

阵列运行自定义操作系统,通常无法安装备份代理软件,因此需要 NDMP 之类的协议。

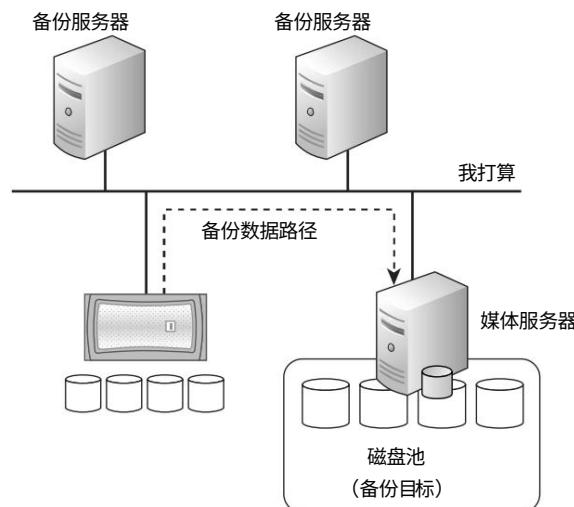
NDMP 备份配置的一个重要特性是直接访问恢复 (DAR)。
直接访问还原可以实现以下功能:

- 选择性文件恢复 (包括单个文件恢复)
- 恢复更快

然而,天下没有免费的午餐!DAR 提供的更快、更精细的恢复功能的副作用是备份时间更长。这是因为备份每个文件时,会记录其在备份中的位置。这样您就不必读取或恢复整个备份映像来查找要恢复的单个文件或目录。

图 11.5 显示了一个简单的 NDMP 配置。

图 11.5 NDMP 备份配置



在图 11.5 中,备份数据通过 LAN 直接从 NAS 设备发送到备份目标磁盘池,如虚线所示。



备份大量小文件可能会非常麻烦 can be而且不仅仅是对于 NDMP 而言。for NDMP.
例如备份 1,000,000 个 1KB 文件所花的时间要比备份 1,000,000 1MB 文件所花的时间长得多。长得多!原因之一是必须对每个单独的文件进行分类,并且在某些备份解决方案中还需要进行验证。这类似于粉刷一面大而平坦的墙壁与粉刷许多面相同面积的小墙壁之间的区别。对于大墙,您只需要粉刷一组角和边,而粉刷许多小墙会大大增加您需要粉刷的角和边的数量。如果您的大型文件系统包含数百万或数十亿个文件,并且备份窗口很小,您很难将备份放入其中,那么基于块的备份解决方案 (如无服务器备份)可能是更好的选择。但是,基于块的备份往往不允许您轻松恢复单个文件。

备份类型

在现代世界中,备份有很多种类型 完整备份、增量备份、差异备份、合成备份 所有这些备份都可以协同工作以满足大多数业务备份需求。应用程序意识是选择备份方法时需要考虑的另一个因素。让我们更详细地了解它们。

完整备份

顾名思义,完整备份会备份所有内容。它们从顶部开始,到底部才结束。存储网络行业协会 (SNIA) 将完整备份定义为“复制一组定义的数据对象的所有数据,无论它们自上次备份以来是否被修改过。”

与所有事物一样,完整备份也有其优点和缺点。

从积极的一面来看,它们在恢复速度和恢复简易性方面提供了最好的服务。例如,如果我们每晚备份包含 250 GB 数据的整个家庭硬盘,并且想要恢复六天前的家庭硬盘内容,我们可能只需要一盘磁带即可执行恢复。该过程将需要识别磁带、将其加载到磁带驱动器中,然后从磁带中恢复数据。

缺点是,完整备份会占用备份目标设备上的大多数备份空间,以及最多的网络带宽和服务器资源。而且它们会耗费最多的时间。使用之前每晚备份家庭硬盘的示例,这意味着我们每晚都会通过网络提取 250 GB 的备份数据并将其存储到我们的备份设备中。虽然这听起来可能不是很多数据,但当您意识到您的组织每晚有 100、500 或超过 1,000 个家庭硬盘需要备份时,它很快就会变成大量数据。它还要求在备份过程中读取我们的整个主目录,从而给托管家庭硬盘的主机和存储阵列带来额外的读取工作负载。因此,完整备份并非全是好事。

注意“增量”的含义!

在一般用法中,并且根据几乎所有商业备份工具,增量备份被定义为备份自上次备份以来挂起的文件的作业……并且“上次备份”可以是增量备份或完整备份。这意味着完整恢复需要最近的完整备份以及自完整备份以来的所有增量备份。

遗憾的是,在撰写本文时,SNIA 对增量备份的定义不同。根据 SNIA 存储词典,累积增量备份 (与差异增量备份相反) 是“复制自上次完整备份以来修改的所有数据对象的备份”(请注意使用完整一词)。由于本书与 CompTIA Storage+ Powered by SNIA 考试相对应,因此本书使用了 SNIA 定义,尤其是在接下来几个部分中。但是,当您在非 SNIA 的现实环境中遇到“增量备份”时,请不要被该定义误导!

累积增量备份

累积增量备份（SNIA 术语，通常称为差异备份）与完整备份不同，但可与完整备份配合使用。完整备份和累积增量备份之间的主要区别在于，累积增量备份仅备份自上次完整备份以来发生变化的数据。因此，首先，您需要进行完整备份，然后才能进行累积增量备份。

大多数累积增量备份解决方案都会备份整个文件，即使自上次备份以来该文件只有一个字节发生了变化。让我们看一个基于 250 GB 家用硬盘的示例。假设我们有一个备份计划，每周五晚上对我们的家用硬盘进行一次完整备份，然后每周每隔一天进行累积增量备份。周五，我们将通过网络拖拽 250 GB 并将其存储在我们的备份设备上。然后在周六，我们将通过网络拖拽并存储自周五晚上完整备份以来更新过的文件到我们的备份设备 - 可能根本没有多少数据。在下一个周五晚上完整备份之前的每个其他备份也是如此；每个累积增量备份仅备份自上一个周五完整备份以来发生变化的数据。除非我们每天更改家用硬盘中的大量数据，否则这种完整备份和累积增量备份的组合将为我们节省大量的备份存储空间和网络带宽。

这很好，但也有缺点。假设今天是星期四，我们删除了主驱动器的所有内容，需要从最新备份中恢复。最新备份是星期三晚上，但星期三晚上的备份只是累积增量备份，因此它只包含自上周五完整备份以来发生变化的数据。

因此，为了恢复主目录的全部内容，我们需要星期五晚上完整备份的磁带以及星期三晚上累积增量备份的磁带。然后，我们的备份软件需要协调该过程，以便恢复的是我们主驱动器在星期三晚上的映像。

它比每晚进行完整备份稍微复杂一些，但当涉及到文件系统备份时，大多数组织认为，在一周内使用累积增量备份的好处远远超过其缺点。

差异增量备份

差异增量备份（再次强调，这是 SNIA 术语）是一种增量备份。与累积增量备份一样，差异增量备份与完整备份配合使用。但差异增量备份不会备份自上次完整备份以来发生更改的所有数据，而是仅备份自上次差异增量备份以来发生更改的数据。这些备份在空间效率方面非常出色，因为与累积增量备份相比，每个差异增量备份需要通过网络向我们的备份设备发送的数据更少。但是，它们也有缺点。在恢复操作方面，差异增量备份是最慢且最复杂的。

让我们通过上一节中的家庭驱动器示例来看一下它是如何工作的。假设今天又是星期四，我们再次删除了家庭驱动器的所有内容，需要从最新备份中恢复它。最新备份是星期三晚上。但这次星期三晚上的备份是差异增量备份，这意味着它只包含自差异增量备份以来发生变化的数据。

前一天晚上（星期二晚上）的备份。而星期二晚上的备份也是差异增量备份，这意味着它只包含自星期一晚上备份以来发生变化的数据。而星期一晚上的备份也是差异增量备份。希望您明白了。最终结果是，为了恢复我们主目录的全部内容，我们需要星期五晚上完整备份的磁带，以及随后每个差异增量备份的所有磁带 - 星期六、星期日、星期一、星期二和星期三！同样，备份软件需要协调该过程，以便恢复的是我们主驱动器在星期三晚上的图像。

这比完整和累积增量备份还要复杂得多。
但是，它更节省空间，并且对网络和主机的负担较小。

合成完整备份

与标准完整备份、增量备份和差异备份相比，合成完整备份是一种较新的备份形式。从高层次上讲，合成完整备份的工作原理是进行一次完整备份，然后从那时起只进行增量备份或差异备份。因此，您可能会听到合成完整备份被称为永久增量备份系统的一种形式。但它比这种描述所暗示的要聪明一些。合成完整备份不是永久进行增量备份，也不是总是必须返回到原始完整备份和最新的差异增量备份，而是定期创建合成完整备份。合成完整备份的外观、感觉和工作方式与普通完整备份完全相同，只是它不是通过搜索我们家庭驱动器中的每个文件、再次通过网络传输每个文件以及再次将每个文件存储到我们的备份设备来创建的。

让我们再看看主驱动器示例。假设我们决定部署合成完整备份方法来备份我们的主目录。在第一个星期五，我们执行传统的完整备份。我们读取所有 250 GB 的数据，通过网络将其提取，并将其作为单个完整备份映像存储在我们的备份设备上。太好了。接下来的每个晚上，我们都会执行增量备份，以便我们仅通过网络提取更改的数据并将其存储到我们的备份设备中。然后，当下一个星期五到来时，我们不会再执行另一个传统的完整备份及其所有相关开销，而是获取上周五的完整备份映像，获取最新的增量备份，然后将上一个完整备份与我们刚刚进行的增量备份相结合。我们最终得到的备份映像与传统的完整备份映像完全一样，只是我们不必从主驱动器读取每个文件并将其通过网络提取。值得注意的是，这个合成完整备份可以作为另一周增量备份的基础。

但是，创建合成完整备份的方式可能会给备份设备带来额外的压力，因为备份软件必须从备份设备读取上次完整备份的全部内容，同时备份设备还要写出新的合成完整备份映像。对于备份设备来说，这种读写工作可能非常辛苦。

应用程序感知备份

应用程序感知备份（有时称为应用程序一致性备份）是一种热备份，如果您希望能够从备份中恢复应用程序，则至关重要。

备份标准文件系统数据（如主驱动器和文件共享）是一回事。但是，如果您尝试备份数据库或应用程序（如 Microsoft Exchange）而不将备份与应用程序协调，那么您几乎肯定无法从备份中恢复应用程序！

应用程序感知备份通常需要在应用程序服务器上安装特殊的代理软件，例如在 Exchange 服务器上安装 Exchange Server 代理。

该代理通常负责与备份服务器进行通信以及向备份设备发送数据，但它还提供在一致状态下备份应用程序所需的智能。



Real World Scenario

确保在进行更改之前进行测试

一家公司经历了快速增长，正在寻找减少备份流量所消耗的网络带宽的方法。由于该公司最初是一家小公司，因此其基于 LAN 的备份设计没有专用的备份 LAN。

他们希望实施合成完整备份。

然而，他们的备份目标设备是基于磁盘的重复数据删除设备。在测试阶段，在推出合成完整备份之前，他们注意到在创建合成完整映像时备份性能受到了显著的负面影响。

经过仔细检查，我们发现，创建合成完整备份映像的方式产生了重复数据删除设备无法处理的工作负载模式；对于基于磁盘的存储系统来说，同时进行大量读取和写入已经够糟糕了，而当重复数据删除加入其中时，情况就更加糟糕了。这是因为当从重复数据删除设备读回数据时，必须对其进行重新水化（取消重复数据删除），这会大大加剧已经发生的磁盘抖动。

在这种情况下，最终结果是恢复到传统的完整备份并考虑实施备份 LAN。

许多应用程序都提供了一种机制，可以对自身进行热备份和一致备份。在 Microsoft 世界中，卷影复制服务 (VSS) 是一种允许 Microsoft 应用程序在热和一致状态下进行备份的技术。一些非 Microsoft 选项也存在类似的选项，例如 Oracle 的 RMAN。但是，由于 VSS 如此流行和知名，让我们快速了解一下它的工作原理。



从技术上讲，任何 Windows 应用程序都可以插入并利用 VSS 框架进行应用程序感知备份。
例如，在 Windows 上运行的 Oracle 可以使用 VSS 进行应用程序感知备份。For example,

Microsoft 卷影复制服务

在 Microsoft VSS 领域,备份软件 (在 VSS 术语中称为请求程序) 指示服务器上运行的 VSS 服务,它想要备份该服务器上的应用程序。例如,如果意图是备份 Microsoft SQL Server,则备份软件 (VSS 请求程序) 将告诉运行 SQL Server 的计算机上的 VSS,它想要备份 SQL Server。然后,VSS 会告诉应用程序 (SQL Server) 为热备份做好准备。

应用程序通过完成实时 I/O、刷新本地缓冲区、应用日志以及设置顺利恢复所需的任何应用程序标志来实现此目的。然后,应用程序将其状态报告给 VSS。然后,VSS 指示应用程序在创建快照时处于静止状态 (冻结应用程序 I/O)。

拍摄快照后,VSS 指示应用程序 (SQL Server) 恢复操作。静止期间冻结的所有应用程序 I/O 均已完成,并且恢复正常操作。然后根据 VSS 创建的快照进行备份。

但重要的是要明白,即使在静止期间,当应用程序保持运行并处理用户请求时,冻结的应用程序 I/O 仍在发生,但直到静止期结束才会提交到磁盘。静止期通常只持续几秒钟,在此期间拍摄快照。

最终结果是应用程序一致的备份 (保证可作为恢复选项),该备份是在应用程序在线且工作时进行的。应用程序的任何用户或客户端都不会知道应用程序刚刚处于热备份模式并进行了备份。

备份目标和设备

所有备份都需要存储在某个地方。从技术上讲,我们将备份的目的地称为备份目标,但有时我们使用备份设备这个术语。

过去,备份目标几乎总是基于磁带,使用磁带驱动器和磁带库等技术。然而,如今有各种各样的选择,包括但不限于以下内容:

- 磁带设备
- 虚拟磁带库 ■ 磁盘池
- 云服务

让我们仔细看看每一个。

备份至磁带

磁带备份 (有时被 SNIA 称为 B2T) 可能是最古老和最常见的备份形式。

胶带的优点

磁带备份包括将数据备份到线性磁带设备。尽管磁带看起来像是一项古老的技术,但现代磁带驱动器和磁带库具有许多优点,包括:

- 容量大
- 高顺序性能
- 低功耗

在性能方面,磁带的性质几乎完全适合备份性能要求。备份作业的典型工作负载是流式传输顺序数据,而磁带在顺序工作负载方面表现良好。然而,磁带的随机性能很差。

但随机性能并不是大多数备份解决方案的首要要求。

胶带的缺点

磁带作为备份介质也有一些缺点:

- 介质降级
- 技术更新
- 对于某些还原类型来说不太方便

作为一种磁性存储介质,磁带会随着时间的推移而退化,尤其是当它们没有在最佳条件下存储时。温度可控、湿度可控等等。在这种理想条件下存储磁带需要花钱。安全的异地磁带存储(有时称为保管)可能更加昂贵。

还有一种担心是,当您在 10 年后从磁带中恢复数据时,即使您将磁带存储在最佳条件下,您也不再拥有任何可以读取它们的磁带驱动器或软件。

此外,磁带非常适合从完整备份中恢复数据,但如果必须基于差异备份执行恢复,它们可能会很麻烦。例如,装载和卸载多个磁带很麻烦,因为要在磁带卷轴上来回快进和倒带以找到您想要的数据在卷轴上的正确位置也很麻烦。

由于这些担忧,以及旋转磁盘介质成本的下降,许多人开始不再使用磁带进行备份。尽管如此,磁带仍然很受欢迎,而且仍然有市场。

LTO 技术

有各种磁带技术可供选择,每种技术都需要特定类型的磁带驱动器,有时还需要磁带库。然而,开放系统领域最流行的磁带格式是 Linear Tape-Open (LTO) 格式。

在 LTO 标准中,LTO 磁带有多种版本,每个版本都用数字表示。当前的 LTO 版本包括 LTO-1、LTO-2、LTO-3、LTO-4、LTO-5 和 LTO-6。

按照LTO标准,它们具有表11.1所示的性能、容量规格和其他特性。

表 11.1 LTO 规格

特征	LTO-1	LTO-2	LTO-3	LTO-4	LTO-5	LTO-6
未压缩容量	100 GB	200 GB	400 GB	800 GB	1.5 TB	2.5 TB
压缩容量	200 GB	400 GB	800 GB	1.6 TB	3 TB	6.25 TB
未压缩速度 (MBps)	20	40	80	120	140	160
压缩速度 (MBps)	40	80	160	240	280	400
压缩率	2:1	2:1	2:1	2:1	2:1	2.5:1
加密	不	不	不	是的	是的	是的
虫	不	不	是的	是的	是的	是的

LTO 标准还要求驱动器和介质之间具有以下向后兼容性：

- LTO 驱动器将能够读取同代磁带上的数据,此外还能读取两台前几代。
- LTO 驱动器能够将数据写入其自身一代以及其前一代的磁带中。

擦鞋

“擦鞋”现象是磁带备份领域的一种现象,当发送到磁带驱动器的数据的速度与磁带驱动器的运行速度不匹配时就会发生。

发生这种情况时,磁带驱动器必须偶尔减速或暂时停止,以便从媒体服务器传入的数据再次积累起来。磁带卷轴的这种启动和停止,或磁带的减速然后加速,看起来有点像用布来回擦鞋的动作。虽然偶尔擦鞋会减慢备份工作的速度,但过度擦鞋会导致磁带盒以及磁带驱动器的磁头和电机过早磨损。

多路复用

多路复用将多个备份作业的数据发送到单个磁带驱动器。目的是提高数据发送到磁带驱动器的速度,使磁带驱动器能够以最大速度运行。这非常有助于避免擦鞋。

但是,多路复用会影响恢复性能。这是因为多路复用到同一磁带驱动器的多个数据流都是交错在一起的。当您从其中一个作业(其中一个流)恢复数据时,必须读取所有多路复用作业的所有数据,并且必须丢弃不进行恢复的作业的数据,从而降低驱动器和恢复操作的整体性能。

一般来说,多路复用可以加快备份速度,但会降低恢复速度。当考虑部署多路复用,您需要权衡恢复速度较慢的缺点与执行多个备份作业的好处。

磁带压缩

大多数磁带驱动器技术(包括LTO)都支持原生压缩。LTO规范使用一种称为流式无损数据压缩(SLDC)的压缩形式。尽管SLDC不是最有效的压缩算法,但它速度很快,并且被认为是良好压缩和良好性能之间的良好平衡。



加密数据的压缩效果不佳。尝试压缩加密数据可能会导致数据集在压缩前更大。幸运的是,大多数在将数据发送到磁带之前加密数据的备份技术都会将其标记为加密。这样就不会尝试对其进行压缩。

如表11.1所示,LTO标准的每个版本都列出了其原生压缩比率作为一项功能。此原生磁带压缩是一项始终启用的功能。

显然,压缩可以显著减少所需的磁带数量支持您的备用资产,帮助降低成本。

磁带加密

过去几年,备份磁带丢失的事件屡见不鲜,引起了行业媒体甚至全国媒体的广泛关注。这凸显了加密所有离开公司数据中心的数据的必要性。磁带备份也不例外。确保对磁带进行加密!

LTO标准定义了从LTO-4开始支持加密的磁带驱动器,但并不强制要求所有LTO-4驱动器都支持加密。因此,如果您计划使用基于磁带驱动器的加密,请务必检查所购买的任何新LTO磁带驱动器的规格。

LTO使用的加密算法是对称分组密码(256位AES-GCM),使用相同的密钥对数据进行加密和解密。对于基于LTO驱动器的加密,需要第三方密钥管理器进行密钥管理。

幸运的是,大多数好的备份应用程序都会执行密钥管理任务。

在性能方面,大多数优质驱动器都能够完全执行加密驱动器内的线速度。

由于加密数据无法压缩,因此应在数据被压缩,这正是基于LTO驱动器的加密所发生的情况。

虚拟磁带库

虚拟磁带库 (VTL) 的兴起是备份行业向前迈出的一大步。VTL 是从基于磁带的备份目标转向磁盘备份目标的第一步,它为将数据备份到基于磁盘的目标的更原生方法铺平了道路。

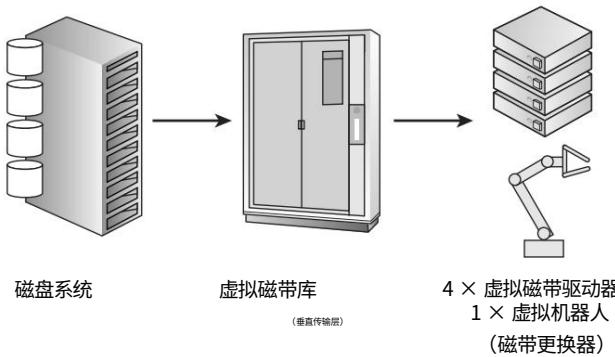
在 VTL 出现时,许多公司已经建立了使用磁带作为备份介质的实践、程序和政策。在这些情况下,VTL 技术帮助这些公司开始使用基于磁盘的备份目标,而无需改变所有现有的流程和实践。

如今,几乎地球上的所有备份应用程序都支持本机备份到磁盘,而无需磁带模拟 (VTL)。事实上,大多数现代备份应用程序都具有比备份到磁带更强大的备份到磁盘选项,有些备份应用程序甚至不支持备份到磁带!

从高层次上讲,VTL 技术将基于磁盘的存储阵列分割成磁带库和磁带驱动器,因此得名虚拟磁带库。这种虚拟化的最初要求是使现有备份软件能够备份到基于磁盘的备份目标,而无需修改以理解磁盘池,基本上就是将基于磁盘的备份目标看起来和感觉起来完全像磁带库,以便现有备份软件可以与其通信。

图 11.6 显示了磁盘阵列被虚拟化为 VTL,该 VTL 配置了四个虚拟磁带驱动器和一个机械手。

图 11.6 VTL 仿真



Strictly speaking, VTL technology is not a true tape backup technology. The main reason is that it adopts disk-based storage and makes it look like a tape library. This is achieved through simulation technology, making it easier to combine with tape backup systems and make it more user-friendly.

VTL 解决方案开创并推动市场接受的一项重要技术
重复数据删除是其中之一。虽然重复数据删除并非 VTL 技术独有的功能,但 VTL 解决方案
几乎总是执行本机重复数据删除。

主要有两个原因。首先,备份工作负载非常适合重复数据删除。我们一次又一次地备份相同的数据集,而且通常大部分数据都没有改变,因此重复数据删除效果会很好。其次,为了与传统磁带库具有价格竞争力,基于磁盘的 VTL 系统需要一种允许它们在更少的磁盘上存储更多数据的技术。



有关重复数据删除技术的详细介绍,请参阅第 10 章“容量优化技术” Chapter 10, “Capacity Optimization Technologies.”

备份至磁盘

备份到磁盘 (B2D) (有时称为磁盘到磁盘 (D2D)) 与备份到磁带 (B2T) 类似,但目标介质是磁盘而不是磁带。磁盘与磁带相比具有以下优势:

- 卓越的随机存取能力
- 卓越的并发读写操作
- 可靠性更高

尽管在备份方面,随机访问并不总是被视为关键的性能考虑因素 (毕竟,许多备份工作负载都是高度连续的),但良好的随机性能仍然有好处。其中一个优势是小文件恢复 (这是最常见的恢复操作类型之一)。使用基于磁盘的恢复,系统几乎可以立即找到所需的数据并开始恢复。使用基于磁带的系统,需要花时间加载磁带,然后快进或倒回到所需位置。

基于磁盘的系统在并发读写操作方面也非常出色。这如果创建合成完整备份,则会很有帮助,其中系统在创建合成完整备份映像期间同时进行读取和写入 (读取上一个完整备份和所需的增量,同时写出合成完整映像)。

另一个可以提供帮助的地方是并发备份 (写入) 和恢复 (读取) 操作,尽管公平地说,同时从同一磁带备份和恢复的情况并不常见。但是,如果您没有很多磁带驱动器,您可能会发现自己需要卸载正在备份的磁带,以便加载需要从中恢复的磁带。基于磁盘的备份目标具有出色的并发读写性能,可以在这里为您提供帮助。

一般而言,备份到磁盘通常可以提供更快的恢复时间。例如,无需等待磁带加载和卷绕到位。此外,某些磁带驱动器的本机顺序性能可能会受到多路复用备份的限制,这意味着基于磁带的恢复并不总是尽可能快。

基于磁盘的备份目标提供比磁带更可靠的备份介质。这是基于这样一个事实:尽管磁盘和磁带都容易发生故障,但磁盘阵列几乎总是受到 RAID 技术的保护。最终结果是,当基于磁盘的系统中磁盘发生故障时,不会丢失任何数据。另一方面,磁带很少受到保护,因为备份作业很少复制到多个磁带上以应对磁带故障的情况。

执行磁盘备份时的一个重要考虑因素是备份软件需要能够理解基于磁盘的备份目标（例如磁盘池），因为它们与磁带驱动器和磁带盒有显著不同。首先，没有磁带盒、临时池、条形码和其他与磁带相关的方面的概念。但是，所有优秀的现代备份应用程序都对基于磁盘的备份目标提供了广泛的支持。

磁盘备份通常通过 IP 网络进行，并为磁盘池提供文件系统接口，例如 NFS 和 SMB/CIFS。它还允许多个备份和恢复操作同时访问磁盘池。

在大多数环境中，使用磁盘备份并不意味着磁带的终结。在实际事实上，磁盘备份正被用于增强现有的基于磁带的备份解决方案，从而显著改善它们。一种常见的方法是将备份分阶段备份到磁盘，然后最终将其复制到磁带，这种方案通常称为磁盘到磁盘到磁带 (D2D2T)。备份被制作到磁盘目标并在那里保存长达一周左右。在此期间，它们最有可能需要用于恢复操作，而且由于它们在磁盘上，因此它们可以提供极好的恢复时间。在此之后，作业被转移到磁带上，然后将它们带到异地并存储以供长期保留。这种方法的一个缺点是需要额外的复制作业来将备份映像从磁盘目标复制到磁带目标。

备份至云

如今，在 IT 领域，云计算正在大行其道。备份领域也不例外。

使用云，您通常使用服务并根据使用服务的量付费。这种商业模式通常被称为消费模式或基于消费的计费。您还可以快速轻松地增加或减少使用云消费的服务金额。

对于备份和恢复，云作为备份目标很有吸引力。从某种意义上说，它几乎是理想的，因为在将备份存储到云后，您不需要快速或频繁地访问它们。然而，最大的问题是如何将备份数据传输到云。一种常见的方法是部署磁盘到磁盘到云 (D2D2C) 架构。这类似于磁盘到磁盘到磁带 (D2D2T)，磁带被云取代。在这种架构中，磁盘到磁盘部分仍然发生在您的数据中心内，利用重复数据删除技术，并仅将重复数据删除的数据传输到云进行长期保留，预计不会像中间磁盘备份平台上的数据那样频繁地进行恢复操作。重复数据删除技术和合成完整备份或其他永久增量方法也有助于此。还应该注意，在这个 D2D2C 架构中，只有一部分备份资产存在于云中。

D2D2C 架构如图 11.7 所示。该过程分为两个步骤：

1. 将数据从磁盘备份到磁盘
2. 将相同的数据从磁盘复制到云端

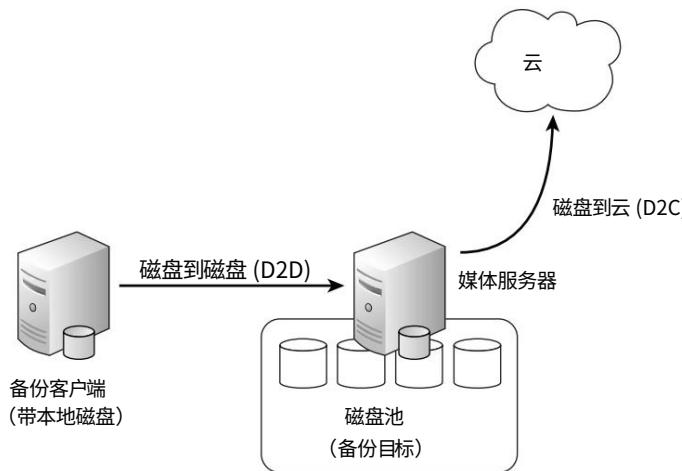
这两个步骤结合在一起，形成完整的 D2D2C 模型。第 1 步和第 2 步之间通常会有几天时间，在此期间，备份映像会保留在现场和

磁盘进行快速恢复。经过一小段时间（例如一周左右），备份映像即可移出到云端。

然而，即使使用基于本地磁盘的备份目标和重复数据删除技术，如果
如果您每个月备份的数据太多，那么使用云作为备份目标对您来说并不合适。将数据上传到云需要花费太长时间。

此外，如果您需要执行大规模恢复，则必须考虑需要多长时间才能恢复。显然，这取决于您的重复数据删除率和与云备份提
供商的网络连接（互联网链接通常不会切断它），但如果每月备份约 10 TB 或更多，则使用云作为备份目标可能会
很有挑战性。

图 11.7 磁盘到磁盘到云



即使考虑到所有这些因素，Gartner 预计，300 多亿美元中的 40%
备份市场将在 2016 年之前转向云端。2013 年末，亚马逊网络服务（全球最大的公共云提供商）在其云存储平台上推出了
VTL 网关，极大地简化了将云存储纳入企业备份环境的途径。

如果你决定使用云作为备份目标，你需要确保加密
您发送到云的所有数据以及您拥有的加密密钥，并将其保存在云服务之外。您还需要一个明确的退出计划，以便在您
对服务不满意的情况下收回您的数据。

备份保留策略

备份环境需要备份保留策略。这些策略规定了备份映像的保留时间。例如，您可能希望将每周的完整备份映像保留
365 天，但将每日增量映像仅保留 28 天。

目前最流行的保留方案是祖父-父亲-儿子 (GFS) 方案。GFS 方案基于三代备份,如下所示:

祖父:每月备份

父亲:每周备份

儿子:每日备份

GFS 方案还针对每一代备份应用保留策略。为了更好地理解其工作原理,您应该逐步了解创建此类策略的步骤。

练习 11.2 展示如何设置基本备份保留策略。

练习 11.2

创建备份保留策略

在本练习中,您将设置基本的备份保留策略。您将使用流行的 GFS 方案。

1. 确定您将使用的三代备份。继续使用此

方案,该方案在许多情况下相当实用:

■ 祖父:每月备份

■ 父亲:每周备份

■ 儿子:每日备份

2. 确定要应用于每一代备份的 GFS 保留策略。

这些是您的备份保留策略:

■ 每月备份:保留 60 个月 (5 年)

■ 每周备份:保留 52 周 (1 年)

■ 每日备份:保留 28 天

3. 根据此特定保留方案设计备份策略。以下是您可以使用的基本方法:

a. 每周五晚进行完整备份,并将每次备份归类为每周备份。

b. 将每个日历月的第一个每周备份指定为每月备份。

根据前面的政策,每个备份都将过期 (每周备份在 52 周后过期,每月备份在 60 个月后过期)。

c. 周六至周四进行增量备份,备份到期时间为

28 天。

d. 如果您使用磁带作为备份介质,则可以添加以下要求:将每周和每月的完整备份存储在异地,但将每日增量备份保留在现场。

这是一个简单但功能强大的备份保留策略的基础,我们可以将其作为服务级别协议的一部分发布给我们的用户,如下所示:

- 1.可以从前一天晚上的备份中恢复过去28天的所有数据（RPO）,并且可以在一个工作日内恢复（RTO）。
- 2.所有28到365天之间的数据都可以恢复到最近的周五晚上的数据备份（RPO）,并且可以在两个工作日内恢复（RTO）。
- 3.所有1至5年之间的数据均可恢复至最近的第一个星期五
个月内即可恢复（RPO）,两个工作日内即可恢复（RTO）。

在此示例 SLA 中,我们重点介绍了 RPO 和 RTO。例如,在第一项中,我们指出,我们可以从过去 28 天恢复数据的时间点是前一天晚上,因此我们的 RPO 不超过 24 小时。我们还指出,恢复这些数据所需的时间不会超过一个工作日,因此我们的 RTO 为一个工作日。

显然,这种备份保留策略并不适合所有情况,但它是一个易于调整的好起点。

当依赖增量和差异备份时,请确保保留它们所依赖的完整备份（增量或差异备份之前的最后一个完整备份）。在基于 GFS 的系统中（如练习 11.2 中描述的系统）,情况将始终如此,因为我们会保留一年的每周完整备份和五年的每月完整备份。



建议 SLA 应与业务部门或指定为业务联络人的 IT 员员协商。协议还应发布给 IT 和业务部门,并应定期审查。as the liaison with the business. The agreements should also

归档

首先,最重要的是,存档和备份是不一样的!这就像说汽车和飞机是一样的,因为它们都是将人们从一个地方运送到另一个地方。

但尝试一下坐飞机去杂货店,或者开车去世界另一端的某个地方。你可以同时做这两件事,但都不是理想的选择。备份和归档技术及其用例也是如此。

备份和归档比较

让我们快速比较和对比备份和归档。

在大多数情况下,备份旨在保护当前或最近使用的数据。备份提供中速到快速的数据恢复,通常是在数据丢失、删除或损坏的情况下。它有点像一份保险单,以防您丢失数据并需要快速恢复数据以继续运营业务。

备份涉及文件、文件夹、应用程序和服务的操作恢复。

典型用例包括：

- 用户请求的恢复
- 灾难恢复 (DR) 和业务连续性 (BC)

关于用户请求恢复备份的话题，我们可能都遇到过。用户打电话或向服务台提交工单，要求恢复特定的文件或文件夹，通常是因为他们刚刚删除了它。然后使用已删除文件或文件夹的先前备份来恢复丢失的数据。我们都交叉手指，希望有一个有效的备份可以恢复。

使用备份进行灾难恢复则有所不同。作为最后的手段，在尝试了可能更快的灾难恢复方法（如复制）之后，可以使用备份来恢复整个系统，无论是在本地还是远程站点。

归档用于高效存储不再使用或很少使用但需要保留的数据。就电子邮件归档而言，电子邮件的副本是在发送和接收时制作的，因此在进入归档之前无法被篡改。归档和检索数据的最常见原因是法律合规性要求。存储在归档中的大多数数据都需要防篡改，以便保证从存储时起数据就不会发生变化。

在将备份与存档进行比较时，最好将备份视为将丢失的数据放回其原始位置，以便将其用于其原始用途。备份的对象根据其位置进行索引。另一方面，存档用于根据其内容定位数据，通常将数据复制到新位置，以便将其用于其他目的 - 通常是企业或法律合规性。存档通常进行全文索引，因此您可以根据其内容而不是其位置来定位项目。



提供防篡改数据存储的档案系统据说可以提供内容保证或内容真实性。术语“不可否认性”通常是指提供内容真实性的档案系统，表示您无法否认数据的真实性。*nonrepudiation* is

尽管备份和档案并不相同，但它们确实相辅相成。

一个例子是，归档数据通常不需要像非归档数据那样进行备份。这是因为归档中的数据不会发生变化。只要备份并存储在可靠的归档平台上，就可能永远不需要再次备份。这意味着将旧的、未使用的数据移动到归档解决方案可以帮助减少您每月备份的数据量。

遵守

法规遵从性无疑是采用归档解决方案的最大驱动因素。大多数发达国家都对数据的存储、检索和最终删除有法律要求，尤其是个人或企业敏感数据。在美国，2002 年的《萨班斯-奥克斯利法案》(SOX) 是其中最受关注的法案之一。

众所周知的例子。其他包括支付卡行业数据安全标准 (PCI DSS) 以及在世界不同国家和地区合法的各种数据保护法。如果您的企业是全球性企业,您可能需要纳入并遵守多项合规法规。

然而,在世界各地,大多数法律要求档案中的数据满足遵循以下标准:

- 防篡改
- 在指定时间后删除
- 提供审计跟踪

在存储领域,各种技术都可以提供防篡改功能,我们通常将它们称为一次写入,多次读取 (WORM) 技术。一些技术,例如某些 DVD 和蓝光光盘技术,本身就提供此功能。其他技术,例如基于磁盘的解决方案,需要以特定方式设计或为其编写特殊固件,使其存储的数据防篡改。在广受欢迎的基于磁盘的合规性归档领域中,一种这样的技术是内容可寻址存储 (CAS)。



有关内容可寻址存储的深入讨论,请参阅第 7 章“文件、NAS 和对象”, see Chapter 7, “Files, NAS, and Objects.”

归档解决方案通常还需要提供详细的审计日志,以显示谁有权访问档案中的数据以及谁访问了这些数据。

此外,根据某些法规,在一定时间后删除敏感数据与保留敏感数据同样重要。许多法规规定,在指定的时间后,必须从档案中清除数据。

公司保留数据副本的时间长短由公司的法律和合规人员决定。但是,数据保留时间过长可能会产生负面影响。例如,如果您的政策规定电子邮件备份仅保留 30 天,那么您显然无法恢复 30 天前的电子邮件。



在现实世界中,越来越多的公司急于尽快删除和清除数据。这种行为主要有两个原因。首先,删除数据可以释放档案中的宝贵空间。其次,出于监管原因,档案中的数据可能(并且经常)在诉讼和其他情况下被用来对付公司,从而对公司处以罚款和类似处罚。出于这两个原因,许多公司都渴望尽快清除数据。

发现

档案需要编入索引并可快速搜索。当请求从前一晚的备份中恢复丢失或删除的文件时,记录呼叫的用户知道文件名

和文件位置。相比之下,当传唤文件或电子邮件时,提供的信息很可能不会那么具体。例如,如果法院传唤电子邮件,它很可能会要求提供类似“XYZ 高管在 2010 年 7 月至 8 月之间发送的有关 XYZ 产品折扣价格的电子邮件”的内容。这比普通用户请求恢复文件要模糊得多。因此,归档解决方案的首要要求是能够快速且广泛地进行搜索。如果您无法在法规规定的时间内找到并检索法院要求的数据,您的企业几乎肯定会受到处罚。

档案媒体

磁盘非常适合用作存档介质。它提供大容量、良好的随机访问和可靠性,当与 RAID 技术结合使用时。而且它的价格越来越便宜。

由于存档中的数据仅存储一次,因此存储存档数据的介质的可靠性至关重要。如果该介质损坏或发生故障,您将丢失存档中该数据的唯一副本,这将使您陷入非常困境,尤其是当您必须响应对该数据的合法请求时。在可靠性方面,基于磁盘的存档几乎总是受到 RAID 类型技术的保护,这使得它们非常可靠,并且比磁带更可靠。

当谈到技术更新时,基于磁盘的存档也比磁带好很多。

尽管磁带和磁盘每隔几年都需要更新,但磁带技术更新尤其痛苦且耗费人力。没人愿意这样做。因此,磁带更新经常被忽视,直到您发现自己在非常旧的磁带上有很多旧的备份,无法再恢复。而且由于磁带通常存储在异地,因此“眼不见,心不烦”这一真理适用,我们经常忘记需要更新它们。

另一方面,基于磁盘的存档位于数据中心的中央,每天都在消耗电力,这意味着您不太可能忘记它们。此外,将数据从基于磁盘的存档迁移到较新的基于磁盘的存档的过程比将旧磁带复制到新磁带要简单得多,工作量也少得多。

关于电力消耗和相关运营支出成本,一些

基于磁盘的档案库支持旋转停止和大量空闲磁盘阵列(MAID)技术,其中未使用的磁盘驱动器会旋转停止(关闭或进入低功耗模式)以降低档案库的整体功耗。当需要它们时,它们会重新旋转。但是,与从异地存储中调用磁带所需的时间相比,这些额外的旋转时间可以忽略不计。

光学介质(例如 DVD 和蓝光)是存档介质的选择,尽管它们在现实世界中很少见。它们都是可移动的,这意味着它们可以保存,并且都提供原生 WORM 功能。它们也可以很便宜。但是,存档解决方案是一项长期投资,您需要非常认真地考虑如何

您的存档解决方案将在 10 年后出现。这是现实世界中光学介质不足的一个领域。这些都是您需要认真考虑的有效问题：

■您选择的光学介质可以使用多长时间？

■10 年后它的可靠性如何？

■付款人会存在多久？

光学介质在容量和性能方面也存在问题。磁带和磁盘可以存储更多数据，并且通常提供更快的访问速度，而且使用磁盘，您无需浪费时间安装和卸载介质。



Real World Scenario

使用备份作为存档的困难

一家美国零售公司接到传票，要求提供一些与调查相关的数据，但他们没有适当的归档解决方案。相反，他们不得不依靠旧的备份来恢复数据。由于被要求的数据是几年前的，数据所在的备份磁带存储在第三方异地，必须召回现场。由于磁带的年代太久远，召回磁带的时间被推迟了一天。然后，当磁带回到现场时，他们使用的是旧格式的磁带 (SDLT)，而该公司的数据中心不再有可以读取该格式的磁带驱动器。幸运的是，在分支机构找到了一个合适的磁带驱动器，并安排尽快将其运送到数据中心。

这意味着分支机构的本地备份必须禁用，直到 SDLT 驱动器返回。驱动器到达数据中心后，必须构建一个新的媒体服务器来容纳磁带驱动器。此外，由于数据太旧，必须读取每个磁带才能索引其内容。由于已召回多盘磁带，这又花了两到三天时间。最后，数据被恢复并提供给法律部门。

除了所有这些开销之外，由于介质退化，磁带无法读取的风险也随之增加。在这个特定情况下，这种情况并没有发生，但很有可能在 IT 人员努力召回磁带、借用磁带驱动器和构建新服务器之后，磁带仍然无法读取。

如果公司在备份数据之后更改了备份应用程序，那么恢复数据也会变得更加复杂！

这种体验与使用专门构建的归档解决方案形成鲜明对比，在专用归档解决方案中，数据存储在经过维护（磁盘经过清理和 RAID 保护）、索引和可搜索的可靠介质上，并且可以通过 SMB/CIFS 和 NFS 等行业标准协议访问数据。

概括

在本章中,我们通过讨论备份的原因以及备份如此重要的原因来设定背景,并定义一些重要术语,例如恢复点目标、恢复时间目标和备份窗口。然后,我们强调了从备份中恢复的重要性。之后,我们详细讨论了主要的备份类型,包括热备份、基于 LAN 的备份、无 LAN 备份、无服务器备份和 NDMP 备份。然后,我们讨论了完整备份、增量备份和差异备份之间的区别,以及如何有效地将它们用作备份保留方案的一部分。然后,我们继续讨论不同的备份目标技术,例如磁带、磁盘和云,并在本章结束时讨论了归档技术和合规性归档。

章节概要

备份和业务连续性 备份是灾难恢复和业务连续性计划的重要组成部分。备份的可靠性和可轻松、可靠地恢复至关重要。花点时间测试您的恢复过程,这样您就不会在紧急情况下试图从特定类型的备份中恢复数据。

热备份 热备份是在系统保持在线时进行的应用程序一致性备份。备份不需要系统停机,并且应用程序保证备份的一致性,因此可用作恢复点。大多数流行的应用程序和数据库都支持热备份模式。

合成完整备份 合成完整备份减少了创建完整备份映像所需的客户端和网络负载。合成完整备份的工作方式是获取已存储在备份目标上的先前完整备份映像并将其与新的增量映像合并,而不是拖曳要备份的客户端的全部内容并通过网络将所有相关数据拖到备份目标。生成的合并映像正是您进行新完整备份时所获得的图像。不同之处在于,此合成完整备份是通过仅对客户端进行增量备份并将其通过网络传输而创建的。

归档 归档不同于备份数据。归档通常采用合规性归档的形式,符合与数据真实性和不可否认性相关的监管要求。归档需要编入索引并可搜索,为数据提供安全可靠的审计跟踪,并建立在可轻松刷新的可靠介质之上。

档案的规划需要非常长远的考虑。

章节 12



存储管理

本章涵盖的主题：

- ✓ 容量管理
- ✓ 退款和返利
- ✓ 绩效管理
- ✓ 警报
- ✓ 存储资源管理
- ✓ 管理协议和接口



存储管理是一个广泛的主题,涵盖与存储资产管理、监控和相关的所有事项。管理的三个主要领域是容量、性能、

性能和可用性。

在容量方面,主要目标很简单:确保您始终有足够的容量来服务您的应用程序,并确保以最佳方式利用容量。良好的容量管理还需要简单明了的报告和趋势。

在性能方面,存储管理包括确保您的存储资产具有高性能(足以满足您的应用程序和要求),以及识别性能瓶颈并适当调整性能。这其中的主要部分包括提供与性能相关的统计数据。

在可用性方面,存储管理是关于监控和管理服务以及组件级冗余,确保存储服务保持正常运行。

这三个方面可以很容易地概括为良好的存储管理就是确保存储始终处于开启状态(可用性)、始终有空间(容量)并且快速(性能)。如果您的资产符合这一描述,那么您做得很好。

良好的存储管理的一个重要规则是做好准备。当一切顺利时,人们会感到高兴和文明。然而,当情况变得艰难,人们认为他们遇到了性能或可用性问题时,他们很快就会开始责怪每个人和每件事。当这种情况发生时,您需要深入挖掘并提供深入分析和证据的能力。

良好的存储管理需要可靠的流程、政策和工具。良好流程的一个例子是生命周期管理,包括服务器和服务的启用和停用。如果您只是关闭服务器并拔下电源,而不回收与其相关的存储,那么您最终将得到一个利用率极低的存储资产,其中有大量的容量被浪费。在政策方面,制定一项政策,规定对生产系统的所有更改都必须由两名有能力的员工执行,这可能会使您的生产环境免于发生重大中断。而在工具方面,当您因高效率性能问题而陷入困境时,能够提供图表和其他支持数据对您来说将非常有价值。

在本章中,您将深入研究所有这些主题以及更多内容。

容量管理

首先,容量管理就是要确保您的环境不会用尽空间。毕竟,提供存储容量(存储文件和数据的空间)是任何存储资产的基本要求。

然而,容量管理还有很多内容。例如,基于 FC SAN 的存储资产的容量管理包括满足以下所有要求:

- SAN 交换机
- SAN 端口
- SAN 带宽
- 交换机间链路 (ISL)
- 前端阵列端口
- 数据中心电源和冷却
- 数据中心占地面积
- 机柜空间可容纳更多驱动器
- 可用容量许可证

这个清单还可以继续列下去。重点是容量管理不仅仅包括 GB 和 TB 的思考。如果你经历了痛苦的采购订单,为你的存储阵列安装 64 TB 的额外容量,却发现该阵列已经满负荷,那么你会显得有点愚蠢。



Real World Scenario

如果你没有规划好升级的每个细节,会发生什么情况

一家公司正兴高采烈地开展业务,因为他们知道其主存储阵列只占用了一半空间,如果需要,可以轻松扩展。在用完阵列中所有现有容量之前不久,该公司下达了购买和安装扩展柜和 160 TB 存储的采购订单。这些都得到了批准,并订购了升级。然而,在计划安装升级的周末,安装工程师发现数据中心地板上没有足够的空间来安装新的扩展柜。原来,自从首次安装存储阵列以来,数据中心已经安装了额外的服务器机架,其中一个机架安装得离存储阵列有点太近了,这意味着没有足够的空间将扩展柜固定在阵列上。

幸运的是,在这个特定情况下,导致问题的服务器机架中只有一台服务器,并且在接下来的周末服务器被关闭,服务器和机架被移动到数据中心的另一个位置。这不是世界末日,但有点令人尴尬,而且很容易变得更加难以解决。

在规划这个数据中心的布局时,从未发现或提到存储阵列的控制柜比标准数据中心机柜宽一两英寸。事实上,过去存储行业一直是数据中心的糟糕公民,经常推出定制机柜尺寸的技术,使数据中心规划成为一场噩梦。幸运的是,如今大多数存储阵列都开始采用标准机架尺寸,但始终值得仔细检查!

虽然容量管理有很多方面,但大部分内容涉及可用千兆字节和兆兆字节。那么让我们来看看一些影响容量管理的技术,然后谈谈我们如何报告容量管理。

容量报告

在现代社会,IT 常常被视为成本的消耗,因此,高管和股东往往希望确保 IT 能够明智地使用资金和资源。这是件好事。但愿我们都不要浪费资源!然而,这通常需要您能够证明您没有浪费资金和资源,而证明这一点的最佳方法之一就是通过精美的报告。高管们喜欢精美的报告以及用于支持他们的原始电子表格数据!

说到容量报告,没有比向企业高层展示所有资金流向更好的方法来证明 IT 部门的存在了。如果这些钱花在了业务上而不是 IT 上,那就更好了。例如,图 12.1 是一份很好的报告,因为它显示大多数存储资源都被业务部门占用。但另一方面,图 12.2 并不是一份很好的报告。图 12.2 显示,与前台业务线用户相比,IT 是存储资源的巨大消耗者(存储资源需要花钱)。

在图 12.1 和 12.2 中,我们将许多共享的公司功能组合在一起。这些功能包括企业电子邮件、电子邮件归档、财务系统、人力资源系统和其他类似服务,这些服务在整个企业范围内共享。这是保持报告可读性的常见做法。但是,请做好准备,因为管理层可能会要求您提供单独的明细,说明哪些内容占用了共享企业空间中的存储空间。

产能报告的其他示例包括:
不同的区域或环境,例如:

- 生产
- 发展
- 分期
- 研究

图 12.1 IT 作为容量的小消耗者

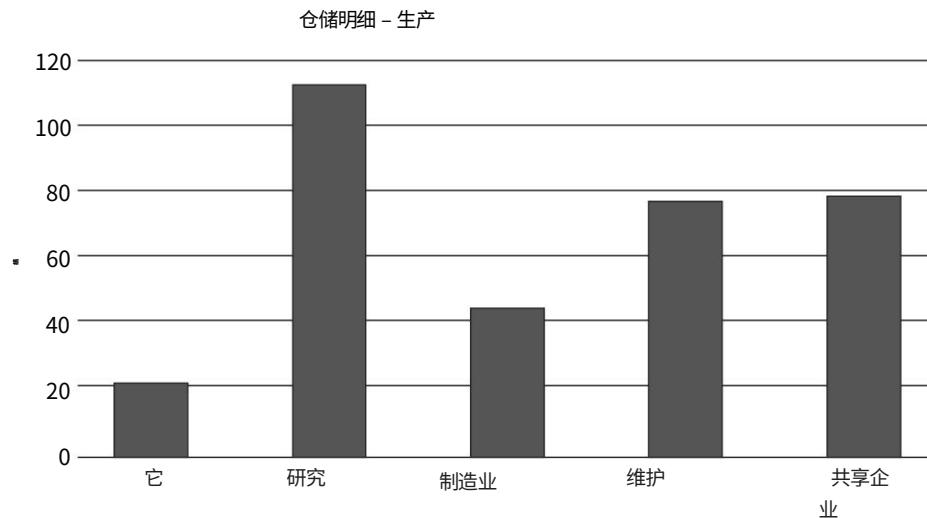


图 12.2 IT 作为容量消耗大户

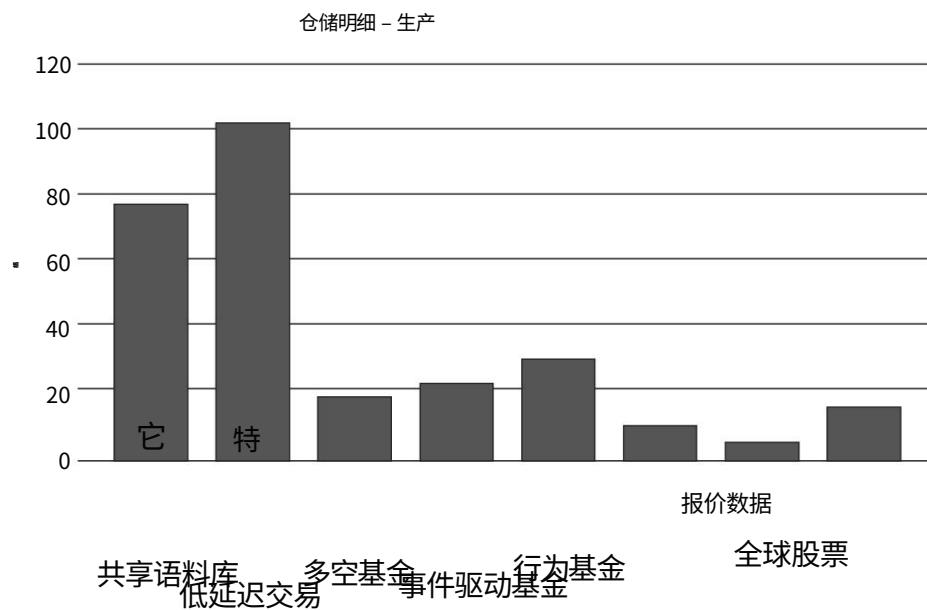


图 12.3 显示了这些环境的细分，并添加了分离。生产分为生产实时和生产灾难恢复。这种对生产实时和生产灾难恢复的额外细分非常常见，因为它突出显示了为灾难恢复 (DR) 目的消耗了多少存储空间。

图 12.3 按环境划分的容量细分

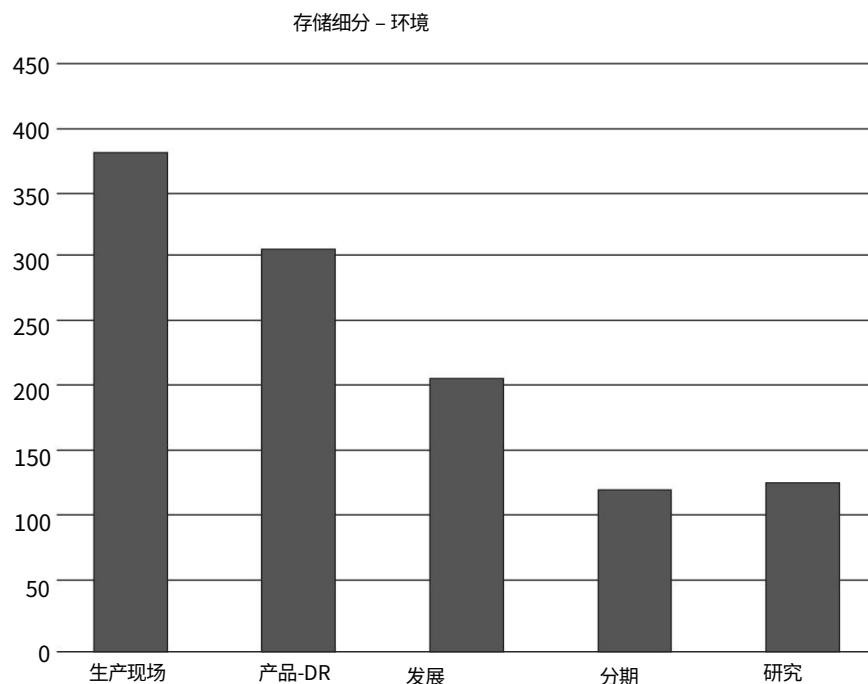


图 12.4 以饼图形式显示了相同的环境数据细目分类。
您可能还会被要求显示年初至今的净容量使用情况等数据，以便管理层了解今年谁的存储消耗速度最快。图 12.5 显示了一家拥有多条业务线的公司此类报告的示例：

- 食品销售
- 家居用品销售
- 保险
- 网上银行

图 12.4 按环境细分的饼图

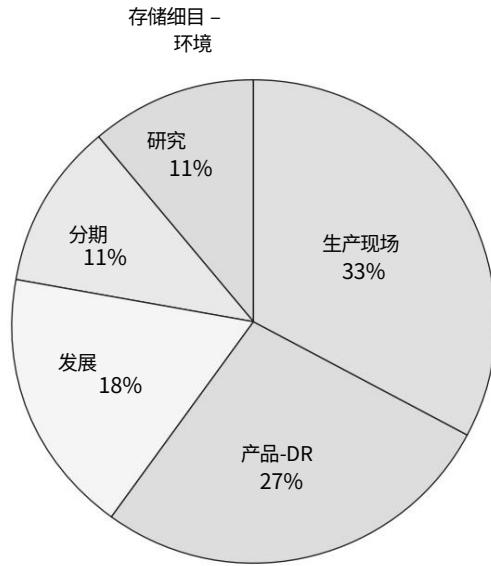
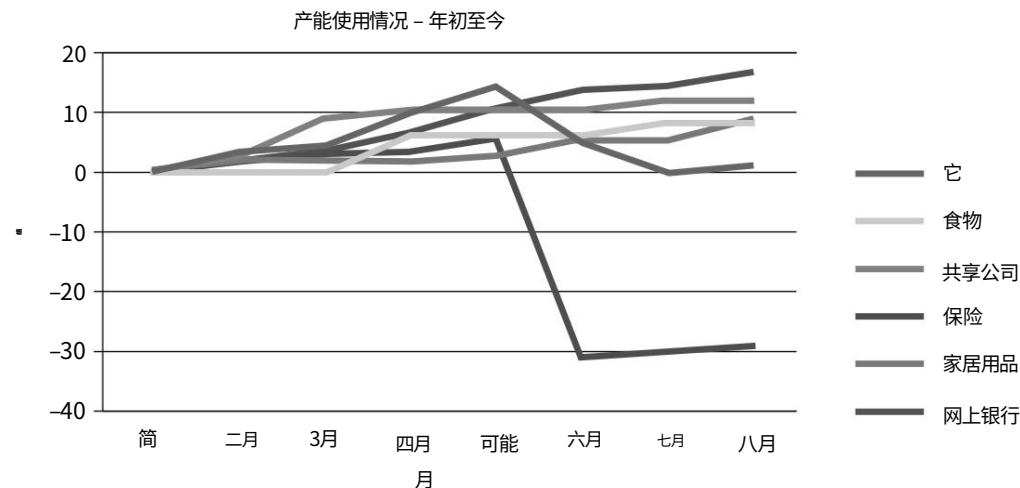


图 12.5 年初至今的净产能使用情况



精简配置注意事项

尽管精简配置 (TP) 是一项出色的技术,具有巨大的容量相关优势,但它使容量管理变得更加困难!现在让我们讨论一下过度配置可能给容量管理带来的挑战,以及它为预测和趋势增加的额外重要性。

过度配置

在容量管理方面,TP 的问题在于它允许我们过度配置阵列。有关精简配置和过度配置的详细讨论,请参阅第 10 章“容量优化技术”。不过,现在我们只能说过度配置允许我们假装拥有比实际更多的存储空间。这当然很好,但它带来了巨大的新风险,如果我们不能妥善管理这些风险,我们可能会失去工作。例如,如果有足够的系统揭穿我们的谎言并要求分配全部存储空间,我们可能没有足够的容量给他们。如果发生这种情况,系统就会开始崩溃。



过度配置有时也称为超额订购或过度分配。术语过度配置也用于固态存储设备的上下文中,表示漫游的操作系统假装其容量小于实际容量!显然,这与精简配置的含义相反。在固态设备(例如 devices 内存驱动器上),该设备可能具有 600 GB 的实际容量,但会使用隐藏的 200 GB 来执行日常操作,从而延长设备的使用寿命。但是,在精简配置的上下文中进行过度配置时,我们会假装拥有比实际更多的容量。

On a solid-state device, such as a 400 GB flash memory drive, the device might have 600 GB of actual capacity but be using the hidden 200 GB to perform housekeeping operations and thus increase the working life span of the device. However, when overprovisioning in the context of thin provisioning, we pretend we have more capacity than we actually do.

图 12-61 显示了过度配置阵列的逻辑表示。该阵列具有 100 TB 的物理容量,但过度配置了 100%。这意味着它假装拥有 200 TB 的容量。当阵列相对未使用时,这种情况是正常的,可能实际使用的数据为 50–70 TB,因为这将需要阵列进行大量运行 array is relatively unused, maybe with 50–70 TB of actual used data, because it would take quite a big run on the array (想像一下“银行挤兑”!) 用完剩余的 30 或 50 TB 数据。但是,如果使用了 90–95 TB 的阵列,您很可能很快就会用完容量。因此,您最好摆脱一些不必要的数据,可能将其移动到另一个系统或删除它们,或者尽快安装更多容量, maybe move it to another system or delete it—or install some more capacity very quickly!

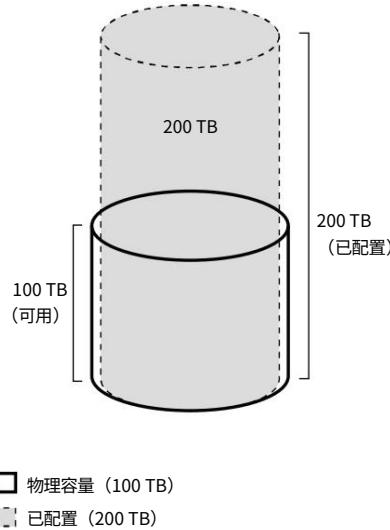
无论如何,精简配置和过度配置将继续存在,所以我们不妨使用工具, so we may as well tool 以便我们可以有效地管理过度配置的阵列的容量。



强烈建议您在过度配置之前先获得相应组织 IT 管理人员的同意。这应该包括在可预测的时间内购买额外容量的协议,因为过度配置会带来额外的风险 of overprovisioning. This should include an agreement to purchase additional capacity in a predictable amount of time, because of the added risk overprovisioning brings.

您最不想做的事情就是在未获得 IT 资产和服务交付总体负责人同意的情况下引入新的风险。 agreement from those who have overall responsibility for the IT estate and service delivery.

图 12.6 过度配置阵列的逻辑表示



趋势的必要性

获得实施过度配置的批准后,您肯定希望慢慢开始。从过度配置一小部分开始,可能是 10% 到 20%。然后确保您对几个月的趋势和预测有把握;我建议至少四到六个月。一旦您很好地了解了您的资产和阵列的容量相关特征,您可能希望将过度配置推向更大胆的水平,例如 30% 到 40%。继续趋势和预测,并重复这个循环。对于您应该或不应该过度配置多少,没有任意的限制;这实际上取决于您的环境。这里的关键是了解您的环境的增长和特征。话虽如此,不建议过度配置太多!

为了了解和报告过度配置环境的趋势,需要了解和关注的关键指标如下:

物理容量 这指的是阵列的可用容量。将其视为您批准的支出预算;它是安全、可靠且经过批准的,您可以根据需要进行支出。

您使用的量不能超过此值。您可以配置量超过此值,但您使用的量不能超过此值。

预置容量 这是您分配给主机的容量,或者换句话说,您假装拥有的容量。当过度配置时,您假装拥有的容量超过阵列的物理容量。

已用容量 这是您的主机实际写入的容量量。

对于每个阵列,都必须了解并单独跟踪这些指标。

让我们用一个简单的例子来概括所有这些。我们可以拥有一个物理容量为 100 TB 的阵列,如果我们启用过度配置,我们可以假装我们拥有的容量超过这个数字。例如,假设我们为连接的主机配置了总计 150 TB 的容量。150 TB 的值将是我们的配置容量数字,使我们的阵列超额配置了 50%,有时也称为 150% 的配置。但是,在我们配置的 150 TB 中,假设我们的主机只写入了其中的 60 TB。这使得 60 TB 成为我们的已用容量数字。总结如下:

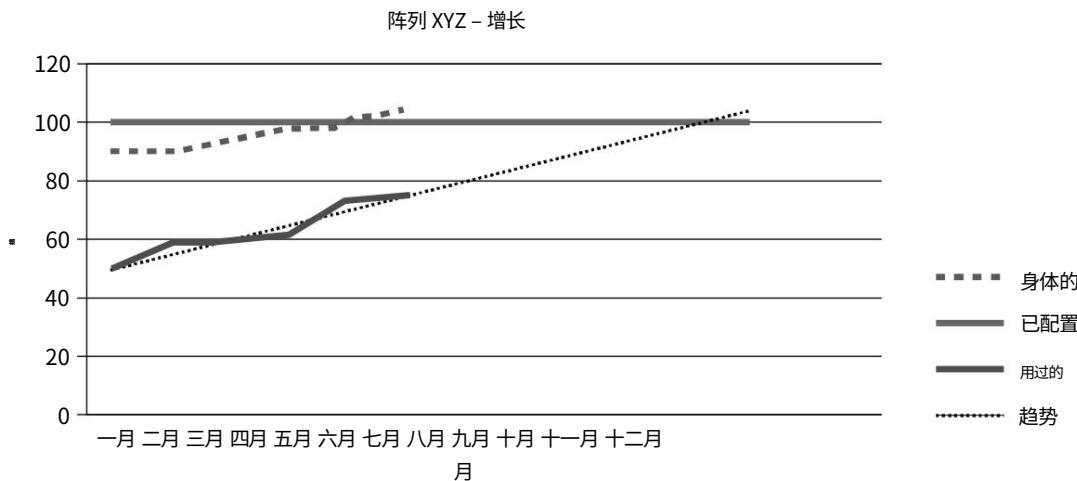
- 物理安装容量 :100 TB
- 预配置容量 :150 TB
- 实际使用容量 :60 TB
- 过度配置百分比 :50%

现在,如果我们使用的容量每月增加 5 TB,那么我们的使用容量需要八个月才能消耗掉所有物理容量。如果发生这种情况,一切都会变得一团糟!我们的主机仍然认为我们有 50 TB 的可用空间,但我们没有,因为我们对它们撒谎了我们有多少空间。因此,进入我们阵列的任何新写入都将失败,应用程序和服务将开始崩溃!

因此,监控和趋势,然后再监控和趋势。确保在可用空间用完之前购买更多容量!实现此目标的一个重要部分是了解您的采购周期以及供应商交付新套件的交货时间。如果您需要三周时间才能获得采购订单批准,而您的供应商需要四周时间才能交付新设备,那么等到银行里剩下一个月的存储空间就毫无意义了!

图 12.7 显示了可用于刚刚举的例子。

图 12.7 过度配置趋势





跟踪业务趋势也很重要，例如收购新公司的计划、开拓新市场的计划、迁移现有应用程序的计划等。所有这些都会对遗产的容量规划产生重大影响。例如，用新应用程序替换旧的应用程序可能需要在短时间内运行这两个应用程序，在此期间您的存储需求可能会翻倍。

如果您的环境足够小且易于管理，另一个好主意可能是确保您的阵列上始终有足够的可用容量（物理容量减去已用容量）以便能够处理任何一台服务器消耗其所有分配容量的情况。例如，如果您的服务器有 5 TB 的分配容量，但该服务器只消耗了 2 TB，它可能会决定写入剩余的 3 TB。为了保护您的阵列免受这种情况的影响，您需要在阵列上至少有 3 TB 的可用容量（*per is consuming only 2 TB, it could potentially decide to write to the remaining 3 TB.* To protect your array from that scenario, you'd need to have at least 3 TB of free capacity on the array. 阵列中的可用容量少于 3 TB 将导致系统崩溃。*would bring the system down.*

关于过度配置，最后说一句：如果在过度配置中要监控一件事，那就是 or in an overprovisioned environment，其实际使用的容量与实际安装的容量之比 *against actual capacity installed!*

重复数据删除和压缩 Compression

重复数据删除和压缩是第 10 章详细介绍的流行容量优化技术。这两种技术都是为了在更小的空间内存储更多的数据。Both technologies deal with storing more data in less space.

追踪这些技术的有效性对于内部 IT 报告很有用。can be useful for internal IT reporting. 然而，IT 部门以外的管理人员不太可能关心重复数据删除和压缩等技术的有效性。geness of technologies such as deduplication and compression.

重复数据删除的有效性以比率来表示，这里列出了几个例子 a couple of examples listed here:

- 4:1
- 10:1

10:1 的重复数据删除率优于 4:1。10:1 表示我们存储的数据量是未使用重复数据删除时的十倍，而 4:1 表示我们存储的数据量只有未使用重复数据删除时的四倍。这些比率也可以用图示形式很好地表示出来，如图 12.8 所示。This information as we would if we hadn't had deduplication. These ratios can also be expressed nicely in graph format, as shown in Figure 12.8.

在图 12.8 中，我们将重复数据删除效率表示为百分比而不是比率。比率的问题在于它乍一看起来可能会产生误导。例如，比率 2:1 和比率 3:1 之间的差异高达 16%。比率 10:1 和 100:1 之间的差异仅为 9%，而比率 100:1 和 500:1 之间的差异仅为 0.8%。因此，随着比率数字的增加，我们获得的收益会大幅减少。only 9 percent, and the difference between 100:1 and 500:1 is a difference of only 0.8 percent. So we get massively diminishing returns as the ratio numbers get higher.



我们将重复数据删除率转换为节省的空间百分比,如前所述 space saved as follows:

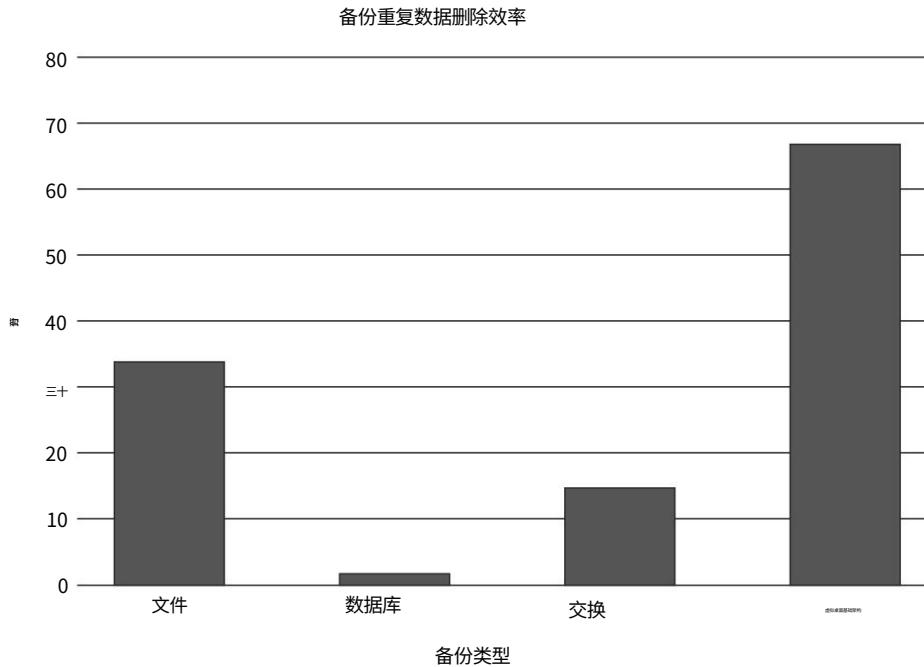
10:1 转换为 $100 - (1/10 \times 100) = 90\% = 90\text{ percent}$

25:1 转换为 $100 - (1/25 \times 100) = 96\%$

2:1 转换为 $100 - (1/2 \times 100) = 50\%$

3:1 换算为 $100 - (1/3 \times 100) = \text{约} 66\%$

图 12.8 重复数据删除效率



请注意,在数据缩减等方面,看似微小的差异可能会产生很大的影响。影响,尤其是在较大的系统上。例如,供应商可能会认为,实现 95% 减少的系统只比实现 90% 减少的系统好 5%。5% 可能看起来是个小数字。但在 100 TB 的数据集上,更好的系统只需要 5 TB 的容量来存储数据,而另一个系统则需要 10 TB。

从容量管理角度来看,在线和后处理重复数据删除技术之间存在很大差异。后处理重复数据删除可能需要您拥有足够的容量来暂存传入数据。这些传入数据未进行重复数据删除。但这些数据稍后会被重复数据删除,而重复数据删除发生后,您将不再需要该空间。例如,假设您的备份作业每个周末都会备份 100 TB 的数据并将其发送到执行端口处理重复数据删除的重复数据删除设备。因此,首先,您需要在重复数据删除设备上至少有 100 TB 的可用空间来存储传入的备份。备份将以未进行重复数据删除的形式保存在那里,直到稍后,当重复数据删除设备将数据删除时,它将被重新存储。

设备运行重复数据删除作业，并将数据集减少到 30 TB。因此，首先，您需要至少 100 TB 的空间，但删除重复数据后，您只需要 30 TB。
这非常浪费。因此，内联重复数据删除是首选方法。



Note 后处理重复数据删除技术设计为工作，您可以安排在忙时安静的时间执行重复数据删除操作 basis, where you schedule a quiet time for the deduplication operation to occur. 一个常见的例子是 NetApp FAS 技术（主存储阵列而非备份目标）上的后处理重复数据删除；将重复数据删除作业安排在凌晨是常见的做法。但是，其他后处理重复数据删除技术会更早地启动重复数据删除过程，这意味着重复数据删除与数据被提取同时发生。在这种情况下，您不需要大型着陆垫，但提取率会受到重复数据删除处理的阻碍。

配额和归档

在文件和电子邮件领域，配额可用于控制和强制限制用户和部门使用的存储空间量。事实上，如果您不限制人们可以访问的存储空间量，他们会不断要求更多，直到没有剩余空间。但是，请注意，限制用户用于存储其工作的存储空间量通常会鼓励他们养成不良习惯，例如将文档存储在公司网络之外的未经授权的位置（例如公共云和 USB 记忆棒）。所有这些都会使数据脱离公司治理的控制，包括备份、归档、防病毒和安全。因此，在考虑使用配额时要小心这一点。

通常会对电子邮件实施不同级别的配额。例如，三层电子邮件配额方法很流行，可能如下所示：

配额级别 1：当用户的邮箱使用量达到此配额级别时，用户将收到每日电子邮件，通知他们已达到配额限制。

配额级别 2：当达到此配额级别时，系统将停止向用户发送电子邮件，并且用户将收到一封通知电子邮件。

配额级别 3：当达到此配额级别时，系统将停止为用户接收电子邮件。

这种三层配额管理方法为用户提供了充足的机会在系统停止为他们发送和接收邮件之前整理他们的邮箱。

然而，许多组织认为强迫员工
删除电子邮件。每当用户删除错误的电子邮件时，IT 也会面临相关负担；通常 IT 必须恢复它。此外，高级

管理层和公司高管几乎肯定会免受此类配额的限制,显然需要一个更好的解决方案。

这时电子邮件归档等功能就派上用场了。例如,电子邮件归档可以配置为悄悄地将超过一定时间的电子邮件移至档案中。

此存档可以放在更便宜的存储中,并且通常不在正常的公司备份计划中,因此它是存储旧电子邮件的理想场所。而至关重要的是,就用户而言,电子邮件仍在他们的收件箱中;它可能只是图标略有不同或需要多花几秒钟才能打开。除此之外,就他们而言,邮件仍在他们的收件箱中。

这类解决方案可以免除手动删除或手动存档旧电子邮件的需要,并且让用户觉得他们的收件箱比实际更大。它还可以通过将存档的电子邮件和文档放在更便宜的存储上来节省 IT 成本,而不会给备份基础设施带来负担。每个人都是赢家。

同样的原则也适用于文件共享和 SharePoint。如果不执行配额,这些文件共享和 SharePoint 网站将成为垃圾场,增长不受控制。当配额失效时(也许高级管理层不喜欢配额),归档技术和信息生命周期管理(ILM)实践就会发挥作用。可以定义以下策略:

- 过去 x 个月未修改的文件可以移至较低的层。
- 过去 y 个月未访问的文件可以移至较低的层级。
- 特定部门所拥有的文件放置在合规性档案中。
- 某个文件系统中的文件被压缩。
- 某种类型的文件会保存一段指定的时间。

此类策略往往遵循以下基本原则:新数据和频繁访问的数据应放在高性能、高成本的存储中,而访问频率较低的数据应放在较低层级、较便宜的存储中。然而,ILM 比这更复杂,它试图根据数据的业务价值对其进行分类,将其放在适当的基础设施上,并在信息的整个生命周期内对其应用适当的策略。

从创建到删除。你可以想象,这是极其困难的。



重要的是要明白,信息生命周期管理(ILM)不是硬件或软件产品,相反,它是管理不断发展的政策和程序的策略。它也是 SharePoint 流行词。ILM 是目前还没有人接近的 strategy, that incorporate

这听起来是不是有点不着边际?是的!在现实世界中,ILM 是一项挑战,大多数组织的 ILM 策略都不成熟。说实话,在大多数组织中,ILM 的优先级最低,甚至根本不算优先级!但是,合规性归档是一些 ILM 原则得到良好遵守的领域。合规性归档将某些数据集和数据类型放入档案中,以防止篡改。

并在法律规定的严格期限内删除。然而,合规归档只是整个 ILM 蓝图的一小部分。

Showback 和 Chargeback

Showback 和 Chargeback 是许多公司的常见做法。它们分别与突出显示成本和计费成本有关。Showback 识别并突出显示某个客户或业务部门正在使用的 IT 服务的成本,但实际上并不向他们收费。它只是显示成本。Chargeback 更进一步,包括账单!

两种方式各有利弊。显然,费用返还包含账单这一事实对行为的影响更大。例如,如果部门知道他们必须为此付费,他们就不太可能要求超过其所需的存储空间。但不利的一面是,费用返还可能会鼓励一些不良行为和做法。例如,为了避免成本,部门可能会要求性能低于其所需的存储(更便宜)。或者他们可能会要求非复制存储,因为它的成本低于复制存储,而实际上他们需要复制。他们还可能开始将数据存储在授权的公司 IT 系统之外,例如公共云等地方。所以这是一个平衡行为。

另一方面,Showback 并不具有太大的威慑力,因为不需要涉及金钱交易。然而,如果高层管理人员审查了反馈报告,则可能会导致高层管理人员向各部门施加压力,要求他们证明或限制其对存储的使用。



Real World Scenario

内部退款导致不良做法的示例

我曾经工作过的一家公司有一个传奇故事,讲的是一个开发系统,为了避免巨额的内部退款账单,开发系统以低成本构建。毕竟,该系统只是一个开发系统,不需要高性能、防弹的存储。因此,开发系统建立在未复制的低层存储上。然后,就像开发系统经常出现的情况一样,一夜之间,在没有人特别注意的情况下,它成为了一个重要的实时生产系统!业务部门依靠它来进行核心交易。

这家公司的政策是复制生产系统上的所有卷。然而,该系统并非如此。相反,部分卷被复制到 DR 站点的第二个系统,但这个副本是一个简单的 rsync 副本,rsync 作业的目标是开发人员有空闲容量的随机服务器。这绝对不是您希望构建生产交易系统的方式。

幸运的是,这次事件没有发生重大事故,但随着越来越多的业务部门开始依赖这个系统,它的性能很快就开始下降。经过调查,发现该系统位于性能最低的存储上,并且没有被复制。

如果使用得当,showback 和 chargeback 会很有用。在消耗昂贵的存储资源时,这两者都可以鼓励更负责任的行为。如果部门知道必须为存储需求付费,它就更有可能认真考虑存储需求。这两种报告都可以显示投入 IT 的资金实际流向的书面记录;这两种报告都显示了谁在使用什么。因此,如果要求 IT 部门说明其资金都花在了哪里,这些报告可以显示业务部门 x 正在以 z 的成本消耗 y 存储。

绩效管理

性能管理是为了确保整个存储资产的最佳性能。它包括阵列、网络、NIC 和 HBA,甚至基于主机的组件（如卷管理器和多路径软件）。它应该涵盖整个存储资产。



值得注意的是应用程序的性能完全有可能非常糟糕,以至于被认为已停止运行。想象一下,如果一个数据库服务器的速度非常慢,以至于来自 web 服务器的查询超时。从技术上讲,数据库可能仍处于运行状态,但对于试图通过网站购买商品的客户而言,网站已停止运行,他们无法购买。因此,性能至关重要。

尽管需要监控的事物似乎无穷无尽,但基准测试是一个至关重要的概念。

基线

在解决疑似性能问题时,有一个参考框架、一个可以比较的东西至关重要。这就是基准测试发挥作用的地方。



一些性能管理工具能够保存历史性能数据的数据仓库在出现性能问题时可以调用该数据库并进行比较。即便如此,我们仍然强烈建议在性能管理工具数据仓库中保留额外的基线数据,例如,以 Excel 电子表格等文档格式保存。主要原因是性能管理工具可能会丢失历史数据。众所周知,性能管理软件升级等事件会丢失数据,或者对后端数据库架构进行重大更新,导致旧数据无法再用于新版本的工具(需要使用旧版本的单独实例,以防您需要读取历史性能数据)。这并不理想。性能管理工具绘制历史数据图表的速度也可能非常慢,以至于在性能问题最严重的时候几乎毫无用处。出于这些原因,我们强烈建议保留基线数据的单独副本。

没有什么比陷入性能问题、查看实时性能统计数据并自问这是好还是坏更糟糕的了。如果您只能根据当前情况判断，就会发生这种情况。您怎么知道 15,000 IOPS 的响应时间约为 15 毫秒是好还是坏？对于某些应用程序来说，这可能是好的，但对于其他应用程序来说可能不是。拥有漂亮的图表和图形固然很好，但当有人问您它们显示的数据是好还是坏时，您却无法自信地回答，您看起来会有点傻。您需要的是没有明显性能问题时的相同统计数据的副本。这样，您就可以比较两者并查看是否存在差异。

例如，假设您当前的统计数据显示，对于据称性能不佳的卷，其 IOPS 为 12,000，响应时间约为 25 毫秒。如果您可以将统计数据与一周前或一个月前的同一天进行比较，并且那时的统计数据也约为 12,000 IOPS，响应时间约为 25 毫秒，那么您就有一组事实表明问题与存储无关。但是，如果一周或一个半月前同一卷的统计数据显示 IOPS 为 40,000，平均响应时间约为 5 毫秒，或者可能是 IOPS 为 8,000，响应时间为 5 毫秒，那么您就知道存储可能出现了问题。

如您所见，性能基准测试可以成为救星。定期（可能每季度一次）重新基准测试存储环境并存储数据以供将来参考是一种很好的做法。如果您不定期进行基准测试，那么至少需要在对环境进行任何重大更改时进行基准测试 - 重大更改包括硬件和软件升级以及系统迁移等。事实上，当对环境进行重大更改时，您应该在前后进行基准测试，然后进行比较。您最不希望看到的是应用程序所有者说自从您升级环境以来，他只获得了正常性能的一半。如果您没有统计数据来证明事实并非如此，那么您的处境就很困难了。



Baseline data before major changes is considered a best practice in storage management. As part of a good storage management plan, you should review current best practices. Many BP, especially those specific to a technology, are pushed by your vendor. However, some BP apply to any technology. For example, backing up existing configurations before a major change is a good practice. Other best practices may include rules such as not deleting LUNs during work hours or changing transactional systems during transaction times. In the end, best practices aim to make your life easier and keep your system running normally. Therefore, maintaining and regularly reviewing a list of best practices is a good idea.

延迟/响应时间

存储环境中最大的性能杀手可能是延迟（有时称为响应时间）。响应时间是发送命令和获得响应所需的时间。对响应时间造成的任何延迟都称为延迟。

高延迟会导致用户愤怒,如果延迟足够高,应用程序可能会变得甚至低于被认为宕机的程度。可以公平地说,在存储环境中,唯一比高延迟更糟糕的事情是数据丢失 (DL) 和数据不可用 (DU)。因此,请认真对待延迟。

延迟几乎可能发生在设计的每个层上。以下是常见的地方:
尽管该列表并不全面:

主机:文件系统、卷管理器、主机总线适配器 (HBA)

网络:SAN 或 IP 网络、交换机间链路

存储阵列:阵列前端端口、缓存、后端、复制

一个常见的误解是,高 IOPS 意味着低延迟。毕竟,IOPS 是一个复杂的名词表示每秒的输入/输出操作数,而提到时间会让人认为 IOPS 和延迟是相关的。在某些方面它们是相关的,但必须了解这两个指标并不是线性相关的。也就是说,如果一个系统可以在 1 毫秒的延迟下执行 10,000 IOPS,也可以在 2 毫秒的延迟下执行 20,000 IOPS,这绝不是说它可以在 10 毫秒的延迟下执行 100,000 IOPS,也可以在 20 毫秒的延迟下执行 200,000 IOPS。事实上,它几乎肯定无法做到后者。IOPS 和延迟并不是线性相关的!以下内容可能有助于澄清:如果您的存储阵列一次只能处理一个请求,那么 1,000 IOPS 就意味着 1 毫秒的延迟。

然而,在现实世界中,存储阵列可以并行处理多个请求,这意味着 IOPS 代表整个系统的总体性能,而延迟则告诉您阵列处理任何单个请求的速度。

查看仅列出 IOPS 的阵列规格毫无意义。如果没有至少与之相关的延迟数字,250,000 IOPS 的数字就毫无意义。10 毫秒延迟的 250,000 IOPS 与 50 毫秒延迟的 250,000 IOPS 是完全不同的应用程序和用户体验。一个有用(但绝不完美)的例子是每天可以处理 100 个订单的订单处理部门。但是,每天处理 100 个订单以及当天通过快递发送货物,与每天处理 100 个订单但除周一外无人在邮局,然后通过最便宜、最慢的邮政服务发送货物的客户体验完全不同。两者都可以每天处理 100 个订单,但其中一个显然比另一个更好。

在基于光纤通道的 SAN 环境中,造成延迟的最大且最常见的因素通常是旋转磁盘驱动器。这是由于在由硅和电子元件主导的世界中,它的机械性质所致。另一方面,FC 交换造成的延迟非常小。这是因为 FC 交换机通常执行直通交换;交换机帧几乎在整个帧进入交换机之前就开始离开交换机(出口)。它甚至在到达之前就已经出去了!直通交换比存储转发交换更快,在存储转发交换中,帧必须在交换机上完全缓冲,然后才能传递到旅程中的下一跳。因此,FC 交换通常造成的延迟非常小——通常为几微秒,尽管一些供应商文档将其列为纳秒。无论哪种方式,与以毫秒为单位的旋转磁盘造成的延迟相比,它都是微不足道的。例如,3 TB Seagate Barracuda 硬盘的规格表编号列出了读写延迟,如下所示:

■ 平均读取性能<8.5毫秒

■ 平均写入性能<9.5 毫秒

而且你可以相信这些数字都是尽可能最好的情况。

显然,固态介质 (如闪存)可以用来替代大量的自旋磁盘并减少旋转磁盘常见的机械延迟。

在通过以太网 IP 网络 (而非 FC 网络)运行的 NAS 环境中,网络延迟可能是一个更重要的因素。这不仅是因为以太网交换机通常采用较慢的存储转发交换方法。还因为 IP 和以太网网络的丢包率要高得多,需要重新传输数据包。总而言之,NAS 和有时 iSCSI 环境中的网络延迟是整体端到端延迟的一个更重要的因素。

其他一些可能会增加存储环境中延迟的因素包括:

- 随机输入/输出
- VMware

VMware 以及其他虚拟机管理程序以创造业界所谓的 I/O 混合器效应而闻名。当来自单个物理服务器上每个虚拟机的所有离散 I/O 流都被虚拟机管理程序混合在一起并作为真正随机的 I/O 进入网络时,就会发生这种情况。这种 I/O 很难快速服务,尤其是对于传统的旋转磁盘,并且会导致延迟增加。其中一个主要原因是它搞砸了阵列识别数据模式和将数据预取到缓存的尝试。这有点像一百个人同时说话。尽管他们可能都在进行完全合理的对话,但由于人数众多,很难分离出个人的声音并理解他们在说什么。幸运的是,固态介质可以在那里提供帮助。

需要极低的延迟

现实生活中,超低延迟是至关重要的,其中包括以下用例:

- 低延迟交易系统
- 必须在处理之前执行欺诈检测的在线购物系统
付款
- 触摸式交通卡系统,例如伦敦牡蛎卡,要求乘客在面板上触摸卡才能通过旋转门进入地铁,并且必须在乘客通过栅栏之前验证卡余额
- 预付费手机系统要求在付款前检查账户余额
接通电话

每秒输入/输出次数

IOPS 是另一个众所周知且流行的存储性能指标,通常用于测量随机工作负载。它是每秒输入/输出操作的缩写,用于衡量系统在一秒钟内正在执行或可以执行的工作量。

然而,IOPS 的概念模糊且被滥用,导致其本身毫无用处。

I/O 到底是什么?所有 I/O 都一样吗?并非所有 I/O 都一样。有些很大;有些很小。有些是读取的;有些是写入的。有些是随机的,有些是顺序的。大多数现实世界的工作负载都包含所有这些的混合:小、大、读取、写入、随机和顺序。

IOPS 数字如果没有相关的延迟数字,那绝对是无用的。如果执行 500,000 IOPS 时的延迟超过 50 毫秒,谁会关心阵列是否可以执行 500,000 IOPS?能够以低于 10 毫秒的延迟执行 500,000 IOPS 的阵列肯定比以 50 毫秒的延迟执行 500,000 IOPS 的阵列更受欢迎。

很多时候,当供应商向您报出他们耀眼的 IOPS 数字时,他们几乎永远不会告诉您有关 I/O 类型或相关延迟的任何信息。事实上,一些供应商会引用荒谬的 IOPS 数字(高达数百万),当被问及这些数字是如何实现的时,他们不得不承认这些数字是微小的读取 I/O,由阵列前端端口上的缓冲区响应。这是一个令人难以置信的不切实际和荒谬的情况。不仅 I/O 不是由后端磁盘提供服务,甚至也不是由 DRAM 缓存提供服务。所有 I/O 都是由阵列端口本身的本地缓冲区提供服务。好像这在现实世界中会发生一样!公平地说,供应商的高级技术人员经常反对在营销材料上公布这些数字,但被营销部门否决了。

总之,如果您打算购买新阵列或调查潜在的性能问题,您需要了解有关 IOPS 的详细信息,例如大小和读写比率,以及与这些 I/O 相关的延迟。如果您只有 IOPS 数字,那就毫无意义了。

MBps 和传输速率

MBps 是另一个常用的存储性能指标。它是每秒兆字节数的缩写,指的是磁盘驱动器或存储阵列在一秒钟内可以传输的兆字节数。它通常用于测量或表达顺序(非随机)工作负载(例如备份和恢复作业)以及流媒体等的性能。



还需要注意的是,MBps 指的是兆字节每秒,而不是兆比特每秒。Mbps 或 Mb/sec 指的是兆比特每秒,not megabits per second. Mbps, or Mb/sec, refers to megabits per second.

从技术上讲,传输速率以 MBps 为单位,这两个术语通常在指磁盘驱动器和固态驱动器的性能时,这两个术语可互换使用。但是,在指存储阵列或整个解决方案(端到端)的吞吐量时,我们往往只使用术语 MBps。

值得注意的是,吞吐量驱动的应用程序(如已经讨论过的备份和流媒体示例)并不像主要由随机驱动的工作负载(如事务数据库)那样担心延迟。

影响存储性能的因素

许多因素都会影响存储性能。有些因素会反复出现。让我们来看看一些比较常见的因素。

袭击

尽管 RAID 可以成为保护数据免遭丢失和不可用的最佳帮手,但它也可能损害我们的性能。它一方面给予我们什么,另一方面又常常夺走我们的什么。

从性能角度来看,通常最令人担忧的 RAID 级别是 RAID 5 和 RAID 6,而数据库管理员(DBA)要求使用非 RAID 5 或 RAID 6 的 SAN 卷的情况并不少见。

关于 RAID 的争论

2003 年,两位数据库管理员 James Morle 和 Mogens Nørgaard 发起了一项名为“反对任何 RAID 5”(BAARF, www.baarf.com)的倡议,以发泄他们对 RAID 5 对数据库性能影响的不满。从他们开明的数据库专家的角度来看,他们谴责任何供应商实施 RAID 5,以及 RAID 4 和 RAID 3(由于 RAID 3 不符合他们的首字母缩略词,他们巧妙地将其称为 RAID free,以便 BAARF 首字母缩略词有效)。

他们对 RAID 5 的态度可以在他们的网站上看到,其中包括以下引言和声明:

够了就是够了……詹姆斯·莫尔(James Morle)和其他人写过书,讨论了 RAID-F [免费,四,五] 的无用性……
我们决定降低我们的血压……永远拒绝再对此进行任何争论。

尽管有些极端和狭隘,但许多 DBA 对 RAID 5 和 RAID 6 技术都有类似的看法。(在这两大技术界支柱发布其权威法令时,RAID 6 还没有流行起来。)

然而,总而言之,RAID 1 和 RAID 10 技术的实施成本往往比 RAID 5 和 RAID 6 要高得多。因此,要求使用 RAID 1 和 RAID 10 是可以的,只要额外的成本是合理的。

基于奇偶校验的 RAID 方案（例如 RAID 5 和 RAID 6）的性能与其他 RAID 方案（例如 RAID 1 和 RAID 10）不同。这是由于一种称为写入惩罚的现象造成的。这可能会导致性能降低，尤其是在工作负载包含大量随机写入活动的情况下（数据库工作负载通常就是这种情况）。写入惩罚发生的原因是小块写入需要大量奇偶校验重新计算，从而导致后端产生额外的 I/O。我们将在第 4 章“RAID：保护您的数据安全”中更详细地讨论 RAID 和 RAID 性能，但既然我们在这里，一个简单的类比可能会有所帮助。

可以将数据写入 RAID 5 或 RAID 6 卷视为类似于在 Microsoft Word 中编辑文档。想象一下，在 Word 文档中，每个句子都相当于一个 RAID 条带。如果我们需要用新句子替换整个句子，这很容易：我们只需突出显示整个句子，然后将新句子粘贴在其顶部，同时删除旧的突出显示句子。这就像大块写入，我们将整个新 RAID 条带写入旧条带，简单又方便。现在假设我们需要对现有句子进行一些修改。也许我们需要删除第三个单词，在第六个单词后添加逗号，交换第七个和第八个单词的位置，在第九个单词和第十个单词之间添加破折号，最后纠正第十个单词中的拼写错误。这比仅仅将新句子粘贴到旧句子上要费力得多。更新现有 RAID 条带也是如此。对于 RAID 5 和 RAID 6 来说，小块写入相对困难（就像我们在句子内进行更改一样），因为它们需要在 RAID 条带内进行更改，这会迫使系统读取条带的其他成员才能重新计算奇偶校验。此外，随机小块写入工作负载需要磁盘上的 R/W 磁头在整个盘片表面上移动，从而导致较长的寻道时间。最终结果是，RAID 5 和 RAID 6 中的大量小块随机写入可能会很慢。即便如此，诸如写入时重定向或随处写入文件系统和大型缓存之类的技术可以在很大程度上掩盖和减轻这种损失。

虽然在理想情况下，我们可能对所有东西都使用 RAID 10（镜像加条带化），但事情并不总是那么简单。主要问题是 RAID 1 和 RAID 10 消耗大量容量，因为它们基于镜像。

这些额外所需空间会推高成本。只要您有数据中心空间和预算中的现金，您就可以拥有任意数量的 RAID 1 或 RAID 10。

另一方面，如果您需要控制成本，RAID 5 和 RAID 6 在某些用例中可能是更好的选择。

缓存

缓存是一种神奇的成分，它几乎使基于旋转磁盘的存储阵列能够跟上数据中心其他技术的速度。如果将 DRAM 缓存和缓存算法从图中移除，基于旋转磁盘的存储阵列实际上会陷入停滞状态。它们的速度会慢得令人难以忍受，世界将会变得不同。如果没有缓存，我们今天所做的许多事情根本无法完成，或者至少我们无法像今天这样出色地完成它们。想象一下，如果没有 80 号州际公路 (I-80)，从芝加哥开车到加利福尼亚就像是一条基于磁盘的

存储阵列没有缓存。添加 I-80 就像添加缓存（使旅程更快）。扩展一下类比，全灰阵列就像乘坐飞机。

为了加快平均响应时间，系统中有足够的缓存非常重要。如果读取或写入 I/O 可以从缓存中得到满足（不必依赖磁盘来完成读取或写入 I/O），那么它将比依赖后端磁盘快得多。但是，并非所有工作负载都能从旋转磁盘前的缓存中受益。有些工作负载会导致高缓存命中率，而其他工作负载则不会。当 I/O 可以从缓存中得到服务时，就会发生缓存命中，而缓存未命中则需要访问后端磁盘。即使在缓慢旋转的磁盘前有一个大缓存，也有一些 I/O 会导致缓存未命中并需要使用后端磁盘。

这些缓存未命中 I/O 导致的响应时间比缓存命中慢得多，这意味着差异（最快和最慢响应时间之间的差异）可能很大，例如从大约 2 毫秒一直到大约 100 毫秒。这与全灰阵列形成了鲜明的对比，全灰阵列的差异通常非常小。

大多数供应商都有标准的磁盘容量与缓存容量比率，这意味着您不必太担心在系统中放入多少缓存。但是，这些供应商方法是一刀切的方法，可能需要根据您的具体要求进行调整。在执行此操作时，咨询您的供应商通常是一个好主意。

精简 LUN

精简 LUN（基于精简配置）采用按需为 LUN 和卷分配空间的概念。因此，在创建 LUN 的第一天，它并没有分配物理空间。只有当用户和应用程序写入时，才会分配容量。这种按需分配模型可以在两个方面产生影响：

- 按需分配过程可能会增加延迟。
- 按需分配过程可能会导致后端布局碎片化。

按需分配过程理论上可以为写入过程增加一点延迟。因为每次写入请求进入精简 LUN 上的新区域时，系统都必须识别空闲区并将其分配给卷。但是，大多数解决方案都经过了优化，以尽量减少这种影响。

可能更令人担忧的是，由于分配给某些精简 LUN 的空间具有伪随机性，它们最终可能会出现严重碎片化的后端布局。这在具有大量连续工作负载的应用程序中尤其明显。如果用户怀疑由于使用精简 LUN 而导致性能问题，请对精简 LUN 和密集 LUN 执行代表性测试并比较结果。

网络跳数

在网络、FC SAN 或 IP 中，流量必须经过的交换机数量会影响响应时间。跨越更多交换机和路由器会增加延迟，通常称为网络引起的延迟。这种延迟通常在使用存储转发交换技术的 IP/以太网网络中更高，此外流量穿越路由器的可能性也更高。

多路径

许多人认为多路径 I/O (MPIO) 解决方案就是为了实现高可用性;如果一条路径发生故障,另一条路径会接管,而应用程序或用户甚至不会注意到。这是真的。但是,MPIO 也会对性能产生重大影响。

例如,通过两个 HBA 和 HBA 端口平衡主机的所有 I/O 可以提供比通过单个端口发送所有 I/O 更多的带宽。它还可以使两个 HBA 的队列和 CPU 处理能力可用。MPIO 还可用于平衡存储阵列上多个端口的 I/O。MPIO 可用于平衡单个主机在多个阵列端口(例如八个端口)上的 I/O,而不是将所有主机 I/O 发送到存储阵列上的两个端口。这可以极大地帮助避免阵列前端端口上的热点,类似于宽条带化避免阵列后端热点的方式。

标准性能工具

现在我们来看一下业界使用的一些常见性能工具。

性能监视器

Perfmon 是一款 Windows 工具,可让您监控种类繁多的基于主机的性能计数器。从存储角度来看,这些计数器非常有用,因为它们为您提供了从主机角度看到的画面。例如,从主机经历的延迟将是端到端延迟,这意味着它将包括基于主机、基于网络和基于阵列的延迟。但是,它只会给您一个数字,并且不会将总体延迟分解为主机引起的延迟、网络引起的延迟和阵列引起的延迟。

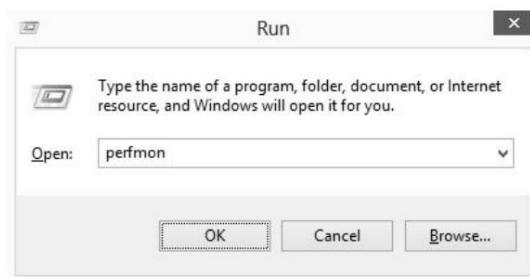
练习 12.1 使用 Windows perfmon 测量 I/O 延迟。

练习 12.1

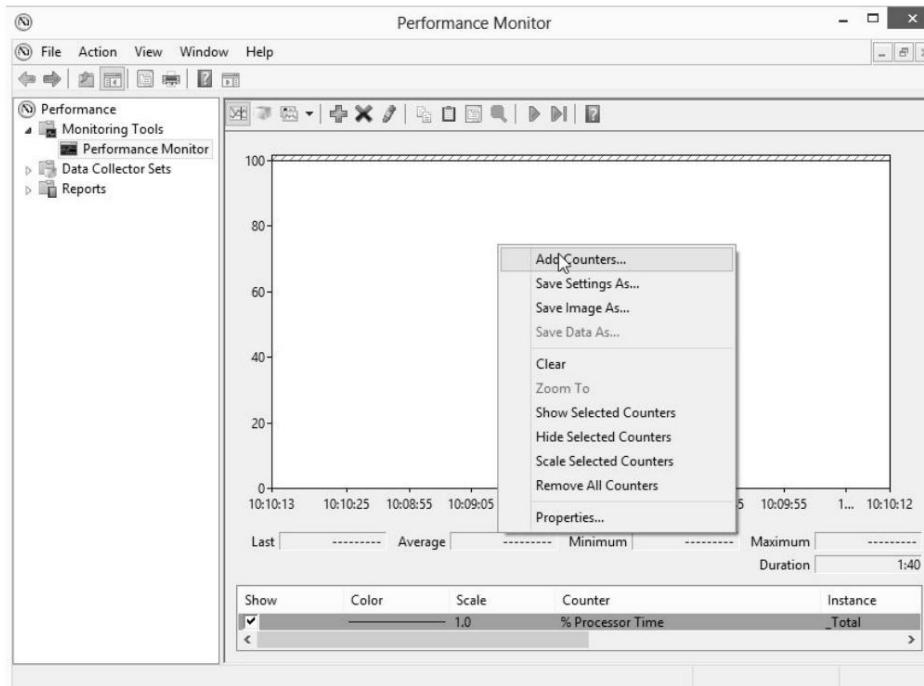
使用 Perfmon

在此示例中,您将使用 Windows perfmon 工具来监视系统上 SAN 卷 (D:) 的平均读写延迟。

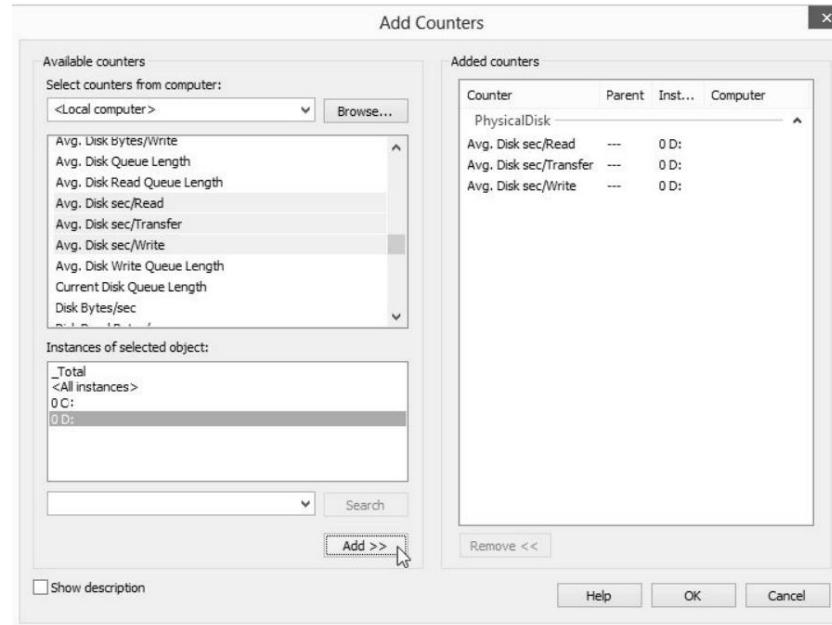
1. 在命令提示符下输入 perfmon, 打开 Windows perfmon 实用程序
运行对话框。



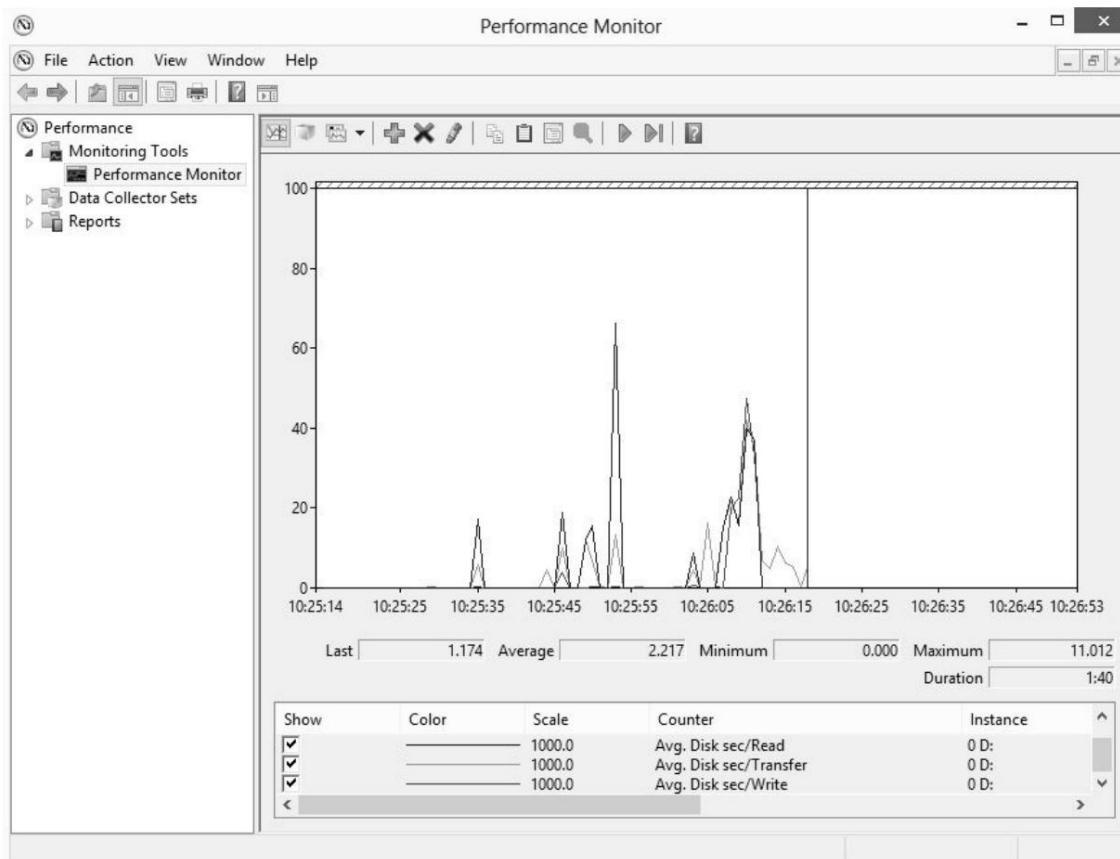
2.在 perfmon 窗口中,右键单击右侧的空白图形区域,然后单击添加计数器。



3.从“添加计数器”屏幕中,选择以下计数器以添加到选择中:平均磁盘秒/读取、平均磁盘秒/写入和平均磁盘秒/传输。对于每个计数器,选择您的 SAN 卷,卷 D:。



现在已经添加了计数器,perfmon 将显示平均读取、平均写入和平均组合读/写延迟的实时延迟数字。



Perfmon 还可用于显示更多与性能相关的计数器。甚至一些非存储相关计数器的统计有助于排除疑似存储相关问题。举一个非常简单的例子,如果服务器上的 CPU 繁忙计数器非常高,则服务器整体性能将受到负面影响,这显然会对服务器的存储性能产生影响。



Perfmon (perfmon.exe)取代了在某些非常旧的Windows版本上使用的旧命令 (sysmon.sys)。

used on some very old versions of Windows.

iostat

iostat 是 Linux 世界中用于监控存储性能的常用工具。练习 12.2 介绍了 iostat 提供的输出类型。

练习 12.2

使用 iostat

在安装了 iostat 的 Linux 服务器上的命令行中,键入以下命令,每 10 秒显示一次该计算机的 I/O 统计信息 (返回的数字是过去 10 秒的平均值) :

```
iostat -x 10
```

以下输出显示过去 10 秒内的 I/O 统计信息:

```
nigelpoulton@ubuntu-02:~$ iostat -x 10
Linux 3.8.0-19-通用 (ubuntu-02)           21/09/13      _x86_64_
(4 个 CPU)
```

平均 CPU	% 用户	% nice	% 系统	% iowait	% steal	% idle
0.18	0.00	0.13	1.41	0.00	98.28	

设备:rrqm/s wrqm/s avgqu-sz await r_await w_await	速度/秒	宽/宽	響B/秒	wkB/s 平均-sz
svctm %util				

星展银行	12.49	4.27 8.77	2.19 298.69	99.02	72.60
1.39 126.62					
135.39 91.54 4.46 4.89					
dm-0	0.00	0.00 11.54	6.44 261.18	99.00	40.05
2.10 116.80					
147.49 61.81 2.57 4.62					
dm-1	0.00	0.00 0.36	0.00 1.42	0.00	8.00
0.00 7.59					
7.59 0.00 5.37 0.19					

虽然本书中命令的输出格式可能不太好,但从 Linux 服务器上的命令行来看,它很容易阅读并提供良好的统计数据。

警报

即使我们当中最优秀的人,偶尔也会分心。当这种情况发生时,警报可以帮到我们。

从高层次上讲,警报是为了通知我们重要事件。这些事件可能是故障、从故障中恢复、超过容量阈值、低于性能阈值或其他重要事件。从高层次上讲,警报基于阈值和条件的原则。您可以配置某些阈值,并将其与超出该阈值时要采取的措施相关联。以下是一些示例:

- 如果可用空间 $\leq 80\%$, 则发送电子邮件警报。
- 如果任何物理组件出现故障,则发送电子邮件警报。
- 如果网络延迟 ≥ 10 毫秒, 则向网络操作中心发送 SNMP 陷阱。

常见的警报形式包括:

- 电子邮件
- 短信/文本信息
- SNMP 陷阱

在某些存储技术中,系统在发生特定事件时进行回呼也是很常见的。当系统向供应商发送包含有关情况的信息的消息时,就会发生回呼。示例包括:

- 可用空间量已降至阈值以下。
- 组件发生故障。

过去,呼叫回家需要将调制解调器等设备连接到阵列和交换机的背面。但现在,呼叫回家通常是通过互联网使用 HTTPS 实现的。值得注意的是,只有与系统相关的信息才会被发送回家;用户数据永远不会被发送回供应商。

根据您的技术,可以针对各种条件配置警报
在容量方面,一些更有用的条件包括:

- 可用空间
- 总使用空间

警报的主要用途是让相关人员(几乎总是包括你)意识到需要解决的某些情况。知识就是力量!没有什么比管理层比您先发现存储环境中的情况更令人尴尬的了。如果您配置了良好的警报,这种情况发生的可能性就会小得多。

更现代化的系统可以回访并报告各种分析数据,甚至可以识别将受已知错误影响的系统并提醒用户。

练习 12.3 将引导您了解如何配置警报。

练习 12.3**配置警报**

在大多数存储系统上,配置警报应该很简单。以下步骤将引导您配置 Brocade 交换机上发生故障的组件的电子邮件警报。

为了简化流程,我们假设交换机已配置电子邮件服务器。您只需配置故障现场可更换单元 (FRU) 的电子邮件警报,这样只要 FRU 发生故障,就会向您发送电子邮件。

1.从交换机命令行键入fwmailcfg命令:

```
LegendarySw01:admin> fwmailcfg
```

2.从出现的文本菜单中,选择选项 5,设置收件人邮件地址

电子邮件提醒:

1:显示邮件配置信息

2:禁用电子邮件警报

3:启用电子邮件警报

4:发送测试邮件

5:设置电子邮件警报的收件人邮件地址 6:中继主机 IP 配置 7:退出选择一个

项目 => : (1..7) [7] 5

3.从邮件配置菜单中,选择选项 12,FRU 类别:

邮件配置菜单

1 :环境类

2:SFP 类

3:端口类

4:面料类

5:电子口岸类

6 :F/FL 端口 (光学)类

7:ALPA 性能监视器类

8:端到端性能监视器类

9:过滤性能监控类

10:安全等级

11:资源监视器类

12:FRU 类

13:退出

选择一个项目 => : (1..13) [13] 12

4.现在系统提示您输入要发送警报电子邮件的电子邮件地址。

输入您的电子邮件地址或电子邮件组地址（我在这里使用我自己的）：

邮寄至:[无] nigelpoulton@hotmail.com 电子邮件警报配置成功！

此消息表明已成功配置用于接收 FRU 相关警报的电子邮件地址。但是，您尚未配置要针对 FRU 状况发送的警报。

5.从主菜单中，选择选项 3 “启用电子邮件警报”：

1:显示邮件配置信息
 2:禁用电子邮件警报
 3:启用电子邮件警报
 4:发送测试邮件
 5:设置电子邮件警报的收件人邮件地址 6:中继主机 IP 配置 7:退出选择一个
 项目 => : (1..7) [7] 3

6.从“邮件启用”菜单中，选择选项 12,FRU 类：

邮件启用菜单

-
- 1 :环境类
 - 2:SFP 类
 - 3:端口类
 - 4:面料类
 - 5:电子口岸类
 - 6:F/FL 端口（光学）类
 - 7:ALPA 性能监视器类
 - 8:端到端性能监视器类
 - 9:过滤性能监控类
 - 10:安全等级
 - 11:资源监视器类
 - 12:FRU 类
 - 13:退出

选择一个项目 => : (1..13) [13] 12

电子邮件警报已启用！

现已配置电子邮件警报，每当发生与现场可更换单元相关的事件时，都会向您的电子邮件地址发送电子邮件。



存储资源管理 应用

首先需要指出的是,本节中提到的 SRM 工具并不是指 VMware Site Recovery Manager。本节中我们讨论的是存储资源管理工具。

在大型存储环境中,如果存在来自多家供应商的多种技术,很容易导致管理工具泛滥。例如,想象一下以下中型环境:

- 3 个 EMC VNX 阵列
- 2 个 Dell Compellent 阵列
- 4 台 Brocade FC 交换机

此环境通常需要至少通过以下管理工具进行管理:

- EMC VNX GUI 或 CLI
- Dell Compellent GUI 或 CLI
- Brocade GUI 或 CLI

在某些情况下,您可能需要这些管理工具的多个实例。您可能需要两个 Dell Compellent GUI 实例来管理两个 Dell Compellent 阵列。这在中大型环境中很容易失控。

为了避免这种管理工具泛滥,第三方公司和存储供应商开发了一些工具,试图通过单一工具来管理异构存储环境。我们称之为存储资源管理 (SRM) 工具。

在底层,SRM 应用程序通常使用存储管理计划规范 (SMI-S) 与异构设备进行通信和控制。因此,它们仅限于每个供应商的设备通过 SMI-S 公开的功能,而且并非所有供应商都很好地支持 SMI-S。当供应商不通过 SMI-S 公开功能时,SRM 工具必须恢复到特定于供应商的 API,从而缩小了 SRM 工具支持的系统范围,并可能推高 SRM 工具的成本。



在存储行业中,我们经常使用“异构”这个词来指代来自多家供应商的存储技术。如果您的存储资产包含来自多家供应商的阵列,那么您就拥有异构环境。如果您的所有存储设备都来自一家供应商,那么您就不是异构环境。

尽管 SRM 工具的初衷是好的,但现实是它们却是一场噩梦去实施和维护。

在现实世界中,SRM 工具往往具有以下要求和属性:

- 专用数据库后端
- 拥有大量 CPU 和 RAM 的强大服务器
- 落后于本地元素管理器
- 复杂的许可
- 需要一位或多位专家来管理它们 (使用 SRM 工具的专家)
- 表现不佳

因此,大多数公司最终部署了 SRM 工具,但仍然必须保留每个平台的原生元素管理器,而不是通过减少管理存储设备所需的工具数量来简化管理。而且他们为这项特权付出了很多钱。这背后的两个主要原因如下:

- 大多数存储管理员不信任 SRM 工具来进行其存储阵列和交换机的日常管理。

- SRM 工具通常比本地元素管理器落后至少 6 到 12 个月
当谈到支持新功能时。

通常,SRM 工具用于整个资产的报告,而本机元素管理器用于环境的管理和故障排除。归根结底,一个通用的工具永远不会像该技术的本机元素管理器那样好,也不会像该技术的本机元素管理器那样对技术有深入的了解。例如,EMC XtremIO 的本机元素管理器总是能够比异构第三方 SRM 工具更好地配置和报告 EMC XtremIO 阵列。

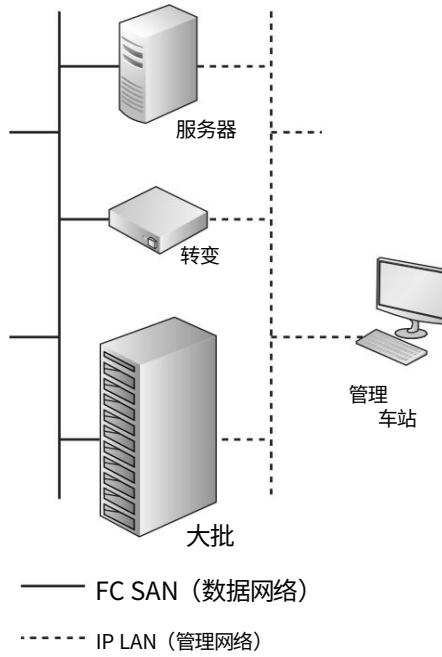
管理协议和接口

有多种方法可以管理您资产中的存储设备。其中大多数是基于网络 (IP) 的。在探索管理存储设备的最常见方法之前,让我们先考虑一下带外和带内这两个术语。

在设备管理方面,带外管理最为常见,需要专用接口来处理管理流量。服务器领域的一个常见示例是无人值守管理接口,例如 HP ProLiant 服务器中的 Integrated Lights-Out (iLO)。通常,这些专用管理接口有自己的专用处理器,即使操作系统关闭,有时甚至电源未打开,也可以管理系统。带外管理通常有自己的专用网络,以便管理流量不会干扰数据流量,并且在主生产数据网络发生故障时提供对设备的管理访问。FC SAN 领域中一个常见且简单的例子是通过 IP 网络而不是用于数据流量的 FC SAN 网络来管理 FC SAN 连接的存储设备。

图 12.9 显示了一个 FC 存储阵列、一个 FC 交换机和一个服务器,它们分别连接到 FC SAN 以传输数据流量,并且还连接到专用于管理流量的单独 IP 网络。

图 12.9 带外管理



另一方面，在带内管理中，管理流量和数据流量共享同一网络，并且可能共享同一网络接口。带内管理还可能要求系统和操作系统正常运行，以便能够管理系统。

FC SAN 环境中带内管理的一个示例是使用所谓的守卫设备或命令设备。这些设备是特殊的 LUN（卷），通过 FC SAN 从阵列呈现给管理主机。然后，管理主机可以向这些设备发送命令，阵列会发送对命令的响应。这实际上允许您通过 FC SAN 而不是 IP 网络来管理阵列。

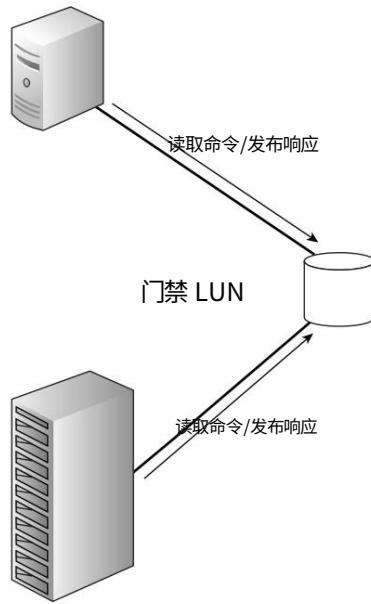
这种配置通常用于复制和快照任务的脚本，其中主机运行脚本来控制阵列上卷的快照和复制状态。

图 12.10 显示了 FC 存储阵列通过 SAN（带内）向主机呈现守卫设备的示例。



An easy way to conceptualize in-band and out-of-band management is to think of the network as a highway for data traffic and a toll road for management traffic. If you need to send management traffic, it's better to use the toll road (out-of-band) because it won't affect the flow of data traffic on the highway (in-band).

图 12.10 带内 FC 管理



命令行界面

命令行界面 (CLI) 一直是硬核管理的代名词,通常情况下,高级配置和高级管理任务只能通过 CLI 执行。因此,有时只有高级或专家用户才会接触 CLI。真男人不点击这个短语偶尔被用来暗示专家不要在 GUI 及其点击界面上浪费时间。

通过命令行进行资产管理时,主要有两个选项:

- 远程登录
- 安全外壳 (SSH)

它们都是基于 TCP/IP 的网络协议,用于提供命令行访问远程系统。但是,SSH 是当今最流行的选择,这是有原因的。Telnet 是一场安全噩梦。它不支持身份验证或加密,并且以明文形式通过网络发送密码!不要使用它。改用 SSH,如果可能的话,使用 SSH-2,因为它更安全。



Although Telnet and SSH work over the network, it is also possible to get a command-line interface to some systems by plugging in a serial cable—

有时也称为控制台电缆。为此,您通常需要随身携带笔记本电脑并站在系统前面。串行电缆 (控制台电缆)的一端插入笔记本电脑上的串行端口,另一端插入要管理的系统上的串行端口。但是,现在很少有笔记本电脑配备串行端口。在这些情况下,您可以使用 USB 转串行转换器/电缆,或连接具有串行端口的设备,然后通过网络访问该设备并远程使用其串行端口。这些控制台连接通常需要执行初始设置任务,例如配置管理 IP 地址等。

SSH 支持通过证书进行相互身份验证以及加密,因此比 Telnet 安全得多。它是一种基于客户端-服务器的协议,存储阵列和 SAN 交换机需要运行 SSH 服务器守护程序,以便您使用 SSH 客户端连接和管理它们。大多数存储阵列和 SAN 交换机都运行 SSH 服务器,因此可以通过 SSH 进行本地管理。

SSH 是基于 TCP 的协议,默认情况下通过 TCP 端口 22 运行。但是,可以通过管理方式将其更改为您希望的任何其他 TCP 端口。



从技术上讲,SSH 是一套由三个安全工具/实用程序组成的套件。Slogin、SSH 和 SCP 都是流行的开放式实用程序。Slogin (远程登录)、SSH (远程 shell) 和 SCP (远程复制) 的安全版本。SSH 中的每个版本都以字母“s”开头,表示它们是安全版本:安全登录 (slogin)、安全 shell (ssh) 和安全复制 (scp)。

CLI 远胜于图形界面的一个方面是任务自动化。存储管理员通常会安排通过 CLI 和脚本 (一个接一个执行的命令集合) 运行的任务。

图形用户界面

大多数存储设备都支持通过图形用户界面 (GUI) 进行管理,以执行基本的日常任务。一般来说,GUI 的功能不如命令行强大,通常是出于以下原因:

- 无法通过 GUI 自动执行任务。
- GUI 通常只提供对最常用的命令和功能的访问。

许多现代 GUI 都是基于 Web 的,允许您通过标准 Web 浏览器连接到它们。在这些情况下,您应该始终将它们配置为仅接受通过 HTTPS 而不是纯 HTTP 的连接。HTTPS 可以被认为是 HTTP 的安全版本,但实际上它只是在安全 SSL/TLS 协议之上运行的普通 HTTP。

只要您的存储设备支持 HTTPS,您就可以从任何受支持的 Web 浏览器通过 HTTPS 连接到它。

HTTPS 连接是通过在 http 部分的末尾添加一个 s 来发起的 URL,如下所示:

<http://technicaldeepdive.com>

<https://technicaldeepdive.com>

中山大学

如前所述,SMI-S 是存储管理计划规范的首字母缩写。这是一项基于 SNIA 的计划,旨在为管理异构存储设备提供基于 Web 的通用标准。您无需让每台存储设备都拥有自己的专有管理界面和管理工具,而是让它们都实现 SMI-S,并可以通过任何基于 SMI-S 的管理软件进行管理。潜在的最终目标是拥有一个单一的管理平台来管理大型异构存储环境(例如 SRM 应用程序)。如果真有涅槃的话,那将是一种极乐状态!

遗憾的是,这一理想从未实现。几乎没有存储管理员会将存储设备的日常管理工作委托给设备附带的供应商专用工具(原生元素管理器)。此外,供应商对 SMI-S 的支持也非常不完善,新功能通常不会添加到 SMI-S 中,或者至少远远落后。

但是,SMI-S 在异构存储环境的集中报告和监控方面非常有用。公司通常使用单一的企业报告和监控工具来管理整个存储资产,即使该资产拥有来自多个供应商的设备。



Real World Scenario

SMI-S 缺陷示例

一家公司是大型存储供应商之一的精简配置的早期实施者,当他们意识到他们用来报告资产的第三方 SRM 工具不支持或不理解精简配置卷时,他们发现自己处于一个尴尬的境地。他们所有的月度容量报告对他们实施精简配置的阵列毫无用处。尽管向 SRM 供应商施加了压力,但大约九个月后,供应商才发布了支持精简配置卷的产品版本。

需要说明的是,SMI-S 是作为客户端-服务器模型实现的。在此模型中,服务器是存储设备,例如阵列、交换机和 HBA。SMI-S 客户端是管理工具。

简单网络管理协议

简单网络管理协议 (SNMP) 是一种流行的基于 IP 的协议,用于管理和监控 IP 网络上的设备。不过,SNMP 更常用于监控网络连接设备的运行状况和状态,而不是管理和更改它们的配置。

SNMP 是基于客户端-服务器的:

- 客户端是受管理和监控的设备,例如阵列、交换机和 HBA。有时客户端也被称为节点或受管理设备。
- 服务器是监控和管理 SNMP 客户端的应用程序。例如,
当组件发生故障时,SNMP 服务器通常会从客户端接收警报(称为陷阱)。此外,在读/写 SNMP 配置中,服务器会将配置更改推送到 SNMP 客户端。有时我们将服务器称为网络管理系统(NMS)。

SNMP 是一种基于 IP 的应用层协议,通过 UDP 端口 161 和 162 进行通信,但通过 TLS 运行的安全版本在端口 10161 和 10162 上进行通信。因此,SNMP 客户端和服务器之间的所有通信都通过 IP 进行。



从安全角度来看,最新版本的 SNMP,第3版(SNMPv3)是最佳选择。使用 SNMP Version 3 (SNMPv3),所有通信(包括密码)都以明文形式发送,相比之下,SNMPv3 支持所有五个主要功能,网络安全功能:身份验证、完整性和隐私(加密)。实际上,SNMPv3 越来越受欢迎,支持 SNMP 的系统也越来越普遍。

特定支持 SNMP 的设备(客户端)上可以管理和监控的内容完全取决于设备本身,因此,某些设备提供广泛的 SNMP 功能,而其他设备则不提供。支持 SNMP 的设备维护一个管理信息库(MIB),该库定义了可以监控和管理哪些设置。

概括

在本章中,我们首先讨论了容量管理和容量报告的重要性,然后简要介绍了一些可能影响容量管理的主要技术。然后我们讨论了性能管理,强调了性能基准的重要性,然后讨论了存储行业使用的主要性能指标。然后我们讨论了可能影响容量管理的常见因素。

影响存储性能并提到了一些用于测量性能的工具。我们通过讨论管理协议（例如 SMI-S 和 SNMP）以及图形用户界面和命令行界面结束了本章。

章节概要

容量管理 首先，容量管理就是确保您的存储空间不会耗尽容量。一旦您掌握了这一点，接下来就是充分利用您拥有的容量。这包括确保需要快速访问的数据位于快速媒体上。报告也是容量管理的主要组成部分。了解谁在使用什么，以及增长预测是什么样的。

性能管理 性能管理是为了确保存储资产的端到端性能足够好。性能不佳会降低应用程序的速度，甚至使其变得毫无用处。性能管理的一个重要部分是基准测试，它使您能够将当前的性能与应有的性能进行比较。

警报 警报是一种确保您和其他人了解存储资产中重要情况和事件的方法。这些情况和事件包括组件故障、服务状态变化以及与容量和性能相关的事件。

章节 十三



更广泛的数据 中心社区

本章涵盖的主题：

- ✓ 数据中心的重要性
- ✓ 数据中心层
- ✓ 机架、机柜和排
- ✓ 电源和冷却
- ✓ 结构化布线
- ✓ 访问控制
- ✓ 防火



本章介绍了数据中心的核心作用和重要性,以及数据中心设计的高级原则。您将了解 Uptime Institute 定义的数据中心层。您还将了解数据中心的主要组件,包括架空地板、机架、电源系统、冷却系统和布线。本章详细介绍了结构化布线方案的主要组件,并概述了最佳实践,这些最佳实践将使数据中心的日常管理尽可能简单,同时仍使数据中心足够灵活,以满足未来的需求。最后,本章总结了一些在数据中心工作的良好实践。

数据中心设计

关于现代数据中心,首先要知道的是,它们对于任何重视数据的组织都至关重要。如果你重视数据,就重视数据中心!如果你不重视数据中心,就不重视数据!

数据中心不仅仅是其组成部分的总和。它本身就是一个系统。数据中心不再是用来隐藏 IT 设备的肮脏备用室。在现代世界中,数据中心不仅仅是一个技术室,而是一个技术引擎 一个集成了计算、存储、网络、电源和冷却的系统,具有自己的性能、弹性和可用性水平。因此,数据中心需要适当的设计,以及适当级别的管理和维护。同样,如果您关心您的数据,您就需要关心您的数据中心。

从许多方面来看,数据中心是计算、网络和存储设备的完美家园 一个从头开始构建的栖息地,以使它们都处于完美的工作状态。

数据中心不仅为IT设备提供物理安全,还提供冷却、加湿、

阳离子、接地和电源馈送都针对 IT 设备进行了高度优化。作为人类,您不会想在 3 级或 4 级数据中心呆太长时间。如果这样做,您的眼睛会酸痛、干涩,耳朵也可能发炎。您还可能会发现温度变化也不舒服,因为数据中心的某些区域太冷而不舒服,而其他区域又太热。数据中心关心的是保持 IT 设备正常工作,而不是人为。

考虑到所有这些,让我们快速看一下我们数据中心的层级
在谈论数据中心正常运行时间和可用性时通常提到。

数据中心层级

最广泛接受的数据中心层级定义来自 Uptime Institute (<http://uptimeinstitute.com>)。Uptime Institute 是 451 集团的一个独立部门,致力于数据中心研究、认证和教育。这些层级也被电信协会 (TIA) 借用和参考,使它们几乎成为事实上的标准。

将数据中心分为不同层级的原因是为了尝试标准化我们如何定义层级以及讨论数据中心的正常运行时间和可用性。



电信行业协会 (TIA) 于 2005 年推出了 10 项与数据中心布线相关的重要标准。这项 TIA-942 标准在标准化和提高数据中心布线水平方面发挥了重要作用。TIA-942 包含一个附录,其中引用了 Uptime Institute 的数据。TIA-942 级别模型 has been instrumental in standardizing and raising the level of cabling within the data center. TIA-942 includes an annex that references the Uptime Institute's data center tier model.

从高到低看,UpTime Institute 将数据中心分为四个层级:four tiers of a data center:

- 第⁴层 4
- 第³层 3
- 第²层 2
- 第¹层 1

如该列表所示,第 4 级为最高级别,第 1 级为最低级别。这意味着 tier 4 是 the highest tier, tier 1 is the lowest tier. This means that tier 4 data centers have the highest availability, reliability, and performance, whereas tier 1 data centers have the lowest.

一方面,四级数据中心将拥有冗余的 cooler 电源、冷却、网络连接,所有一切。另一方面,一级数据中心可能只不过是一您的车库只有 1 条供电电路、1 个移动空调机组和 1 个宽带连接。hand, a tier 1 data center could be little more than your garage with a single power circuit, a mobile air-conditioning unit, and a broadband connection.



并非所有数据中心都经过 UpTime Institute 认证,因此即使是数据中心,也可以按照 UpTime Institute How-to Guide 和层级相对应的标准进行构建和维护。要正式将数据中心称为 tier 3 或 tier 4,数据中心必须经过正式认证或 prep to formal UpTime Institute tiers. To formally refer to a data center as, for example, tier 3 or tier 4, the data center should be formally certified.

让我们来讨论一下 UpTime Institute 指定的数据中心层级的某些高级功能和要求。

Tier 1

第 1 层是最基本的(可用性最低)的数据中心。它有点像您家里的车库。不需要多个电源或现场柴油发电机等,也不需要多个不同的网络连接。基本上 requirement for things such as multiple power feeds or an on-site diesel generator. There's also no requirement for multiple diverse network connections. Basically,

不需要任何级别的组件冗余。您几乎可以自己在家建造一个!因此,一级数据中心极易受到意外停机的影响,因此您几乎肯定不想在其中运营业务。

2 级

2 级数据中心比 1 级数据中心高出很多。2 级数据中心拥有冗余的现场发电机和不间断电源,以及一些组件级冗余。但是,2 级数据中心仍然需要停机以进行某些计划的维护活动,例如年度电力维护和测试。

第 3 级

第 3 层数据中心以第 2 层为基础,并添加 N+1 冗余电源和冷却路径(例如,所有设备必须具有多个电源和电源馈送)。

这些冗余路径只需在主动/被动模式下工作,但每条路径必须能够独立处理整个负载。这种主动/被动操作模式允许您执行计划维护而不会导致停机,但不一定能在主动电源路径发生意外问题时为您提供保护。Tier 3 数据中心需要能够长时间为整个数据中心供电的现场发电机。

第 4 层

4 级数据中心是目前定义的最高级别。4 级数据中心可以安然度过电力、冷却和网络等基本服务的意外中断,而不会导致停机。所有这些主要系统都以多个 N+1 冗余和物理隔离的形式实施,这意味着电力系统可能有两个独立的路径,每个路径都是 N+1 冗余的。4 级数据中心还需要现场发电机,能够长时间为整个数据中心供电。

如果您能够负担得起,您可能希望在 4 级数据中心开展您的业务。

如果您与私有云供应商合作,您需要确保他们的数据中心是 3 级或 4 级数据中心。

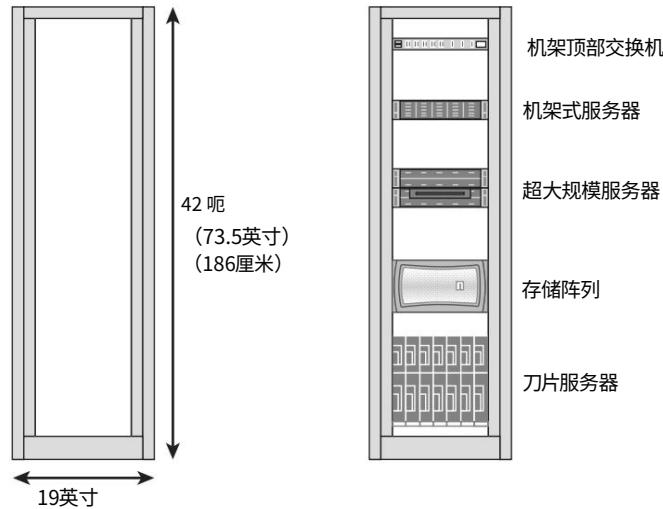
数据中心概览

现在让我们来看看数据中心的一些主要组件。

机架

数据中心的几乎所有 IT 设备都安装在标准化机架中,称为 19-英寸机架。图 13.1 显示了两个并排的行业标准 19 英寸机架的正面视图。一个机架是空的,另一个机架部分装有各种计算、存储和网络硬件。

图 13.1 19 英寸机架



这些 19 英寸机架之所以得名,是因为无论从正面还是背面看,它们的宽度都约为 19 英寸。这种标准宽度加上标准化深度,使服务器、存储阵列和网络交换机等 IT 设备能够采用标准尺寸,并且几乎可以适应所有数据中心机架。这使得在数据中心安装设备比制造商在 19 英寸机架标准化之前容易得多。



According to the EIA-310-D standard, 19 inches is the width of the mounting rails on the equipment. The space between the rails is 17.72 inches (450 毫米).

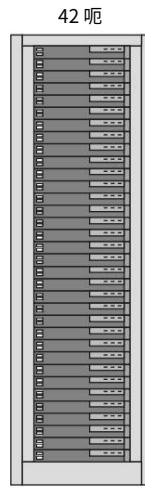
在谈论 19 英寸机架的高度时,我们通常以机架单位 (RU) 为单位,标准 RU 为 1.75 英寸。通常,我们缩短术语机架单位
简单地说是单位或 U。因此,流行的术语和机架尺寸是 42 U,即高度为 42 个机架单位,或 73.5 英寸。我们还用 RU 或 U 来测量服务器、交换机和存储阵列的高度。例如,单个 42 U 机架可容纳多达 24 个 2-U 服务器,直接堆叠在一起,如图 13.2 所示。但应注意的是,虽然单个机架可以物理上完全填充,但数据中心可能无法提供足够的电力和冷却来让您完全填充机架。

在 19 英寸机架中安装设备的过程称为上架,
移除设备的过程称为卸下机架。



确保在将设备放入或取出时使用正确的起重技术。通常情况下,应该有两个人参与,这样您就不会受伤或损坏设备。More often than not, there should be two people involved, so you don't injure yourself or cause damage to equipment.

图 13.2 个 19 英寸机架,装有 24 台 2U 服务器



重型设备应放置在机架底部。当重型设备靠在机架上或从机架上滑出时，重型设备不会翻倒。机架应固定在坚固的墙或安装特殊的滑轨架，以提供额外的稳定性和抗倾倒性。 *Handling heavy equipment in and*

遗憾的是,存储阵列是需要非标准专用机架的罪魁祸首之一。数据中心将整排机架专门用于存储设备的情况并不少见,因为存储阵列通常安装在定制机架中,无法安装在标准数据中心机架中。不过幸运的是,大多数存储供应商都开始支持将其设备安装在行业标准的 19 英寸机架中。

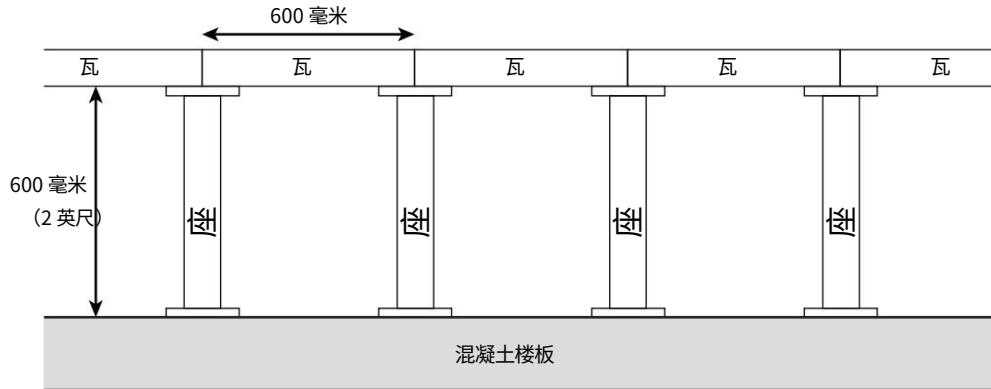
美国电气规范要求将机架接地,有时也要求将机架内的设备接地,以保护人员和设备免受电击和静电。出于同样的原因,强烈建议在数据中心使用设备时佩戴防静电腕带,尤其是在打开设备盖子时。

地板

现代数据中心普遍采用架空地板。架空地板是一种人工地板,通常由 600 毫米 × 600 毫米的地砖制成,并放置在基座上。

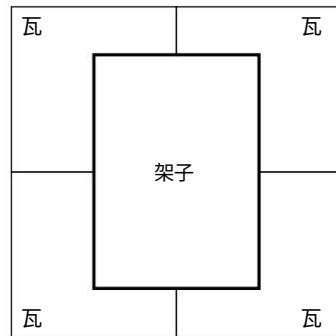
地板砖是可拆卸的,工程师可以将其抬起,进入架空地板和下方实心地板之间的空隙。该地板空隙通常用于布线网络电缆和电源线,以及可能引导冷空气。图 13.3 显示了一个人工架空地板,它位于支撑基座上的混凝土地板上方 2 英尺处。

图 13.3 架空地板和空隙



了解高架地板能承受多大的重量很重要,这样你就不会承受太大的机架的重量超过地板的额定承载能力。当涉及到您的数据中心架空地板的承载能力时,您应该始终咨询专家。通常的做法是安装机架,使其跨越多个地板砖。通常,每个地板砖都能承受一定的重量,并且将机架放在四块地板砖上(如图 13.4 所示),有时可以为您提供所有四块地板砖的强度和承重能力。但是,在做出此假设之前,请务必咨询知识渊博的专家。

图 13.4 分布在四块地砖上的数据中心机架的鸟瞰图



行

在大多数数据中心,机架被排列成行,如图 13.5 中的鸟瞰图所示。

通过这种方式,可以通过给出货架的行位置以及货架的架子位置。例如,图 13.6 中突出显示了 A 行中的架子 3。

图 13.5 5 个机架按行排列

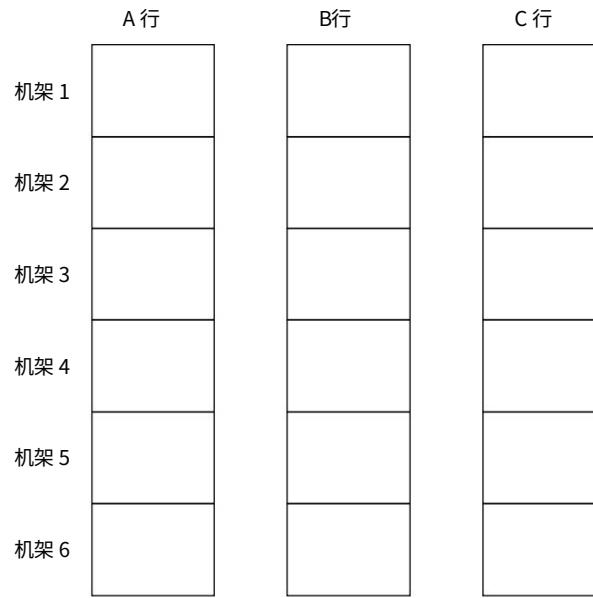
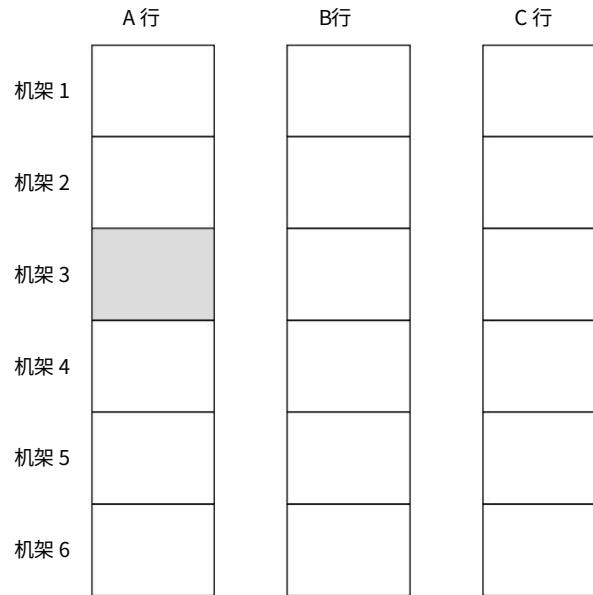


图 13.6 将架子 3 放置在 A 排



在本章后面,您将看到机架排在供电和冷却方面的重要性。现在,只要知道如果将机架整齐地排列,数据中心的设计和管理就会简单得多就足够了。

力量

电力是现代数据中心最重要的考虑因素之一!如果您认为世界对越来越多的存储的无限需求超出了范围,那么数据中心对电力的需求也同样如此!

数据中心需要电力,而且往往无法获得足够的电力。事实上,一些大型数据中心消耗的电力比一个普通城市还要多,这并不罕见。但消耗电力的不仅仅是服务器、存储和网络设备。供暖、通风和空调 (HVAC) 系统消耗大量电力。在世界上的某些地方,HVAC 消耗的电力比 IT 设备还要多,这意味着许多数据中心维持适当温度所消耗的电力比维持设备通电所消耗的电力还要多。

但不仅仅是整个数据中心都在为电力而苦苦挣扎。许多数据中心都在努力为单个机架提供足够的电力和冷却。例如,现代服务器消耗的电力和产生的热量比以往任何时候都多,而只能为每个机架提供 5-6 千瓦左右的数据中心根本无法提供足够的电力来装满整个机架。如果您的数据中心无法提供足够的冷却电力来让您装满数据中心机架,那么您的机架可能不得不半空运行。许多现代数据中心现在为每个机架提供 20 千瓦或更高的电力,因此每一寸机架空间都可以得到利用!



In 2011, Google announced that its data centers had consumed 260 megawatts (MW) of power, which is equivalent to 260,000 households. This is approximately twice the electricity usage of a medium-sized city. Although Google is not the world's most energy-consuming company, it has committed to taking many measures to offset and reduce its impact on the environment.

由于我们目前所处的电力限制,供应商正在放置重要的不能专注于提高服务器、存储和网络设备背后的电源 (PSU) 的效率。例如,如果您每年为 IT 设备供电的成本为 500,000 美元,那么将 PSU 效率提高 20% 可以显著减少您的年度电费。

业界使用电源使用效率 (PUE) 等级来衡量
数据中心的 PUE 由以下公式确定:

$$x = \frac{\text{数据中心的总供电量}}{\text{IT设备消耗的总电量}}$$

其中 x 是效率等级。PUE 为 2 表示您为实际 IT 设备 (服务器、存储、网络) 供电的每瓦特电力会产生 1 瓦特的开销。PUE 为 4 表示您为设备供电的每瓦特电力会产生 3 瓦特的开销。因此,显然 PUE 等级越低越好。

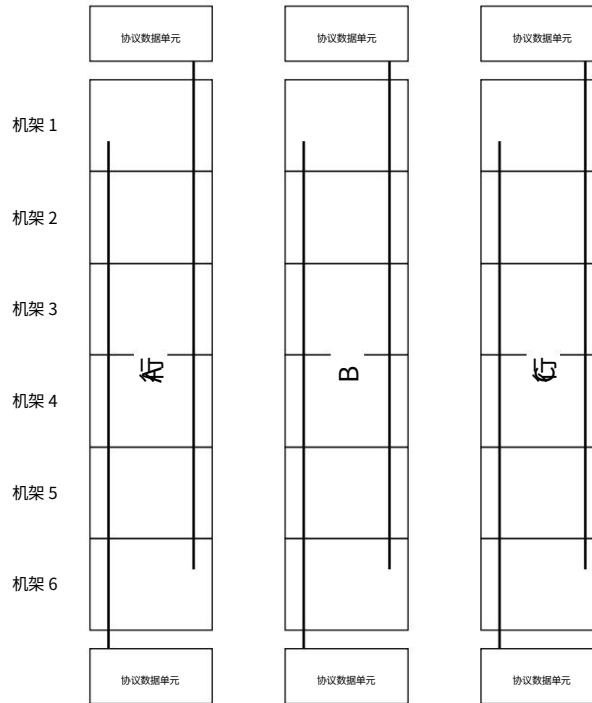
数据中心通常有多个独立的电源供应（如果可能的话，最好来自不同的公用事业公司）。这样，即使没有一家公用事业公司的电力供应，数据中心也不必依靠现场发电机供电。



为了在公用电源断电时为数据中心供电，大多数数据中心都配备了基于柴油发动机的备用电源系统。如果公用电源断电，这些现场柴油发电机将自动为数据中心供电。确保公用电源恢复的通常做法是确保保留足够的燃料，以便能够为数据中心提供至少 48 小时的满负荷供电。显然，定期（至少每年一次）测试这些现场发电机非常重要，这样您就知道在需要时可以依赖它们。作为此类发电机重要性的一个例子，当 2012 年飓风桑迪袭击时，纽约和纽瓦克地区的许多数据中心不得不长期依赖发电机供电。

在数据中心内部，通常为每个机架提供两个独立的电源。该电源通常通过每排末端的独立配电单元 (PDU) 供电，如图 13.7 所示。

图 13.7 数据中心行的配电

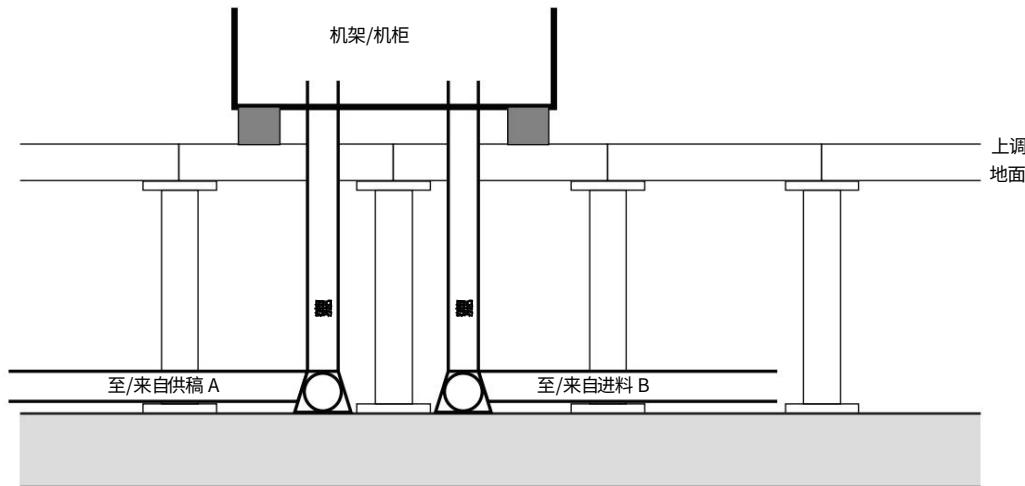


在图 13.7 中,每行均由 PDU A 和 PDU B 供电。对于每一行,PDU A 都独立于 PDU B,这意味着一个 PDU 可能会发生故障,或者为其供电的电源可能会发生故障,而不会影响另一个 PDU。例如,如果 C 行的 PDU B 发生故障,PDU A 将不受影响,并将继续为 C 行机架上的设备供电。这种冗余电源配置允许计算、存储和网络设备使用双电源供电,并能经受住许多电源故障情况。这还允许您使用一个电源进行维护,而无需关闭设备电源或使用现场发电机供电。

所有电源通常都由不间断电源 (UPS) 保护,以确保电源干净且无尖峰。如果公用电源短暂中断,UPS 系统通常会提供几秒钟或几分钟的电力。这意味着计算、存储和网络设备不会受到短暂的电力尖峰或下降的影响。

电力通常通过地板空隙输送到机架,如图 13.8 所示。

图 13.8 通过地板空隙为机柜供电



每个机架的电源通常在排尾的 PDU 上都有自己的断路器,允许您逐个机架地隔离电源。

在没有架空地板的数据中心设施中,也可以将电源线架空,并将工业和多相电源插头和插座悬挂在天花板上,以便您可以轻松地将它们连接到机架式电源板上。与在地板空隙中布线一样,将它们架空布线可防止在数据中心进行日常工作时妨碍它们。



电源通常被称为数据中心的机电 (M&E) 子系统或 the mechanical and electrical (M&E) subsystems of the data center.

冷却

冷却是数据中心的一项重要业务,需要耗费大量电力来保持数据中心的冷却。为了减少保持设备冷却所需的能量,大多数数据中心都采用热/冷通道原理。大多数现代 IT 设备的设计方式是风扇从前面吸入冷空气,从后面排出热空气。在热/冷通道配置中,机架的安装方式是,一排中的所有机架都面向同一方向,而相对的行则以背对背的配置彼此面对。这意味着相邻行中的设备始终面向背对背或正面对正面,如图 13.9 所示。



冷却与功率有关,也与设备位置有关。Uptime Institute 报告表明,机架上部的设备比机架下部的设备更容易发生故障。Uptime Institute report suggests that equipment

这是因为冷空气上升时会变热,从而使得机架顶部的温度高于底部。

图 13.9 热/冷通道配置

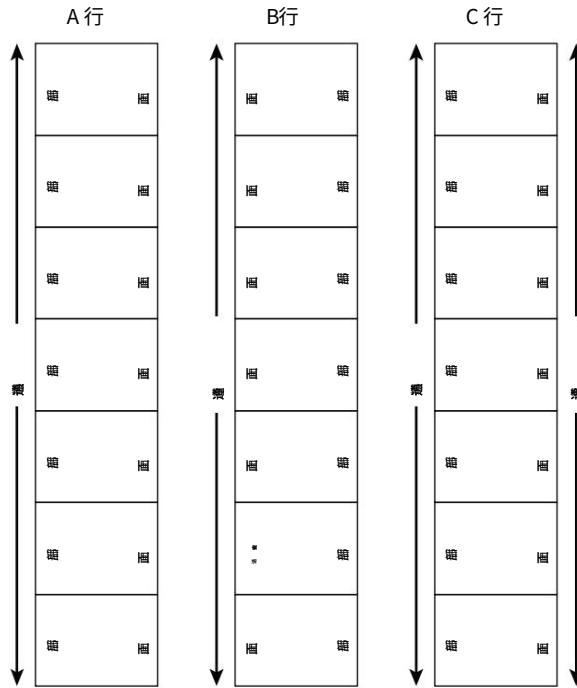
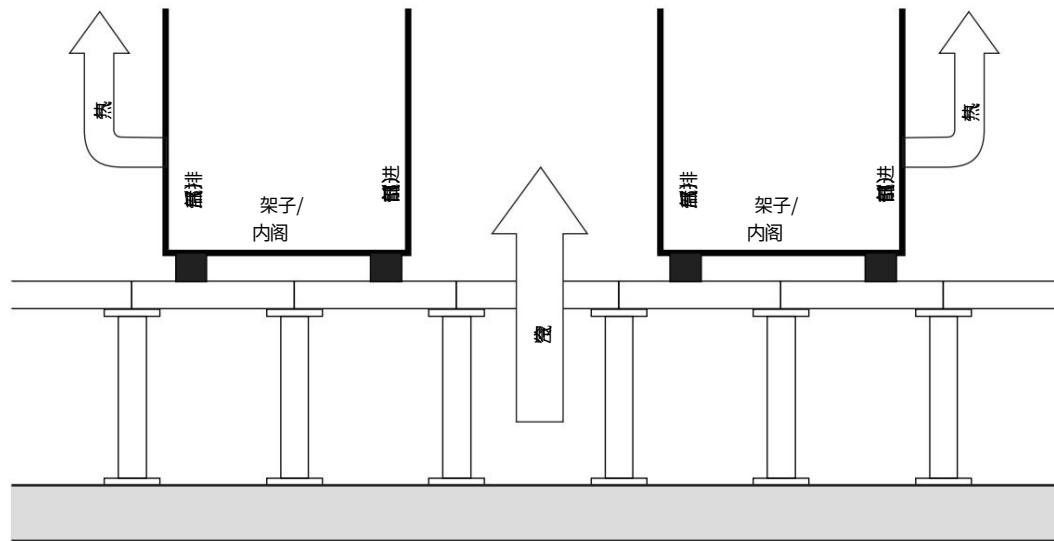


图 13.9 所示的热/冷通道配置有助于保持热空气和冷空气分离。与将一些设备吹出的热空气吸入其他设备前部相比,这可以使数据中心冷却系统更加高效。

设备；试图通过使用热空气来保持设备凉爽并不是一个好主意。然而，即使是热/冷通道数据中心，热空气和冷空气仍然会混合，因此，一些数据中心决定封闭热通道或冷通道。例如，在将冷空气保持在封闭通道内的数据中心中，来自外部的热空气无法进入封闭区域并升高温度。通过地板空隙将冷空气泵送到数据中心周围，并通过穿孔地板砖将其进入冷通道，这也是一种流行的做法，如图 13.10 所示。

图 13.10 冷空气通过地板空隙



大多数现代数据中心机架都配有穿孔的前门和后门，因此机架可以呼吸；空气可以通过前门和后门上的微小穿孔进出机架。如果您的机架没有穿孔的前门和后门，您应该认真考虑将门完全拆除。事实上，在大型数据中心，拆除前门和后门并不罕见。



有些设备，尤其是较旧的网络设备，不符合前向气流。有些设备从左侧吸入冷空气，从右侧吹出热空气，有些设备甚至从下方吸入冷空气，从顶部排出热空气。如果您运营的数据中心是围绕热/冷通道设计建造的，则应尽量避免在数据中心的同一区域使用这种非标准设备。

数据中心通常位于战略地理区域，这有利于提供机会利用当地的环境条件，使数据中心尽可能绿色环保。例如，凉爽的地区

空气可以被引入数据中心并用于冷却设备或阳光充足的地区,从而能够充分利用太阳能。

在所有有关机架、排、电源和气流的讨论中,人们很容易忘记所有这些的目的都是为服务器、存储和网络设备创造最佳的运行环境,同时也要以最有效的成本来实现。

数据中心机架中的服务器安装

现在我们已经讨论了很多关于机架、电源冷却和相关事项,练习 13.1 将引导您完成将新服务器安装到数据中心机架所需的常见步骤。

练习 13.1

安装新的 2U 服务器

本练习概述了在数据中心机架中安装新设备所需的一些步骤。

1.确定要安装新设备的机架是否有足够的空间容纳该设备。请检查以下内容:

- 机架中有足够的物理空间吗?机架至少需要两个空闲机架单元。
- 机架中的电源板上是否有足够的空闲插座来接通电源
新服务器的电缆?
- 机架是否有足够的备用电源来处理新服务器的功耗?
- 机架是否有足够的冷却能力来承载你的新服务器?
- 机架是否有合适的铜缆和光缆?
- 进行网络或电源线布线工作时是否需要拆除地砖?
- 地板和机架能否承受添加新的 2U 服务器的重量?

2.确定有足够的空间、电源、冷却等后,您需要规划新服务器的安装。此规划通常包括以下许多注意事项:

- 您的新服务器何时、由谁交付?大多数数据中心不允许在未事先通知的情况下交付设备。您可能需要预先通知数据中心的安全办公室,并向他们提供设备描述以及交付设备的公司名称。
- 谁来负责将设备架起来?您需要两个人来负责将设备架起来
设备。

■设备何时可以上架?如果要在生产车间内将设备上架到机架中,
数据中心的区域,物理机架工作可能需要在周末进行,并且可能需要变更管理委员会正式同意授权的变更号码。

■您是否有轨道设备来将您的服务器物理地装入机架?

■是否要求使用防静电腕带、防静电垫或其他特定
设备?如果有,它们存放在哪里?

■是否有需要你遵守的指导方针或协议?请务必
提前审查它们,并就任何疑问与相关方进行跟进。

■您需要提前填写哪些文件或表格,以及提前多久
需要提前填吗?

■您是否需要与网络和存储团队合作,以便在 IP 和存储网络上对新服务器进行物理和逻辑配置?

3.一切准备就绪后,安装新设备的日子就到了。在前往数据中心之前,请确保所有电子文件都已整理好。确保数据中心的保
安人员正在等候您的到来,并允许您进入现场。还要确保您的设备已送达,并了解其当前存放在哪里。

4.当您到达并可以访问站点和拥有新服务器的计算机房时,请按照类似以下步骤操作:

a.找到要放置服务器的机架。

b.打开机架,确保安装新服务器时没有障碍物。小心连接现有服务器和存储设备的电源线和网线。你肯定不想弄坏
任何一根。

c.打开服务器包装。大多数数据中心的标准做法是
在建造室内拆箱新设备,而不是在数据中心的计算机房中。

d.安装机架设备(导轨)。在工作过程中,您可能需要佩戴防静电腕带或使用防静电垫。

e.将服务器抬到导轨上,然后轻轻地将其滑入到位,确保不会
在执行此操作时,请拔下或损坏任何其他电缆。

f.连接网络和电源线,包括移除所有所需的地砖。

g.启动您的服务器。

h.如果您必须抬起任何地砖,请将其放回原位。

i.确保机架中的所有其他系统都正常,然后关闭机架门。

j. 测试服务器，并遵循所需的任何协议来帮助网络和存储团队在 IP 和存储网络上配置您的服务器。

k. 处理好设备交付时的所有包装。

这就是全部内容。当然，每个场景都是独一无二的。在执行这些步骤时，您可能会发现您的流程在几个点上有所不同，但此练习是您遵循的一般步骤的良好蓝图。

到目前为止，我们讨论的一切都是为了创造最佳的商业环境
应用程序和 IT 设备在其中运行。现在让我们来看看数据中心布线。

数据中心布线

如果您还不知道电力和 HVAC 对数据中心有多么重要，希望您读完前面的部分后能够明白。让我们继续了解良好的布线同样重要。

遵循标准原则

对于任何数据中心来说，良好、深思熟虑且正确安装的布线都是绝对必要的。这将使数据中心的日常管理和未来发展变得比以前简单得多。在为数据中心布线时，需要考虑几个总体原则。它们包括以下内容：

- 规划未来
- 测试
- 记录

在规划未来时，请尝试使用使用寿命为 10 年的电缆。这包括电缆的坚固性以及其传输未来 10 年预期数据速率（带宽）的能力。例如，如果您正在安装铜质双绞线电缆，请确保安装最新类别，以便它有传输 40G 以太网的机会。光纤电缆也是如此；安装最新的光纤模式（OM），例如激光优化的 OM3 或 OM4，它们都能够支持 40G 和 100G 以太网（包括 FCoE）。通过安装廉价电缆来偷工减料和降低成本可能会为您节省一些现金，但几年后，它就会给您带来麻烦。在数据中心维护中，没有什么比重新布线更艰巨和昂贵的任务了。

规划未来不仅仅是选择合适的电缆类型。它涉及到诸如拥有足够大的地板空隙以及在建筑物内不同楼层之间拥有足够的空间用于导管等事项，以便您今天能够进行电缆布线（这样您就可以轻松地保持在建议的弯曲半径内等等），以及将来增加电缆容量。



Real World Scenario

布线占用空间

一家公司在公司总部安装计算机房时没有考虑周全。该公司没有留出足够的空间在楼层之间铺设电缆，最终不得不拆除整个电梯井并将其重新用作楼层之间的电缆管道。

测试对于布线，尤其是结构化布线至关重要。如果您要自己进行测试，请帮自己和您的组织一个忙，使用高质量的布线和高级测试设备。如果第三方将为您进行电缆测试，请确保只接受最高分作为及格分数。您绝对不想为廉价且未经适当测试的结构化布线负责。结构化布线系统中的故障电缆通常无法更换；您只能忍受无法使用的电缆，并确保不要通过它连接任何东西。因此，请务必遵循木匠的座右铭“测量两次，切割一次”。



Copper cabling is susceptible to interference from nearby devices, techniques, and pathways. Some best practices include not mixing CAT 6A cables with CAT 5 or 5e cables, and avoiding mixing them in the same ducts and pathways so that you can implement cabling best practices. Some cabling best practices include things like not mixing CAT 6A cables running 10G Ethernet in the same ducts and pathways as CAT 5, CAT 5e, and CAT 6. Excessive pair density may also lead to AXT in copper patch panels.

记录布线也很重要。这包括容易被忽视的事情，例如在配线架上使用标签，以及在电缆末端贴标签。如果不记录布线，布线就会失控，变得一团糟，难以管理。Cabling is something that will get out of control and become an unmanageable mess if you don't document it.

使用结构化布线 Structured Cabling

最常见且最推荐的数据中心布线形式是结构化布线。结构化布线是指它遵循一套定义明确的标准。这些标准多年来一直致力于提高数据中心内布线的质量和可管理性。These standards have been established over the years to improve the quality and manageability of cabling within the data center.



结构化布线的对立面是点对点布线或简单的临时布线。这两种布线都无法扩展。点对点布线是指将电缆直接从服务器连接到交换机再连接到存储。电缆通常在机架内和机架之间布线，而无需使用配线架。

这些解决方案可以在一些小型数据中心（通常称为计算机房或设备室）中发挥作用，这些中心没有空间或不需要严肃、长期、可扩展的布线解决方案。

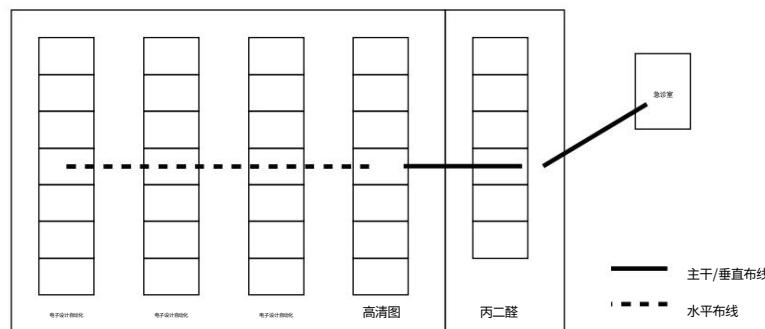
数据中心布线的事实标准是 TIA-942，由 TIA 于 2005 年制定了该标准。除了建议使用尽可能高容量的电缆（面向未来）外，该标准还概述了与电缆相关的数据中心的以下主要组件：

- 入口房间（ER）
- 主要分布区（MDA）
- 水平分布区（HDA）
- 设备分布区（EDA）
- 区域分布区（ZDA）
- 垂直布线/主干布线
- 水平布线

从非常高的层次来看，它们协同工作的方式如下。所有外部网络连接都到达数据中心的入口室（ER）。从这里，主干布线将 ER 中的设备连接到网络核心所在的主分布区（MDA）。

然后，更多的主干电缆将 MDA 中的设备连接到水平分布区（HDA），分布层交换机（通常是核心 SAN 交换机）就位于此处。水平电缆从 HDA 延伸到设备分布区（EDA），服务器、存储阵列和边缘交换机就位于此处。这些都显示在图 13.11 中。

图 13.11 高级结构化布线



但这只是一张高层次的图片。让我们更深入地看看每个组件。

入口房间

根据 TIA-942,每个数据中心必须至少有一个 ER,但可以根据需要拥有更多。ER 是您所有外部网络连接的入口 来自网络提供商的连接以及与其他园区网络的所有连接。基本上,任何想要进入数据中心的网络都必须通过 ER 进入。ER 实际上是数据中心内部与外部世界之间的缓冲区。在网络术语中,我们将其称为分界点 您在此划定网络提供商的网络设备和您自己的网络设备之间的界线。在 ER 内,您将您的设备连接到网络提供商的设备,并且您的网络提供商在 ER 之外不应有任何网络电路或设备。

急诊室以外的任何地方都应由您拥有和管理。

拥有多个 ER 可以为外部链接提供更高的可用性。例如,如果您的所有外部链接都汇入一个 ER,而该房间发生电源故障、洪水或其他事故,您就有失去与数据中心的所有外部连接的风险。

急诊室可以位于您的计算机室内或室外。



计算机房是您的 MDA 和 HDA (包括服务器、存储和网络设备所在的地方 servers, storage, and network equipment) reside.

主要分布地区

主分布区 (MDA) 位于计算机房内,通常容纳核心网络交换机和路由器。这些是您的网络交换机和路由器,由您拥有和管理。

MDA 的朝外侧连接到 ER,朝内的一侧连接到 HDA。这两个连接都是通过主干布线 (有时称为垂直布线) 进行的。主干布线通常是 ber 布线。

严格来说,MDA 的内部可以直接连接到您的 EDA,也可以通过 HDA 连接到 EDA。您选择哪种设计将取决于规模。较小的数据中心通常会省略 HDA 和从 MDA 到 EDA 的直接电缆。但是,较大的数据中心和计划扩展的数据中心通常会实施一个或多个 HDA。除了规模之外,HDA 还可以使进行与电缆相关的更改变得更容易且风险更低,因为它们通常允许您减少 MDA 内所需的电缆更改次数。

水平分布区域

水平分布区 (HDA) 是可选的。如果您的数据中心很大或者您计划将其扩大,则需要实施一个或多个 HDA。如果您希望数据中心始终保持较小规模,则可以跳过 HDA,直接将 MDA 连接到各个 EDA。



在谈论数据中心的大小或规模时,高端口数 (数据中心的网络端口数量) 并不总是决定 MDA 和 HDA 的必要性。如果 MDA 到各个 EDA 的距离超过建议的电缆长度,HDA 也很有用。这是因为 HDA 中采用的交换技术是主动的,这意味着它们会重复信号。重复信号允许信号传输到很远的距离。

如果您在数据中心实施 HDA,有时核心 SAN 交换机可能就位于此处,因为它们充当存储阵列和服务器所连接的边缘交换机之间的整合点。不过,在数据中心的 EDA 中安装 SAN 交换机也很常见。在 IP 网络中,HDA 通常映射到聚合层。

MDA 和 HDA 之间的布线是主干布线。



非常大的数据中心还可以实现位于 MDA 和 EDA 之间的区域分布区 (ZDA),并充当整合点,从而能够提供更大的灵活性。on the HDA and EDA and act as consolidation points to

设备分布区

设备分布区 (EDA) 是大多数服务器和存储活动发生的地方,因为这里是服务器、存储和架顶式 (ToR) 网络交换机所在地。

HDA 和 EDA 之间的布线被视为水平布线。虽然过去铜线在水平布线中占主导地位,但很快它就被光纤布线取代,因为光纤布线能够提供 40G 以太网,并且正在规划处理 100G 以太网。

铜双轴直连电缆 (DAC) 是

EDA,尤其是单机架内的布线。例如,铜双轴电缆通常用于服务器和 ToR 交换机之间的布线,因为它能够在长达 5 米的距离内传输 10G 以太网。铜双轴电缆解决方案也比其他铜线替代方案 (如非屏蔽双绞线 (UTP) 和屏蔽双绞线 (STP)) 更节能,因为它们在服务器和交换机中使用了更节能的 PHY 芯片。

除了上述概念外,TIA 还建议使用尽可能高容量的布线介质 (例如激光优化的 OM4 和可用的最高类别规格),以确保数据中心的未来发展并减轻未来潜在的干扰。这是合理的建议。请仔细考虑布线的预期寿命。

在数据中心工作

有一些规则和礼节可以帮助您完成在数据中心开展的任何工作。

访问控制

访问控制级别因数据中心的规模和性质而异。一般来说,只有经过授权的人员才可以访问您的数据中心。即使是您作为 IT/存储管理员,也可能不允许不受控制地访问数据中心。数据中心通常设有高围栏,并设有由专职安保人员看守的安全门。要通过前门,您通常需要获得预授权。



Real World Scenario

数据中心安全应受到重视

一家大型全球金融机构的 CIO 突然出现在该公司的一个主要数据中心。由于他的名字不在当天有权访问数据中心的人员名单上,因此他被拒绝进入。

一旦通过安全门,就可以正常进入急诊室和计算机房间必须配备电子访问系统,例如刷卡访问或生物识别访问。此外,一些数据中心将其服务器、网络和存储设备放在带锁的柜子里,使用徽章或密码访问。打开柜子需要预授权。

急诊室通常设在计算机室外,这样工程师您的网络提供商无需访问您的计算机房。

防火系统

显然,数据中心发生火灾可能会给您的组织带来灾难性的后果。数据中心需要一流的火灾预防系统,包括以下内容:

- 火灾报警系统
- 灭火系统

说到灭火,数据中心有各种各样的灭火系统,从小型手持灭火器到昂贵的顶级气体灭火系统。如果可能的话,应避免使用水基灭火系统,因为水和计算机设备通常不太相容!然而,通过使火灾缺氧来控制和扑灭火灾的气体抑制系统可能很昂贵。

一般行为准则

所有数据中心都应禁止饮食。IT 人员在控制台上工作时将半杯饮料放在服务器上的时代已经一去不复返了。事实上,在许多数据中心,如果你被发现携带食物和饮料,你会被立即带离现场并被要求不要再回来。

自己离开后清理干净也是礼貌和最佳做法。这意味着离开时物品的状态应与到达时相同或更好。离开后,不要留下任何空电缆袋、备用螺丝或其他任何东西。务必记住将地板砖安全牢固地放回原位!毕竟,架空地板对于保持数据中心的适当气流通常很重要。



Real World Scenario

使用架空地板时要小心

一名工人在一间装有架空地板的小型计算机房中,将一排瓷砖全部抬起,以便轻松地在地板空隙中铺设一些电缆。但他没有意识到,由于整排瓷砖被抬起,架空地板被严重削弱。当他拆除这排瓷砖时,部分地板塌陷,损坏了地板空隙中的电缆和设备,这需要进行大量工程来修复和重新认证地板。

确保在数据中心内只执行授权的工作。不要过度兴奋,并决定在数据中心做其他工作。例如,不要未经授权就掀起地板砖。大多数数据中心都会配备闭路安全摄像头,以监控数据中心内外的情况,以及人们在数据中心时发生的情况。

最后还有一点需要牢记。不要按下大大的红色紧急断电 (EPO) 按钮!大多数数据中心都有一个大大的红色 EPO 按钮,用于在紧急情况下切断数据中心的电源。虽然按下此按钮后会发生什么可能非常诱人,但除非真的有紧急情况,否则请不要按下它!

概括

在本章中,我们了解了数据中心的重要性,以及并非所有数据中心都是平等的。我们了解了 Uptime Institute 维护的广泛接受的数据中心层级模型,并且第 4 层数据中心是最高层级,可提供最佳可用性和冗余性。我们还了解了数据中心机架、架空地板、气流和冷却以及电源。我们还了解了布线的重要性,以及它如何影响当今数据中心的可靠性,以及它如何为新技术提供未来保障。我们在本章的最后讨论了物理数据中心安全、重新检测和抑制系统,以及所有数据中心都应遵守的一些一般行为准则。

章节概要

机架 数据中心机架采用标准尺寸,应接地,并需要按照规划的布局进行安装,以促进数据中心内有效的空气流动和冷却。应尽可能避免使用非标准尺寸的机架,并且机架内的所有工作都应经过批准并极其谨慎地进行。

电力 现代数据中心消耗大量电力,并且通常配备备用发电机,以便在市电断电时提供电力。数据中心的大多数机架和行列都具有 N+1 冗余电源。

冷却 现代 IT 设备会释放大量热量,如果不保持冷却,就会发生故障。许多数据中心都设有热/冷通道,有时甚至会封闭通道以防止冷热空气混合。将热空气和冷空气分开可以大大降低保持数据中心冷却的成本。

结构化布线 布线是数据中心的重要组成部分。在结构化布线方面,不要偷工减料,确保您规划和部署的布线解决方案可扩展性好,并且能够满足您数据中心 10 年后的需求。

紧急断电按钮 每个数据中心至少有一个非常酷的紧急断电 (EPO) 按钮。如果发生真正的紧急情况,此按钮可以切断计算机房的所有电源,从而挽救生命。仅在真正的紧急情况下使用它。如果您不当使用 EPO 按钮,您应该会被解雇。

Machine Translated by Google

章节 14



融合 联网

本章涵盖的主题：

- ✓ 两大主流数据中心网络
- ✓ 存储网络要求
- ✓ 数据中心桥接/融合增强以太网
- ✓ 无损网络
- ✓ 拥塞通知
- ✓ 增强传输选择
- ✓ 以太网光纤通道



本章介绍了几乎所有数据中心中都存在的两种主要网络：以太网和光纤通道。您将了解如何通过协作构建更高效的数据中心。

将这两个网络结合在一起。您还将了解存储流量对网络的具体要求，以及传统以太网如何无法满足这些要求。然后，您将看到在新版本的以太网中实现的重大变化，该版本被称为数据中心桥接 (DCB) 或融合增强以太网 (CEE)。本章还讨论了光纤通道帧是如何在这些新的 DCB/CEE 网络上封装和传输的。本章最后介绍了这对存储网络设计的影响以及它在现实世界中的实现方式。

两个网络的故事

大多数数据中心至少有两个主要网络：

- 以太网
- 光纤通道网络

让我们快速地看一下每一个。

以太网

以太网是一种通用的万能网络。它兼具多功能性和可扩展性，是世界上最常用的网络。事实上，地球上几乎每个数据中心都会部署至少一个以太网网络。大多数数据中心都会部署多个以太网网络。

缺点是，以太网在性能方面并不确定，尤其是延迟。不过，最近的 10G 和 40G 以太网技术已在解决以太网的性能问题方面取得了一定进展。

以太网也是我们所说的有损网络。有损告诉我们，当以太网开始丢帧。这种情况发生在拥塞等情况下，当交换机正在接收帧但因为下一跳繁忙而无法转发时。当发生此类数据包丢失时，堆栈中较高级别的协议（如传输控制协议 (TCP)）会负责重新传输无法到达目的地的帧。



理解术语并尝试去理解和掌握它们的使用方法。本章将介绍一些常用的术语，包括但不限于以下这些：帧和包、有损和无损、交换机和路由器、以及层2和层3。在讨论帧和包时，我们经常将它们互换使用，因为它们在很多情况下是等价的。在讨论有损和无损时，我们也将它们互换使用，因为它们在很多情况下是等价的。在讨论交换机和路由器时，我们也将它们互换使用，因为它们在很多情况下是等价的。在讨论层2和层3时，我们也将它们互换使用，因为它们在很多情况下是等价的。

总而言之，虽然以太网不是世界上性能最好的网络技术，但它 technology in the world, Ethernet 几乎可以肯定它是用途最广泛、应用最广泛的技术之一。widely deployed.

财经频道 Channel

光纤通道(FC)网络是一种高速、低延迟的网络，几乎专用于传输存储流量。光纤通道很少用于存储网络以外的用途。FC还运行链路层信令，用于传达缓冲区信用和状态，使其成为一种无损网络技术，这意味着它不会因拥塞而丢弃数据包。这种低延迟和无损能力的结合使FC成为传输 SCSI 流量的理想选择，而 SCSI 流量无法很好地处理不可靠的传输。



At a high level, when we talk about link-layer technologies, what has been discussed so far is point-to-point technologies. These technologies are characterized by the fact that they are end-to-end technologies. This means that the connection between two devices is established directly without any intermediate nodes. Point-to-point technologies are typically implemented using optical fiber or copper cables. They are used in various applications such as local area networks (LANs), metropolitan area networks (MANs), and wide area networks (WANs). Point-to-point technologies are also used in storage area networks (SANs) and network-attached storage (NAS) systems.

缺点是，FC 网络的通用性和可扩展性不如以太网，更不用说更昂贵了。但它们是可靠的。事实上，光纤通道中的通道一词意义重大。该术语至少部分源于 SCSI 是一种通道技术，而通道技术与网络技术略有不同。通道技术的主要目标是提供高性能和低开销的互连，而网络技术的主要目标通常是提供通用性和可扩展性。光纤通道显然是同时提供低延迟和可靠性的另一种尝试。它做得不错，但远不如以太网等纯网络技术那样可扩展。最终，FC 提供了一个高速、低延迟的存储网络（通道），而且相对无竞争。head interconnects，whereas the main goal of networking technologies is often to provide versatility and scalability. Fibre Channel is obviously an attempt at providing both low latency and scalability. It does a decent job, but it is nowhere near as scalable as a pure networking technology like Ethernet. Ultimately, FC provides a high-speed, low-latency storage network (a channel) that is relatively contention free.

单一融合数据中心网络

两个物理和逻辑上独立的网络（以太网和光纤通道）的存在为整合创造了良好的机会。这正是融合网络的意义所在：通过基础设施合理化和整合来降低成本。以下是整合以太网和光纤通道网络的一些明显好处：

- 减少物理服务器中的网络适配器数量
- 减少每台物理服务器的线缆数量
- 减少基础设施中的交换机数量

因此，无需使用一组专用于光纤的网络适配器、电缆和交换机。通道和另一组适配器、电缆和交换机专用于以太网，而使用融合网络，我们只有一组网络适配器、电缆和交换机，并通过它们运行以太网和 FC 存储流量。这几乎有望使曾经的理想成为现实。您只需在第一天将一台服务器安装其中，并在其中安装融合网络适配器（CNA），就可以进行任何所需的网络。您可以进行通用 IP 网络、FC 存储网络、iSCSI 存储、NAS，甚至低延迟、高性能计算。一个网络适配器和一根电缆就可以完成所有工作！

这些减少是有好处的，特别是在大型数据中心和部署密集服务器的数据中心。随着数据中心机架上服务器和存储设备的密度越来越高，减少网络适配器、电缆和交换机数量的前景突然变得非常有吸引力。它还具有降低电力和冷却成本的积极作用。总而言之，它可以降低数据中心的运行成本并降低总拥有成本（TCO）。

然而，正如通常的情况一样，这并不像将您的 FC 存储流量放在任何旧以太网上那么简单。有一些特定的要求，如果不解决，将使您作为存储管理员的生活变得非常困难。让我们来看看这些要求。

存储网络要求

人们很容易忘记，作为一种存储网络技术，FC 的主要作用是促进启动器和目标之间可靠、高效地传输 SCSI 命令。它还尽可能地模拟旧的内部 SCSI 总线，而 SCSI 协议最初就是设计用于该总线的。

让我们回顾一下 SCSI。SCSI（小型计算机系统接口）最初设计用于物理服务器硬盘的连接，在相对较短的并行电缆上无争用地运行。无争用表示没有其他协议或流量

c 正在争夺相同的带宽。实际上不存在争用。因此，SCSI 的设计无法很好地处理延迟、拥塞或传输错误。事实上，当这些情况发生时，SCSI 处理得非常糟糕。因此，游戏的目标是不惜一切代价避免它们。

光纤通道在传输 SCSI 流量方面如此出色的原因之一是它具有某些功能和特性,可提供与服务器机箱内的本地 SCSI 电缆类似体验。其中一些功能和特性包括:

■简单的拓扑结构

■良好的带宽

■低延迟

■可靠性 (无损)

■确定性性能

另一方面,传统以太网与 SCSI 不太匹配。首先,以太网上经常出现拥塞。当存在争用时,您可以告别任何确定性性能或低延迟的希望。如前所述,以太网也是有损的。当发生拥塞时,它们会丢弃帧。这些对 SCSI 来说都不是好兆头。

总之,您当然不会简单地将 FC 帧包装在常规以太网帧内,然后将其丢弃到现有的千兆以太网网络上。显然,为了让以太网成为可行的 FC 存储流量传输方式,必须做出一些改变!

增强型以太网

为了创建一个能够传输 IP 和 FC 存储流量的单一数据中心网络,以太网必须进行大幅增强和升级。这些变化如此广泛,以至于许多人觉得它几乎不像我们大多数人初次接触网络时的以太网。



有时我们将数据中心网络称为统一结构,这两个术语的含义基本相同:可以传输 LAN 和 SAN 流量的单一网络。Terms mean basically the same thing: a single network that can transport

为了创建这种新的增强型以太网,电气和电子工程师协会 (IEEE) 在 802.1 工作组内成立了一个新的任务组。这个新任务组称为数据中心桥接 (DCB) 任务组,负责开发能够传输所有常见数据中心网络流量类型的数据中心以太网网络:

■ IP LAN 流量

■ FC 存储流量

■ InfiniBand 高性能计算流量

根据您与谁交谈,这种增强型以太网通常被称为数据中心桥接 (DCB) 或融合增强型以太网 (CEE)。我们有时也称之为数据中心结构或统一结构。

如果 CEE 想要传输 FC 存储流量,它需要进行一系列增强,包括以下内容:

- 增加带宽
- 服务等级
- 优先事项
- 拥塞管理
- 增强型传输选择 (ETS)

除了这些逻辑变化之外,还有一大堆新的硬件要求
电缆、网络适配器、交换机端口和交换机 您可能会怀疑这是否值得。让我们来看看我们之前列出的各项增强功能。

增加带宽

为了满足统一数据中心结构的带宽和其他需求,CEE 是一个无冲突的全双工网络,最低运行速度为 10 Gbps。如果我们想将多个 1 Gbps LAN 以及 2 Gbps 和 4 Gbps SAN 整合到一条线路上,这一点至关重要。10G CEE 还可以安全地传输 8 Gbps FC 流量,并允许最新一代 CNA 使用整个 10 Gbps 链路带宽。

10 Gbps 增强型以太网只是一个开始。40 Gbps 以太网和 100G 以太网已经问世,而且随着时间的推移,这两种以太网的价格都将变得实惠。像这样的强大技术路线图将以太网定位为未来的数据中心网络。

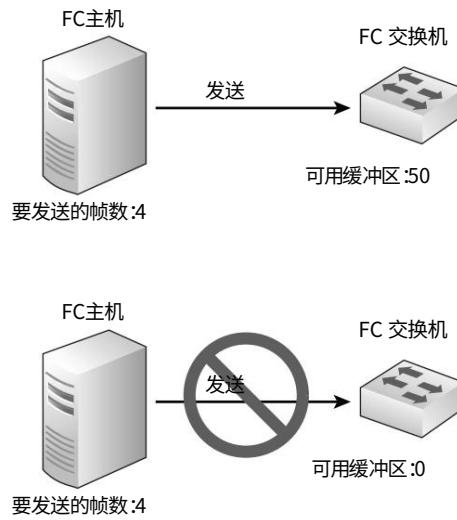
优先级和服务等级

我之前提到过,FC 存储流量需要无损网络。但无损网络到底是什么?简单地说,无损网络不会因为拥塞而丢弃帧。这与有损网络相反,有损网络通常会在发生拥塞时丢弃帧。请考虑以下高级且过于简单的示例:交换机接收到 100 个帧,但无法立即转发它们,因为下一跳的交换机已饱和并且无法接受更多帧。因此,交换机缓冲了这 100 个帧,并出现拥塞。交换机的缓冲区已满后,它无法再接受任何帧并开始丢弃(丢弃)它们。

网络通常通过链路层信令来实现无损行为,该信令跟踪远端的缓冲区。我们通常将此称为流控制。例如,在 FC 世界中,流控制是通过称为缓冲区到缓冲区 (B2B) 流控制的链路层系统实现的,其中发送方不允许发送帧,除非它明确知道接收方有足够的缓冲区来接收帧。它简单而有效地确保不会发生拥塞并且不会丢弃数据包。FC 流控制如图 14.1 所示。

数据中心桥接任务组已决定在
CEE 通过一种称为基于优先级的流量控制(简称 PFC)的机制实现。

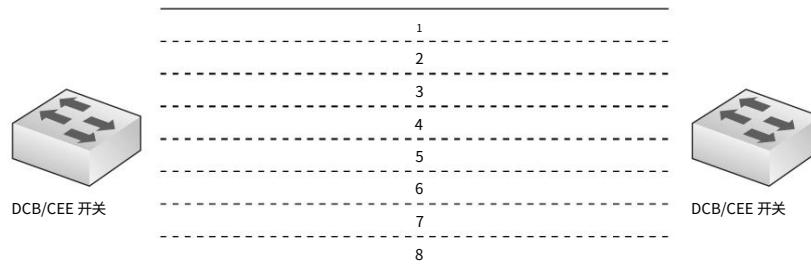
图 14.1 光纤通道 B2B 流量控制



In some older documentation, you may find PFC referred to as *per priority pause* or *class-based flow control*.

在深入研究 PFC 之前,值得花点时间简要介绍一下以太网优先级。CEE 定义了八个优先级,允许在链路层提供八类服务。这是通过用编码优先级标记帧来实现的,从而有效地允许将 CEE 网络的带宽划分为八个逻辑通道或虚拟链路。图 14.2 显示了两个 CEE 交换机之间的链路,该链路划分为八个逻辑通道,标记为 1 至 8。

图 14.2 CEE 网络分为 8 个逻辑通道



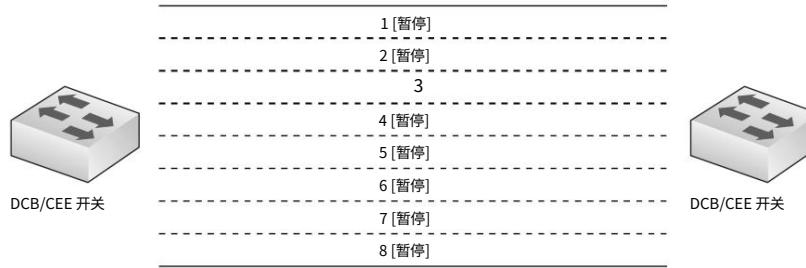
现在您已经了解了 CEE 网络中的八个优先级,让我们来看看 PFC 的工作原理。PFC 利用这八个逻辑通道(优先级),并可以有选择地为每个通道强制执行 PAUSE 条件。这是通过在网络上发出特殊的 PFC PAUSE 帧来实现的,指定以下内容:

- PAUSE 条件适用于哪个通道
- PAUSE 状态持续多长时间

通过发出暂停帧,指示除一条通道外的所有通道停止发送,您可以有效地优先考虑尚未暂停的车道。同样,如果您在单个车道上强制执行暂停条件,则实际上会降低该车道的优先级,并优先考虑所有其他车道。

图 14.3 显示了 CEE 的 8 个逻辑通道,其中除通道 3 之外的所有通道均暂停。

图 14.3 暂停条件应用于除 3 号车道外的所有车道



还可以选择性地消除暂停条件,例如如果拥塞比预期更快消散,则无需等到暂停超时到期。

拥塞管理

拥塞管理是一种尝试避免网络拥塞的方法,最好是在拥塞开始对网络造成问题之前从源头上进行处理。拥塞管理的常用方法包括速率限制和选择性暂停帧传输。

然而,试图解决大型网络中的拥塞问题是一项艰巨的任务,而且充满危险。拥塞管理系统可能会使情况变得更糟。到目前为止,标准化拥塞通知的尝试都没有得到广泛欢迎。

增强传输选择

增强传输选择(ETS)是标准化 CEE 网络上的服务质量(QoS)的一种尝试。ETS 允许在 CEE 定义的八个服务类别之间对网络带宽进行分割。这种带宽分配是动态的,这意味着如果某个特定服务类别未使用其全部带宽分配,则其剩余分配可供其他类别借用。但是,一旦分配的带宽被分配的类别需要,借用带宽的其他类别就必须归还。

这种动态分配和重新分配带宽的方式可以智能地使用可用资源,与以前的模型相比有了很大的改进,以前的模型中不同类别的流量都有自己固定的带宽分配,而不能共享。

数据中心桥接能力交换

有了 CEE 和 FCoE 带来的所有新功能和配置选项 (我们将在下一节中讨论) ,如果有一个简单、自动化且有效的机制来协商和配置这些选项,生活就会变得容易得多。

例如,当新设备添加到 CEE 网络中时,如果能够自动协商新添加的设备支持哪些功能和选项,那就太好了。如果能够将标准配置推送到这些新添加的设备,那就太好了。考虑到所有这些,我们设计了数据中心桥接能力交换协议 (DCBX)。

设备可以使用 DCBX 向网络中的其他设备 (如 CNA 和其他交换机)发出配置。这可以通过允许集中管理配置以及减少错误配置的机会来简化网络管理。



DCBX 是链路层发现协议 (LLDP) 的基础扩展 of the Link Layer Discovery Protocol (LLDP).

PFC 和 ETS 都是可以通过 DCBX 协商和配置的链路层功能的示例。毕竟,如果 CNA 不支持或不理解 PFC,交换机向新连接的 CNA 发出 PFC 配置和指令就没有什么意义了。

以太网光纤通道

以太网光纤通道 (FCoE) 的字面意思就是:通过以太网传输的光纤通道帧。用更专业的术语来说,光纤通道帧无状态地封装在第 2 层以太网帧中,并通过 IEEE 802.3 有线以太网网络传输。

光纤通道和以太网如何相互作用?就以太网而言,FCoE 与任何其他上层协议 (ULP) (如 IP 或 MPLS)没有什么不同。因此,以太网很乐意传输 FCoE 流量。在 IP 方面,就光纤通道而言,以太网只是一种新的物理传输介质,取代了 FC 堆栈的较低层,即 FC-0 和 FC-1。



The 标准由国际信息与通信技术委员会 (driven by the International Organization for Standardization's Technical Committee 113) 相同的 T11 技术委员会制定。该技术委员会有定义技术标准的职责。

为了使 FC 帧能够通过以太网传输（请记住，必须是 CEE 以太网网络），必须将 FC 帧封装在以太网帧内。让我们来了解一下封装。

FCoE 封装

每个 FC 帧的长度都略大于 2 KB。但是，标准以太网帧的长度只有 1.5 KB。不需要火箭科学家就能知道不能在 1.5 KB 内包含 2 KB。所以需要采取一些措施。定义 FCoE 标准的人有两个明显的选择：

■ 片段 FC 框架

■ 使用更大的以太网帧

从性能和简单性的角度来看，应不惜一切代价避免对帧进行碎片化。碎片化会增加封装过程的复杂性和开销，最终导致延迟和性能下降。这对 FC 存储流量来说不是好事。

使用更大的以太网帧需要使用一种称为巨型帧的以太网技术，该技术允许以太网帧的大小高达 9 KB。显然，高达 9 KB 的巨型帧完全能够封装 2 KB FC 帧，而无需进行分段。这使得封装过程尽可能简单、轻量和快速。每个 FC 帧对应一个 FCoE 帧。没有分段和重组开销，因此没有不必要的协议相关延迟。所有这些对于传输 FC 存储流量都很重要。因此，出于这些原因，标准机构决定采用选项 2，即使用巨型帧。



存储网络行业协会 (SNIA) 将巨型帧定义为长度为 2.5 KB 的帧，并指出 FCoE 封装至少需要小型巨型 2.5 KB 的以太网帧。

现在让我们进一步了解一下 FCoE 封装过程。

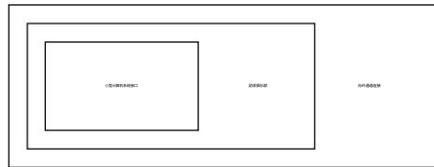
FC 是为了传输 SCSI 数据而存在的，SCSI 命令和数据已经封装在 FC 框架内，如图 14.4 所示。

图 14.4 封装在 FC 中的 SCSI



FCoE 通过添加封装层，进一步提高了这种封装的效率。将 FC 帧封装在 FCoE 帧内，如图 14.5 所示。

图 14.5 封装在 FCoE 中的 FC



FCoE 帧的封装是一个逐跳过程,这意味着网络中的每个 FCoE 交换机在收到 FCoE 帧时都必须剥离 FCoE 封装,然后在转发到下一跳之前重新封装该帧。如果这一过程不能快速有效地完成,很快就会开始影响性能。

因此,CNA 在 ASIC 上的硬件中执行这种封装,就像大多数 FCoE 交换机一样。



■ 面上有软件 FCoE 堆栈,任何需要昂贵的专用硬件即可实现 iSCSI。显然,这种设计需要使用主机资源 (如 CPU 周期) 来执行 iSCoE 处理。Virtualized CNAs. Obviously, this design requires the

尽管 iSCSI 在很大程度上走的是软件 iSCSI 启动器而非硬件 iSCSI 启动器的道路,但 iSCSI 的部署规模几乎总是比 FC (以及 FCoE)小得多。虽然不一定推荐用于生产用例,尤其是高性能要求,但软件 FCoE 堆栈对于实验室和测试环境非常有用。在 vSphere 5 中,VMware 引入了可与某些经过认证的 10G 卡配合使用的软件 FCoE 适配器。只要实验室中有支持 FCoE 的交换机,这就可以成为在实验室中试用 FCoE 的好方法。

关于 FCoE 封装,还有最后一个关键点。FCoE 封装过程保持 FC 帧完好无损。Source_ID (S_ID)、Destination_ID (D_ID) 和全局端口名称 (WWPN) 等信息均未受影响。这意味着 FCoE 不会放弃现有的 FC 概念和机制,例如分区、WWPN 和简单名称服务器。这些都保持不变,使我们能够继续使用多年来一直使用的熟悉的 FC 分区概念。



■ 值得明确指出的是,FCoE 不涉及将 IP 数据包封装在 FC 帧中。FC 帧封装在 IP 数据包中。IP 帧在以太网第 2 层以太网帧中,这意味着 FCoE 网络是典型的扁平第 2 层不可路由的网络。Protocols 无法通过第 3 层的 IP 路由。这与 iSCSI 形成对比,iSCSI 是封装在 IP 数据包中的 SCSI,可通过 IP 在第 3 层路由。

融合网络适配器

传统的以太网是通过网络适配器（称为网络接口卡 (NIC)）来访问的。每个想要连接到以太网的主机至少需要一个网络接口卡。

相反，传统的光纤通道网络是通过每个主机上的网络适配器（称为主机总线适配器 (HBA)）访问的。访问 FCoE 网络需要一种新型网络适配器，称为融合网络适配器 (CNA)。这三种网络适配器卡均以 PCI 适配器卡的形式实现。它们可以是扩展卡，也可以直接安装在服务器主板上，即所谓的主板上的 LAN (LOM)。

CNA 顾名思义就是：NIC 和 HBA 融合成一个网卡。出于性能原因，CNA 在硬件（通常是 ASIC）中提供 NIC 和 HBA 功能，以便 FCoE 封装等功能可以快速完成，而不影响主机 CPU 资源。

现在让我们看看 CNA 如何实现基础设施整合。在具有不同以太网和 FC 网络的传统数据中心中，物理服务器至少具有以下网卡的情况并不少见：

- 两个 1 Gbps 以太网 NIC，用于您的生产网络
- 两个用于 vMotion 的 1 Gbps 以太网 NIC
- 两个用于存储的 8 Gbps FC HBA
- 两个 1 Gbps 以太网（专用于管理流量）

八张网卡显然需要八根电缆。那只是电缆而已。拥有独立的以太网和 FC 网络也需要大量的网络硬件。您需要专用的以太网交换机和专用的 FC 交换机。但是，使用融合数据中心网络，同一台服务器中的网卡数量可以轻松减少，如下所示：

- 两个 10 Gbps CNA
- 一个 1 Gbps NIC（专用于管理流量）

两个 10 Gbps CNA 可以处理所有 IP 和 FC 流量要求。

为了融合网络，你可能需要克服的一个担忧是

共享单个 CNA 以实现多种功能。一些组织对网络适配器隔离有严格的政策，每个应用程序使用专用的网络适配器卡。因此，具有两个应用程序的服务器将具有两个 NIC，外加一个单独的 NIC 用于管理，等等。如果您要融合网络，就必须克服这种思维定势，因为您将在同一个网络适配器上运行 FC 和 IP！

当然，尽管与 HBA 没有太大区别，但 CNA 往往比 NIC 更昂贵。

FCoE 和无损网络功能

由于 FCoE 承载的是封装的 FC 帧，因此需要无损网络。如果网络/传输不是无损的，恢复和重新传输必须由 SCSI 处理，这不是好事。SCSI 的设计从来就不擅长恢复。

如本章前面所述，CEE 定义了 8 个逻辑通道，使我们能够在链路层，通过为帧添加标记来创建八个潜在的服务类别

编码优先级。例如,FCoE 帧可以标记为优先级 3,从而有效地将 FCoE 帧分配给通道 3 或服务等级 3。然后,CEE 技术 (例如 PFC)可以利用此优先级来确保在发生拥塞时可以优先处理或取消 FCoE 帧传输的优先级。



设置优先级P(服务类别)通常通过设置VLAN标签中的服务优先级位来完成的。
service priority bit in the VLAN tag.

让我们看一个简单的例子。图 14.6 显示了具有两个 FCoE 交换机的网络。两个 FCoE 交换机之间的链路被划分为八个通道 (优先级),由于通道 2 上的流量极高,网络正在经历拥塞。如图 14.7 所示,网络已向通道 2 发出 PFC PAUSE 条件,有效地降低了通道 2 的优先级,并优先考虑所有其他七个通道。这样做的好处是允许通道 3 上的 FCoE 帧继续传送,并使网络有机会从拥塞中恢复。

图 14.6 PFC 和 8 个逻辑通道

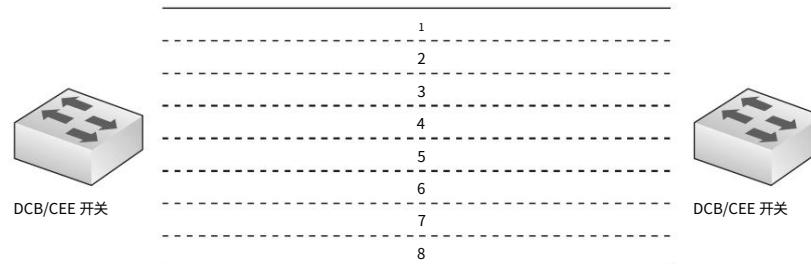
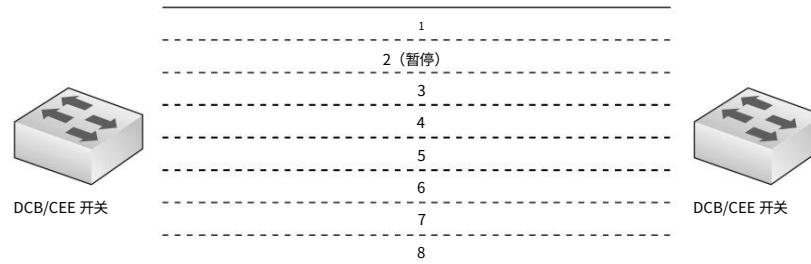


图 14.7 通道 2 上施加的 PFC 暂停条件



在图 14.7 中,PFC 选择性地将暂停条件应用于通道 2。此暂停条件指示通道 2 在一定时间内停止传输帧。一旦这段时间过去,帧就可以再次在通道 2 上传输。如果拥塞在暂停到期之前消散,网络也可以移除暂停条件。



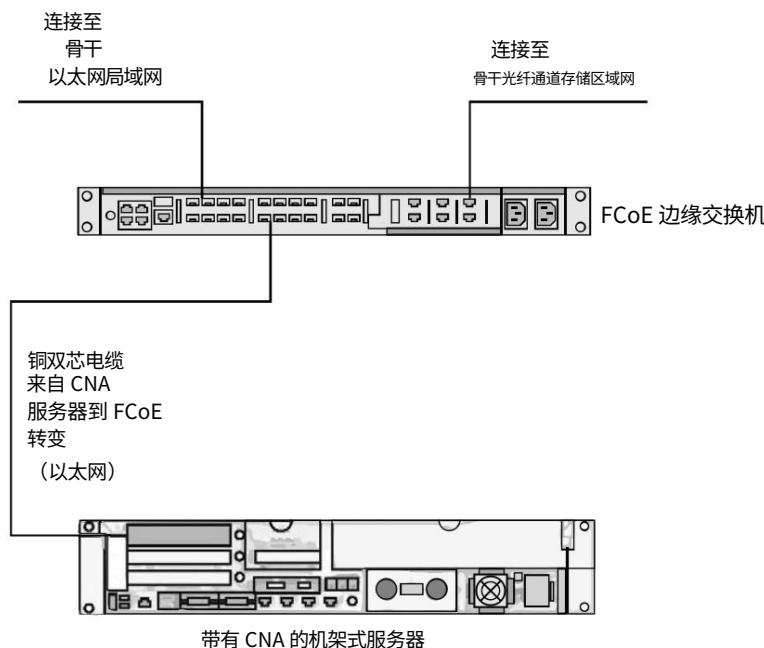
尽管 CEE 包含拥塞通知规范，但 FCoE 并不需要该规范。在 FCoE 网络中，流量控制由 PFC 处理，而不是 FC 级信用系统。 In CoE network, flow control is handled by PFC

FCoE 交换机

随着 CEE 和 FCoE 中出现各种新技术，必然需要新的交换机。幸运的是，市场上已有多家供应商提供大量此类交换机。

FCoE 交换机连接到 LAN 和 SAN。它既有 CEE 端口，也有 FC 端口。这对于已经在 FC SAN 技术上投入大量资金的公司非常有用，因为它允许将 FCoE 交换机无缝添加到现有的 SAN 环境中。图 14.8 显示了连接到企业 LAN 和企业 SAN 的 FCoE 交换机。

图 14.8 台 FCoE 交换机连接至传统 FC 设备



尽管 FCoE 帧通过 CEE 网络传输，但 FCoE 帧的转发并不是以太网转发的职责。这意味着典型的以太网第 2 层多路径技术（例如生成树协议 (STP)、以太通道和多链路透明互连 (TRILL)）对 FCoE 帧的转发没有影响。相反，FCoE 转发是 FCoE 交换机中称为 FCoE 转发器 (FCF) 的功能的职责。FCF 根据封装的 FC 帧的 FC Destination_ID (D_ID) 字段以及由 Fabric 最短路径优先 (FSPF) 协议创建的路由表转发 FCoE 帧。

现实世界中的 FCoE

与大多数技术一样,人们喜欢先试用一下,然后再将其部署到支持其最重要的应用程序和数据的生产环境中。此外,随着 IT 预算越来越紧张,大多数组织根本没有足够的资金来拆除和更换现有的以太网和 FC 网络,换成崭新的 FCoE 网络 即使它们能带来长期成本效益。因此,现实世界中 FCoE 的采用充其量只是不温不火。

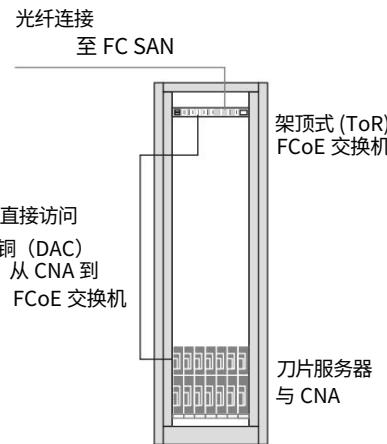
那些部署了 FCoE 的人通常采用以下两种方式之一或两种方式进行部署:

- 作为预先打包的融合计算网络和存储包的一部分
- 在接入层 (主机和接入层交换机之间的第一跳,例如
机架顶部)

预包装产品包括所谓的融合块或融合 pod,例如 VCE Vblock 或 NetApp FlexPod。在这些包装中,您将整个计算、网络和存储堆栈作为单个包装产品购买。这些块往往随具有 CNA 并从 FCoE SAN 存储启动的服务器一起提供。您通常没有选择如何配置它们,并且被迫在服务器和 FCoE 交换机中部署 CNA,这是从服务器到机架顶部交换机的第一跳 (并不是说这种预先编写的设计总是一件坏事)。

人们部署 FCoE 的另一个主要领域是网络的接入层。这意味着进入网络的第一跳,位于服务器中的网络适配器和通常位于安装服务器的同一机架顶部的接入层网络交换机之间。如图 14.9 所示。

图 14.9 网络接入层的 FCoE



网络接入层是部署任何新技术的自然且良好的起点。首先,在网络接入层部署新技术的风险远低于将其直接置于核心层。在接入层部署可以让您

缓慢而系统地部署 FCoE（一次一个机架），并在开始在网络内更深入地部署之前对该技术获得信心。

因此，除了 NetApp FlexPod 和 VCE Vblock 等，大多数 FCoE 部署都倾向于在网络接入层部署，即使用 CNA 而不是 NIC 和 HBA 部署新服务器。这减少了与网络相关的布线和机架中的网络交换机数量。在服务器中部署 CNA 还可以让您的服务器应对未来的任何 FCoE 部署，因为 CNA 现在就可以开始使用 10G 以太网，但如果您在未来部署 FCoE，您也可以处理 FCoE。

在现实世界中部署 FCoE 时需要考虑的另一个问题与谁拥有融合基础设施有关。谁拥有 CNA，而且通常更重要的是，谁拥有 FCoE 交换机？在做出此决定时，应考虑以下因素：

- 谁决定购买哪种 FCoE 交换机 供应商和型号？
- 谁支付交换机和支持的费用？
- 谁决定何时部署什么版本的固件？
- 谁拥有交换机的根访问权限？

这些并不总是容易的决定。在大多数组织中，趋势似乎是服务器团队拥有 CNA，网络团队拥有 FCoE 交换机，存储人员就 CNA 和 FCoE 的配置以及 FCoE 交换机内的 FC 设置提供专家建议。

尽管 FCoE 在某种程度上与服务器虚拟化互补且类似，但它需要指出的是，FCoE 的普及程度与服务器虚拟化的普及程度完全不同。这在很大程度上是因为 FCoE 带来的灵活性和敏捷性远不及服务器虚拟化。事实上，软件定义网络 (SDN) 似乎更有可能带来与服务器虚拟化技术类似的革命性灵活性和新模式。

概括

本章介绍了全球大多数数据中心采用的两种主流网络技术：以太网和光纤通道。您了解了它们之间的差异以及将它们组合成单个数据中心网络的潜在优势。

本章展示了以太网如何成为自然选择，因为它被广泛采用，并且具有更高的多功能性和灵活性。本章还概述了以太网所缺少的功能，以便使其成为可行的存储流量传输方式，然后详细介绍了新版以太网如何实施重大更改。

称为数据中心桥接 (DCB) 或融合增强型以太网 (CEE)。这使得它适合传输存储流量。然后，您了解了 FC 如何在 DCB/CEE 之上分层，并在本章结束时了解了它如何影响未来的网络设计以及它如何在当今现实世界中实施。

章节概要

数据中心桥接 数据中心桥接 (DCB) 是电气和电子工程师协会 (IEEE) 802.1 工作组下属的一个任务组。该任务组负责对 802.1 网络进行增强,以使其成为未来数据中心网络的自然选择。DCB 任务组负责的增强型以太网版本有时称为 DCB 以太网。

融合增强型以太网 融合增强型以太网 (CEE) 是一种新的以太网增强版本,适合传输存储流量以及传统 IP 流量。

这使其成为数据中心网络整合的自然选择。

无损网络 无损网络在发生拥塞时不会丢帧。对于希望传输 FC 流量的任何网络来说,这都是一项关键网络功能。

以太网光纤通道 以太网光纤通道 (FCoE) 是将光纤通道帧封装在以太网帧内,允许光纤通道流量通过融合增强型以太网网络传输。FCoE 是实现数据中心网络整合的重要技术。

章节 15



云储存

本章涵盖的主题：

- ✓ 定义云
- ✓ 云计算模型
- ✓ 公共云
- ✓ 私有云
- ✓ 混合云
- ✓ 云存储
- ✓ 最终一致性
- ✓ 数据持久性



本章深入探讨日益流行和重要的云主题。本章首先从高层次定义云，然后详细介绍。由于云计算的概念可以代表许多人的思维方式发生重大变化，因此您将亲身体验一些最流行的云服务，包括启动虚拟机 (VM) 以及公共云中的对象存储。

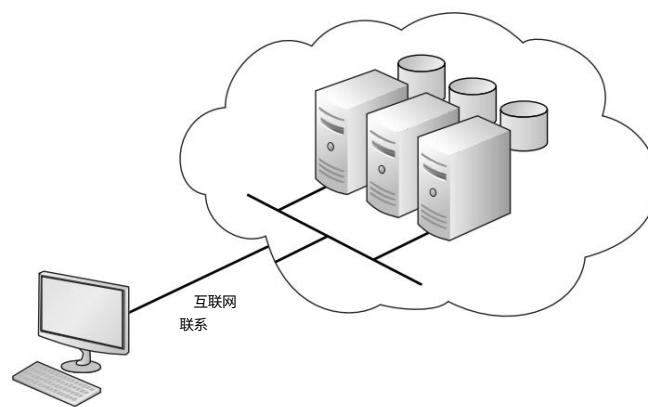
本章介绍了三种主要的云模型：公共云、私有云和混合云。了解云计算后，您将重点了解云存储的重要概念和理论，以及在云中创建存储和上传对象时的一些动手活动。您将探索公共云存储、私有云存储和混合云存储以及每种存储的一些潜在用例。

云概述

对于云到底是什么，实际上并没有任何正式的定义，但不要因此而阻碍你。云是真实存在的，而且未来会越来越大。幸运的是，它非常简单。

云，或云计算，是一种计算模式。该模式的主要特点是 IT 资源通过互联网远程访问，而不是通过局域网 (LAN) 本地访问。如图 15.1 所示。

图 15.1 高级云计算模型



从很多方面来看,云计算是一种极端的虚拟化形式,作为消费者,您无需关心底层硬件或软件。您不再需要考虑硬件维护、硬件升级、技术更新、安全补丁、防病毒、支持人员或任何拥有该技术时必须考虑的典型事项。您实际上只是消费服务,而不必关心底层规格。再简单不过了。云计算的一个主要目标正是:让 IT 资源的消耗尽可能简单。



The “云”这个术语可能来自中世纪的云图像,用在 the network infrastructure diagram 中表示互联网将互联网表示为云。其原因是,在大多数情况下,您不需要了解互联网和互联网路由的复杂性;您只需要知道它是一个可靠的传输数据的网络。云计算也是如此:您支付并接收服务,而无需关心细节。您只需理解它是否有效。

“Cloud”这个术语可能来自中世纪的云图像,用在 the network infrastructure diagram 中表示互联网将互联网表示为云。其原因是,在大多数情况下,您不需要了解互联网和互联网路由的复杂性;您只需要知道它是一个可靠的传输数据的网络。云计算也是如此:您支付并接收服务,而无需关心细节。您只需理解它是否有效。

提供云服务的公司通常被称为服务提供商或云提供商。他们拥有服务背后的技术,并负责确保一切服务都能无缝地为您服务。他们也是您通常按月支付使用其服务的费用的人。出于这些原因,我们通常将云解决方案称为完全托管服务。一些重要的公共云提供商包括亚马逊、微软、谷歌和 Rackspace 等公司。They’re also the people that you pay, usually on a monthly basis, for the privilege of using their services. For these reasons, we often refer to cloud solutions as *fully managed services*. Some of the major public cloud providers include companies such as Amazon, Microsoft, Google, and Rackspace.

云计算带来的最大优势是成本。云计算服务几乎总是设置为按月付费,具体取决于当月的用量。这种运营支出 (op-ex) 模式意味着进入云计算的财务之路变得轻松。*that you have consumed that month. This operating expenditure (op-ex) model means that the financial road into the cloud is made easy.* 无需像非云模式那样预先投入大量资本支出 (cap-ex)。在非云模式中,如果您想要 10 台服务器,每台服务器有 64 GB RAM 总共有 64 TB 的存储空间,则必须预先支付现金购买这些服务器 (cap-ex), with 64 GB RAM each and a total of 6 TB of storage, you have to lay down the cash to buy these up front (cap-ex)。云计算则不是这样。在云模型中,你可以启动 10 台虚拟服务器,并为其配备所需的 RAM 和存储空间。通过提供商的自助服务工具,然后开始按月付费。一两个礼拜后,如果你意识到你真的只需要其中 5 台服务器,你可以关闭不需要的 5 台服务器并立即停止付费。然后,在以后,如果你需要另外 20 台服务器,你可以启动另外 20 台服务器并让它们立即可用。同样,没有前期资本支出成本,服务器交付也不需要很长的准备时间。When you don’t want and immediately stop paying for them. Then, at a later date, if you need another 20 servers, you can spin up another 20 servers and have them instantly available. Again, there is no up-front cap-ex cost and no long lead-time on delivery of the servers.

这种随时增加或减少消耗的能力被称为弹性计算。您可以根据需要使用尽可能多的云资源,也可以每天进行更改。第一天没有必要过度配置,以防将来可能需要更多容量,但最终却发现未来的需求永远无法实现。You can change it on a daily basis. There is no need to over-spec on day one, just in case you might need more capacity in the future, only to find that that future requirement never materializes. 云模型与更传统的 IT 方法完全不同,它更加灵活。From more-traditional approaches to IT. It is way more flexible.

云的最大优势是服务往往是自助式的。云用户可以通过一个简单的 Web 网站订购自己的配置,在完成配置后,order their own configurations via a simple web portal, and after they’ve completed their

订购后,服务已配置完毕,几乎始终可立即使用。例如,云服务可能会提供一个 Web 门户,用于订购符合您确切规格的虚拟机(VM)。在门户上,您需要为您的 VM 输入以下所有内容:

- 姓名
- CPU 数量
- 内存容量
- 网卡数量和速度
- 磁盘数量和大小

从自助服务门户订购虚拟机后,单击“创建”,虚拟机便会创建并立即提供给您。这很像为您的 Amazon Kindle 设备购买一本新书。您选择所需的书籍,将其添加到购物车,继续付款,然后书籍便会立即发送到您的设备。

现在我们简单谈谈三大云模型:

- 公共
- 私人
- 混合型

公有云

大多数人在考虑云时都会想到公共云。从很多方面来看,它是最纯粹的云形式。

公共云通常通过互联网访问,并提供最高级别的抽象。我的意思是,它是不透明的:客户无法看到支持服务的技术、服务器硬件以及所涉及的网络和存储技术。

关于公共云,我们都需要理解和接受的一件事是,它多租户,意味着所有底层基础设施由多个客户共享。您绝对不会拥有自己的物理服务器或物理存储阵列!如果您将服务或数据放在公共云中,它将与数千个其他客户的服务和数据位于同一个共享基础设施上。这是真正的多租户!

此外,作为公共云客户,您几乎不知道您的数据实际存放在哪里。当然,您可能知道数据存放在哪个地区,例如北美、欧盟(EU)、亚太地区(APAC)等等。但您不会知道确切的位置,而且您肯定无法出现在云提供商数据中心的前门并进入其中查看。

所有这些对于提供我们期望云服务提供的许多功能都很重要,例如以下内容:

- 成本低
- 弹性

■高度可扩展

■自助服务

■可访问性

然而,公共云服务的超级共享性质对于某些用户来说可能是一个挑战

案例和组织。您可能为一家拥有严格安全政策的组织工作,该政策规定服务是否可以托管在公共云中。在这些情况下,您可能不被允许使用公共云服务。

托管在公有云中的服务可以从世界上任何有互联网连接的地方访问,这使得基于公有云的服务成为世界上最容易访问的 IT 服务之一。只要您有互联网连接,就可以随时从世界上任何地方使用任何设备访问它们。

这引出了公共云的一个关键点:安全性!使用公共云服务时,您需要具备安全性意识。确保加密数据,并确保由您(而不是云提供商)持有加密密钥。并确保只将您希望公开查看的任何数据标记为“公开”。

最受欢迎的公共云计算服务可能是 Amazon Elastic Compute 云 (EC2),是 Amazon Web Services (AWS) 的一部分。其他流行的公共云计算服务包括 Windows Azure 和 Rackspace Cloud Servers。

您需要在云中执行的第一件基本事情之一就是创建虚拟机。

练习 15.1 引导您了解如何使用 Windows Azure 创建 Windows 虚拟机,练习 15.2 向您展示如何使用 AWS 创建 Linux 虚拟机。

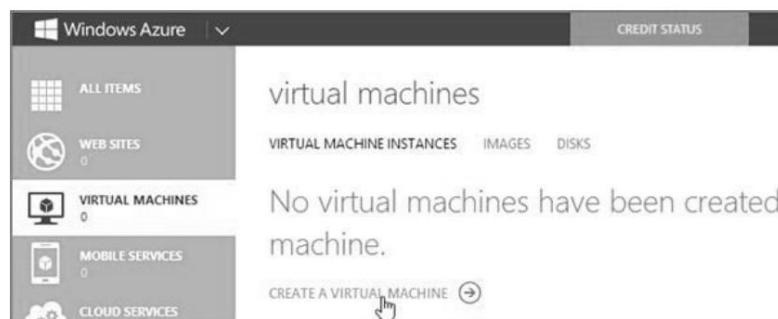
练习 15.1

使用 Windows Azure 在云中创建新的虚拟机

在本练习中,您将逐步完成在 Windows Azure 公有云中创建 Windows VM 并登录的过程。本练习假设您拥有 Windows Azure 帐户。

1.在 Windows Azure Web 用户界面中,选择虚拟机,然后单击

创建虚拟机链接。



2. 出现创建新虚拟机的属性表, 要求您提供

您的虚拟机的以下信息:

■ DNS 名称

■ 操作系统 (简称镜像)

■ VM 的大小 (指定 CPU 数量和内存量)

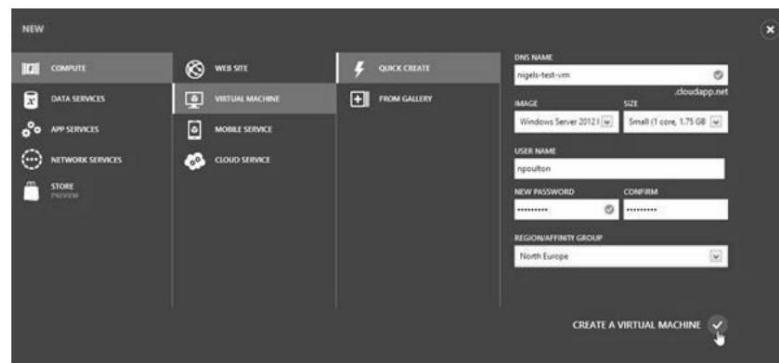
■ 用户名

■ 密码

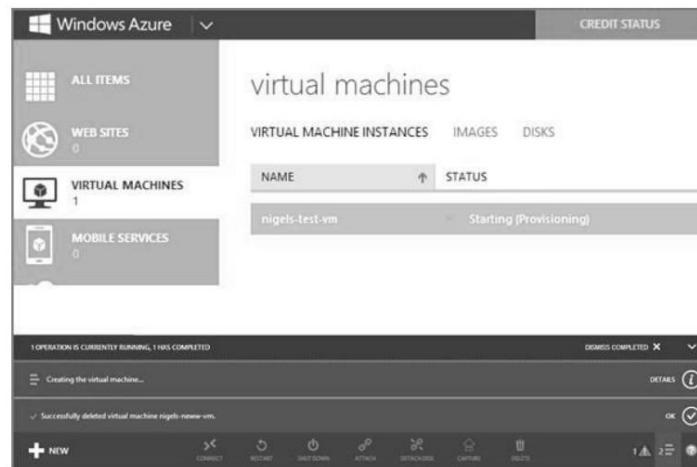
■ 存储虚拟机的区域

输入创建新 VM 所需的信息, 然后单击“创建虚拟机”旁边的复选标记。我正在为名为nigels-test-vm的新 VM 选择配置, 它将在一个小型 VM 映像上运行 Windows Server 2012, 该映像具有单核和 1.75 GB RAM, 用户名和密码以及北欧地区。

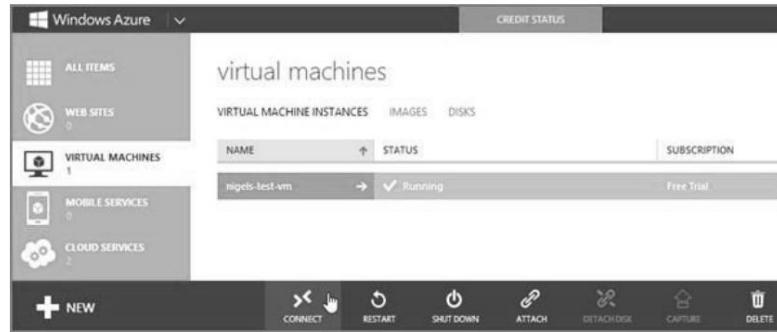
您应该使用适合您场景的任何设置。



单击“创建虚拟机”后, Windows Azure 将创建您的新 VM。该过程只需一两分钟。



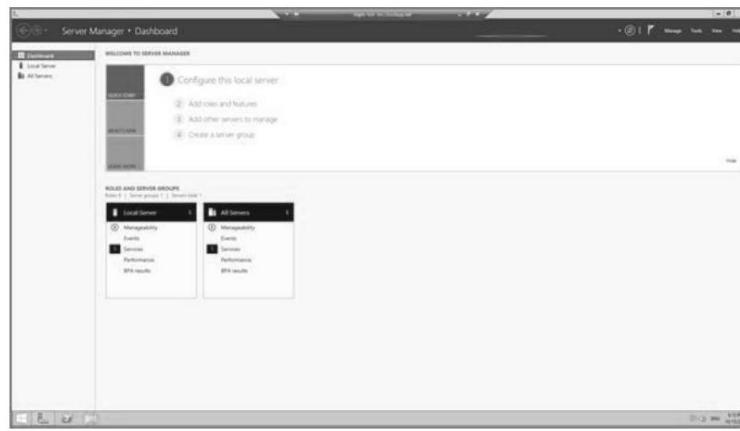
3.现在您的 VM 已创建并正在运行,您可以登录。要登录 Windows Azure VM,请在可用 VM 列表中突出显示 VM,然后单击“连接”。



Windows Azure 门户现在下载远程桌面协议 (RDP) 应用程序,该应用程序将允许您登录到新创建的 VM。打开 RDP 应用程序会弹出“远程桌面连接”对话框。



4.在远程桌面连接对话框中单击连接按钮。然后输入
在出现的 Windows 安全对话框中输入您的用户名和密码。这样您就可以登录到新的 VM。



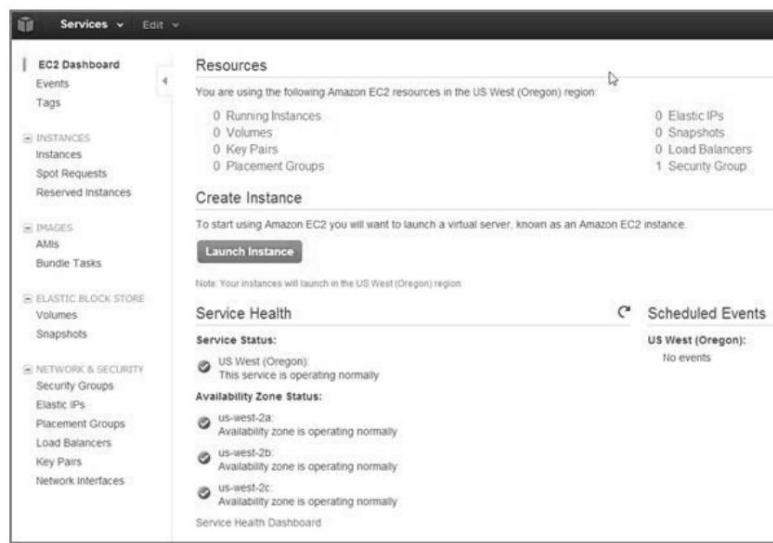
练习 15.2

使用 Amazon Web Services 在云中创建新的虚拟机

在本练习中,您将逐步了解在 AWS 的 EC2 组件中创建新 Linux VM 的过程。本练习假设您有一个 Amazon Web Services 帐户。

1. 登录 AWS 管理控制台后,首先要导航至云仪表板中的 EC2 虚拟服务器。

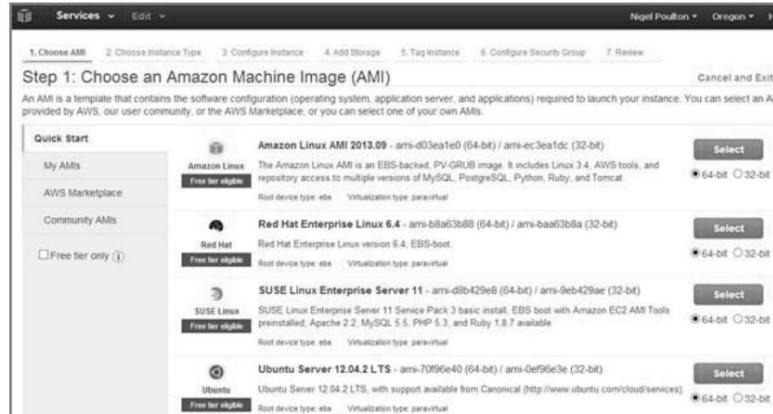
2. 在 EC2 仪表板中,您需要创建一个新的 VM。在 EC2 中,VM 称为实例。在 EC2 仪表板 Web 用户界面中,选择启动实例选项。



3. 从出现的屏幕中,选择您要创建的虚拟机类型。选项

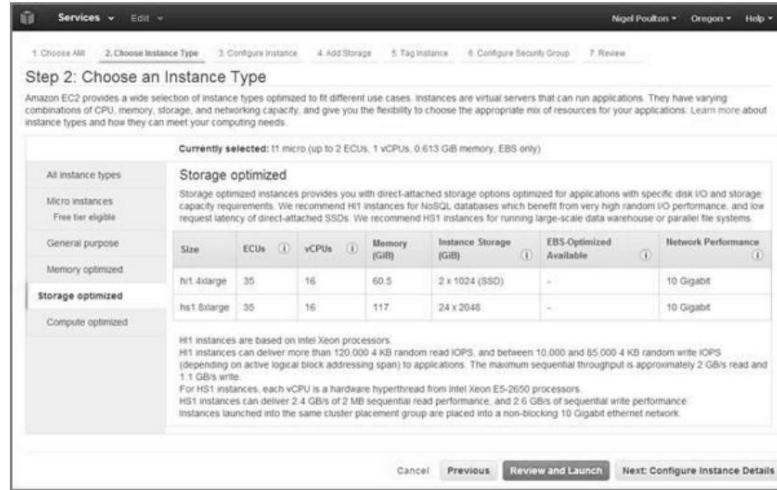
包括大多数流行的操作系统以及 32 位或 64 位的选择。

以下屏幕截图显示了一些选项。根据您的操作系统详细信息选择您的实例类型。

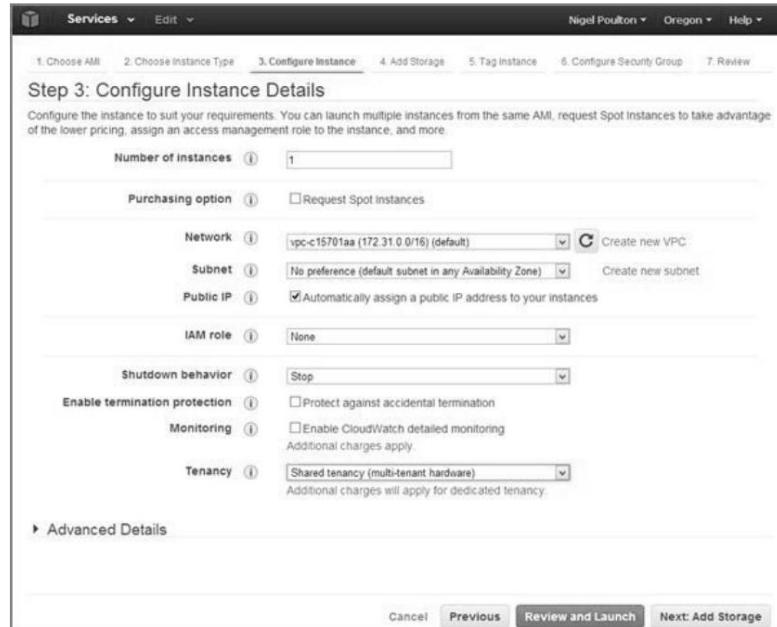


4.现在系统会提示您输入与虚拟机相关的详细信息。EC2 为您提供

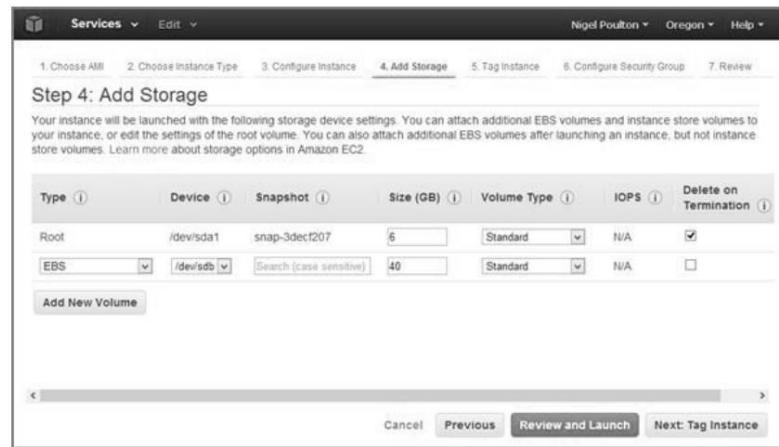
针对某些特性进行了优化的几个选项,例如内存优化、计算优化和存储优化。选择要创建的 VM 类型,然后单击下一步:配置实例详细信息。以下屏幕截图显示了存储优化的 VM 配置。



5.点击剩余的屏幕来配置您的新虚拟机。这包括配置网络和关机详细信息,其中一些显示在此处。



6. 您还可以向您的虚拟机添加存储设备。



7. 提供所有信息后,单击“审核并启动”,然后会创建您的新虚拟机。

私有云

私有云可以采用以下两种设计之一:

- 您自己的内部云,拥有您仍拥有和管理的本地设备,但用于向您的内部客户提供类似云的服务。
- 由第三方托管的服务,在本地之外,但为客户提供他们自己的专用基础设施(无多租户)。



The term *on-premise* refers to hosting IT equipment in your own data center facilities (*on-premises*). The term *off-premise* refers to hosting IT

在第一种模式中,您拥有本地设备并提供类似云的服务
对于您的内部客户,这些服务包括允许用户自行配置的 Web 门户以及每月分项计费服务。这是大型组织中的常见方法,其中内部 IT 部门面临来自公共云的压力,需要在内部提供类似云的服务,以阻止用户选择公共云服务而不是内部 IT 部门提供的服务。

在第二种模式中,第三方为客户提供专用的场外设备
这提供了云服务的一些优势,而没有一些风险

人们对公共云的看法。例如,一个组织可能不想再拥有和管理自己的所有技术基础设施,但也可能不愿意将其所有数据和服务放在大规模共享平台上的公共云上。

私有云选项可能是一个不错的解决方案,即另一家公司负责所有基础设施管理,但提供专用平台。这些选项通常允许客户访问数据中心设施并查看他们付费使用的设备。

混合云

混合云,顾名思义,是公有云和私有云的组合。许多组织并不热衷于在第一天就将所有数据和服务都投入到公有云中。相反,他们更愿意先将一些不太重要的应用和服务放入公有云中,同时将关键任务和敏感应用保留在内部。这种公有云和私有云的混合有时被称为混合云。

混合云还允许客户在业务需要额外资源时进入云端计算资源。例如,假设您的企业在一年中的某个特定时间需要更多的计算能力。您每年举办一次活动,而您的网站访问量会大幅增加。利用公共云,您可以每年增加一次计算能力,然后在不再需要时将其关闭(并停止付费)。

存储和云

您已经了解了云计算。现在让我们更具体地了解一下云存储。

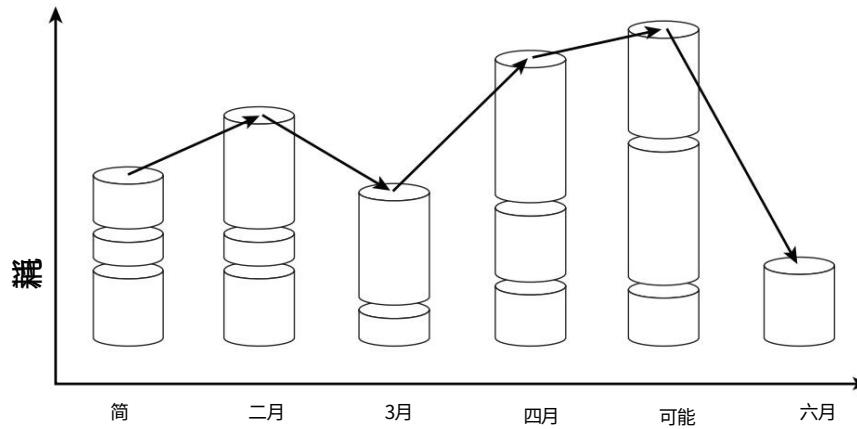


存储即服务有时写为 SaaS, is sometimes written as *SaaS*.

云存储可通过互联网远程访问。您可以使用账户每月支付费用。
您只需为您所使用的服务付费的消费模式。

弹性存储和基于消费的计费的组合如图 15.2 所示。

图 15.2 弹性和消费计费



云存储具有云计算的大部分功能和优势,其中包括:

- 简单
- 弹性
- 高度可扩展
- 可访问性
- 快速自助服务配置
- 内置保护功能
- 无资本支出成本

云存储并不适合所有用例。云存储服务有一些需要考虑的潜在缺点。这些缺点包括与非基于云的存储解决方案相比性能较低。例如,通过互联网连接访问存储永远不会像通过位于您自己的数据中心网络内靠近您的服务器的短距离、低延迟光纤通道 SAN 访问存储那样快。

同样,您也不会希望将任何对延迟敏感的应用程序放在公共云存储上。

所以这意味着云存储上没有基于 SQL 的关系数据库。

大多数云存储解决方案都是基于对象存储的。虽然第 7 章“文件、NAS 和对象”详细介绍了对象存储(对象存储),但我还是要提醒您,对象存储不是为高性能或结构化数据(数据库)而设计的,并且不适合经常变化的数据。但是,它非常适合流式/顺序读取工作负载,因此非常适合存储和检索富媒体和其他 Web 2.0 类型的内容,例如照片图像、视频、音频和其他文档。因此,许多流行的社交媒体网站都使用对象存储系统来托管其上传的内容(例如图像和视频),因为这些内容上传一次(写入一次),播放(读取)多次,并且上传后几乎从未更改过。



Makes sure the vendor knows whether you are using a base unit of measurement for provider charges when based on MB or 1024 KB. Although the numbers may differ slightly depending on the provider, 1 MB is 1,000 KB, so the difference may be significant.

数据持久性

数据持久性概念是云存储的关键。数据持久性是指存储在云中的对象在云中发生故障时仍能生存的能力。您可以将其视为云的 RAID。



The concept of durability is borrowed from the database world, where the D in the well-known ACID acronym stands for durability. ACID,

在撰写本文时,Amazon S3 声称其耐用性达到 99.99999999% 可靠性。亚马逊还表示,这种级别的耐用性预计每年平均丢失对象率为 0.000000001%。基本上,丢失存储在 S3 中的对象是非常不幸的,或者正如亚马逊所说,如果您在 Amazon S3 上存储了 10,000 个对象,则平均每 10,000,000 年可能会丢失一个对象。

显然,这是指由于 S3 内部故障而丢失对象的可能性,并且它不包括因用户错误(例如意外删除对象)造成的损失。

这种极高的数据持久性是通过云存储服务在整个云中存储每个对象的多个副本实现的。例如,如果您将数据保存在所选云的美国区域内,则数据的多个副本将保存在美国区域内的多个数据中心。这意味着,如果云的美国区域内的一个节点甚至整个数据中心发生故障,您的数据仍可从同一区域内的其他数据中心获取。

许多云存储解决方案还提供地理复制,其中对象也会复制到区域之外,即全球范围内。在地理复制场景中,您的数据会受到区域内多个设施以及多个区域之间保护。这甚至提供了更高的保护。

数据持久性主题中的一个重要概念是原子更新 记住,ACID 中的 A 代表原子性!原子更新可确保在云确认更新成功之前,对云存储的所有更新(新对象或对现有对象的更新)都存储到云中的多个位置。因此,您知道,一旦某个对象或对某个对象的更新被信号表示已成功提交到云,该对象就已存储到多个位置。



尽管云存储服务宣传其数据持久性非常出色,但您仍然需要备份数据。数据持久性本身并不能保护您免受意外删除或无意更新对象的影响。如果云存储解决方案提供版本控制(即每次更新或删除对象时都会保留对象的先前版本),那么这可能允许您停止备份存储在云中的数据。但是,在做出像更改备份实践这样重要的决定之前,您需要进行彻底的研究。

最终一致性模型

云存储平台的某些方面基于最终一致性模型。在此模型中,对象和对象更新需要时间才能在整个云中传播。正如您在数据持久性部分中了解到的那样,存储云中的对象会被复制或传播到云中的多个节点和位置,以防止云中发生故障。有时,对象和对象更新的传播是异步发生的(例如传播到区域外),这意味着在更改复制到整个云之前,您对云的更新会发出完成信号。

例如,在配置了地理复制的云存储解决方案中,对象可能在上午 10 点在美国加利福尼亚州更新,而某人在上午 10:01 (加利福尼亚时间)从亚太地区的副本访问该对象时,仍然可以看到该对象的先前版本,而更新的对象传播到亚太地区时,时间是上午 10:02。话虽如此,原子更新仍然受到尊重,这意味着如果您的云存储解决方案配置为在区域内制作三个副本,则这三个副本将在您的对象的任何更新被信号化为完成之前更新;只是区域外的传播应该是异步的,可能需要一段时间才能传播。只要您离开系统足够长的时间而不进行更改,云中数据的所有副本最终都会保持一致。显然,如果快速的全局一致性对您的组织很重要,那么基于云的对象存储可能不是最佳解决方案。



最终一致性模型(有时称为松散一致性)的一些示例包括Amazon S3、数据库和一些社交媒体和照片共享网站。在这些解决方案中,更改和更新只需很短的时间即可在整个系统中传播。

公共云存储

大多数人在考虑云存储时,都会想到公共云存储,因为它是最流行的云存储形式。Amazon S3、Windows Azure 和 Rackspace Cloud Files 等服务都是公共云存储服务的例子。

从高层次上讲,公共云存储是通过互联网访问的存储,通常通过 RESTful 应用程序编程接口 (API) 进行存储,并根据您当月的用量按月付费。公共云存储本质上也是多租户的,这意味着它有多个客户共享相同的通用基础设施和存储平台。

为了帮助你理解它的工作原理,练习 15.3 向你展示了如何配置一些 Amazon S3 中的云存储。

练习 15.3

开始使用 Amazon S3 公共云存储

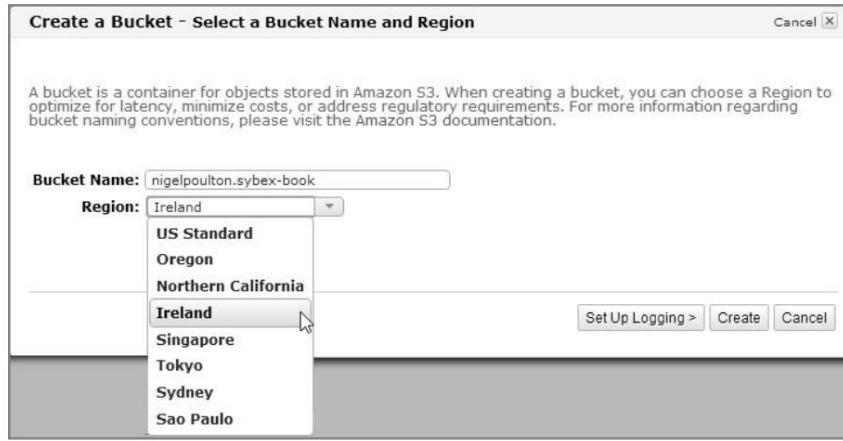
在本练习中,您将逐步了解启动和运行 Amazon Simple Storage Service (S3) 所需的步骤。在开始使用 Amazon S3 之前,您需要一个 Amazon Web Services 账户。在本示例中,我们假设您已经有一个账户。现在,您将创建一个云存储桶、上传一个对象并配置自定义元数据。

- 1.从 AWS Web 控制台中导航到 S3 服务。

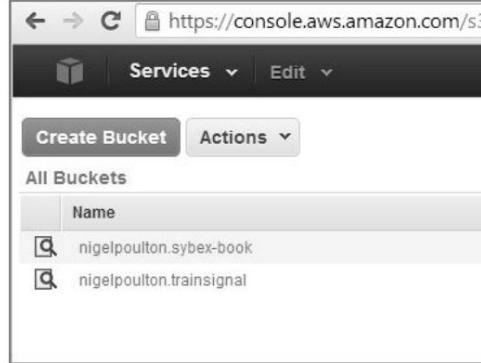


- 2.由于 Amazon S3 中的所有对象都存储在存储桶中,因此让我们创建一个存储桶。单击创建在 S3 网络界面中点击 Bucket 按钮。

- 3.在随后的“创建存储桶”屏幕中,为您的存储桶选择一个名称,并选择您想要存储存储桶的区域。在此示例中,我创建了一个名为 nigelpoulton.sybex-book 的存储桶,并选择将其存储在爱尔兰地区。

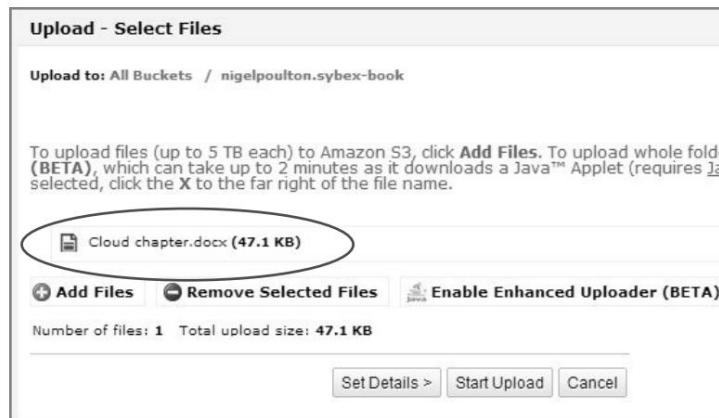


4.单击创建。新创建的存储桶将显示在存储桶列表中。



5.现在让我们将一个对象（文件）添加到存储桶。选择您的存储桶并单击上传。

6.在“上传”屏幕上,单击“添加文件”,然后从对话框中选择要上传到存储桶的任何文件。如下图所示,我选择上传本章的文件,当前名为Cloud chapter.docx。



7.选择文件后,单击“开始上传”。下图显示我的
云章节.docx对象已上传至mynigelpoulton.sybex-book
桶。

The screenshot shows the AWS S3 console interface. At the top, the URL is https://console.aws.amazon.com/s3/home?region=us-west-2. Below the header, there are buttons for 'Upload', 'Create Folder', and 'Actions'. The main area displays 'All Buckets / nigelpoulton.sybex-book'. A table lists one object: 'Cloud chapter.docx' with a size of 47.1 KB, last modified on Wed Oct 09 21:13:06 GMT+100 2013, and a storage class of Standard.

8.右键单击您上传的对象。您将看到一个上下文菜单,其中列出了您可以对该对象执行的所有常用选项。

The screenshot shows the AWS S3 console with a context menu open over the 'Cloud chapter.docx' file. The menu options are: Open, Download (with a cursor icon), Make Public, Rename, Delete, Initiate Restore, Cut, Copy, and Properties.

9.从右键菜单中,选择“属性”。将打开一个页面,您可以在其中设置 com-
与存储在 Amazon S3 云中的对象关联的 mon 属性,其中包括:

- 权限
- 加密设置
- 冗余
- 自定义元数据

▼ Permissions

Grantee: nigelpoulton	<input checked="" type="checkbox"/> Open/Download	<input checked="" type="checkbox"/> View Permissions	<input checked="" type="checkbox"/> Edit Permissions	x
Grantee: []	<input type="checkbox"/> Open/Download	<input type="checkbox"/> View Permissions	<input type="checkbox"/> Edit Permissions	x

+ Add more permissions

▼ Details

Storage Class : Standard Reduced Redundancy

Server Side Encryption: None AES-256

▼ Metadata

Key: Content-Type	Value: application/msword	x
Key: project	Value: sybex	x
Key: topic	Value: cloud	x

+ Add more metadata **- Remove selected metadata**

Save **Cancel**

就这样。您刚刚创建了一个云存储桶,上传了一个对象,并配置了自定义元数据。

API 访问

从练习 15.3 中可以看出,可以通过 Web 界面访问 Amazon S3 存储。

上传对象、下载对象、删除对象、管理对象访问等操作都可以通过 Web 浏览器完成。您不需要像使用 SMB/CIFS 共享或 NFS 导出那样挂载云存储。事实上,您无法这样做。您也不需要像使用 SCSI LUN 那样使用文件系统对其进行格式化。同样,您无法这样做。相反,您可以直接通过 Web 浏览器或使用 REST API 直接与其通信的应用程序来访问它。

如前所述,API 是应用程序编程接口的首字母缩写词,API 有点像一组用于访问和使用基于 Web 的服务的已发布指令。

在公有云环境中,API 几乎总是 RESTful。我所说的 RESTful 指的是使用简单的 put/get 样式语义(例如常见的 HTTP PUT 和 GET)的系统。

在万维网上广泛使用的动词。大多数基于 REST 的系统使用 HTTP(尽管它们不必如此),并且大多数(尽管不是全部)对象存储系统都是通过 RESTful API 访问的。

如果您不熟悉 API,那么可以将它们视为完成某项任务时应遵循的一组规则和协议。例如,在五星级餐厅点餐的协议与在麦当劳点餐的协议截然不同。在五星级餐厅,服务员会恭敬地将您带到座位并给您一份菜单,然后您悠闲地浏览菜单,等待服务员过来为您点饮料,然后等饮料送来后再点餐。您通常会保留账单,并在用餐后结账。了解这一点有助于让您享受用餐体验。但在麦当劳餐厅尝试同样的方法,您会发现自己非常饥饿。没有人等着为您安排座位,也没有服务员来到您的餐桌为您点餐。在付款之前,您甚至不要费心尝试吃您的食物!API 并没有太大的不同。Amazon S3 API 为您提供访问 S3 云中数据的说明。遵循 S3 API,您的 S3 体验可能会很好。不遵循 S3 API,您的 S3 体验就不会那么好。

API 也是软件到软件的接口,而不是用户界面。它们旨在可用于以编程方式或系统方式访问系统,无需用户协助。关于这一点,许多 API 都与许多流行的编程语言(如 Python、Java、PHP、C# 等)进行了语言绑定。这些语言绑定比纯 REST 更高级别(更简单),允许您在编程语言的本机上下文中执行 RESTful 操作,从而使编程 API 变得简单得多。

公有云存储性能

就性能而言,云存储接近垫底,这是有道理的。它与本地系统或应用程序距离较远,这意味着云存储的表现不会很出色。存储性能几乎总是随着距离 CPU 越远而变慢。例如,主内存(RAM)靠近 CPU,速度极快。本地连接磁盘距离 CPU 稍远(您通过系统主板上的 PCI 总线访问它),速度较慢。网络附加存储(SAN 和 NAS)距离更远,速度也更慢。而云存储是所有存储中距离最远的!因此,毫不奇怪,它是所有存储中最慢的。

值得注意的是,距离 CPU 越近,可用容量越小,而距离 CPU 越远,可用容量越大。例如,RAM 比硬盘更靠近系统上的 CPU,而硬盘又比 NAS 阵列更靠近。由此我们应该知道,RAM 容量通常比硬盘容量小很多,而硬盘容量通常比 NAS 阵列提供的容量小很多。云存储也符合这一规则,因为它是当今最具可扩展性的存储形式。

虽然我们谈论的是与 CPU 的距离,但事实似乎也如此:存储距离 CPU 越远,其可访问性和可共享性就越高。例如,与共享本地连接的硬盘上的空间相比,共享主内存(RAM)相对困难。同样,NAS 阵列在设计时也考虑到了可共享性,云存储也是如此。云存储提供了极致的存储可访问性和可共享性。

尽管如此,云存储还是有其用例的。在某些用例中,例如访问照片和视频文件等大型对象,云存储可以表现出色。这是因为访问这些类型的对象需要良好的吞吐量而不是低延迟;单个 HTTP GET 可以传输大量数据。而且互联网带宽越来越便宜。

原子上传

所有好的公共云存储解决方案都是基于原子上传的原理运作的。

这是一件好事。原子上传可确保在上传操作 100% 完成之前,无法从云中访问对象。该原子上传过程通常涉及在发出上传完成信号之前制作对象的多个副本。这意味着云中的故障绝不会导致您丢失对象。原子上传包括任何相关的元数据和安全属性。

地理复制

公共云服务通常提供全球复制功能,有时也称为地理复制。在这种情况下,对象可以在云的全球基础设施中复制。

这显然可以提高数据持久性,即数据抵御大规模局部灾难和故障的能力。

地理复制的另一个主要用途是本地访问数据。例如,如果您的公共云提供商在美国、英国和新加坡设有数据中心,而您在美国、英国和澳大利亚设有办事处,那么您可能希望将对象复制到云提供商的所有设施,以便您在世界各地的每个地点的员工都可以访问数据,而不必总是返回美国。但是,出于法律或监管目的,您可能不允许将数据离开某个区域或法律管辖区。例如,欧盟 (EU) 有法律禁止某些数据离开欧盟。此外,您可能担心政府访问存储在各个国家/地区的数据。

内容分发网络

关于本地访问数据的话题,您的数据离用户越近,性能就越好,获得良好用户体验的机会就越大。例如,如果您的公共云提供商在美国、英国和新加坡设有数据中心,而您在香港有用户,那么您可能希望香港的用户能够通过新加坡云提供商的设施访问数据。但是,从香港到新加坡仍然很远。这时,内容分发网络 (CDN) 就可以派上用场了。CDN 的工作原理是将对象缓存在离用户更近的地方。它们能够通过在全球各地的边缘位置设置缓存节点来实现这一点。

例如,与 Rackspace 合作的 CDN 提供商 Akamai Technologies 可以

在香港缓存对象,这样香港的用户就可以在本地访问数据,而不必每次想要访问某个对象时都要在互联网上来回穿越到新加坡。



有时边缘位置被称作边缘点 *edge points* are referred to as *points of presence (POP)*.

因此,CDN 采用了从世界任何地方的任何设备访问数据的理念
云存储更进一步,使其可以从世界任何地方的任何设备快速访问您的数据!



有时您可能会看到 CDN 被称为内容分发网络,而不是内容交付网络或 content delivery network, rather than a content delivery network. Both terms mean the same thing.

一些云解决方案拥有自己的 CDN,例如 Amazon S3,而其他云解决方案则与第三方 CDN 专家合作,例如 Rackspace 与 Akamai 合作推出其 Cloud Files 云存储平台。例如,Akamai (一家 CDN 提供商)在全球 200 多个边缘位置拥有超过 100,000 台服务器,可以缓存数据,使其更接近您的用户。在成本方面,您还需要检查使用 CDN 是否需要额外付费。

虽然公共云存储可能会给您和您的组织带来挑战和风险,但组织内部肯定存在这样一种趋势,即将其较低层级的需求以及较低风险的数据和服务转移给其他人,而其他人正日益成为公共云。



Real World Scenario

小型办公室使用公共云存储

许多组织开始利用公共云存储产品来满足没有 IT 人员的远程分支机构的存储需求。一种流行的趋势是将备份推向云端,因为小型办公室的数据量往往不大,因此非常适合备份到云端。例如,许多组织开始通过 AWS VTL 接口 (网关-VTL)等接口将区域办事处的备份发送到云端,而不是在远程分支机构使用物理磁带或甚至小型专用磁盘单元/VTL 来满足备份存储的需求。这样就无需在远程分支机构部署专用硬件,也无需手动处理磁带等。

私有云存储

私有云存储主要有两大特点：

- 现场
- 场外

现场和场外的一个共同特点是私人客户

云存储拥有自己的私有基础设施 自己的专用存储系统。

这与公共云存储形成了鲜明对比，在公共云存储中，成千上万的客户共享相同的底层基础设施和存储系统。

尽管在租赁方面与公共云存储截然相反，但私有云存储服务旨在模仿公共云存储的许多其他功能，包括：

- 简单
- 弹性
- 可扩展性
- 快速自助服务配置
- 内置保护/数据耐用性
- 基于运营成本的消费计费

在本地私有云存储解决方案中，您可以通过企业 LAN 访问由第三方拥有但位于您的数据中心的存储。在异地私有云存储解决方案中，存储被移动到服务提供商的数据中心，并通过 Internet 或私有 WAN 链接进行访问。

对于那些担心将公司数据从自己手中转移到大规模共享的公共平台上的公司来说，私有云存储可能是一种比公共云存储更可接受的解决方案。在公司逐渐接受公共云存储理念的过程中，私有云存储可能是一个很好的垫脚石或折中方案。

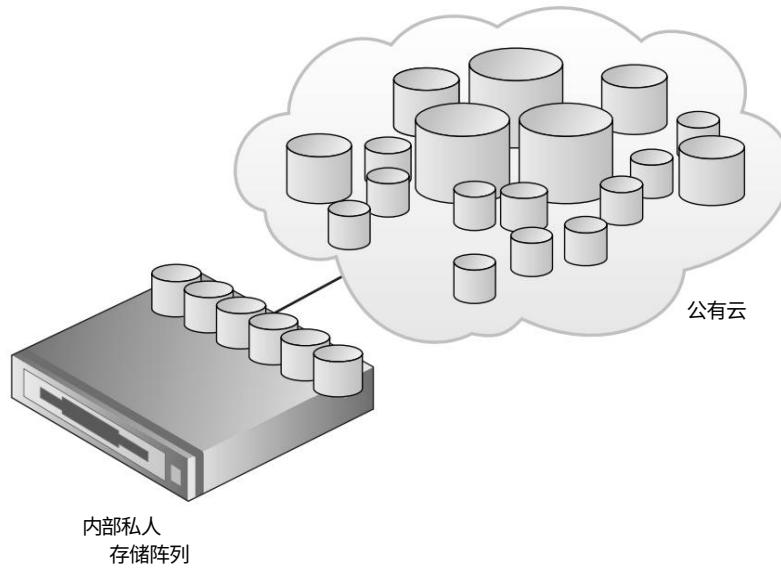
混合云存储

在混合云存储中，单个应用程序或产品使用公共云存储和私有存储的混合，从而兼具两者的优势。一些商用混合存储解决方案的示例包括 Microsoft StorSimple、Nasuni、Panzura 和 TwinStrata。例如，Microsoft StorSimple 是一个硬件和软件存储阵列，您可以购买并放置在数据中心，但可以本机使用 Windows Azure 云作为低层存储。在高层次上，StorSimple 产品包括一个内部硬件存储阵列，带有内部 SSD 和 SAS 驱动器，以及与 Windows Azure 云存储的本机集成。这可以提供

传统的现场存储系统采用 SSD 和 SAS 驱动器,可快速访问经常引用的数据,同时使 Windows Azure 成为存储不经常访问的数据的存储层。微软将 StorSimple 称为云集成存储 (CIS),并将 Windows Azure 云的使用称为层 (CaaT)。

图 15.3 展示了一个简单的混合云存储系统,例如 Microsoft StorSimple。

图 15.3 混合云存储设备



混合云备份是日益常见的混合云存储解决方案的另一个常见示例。在混合云备份解决方案中,暂存设备位于您的数据中心,用于在本地存储一到两周的备份,以便可以使用它们执行快速恢复。这些备份会在一到两周后移至云端。

计算机辅助制造研究所

您可能遇到的与云存储相关的基于标准的 API 是云数据管理计划 (CDMI)。SNIA 的这项计划旨在为云存储领域带来一些标准 用于存储、访问和管理存储在云中的对象的标准。但是,由于 S3、OpenStack Swift (云文件)和 Azure 等其他 API 的广泛流行,CDMI 实际上无关紧要。大多数为云存储编写代码的人通常都是针对 S3、Swift 和 Azure 编写代码。针对 CDMI 编写代码的人并不多。但是,越来越多的对象存储系统开始支持它。

事实标准和法律标准

有两种主要类型的标准需要注意：

■ 法定标准

■ 事实上的标准

法定标准是由标准机构和委员会（如 SNIA、IEEE 和 ISO）颁布给我们的标准。事实上的标准不受标准机构控制和颁布给我们，但在现实世界中应用非常广泛，因此可能比法定标准更重要。在云存储领域，CDMI 就是一个法定标准的例子，它由标准机构编写和监管，但并未得到广泛采用和实践。相反，S3 和 Swift 的应用非常广泛，以至于它们几乎成为每个人都遵守和实践的非官方标准。

两者各有利弊。

从架构上讲，基于 CDMI 的解决方案看起来和感觉起来都像任何其他基于云的对象存储。CDMI 将对象存储在容器中，容器的命名空间类似于 Amazon S3 中的存储桶。对象存储在容器内，容器不能嵌套。CDMI 云存储中的所有对象都分配有一个 URL，通过该 URL 可以使用常见的 HTTP 动词（例如 PUT、GET、DELETE 等）通过 REST 访问和操作它们。CDMI 还支持其他必要的云存储功能，例如安全性和身份验证。

CDMI 以客户端-服务器架构实现。CDMI 客户端向云存储服务发出 CDMI 请求。然后，云存储解决方案中的 CDMI 服务器将数据返回给 CDMI 客户端。这一切都通过 HTTP/HTTPS 进行，并且非常适合用于 Web 浏览器和 Web，因为 CDMI 服务器是建立在 HTTP 之上的。以下是客户端请求在 <http://cloud.provider.com> 访问的云平台上查看名为 sybex-book 的容器中名为 cloud-chapter.pdf 的 PDF 文件的示例：

获取/sybex-book/cloud-chapter.pdf HTTP/1.1

主机:cloud.provider.com

接受： */*

除了能够在云中创建、检索、更新和删除数据之外，CDMI 还支持端口可以发现特定云存储解决方案的功能。

概括

在本章中,您学习了云计算模型的基础知识以及三种流行的云计算形式（公共云、私有云和混合云）之间的差异。您了解了云的简单、弹性、可扩展、自助服务以及按使用量付费的模式。然后,您了解了存储特定的云概念,包括性能、原子更新、数据持久性和 API 访问。在整个章节中,您谈到了云可能和可能不是一个好的解决方案。您还了解了在一些最流行的公共云服务中创建虚拟服务器和对象存储的示例。

章节概要

云计算 在云计算模型中,CPU、网络和存储等计算资源可通过互联网远程访问。此模型比传统的计算模型具有多项优势,例如弹性、可扩展性和可访问性。

云存储 云存储通过互联网通过 API 进行访问。它具有云计算的许多优点,包括弹性、可扩展性和可访问性。

公有云 公有云服务,包括云计算和云存储,在共享基础设施上运行。多个客户共享同一基础设施称为多租户。

API 访问 应用程序编程接口 (API) 允许从应用程序内部访问云计算和云存储等服务。只要云服务的 API 是公开的,您就可以编写自己的代码 (软件) 来直接与该云服务对话。

附录 A



美国计算机技术协会 存储+考试

本附录涵盖的主题：

- ✓ 准备 CompTIA Storage+ 考试
- ✓ 报名考试
- ✓ 参加考试
- ✓ 回顾考试目标



正如本书的介绍章节中提到的,IT 在我们所做的一切事情中都变得越来越重要。在企业和商业 IT 领域,存储是一个关键的 IT 系统,需要熟练且训练有素的专业人员来设计和管理高可用性和高性能系统。证明您在这方面的能力的一种方法是获得 CompTIA Storage+ Powered by SNIA 认证。

要获得此认证,您需要通过以下考试:

CompTIA Storage+ 由 SNIA 考试 SGO-001 提供支持 Storage+ 考试是 CompTIA 与存储网络行业协会 (SNIA) 联合提供的唯一存储认证。该考试涵盖配置基本存储网络和系统所需的所有主题,为构建更深入的存储技能奠定了坚实的基础。该考试面向至少具有一年实践经验的专业人士。

本书涵盖了所有考试目标以及更多内容,使您能够通过考试并在存储和系统管理员的职业生涯中取得成功。

准备考试

准备是成功的关键。既然您已经阅读了本书并准备通过考试,那么考虑一下考试的样貌和感受是个好主意,这样就不会让您措手不及。



强烈建议您参加考试之前访问 [CompTIA](http://certification.comptia.org/getcertified) 网站,因为有关者页面详细信息可能会发生变化。此考试的 CompTIA 网站是 <http://certification.comptia.org/getcertified>。The CompTIA website for this exam is <http://certification.comptia.org/getcertified/certifications/storage.aspx>。

首先要注意的是考试与供应商无关。考试中不会有任何 Dell Compellent 或 HP 3PAR 问题。它们都更加通用和概念化。There are no Dell Compellent or HP 3PAR questions on the exam. They're all more universal and conceptual. 这是一件好事,因为您在本书中了解到的信息可以转移到大多数(如果不是全部)供应商。transferrable across most, if not all, vendors.

考试时间为 1 小时 30 分钟,包含 100 道多项选择题,且 consists of 100 multiple-choice questions. 没有动手练习或实验室模拟。要通过考试,您需要获得 70% 的分数。To pass, you need a score of 70% 或更高。

考试分为表 A.1 所示的专业领域。

表 A.1 考试目标类别

专业领域	考试百分比
1.0 存储组件	20%
2.0 连接性	24%
3.0 存储管理	26%
4.0 数据保护	17%
5.0 存储性能	13%

查看目标表,如果你觉得自己在某个特定领域比较弱,那就做一些
做更多的作业,这样你就可以尽可能自信地参加考试。

参加考试

截至本书出版时,CompTIA Storage+ Powered by SNIA 考试只能在 Pearson VUE 考试中心进行。Pearson VUE 是一家领先的全球培训公司,在全球拥有 5,000 多个考试中心。

要预约考试,您只需访问 www.pearsonvue.com/comptia 并按照在线程序操作即可。该网站可让您找到最近的考试中心、预约考试并付款。要完成此过程,您需要一个 Pearson VUE 帐户 (可在网站上设置) 和一种电子支付方式。考试费用约为 200 美元 (或您所在国家/地区的当地等值金额) 。

预约考试后,您还需要整理一些其他重要事项,例如:

- 规划前往考试中心的路线 您一定不想迟到!
- 请务必携带所需的身份证件。如果没有提供带有您的照片和您当前有效签名的身份证件文件,您将无法参加考试。

一旦你了解了考试中心的位置,并且获得了所需的身份证件,
您已经为考试当天做好了准备。

考试当天,请尽早到达考试中心。您最不需要的就是迟到的压力。考试期间,请准备好交出所有个人电子设备,并将它们存放在安全的储物柜中。

在考试中心签到后,您将被带到一个安全的房间,在那里您将登录到一台计算机并被要求完成一份简短的计算机调查。完成这项调查所需的时间不会影响您参加考试的时间,因此请花点时间完成调查。您在调查中给出的答案不会影响您的考试成绩。

完成调查后,就到了考试的时间。如前所述,考试是多项选择题。您可以标记问题以供以后复习;如果您不确定某个问题的答案,您可以将该问题标记为待复习,在您回答完所有其他问题后,系统将允许您返回到您标记的任何问题。尽最大努力回答所有问题!留下一个未回答的问题没有任何好处。未回答的问题自动被视为不正确。如果您猜测答案,您就有机会猜对。所以请回答所有问题!

回答完所有问题并点击“完成”按钮后,系统会显示您的分数以及是否通过了考试。希望您通过了考试,系统会自动打印出您的成绩单。然后,您会收到一封电子邮件,其中列出了您的成绩,您可以选择将证书邮寄给您。如果您没有通过考试,您可以在任何方便的时间和您选择的考试中心重新参加考试 - 第一次和第二次考试之间无需等待。但是,如果您第二次考试没有通过,则每次重新考试之间至少需要等待 14 天。

回顾考试目标

表 A.2 提供了 SGO-001 考试的目标映射。您还可以在本书的配套网站 <http://www.sybex.com/go/datasstoragenetworking> 上的灰卡和练习考试中找到考试目标的介绍。

表 A.2 考试 SG0-001 CompTIA Storage+ 由 SNIA 提供支持 目标图

1.0 存储组件	对应书籍章节
1.1 描述磁盘类型、组件和特性。	所有内容均在第 2 章中介绍
• SATA	
• Fibre Channel	
• SAS	
• SCSI	
• SSD	
• Spindle	
• Platter	
• Cylinder	
• Heads	
• Speeds	
7,200 转/分	
10,000 rpm	
15,000 rpm	
• I/O vs. throughput	
• Capacity vs. speed	
1.2 比较可移动媒体的类型、组件和特性。	第2章
• Tape	第十章
o 尺寸与速度	第十一章
o 多流和多路复用（优点和缺点）	
o 擦鞋	
o LTO 版本（LTO1、LTO2、LTO3、LTO4、LTO5）	
o 压缩和加密（硬件/软件）	
NDMP	
• Other removable media	
DVD 光盘	
o 蓝光光盘	
o 闪存驱动器	
蠕虫病毒	

表 A.2 考试 SG0-001 CompTIA Storage+ Powered by SNIA 目标图（续）

1.3 给定一个场景,安装和维护连接器和电缆类型（记住它们的属性）。	第十三章 第十四章 第五章
• Fiber cables	
o 多模（短波）与单模（长波） o 长度、速度和距离限制 o 连接器:LC、SC、SFP o 电 缆保养:弯曲半径、应力	
• Copper cables	
o CAT 5 o CAT 5e o CAT 6 串行 o 双轴 SAS 连接器:RJ-45、BD-9	
• SAS1 and SAS2 port speeds	
1.4 描述物理网络硬件的用途。	第五章
• Switch and features	第六章
o 中继 Ø 岛链 o 端口通道 o 端口类型:G_Ports、F_Ports、N_Ports、E_Ports, U_Ports o 董事 o 热插拔	
• HBA	
• CNA	
• Router	

1.5 给定一个场景,安装和维护模块化存储阵列组件。

第3章

第五章

- Controller head

- 单 双

- 网格

- 缓存

- 扩展适配器 阵列端口类型

- 和布线:光纤通道、

- FCoE,iSCSI,SAS

- Disk enclosure

- 机柜控制器

- 监控卡 机柜寻址 布

- 线

- Hot-pluggable

1.6 确定以下环境问题及其相关影响。

第十三章

- HVAC

- 冷却不当 湿度控制不足

- Fire suppression

- Floor and rack loading

- Adequate power

- 足够的容量 适当的电路划

- 分 接地

1.7 在安装和维护存储设备时采用适当的安全技术。

第十三章

- Proper lifting techniques

- Weight considerations

- Antistatic devices

- Rack stabilization

表 A.2 考试 SG0-001 CompTIA Storage+ Powered by SNIA 目标图（续）

2.0 连接性

2.1 识别常见的存储网络行业术语。第 5 章

- Link
- Oversubscription
- Worldwide node name
- Worldwide port name
- Flow control
- N_Port ID
- Buffer-to-buffer credit

2.2 解释以下存储网络行业术语。

- Alias
- Name service
- Link
- Connection
- Initiator
- Target
- Fabric

2.3 给定一个场景,实现以下光纤通道技术。

第五章

- **Topologies**

- 点对点 ◦ 仲裁环路
- 单一结构 ◦ 冗余结构

- **Implementations**

- 分区最佳实践 ◦ 分区别名

区域

- 区域设置

- 硬分区 ◦ 软分区

- 域 ID

- NPIV

- SCSI ID

- **Multipathing**

负载平衡

故障转移

- 磁盘路径数量 ◦ 互操作性

- **Physical connections vs. logical connections**

- **Protocols**

SCSI 接口

氯氯乙烯

光纤通道IP

2.4 给定一个场景,实现以下以太网网络技术。

第五章

- **Features**

VLAN

广域网

城域网

局域网

- **Multipathing**

- iSCSI

◦ 多点输入输出系统

链路聚合

- **Protocols**

- iSCSI

- NFS

- CIFS

表 A.2 考试 SG0-001 CompTIA Storage+ Powered by SNIA 目标图（续）

2.5 了解融合存储网络技术的基础知识。	第五章 第3章 第六章 第7章
<ul style="list-style-type: none"> • FCoE • DCB (DCE, CEE) • LLDP • Class of service • Priority tagging • Baby jumbo frames • 10G Ethernet 	
2.6 给定一个场景,使用适当的网络工具。	第十四章
<ul style="list-style-type: none"> • TCP/IP network <ul style="list-style-type: none"> ◦ ping ◦ ◦ tracert/traceroute ◦ ipconfig/ifconfig ◦ ◦ nslookup • Fibre Channel network <ul style="list-style-type: none"> ◦ 端口错误计数器 ◦ fcping ◦ 名称 ◦ 服务器 ◦ 重新扫描 	
2.7 解决以下常见的网络问题。	第五章 第六章
<ul style="list-style-type: none"> • Bad cables • Bad ports • Bad connectors • Incorrect configuration on NIC • Incorrect VLAN • Bad NIC • NIC improperly connected • Incorrect firewall settings 	
2.8 解决以下常见的光纤通道问题。	第十三章 第五章 第六章
<ul style="list-style-type: none"> • Zoning errors • Zoning misconfiguration • Failed GBIC or SFP • Failed HBA • Connectivity • Interoperability issues • Hardware/software incompatibility • Outdated firmware/drivers • Failed cable • Misconfigured Fibre Channel cable 	

2.9 比较和对比常见的存储基础设施。

第十三章

• SAN

- 光纤通道
- 块模式
- 主机上的文件系统 ○
- FC 或 iSCSI 协议
- 面料

第五章

第六章

• NAS

- 基于 TCP/IP
 - UNC 可寻址存储 ○ 存储上的文件系统 (NFS 或 CIFS)
 - 基于以太网
- DAS
- 直连存储 (SAS、SATA、SCSI) ○ 主机上的文件系统

3.0 存储管理

3.1 解释以下 RAID 级别和相关属性。

第3章

• Levels

- 0
- 1
- Ø 5
- 6
- 1+0 (10)

第五章

第六章

第7章

[0347] ○ 0+1

• Properties

- 高读取 ○ 高写
- 入
- 容错
- 重建时间
- 性能
- 故障模式
- 容量开销

表 A.2 考试 SG0-001 CompTIA Storage+ Powered by SNIA 目标图 (续)

3.2 给定一个场景,执行存储配置技术。	第四章
• LUN provisioning <ul style="list-style-type: none">◦ LUN ID	
• LUN masking and sharing <ul style="list-style-type: none">◦ 基于主机与基于存储 (磁盘/磁带) ◦ 负载平衡	
• Thin provisioning <ul style="list-style-type: none">◦ 薄层填海	
• Best practices for disk provisioning	
3.3 解释卷管理概念。	第3章
• File- vs. block-level architecture	第十章
• Configuration layer <ul style="list-style-type: none">◦ LVM	
• Logical volume	
• Volume group	
• Filesystem	
• Mount point	
3.4 描述一般的虚拟化概念。	第四章
• Virtual storage <ul style="list-style-type: none">◦ 磁带 ◦ 磁盘	第3章 第7章
• Virtual provisioning of the host, array, and fabric	
• LVM	
• VSAN/virtual fabric	
• VLAN	
• NPIV	

3.5 给定一个场景,实现监控、警报和报告。	第九章 第十一章 第五章 第六章
• Setting thresholds	
• Trending	
• Forecasting/capacity planning	
• Recording baseline	
• Setting alerts	
• Auditing log files	
• Alerting methods	
手机	
o 电子邮件	
简单网络管理协议	
o 打电话回家	
3.6 解释管理协议、接口和相关用途。	第十章 第十二章
• Management protocols	
简单网络管理协议	
o 中型	
o 网页管理模块	
• Administration	
命令行界面	
串行	
远程登录	
SSH	
HTTP/S	
• In-band vs. out-of-band management	
3.7 解释信息生命周期管理概念。	第十二章
• Data migration strategies (HSM) and storage tiers	
• Archiving and purging	
• Compliance and preservation	
• Content-addressable storage (CAS)/object-oriented	
储存 (缺货)	
• Value of data based on frequency of access	
3.8 解释重复数据删除和压缩的各种功能和区别。	第7章 第3章 第十章 第十一章
• Inline and post-process deduplication	
• Software based vs. appliance based	
• Single-instance storage	
• Performance and capacity implications	
• Reduction ratios vs. data type	

表 A.2 考试 SG0-001 CompTIA Storage+ Powered by SNIA 目标图 (续)

4.0 数据保护

4.1 解释冗余概念、相关目的和组件。

第十章

- High availability
- Single point of failure
- Component redundancy

电源

控制器

- 磁盘 (热备用) ◦ 路径/总线

◦ 开关

◦ 羟丁酸

◦ 网卡

◦ 数组

- Cache battery backup and cache mirroring

4.2 比较和对比不同的复制方法和属性。

第3章

- Synchronous and asynchronous
- Local vs. remote
- Site redundancy
- Snapshots and clones
- Replication consistency

第四章

第五章

4.3 解释长期存储的数据备份概念的基础知识。

第八章

- Recovery point objective (RPO) and recovery time

目标 (RTO)

- Backup and restore methods

- 全部

- 增量

- 差速器

- 渐进式

- Backup implementation methods

- 无需 LAN

- 无服务器

- 基于服务器

- Backup targets

- 磁盘到磁盘

- 磁盘到磁带

- 虚拟磁带库

- D2D2T

- Vaulting vs. e-vaulting

- Verify backups

- 数据完整性 ○ 校验

- 和

- 应用程序验证

- Data retention and preservation policy

- 轮换方案 (GFS 祖父方案,
父子)

- 企业和法律合规性 ○ 异地磁带存储/灾难恢复
计划

表 A.2 考试 SG0-001 CompTIA Storage+ Powered by SNIA 目标图（续）

4.4 解释数据安全的基本概念和重要性。 第十一章

- **Access management**

访问控制列表

- 物理访问 ○ 多协议/互

操作性

- **Encryption**

- 磁盘加密 ○ 磁带加密

- 网络加密 (IPsec) ○ 主

机加密 ○ 加密密钥

- **Storage security**

- 共享访问 (NFSv3 与 NFSv4) ○ 共享访问 (CIFS)

- 文件权限与共享/导出权限

5.0 存储性能

5.1 解释延迟和吞吐量如何影响存储性能。

第3章

第十一章

第7章

- **Cache**

- 读取与写入流量

- 取消暂存

- 缓存命中和未命中

- **RAID type and size**

磁盘数量

- **IOPS calculations**

- **Random vs. sequential I/O**

- **Impact of replication**

5.2 识别调整和工作负载平衡概念。

第3章

- **Application to storage data profiling**

第四章

- **Tiering**

- 自动

手动

高速加工

- **Partition alignment**

- 碎片化和对性能的影响

- **Queue depth**

第八章

5.3 描述存储设备带宽属性和功能。	第3章 第十章
<ul style="list-style-type: none">• Bus bandwidth/loop bandwidth• Cable speeds• Disk throughput vs. bus bandwidth vs. cache• Embedded switch port speed• Shared vs. dedicated• Multipathing for load balancing	
5.4 描述网络设备带宽属性和功能。	第3章 第五章
<ul style="list-style-type: none">• Shared vs. dedicated• Teaming/link aggregation• Class of service<ul style="list-style-type: none">◦ 巨型帧• TOE	
5.5 解释存储/主机工具的性能指标、参数和用途。	第五章 第六章
<ul style="list-style-type: none">• Baseling and data capture• Switch<ul style="list-style-type: none">◦ 端口统计◦ 阈值◦ 跳数 ◦◦ 端口组◦ ISL/中继线◦ 带宽• Array<ul style="list-style-type: none">◦ 缓存命中率◦ CPU 负载◦ 端口统计◦ 带宽<ul style="list-style-type: none">◦ 吞吐量 ◦ I/O 延迟• Host tools<ul style="list-style-type: none">◦ sysmon ◦◦ perfmon ◦ iostat	



考试目标可能随时更改，恕不另行通知。
考试目标可能随时更改，恕不另行通知。
确保您访问<http://certification.tutorialspoint.com/getcertified/comptia.org/getcertified/certifications/storage.aspx>。如果考试的 URL 已移动，Google 可能知道它在哪里！

祝考试顺利！

附录

乙

关于附加 学习工具





本附录包含以下内容：

- ✓ 额外的学习工具
- ✓ 系统要求
- ✓ 使用学习工具
- ✓ 故障排除

额外的学习工具

以下部分按类别排列，并总结了您将从配套网站找到的软件和其他好东西。如果您需要安装这些项目的帮助，请参阅本附录“使用学习工具”部分中的安装说明。



您可以在 www.sybex.com/go/otherstudytools 找到其他学习工具。在 www.sybex.com/go/storageplus，您将获得有关如何将文件下载到硬盘的说明。

Sybex Test Engine

文件包含 Sybex 测试引擎，其中包括 CompTIA Storage+ 考试的一次奖励模拟考试 practice exam for the CompTIA Storage+ exam.

Electronic Flashcards

这些方便的电子卡片就像它们的名字一样：一面写着问题，另一面则显示答案。One side contains a question, and the other side shows the answer.

PDF of Glossary of Terms

我们已提供 PDF 格式的电子版词汇表。您可以使用 Adobe Reader 查看电子版词汇表。You can view the electronic version of the Glossary with Adobe Reader.

Adobe Reader

我们还提供了可用于下载 Adobe Reader 副本的链接，以便您可以查看随书内容提供的 PDF 文件。有关 Adobe Reader 的更多信息或查看更新版本，请访问 Adobe 网站 www.adobe.com/products/reader/，或访问 Adobe 网站 www.adobe.com/products/reader/，以获取更新版本。

系统要求

确保您的计算机满足以下列表中显示的最低系统要求。如果您的计算机不符合其中大部分要求,则您在使用软件和文件时可能会遇到问题。有关最新和最有用的信息,请参阅下载中的自述文件。

- 运行 Microsoft Windows 98、Windows 2000、Windows NT4（带 SP4 或更高版本）、
Windows Me、Windows XP、Windows Vista 或 Windows 7 的 PC
- 互联网连接

使用学习工具

要安装这些项目,请按照下列步骤操作:

1.将.ZIP文件下载到您的硬盘,然后解压到适当的位置。有关下载此文件位置的说明可在此处找到:

www.sybex.com/go/datasstoragenetworking。

2.单击Start.EXE文件打开学习工具文件。

3.阅读许可协议,如果要使用,请单击“接受”按钮
学习工具。

主界面出现。该界面允许您只需单击一两次即可访问内容。

故障排除

Wiley 已尝试提供可在大多数计算机上运行且满足最低系统要求的程序。但您的计算机可能有所不同,某些程序可能由于某些原因无法正常运行。

最有可能出现的两个问题是,您没有足够的内存 (RAM) 来运行您想要使用的程序,或者您正在运行其他程序,而这些程序影响了程序的安装或运行。如果您收到“内存不足”或“安装程序无法继续”等错误消息,请尝试以下一个或多个建议,然后再次尝试使用该软件:

关闭计算机上运行的所有防病毒软件。安装程序有时会模仿病毒活动,并可能让您的计算机误以为它已被病毒感染。

关闭所有正在运行的程序。运行的程序越多,其他程序可用的内存就越少。安装程序通常会更新文件和程序;因此,如果您让其他程序继续运行,安装可能无法正常工作。

让当地的计算机商店为您的计算机添加更多 RAM。诚然，这是一个激进且成本较高的步骤。但是，添加更多内存确实可以提高计算机的速度并允许同时运行更多程序。

客户服务

如果您在使用本书的配套学习工具时遇到问题，请拨打 Wiley 产品技术支持电话 (800) 762-2974，分机 74，或发送电子邮件至<http://sybex.custhelp.com/>。

指数

读者须知：本索引中粗体页码表示对某一主题的主要讨论。斜体页码表示插图。

A

接受 (ACC) 框架, 213 访问控制云存储 API, 518–519 数据中心, 479
NFSv4, 261 访问控制条目 (ACE), 261 访问控制列表 (ACL), 261
接入网关 (AG) 模式, 210–211
ACID 缩写, 513 采购成本, 50 双控制器架构中的活动控制器, 61 自动分层中的活动监控, 381–382, 382 执行器组件, 21 临时布线, 476
添加计数器屏幕, 445, 445 寻址磁盘驱动器
磁头 中学生, 14–15, 15
LBA, 20–21, 21
高级格式 标准, 19, 19
AFR (年故障率), 35
AG (访问网关)模式, 210–211 聚合层, 478 警报, 448–450 别名 iSCSI, 239
SAN 区域, 189–192 alicreate 命令, 190 外来串扰 (AXT), 475 全闪存存储阵列, 65–67 按需分配模型, 356 写入时分配快照, 101, 325 备选 iSCSI 命名约定, 239

ALUA (非对称逻辑单元访问), 61–62 永远在线的企业, 59

Amazon Glazier 档案, 284
Amazon 简单存储服务 (S3) 服务 API, 519 对象存储, 284 公共云存储, 515–518, 515–518
亚马逊网络服务 (AWS), 505, 508–510, 508–510 年故障率 (AFR), 35 防病毒程序, 278

API (应用程序编程接口) 云存储, 518–519 对象存储, 281–282 应用程序感知备份, 404–405 应用程序感知快照, 100, 324 基于应用程序的复制, 86 基于应用程序的快照, 101–102 应用程序一致性副本, 300 应用程序一致性快照, 100 应用程序层复制, 299–301 应用程序编程接口 (API) 云存储, 518–519 对象存储, 281–282
专用集成电路 (ASIC), 268 不了解
应用程序的副本, 300 仲裁环路, 196 体系结构 备份, 394–396, 395 对象存储, 282–283, 282 存储阵列 双控制器, 60–62 网格存储, 62–63 概述, 58, 59 冗余, 59 存档, 415 备份为, 419 与备份, 415–416 容量管理, 433–435 合规性, 416–417 发现, 417–418 介质, 418–419 面密度, 17 N_Port ID 区域, 208 基于阵列的复制, 86, 306–310, 310 基于阵列的快照, 97–102, 318–319 阵列 自动分层, 384 NAS, 55, 267–270, 269

横向扩展, 272–275, 274 传统, 271–272 统一, 272 存储。请参阅存储阵列 虚拟化, 343 基于控制器, 343 功能, 345 延长使用寿命, 344

ASIC (专用集成电路), 268 非对称逻辑单元访问 (ALUA), 61–62 非对称虚拟化, 337, 337 异步复制 基于日志, 92, 308–310, 310 概述, 298–299, 299 基于快照, 307–308 存储阵列, 89–92, 90, 93

原子测试和设置 (ATS), 122 原子更新, 513
原子上传到云存储, 520

ATS (原子测试与设置), 122 多模光纤中
的衰减, 180 iSCSI 中的身份验证, 241–245, 242–
245 内容真实性, 285 自动分层, 112 活动监控, 381–382, 382
阵列大小, 112, 384 概述, 380 策略,
383 远程复制, 115,
384 计划, 114–115, 382 存储虚拟化, 119–120,
120 子 LUN, 112–113,
380–381 层, 113–
114, 114, 381 虚拟化, 345–346, 345 自动设备发现, 240
自动协商 FC SAN 端口速度, 176
可用性

备用电源系统, 468 备份服务器, 394–
396, 395 备份目标, 395 备份到磁盘 (B2D),
411–412 备份, 388 应用感知
知, 404–405 架构, 394–396, 395 档案。请参阅
存档 备份窗口,
393 业务连续性, 389 云, 412–413, 413
差异, 403 磁盘, 411–412 完整, 402 热, 396 增量, 402–403 基于 LAN,
397, 398 无需 LAN, 397–398
398 NDMP, 400–401, 401 脱机, 397 目的, 388–389 恢复点目
标, 389–390 恢复时间
目标, 390 保留策略, 413–415 调度, 390–
393, 391–392 无服务器, 398–
400, 399 与快照相比, 317 交换
机配置, 217 合成完整, 404 磁带, 406–411, 410 重复数据删除用例, 377–378,
377 VSS, 405–406 坏块
映射, 22 带宽 异步复制, 92, 93, 310, 310 增强以
太网, 488 ETS, 490–491 同步复制, 297, 298

NAS 阵列, 270 性能,
4–5 平均延迟, 29

AWS (亚马逊网络服务), 505, 508–510, 508–510

AWS 管理控制台, 508
AXT (外部串扰), 475

乙

B2B (缓冲液到缓冲液)控制, 212–213, 488, 489
B2D (备份到磁盘), 411–412
BAARF (对抗任何 RAID 五), 441 小型巨型帧, 492 主干布线,
477 后端对象存储, 282, 282

SDS, 348 存储阵列, 58, 59, 83–85 后台闪存磨损均衡, 43–44 备份客户端, 395

基准性能, 436–437 基本头段 (BHS), 235 电池供电缓存, 81 内部 RAID 控制器, 145 写回, 145

对抗任意 RAID 五 (BAARF), 441
BBWC (电池供电的写回缓存), 145
BC (业务连续性)备份, 389 计划, 290–293 存储阵列, 85–
86 光纤弯曲半径, 182

BHS (基本头段), 235 绑定, 语言, 519

指纹识别系统中的逐位比较 - CAS (内容可寻址存储) 551

指纹识别系统中的逐位比较,374 重建域中的爆炸区,143 基于块的备份,397 基于块的重复数据删除,371–374,373 基于块的带内虚拟化,339 具有分布式奇偶校验的块级条带化,153–157,155–157 块,55,256 融合,497 光纤 SAN, 169 闪存, 37, 37 VAAI 卸载, 121 链路聚合中的绑定, 277 从 SAN 启动, 172 BPDU Guard 工具, 234 桥接拓扑, 252 数据中心内的桥接能力交换, 491 Brocade 交换机 警报, 449 检查, 174–175 虚拟结构, 204 存储桶 Amazon S3, 284, 515 对象存储, 283 缓冲区到缓冲区 (B2B) 流量控制, 212–213, 488, 489 异步复制中的缓冲区, 309 ODX 中的批 量归零, 127 防弹阵列, 64 掩体站点, 93, 311 突发速率, 34 业务连续性 (BC) 备份, 389 计划, 290–293 存储阵列, 85–86 缓存命中压缩, 369 IOPS, 32 RAID, 443 存储阵列, 81 缓存未命中压缩, 369 IOPS, 32 RAID, 443 存储阵列, 81 缓存算法 和技术, 82 全闪存阵列, 65–66 数据和控制, 81 闪存, 44, 82–83, 145 概述, 46 性能, 80, 442–443 持久性, 82 保护, 81–82 读取和写入, 80 存储阵列, 58, 59, 79–83 虚拟化, 347 写回, 133 磁盘驱动器笼, 34 算磁盘驱动器 IOPS, 33 通过警报呼叫家庭, 448 容量管理容量报告, 424–426, 425–427 重复数据删除和压缩, 431–433, 432 概述, 423–424 配额和归档, 433–435 showback 和 chargeback, 435–436 精简配置注意事项, 427–431, 429–430 机械驱动器, 25–26 优化技术 自动分层。请参阅自动分层 压缩。请参阅压缩 概述, 354 性能影响, 355 精简配置。请参阅精简配置 (TP) 使用案例, 356 过度配置, 105–106 存储阵列, 67 容量岛, 67 资本 支出, 51 CAS (内容可寻址存储) 归档, 417 对象存储, 285–286

C

CaaS (云作为一层), 523 电缆数据中心, 474 水平分布区, 477–478 主要分布区, 477 标准原则, 474–475 结构化, 475–477, 476

FC SAN, 178, 179–180 弯曲半径, 182 连接器, 183–184, 184 多模光纤, 180–181, 180 单模光纤, 182 缓存命中率, 443

级联拓扑描述, 201,
 201 三站点, 93–94, 93,
 311–312, 311
 “廉价磁盘冗余阵列案例
 (RAID)”论文, 130 个数
 据中心布线类别, 474–475
 CBT (更改块跟踪)引擎, 305
 CDB (命令描述符块), 235
 CDMI (云数据管理)
 倡议), 523–524
 CDN (内容分发网络)云存储, 520–521 提供
 商, 283
 CEE (融合增强型以太网), 487–490, 489 单元管理器, 394 闪存
 中的单元, 37–38, 37
 cfgactvshow 命令, 190 cfgadd 命令, 190
 cfgsave 命令, 190
 质询握手认证协议
 (CHAP), 241–245, 242–245, 251 更改块跟
 踪 (CBT) 引擎, 305 更改, 测试, 358–359, 383–384, 405 通道
 FCP 堆栈, 185 光纤通道。请参阅光纤通道 SAN 闪存, 43
 CHAP (质
 询握手身份验证协
 议), 241–245, 242–245, 251 CHAP 凭证屏幕, 244–
 245, 244–245 容量管理
 中的退款, 435–436 CHS (柱面-头部-扇区)寻址方案, 15, 15 RAID
 中的块, 140–141 CID (连接 ID), 247 CIFS (通
 用 Internet 文件系统), 55, 257
 权限, 265–266 报告份额安
 装, 265 优势, 263–265
 CIS (云集成存储), 523
 Cisco VSAN, 203–204 光纤包
 层, 179–180、180 基于类别的流量控制, 489 服务类
 别 (COS) 增强型以太网, 488–490、
 489–490
 SAN, 214–215
 CLI (命令行界面), 454–455 客户端重复数据删除,
 374 客户端备份, 395
 CIFS, 263
 光纤 SAN, 169
 iSNS, 250
 SNMP, 457

克隆快照, 320–322, 322 云作为层
 (CaaS), 523
 云数据管理计划 (CDMI), 523–524 云集成存储 (CIS), 523 云对象规
 模, 280 云存
 储, 5, 502
 API 访问, 518–519 原子上
 传, 520 备份, 412–413,
 413
 CDMI, 523–524
 CDN, 520–521 数据
 持久性, 513–514 最终一致性模
 型, 514 特性和优势, 511–513, 512
 地理复制, 520 混合, 511,
 522–523, 523 影响, 6 对象, 283 概
 述, 502–504 私
 有, 510–511,
 522 公共, 504–505, 514–
 520, 515–518 小型办公室,
 521 重复数据删除用例, 378–379 虚拟机
 亚马逊网络服务, 508–510, 508–510
 Windows Azure, 505–507, 505–507 集群感知
 文件系统, 273 集群架构, 62 集群
 LUN, 79
 NAS, 273
 CmdSN (命令序列号), 236
 CNA (融合网络适配器), 170
 FCoE, 494
 光纤 SAN, 172 iSCSI
 SAN, 229 冲突指纹识
 别系统, 374
 对象存储, 286 IQN 中的冒号 (:), 237–
 238 命令复合, 264 命令描
 述符块 (CDB), 235 命令设备, 453 命令
 行界面 (CLI), 454–455 命令序列号
 (CmdSN), 236
 通用互联网文件系统 (CIFS), 55, 257
 权限, 265–266 报告共享安
 装, 265 优势, 263–265 符合档案规定,
 285–286, 416–417 复合
 RPC, 259–260 复合, 命令, 264

压缩, 109–111 容量管理, 431–433, 432 重复数据删除, 379 描述, 355 未来, 369 内联, 368–369 操作, 367–368 概述, 366–367 性能问题, 369 后处理, 369 磁带备份, 409 计算机房, 476 交换机配置转储, 220–221 拥塞管理, 490 连接忠诚度, 247 连接 ID (CID), 247 连接

光纤通道 SAN, 217–219 存储阵列, 70–72, 71–72, 84 FC SAN 连接器, 183–184, 184 控制台电缆, 455 整合本地快照, 316

NAS 阵列, 270–271 消费级闪存驱动器, 42 消费模型, 412 容器 CDMI, 524 对象 存储, 283 内容可寻址存储 (CAS) 档案, 417 对象存储, 285–286 档案中的内容保证, 416 内容真实性, 285 内容交付网络 (CDN) 云存储, 520–521 提供商, 283 控制缓存, 81 控制平面, 349 控制站, 250 基于控制器的虚拟化, 115–116, 116, 336 配置, 341–344, 342 概述, 339–341, 339–341 控制器, 22 双控制器架构, 61 结构服务器, 187

NAS, 57岁 RAID, 133–134, 144–145 SDS, 348 无 缝故障, 62–63 虚拟化, 18 融合块, 497 融合增强以太网 (CEE), 487–490, 489

融合网络适配器 (CNA), 170 FCoE, 494 光纤 SAN, 172 iSCSI SAN, 229 融合网络数据中心, 486 增强型以太网, 487–491, 489–490 FCoE, 491–498 存储 网络要求, 486–487 以太网, 484–485 光纤通道, 485 冷却数据中心, 470–472, 470–471 铜双轴电缆, 478 写时复制 (CoW) 快照, 101, 324–325 核心边缘拓扑, 199–201, 200 光纤芯, 179–180, 180

COS (服务类别) 增强以太网, 488–490, 489–490 SAN, 214–215 成本和开销 闪存, 44–45 iSCSI SAN, 227–228

RAID 0, 146 RAID 1, 150 RAID 5, 153–154 RAID 6, 157–159 showback 和 chargeback, 435–436 存储阵列, 69 存储设备, 50–51 磁带备份, 407

计数器 Perfmon, 444–446, 444–446 端口错误, 221–224 CoW (写时复制) 快照, 101, 324–325 存储阵列的 CPU, 74–76, 83–84 崩溃一致性磁盘驱动器, 11 副本, 307 快照, 100, 324

CRC (循环冗余校验) EDP, 16 FC 框架, 212 创建存储桶屏幕, 515, 516 创建支持命令, 220 SAN 流量控制中的信用系统, 212 串扰, 475 累积增量备份, 402–403 自定义硬件, 268 直通交换技术, 57 循环冗余校验 (CRC)

EDP, 16 FC 框架, 212 柱面头 扇区 (CHS) 寻址方案, 15, 15 柱面, 19–20, 20

德

D2D2C (磁盘到磁盘到云)备份, 412 D2D2T (磁盘到磁盘到磁带)备份, 412 DAC (直连电缆), 478 DAG (数据可用性组), 302–303, 302 DAR (直接访问恢复)功能, 401 DAS (直连存储)与 NAS, 271 IQN 中的破折号 (-), 237–238 数据可用性组 (DAG), 302–303, 302 数据缓存, 81 数据中心桥接 (DCB), 487 数据中心桥接能力交换协议 (DCBX), 491 数据中心访问控制, 479 桥接能力交换, 491 布线, 474 水平分布区, 477–478 主分布区, 477 标准原则, 474–475 结构化, 475–476 冷却, 470–472, 470–471 防火系统, 479–480 地板, 464–465, 465 概述, 460 电源, 467–469, 468–469 机架, 462–464, 463–464 行, 465–466, 466 行为准则, 480 服 务器安装, 472–474 单一融合网络, 486 增强型以太网, 487–491, 489–490 FCoE, 491–498 存储网络要求, 486–487 层, 461–462 云存储中的数据持久性, 513–514 数据完整性字段, 16 Data ONTAP, SDS 中的 160 个数据平面、IQN 中的 349 个日期, 238 个 DCB (数据中心桥接), 487 个 DCBX (数据中心桥接能力交换协议), 491 个 DDS (发现域集), 250 个事实 上和法律上的标准, 524 个退役服务器, 77 个专用 VLAN, 233 个重复数据删除全闪存阵列, 66 个基于块的阵列, 371–374, 373

容量管理, 431–433, 432 CAS, 286 带压缩, 379 描述, 355 联合, 375 内联, 375–376 概述, 370–371, 370 后处理, 376, 433 主存储和云用例, 378–379 基于源, 374 存储阵列, 111–113 摘要, 379–380 基于目标, 375 用例, 377–379, 377 碎片整理文件系统, 48–49, 48–49 降级模式 RAID 1, 148–149 RAID 5, 155 删除文件系统, 357–358 空间回收, 363–365 节省空间的快照中的增量, 319 分界点, 477 RAID 条带深度, 140 倾斜过程, 197 设备特定模块 (DSM), 73 iSCSI 中的设备发现, 239–240 登录, 241, 241 DFS (分布式文件系统), 269 差异备份, 403 直接访问恢复 (DAR) 功能, 401 直连电缆 (DAC), 478 直连连接, 70–71 直连存储 (DAS) 与 NAS, 271 FC SAN 中的控制器, 173–178, 175 灾难恢复 (DR), 290 备份, 416 概述, 393–394 RAID-Z, 164 丢弃数据包, 485 发现存档, 417–418 iSCSI, 239–240 发现域集 (DDS), 250 发现域, 248–249, 249 磁盘驱动器 备份, 411–412 描述, 10 碎片, 23, 49 混合, 45

测绘,258 机械。请参阅机械磁盘驱动器 可靠性,34–35 存储阵列,84–85 磁盘管理插件,73, 73 磁盘到磁盘到云 (D2D2C) 备份,412 磁盘到磁盘到磁带 (D2D2T) 备份,412 diskpart 命令,219–220 复制中的距离,88, 295–296 分布式文件系统 (DFS),269 分布式文件系统,273 分布式 RAID,156–157, 156–157, 161 分布式备用,142 分发网络 在 CDN 中,521 布线文档,475 域 ID 在 FC SAN 中,176 域发现,248–250, 249 在 IQN 中,237 N_Port ID,208 重建,143 双奇偶校验 RAID 6, 157 RAID-X, 163 驱动器副本重建, 142–143 驱动器。请参阅磁盘驱动器 丢包,485 DSM (设备特定模块), 73 双控制器架构, 60–62 云存储的耐久性, 513–514 单模光纤的灰尘, 182, 183 ETS 中的动态带宽分配,490–491 封装 FCoE,492–493, 492–493 IP 网络,235–236, 235–236 SCSI,492,492 磁带备份加密,409 光纤 SAN 中的终端设备,171 帧结束 (EOF) 有序集,212 端到端 (E2E) 信用系统,212–213 端到端数据保护 (EDP),16 增强型以太网,487–488 带宽,488 拥塞管理,490 数据中心桥接能力交换,491 优先级和服务类别,488–490,489–490 传输选择,490–491 增强型传输选择 (ETS),490–491 企业级阵列,63–64 企业级固态术语,42 企业级闪存,40, 42 个入口房间,477 个 EOF (帧结束) 有序集,212 个 EPO (紧急断电) 按钮,480 个设备分布区 (EDA),478 个设备室、476 个擦除闪存,37–38 个 RAID 的擦除码,164–165 个错误校正码 (ECC),19 个 SAN 流控制中的错误恢复,213 个 /etc/exportfs 文件,262 个以太网描述,484–485 增强。请参阅增强型以太网 ETS (增强型传输选择),490–491 EUI (扩展唯一标识符),239 最终一致性模型,514 EX_Port 模式,175 SAN 流量控制中的交换,212 自动分层中的排除,115,383 排他或 (XOR) 运算奇偶校验,138–139 RAID 5,155 exportfs 命令,262 导出 NFS 共享,262 导出,NAS,257 EXTENDED COPY 命令,124,399–400 扩展唯一标识符 (EUI),239 可扩展访问方法 (XAM) API,279 基于扩展区的锁定,ATS,122 扩展区大小

埃

E_Port 模式, 175
E2E (端到端) 信用系统, 212–213 急切归零 (EZT)
卷, 122
EC2 (弹性计算云), 505
ECC (错误更正码), 19
EDA (设备分布区)、478 个 CDN 中的边缘位置, 520–521
编辑服务器屏幕, 243, 243
EDP (端到端数据保护), 16
弹性计算云 (EC2), 505 弹性计算, 503

EMC PowerPath 软件, 74 紧急缓存降级模式, 80 紧急断电 (EPO) 按钮, 480 eMLC (企业级 MLC) 闪存, 40

EMX VMAX 存储阵列, 54, 54

基于快照的异步复制, 308 快照, 98–99, 322–323 子 LUN, 113 精简配置, 103, 358

外部阵列,使用寿命,119 外部闪存缓存,
83 外部 RAID 控制器,134 虚拟化
控制器的外部存储,118

EZT (急切归零厚)卷,122

F

F_Port 模式, 175 结构登
录 (FLOGI) 处理功能, 188 NPV 交换机,
210–211 结构最短
路径优先 (FSPF) 协议, 187, 496
结构 FC, 171, 185–186, 185 重新配置, 216–217 SAN 分区,
187–188 服
务, 175
虚拟, 204 工厂格式, 18 Oracle
Data Guard 中的故障转移, 301 仅故
障转移负载平衡, 72 闪存故障,
41 存储阵列中的
扇入, 72, 72 快
速文件克隆, 125 容错, 137
FBA (固定块架构), 19 FC-AL 拓扑, 196 FC-P2P
拓扑, 194–196, 195 FC SAN。请参阅光纤通道
SAN FC-SW 拓扑, 196–197 FCoE。请参
阅以太网光纤通道 (FCoE)

FCoE 转发器 (FCF), 496 FCP (光纤通道
协议), 168 fcping 命令, 217–218 FCR (光纤通
道路由)技术, 204 fctrace 工具, 219
FEC (前向纠错), 164–165 联合重复数据删除, 375 FFP (全功能
阶段), 246–247, 246
以太网光纤通道 (FCoE), 491–492 融合网络适配器, 494 封
装, 492–493, 492–493 无损网络功能,
494–495, 495 现实世界, 497–498 交换机, 496, 496 光
纤通道协议 (FCP), 168 光纤通道路由 (FCR) 技术, 204 光纤通道
SAN, 168 电缆, 178, 179–180 弯曲半径, 182
连接器, 183–184, 184

多模光纤, 180–181, 180 单模光纤, 182 驱动
器, 16, 24 结构, 185–187, 185 逻
辑, 请参阅逻辑
SAN 组件 命名, 205 N_Port
ID, 206–210, 209 全球名称, 205–206 网络, 57, 485
概述, 168–169,
169 物理组件, 171–172 用途, 170–171
路由, 202–204, 203 存储阵列, 178 交換
机和控制器, 173–178,
175 拓扑 仲裁环路, 196 级联,
201, 201 服务类别, 214–215 核心-边缘拓
扑, 199–201, 200 结构重
新配置, 216–217 FC 命名, 205–
211, 209 FC 路由, 202–
204, 203 流量控制, 212–214 成帧, 211–212, 211
交換机间链路,
197–199, 199 局部性, 215–
216 网状, 201–202, 202
超额认购, 215 点对点, 194–196, 195
冗余, 194 环, 201 交換结构, 196–197 分层, 202
故障排除, 217 连接问题, 217–219 端口错误计
数器, 221–224 SCSI 总线重新扫描,
219–220 交換机配置备份, 217 交換
机配置转储, 220–221 现场可编
程门阵列 (FPGA), 268 50 微米
多模光纤, 180–181 文件级重复数据删除,
371–372 文件级单实例,
286 文件服务器与 NAS 阵
列, 269–271 文件传输协议
(FTP), 267–268 文件管理器, 55, 57 文件
概述, 256–258 文件系
统数据删除,
357–358 碎片, 47–49, 47–49 指纹
CAS 对象, 285
重复数据删除, 373–374

防火系统,479–480 存储阵列固件,74 首遍重复数据删除,374 固定块架构 (FBA),19 定长块重复数据删除,372–373,373 闪存,5,37 缓存,44,82–83,145 通道和并行性,43 成本,44–45 垃圾收集,42

NAND,38–41,39–41 过度配置,42 可靠性,41 磨损均衡,43–44 写入放大,43 写入悬崖,45 写入耐久性,43 写入,37–38,37 平面命名空间,278,279

FLOGI (结构登录)进程函数,188

NPV 交换机,210–211 数据中心的地板,464–465,465 流量控制,212–214,488 磁盘驱动器的飞行高度,13–14,14 数据中心的食物和饮料,480 强制刷新模式,80 机械驱动器的外形尺寸,25 前向纠错 (FEC),164–165

FPGA (现场可编程门阵列),268 碎片,磁盘驱动器,23,47,49 碎片整理,48–49,48–49 文件系统,47–49,47–49 固态驱动器,50 帧 SAN 拓扑,211–212,211

Frankenstorage, 58, 272 前端对象存储, 282, 282

端口速度, 74 存储阵列, 58, 59 存储阵列, 69–74, 71–74

FSPF (结构最短路径优先)协议,187,496
FTP (文件传输协议),267–268
FTPS (带 SSL 的 FTP),267
完整备份, 402, 404 完整克隆, 98, 320–322, 322 完整功能阶段 (FFP), 246–247, 246 完整文件克隆, 125 RAID 1 中的完全冗余, 151 云存储中的完全托管服务, 503 存储阵列的功能, 68 fwmailcfg 命令, 449

G

闪存中的垃圾收集, 42 网关守卫设备, 453 网关 iSCSI, 252

NAS,268,275–276,276 地理复制,513,520 GET 命令, 281 GFS (祖父-父亲-儿子)备份方案,413–415

Gibson, Garth, 130 千兆字节, 26 全局重复数据删除, 112 全局命名空间, 273–274, 274 Google 数据中心功耗,467 祖父-父亲-儿子 (GFS) 备份方案,413–415 快照粒度,98 图形用户界面 (GUI),455–456 网格存储架构, 62–63 组,RAID,135,135 保护字段,16

GUI (图形用户界面),455–456

H

业务连续性计划中的 HA (高可用性),293 FC SAN 中的硬编码交换机端口速度,176 硬盘驱动器 (HDD) 存储。请参见磁盘驱动器;机械磁盘驱动器 SAN 中的硬分区,193 硬件加速复制,124,124 硬件加速归零, 122–123,123 硬件辅助锁定, 122 基于硬件的重复数据删除,375 硬件卸载,121 Microsoft ODX,125–127,126 VMware VAAI,121–125

硬件 RAID, 132–134 哈希重复

数据删除, 373–374 对象存储, 286 hashow 命令, 174

HBA (主机总线适配器),170

FCoE,494 iSCSI SAN,229–230 SAN,172,173

HDA (水平分布区域),477–478 次磁头碰撞,FC 帧中的 13–14 个标头,212

磁盘驱动器

动器磁头,13–15,14–15
 CHS 寻址,14–15,16
 LBA 寻址,20–21,21 抖动,28–29
 NAS,57,270 19
 英寸机架高度,463 异构存储,451
 业务连续性计划中的高可用性 (HA),293 高写入待处理水印,80 ISL 中的跳数和跳数,198,199 水
 平分布区域 (HDA),477–478 NAS 阵列的水
 平可扩展性,272–273 基于主机的复制,86–87
 基于主机的快照概述,101–102 和卷管理器,326–327,327
 基于主机的虚拟化,334–336,335 主机总线适配器 (HBA),170

启动器模式,117 启动器会话

ID (ISID),247 启动器 光纤 SAN,172 iSCSI SAN,228–232,230–232 SAN 区域,189 内联压缩,110,368–369 内联重复数据删除,111–112,375–376
 内联空间回收,107,365–366 每秒输入/输出操作数 (IOPS),19,30–31,31 缓存命中和未命中,32 计算,33–34
 性能管理,438,440 随机与顺序,31–32,31 实例化,111,286 瞬时克隆,321 智能,存储阵列。查看存储阵列智能磁盘控制器,21 智能直通,210 智能
 存储阵列,75 进程间通信 (IPC),264 交换机间链路 (ISL)

FCoE,494

iSCSI SAN,229–230
 SAN,172,173 主
 机集成空间回收,107 热备份,396 热/冷通道原理,470–471,470
 RAID 中的热备用,141–143,144
 热备用路由器协议 (HSRP),201 存储阵列中的热插拔,59–60,84 FC SAN 中的集线器,177–178 混合云,511,522–523,523 混合驱动器,45 基于虚拟机管理程序的复制,86–87,304–306 基于虚拟机管理程序的快照,315–318,316–318

概述,197–199,199 中继,197 VSAN

间路由 (IVR),204
 接口和协议,22–24 API 云存储,518–519 对象存储,281–282 iSCSI SAN,228–230 NAS,258–263,260–261 存储管理,452–457,453–454 内部存储驱动器寻址,14–15,15 虚拟化控制器,118 互联网小型计算机系统接口。请参阅 iSCSI (Internet 小型计算机系统接口) SAN Internet 存储名称服务 (iSNS),247 发现
 域,248–250,249 状态更改通知,250–251 FC SAN 中的互操作性模式,177 内部网 SCSI,234 调用 DR,90,293,299 Iometer 工具,33 IOPS (每秒输入/输出操作数),19,30–31,31 缓存命中和未命中,32 计算,33–34 性能管理,438,440 随机与顺序,31–32,31 iostat 工具,447 IP 网络,232–233 封装,235–236,235–236 通过 WAN 和 LAN 的 iSCSI,234

我

I/O 混合器效果,67,439
 I/O 操作,30
 IBM RAMAC 350 磁盘存储单元,5,12
 IBM XIV RAID,161
 IDA (信息传播算法),164–165
 ILM (信息生命周期管理),434–435 不可变数据,285 带内管理,453,454 带内虚拟化,336–337,337 闪存的动态磨损均衡,43 增量备份,402–403 永久增量备份系统,404 不可破坏的企业级阵列,64 信息分散算法 (IDA),164–165

信息生命周期管理 (ILM),434–435 信息技术的重要性,2–3 存储在信息技术中的作用,3

iSCSI PDU, 235 选件, 233–234
 IPC (进程间通信), 264
 IQN (iSCSI 限定名称), 77, 237–238, 238 iSCSI 启动器属性屏幕, 232, 232 iSCSI 限定名称 (IQN), 77, 237–238, 238 iSCSI (Internet 小型计算机系统接口)
 SAN, 226 别名, 239 成本, 227–228 设备发现, 239–240
 FCoE, 493 个
 网关, 252 个发起方和目标, 228–232, 230–232 个接口, 228–230
 IP 网络注意事项, 232–236, 235–236 iSNS, 247–251, 249 名称, 236–239, 238
 概述, 226–227, 227
 PDU, 227, 235 安全性, 251–252 会话, 240 身份验证, 241–245, 242–245 设备登录, 241, 241 完整功能阶段, 246–247, 246
 通过 WAN 和 LAN, 234
 ISID (发起者会话 ID), 247
 ISL (交换机间链路)
 概述, 197–199, 199 中继, 197 *isllshow*
 命令, 198 iSNS (Internet 存储名称服务), 247 发现域, 248–250, 249 状态更改通知, 250–251
 iSNS 管理信息库 (iSNSMIB), 251 iSNS 协议 (iSNP), 248
 IVR (VSAN 间路由), 204

J

光纤护套, 179–180, 180 基于日志的复制, 92, 308–310, 310 日志卷, 92, 308–310

JPEG 图像压缩, 367 巨型帧, 277, 492

钾

兰迪·卡茨 130
 Kerberos v5, 261
 Amazon S3 中的密钥, 284

大号

L1 缓存, 83
 L2 缓存, 83
 LACP (链路聚合控制协议), 277
 基于 LAN 的备份, 397, 398
 无需 LAN 的备份, 397–398, 398
 主板上的 LAN (LOM)
 FCoE, 494
 光纤 SAN, 172 语义绑定, 519
 LAN (局域网) 定义, 260 iSCSI SAN, 234 大型文件副本, ODX, 126, 126
 激光优化多模光纤, 181 延迟块大小, 141 网络导致, 443 性能管理, 437–439 位置, 28 复制, 88, 295–296, 298 旋转, 29
 存储阵列, 68–69
 LBA (逻辑块寻址), 20–21, 21
 LC (透明连接器), 183–184, 184
 最小队列深度负载平衡, 72 最近最少使用 (LRU)
 排队算法, 82–83
 Lempel-Ziv-Oberhumer (LZO) 压缩算法, 110 自动分层限制, 383
 线性磁带开放 (LTO) 格式, 407–409
 链路聚合控制协议 (LACP), 277
 链路层发现协议 (LLDP), 491 个链路
 ISL, 197–199, 199 复制, 88–89, 296–297
 LIP (循环初始化原语) 风暴, 196 列出磁盘命令, 219–220
 LLDP (链路层发现协议), 491 多路径 I/O 中的负载平衡, 72–73, 73 局域网 (LAN) 定义, 260 iSCSI SAN, 234 本地快照, 96–102, 96, 314, 314 应用程序感知, 101–102, 324 基于阵列, 97–102, 318–319 写时复制, 101, 324–325
 崩溃一致性和应用程序感知, 100 扩展大小, 98–99, 322–323 完整克隆, 98, 320–322, 322 基于主机, 101–102, 326–327, 327 基于虚拟机管理程序, 315–318, 316–318
 写入时重定向, 101, 325–326 保留空间, 99, 323

节省空间,97–98,319–320,319–320 分层,100,323–324 SAN 拓扑中的局部性,215–216 对象的位置,280 存储阵列中的锁定,69 锁定,ATS,122 逻辑块寻址 (LBA),20–21,21 逻辑块保护,16 RAID 中的逻辑驱动器,131–132,131 逻辑 SAN 组件

结构服务,186–187 FC 结构,185–186,185 FCP 堆栈,184–185,185 分区。请参阅 SAN 中的分区 Oracle Data Guard 中的逻辑备用数据库,301 逻辑存储区域网络 (LSAN) 区域,204 逻辑单元号 (LUN) 群集配置,79 基于控制器的虚拟化,341–343 双控制器体系结构,61–62 光纤 SAN,172,178 掩蔽,77–78 MPIO,74 SCSI,56 无服务器备份,399–400 共享,78–79 存储阵列,76–79 厚和薄,79 真正的主动/主动所有权,62 虚拟,332 逻辑卷管理器 (LVM) 基于主机的虚拟化,334–335,335 RAID,132 复制,303–304 快照,101–102 登录服务器,结构,187 登录 iSCSI,241,241 SAN 分区,187–188 LOM (主板上的 LAN)

FCoE,494 光纤 SAN,172 长距离,流量控制,213–214 循环初始化原语 (LIP) 风暴,196 云存储中的松散一致性,514 FCoE 中的无损网络功能,494–495,495 有损压缩,367 有损网络,484–485 低级格式,18 LRU (最近最少使用) 排队算法,82–83 LSAN (逻辑存储区域网络) 区域,204 LTO (线性磁带开放) 格式,407–409 透明连接器 (LC),183–184,184 LUN。请参阅逻辑单元号 (LUN)

lvcreate 命令
RAID 0,147
RAID 1,150
lvdisplay 命令,150
LVM (逻辑卷管理器) 基于主机的虚拟化,334–335,335 RAID,132 复制,303–304 快照,101–102 LZO (Lempel-Ziv-Oberhumer) 压缩算法,110

米

M&E (机械和电气) 子系统,469 磁带,10,46。另请参阅磁带备份 MAID (大量空闲磁盘阵列) 技术,418 主分布区 (MDA),477 SNMP 中的受管设备,457 管理 NAS 阵列,270 性能。请参阅性能存储。请参阅存储管理 管理信息库 (MIB),251 iSNS 中的管理 SCN,250 结构管理服务器,187 iSNS 中的管理站,251 MAN (城域网),260 手动设备发现,240 手动 VM 快照,317–318,317 多对多操作,大规模并行,162 映射驱动器,258 屏蔽,LUN,77–78 大规模空闲磁盘阵列 (MAID) 技术,418 大规模并行重建,162 主服务器,394 FC SAN 中的配对,183 最大传输速率,34 最大传输单元 (MTU) 大小,236 MBps (兆字节/秒),33–34,440–441 MDA (主分布区),477 平均故障间隔时间 (MTBF),35 测量成本,51 机械和电气 (M&E) 子系统,469 机械磁盘驱动器结构,12,12 执行器组件,21 控制器,22 柱面,19–20,20 逻辑块寻址,20–21,21 盘片,13,13 读/写头,13–15,14–15

主轴,21,22 磁道和扇区,
16–19,16–17,19 容量,25–26 接口和协议,22–24 概述,11–
12 性能 IOPS,30–34,31
旋转延迟,29 寻道时间,28–29 可靠性,34–35 大小,
25–26 速度,26–27 机械存
储,4 存档介质,418–
419 媒体服务器,395 兆字
节/秒 (Mbps),33–34,440–441 内
存描述,4 闪存。请参阅闪存
易失性,11 内存传输缓冲区,
309 合并结构,186
网状拓扑,201–202,
202 元数据

MPIO (多路径 I/O) 软件,72 性能,444 存储阵
列,71–74,72–74
MTBF (平均故障间隔时间),35
MTU (最大传输单元) 大小,236 多光纤终端推入式 (MTP) 电缆,
184, 184 NAND 闪存中的多层单元 (MLC),39–40,40
多模光纤,180–181,
180 多协议阵列,272 多路径 I/O (MPIO) 软件,72 性能,444 存储阵
列,71–74,72–
74 多路复用磁带备份,408–409 多协议阵列,
56, 252 多租户云存储,504

否

对象存储,281 写入时重定
向快照,325 域域网 (MAN),260

MIB (管理信息库),存储阵列中的 251 微码,74–75 微秒,28

Microsoft Exchange Server 复制,302–303,302

Microsoft ODX, 125–127, 126
Microsoft StorSimple,523,523
Microsoft Windows 用于 MPIO 负载平衡,72–
73, 73 中端阵列, 61, 64–65
毫秒,28 单模光纤中的最小弯曲半径,
182 RAID 1 中的镜像
对, 148 镜像

RAID 1,148–151,151
RAID-X, 163 远程
复制, 294
镜像缓存,80–82 备份配置错误,
抖动,29
NAND 闪存中的 MLC (多层单元),39–40,40 监控自
动分层,114,381–382,382

Morle, James, NAS
中有 441 个挂载点,其中 257–258
个挂载
NFS 共享, 262 远程
CIFS 共享, 265

N_Port ID 虚拟化 (NPIV),208–210,209 N_Port ID,206–208 FC SAN 中的
N_Port 模式,175 N_Port 虚拟化
器 (NPV),210–211 光纤 SAN 中的 N_Port,171
NAA (网络地址授权),239 结构的名称服务器,187 iSCSI
中的名称,236–239,238 命名空间对象,278,
279 单个全局,273–274,274 命名授权,iQN,238 NAND 闪存,37–
38 多层单元,39–40,40 单层单元,38–39,39 三
层单元,40–41,41 NAS。请参阅网络附加存储
(NAS)

NAS 蔓延, 272 光纤通
道 SAN 中的本机模式, 177 NDMP (网络数据管理协
议), 400–401, 401 近线存储, 368 近线 SAS (NL-SAS) 驱动器,
24 net use 命令,
265 NetApp Filer, 271 网络
地址授权机构 (NAA), 239 网络连接存储 (NAS),
57, 256 阵列, 55, 267–270,
269

横向扩展,272–275, 274 传统,
271–272 统一, 272 集群,
273 与 DAS 相
比, 271 快速文件
克隆, 125

完整文件克隆, 125 网关设备, 275–276, 276 网关, 268 概述, 256–258, 257 性能, 277–278 协议, 258–260, 260 与 SAN, 271 SMB/CIFS, 263–266 VAAI 卸载, 121 网络数据管理协议 (NDMP), 400–401, 401 网络文件系统 (NFS), 258–259

NFSv3, 259, 260
NFSv4, 259–261, 261 共享
导出和装载, 262
网络性能中的网络跳跃, 443 网络接口卡 (NIC), 494 网络管理系统 (NMS), 457 网络往返时间 (RTT), 87 网络融合。请参阅融合网络 LAN 定义, 260 iSCSI SAN over, 234 NAS 性能, 277–278 门户 iSCSI, 243 iSNS, 250 复制链接, 296–297 NFS (网络文件系统), 258–259

NFSv3, 259, 260
NFSv4, 259–261, 261 共享
导出和装载, 262 nfsd 守护进程, 262

NFSv4 安全性, 261
NIC (网络接口卡), 494 19 英寸机架, 463, 463
NL-SAS (近线 SAS)驱动器, 24
NMS (网络管理系统), 457 节点端口, 171, 206–211, 209 SNMP 中的节点, 457 噪声, 盘片, 13 非持久性存储, 3–4 非主存储, 368 档案不可否认性, 416 非易失性随机存取存储器 (NVRAM), 82 非易失性存储, 3–4, 11

NOR 闪存, 37
Nørgaard, Mogens, 441

NPIV (N_Port ID 虚拟化), 208–210, 209
NPV (N_Port 虚拟化器), 210–211 nsshow 命令, 206
NTFS 权限, 266
NVRAM (非易失性随机存取存储器), 82

哦

基于对象的 RAID, 161 基于对象的重建, 163 对象 ID (OID), 279–280 对象存储 Amazon S3, 284 API 访问, 281–282 架构, 282–283, 282 合规性 存档和内存可寻址存储, 285–286 元数据, 281 OID, 279–280 OpenStack Swift, 284–285 概述, 278–280, 279 公有云 和私有云, 283 安全性, 284 摘要, 286 对象存储设备 (OSD), 278 对象概述, 256–258 奇校验, 137 ODX (卸载数据传输), 125–127, 126 场外私有云, 510, 522 离线备份, 397 离线存储, 368 卸载数据传输 (ODX), 125–127, 126 OID (对象 ID), 279–280 老式空间 回收, 106 本地私有云, 510, 522 联机备份, 396 联机备件, 141 OpenStack Swift, 283–285 运营成本, 51 光纤, 178–179 多模式, 180–181, 180 单模式, 182 优化, 容量。请参阅容量 Oracle Data Guard, 300–301, 301 组织唯一标识符 (OUI), 239 OSD (对象存储设备), 278 带外管理, 452–453, 453 带外虚拟化, 336–337, 337 重叠 I/O 块, 141

过度配置容量, 105–
 106, 429–431, 430 财务优势, 104 闪存, 42
 风险, 104–105 空间回收, 107
 存储阵列, 104–106,
 105 精简配置, 359–
 360, 428, 429 趋势, 105, 105

SAN 拓扑中的超额认购, 215

磷

P/E (编程/擦除)周期
 NAND 闪存, 39 写入耐久性, 43

页面

闪存, 37, 37 精简配置, 103 闪
 存中的并行性, 43 奇偶校验
 RAID, 137–139, 138–139 RAID 5, 153–157,
 155–157
 RAID 6, 157 RAID-X, 163 重建, 143
 分区, RAID, 132 双控制器架构中的被动
 控制器, 61 pathinfo
 工具, 219 Patterson,
 David, 130 PAUSE
 条件, 490, 490, 495, 495 FC 帧
 中的有效载荷, 212 支付卡行业数据安全标准 (PCI
 DSS), 417 PCIe 卡/固态
 卡, 35, 36 PDU (配电单
 元), 468–469, 468 PDU (协
 议数据单元), 227, 235 每个优先级暂停, 489 Perfmon
 工具, 444–446, 444–446 性能, 4–5, 436
 基准, 436–437 缓存, 80, 442–443 容量优化, 355 块大小, 141 压缩,
 110, 369 IOPS, 440
 iostat 工具, 447 延迟/响应时间, 437–439 MBps 和
 传输速率, 440–441 机械驱动器。请参阅机械磁盘驱动器

多路径, 444
美国国家科学院院刊, 277–278
 NAS 阵列, 270 网络跳
 数, 443
 Perfmon 工具, 444–446, 444–446 池化,
 108 公共云存储,
 519–520
 RAID 因素, 149, 441–442 重建, 142–
 143 存储阵列, 68 精简
 LUN, 443 精简配置,
 103–104

IQN 中的句点 (.), 237–238 SMB/CIFS
 中的权限, 265–267 持久性存储, 3–4, 11, 82
 PFC (基于优先级的流量控制), 488
 PFC PAUSE 条件, 490, 490, 495, 495 过度配置中
 的物理容量, 429–431, 430 FC SAN 中的物理组件, 171–
 172 Oracle Data Guard 中的物理备用数据库, 301 流水线, 264
 PIT (时间点)副本, 96, 314 规划数据中心布线, 474 盘
 片, 13, 13 PLOGI (端口登录)过程, 188 NPV 交换机,
 210–211 SAN
 分区, 187–188 时间
 点 (PIT) 副本, 96, 314 时间点快照, 90 点对点
 布线, 476 点对点 SAN 拓扑, 194–196, 195
 基于指针的快照,
 97, 319, 319 内容交付网络中的接入点
 (POP), 521 策略自动分层, 115,
 383 保留, 413–415 池化存
 储阵列, 108–109, 109 内容交付网络中的 POP
 (接入点), 521 端口登录 (PLOGI)
 过程, 188 NPV 交换机, 210–211
 SAN 分区, 187–188 端口映射器服务, 259 FC SAN 中的端口
 速度, 176 端口全球名称 (pWWN), 205 SAN 中的
 端口分区, 192–193 iSNS 中的门户组, 250 交换机间链路的
 PortChannels,
 197
 porterrshow 命令, 222
 PortFast 工具, 234

端口错

误计数器,221–224
 光纤 SAN, 171, 175–176, 176
 N_Port ID,208 存储
 阵列,70–72,71–72,84
portstasshow 命令, 223–224 portzoneshop
 命令, 193 位置延迟, 28, 141 后处理
 压缩, 110, 369 后处理重复数据删
 除, 112, 376, 433 后处理空间回收, 107

电源 数据

中心, 467–469, 468–469 存储阵列, 85 配电单元
 (PDU), 468–469, 468 电源
 使用效率 (PUE) 等级, 467 预测性故障, 143 预取, 缓存, 82 闪存性价比, 44
 具有本地快照的主阵列, 314 Oracle Data Guard 中的主数据库, 300
 主存储, 109 重复数据删除, 378–
 379 描述, 368

RAID 适用于,165

具有本地快照的主卷,96,314 FC SAN 中的主要交换机,176–
 177 优先级自动分层,383 增强型以太网,488–490,
 489–490 重
 建,143,144 基于优先
 级的流量控制 (PFC),488 私有云,283,510–511,
 522 存储阵列的处理
 器,58,59 编程/擦除 (P/E) 周期

NAND 闪存,39 写入耐久性,43

闪存中的程序操作,38 保护缓存,81–82 协议数据单
 元 (PDU),227,235 协议。请参
 阅接口和协议 云存储中的提供商,503 已配置容
 量,429–431,430 公共云,504–505,514–518,515–
 518 对象存储,283 性能,519–520 PUE
 (电源使用效率)评级,467 采购订单触发点,
 361–363 PUT 命令,281 pvcreate 命令 RAID 0,146
 RAID 1,149 pWWN (端
 口全球名称) ,205

问

队列,磁盘驱动器,24,32
 快速连接屏幕, 230–231, 231 静止虚拟机, 315
 容量管理中的配额, 433–
 435

R

R/W (读/写)头,13–15,14–15 机架单元 (RU),463 机架
 冷却,470–471 数据中心,462–
 464,
 463–464 服务器安装,
 472–474

RAID (廉价冗余阵列)

磁盘) ,130 控
 制器,144–145 擦除代码,
 164–165 未来,160–161 组,
 135,135 硬件和软
 件,132–134 历史和原
 因,130–131 热备用和重建,141–143,144

IBM XIV 和 RAID-X, 161–163 概述, 131–
 132, 131 奇偶校验, 137–139,
 138–139 性能, 441–442

RAID 0,145–147,148
 RAID 1,148–151,151
 RAID 1+0 (RAID 10),151–153,152
 RAID 2,159
 RAID 3,159
 RAID 4,159–160,160
 RAID 5,153–157,155–157
 RAID 6,157–159
 RAID 7,164
 RAID-TM,164
 RAIN,164 条
 纹和条纹大小,139–141,140 条纹,136–137,136

ZFS 和 RAID-Z,163–164
 RAIN (廉价节点冗余/可靠阵列) ,164,275 数据中心的高架
 地板,480 随机存取
 存储器 (RAM),4 随机 IOPS,31–32,31 随
 机读取,18

快速生成树协议 (RSTP),234 原始块设备,256

RDP (远程桌面协议) ,507 读取缓存,80

- 读取命中, 81
 读取未命中, 81 读
 取操作闪存, 38
- IOPS, 31–33, 31 盘
 片, 15 读/写
 (R/W) 磁头, 13–15, 14–15 重新平衡存储阵列, 109 重建 RAID, 141–143, 144 恢复, 290 备份, 416 概述, 393–394
- RAID-Z, 164 恢复
 复点目标 (RPO) 异步复制, 91, 308–309
 备份保留策略, 415 业务连续性计划, 292, 389–390 同步复制, 88 恢复时间目标 (RTO)
- 备份保留策略, 415 业务连续性计划, 292, 390 写入时重定向 (RoW)
- 快照, 101, 325–326 ZFS, 163
 冗余多路
 径 I/O, 71
 SAN 拓扑, 194 存储
 阵列, 59, 84 廉价磁盘冗余
 余阵列。请参阅 RAID (廉价磁盘冗余阵列) 廉价节点冗余/可靠阵列 (RAIN), 164, 275
 引用位置, 81 注册状态更改通知 (RSCN) 结构, 187 WWPN/N, 208 业务连续性演练, 86, 291 可靠性 磁盘驱动器, 34–35 闪存, 41 存储阵列, 68 远程 CIFS 共享挂载, 265 远程桌面连接对话框, 507, 507 远程桌面协议 (RDP), 507 远程复制, 115, 384 NAS 中的重新解析点, 257 重复信号, 478 副本, 85 复制, 290 基于应用
 程序, 86 应用层, 299–301 基于阵列, 86, 306–310, 310
- 异步基于日志, 92, 308–310, 310 概述, 298–299, 299 基于快照, 307–308 存储阵列, 89–92, 90, 93 自动分层, 384 业务连续性, 290–293 基于主机
 和基于虚拟机管理程序, 86–87, 304–306 本地快照。
 请参阅本地快照 基于 LVM, 303–304 Microsoft Exchange Server, 302–303, 302 Oracle Data Guard, 300–301, 301 概述, 294 基于存储阵列, 306 同步, 87–88, 87 距离和延迟, 88, 295–296 链接注意事项, 88–89, 89, 296–297, 297 拓扑, 93 概述, 310–311 三站
 点级联, 93–94, 93, 311–312, 311 三站点多目标, 94, 95, 312, 312 三站点三角, 94–96, 95, 313, 313 报告, 容量, 424–426, 425–427 数据表示(ROD) 令牌, 126–127 重新保护 RAID-X, 161–163 VSS 中的请求者, 405 重新扫描命令, 220 重新扫描 SCSI 总线, 219–220 保留快照空间, 99, 323 弹性业务连续性计划, 293 奇偶校验, 137 响应时间, 437–439 RESTful API, 281, 518 恢复
- 测试, 394 用户请求, 416 备份保留, 390, 413–415 IQN 中的反向格式, 238 每分钟转数 (RPM) 磁盘驱动器, 27 固态介质, 36
- 盘片刚性, 13 环形拓扑, 201
 过度配置的风险, 104–105
 ROD (数据表示)令牌, 126–127 根压缩, 263 旋转延迟, 29
- 循环负载平衡, 72 往返时间 (RTT), 87
 FC 中的路由, 202–204, 203

- RoW (写入时重定向)
 - 快照,101,325–326
 - ZFS,数据
 - 中心中 163 行,465–466,466
 - RPM (每分钟转数) 磁盘驱动器,27 固态介质,36
 - RPO (恢复点目标), 88, 91
 - 异步复制, 91, 308–309 备份保留策略, 415 业务连续性计划, 292, 389–390 同步
 - 复制, 88
 - RSCN (注册状态变化通知)结构,187
 - WWPN/N,208
 - RSTP (快速生成树协议), 234
 - RTO (恢复时间目标)
 - 备份保留政策, 415 业务连续性计划, 292, 390
 - RTT (往返时间), 87 数据中心的行为准则,480
 - RU (机架单元),463
-

年代

- S3 (简单存储服务)
 - API,519 对象存储,284 公共云存储, 515–518,515–518 SaaS (存储即服务), 511 SAN 存储, 56–57 阵列,55 备份,397–398,398 FC。请参阅光纤通道 SAN iSCSI。请参阅 iSCSI
 - (Internet 小型计算机系统接口) SAN 与 NAS,271 SAN 连接连接, 71,71 萨班斯-奥克斯利 (SOX) 法案,285,416–417 SAS (串行连接 SCSI)驱动器,16,23–24
 - SATA (串行高级技术附件)驱动器,16,23 SC (标准连接器), 183–184,184 横向扩展架构,62 功能,63 定义,62 NAS 阵列, 272–275,274
 - 扩展 NAS 阵列, 272
- 调度自动分层, 382 备份,390–393, 391–392 分层操作,114–115 SCN (状态更改通知),250–251 SCP (安全复制),268 SCSI (小型计算机系统接口),486 总线重新扫描, 219–220 封装,492,492 光纤 SAN,169 iSCSI。请参阅 iSCSI (Internet 小型计算机系统接口)
 - SDN 逻辑单元号,56 UNMAP 标准, 365–366 SCSI 附件服务 (SES), 56 SDN (软件定义网络),268 SDS (软件定义存储),268,347–348 未来,351 概述,349–351,349 传统,348 无缝控制器故障,62–63 扇区,磁盘驱动器,16–19,16–17,19 安全复制 (SCP), 268 安全 Shell (SSH),454–455 安全 Shell FTP (SFTP),267 安全性 云存储,505 数据中心,479 iSCSI,251–252 NFSv4,261 对象存储,284 寻道时间
 - 磁盘驱动器, 28–29
 - 固态介质, 36 分段结构, 186
 - 选择备份配置屏幕,391,391
 - 选择目标磁盘屏幕,392
 - 选择目的地类型屏幕,392
 - 选择项目屏幕,391,391
 - 自我监测、分析和报告技术 (SMART),46–47
 - SendTargets 在 iSCSI 设备发现中, 240 SAN 流控制中的序列, 212 顺序 IOPS, 31–32, 31
 - 顺序读取, 18
 - 串行高级技术附件 (SATA) 驱动器, 16, 23
 - 串行连接 SCSI (SAS) 驱动器, 16, 23–24
 - iSCSI SAN 中的服务器管理器, 230, 230
 - 服务器消息块 (SMB) 协议, 55, 263–266

- 无服务器备份, 398–400, 399
- 服务器
 - 备份, 394–396, 395
 - CIFS, 263 数
 - 据中心安装, 472–474 结构, 187
 - 光纤 SAN, 169 文件, 269–271 iSNS, 250 主, 394 介质, 395
 - SNMP, 457 个服务级别协议 (SLA)
 - 业务连续性计划, 293
 - RPO, 389 云存储服务提供商, 503 服务业务连续性计划, 293 结构, 186–187
 - 伺服信号, 18
 - SES (SCSI 机箱服务), 56 会话 ID (SSID), 247
 - 会话, iSCSI, 240 身份验证, 241–245, 242–245 设备登录, 241, 241 完整功能阶段, 246–247, 246
- SFP (小型可插拔), 172, 173, 182, 183 sfpshow 命令, 175
- SFTP (安全外壳 FTP), 267
- 用于虚拟化的共享存储模型 (SSM), 332–334, 333 基于主机, 334–336, 335 带内和带外, 336–337, 337 基于网络, 336 基于存储, 336 共享
- CIFS, 265–266
- NFS, 262 权限, 266 存储阵列, 76–79
- 共享 LUN, 78–79 屏蔽双绞线 (STP) 电缆, 478 磁带备份中的擦鞋, 408 短行程磁盘, 18 show port-channel database 命令, 198 show tech-support brief 命令, 220–221 showback 在容量管理中, 435–436 复制中的 sidefiles, 309
- 多模光纤中的信号损失, 180 简单名称服务器 (SNS), 187
- 简单网络管理协议 (SNMP), 457
- 简单存储服务 (S3)
 - API, 519 对象存储, 284 公共云存储, 515–518, 515–518 单一融合数据中心网络, 486 增强型以太网, 487–491, 489–490
 - FCoE, 491–498 存储
 - 网络要求, 486–487
 - 单个全局命名空间, 273–274, 274 单个启动器分区, 191 单实例存储 (SIS), 371–372 单实例, 111, 286
 - 单个大型昂贵驱动器 (SLED), 130 NAND 闪存中的单级单元 (SLC), 38–39, 39 单模光纤, 182 SIS (单实例存储), 371–372 站点恢复管理器 (SRM), 87, 304, 306
 - 62.5 微米多模光纤, 180–181 大小磁盘驱动器, 25–26 扩展。请参阅扩展大小层, 113–114, 114 SLA (服务级别协议)
- 业务连续性计划, 293
- RPO, 389
- NAND 闪存中的 SLC (单层单元), 38–39, 39
- SLDC (流无损数据压缩), 409
- SLED (单个、大型、昂贵的驱动器), 130 slotshow 命令, 174
- 小型阵列, 自动分层, 115
- 小型计算机系统接口 (SCSI), 486 小型可插拔 (SFP), 172, 173, 182, 183 小型办公室, 云存储, 521
- SMART (自我监测、分析和报告技术), 46–47
- SMB (服务器消息块) 协议, 55, 263–266
- SMB/CIFS 权限, 265–267
- SMB Direct 功能, 264
- SMI-S (存储管理计划规格), 451, 456–457 盘片光滑度, 13

基于主机的快照

照,326–327,327 本地。请参阅本地快照
照
SNIA (SGO-001) 考试,7 SNIA (存储
网络行业协会)
虚拟化的共享存储模型,332–334,333 基于主机,334–336,
335 带内和带外,
336–337,337 基于网络,336 基于存储,
336 XAM API,279 SNMP (简单网络管理协议),457 SNS (简
单名称服务器),187 FC 帧中的
SOF (帧起始),212 软分区,
193 软件定义网络
(SDN),268 软件定义存储 (SDS),268,347–348 未来,351
概述,349–351,349
传统,348 软件启动器,229 软件 RAID,132–
134 固态卡 (SSC),35,36 固态介质,35–36,36 描述,10 闪
存。请参见闪存 随机访问,
31 固态硬盘 描述,35 碎片,50 固态存储 (SSS),4–5 本地快照的
源阵列,314 基于源的重复数据删除,374 源本地快照,314 三站点级联拓
扑中的源站点,
311 本地快照的源卷,96,314 SOX (萨
班斯·奥克斯利)法案,
285,416–417 节省空间的快照,97–
98,319–320,319–320 空间回收方法,
106–107 精简配置,363–365 生成树协议
(STP) 禁用,234 以太网,201 FCoE 交换机,
496 备用进程,142 本
地快照中的稀疏文件,316

指定备份时间屏幕,392,392 速度

光纤通道 SAN 端口,176 前端端口,74 机
械驱动器,26–27 主轴,21,
22 拼接光纤电缆,184
SRM (站点恢复管理器),87,304,306
SRM (存储资源管理)应用程序,451–452 工具,451–
452
SSC (固态卡),35,36
SSD (固态硬盘) 描述,35 碎片,50

SSH (安全外壳),454–455

SSID (会话 ID),247
用于虚拟化的 SSM (共享存储模型),332–334,333 基于主机,334–
336,335 带内和带
外,336–337,337 基于网络,336 基于
存储,336

SSS (固态存储),4–5 标准连接器 (SC),
183–184,184 事实上和法律上的标准,524 Oracle Data
Guard 中的备用数据库,300–301 FC 帧中的帧起始
(SOF),212 状态更改通知 (SCN),250–251 闪存中的稳态数字,45 存储
阵列,54 全闪存,65–67 架构双控制器,60–62 网格存储,
62–63 概述,58,59 冗余,59 组件,69,70 后端,83–85 缓存,79–
83

CPU,74–76 前端,
69–74,71–74

LUN、卷和共享,76–79 电源,85 缺点,68–69 企业
级,63–64

FC SAN,178 智能,
85 自动分层,112–
115,114 业务连续性,85–86

压缩,109–111 重复数据删除,111–113 硬件卸载,121–127,123–124,126 本地快照,96–102,96 过度配置,104–106,105 池化和宽条带化,108–109,109 复制。请参阅复制 空间回收,106–107 精简配置,102–104 虚拟化。请参阅虚拟化 限制,99–100 中端,64–65 NAS 存储,57 概述,54–56,54 优点,67–68 机架,464 复制。请参阅 SAN 存储,56–57 统一存储,58 存储即服务 (SaaS),511 基于存储的虚拟化,336 存储设备 缓存,46 成本,50–51 碎片化,47–50 混合驱动器,45 机械磁盘驱动器。参见机械

存储网络行业协会 (SNIA)
用于虚拟化的共享存储模型,332–334,333 基于主机,334–336,335 带内和带外,336–337,337 基于网络,336 基于存储,336 XAM API,279 LUN 上的存储池,332 存储资源管理 (SRM) 应用程序,451–452 工具,451–452 存储空间,132–134 存储虚拟化。
请参阅虚拟化 StorSimple 系统,523,523 STP (屏蔽双绞线) 电缆,478 STP (生成树协议) 禁用,234 以太网,201 FCoE 交换机,496 STR (持续传输速率),34 存储阵列中的闲置容量,67 流式无损数据压缩 (SLDC),409 RAID 中的条带化,136–137,136 条带和条带大小,139–141,140 无奇偶校验的条带化,145–147,148

磁盘驱动器概述
10–11 SMART,46–47 固态介质。请参阅固态介质磁带,46 存储虚拟机管理程序,341 存储管理警报,448–450 容量。请参阅容量 命令行界面,454–455 图形用户界面,455–456 概述,422 性能。请参阅性能 协议和接口,452–457,453–454 单一融合数据中心网络,486–487 SMI-S,456–457 SNMP,457 存储管理计划规范 (SMI-S),451,456–457

结构化布线,475–476,476 结构化数据,269 子 LUN 自动分层,112–113,380–381 重复数据删除中的子文件,371 supportshow 命令,220 持续传输速率 (STR),34 FC SAN 中的交换机端口,175,176 交换机配置备份,217 配置转储,220–221
FCoE,496,496 光纤 SAN,171、173–178、175、196–197 Oracle Data Guard 中的切换,297 switchshow 命令,205 同步重复数据删除,375–376 同步复制,87–88,87
距离和延迟,88,295–296 链接注意事项,88–89,89,296–297,297 合成完整备份,404 sysmon 工具,446

电视

T10 UNMAP 技术,365–366
T11 委员会

FCoE,491

网络地址授权,239 磁带备份,406 优点,407
压缩,409 缺点,407 加密,409

LTO 格式,407–409 多路复用,
408–409 擦鞋,408 虚拟磁带库,
410–411 磁带存储,10,
46 基于目标的重复数据删除,375 目标门

户组标签 (TPGT),247,250
目标备份,395 iSCSI SAN,228–232,230–
232

SAN 区域,189 三站
点级联拓扑,311 虚拟化,337,340

TCP/IP 协议,233

TCP 卸载引擎 (TOE),229

电信行业协会
(TIA),461,476–477
Telnet,454–455
TB,26 测试布
线,475
更改,358–359,
383–384,405 恢复操作,394 密集 LUN,
79 密集卷,102,356–357 精简 LUN,79,443 精简配置 (TP),102–104,356

容量管理,427–431,429–430 关注点,366 描述,355
过度配置,359–
360,428,429 空间
回收,363–365 厚卷,356–357 薄卷,357–358
趋势,360–361,361,428–430,429
采购订单触发点,361–363

UNMAP SCSI 标准,365–366

精简配置 stun (TP Stun),124–125 精简卷,103–104
抖动,磁头,28–29 3.5 英寸驱动器,25–26 三站点级联拓扑,
93–94,93,311–312,311 三站点多目标拓扑,94,95,312,312 三站点三角拓扑,
94–96,95,313,313 TIA (电信行业协会),461 用于布线的 TIA-942,476–477
第 0 层和第 1 层性能,67 分层拓扑,202 自动分层。请参阅自动分层核心边缘,202 数据中心,461–462 说明,355
快照,100,323–324 结构时间服务器,187 NAND 闪存中的 TLC
(三级单元),40–41,41 TOE (TCP 卸载引擎),229 复制中的拓扑,93
概述,310–311 三站点级联,93–94,93,311–312,311 三站点多目标,94,95,312,312 三站点三角,94–96,95,313,313 TP。请参阅精简配置 (TP)

TP Stun (精简配置 Stun),124–125 TPGT (目标门户组标签),247,250 磁道 0,19 磁盘驱动器上的磁道,16–19,16–17,19 传输速率最大和持续,34 性能管理,440–441 多链路透明互连 (TRILL),496
SNMP 中的陷阱,457 趋势过度配置,105,105 精简配置,360–361,361,428–430,429

采购订单的触发点,361–363
TRILL (多网透明互联)
连接),496
NAND 闪存中的三级单元 (TLC),40–41,41 故障排除 FC SAN,217

连接问题,217–219

端口错误计数器,221–224

SCSI 总线重新扫描, 219–220 交换机配置
备份, 217 交换机配置转储, 220–221

真正的主动/主动 LUN 所有权,62 中继 Cisco VSAN,
203 交换
 机间链路,197
 trunkshow 命令,198 双轴电
缆,478 2.5 英寸驱动器,25–26 2.5 微
米多模光纤,180–181

乌

U_Port 模式, 175 存储阵
列的 ucode, 75 umount 命令, 262

UNC (通用命名约定) ,263 无争议协议,486 统一结构,487
统一 NAS 阵列,272 统一存储,58 统
一存储阵列,56 不间断电
源 (UPS) ,469

通用命名约定 (UNC), 263
UNMAP 命令, 107, 365–366 卸下机架, 463 不
可靠网络, 485 非屏
蔽双绞线 (UTP) 电缆, 478

UPS (不间断电源) ,469

Uptime Institute 冷

却报告, 470 数据中心层,
461 磁盘驱动器的可用容量,

26

用例

容量优化技术, 356 重复数据删除, 377–379, 377 已用
容量过度配置, 429–431, 430 触发点,

361

用户请求的备份恢复,416

UTP (非屏蔽双绞线) 电缆,478

五

虚拟助手。请参阅 vStorage API for Array Integration (VAAI) 重复
数据删除
中的可变块大小,373 磁带备份保管,407 SDS 中的 vCPU,
350 云存储版本控制,514 垂直布线,477
NAS 阵列的垂直扩展,272
vgcreate 命令 RAID 0,147 RAID 1,149
vgdisplay 命令,147 虚拟结构
Brocade,204 VSAN,203–204 专用虚拟 LAN (VLAN),
233 基于 LAN 的备份,397
 中继,197 虚拟机磁
 盘 (VMDK) 级别,
304–305 虚拟机文件系统 (VMFS),
122 云存储中的虚拟
 机 Amazon Web
 Services,508–510,508–
510 Windows Azure,505–507,
 505–507 全闪存中的
 虚拟服务器环境阵列,67 虚拟存储设
 备 (VSA), 349–350
虚拟磁带库 (VTL), 410–411 虚拟化, 330 阵列功能, 345 延长阵列寿
命, 344 自动分层, 345–346, 345 注意事项, 346–347 基于控制
器, 115–116, 116, 336 配置, 341–344, 342 概
述, 339–341, 339–341

N_Port ID,208–210,209 概述,
330–332,331 规划,347 目
的,338–339

SNIA 共享存储模型, 332–337, 333–335, 337 软件定义存储,
347–351,
349

存储阵列,115–116,116 自动分层,119–120,
120 注意事项,120–121 配置,116–118,117
外部,119 功能,119 虚拟化模
式,117 重复数据删除的虚拟化用例,378 VLAN
(虚拟 LAN)专用,
233 基于 LAN 的备份,397 中
继,197 VMDK (虚拟机磁盘)级别,304–
305 VMFS (虚拟机文件系统),122 VMware 快照,315–318,316–318
VAAI。请参阅 vStorage API for Array
Integration (VAAI) 地
板空隙,464,465 易失性存储,3–4 卷管理
器 基于主机的快照,
326–327,327 基于主机的虚拟化,334–335,335 LVM。请参阅 LVM (逻辑
卷管理器)
卷影复制服务 (VSS),99,405–406 卷 FC SAN,178 本地快照,96 存储阵列,76–79
厚,102,356–
357 薄,103–104 SDS
中的 vRAM,350 VSA (虚拟存
储设备),349–350 Vsphere
Replication,304–306 VSS (卷
影复制服务),99,405–
406 用于阵列集成的 vStorage
API (VAAI),121–122
NAS 快速文件克隆,125 NAS 完整文
件克隆,125 卸载,121–122
ATS, 122 硬
件加速复制, 124, 124 硬件加速归零, 122–123, 123
精简配置关闭, 124–125 零空间回收, 125

VTL (虚拟磁带库),410–411

西

WAFL (任意位置写入文件布局)文件系统,160 向前
行走行为,358 闪存中
的磨损均衡,43–44
众所周知的零令牌 ROD 令牌,129 广域网 (WAN) 定义,260
iSCSI SAN,234 宽条带化,85,108–109,
109 RAID 条带宽
度,140
Windows Azure 云存储,505–507,505–507
Windows Server Backup 工具,390 磁盘驱动
器抖动,27 全球名称 (WWN),205–
206 全球端口名称 (WWPN),77,205–206
WORM (一次写入,多次读取)技术,285,417 闪存写入放大,43
随处写入文件布局 (WAFL) 文件系统,160 写回缓存
描述,80 闪存支持,
133 内部 RAID 控制器,
144–145 虚拟化,
347 写入缓存,80 闪存中
的写入悬崖,38,45 写入合并闪存,44
RAID 5,156 闪存
中的写入组合,44 闪存中的写入耐久性,39,43 写入
意图日志,92,308 写入 IOPS,31–33,31 一次写入,多次
读取 (WORM) 技术,285,417 写入
惩罚奇偶校验,139
RAID,155–157,156–157,442
直写模式,61–62,80 写入闪存,37–38,37 盘
片,15
WWN 分区,192–193
WWN (全球通用名称),205–206
WWPN (全球港口名称),77,205–206

X

XAM (可扩展访问方法)API, 279
XCOPY LITE 原语, 126
XOR (异或)运算奇偶校验, 138–139

RAID 5, 155

是

ZDR (分区数据记录), 17, 17 复制中零数据丢失, 295, 298 零位置延迟, 32

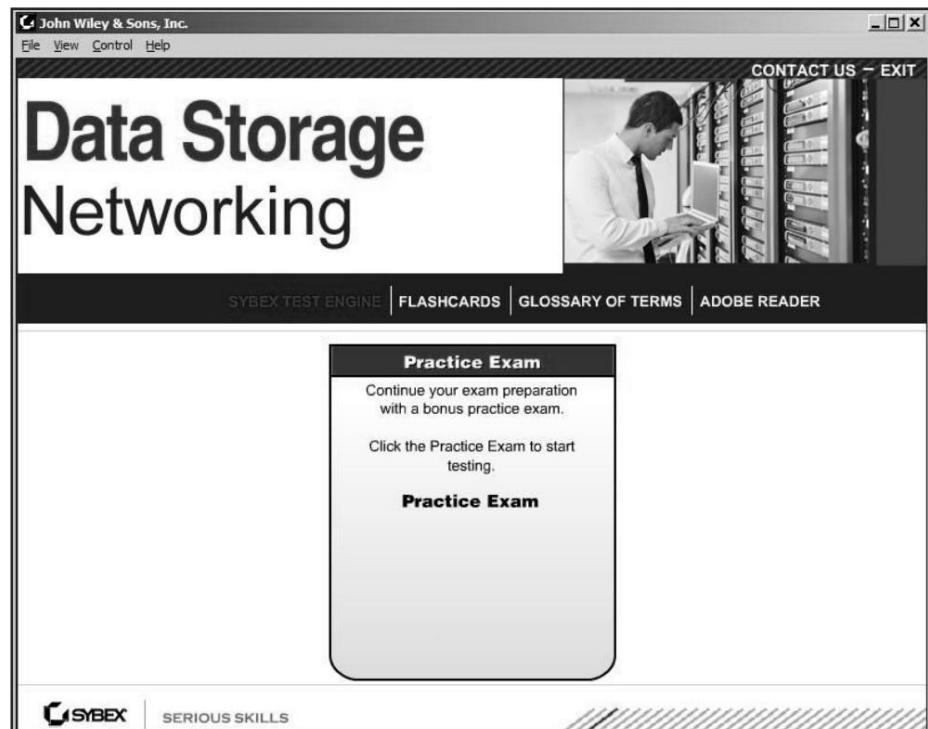
零空间回收, 106, 125, 363–365
ZFS RAID, 163–164 区域集, 189, 191–192 zonecreate 命令, 190 分区数据记录 (ZDR), 17, 17 SAN 中的分区, 187 别名, 192 最佳实践, 191 配置, 190–191
示例, 189,
189 结构和端口登录, 187–188 硬和软, 193 重要性, 188 端口和 WWN, 192–193 区域集, 191–192

Free Online Study Tools

Register on Sybex.com to gain access to a complete set of study tools.

综合学习工具包包括：

- 一个 CompTIA Storage+ practice exam to test your knowledge of the material.
- 电子抽认卡 to reinforce your learning.
- 可搜索的词汇表 gives you instant access to the key terms you'll need to know as a data storage professional.



请访问www.sybex.com/go/datastoragenetworking进行注册并获取此综合学习工具包的访问权限。

上传者:[StormRG]