAN 网络架构中的三种常用协议,FC 协议通常和 iSCSI 协议用于现代的 SAN 架构中,而 FCoE 协议主要用于 SAN 和 LAN 业务融合场景,从协议、应用场景、优缺点等几方面来对比 DAS 和 SAN 两种存储架构,如表 5.1 所示。

表 5.1 DAS 和 SAN 的区别

指 标 对 比	DAS	SAN
成本	低	高
扩展性	不易于扩展	易于扩展
是否集中管理	否	是
备份效率	低	高
网络传输协议	无	光纤通道协议

从连接方式上对比,DAS 采用了直接连接,即存储设备直接连接应用服务器,但是扩展性较差;SAN 网络则是通过多种技术来连接存储设备 and 应用服务器,具有很好的传输速率和扩展性。SAN 不受现今主流的、基于 SCSI 存储结构的布局限制。特别重要的是,随着存储容量的爆炸性增长,SAN 允许独立地增加它们的存储容量。SAN 网络的结构允许任何服务器连接到任何存储阵列,这样不管数据放置在哪里,服务器都可以直接存取所需的数据。因为采用了光纤接口,SAN 还具有更高的带宽。

SAN 的优点:将存储和服务端隔离,简化了存储管理,能够统一、集中地管理各种资源,使存储更为高效。通常的网络中,可能一个服务器可用空间用完了,另一个服务器还有很多可用空间。SAN 把所有存储空间有效地汇集在一起,每个服务器都享有访问组织内部的所有存储空间的同等权利,SAN 能屏蔽系统的硬件,可以同时采用不同厂商的存储设备。

SAN 的不足:跨平台性能没有 NAS 好,价格偏高,搭建 SAN 比在服务器后端安装 NAS 要复杂得多。

DAS 存储一般应用在中小企业,与计算机采用直连方式;NAS 存储则通过以太网添加到计算机上,SAN 和 NAS 的区别主要体现在操作系统在什么位置。NAS 和 SAN 混合搭配的解决方案为大多数企业带来最大的灵活性和性能优势。服务器环境越是异构化,NAS 就越是重要,因为它能无缝集成易购的服务器。而企业数据量越大,高效的 SAN 就越重要,SAN 网络则使用光纤接口,提供高性能、高扩展性的存储,其应用场景如下。

- (1) 对数据安全性要求很高的企业,如金融、证券和电信。
- (2) 对数据存储性能要求高的企业,如电视台、测绘部门和交通部运输部门。
- (3) 具有本质上物理集中、逻辑上又彼此独立的数据管理特点的企业,如银行、证券和电信等行业。

5.2.2 FC SAN 概述

随着当今社会对信息存储需求的空前增加,对信息存储系统的性能、信息网络的利用率和信息的备份、容灾能力都有更高的要求,SAN 可以很好地满足数据统一存储、企业数据共享、远程数据容灾等的需要。随着 IT 技术的迅速发展及各种数据的集中化,建立一个基于 SAN 的存储体系结构也已经成为信息化的必然之路。FC SAN 是当今 SAN 网络中的主流,在高性能应用环境中占主要份额。

1. FC SAN 基础知识

20 世纪 80 年代,随着计算机处理器运算能力的提高,外部设备的 I/O 带宽成为整个存储系统的一大瓶颈。为了解决 I/O 瓶颈对整个存储系统所带来的消极影响,提高存储系统的存取性能,

美国国家标准委员会(American National Standard Institute,ANSI)的 X3T11 工作组于 1988 年开始制定一种高速串行通信协议——光纤通道协议。FC 协议制定的初衷是用来提高硬盘传输带宽,侧重于数据的快速、高效、可靠传输。随着技术发展,该协议将快速可靠的通道技术和灵活可扩展的网络技术有机地融合在一起,既提供通道所需要的指令集,也提供网络所需要的各种协议,因此,它不仅能够进行数据的高速传输、音频和视频信号的串行通信,而且为网络、存储设备和数据传送设备提供了实用、廉价和可扩展的数据交换标准,并能广泛用于高性能大型数据仓库、数据存储备份和恢复系统、基于网络的存储、高性能的工作组、数据的视/音频网络等。这些特点使得 FC 协议在整个 20 世纪 90 年代都得到了人们的认可,并且从 20 世纪 90 年代末开始,FC SAN 得到广泛应用。目前,该协议被用在绝大多数大容量、高端直连存储设备卡上。

FC SAN 是指使用 FC 协议的 SAN 网络。作为 SAN 网络中第一个成功的千兆位串行传输技术,FC 已成为最适合块 I/O 应用的体系结构。FC 满足存储网络对传输技术的下列需求。

- (1) 高速长距离的串行传输。
- (2) 大规模网络应用中的异步通信。
- (3) 较低的传输误码率。
- (4) 较低的数据传输延迟。
- (5) 模块化和层次化结构。
- (6) 传输协议可在 BA 上以硬件方式实现,减少对服务器 CPU 的占用。

2. FC 协议栈

光纤通道协议其实并不能翻译成光纤协议,只是协议普遍采用光纤作为传输线缆,因此很多人把 FC 称为光纤通道协议。在逻辑上,可以将 FC 看作一种用于构造高性能信息传输的、双向的、点对点的串行数据通道。在物理上,FC 是一到多对应的点对点的互连链路,每条链路终结于一个端口或转发器。FC 的链路介质可以是光纤、双绞线或同轴电缆。

光纤通道是一种通用的传输通道,它能够为多种高层协议(Upper Level Protocols,ULP)提供高性能的传输通道,协议包括智能外设接口(Intelligent Peripheral Interface,IPI)命令集、小型计算机系统接口(Small Computer System Interface,SCSI)命令集或高性能并行接口(High-Performance Parallel Interface,HiPPI)数据帧、互联网协议(Internet Protocol,IP)、IEEE 802.2 等。

光纤通道是一种基于标准的网络结构。它的标准定义了物理层的特征、传输控制方法以及与 TCP/IP、SCSI-3、HiPPE 和其他一些协议的上层接口。光纤通道是一种千兆位传输技术,目前的实现支持最高可达 64Gb/s 的传输速率。

光纤通道标准定义了一个通过网络移动数据的多层结构。它的协议被划分为 5 个层次,如图 5.3 所示。

FC-0 层描述物理接口,包括传送介质、发射机和接收机及其接口。FC-0 层规定了各种介质和与之有关的能以各种速率运行的驱动器和接收机。

FC-1 层中定义了 FC 的底层传输协议,包括串行编码、解码和链路状态维护。它描述了 8B/10B 的编码规则,使控制字节与数据字节分离且可简化比特,字节和字同步,该编码还具有检测某些传送和接收误差的机制。在 FC-1 层中由几个专用字符组合在一起,并通过字符命令集来表示一定的特殊含义,如帧边界、简单传输请求或通过周期性的交互维持链路传输状态。

FC-2 层是信令协议层,它规定了需要传送成块数据的规则和机制。在所有协议层中,FC-2 层

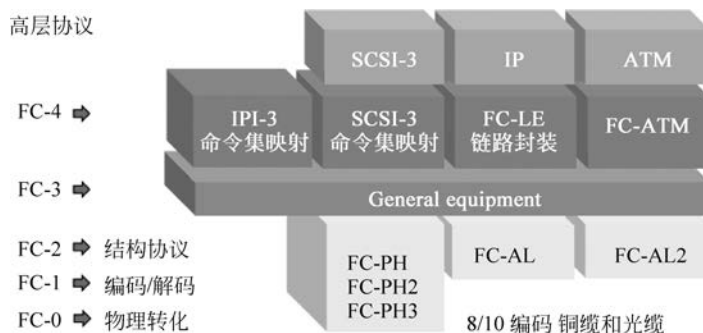


图 5.3 光纤通道协议栈

是最复杂的一层,它提供不同类型的服务、分组、排序、检错、传送数据的分段重组,以及协调不同容量的端口之间的通信需要的注册服务。

FC-3 层提供的一系列服务,是光纤通路节点的多个端口所共用的。尽管这一层没有明确定义,但是它所提供的功能适用于整个体系结构未来的扩展,例如,多路复用和地址绑定功能。

FC-4 层提供了光纤通路到已存在的更上层协议的映射,这些协议包括 IP、SCSI 协议、HiPPI 等。例如,串行 SCSI 必须将光纤通道设备映射为可被操作系统访问的逻辑设备。对于主机总线适配器,这种功能一般要由厂商提供的设备驱动程序程序来实现。

FC 协议数据帧及数据包的发送和接收是在 FC-2 层实现的,每个光纤通道帧由多个 4B 的传输字组成。一个光纤通道帧最多由 537 个传输字组成,最大传输 2148B 的数据。

3. FC 与 SCSI 协议关系

总线是计算机与存储系统间进行数据通信的主要通道,是源设备到目标设备的数据传输路径。在数据通信过程中,控制器首先向总线处理器发出请求信号,请求使用总线,该请求被接受后,控制器高速缓存开始进行数据的发送。在整个过程中,控制器占用总线,总线上连接的其他设备都无法占用总线。但是,总线具备中断功能,因此总线处理器可以随时中断传输过程并且将总线控制权转交给其他设备,以便其他拥有更高优先级的设备执行操作。

例如,将移动手机或数码相机连接到计算机时,一般使用通用串行总线(Universal Serial Bus, USB)端口。对于存储音频文件、图元文件等的小型电子设备,如 MP3 或移动手机,USB 端口已经可以满足传输数据和充电的工作。然而,USB 串行总线并不足以同时支持整台计算机和服务器以及其他多台设备的数据传输使用。在这种情况下,计算机就需要使用 SCSI 这种并行总线。

(1) SCSI 协议。

小型计算机系统接口(SCSI)是一种用于计算机及其周边设备之间(如硬盘、光驱、打印机、扫描仪等)系统级接口的独立处理器标准。如图 5.4 所示,SCSI 总线上的数据操作和管理是由 SCSI 控制器控制的,SCSI 控制器可以看成是一块小型 CPU,它有自己的命令集和缓存空间。SCSI 总线结构可以对计算机中连接到 SCSI 总线上的多个设备进行动态分工操作,并可以对系统中的多个工作灵活地进行资源分配,动态完成。

SCSI 协议是主机与存储磁盘通信的基本协议,也是计算机和外围存储设备之间进行数据传输的通用接口标准,支持并行数据传输。SCSI 协议除了被 DAS 用于实现主机服务器与存储设备的互连,也是 SAN 网络传输的基本协议,承载在 FC 协议和 iSCSI 协议中进行传输。

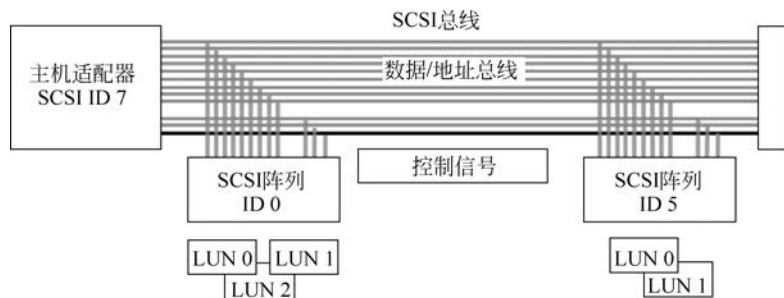


图 5.4 SCSI 总线模型

(2) 并行 SCSI 的演变。

SCSI 协议最初由 Shugart Associates 和 NCR 公司在 1981 年开发出来,并命名为 SASI (Shugart Associates System Interface)。两家公司开发 SASI 的目的是建立一个专用的、高性能的系统接口标准。后来,为了增加 SASI 在行业的接受度,将 SASI 协议更新升级成了一个更强大的接口协议,并更名为 SCSI。1986 年,美国国家标准委员会(ANSI)认可 SCSI 作为行业标准。SCSI 经历了 SCSI-1、SCSI-2、SCSI-3 的演变过程。

SCSI-1 是最初的 SCSI 标准。SCSI-1 又称为 Narrow SCSI,它定义了线缆长度、信号特性、命令和传输模式。其支持的最大数据传输率为 5MB/s,使用 8 位窄总线,最大支持接入 7 个设备。SCSI-1 是在 1986 年开发的原始规范,现已不再使用。

SCSI-2 标准是 1992 年制定的,SCSI-2 是 SCSI-1 的发展,在 SCSI-1 标准中加入一些新功能。SCSI-2 提供了两种传输选择:一种为 Fast SCSI,同步传输速率可达 10MB/s;另一种是 Wide SCSI,最大同步传输速率为 20MB/s,并且由 8 位窄总线扩展到 16 位,最大支持接入 15 个设备。

SCSI-3 是 SCSI 最新版本,也称为 UltraSCSI,由多个相关的标准组成。SCSI-3 最大支持 15 个设备的接入,最大传输速率可达 640MB/s。同时,SCSI-3 大大地提高了总线频率,降低了信号干扰,增强了数据传输的稳定性。

(3) SCSI 通信过程。

从本质上讲,SCSI 是一个智能传输协议。多个设备连接到同一组总线上的并行通信通道,这些设备可以相互进行通信。也就是说,两个连在同一组总线上的设备可以互相通信,不需要 CPU 或者特别的适配卡协助。

SCSI 协议在传输过程中经历以下 5 个阶段。

① 总线忙。在总线通信开始之前,总线必须处于空闲状态。发起连接的设备(启动器)首先会发一个测试信号来确认总线是否空闲。

② 寻址:通过发送方的地址和接收方的地址来确认通信的双方。

③ 协商:通信双方协商确定后面数据包的大小和数据包发送的速度。

④ 连接:数据包传输阶段。

⑤ 断开连接:数据传输完成,释放总线。

一旦启动器监测到总线处于空闲状态,则该启动器设备就获得了该总线的传输数据专有权,从而占有总线。然后,启动器通过寻址来确定目标器设备。由于 SCSI 协议拥有多个版本,不同版本在数据包发包速度和支持的设备地址位数等方面存在不同,因此需要两个设备之间事先协商好通信参数,协商内容包括数据包发包的速度和地址的位数等。尽管这种协商过程比较耗时,但只

有协商成功之后,启动器和目标器设备才能够进行真正的数据内容传输。

两个 SCSI 设备的每次连接通信都要经历以上 5 个阶段。由于协商阶段的时间较长,影响了整体的传输效率,但是协商阶段是必需的,只有协商成功之后,启动器和目标器设备才能进行数据传输。为了保障传输性能又能保证通信连接,下面介绍一种断开重连技术。

断开重连技术有助于缩短设备通信连接时间。采用断开重连技术时,只有在第一次连接时需要执行“总线忙→寻址→协商→连接→断开连接”这五个步骤,之后当同一个启动器跟同一个目标器进行通信数据传输时,可以直接省略协商这一步骤,这是由于双方之前在建立通信时已经协商好相关参数,因此再次连接时可以使用上一次的协商结果。

为了提升整体性能效率,SCSI 还引入了标签指令队列技术,其工作方式如下:启动器设备在发送数据时一次发送多个 SCSI 数据包,目标器设备收到这些数据包后进行内部处理,然后再将数据包的内容写到相应的物理存储介质中。当目标器设备接收到数据包并缓存到内部存储之后,立即释放其对总线的使用权,便于其他设备可以使用总线。利用标签指令队列技术,既可以减少设备间建立连接的次数,也可以减少对总线的占用次数,增加了总线的整体利用率。

(4) SCSI 数据传输原理。

如图 5.5 所示,SCSI 协议中数据传输原理具体如下。

① 当设备 B 要向设备 D 传输数据时,数据的发起端(即设备 B)以电信号的方式将数据发送出去,数据从设备 B 与总线的接入点发送到总线上。

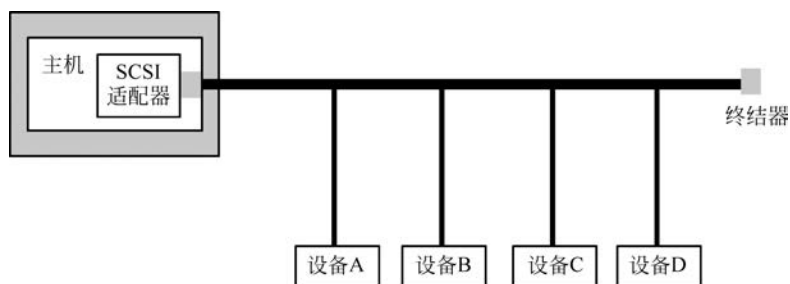


图 5.5 SCSI 协议数据传输原理

② 信号到达总线上的交叉点后被分成两份,分别朝分叉的两个方向继续传输,并且在每个交叉口分份,再沿着各个分叉的方向进行传输。因此,当承载着数据的电信号从设备 B 传输到设备 B 与总线的交叉点后,信号将分成两份,分别向总线的两个相反的方向进行传输。当信号到达设备 A 或设备 C 与总线的交叉点时,信号再次会分成两份,分别沿着总线方向和设备方向进行传输。当信号到达设备 D 与总线的交叉点时,也会以同样的方式进行传输。

③ 当设备 D 收到信号时,其中一份信号向设备 D 传输,另一方则继续向前传输到达总线的尽头。在传输的信号里,包含数据包目标设备等信息。因此,当设备 A 和设备 C 接收到这个数据包后,发现这个数据包不是传输给它们的,数据包将被丢弃。而设备 D 发现这个数据包是传输给自己的,就会接收并处理这个数据包的数据。

④ 除了设备 D 收到的数据包,还有一份数据继续往下传输并到达总线的尽头。为了避免信号被反射回总线,需要在总线的尽头安装一个终结器以吸收信号。

终结器位于 SCSI 总线的尽头,用于吸收接收到的信号以防止反射,减小信号相互影响,维持 SCSI 总线上的电平稳定。因此,每一个 SCSI 总线系统都需要安装终结器,以保证正常地进行信号传输。另外,在传输过程中会有一份信号传输到总线的另一端,即 SCSI 适配器端,与终结器一样,传输到这一端的信号同样也会被吸收,防止信号被反射回去。

(5) SCSI 数据传输方式。

SCSI 通信传输方式有两种：异步方式和同步方式。

在异步传输方式下,两组数据传输之间没有固定的时间间隔。SCSI 协议发送额外的信息或者命令来发起通信。例如,在发送内容数据前,首先由发起方发送状态信息,接收方根据状态信息获悉马上要发送内容数据。额外状态信息或者命令的发送时间可以不是固定的,因此,内容数据之间的传输间隔也可以不是固定的。这是异步传输方式的重要特点。

同步传输方式下,数据包会按照定时器设定的时间间隔进行传输。首先,通信双方通过异步方式来确定对方设备是否已经准备好接收数据;建立连接之后,通信双方会采用最高效的传输方式进行实际内容数据的传输,这种方式就是同步传输方式。在同步传输方式中,发起方发送数据的时间间隔是固定的,而接收方知道这个时间间隔,就能快速地接收和处理数据。

(6) SCSI ID。

在 SCSI 总线体系结构中,任何连接到总线的设备都可以互相通信。为实现这一点,信号从发送端设备发出后,最终会在多点总线(多分支总线)上结束。在多点总线上,信号将被传输到目标设备上。在这个通信流程中,需要保证总线上的多个并发用户不会互相干扰。如果总线上多个设备同时发送信息就会产生线路拥塞,发生线路拥塞时,多个设备发送的信息之间会互相冲突、干扰,最终导致所有设备的发送操作都不成功,而发送端必须重新进行发送,这就导致数据的发送效率变低,因此需要保证连接到总线上的多个设备不会同时发送信息。

为了保证 SCSI 总线系统同一时刻在整条总线上只有一个设备在发送信息,开发者们设计出了一种带优先级的等候机制,即总线上的每个 SCSI 设备都有不同的优先级。SCSI ID 用于唯一标识总线上的设备,即标识着数据的发送方和接收方,这里采用设备的 SCSI ID 来标记设备的优先级,用于决定每个设备在检测到总线忙时需要等待多久再尝试发送数据。

如果是 8 位窄线,则优先级从高到低为 $7 > 6 > 5 > 4 > 3 > 2 > 1 > 0$ 。

如果是 16 位窄线,则优先级从高到低为 $7 > 6 > 5 > 4 > 3 > 2 > 1 > 0 > 15 > 14 > 13 > 12 > 11 > 10 > 9 > 8$ 。

无论是 8 位窄线还是 16 位窄线,能连接的设备数都是 $n-1$ (n 表示总线宽度),其中一位被 SCSI 控制器占用。由于控制器需要控制整条总线,因此控制器的优先级必须是最高 7。

当一个设备需要发送数据时,它必须要检测总线是否在忙,即是否有另外一个设备正在发送数据。当设备检测到总线在忙(例如,有其他设备正在发送数据),它就需要等待一定的时间再尝试发送。这个等待的时间长短是由其 SCSI ID 决定的。设备的优先级越高,则它等待的时间就减少,因而,在等待时间结束后,能够成功发送数据的概率就越大。通常来说,总线上的设备中,速度快的设备(如硬盘)比速度慢的设备(如磁带库)拥有更高的优先级。

(7) SCSI 协议寻址。

在 SCSI 总线的通信过程中,除了保证总线上的多个并发用户不会互相干扰之外,还要组织通信流程,以保证数据最终到达总线上正确的目标设备。借助 SCSI 寻址,将信息准确无误地发送到正确的目标设备上。上文提过,SCSI ID 用于唯一标识总线上的设备,但是,仅依靠 SCSI ID 是不能寻找到目标设备的。如图 5.6 所示,SCSI 总线的寻址过程是通过总线号(Bus ID)—设备号(SCSI ID)—逻辑单元号(LUN ID)来实现的。

总线号用于区分每一条总线。传统的 SCSI 适配卡连接单个总线,相应地,只有一个总线号。一个服务器可能配置了多个 SCSI 控制器,从而可能有多条 SCSI 总线。

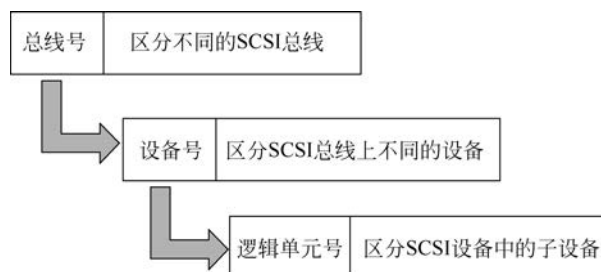


图 5.6 SCSI 协议寻址内容

设备号用于识别某 SCSI 总线上的每一个存储设备,每条总线最多可允许有 8 个或者 16 个设备号。服务器中的主机总线适配器也拥有独立设备号。

逻辑单元号用于识别某存储设备上的每一个子设备。子设备包括虚拟磁盘、磁带驱动器和介质更换器等。利用逻辑单元号,可以对存储设备中的子设备进行寻址。

(8) FC 与 SCSI 协议对比。

FC 通道并不是替代 SCSI 的,FC 可以通过构建帧来传输 SCSI 指令、数据和状态信息单元。SCSI 是位于光纤通道协议栈 FC-4 的上层协议,SCSI 是 FC 协议的子集,FC 与 SCSI 协议之间的关系如图 5.7 所示。

4. FC 典型组网拓扑和连接设备

FC(Fibre Channel)可称为 FC 协议,或者 FC 网络、FC 互联。由于其性能较高,逐渐发展到前端作为主机接口,并逐渐发展出点对点、交换机等组网。存储中的 FC 如图 5.8 所示。



图 5.7 FC 与 SCSI 协议之间的关系

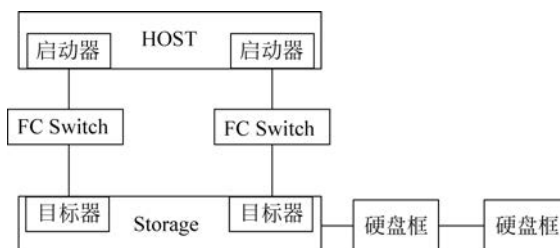


图 5.8 存储中的 FC

存储网络引入 FC 后,具有的优势:提高扩展性,增加传输距离,解决安全性问题。

光纤通道的层次基本上相当于 OSI 参考模型的较低层,并且可以看成是链路层的网络。光纤通道呈现为单个不可分割的网络,并在整个网络中使用统一的地址空间。虽然在理论上这个地址空间可以非常大,在单个网络中可以有千万个地址,但实际上光纤通道通常在一个 SAN 中只支持数十台设备,或者在某些大型数据中心应用中支持上百台设备。

光纤通道用拓扑结构来描述各个节点的连接方式。光纤通道术语中的“节点”是指通过网络进行通信的任何实体,而不一定是一个硬件节点,这个节点通常是一个设备,如一个磁盘存储器、服务器上的一个主机总线适配器或者是一个光纤网交换机。如图 5.9 所示,列出三种光纤通道的拓扑结构。

点对点(Fibre Channel Point-to-Point): 两个设备直接连接到对方,这是最简单的一种拓扑,

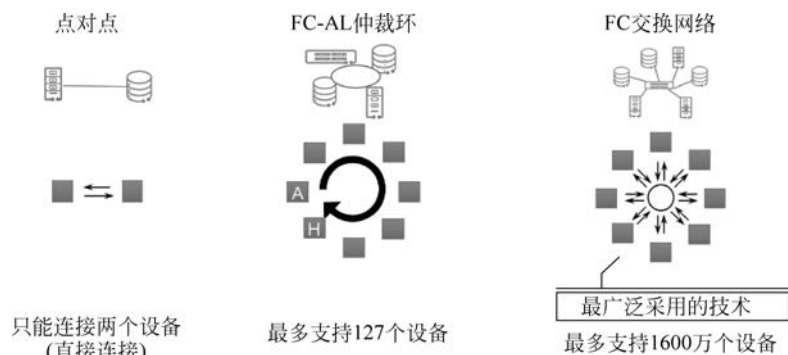


图 5.9 光纤通道的拓扑结构

连接能力有限。

FC-AL 仲裁环(Fibre Channel Arbitrated Loop)：在这种连接方式中，所有设备连接在一个类似于令牌环的环路上。在这个环路中添加或者移除一个设备会导致环路上所有活动中断。环路上一个设备的故障将导致整个环路不能进行工作。通过添加光纤通道集线器的方法，能够将众多设备连接到一起，形成一个逻辑上的环路，并且能够旁路故障节点，使得环上节点的故障不会影响到整个环路的通信。仲裁环曾经用于小型的 SAN 环境中，但是现在已经不再使用，其原因在于一个仲裁环最多只能容纳 127 个设备，而现在 SAN 环境中使用的设备基本上都多于 127 个设备。

FC 交换网络(Fibre Channel Switched Fabric)：这是构建现代 FC SAN 所采用的连接方式。它使用 FC 交换机连接主机和存储设备。在现代的 SAN 中，最好使用两个交换机来连接主机和存储设备，这样可以形成链路冗余，增强 SAN 的可靠性。

光纤通道既支持光纤介质，也支持铜缆介质。由于光纤介质对噪声不敏感，用它来作传输介质是最好的，例如，光缆具有如下优点：可以达到更长的距离，对电磁干扰不敏感，无电磁辐射，在设备之间无电连接和无交叉干扰的问题。但是，铜介质也得到了许多的使用，尤其是对小型光纤通道磁盘驱动器的连接，原因在于与光缆相比，铜缆较为便宜。

光缆及其接插件也有多种不同的类型。用于长距离传输的光缆比用于短距离的光缆更为昂贵。

人们通常用模来区分光纤类型。多模光纤使用短波激光，其纤芯直径为 50 μm 或 62.5 μm ，包层直径为 125 μm 。其中，纤芯为光通路，包层用来把光线反射到纤芯上。由于短波激光流是由数百种模(即所传输的光波的波长)组成的，它们在光纤内以不同的角度发生全反射，因此称为多模。光的散射效应限制了原始信号所能达到的总长度。如表 5.2 所示，可以看到多模光纤的最长距离是 500m。

表 5.2 常用的光纤连接介质

介质类型	发射器	速率	距离
9 μm 单模光纤	1550nm 长波光激光器	1Gb/s	2m~50km
		2Gb/s	2m~50km
	1300nm 长波光激光器	1Gb/s	2m~10km
		2Gb/s	2m~2km
		4Gb/s	2m~2km

续表

介 质 类 型	发 射 器	速 率	距 离
50μm 多模光纤	850nm 短波光激光器	1Gb/s	0.5~500m
		2Gb/s	0.5~300m
		4Gb/s	0.5~170m
62.5μm 多模光纤		1Gb/s	0.5~300m
		2Gb/s	0.5~150m
		4Gb/s	0.5~70m

不管是光纤介质还是铜缆介质,光纤通道都要求误码率达到每传递 10^{12} b 不超过 1b 错。这就意味着对于一条 1000Mb/s 的连接,在全负荷情况下平均每 16.6min 最多可以发生 1 位错。这一误码率要求也适用于光纤通道中所有的部件,如中继器和交换机。

光纤通信的原理:在发送端首先把要传送的信息(如话音)变成电信号,然后调制到激光器发出的激光束上,光的强度随电信号的幅度(频率)变化而变化,光通信利用全反射原理,当光的注入角满足一定的条件时,光便能在光纤内形成全反射,从而达到长距离传输的目的。基于光射线在纤芯和包层界面上的全反射,使光线限制在纤芯中传输。如果光以一个不正确的角度射到界面,光将离开包层,那么,这部分光信号将丢失。这意味着,由于光不太亮,导致信号微弱,最后光电传感器不能检测到。光纤中有两种光线,即子午光线和斜射光线,子午光线是位于子午面上的光线,而斜射光线是不经过光纤轴线传输的光线。

从上述光纤通信原理可知,光纤传输对线缆是有要求的,因此处理线缆是非常重要的。例如,工程师在处理光纤线时,不能将光纤线弯曲太多,并且需要保持线缆两端的收发器没有灰尘。很多原因会导致光信号的衰减,需要在使用中注意,这些原因如下。

- (1) 最大弯曲度:光纤路径的弯曲度即使与规格有微小的差异,也会导致信号失真。
- (2) 最小弯曲度:将光纤线扎得过紧会导致信号丢失。
- (3) 散射:杂质有不同的折射率。当光通过杂质时,光将发生散射。
- (4) 吸收:当光以不合适的角度射到包层时,在包层光将被吸收。

光纤收发器,通常又称为 FC 光模块,光纤收发器就是光纤发射器加上光纤接收器,包含一个激光器或者发光二极管以创建光脉冲,包含探测光的一个光学传感器,如图 5.10 所示。光纤收发器存在于存储设备、交换机和服务器主机总线适配器(Host Bus Adapter, HBA)卡上,可以单独地移除和更换。



图 5.10 FC 光模块

主机总线适配器,是一种能插入计算机、服务器或大型主机的板卡或集成电路适配器,在服务器和存储设备之间提供 I/O 处理和物理连接。支持连接 FC 网络的应用,实现高带宽高性能存储组网方案。光纤通道 HBA 卡是将主机接入 FC 网络必不可少的设备,如图 5.11 所示。HBA 能减轻主处理器在数据存储和检索任务方面的负担,提高服务器的性能。



图 5.11 HBA 卡

存储设备的 FC 接口模块与 FCoE 接口模块,如图 5.12 所示。

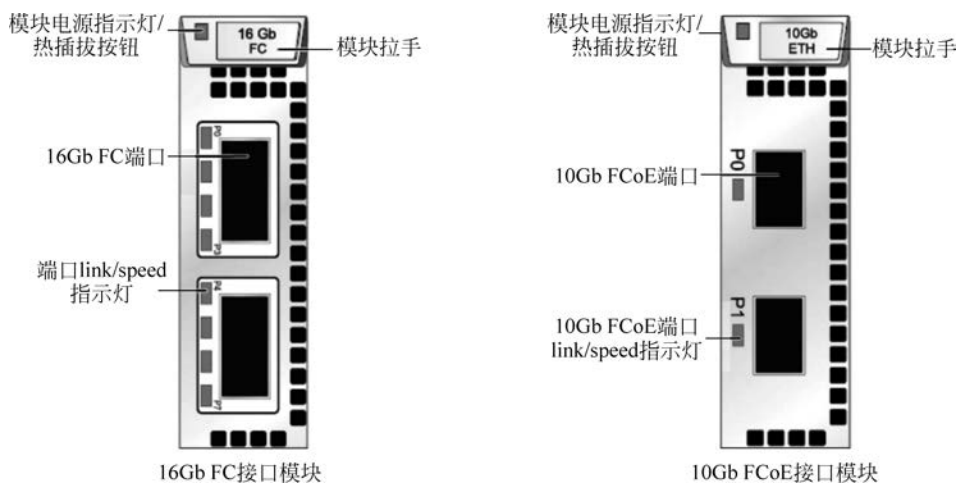


图 5.12 FC 接口模块与 FCoE 接口模块

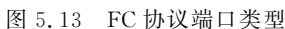
5. FC 端口

在所有的拓扑结构中,设备(包括服务器、存储设备和网络连接设备)都必须配置一个或多个光纤通道端口。在服务器上,端口一般借助主机总线适配器实现。一个端口总是由两个通道构成:一个是输入通道;另一个是输出通道。在两个端口之间的连接称作链路。在点到点和交换网拓扑中,链路总是双向的。在交换网的情况下,链路所涉及的两个端口的输出通道和输入通道通过一个交叉装置连接在一起,使得每一个输出通道都连接到一个输入通道。另外,仲裁环拓扑的链路是单向的,每个输出通道都连接到下一个端口的输入通道,直到圆周闭合为止。仲裁环的线缆连接可以借助一个集线器简化,形成星状环,此时,端点设备双向连接到集线器,在集线器内部的线缆连接保证在仲裁环内部维持单向的数据流。FC 协议通过对端口标识,即端口标识符进行寻址并在不同设备间互连时定义了不同的接口类型。

如图 5.13 所示,根据不同的功能,如自适应式管理员设定可以将 FC 端口分为以下几种类型。

N 端口(节点端口): 光纤通道通信是围绕 N 端口和 F 端口开发的,这里的 N 表示 Node(节点),F 表示 Fabric(交换网)。N 端口描述一个端点设备(服务器,存储设备),也称节点,具有加入交换网拓扑或点到点拓扑的能力。

F 端口(交换端口): F 端口是 N 端口在光纤通道交换机中的对接点。F 端口知道怎样把一个



NL 端口(节点和环端口): NL 端口同时具有 N 端口和 L 端口的能力。一个 NL 端口既可以连到一个交换网,也可以连到一个仲裁环。

E 端口(扩展端口): 两个光纤通道交换机通过 E 端口连接在一起。E 端口使得连接到两个不同交换机的端点设备可以互传数据。光纤通道交换机通过 E 端口在整个光纤通道网络上中转信息。

在一个光纤网络环境中,可能会有成千上万的组件,所以必须用一个唯一的标识符来识别每一个设备。光纤通道使用的标识符是全球唯一名称(World Wide Name, WWN)。光纤网络里,每一个设备(包括光纤通道兼容的设备)的唯一标识就是 WWN,用于标识存储设备中 I/O 模块的单个接口。如图 5.14 所示,WWN 有两种不同的定义: WWNN 和 WWPNN。



全球唯一端口号(World Wide Port Number, WWPN)是分配给每一个光纤通道端口的一个唯一的 64 位标识符。每个 WWPN 被该端口独享,WWPN 在存储区域网络中的应用就等同于 MAC

地址在以太网协议中的应用。

6. FC 分区

光纤通道协议提供有安全机制,例如,它支持分区(Zone)功能,可以规定不同分区的设备能否交互。一个设备节点或 WWN 处于单独的分区,也可以同时处于多个分区之中。

FC 分区有两种类型:软分区和硬分区,如图 5.15 所示。

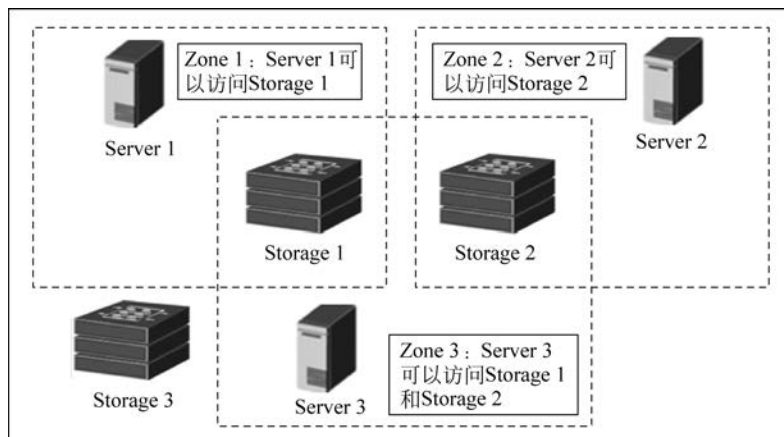


图 5.15 FC 分区

软分区意味着交换机将设备的 WWN 分配到一个区之中,这和设备所连接的端口无关。例如,如果 A 和 B 在同一个软分区内,它们就能够彼此互访。同理,如果 A 和 C 处于另一个软分区之中,此时,A 和 B 就能够看到对方,但是 A 看不到 C。软分区要依赖光纤通道中节点的 WWN。通过使用软分区,可以连接交换机上的任何端口,可以看到应该看到的其他节点。

硬分区和以太网里的 VLAN 类似。当一个端口被分配到一个分区里,任何连接到这个端口的设备都属于这个分区。为了防止有人破坏光纤连接,在交换机上进行硬分区时,将 WWN 和目标上的 LUN 地址绑定。通过使用存储阵列的 WWN masking,多个 Initiator 都能够看到这个 Target。

5.2.3 IP SAN 概述

早期 SAN 网络采用光纤通道进行块数据传输,因此早期 SAN 指的是 FC SAN。在实际应用中,如果企业要使用 SAN 网络进行数据存储,需要购买 FC SAN 存储网络相关的设备组件,其昂贵的价格和复杂的配置限制了中小型企业,尤其是小型企业的部署使用。因此,为了提高 SAN 存储网络的使用,满足中小型企业的请求,工程师们提出并设计了 IP SAN 方案。

IP SAN 指基于 IP 传输的网络存储系统,其使用标准的 TCP/IP,可在以太网上进行块数据的传输,无须配置专门的 FC 网络。如图 5.16 所示为 IP SAN 的拓扑结构。

1. IP SAN 网络架构的优缺点

IP SAN 具有如下优点。

(1) 接入标准化。IP SAN 的部署不需要专用的光纤

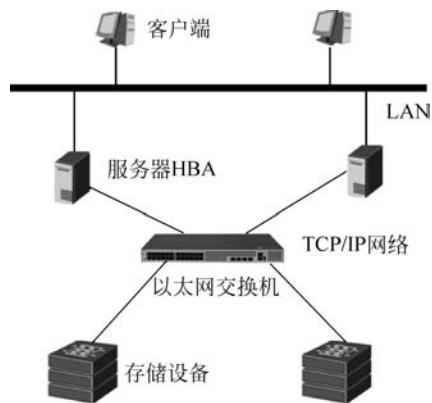


图 5.16 IP SAN 的拓扑结构

HBA 卡和光纤交换机,可直接利用现有网络中的以太网卡和以太网交换机。

(2) 传输距离远。只要 IP 网络可达的地方,就可以部署 IP SAN 存储网络。

(3) 可维护性好。IP 网络的维护工具非常发达,具有较多的专业技术人员支持。

(4) 带宽扩展方便。iSCSI 承载于以太网,现在以太网已经发展为 10Gb/s 速率,目前已知的最高速率可达 400Gb/s。

(5) 成本低。整体降低产品的总体拥有成本。

IP SAN 的缺点如下。

(1) 数据安全性。数据在 IP SAN 网络中传输时,尽管 IP 协议可以应用 IPSec 以保障数据的安全性,但也只能提供数据在网络传输过程中的动态安全性,并不能保证数据被保存在存储设备上的静态安全性。另外,使用 IP 网络构建的 IP SAN 和传统的 IP 业务很难从物理上完全隔离,而 IP 网络是开放式网络,仍然存在众多安全漏洞,这对 IP SAN 也构成安全性威胁。

(2) TCP 负载空闲引擎。由于 IP 是无连接不可靠的传输协议,数据的可靠性和完整性是由 TCP 来提供的,而 TCP 为了完成数据的排序工作需要占用较多的主机 CPU 资源,导致用户业务处理延迟的增加。

(3) 占用 IP 网络资源。由于 IP SAN 是直接部署在现有的网络资源上,而 IP 网络尤其是以太网络的效率和 QoS 都较低,因此 IP SAN 网络将占用系统资源。

2. IP SAN 组网形式

IP SAN 组网形式可分为直连组网、单交换组网和双交换组网。

(1) 直连组网。

主机与存储设备之间直接通过以太网卡、iSCSI HBA 卡连接,如图 5.17 所示。

直连组网方式具有构建简单、经济省钱的优点,但存在主机存储资源分享比较困难的问题。

(2) 单交换组网。

主机与存储设备之间通过一台以太网交换机进行通信,同时主机安装以太网卡、iSCSI HBA 卡实现连接,如图 5.18 所示。

单交换组网结构使多台主机能共同分享同一台存储设备,扩展性强,但存在单点故障问题。



图 5.17 直连组网



图 5.18 单交换组网

(3) 双交换组网。

主机与存储设备之间通过两台以太网交换机进行通信,同时主机安装以太网卡、iSCSI HBA 卡实现连接,如图 5.19 所示。同一台主机到存储设备端由多条路径连接,可靠性强,避免了在单

交换组网中以太网交换机处存在的单点故障。

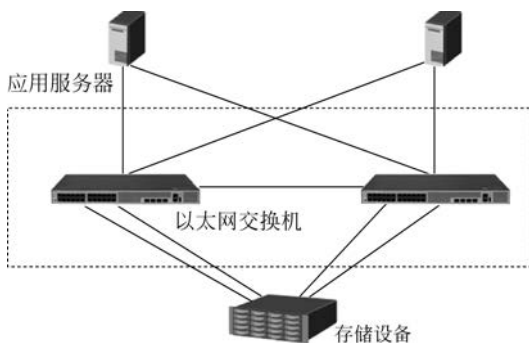


图 5.19 双交换组网

3. FC SAN 和 IP SAN 对比

下面从网络速度、网络架构、传输距离、管理维护、兼容性、性能、成本、容灾、安全性等方面对 FC SAN 和 IP SAN 进行分析和对比。

- (1) 网络速度。FC SAN 支持 4Gb/s、8Gb/s、16Gb/s；IP SAN 支持 1Gb/s、10Gb/s。
- (2) 网络架构。FC SAN 需要单独建设光纤网络和 HBA 卡；IP SAN 可直接使用现有 IP 网络。
- (3) 传输距离。FC SAN 的传输距离受到光纤传输距离的限制；IP SAN 理论上没有距离限制，只要 IP 网络可达的地方，都能部署。
- (4) 管理维护。FC SAN 技术和管理较复杂；IP SAN 的管理维护与 IP 设备一样操作简单。
- (5) 兼容性。FC SAN 的兼容性差；IP SAN 与所有 IP 网络设备都兼容。
- (6) 性能。FC SAN 具有非常高的传输和读写性能；IP SAN 目前主流 1Gb、10Gb 正在发展。
- (7) 成本。FC SAN 网络的搭建需要购买光纤交换机、HBA 卡、光纤磁盘阵列等，同时需要培训人员、系统设置与监测等，成本高；IP SAN 购买与维护成本都较低，有更高的投资收益比例。
- (8) 容灾。FC SAN 搭建容灾的硬件、软件成本都比较高；IP SAN 本身可以实现本地和异地容灾，且成本低。
- (9) 安全性。FC SAN 与传统业务 IP 网络从物理上隔离，保证了 SAN 网络下传输和存储的数据安全性；IP SAN 网络中，尽管 IP 协议可以应用 IPSec 以保障数据的安全性，但只能提供数据在网络传输过程中的动态安全性，并不能保证数据在存储设备上的静态安全性。由于 IP 网络是开放式网络，仍然存在众多安全漏洞，这对于使用传统 IP 网络构建的 IP SAN 是一个安全威胁。

5.2.4 iSCSI 协议栈

SCSI 允许连接设备数量较少，SCSI 连接设备距离非常有限，基于 IP 网络的 iSCSI 诞生。iSCSI 是层次协议，iSCSI 节点将 SCSI 命令描述符块 (Command Descriptor Block, CDB) 和数据封装成 iSCSI 包协议数据单元 (Protocol Data Unit, PDU) 后传送给 TCP/IP 层，再由 TCP/IP 将 iSCSI 封装成 IP 数据，然后发送到以太网上进行传输，如图 5.20 所示。

所有的 SCSI 命令都被封装成 iSCSI 协议数据单元 PDU，利用 TCP/IP 协议栈中传输层的 TCP 封装 TCP/IP 数据包，提供可靠的传输机制。



图 5.20 iSCSI 协议栈

1. iSCSI 协议栈工作原理

iSCSI 架构是基于 C/S 模型进行数据传输的,如图 5.21 所示,在 iSCSI 系统中,用户在一台 SCSI 存储设备上发出 I/O 请求,操作系统对该请求进行处理,并将该请求转换成 SCSI 指令,再传给目标 SCSI 控制卡。iSCSI 节点将指令和数据封装形成一个 iSCSI 包,然后将数据单元传送给 TCP/IP 层,由 TCP/IP 将 iSCSI 包封装成 IP 数据,以适合在网络中传输。此过程也可以对封装的 SCSI 命令进行加密处理,以保证在不安全的网络上传送的安全性。

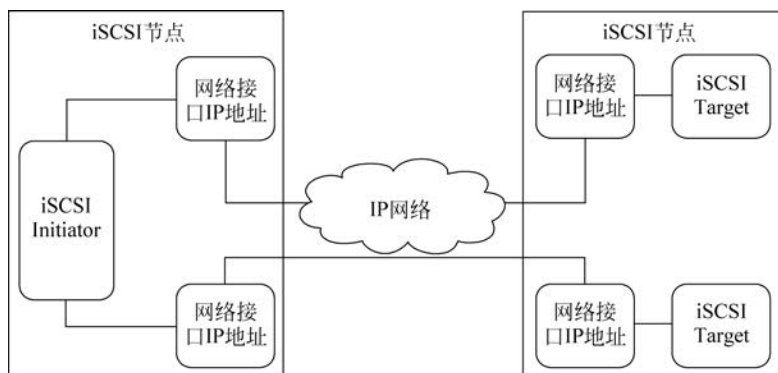


图 5.21 iSCSI 体系结构

由于 iSCSI 的通信体系仍然继承了 SCSI 的部分特性,因此在 iSCSI 通信中,需要一个发起 I/O 请求的启动器和一个响应并执行实际 I/O 操作的目标器(Target)。在 Initiator 和 Target 建立连接后,Target 在操作中作为主设备控制整个工作过程。

iSCSI 具体的传输过程如图 5.22 所示。

发起端(Initiator): SCSI 层负责生成命令描述符块 CDB,将 CDB 传给 iSCSI 层,iSCSI 层负责生成 iSCSI 协议数据单元 PDU,并通过 IP 网络将 PDU 发给 Target。发起端可以是软件 Initiator 驱动程序、硬件的 TCP 卸载引擎(TCP/IP Offloading Engine, TOE)网卡,或者是 iSCSI HBA 卡等。

目标器(Target): 与发起端的处理流程相反,Target 端点通过 TCP/IP 收到 iSCSI PDU 后,执行解封装操作,并将解封后的 CDB 传给 SCSI 层,SCSI 层负责解释 CDB 的意义,必要时发送响应,目标器可以是磁盘阵列、服务器上的硬盘或磁带库。

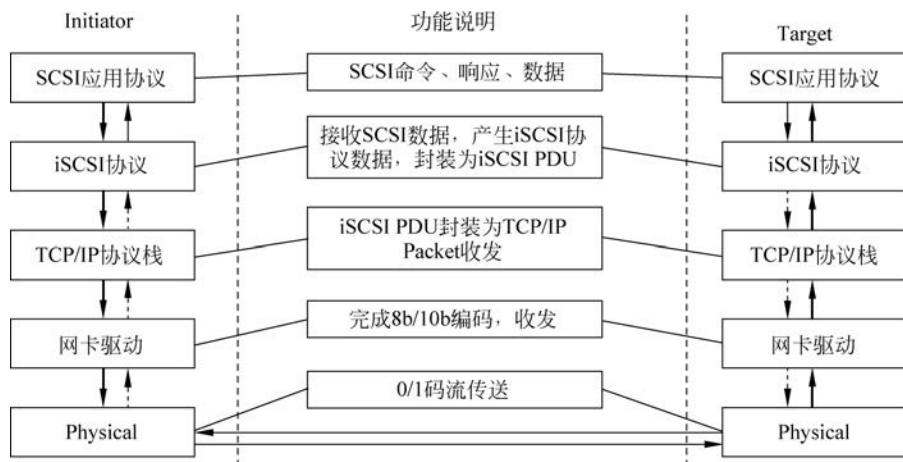


图 5.22 iSCSI 传输过程

2. iSCSI 存储设备与主机连接方式

实现 IP SAN 的典型协议是 iSCSI,它定义了 SCSI 指令集在 IP 网络中传输的封装方式。iSCSI(Internet SCSI)是因特网工程特别工作组在 2003 年制定的一项标准,iSCSI 协议是建立在 SCSI 协议和 TCP/IP 基础上的标准化协议,用于将 SCSI 数据块封装成 IP 数据包,并在以太网中进行传输。

iSCSI 设备一般采用 IP 接口作为主机接口,连接到以太网交换机,构建一个基于 TCP/IP 的 IP SAN 存储网络。根据主机端所采用的不同连接方式,iSCSI 设备与主机的连接有以下三种形式。

(1) NIC+Initiator 软件。

如图 5.23 所示,主机使用标准的以太网卡(Network Interface Card,NIC)与网络连接。由主机 Initiator 软件完成 iSCSI 报文到 TCP/IP 报文转换。NIC+Initiator 软件的方式直接使用传统主机系统通用的 NIC 卡,其成本最低。但由于 iSCSI 协议和 TCP/IP 处理需要占用主机 CPU 资源,降低了主机系统性能。

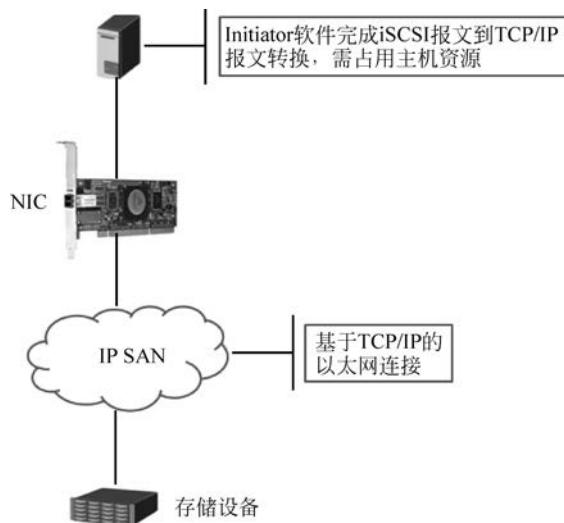


图 5.23 NIC+Initiator 软件

(2) TOE+Initiator 软件。

如图 5.24 所示,在 TOE+Initiator 软件方式下,主机使用 TCP 卸载引擎(TOE)来专门处理 TCP/IP 转换,而 iSCSI 报文的转换由主机的 CPU 实现,与 NIC+Initiator 软件的方式相比,TOE+Initiator 软件将主机 CPU 实现 TCP 处理功能下放给 TOE 网卡,降低了主机的运行开销,同时提高了数据的传输速率。

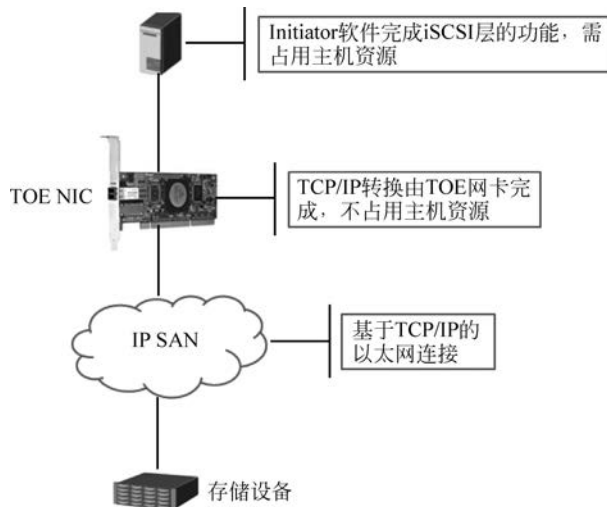


图 5.24 TOE+Initiator 软件

(3) iSCSI HBA。

如图 5.25 所示,在主机上安装一个 iSCSI HBA 卡,iSCSI HBA 卡完成 iSCSI 报文转换和 TCP/IP 报文转换功能,最大限度释放主机 CPU 资源,使得 IP SAN 操作对主机的开销占用最小,相比上述两种实现方式,iSCSI HBA 方式能获得最好的传输性能。而 iSCSI HBA 方式的代价是系统构建成本高。

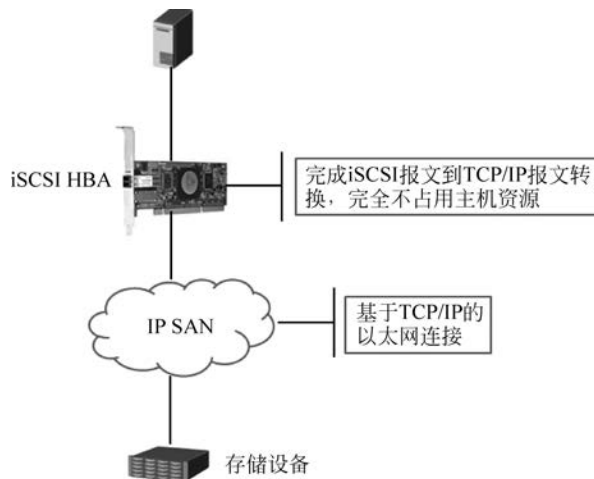


图 5.25 iSCSI HBA

3. iSCSI 应用场景

iSCSI 构建的 IP SAN 存储网络广受中小企业的欢迎,因为中小企业大都以 TCP/IP 为基础构

建网络环境,iSCSI可以直接部署在IP网络上,降低搭建成本。

基于iSCSI存储技术的IP存储主要用于解决远程存储问题,实现异地间的数据传输,两个典型的iSCSI应用场景是异地数据交换和异地数据备份。

华为大多数的存储设备都支持iSCSI协议。以OceanStor S5500存储阵列为例,其支持两种不同传输性能的iSCSI接口模块,包括1Gb/s和10Gb/s。

(1) 1Gb/s iSCSI接口模块提供了存储设备接收应用服务器读写请求的服务。每一个1Gb/s iSCSI接口模块有4个端口用于接收应用服务器发出的数据交换命令。

(2) 10Gb/s TOE接口模块提供了存储设备接收应用服务器读写请求的服务。每一个10Gb/s TOE接口模块有4个端口用于接收应用服务器发出的数据交换命令。

5.2.5 FCoE 概述

以太网光纤通道(Fibre Channel over Ethernet, FCoE)是由美国国家标准委员会ANSI定义的一种融合网络技术,是以光纤通道FC存储协议为核心的I/O整合方案。FCoE是将FC帧封装到以太网帧中,以实现在以太网基础设施上传输光纤信道信号的功能。

1. FCoE 产生背景

通常情况下,数据中心运行的网络包括前端业务网络 and 后端存储网络。前端业务网络通常是以太网网络Ethernet,用于客户端到服务器、服务器到服务器的通信;后端存储网络可以是FC SAN,也可以是IP SAN,用于服务器和存储设备、存储设备和存储设备之间的通信。一个数据中心运行多个独立的网络,服务器需要为每种网络配置单独的接口,包括连接以太网的网络接口卡NIC和连接光纤通道网络的主机总线适配器HBA。

多个独立网络并存的设计方案满足了数据中心的性能追求,同时也带来了一系列问题:首先,数据中心服务器需为每种网络专门配置一块甚至多块HBA卡,每种网络需要部署专用的交换机、线缆等硬件设备,投资成本高;其次,多个网络相互独立,彼此隔离,管理维护过程复杂,需要更多的人员运行维护,增加了人才培养成本投入;再次,服务器部署难度大,多物理接口卡造成软硬件之间耦合性强,削弱了业务灵活性,造成业务迁移复杂、困难;最后,存储网络和业务网络相互独立这种设计方式,难以充分利用以太网的扩展性的同时保留光纤通道的高可靠性和高效率,同时,相隔甚远的两个存储局域网也难以通信交互。

FC通道和以太网各有其优势,加之以太网迅速发展,人们自然想到把两种网络相融合,于是工程师提出了FCoE的设计构想,即将FC帧封装到以太网帧中借助以太网链路进行传输。FCoE技术的产生,极大程度上降低了数据中心基础设施和运行维护的投资成本,实现了数据中心在以太网和FC基础设施的无缝互通,使用户享受融合网络带来的优势。

FCoE协议是指在增强型以太网基础设施上传输光纤信道信号功能的规范,它目前已被大部分网络和存储供应商支持,FCoE将光纤通道协议映射到以太网上,这是一个独立的以太网转发协议,被视为极具应用前景的新一代存储区域网协议。当一个生产系统的前端业务网和后端存储网络融合后,存储网络和以太网可共享同一个单一的、集成的网络基础设施,实现不同类型网络的共存和网络基础设施精简整合的目标。

2. FCoE 的优点

FCoE的优点如下。

（1）精简网络结构，增强业务灵活性。

在存储网络 SAN 和业务网络 LAN 分开部署的情况下，组网比较复杂。当现有 LAN 网络中的主机需要增加 FC SAN 进行存储连接时，必须对主机进行停机，以另外部署 FC 交换机、光纤线缆以及在主机上安装 FC HBA 卡，这种双重组网部署结构相对复杂而且增加了网络管理难度。由于两种网络相互独立，业务运作时主机需要通过不同的网络进行客户端的交互和存储的访问，对于没有接入存储网络的主机而言，无法访问到 SAN 中的存储设备，业务灵活性受到了一定的限制。

而在 SAN 和 LAN 网络融合部署的情况下，组网结构相对简单。当现有 LAN 网络中的主机需要进行存储连接时，直接通过 FCoE 交换机将流量发送给 SAN 中的存储阵列即可；与此同时，基于以太网的业务数据同样可以通过 FCoE 交换机传送至 LAN 上的客户端。由此可见，FCoE 组网的部署简化了生产系统的组网结构，不仅网络设备数量减少，而且管理和维护也变得方便。此外，原本以太网和 FC 网络领域的架构依然可以延续，连入融合网络中的所有的服务器，既能与客户端交互通信，又能访问存储设备，特别是在虚拟机迁移的应用场景下，可为虚拟机提供一致的存储连接，提高了系统的灵活性和可用性，增强了业务灵活性。

FCoE 交换机包含传统的 FC 模块与接口、增强型以太网接口，可以实现传统以太网设备、传统 FC 设备和 FCoE 设备之间的交互通信。FCoE 卡也叫 CNA 卡，即硬 FCoE HBA 卡，是 FC HBA 卡和以太网卡的融合网卡，通过卡上的控制芯片，可以实现 FCoE 协议、FC 协议和以太网协议的处理。

（2）节约资源，降低成本。

LAN 和 SAN 网络通过 FCoE 技术共享网络资源，更有效地整合和利用分散的资源。双重网络分开部署情况下，需要投入以太网卡、以太网线缆、以太网交换机、FC HBA 卡、FC 线缆和 FC 交换机，所有的网络设备都要双重部署，而利用 FCoE 技术进行融合网络部署的情况下，只需投入 FCoE 卡、FCoE 交换机和以太网线缆即可。网络的融合不仅减少了网络基础设施的投资，而且简化了网络复杂度，降低了网络的管理和维护成本；同时，服务器采用融合网络适配器，一定程度上减少了生产系统的电力和冷却成本；FCoE 可以和现有的以太网及 FC 基础设施无缝互通，现有网络设施投资得到了保护。

（3）兼备以太网的扩展性，保留光纤通道的高可靠性。

FCoE 技术实现了在增强型以太网基础设施上传输光纤信道信号的功能，获得了光纤通道存储网络所具有的高性能和高可靠性优势，达到了将存储网络融入以太网架构的目标。FCoE 依然可以提供标准的光纤通道原有的多种服务，而且这些服务都可以按照原有标准运作，保有 FC 网络的低延迟性、高性能和高可靠性等特点，为服务器提供访问存储设备的后端存储网络。FCoE 采用增强型以太网作为物理网络传输通道，可以传输以太网数据帧，可以为前端的业务提供数据传输通道。FCoE 并不是要代替传统的光纤通道技术，而是对光纤通道进行拓展。如图 5.26 所示为多台 FCoE 交换机组网情况。

3. FCoE 协议栈

光纤通道和以太网都是使用数据链路层协议在网络节点间进行数据帧传输的，如果要实现将 FC 帧封装在以太网帧中通过成熟的以太网络来完成终端到终端的数据传输，必须要有相关协议的支撑，即 FCoE 协议。

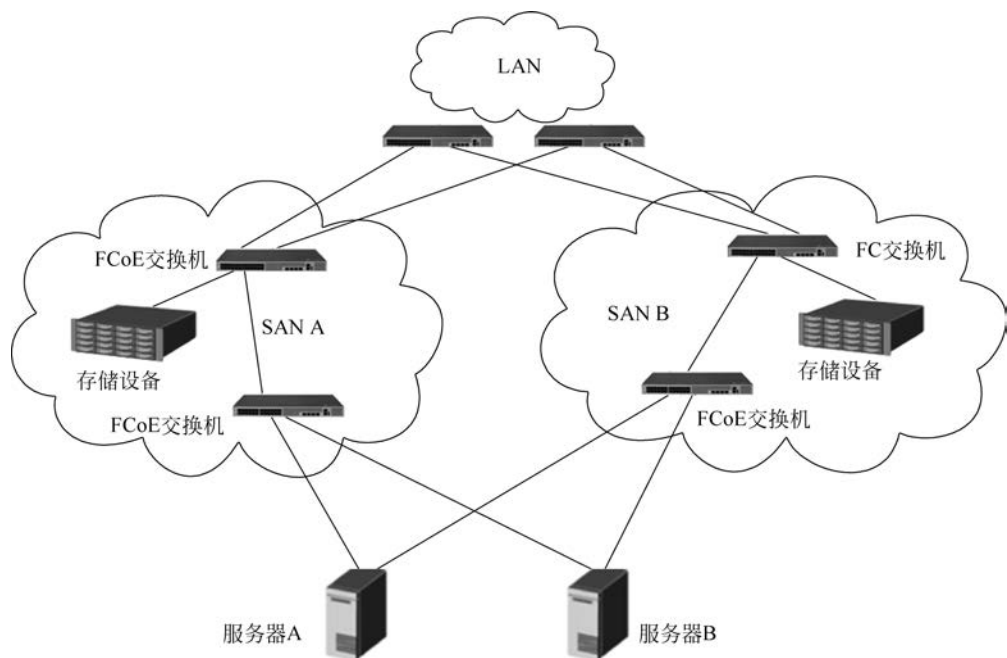


图 5.26 多台 FCoE 交换机组网

开放系统互连(Open System Interconnect,OSI)参考模型把网络通信的工作分为 7 层,从下往上分别是物理层、数据链路层、网络层、传输层、会话层、表示层和应用层。物理层是传输网络信号的物理媒介,在设备之间传输比特流,规定了电平、速度和电缆针脚;数据链路层是帧协议层,将比特组合成字节,再将字节组合成帧,使用链路层地址(以太网使用 MAC 地址)来访问介质,并进行差错检测。在 FCoE 协议栈中,FC-0 和 FC-1 被映射成为 Ethernet 协议的数据链路层和物理层,把 FC-2 层以上的内容封装到以太网报文中进行承载,并添加了 FCoE 映射层作为上层 FC 协议栈与底层 Ethernet 协议栈之间的适配层。

FCoE 协议实现了将一个完整的 FC 帧封装在以太帧中的功能,其报文封装格式如图 5.27 所示。其中,以太帧头中指定了报文的源 MAC 地址、目的 MAC 地址、以太帧类型和 FCoE 的 VLAN,FCoE 帧头指定了 FCoE 帧版本号和 控制信息。FC 帧头和传统 FC 帧头相同,指定了 FC 帧的源地址、目的地址等信息,FC 帧内容即为 SCSI 的指令、数据和状态信息单元。

以太网帧头	FCoE帧头	FC帧头	光纤通道的有效载荷	CRC校验	帧结束	总线控制系统
-------	--------	------	-----------	-------	-----	--------

图 5.27 FCoE 报文封装格式

通常情况下,一个以太网的帧最大为 1500B,而一个典型的 FC 帧最大约为 2112B,因此在以太网上打包 FC 帧时往往需要进行分段发送,然后在接收方进行重组。这种分段再重组的传输方式会产生额外的处理开销,影响 FCoE 端到端的传输效率。既然以太网和光纤通道各自所传输的帧之间存在这种差异,那自然需要一个更大的以太网帧来平衡差异,于是便出现了 FCoE 以太网巨型帧,尽管这种巨型帧不是正式的 IEEE 标准,但它允许以太网帧在长度上达到 9KB。

4. FCoE 面临的挑战

FCoE 面临的挑战如下。

(1) 扩展性。FCoE 技术将 FC 帧封装在以太网帧中,承载在以太网二层链路上,实现了两个相隔甚远的存储区域网络的互通,让业务主机可以访问到距离更远的存储设备,极大地提高了存储区域网络的扩展性能,理论上,遍布世界各地的 IP 网络可达之处,存储网络便可达。

(2) 可靠性。FC 协议不允许出现丢包,而以太网可以容忍网络丢包,那么 FCoE 借助现有普通以太网链路来传输 FC 帧,是存在网络丢包现象的,因此需要对以太网做一定的增强来避免丢包。融合增强型以太网(Converged Enhanced Ethernet,CEE)作为 FCoE 物理网络传输架构,不仅能够提供标准的光纤通道有效内容载荷,避免类似 TCP/IP 的开销和数据包损失,而且通过基于优先级的流量控制、增强传输选择和拥塞通告,达到 FCoE 对以太网提出的无丢包要求。

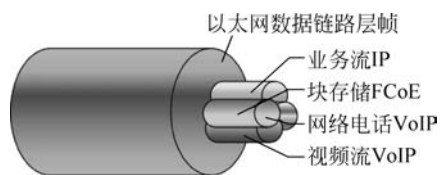


图 5.28 融合增强型以太网

基于优先级的流量控制(Priority-based Flow Control, PFC)是对以太网 Pause 机制的一种增强。以太网 Pause 机制能够实现网络不丢包的要求,但它会阻止一条链路上的所有流量。如图 5.28 所示,PFC 可以在一条以太网物理链路上创建多个独立的虚拟链路,并允许暂停和重启其中任意一条虚拟链路,通过对单个虚拟链路创建无丢包类

别的服务供 FCoE 使用,实现同一物理链路上多种类型流量的共存,如业务流 IP、块存储 FCoE、网络电话 VoIP、视频流 VoIP 等。

增强传输选择(Enhanced Transmission Selection,ETS)通过为不同的业务流量设定优先级,从而保证了高优先级业务的带宽,也允许低优先级流量使用高优先级队列的闲置带宽,提高整个网络的效率。

拥塞通告是一种在二层网络对持续拥塞流量的端到端管理方法。当网络中发生拥塞时,由拥塞点向数据源发送指示来限制引起拥塞的流量,并在拥塞消失时通知其取消限制。

5. FCoE 应用场景

存储区域网络的一个重要应用场景是大型数据中心。在传统的数据中心组网中,服务器与服务器、客户端之间的通信基于以太网 LAN,服务器与存储设备之间的通信基于存储区域网络 SAN。而 LAN 网络和 SAN 网络的部署和维护都是相互独立的。随着数据中心的迅猛发展,数据量和服务器数量日益剧增,LAN 和 SAN 独立部署方式已经无法满足企业的需求。

(1) 设备的增加使网络越来越复杂,同时,LAN 和 SAN 网络的独立部署使得业务部署的灵活性差,网络扩展困难,维护和管理成本高。

(2) 能效比低:服务器上需配置多块网卡,用于接入 LAN 的网络接口卡 NIC 和 SAN 的主机总适配器 HBA,配置多类型的网卡增加了整个数据中心的电力消耗和冷却成本。

如果大型数据中心采用以太网光线通道 FCoE 构建网络,FCoE 组网方式既支持 LAN 网络的数据传输,也支持 FC 网络的数据传输。

采用以太网光线通道 FCoE 构建网络具有如下优势。

(1) 降低总体拥有成本 TCO。

FCoE 技术共享网络资源,整合 LAN 和 SAN 网络,并有效地利用资源,减少对于 SAN 网络基础设施的投资,简化网络复杂度,降低网络的管理和维护成本;同时,服务器采用融合网络适配器

(即 CNA 卡),无须像传统网络去配置 LAN 网络接口卡 NIC 和 SAN 主机总适配器 HBA,减少数据中心的电力和冷却成本。

(2) 强大的投资保护。

FCoE 可以和数据中心现有的以太网及 FC 基础设施实现无缝互通,同时保护了客户在现有以太网和 FC 网络上的投资。

(3) 增强的业务灵活性。

FCoE 使得所有服务器共享存储资源。特别是在虚拟机迁移的应用场景下,可为虚拟机提供一致的存储连接,提高了系统的灵活性和可用性,增强了业务灵活性。

5.3 项目实施

5.3.1 iSCSI 服务器配置与管理

存储服务器的在线扩容、数据快照、NAS 等技术均可应用服务器提供磁盘空间服务,但是一些关键应用需要服务器提供本地磁盘服务,这就需要使用存储提供的服务器本地磁盘空间服务存储区域网络 SAN 服务,SAN 服务可以为应用服务器提供本地的磁盘服务,并支持在线扩容、容灾备份等功能。SAN 是一种在服务器和存储服务器之间实现高速可靠访问的存储网络。存储服务器基于 SCSI 协议将卷上的 1 个存储区块租赁给服务器,服务器通过 SCSI 客户端将这个区块识别为 1 个本地硬盘,然后初始化该硬盘后即可用于存取数据。SCSI 的主要功能是在主机和存储设备之间传送命令、状态和块数据。

当多数企业由于 FC SAN 的高成本而对 SAN 敬而远之时,iSCSI(Internet SCSI)技术的出现,推动了 IP SAN 在企业中的应用。大多数中小企业都以 TCP/IP 为基础建立了网络环境,iSCSI 可以在 IP 网络上实现 SCSI 的功能,允许用户通过 TCP/IP 网络构建存储区域网,为众多要求经济合理和便于管理的中小企业的存储设备提供了直接访问的能力。

由此可见,IP SAN 实际上就是使用 IP 将服务器与存储设备连接起来的技术,基于 IP 网络实现数据块级别的存储。在 IP SAN 的标准中,除了已获通过的 iSCSI,还有 FCIP、iFCP 等协议标准。其中,iSCSI 发展是最快的,它已经成为 IP 存储技术的一个典型代表。基于 iSCSI 的 SAN 的目的就是要使用本地 iSCSI 导向器(Initiator)和 iSCSI 目标(Target)来建立 SAN。iSCSI 的两个组件如下。

目标器(Target,服务端):存储设备上的 iSCSI 服务,用于转换 TCP/IP 包中的 SCSI 命令和数据,服务端的端口号默认为 3260。

导向器(Initiator,客户端):iSCSI 客户端软件,一般安装在应用服务器上,它接收应用层的 SCSI 请求,并将 SCSI 命令和数据封装到 TCP/IP 包中发送到 IP 网络中。

1. iSCSI 目标服务器安装

打开“服务器管理器”窗口,选择“管理”→“添加角色和功能”选项,弹出“添加角色和功能向导”窗口,持续单击“下一步”按钮,直至出现“选择服务器角色”窗口,勾选“iSCSI 目标服务器”复选框,如图 5.29 所示,持续单击“下一步”按钮,直至出现“确认”窗口,单击“安装”按钮,完成 iSCSI 目标服务器的安装,如图 5.30 所示。



视频讲解

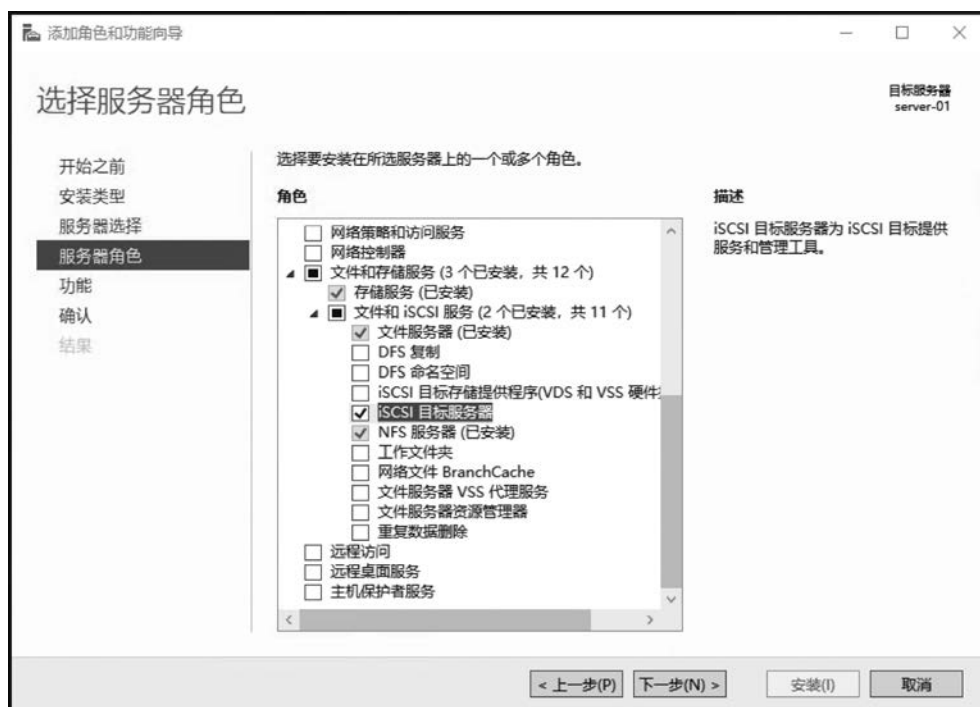


图 5.29 添加“iSCSI 目标服务器”

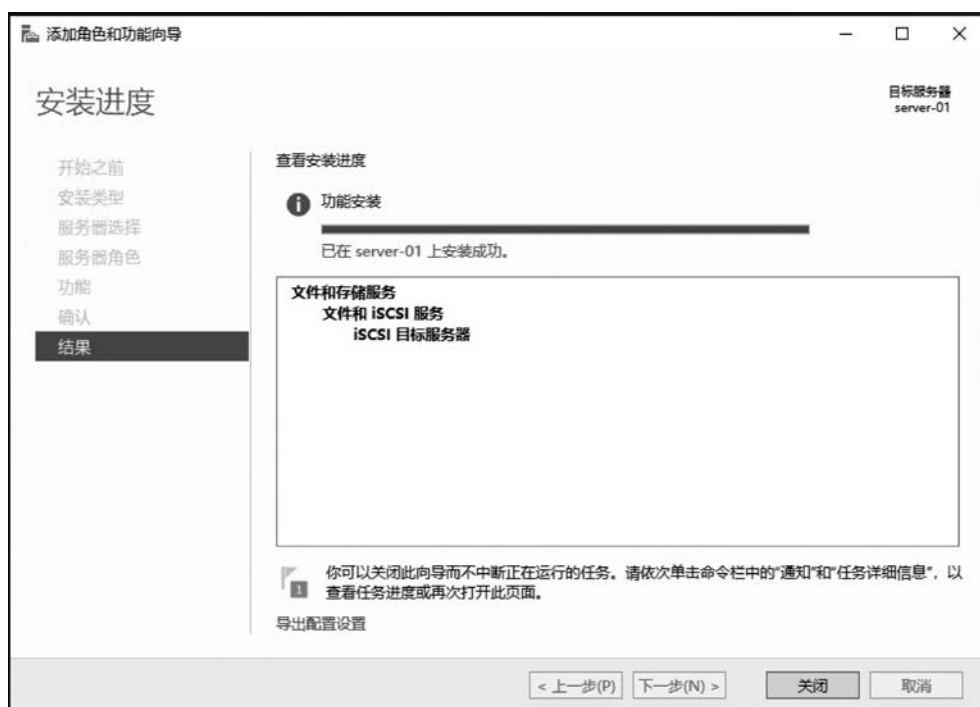


图 5.30 完成“iSCSI 目标服务器”安装



视频讲解

2. iSCSI 服务端虚拟磁盘配置与管理

创建一块 iSCSI 虚拟磁盘,操作过程如下。

(1) 打开“服务器管理器”窗口,在“仪表板”主窗口中,选择“文件和存储服务”→iSCSI 选项,单击 iSCSI 管理界面中的“任务”下拉列表,选择“新建 iSCSI 虚拟磁盘”选项,如图 5.31 所示,弹出“选择 iSCSI 虚拟磁盘位置”窗口,如图 5.32 所示。



图 5.31 “新建 iSCSI 虚拟磁盘”选项



图 5.32 “选择 iSCSI 虚拟磁盘位置”窗口

(2) 在“选择 iSCSI 虚拟磁盘位置”窗口中,选择相应的服务器和存储位置,单击“下一步”按钮,弹出“指定 iSCSI 虚拟磁盘名称”窗口,输入虚拟磁盘名称“iSCSI-01”,如图 5.33 所示,单击“下一步”按钮,弹出“指定 iSCSI 虚拟磁盘大小”窗口,输入虚拟磁盘大小为 30GB,选择“固定大小”单选按钮,如图 5.34 所示。

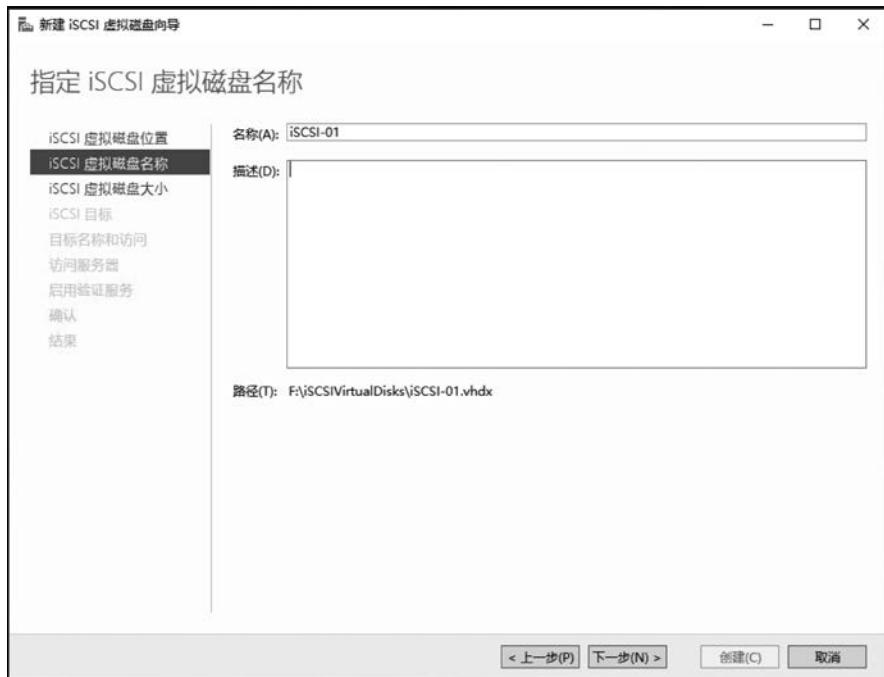


图 5.33 “指定 iSCSI 虚拟磁盘名称”窗口

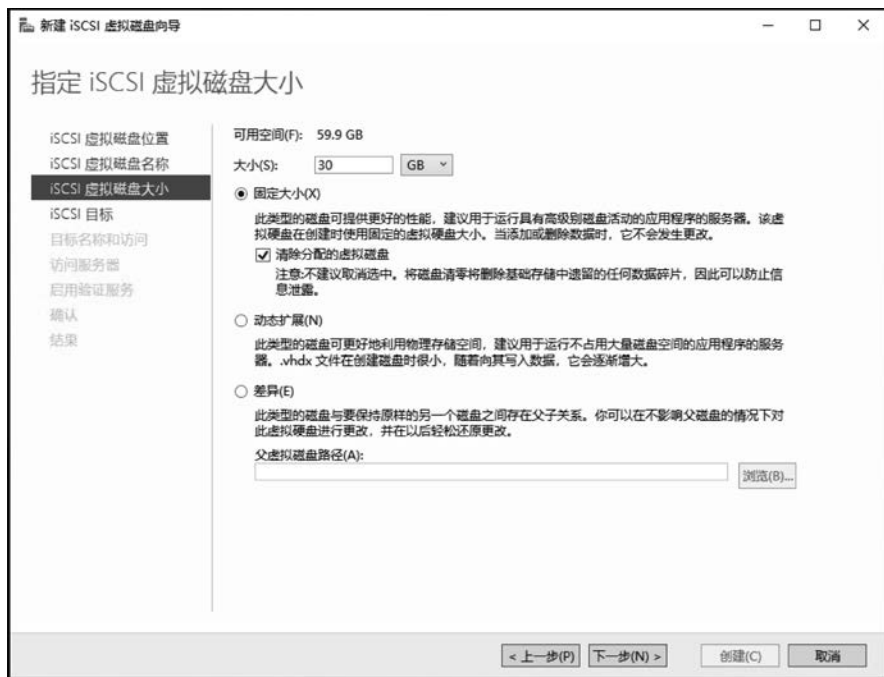


图 5.34 “指定 iSCSI 虚拟磁盘大小”窗口

(3) 在“指定 iSCSI 虚拟磁盘大小”窗口中,单击“下一步”按钮,弹出“分配 iSCSI 目标”窗口,如图 5.35 所示,单击“下一步”按钮,弹出“指定目标名称”窗口,输入名称“target-01”,如图 5.36 所示。

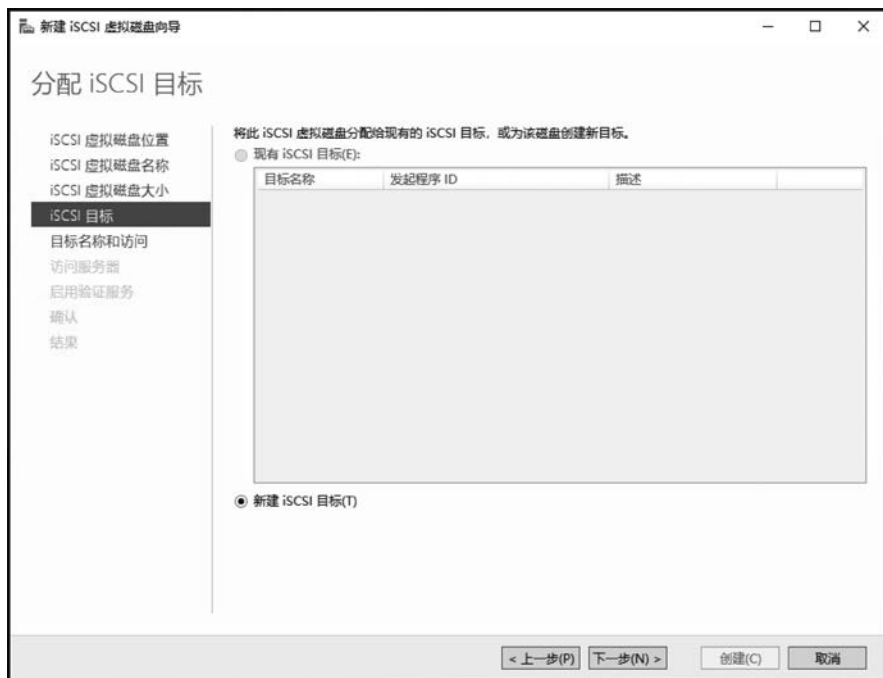


图 5.35 “分配 iSCSI 目标”窗口

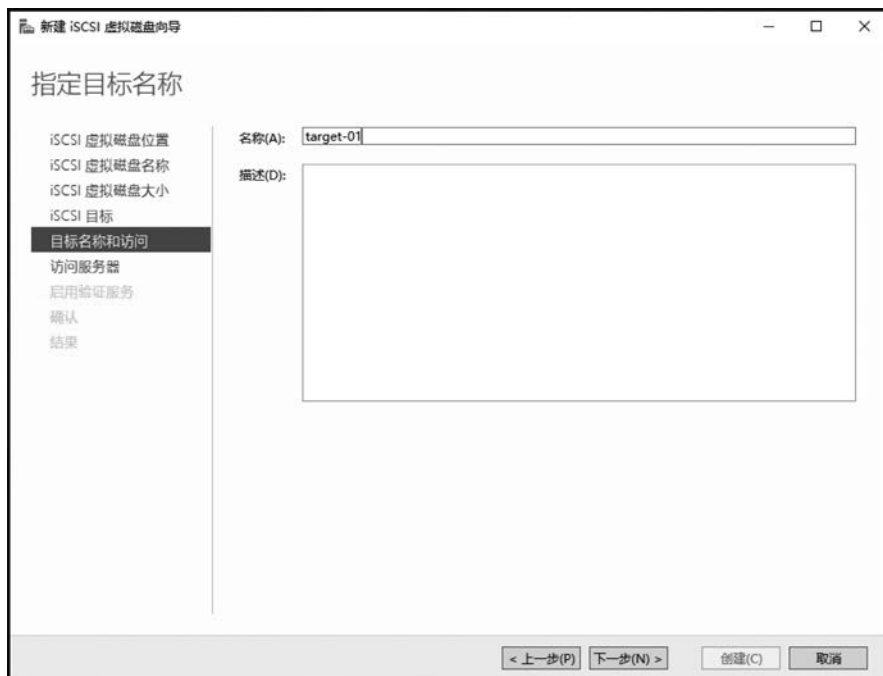


图 5.36 “指定目标名称”窗口

(4) 在“指定目标名称”窗口中,单击“下一步”按钮,弹出“指定访问服务器”窗口,单击“添加”按钮,弹出“添加发起程序 ID”窗口,选择“输入选定类型的值”单选按钮,在“类型”下拉列表中选择“IP 地址”选项,“值”输入“192.168.100.100”,如图 5.37 所示,单击“确定”按钮,返回“指定访问服务器”窗口,如图 5.38 所示。

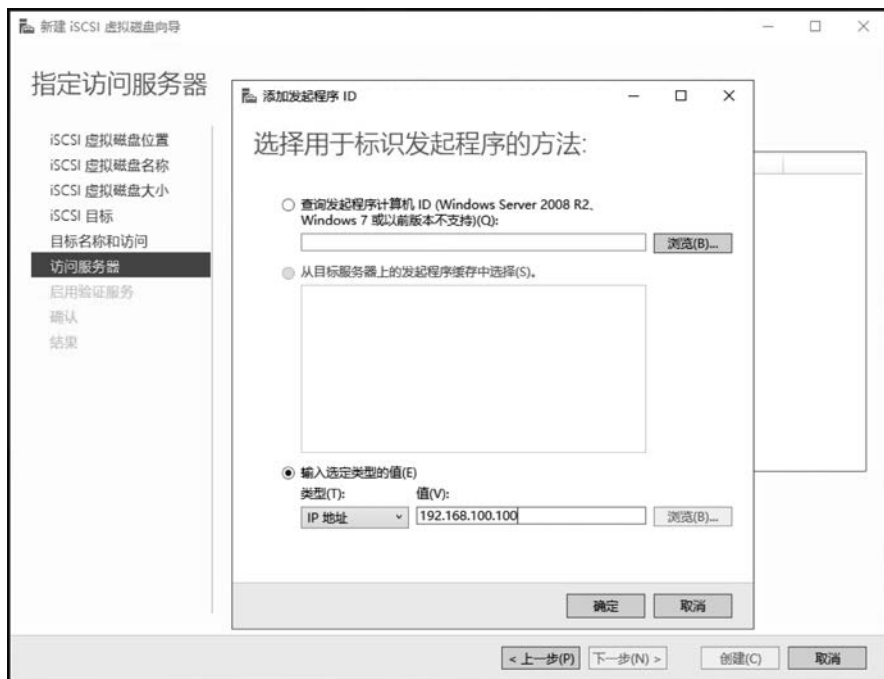


图 5.37 “添加发起程序 ID”窗口



图 5.38 “指定访问服务器”窗口

(5) 在“指定访问服务器”窗口中,单击“下一步”按钮,弹出“启用身份验证”窗口,如图 5.39 所示,单击“下一步”按钮,弹出“确认选择”窗口,如图 5.40 所示。



图 5.39 “启用身份验证”窗口



图 5.40 “确认选择”窗口

(6) 在“确认选择”窗口中,单击“创建”按钮,弹出“查看结果”窗口,如图 5.41 所示,单击“关闭”按钮,完成 iSCSI 虚拟磁盘的创建,返回“iSCSI 虚拟磁盘”窗口,如图 5.42 所示,此时查看文件资源管理器,查看 F 盘文件,如图 5.43 所示。



图 5.41 “查看结果”窗口



图 5.42 “iSCSI 虚拟磁盘”窗口

(7) 扩展 iSCSI 虚拟磁盘。在“iSCSI 虚拟磁盘”窗口中,选择相应的虚拟磁盘,单击鼠标右键,在弹出的快捷菜单中,选择“扩展 iSCSI 虚拟磁盘”选项,如图 5.44 所示,弹出“扩展 iSCSI 虚拟磁盘”对话框,如图 5.45 所示,输入新虚拟磁盘大小为 40GB,单击“确定”按钮,返回“iSCSI 虚拟磁



图 5.43 查看 F 盘文件



图 5.44 “扩展 iSCSI 虚拟磁盘”选项



图 5.45 “扩展 iSCSI 虚拟磁盘”对话框

盘”窗口,可以看到新虚拟磁盘大小已经变为 40GB,如图 5.46 所示。

3. iSCSI 客户端虚拟磁盘配置与管理

通过 iSCSI 发起程序连接 iSCSI 虚拟磁盘,并对连接的虚拟磁盘进行分区及格式化操作。

(1) 打开“服务器管理器”窗口,选择“工具”→“iSCSI 发起程序”选项,如图 5.47 所示,弹出 Microsoft iSCSI 对话框,如图 5.48 所示。

(2) 在 Microsoft iSCSI 对话框中,单击“是”按钮,弹出“iSCSI 发起程序属性”对话框,在“目标”区域中输入“192.168.100.100”,单击“快速连接”按钮,弹出“快速连接”对话框,如图 5.49 所示。单击“完成”按钮,返回“iSCSI 发起程序 属性”对话框,如图 5.50 所示。



视频讲解



图 5.46 返回“iSCSI 虚拟磁盘”窗口



图 5.47 “iSCSI 发起程序”选项

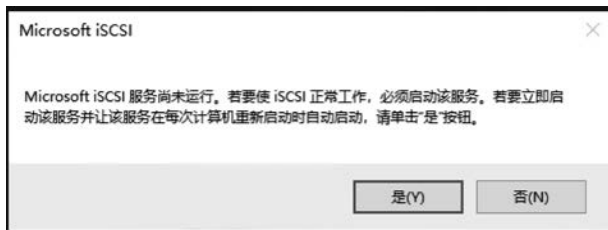


图 5.48 Microsoft iSCSI 对话框

(3) 打开本地磁盘管理工具,查看本地磁盘,可以看到系统已经识别到该 iSCSI 磁盘,大小为 30GB,如图 5.51 所示。

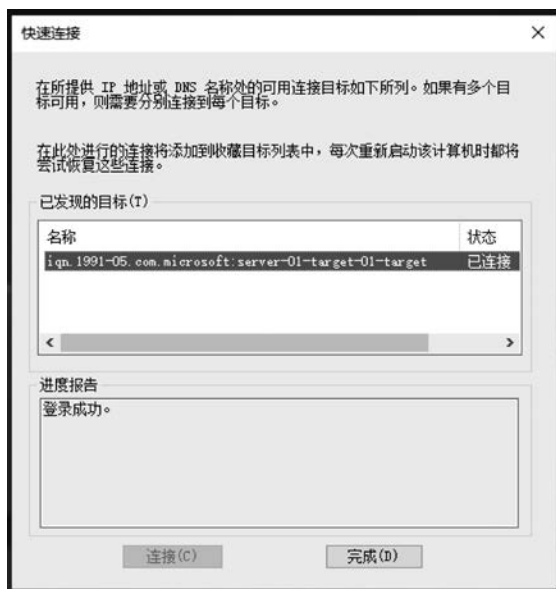


图 5.49 “快速连接”对话框

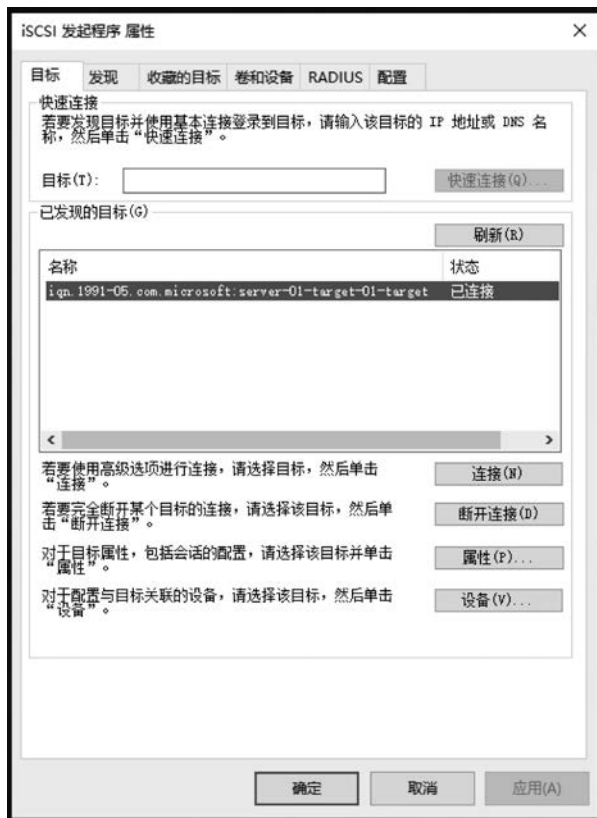


图 5.50 “iSCSI 发起程序 属性”对话框



图 5.51 查看本地磁盘

(4) 对磁盘进行新建简单卷操作，新建一个简单卷 G 盘，如图 5.52 所示。



图 5.52 对磁盘进行新建简单卷操作

5.3.2 多路径链路 iSCSI 虚拟磁盘应用部署

多路径解决方案使用冗余的物理路径组件(适配器、电缆和交换机)在服务器与存储设备之间创建逻辑路径。如果这些组件中的 1 个或多个发生故障,导致路径无法使用,多路径逻辑就使用 I/O 的另一条路径以使应用程序仍然能够访问其数据。

Windows Server 2019 中的多路输入/输出(Multi Path Input Output, MPIO)功能包含 1 个设备特定模块分布式存储管理器(Distributed Storage Manager, DSM),该模块提供以下负载平衡策略。

(1) 故障转移。不执行负载平衡。应用程序需要指定 1 个主路径和 1 组备用路径,主路径用于处理设备请求,备用路径阻塞状态。如果主路径发生故障,服务器会自动启用其中的备用路径。

(2) 故障恢复。故障恢复是指只要首选路径有效,所有 I/O 都指向首选路径。如果首选路径发生故障,I/O 将被定向到备用路径,直到首选路径功能恢复为止。

(3) 循环。DSM 以轮询方式使用 I/O 的所有可用路径。

1. 多路径 I/O 服务安装

打开“服务器管理器”窗口,选择“管理”→“添加角色和功能”选项,弹出“添加角色和功能向导”窗口,持续单击“下一步”按钮,直至出现“选择功能”窗口,勾选“多路径 I/O”复选框,如图 5.53 所示,持续单击“下一步”按钮,直至出现“确认”窗口,单击“安装”按钮,完成“多路径 I/O 服务”安装,如图 5.54 所示。

2. 多路径 iSCSI 虚拟磁盘部署

在存储服务器 Server-01 上安装 iSCSI 目标服务器,设置服务器的两个网络接口的 IP 地址的相关信息,如图 5.55 所示。

(1) 打开“服务器管理器”窗口,选择“工具”→“iSCSI 发起程序”选项,弹出“iSCSI 发起程序属



视频讲解

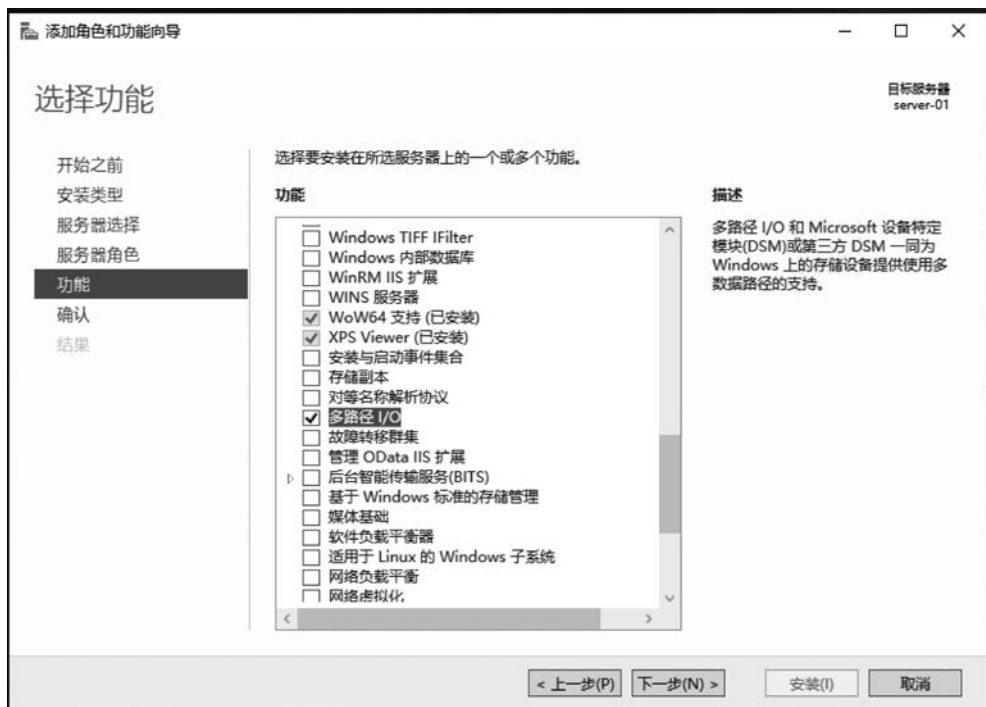


图 5.53 “多路径 I/O”选项

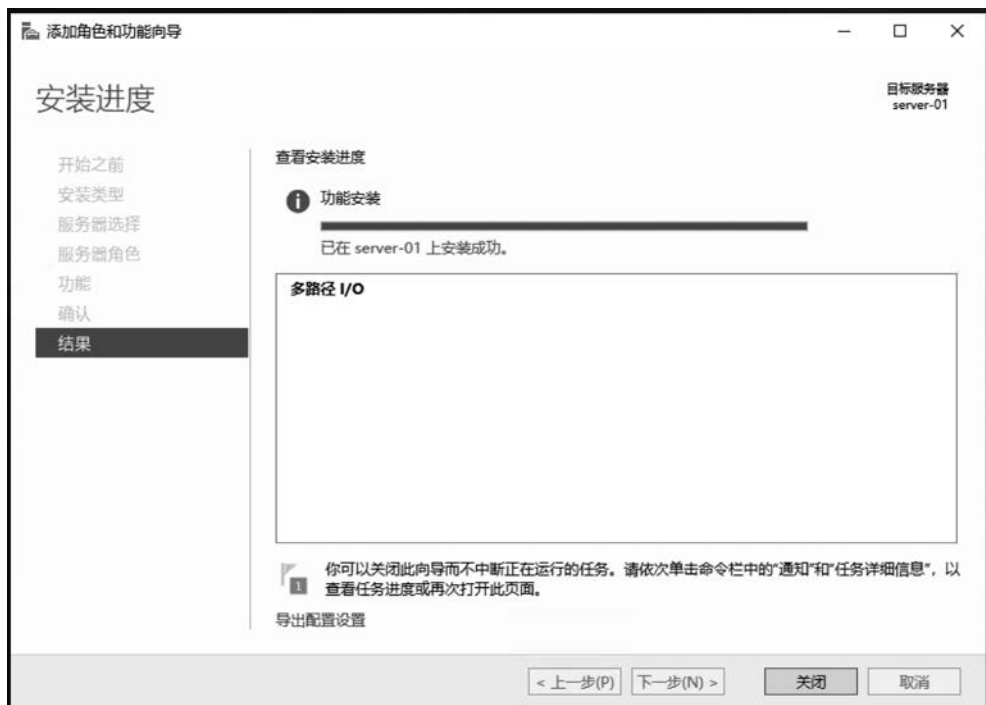


图 5.54 完成“多路径 I/O”服务安装

性”对话框,选择“发现”选项卡,如图 5.56 所示,单击“发现门户”按钮,弹出“发现目标门户”对话框,输入 IP 地址“192.168.200.100”,端口“3260”,如图 5.57 所示。

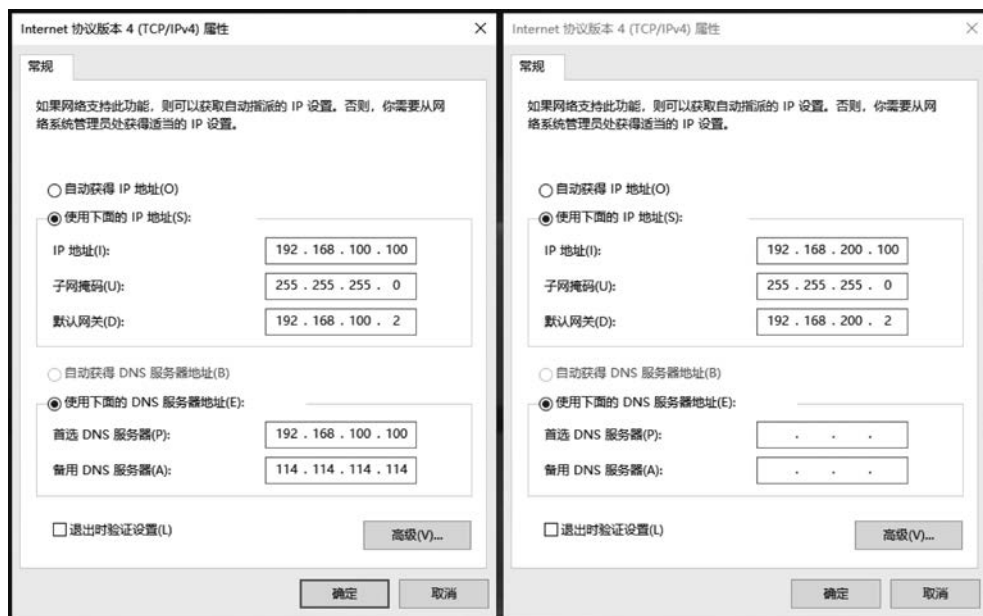


图 5.55 网络接口 IP 地址信息

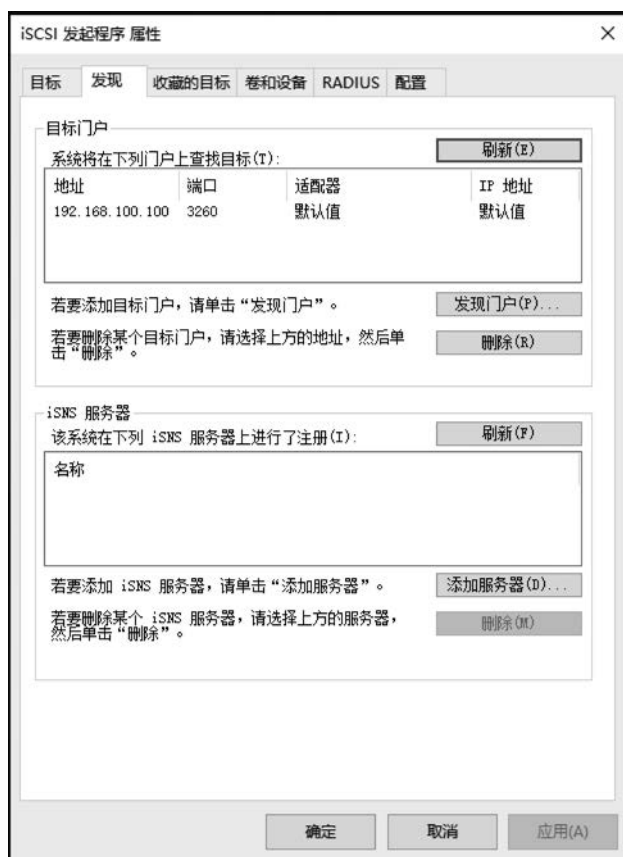


图 5.56 “iSCSI 发起程序 属性”对话框

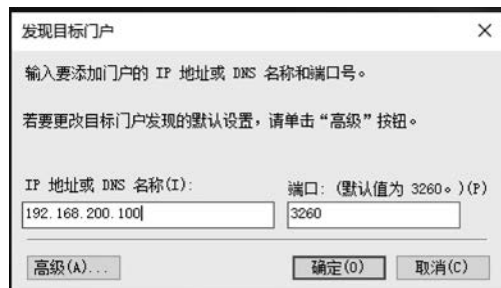


图 5.57 “发现目标门户”对话框

(2) 在“发现目标门户”对话框中,单击“确定”按钮,返回“iSCSI 发起程序 属性”对话框,如图 5.58 所示,打开“卷和设备”选项卡,单击“自动配置”按钮,在“卷列表”中,自动挂载磁盘,如图 5.59 所示。

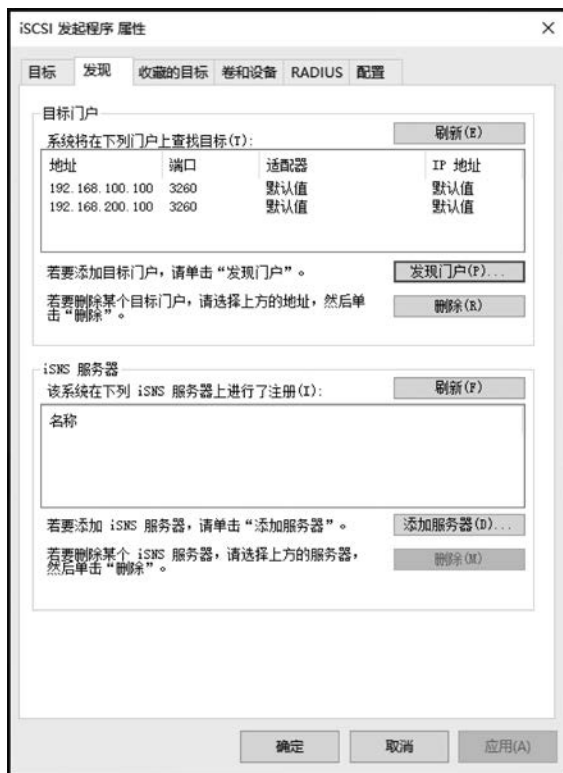


图 5.58 “发现”选项卡

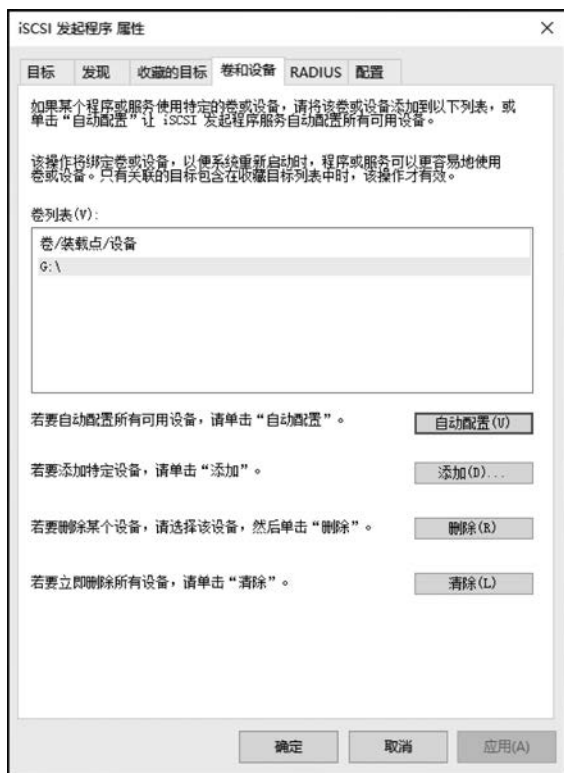


图 5.59 “卷和设备”选项卡

(3) 打开“服务器管理器”窗口,选择“工具”→MPIO 选项,如图 5.60 所示,弹出“MPIO 属性”对话框,如图 5.61 所示,选择“发现多路径”选项卡,勾选“添加对 iSCSI 设备的支持”复选框,如图 5.62 所示。

(4) 在“发现多路径”选项卡中,单击“添加”按钮,弹出“MPIO 操作: 成功”对话框,如图 5.63 所示,重启完成后,在“MPIO 属性”对话框中可以看到新增加的 iSCSI 设备信息,如图 5.64 所示。

(5) 打开“服务器管理器”窗口,选择“工具”→“iSCSI 发起程序”选项,弹出“iSCSI 发起程序属性”对话框,选择“目标”选项卡,如图 5.65 所示,单击“连接”按钮,弹出“连接到目标”对话框,勾选

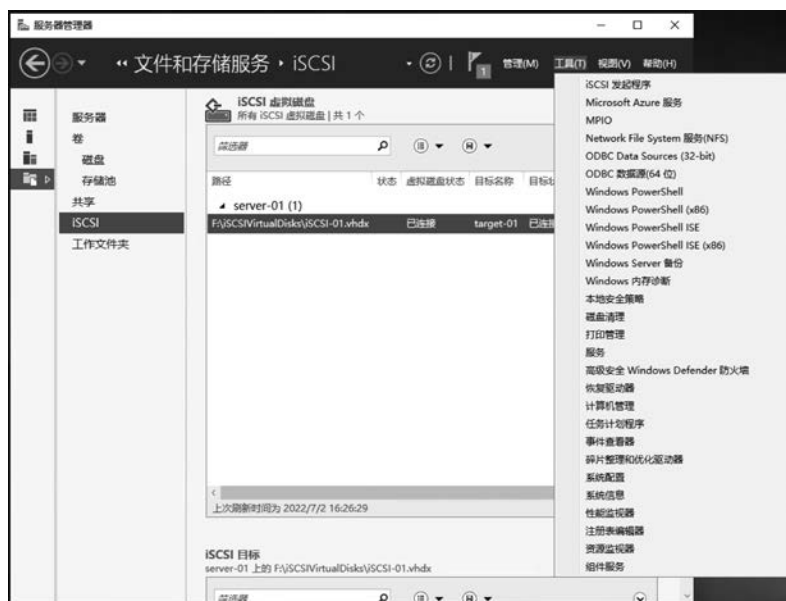


图 5.60 MPIO 选项

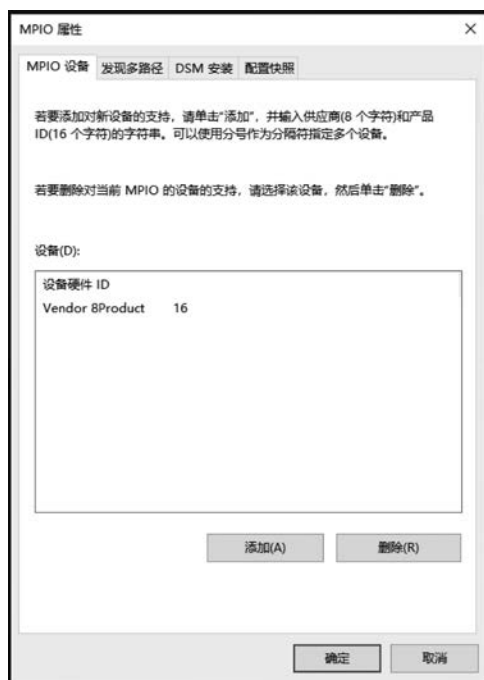


图 5.61 “MPIO 属性”对话框



图 5.62 “发现多路径”选项卡



图 5.63 “MPIO 操作: 成功”对话框

“启用多路径”复选框,如图 5.66 所示。



图 5.64 “MPIO 设备”选项卡

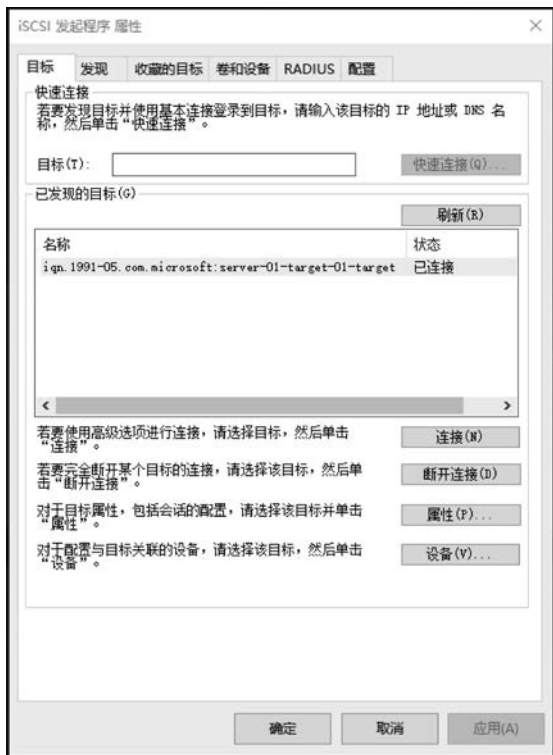


图 5.65 “目标”选项卡



图 5.66 “连接到目标”对话框

(6) 在“连接到目标”对话框中,单击“高级”按钮,弹出“高级设置”对话框,选择“目标门户 IP”为 192.168.100.100,单击“确定”按钮,如图 5.67 所示。以同样的方法,选择“目标门户 IP”为 192.168.200.100,单击“确定”按钮,如图 5.68 所示。

(7) 在“iSCSI 发起程序属性”对话框中,选择“目标”选项卡,单击“属性”按钮,弹出“属性”对话框,如图 5.69 所示,勾选“标识符”复选框,单击“设备”按钮,弹出“设备”对话框,如图 5.70 所示。

(8) 在“设备”对话框中,单击 MPIO 按钮,弹出“设备详细信息”对话框,如图 5.71 所示,单击“确定”按钮,返回“iSCSI 发起程序属性”对话框,选择“卷和设备”选项卡,可以看到卷列表情况,如图 5.72 所示。

(9) 打开磁盘管理器和文件资源管理,对连接的磁盘进行联机/初始化/新建简单卷操作,新建磁盘 G 盘,如图 5.73 所示。



图 5.67 “目标门户 IP”对话框 1



图 5.68 “目标门户 IP”对话框 2

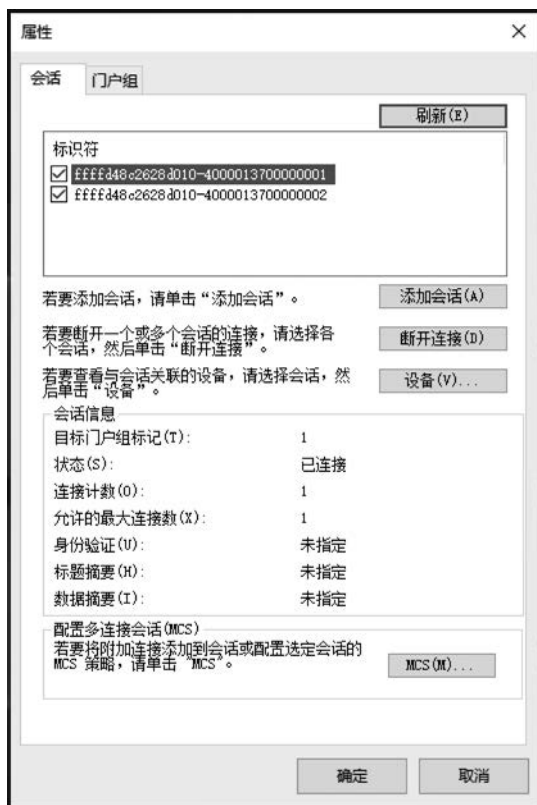


图 5.69 “属性”对话框



图 5.70 “设备”对话框



图 5.71 “设备详细信息”对话框

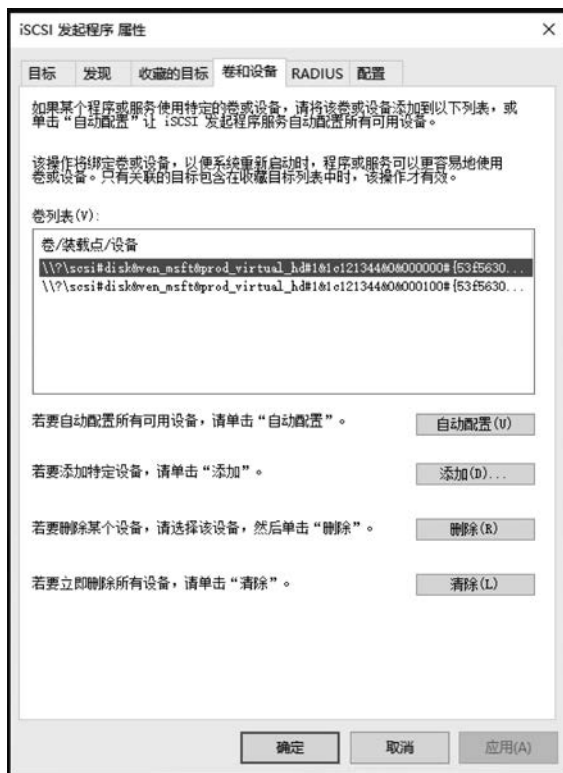


图 5.72 “卷和设备”选项卡

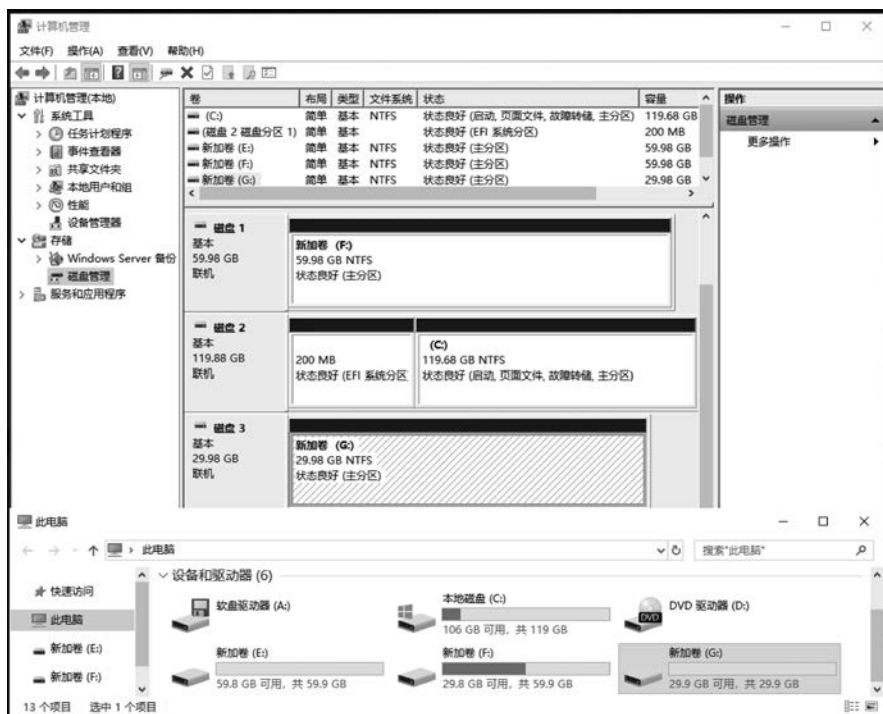


图 5.73 新建磁盘 G 盘

课后习题

1. 选择题

- (1) FC-1 协议的作用是()。
- A. 物理变换 B. 编码/解码 C. 结构协议 D. 通用服务
- (2) FC-2 协议的作用是()。
- A. 物理变换 B. 编码/解码 C. 结构协议 D. 通用服务
- (3) FC-3 协议的作用是()。
- A. 物理变换 B. 编码/解码 C. 结构协议 D. 通用服务
- (4) SCSI 协议传输过程：()。
- A. 总线忙—协商—寻址—连接—断开连接
B. 总线忙—寻址—协商—连接—断开连接
C. 协商—总线忙—寻址—连接—断开连接
D. 寻址—总线忙—协商—连接—断开连接
- (5) 【多选】SAN 的基本组件组成：()。
- A. 服务器 B. 网络基础设施 C. 打印机 D. 存储
- (6) 【多选】SAN 的存储结构特征：()。
- A. 具有可伸缩能力 B. 非常可靠
C. 提供尽可能高的传输速度 D. 易于管理

2. 简答题

- (1) 简述传统数据存储 DAS 存在的不足。
- (2) 简述 DAS 和 SAN 的区别。
- (3) 简述 FC 与 SCSI 协议的关系。
- (4) 简述 IP SAN 网络架构的优缺点。
- (5) 简述 iSCSI 协议栈工作原理。