

致谢



NetApp 专版 (第二版)

闪存存储

FOR
DUMMIES®

学习如何:

- 优化存储性能
- 将服务器闪存用作存储缓存
- 减少存储基础架构在数据中心内的占用空间
- 减少存储基础架构在数据中心内的占用空间

Lawrence C. Miller, CISSP



关于 NetApp

NetApp 致力于提供具有创新性的存储和数据管理解决方案，帮助客户加速实现业务突破，显著降低成本并提高效率。了解 NetApp 倾力帮助全球各地的公司持续快速发展的详情，请访问 <http://www.netapp.com/cn/>。

闪存存储 FOR DUMMIES[®]

NetApp 专版 (第二版)

作者：**Lawrence C. Miller , CISSP**

WILEY

《闪存存储 For Dummies》(Flash Storage For Dummies®), NetApp 专版 (第二版)

出版商

John Wiley & Sons, Inc.

111 River St.

Hoboken, NJ 07030-5774

www.wiley.com

版权所有 © 2014 John Wiley & Sons, Inc., Hoboken, New Jersey

除美国 1976 年版权法第 107 条或第 108 条规定之外, 未经出版商事先书面许可, 不得以电子、机械、影印、录音、扫描或其他任何形式或方式复制或传播本书的任何部分, 也不得存储于检索系统中。如需向出版商请求许可, 请致函 Permissions Department, John Wiley & Sons, Inc., 地址: 111 River Street, Hoboken, NJ 07030, 电话: (201) 748-6011, 传真: (201) 748-6008, 或访问网站 <http://www.wiley.com/go/permissions>。

商标: Wiley、Wiley 标识、For Dummies、Dummies Man 标识、A Reference for the Rest of Us!、The Dummies Way、Dummies.com、Making Everything Easier 和其他相关商业外观是 John Wiley & Sons, Inc. 和/或其子公司在美国和其他国家或地区的商标或注册商标, 未经书面允许, 不得使用。NetApp、NetApp 标识、Go further, faster、Data ONTAP、Flash Cache 和 Flash Pool 是 NetApp, Inc. 在美国和/或其他国家或地区的商标或注册商标。所有其他商标均为其各自所有者的财产。John Wiley & Sons, Inc. 与本书所提到的任何产品或供应商没有关系。

责任限制/免责声明: 出版商和作者不对本书内容的准确性或完整性做任何陈述或担保, 且明确声明不做任何担保, 包括但不限于针对特定用途的适用性的担保。销售或促销材料不产生或延长任何担保。本书所含的建议和策略不一定适合所有情况。出版商保证不借助本书内容提供任何有关法律、会计或其他方面的专业服务。如需专业帮助, 请寻求能够胜任的专业人士的服务。出版商和作者均不为此产生的损害负责。本书提及某个企业或网站作为引用和/或补充信息的潜在来源并不意味着本书作者或出版商认可该企业或网站可能提供的信息或可能给出的建议。此外, 读者应知悉, 本书所列网站可能在本书写成后的时间里发生改变甚至失效。

要了解我们的其他产品和服务以及如何为您的业务或企业定制“傻瓜系列”丛书的一般信息, 请联系我们美国的业务开发部, 电话: 877-409-4177, 电子邮件: info@dummies.biz, 或访问网站 www.wiley.com/go/custompub。要了解为产品或服务获取“傻瓜系列”品牌许可的信息, 请发送电子邮件至 BrandedRights&Licenses@Wiley.com。

ISBN 978-1-118-99423-8 (pbk); ISBN 978-1-118-99626-3 (ebk)

中国印刷

10987654321

内容概览

引言	1
傻瓜式设想	1
关于本书	1
本书中使用的图标	2
阅读顺序	3
第 1 章：闪存存储技术 — 什么是闪存 存储技术以及为什么要选择该技术	5
闪存概念	5
SLC、MLC 和部分 TLC	6
闪存技术的发展	8
闪存在企业中的应用	10
第 2 章：闪存存储技术 — 部署位置和部署方式	17
混合解决方案与纯闪存解决方案	18
存储阵列闪存用作缓存	19
SSD 用作缓存	23
SSD 替代 HDD	24
第 3 章：了解闪存存储技术的优势	27
性能	27
可靠性	28
效率	30
成本	34
第 4 章：闪存在未来存储项目中的十大 (或五大) 使用策略	39

出版商声明

为本书的发行提供帮助的人员包括：

高级项目编辑：Zoë Wykes

项目协调员：Melissa Cossell

编辑经理：Rev Mengle

NetApp 特别协助人员：

策划编辑：Kyle Looper

*Larry Freeman、Paul Feresten、
Darrell Suggs、Duncan Moore、*

高级业务推广代表：

Cesar Orosco、Rip Wilson

Karen L. Hattan

定制出版项目专员：

Michael Sullivan

引言



閃

存技术的出现在存储行业引起了不小的轰动。

相比传统的硬盘驱动器，闪存存储在速度和可靠性上更具优势，它是一项灵活且成本效益日益提升的技术，可用来优化运行任务关键型、I/O 密集型应用程序的企业存储环境。

《傻瓜系列之闪存存储》NetApp 专版第二版将为您介绍闪存存储技术在企业中的多种用途和优势。从闪存加速存储到纯闪存阵列，闪存技术大大提升了存储基础架构的性能和可靠性。同时还降低了数据中心的能耗和空间成本。

傻瓜式设想

曾有人说过，大多数设想皆是徒劳，尽管如此，我还是要做一些设想。作为本书的主要前提，我设想您是一位对存储技术（例如，网络连接存储 (NAS) 和存储区域网络 (SAN)）有一定了解的技术经理或分析师。因此，本书主要为那些正在评估新存储技术的技术人员而著。

关于本书

本书中信息的丰富程度堪与《美国国会记录》或完整版《大英百科全书》等杰作相媲美，由四章精悍内容构成，

将您所需的信息进行提炼，方便您阅读使用。下面简要介绍了各个章节的内容。

第 1 章：闪存存储技术 — 什么是闪存存储技术以及为什么要选择该技术。本章首先介绍闪存技术：什么是闪存技术，该技术的发展历程，以及为什么要考虑在存储基础架构中采用该技术。

第 2 章：闪存存储技术 — 部署位置和部署方式。在本章中，您将了解以下 NetApp 闪存存储解决方案：NetApp Flash Cache、Flash Pool、纯闪存 FAS 以及 EF 系列。

第 3 章：了解闪存存储技术的优势。本章将介绍闪存存储能够为企业带来何种优势。

第 4 章：闪存在未来存储项目中的十大（或五大）使用策略。最后，按照一贯的“傻瓜系列”风格，我将就如何在存储环境中使用闪存提供一些有用的提示。

本书中使用的图标

本书在某些情况下使用了图标，旨在提醒读者对重要信息予以注意。不是笑脸或者其他可爱的表情符号，而是真正会引起注意的图标。以下是将用到的图标。



这个图标表示值得您铭记在心或永久存放于脑海的内容，要像记纪念日和生日一样记住。



如果您一直以来梦寐以求的只是征服魔兽世界的神秘笈，请注意，我们无能为力。此图标说明各类

术语行话，表示内容专门针对精英人士，或者至少是网虫。



感谢阅读，希望您喜欢这本书，敬请关注作者！事实上，这个图标表示一些有用的建议和有益的宝贵信息。

阅读顺序

在此向 Lewis Carroll（刘易斯·卡罗尔）、Alice（爱丽丝）和 Cheshire Cat（笑脸猫）致以歉意：

“可以告诉我该走哪条路吗？”

笑脸猫（也指傻瓜书作者）回答说“这得看你想往哪里去”。

“我并不太在乎往哪去...”，爱丽丝说。

“那么，随便你选择哪一条路都一样！”

用这段对话来形容《傻瓜系列之闪存存储》NetApp 专版第二版再贴切不过了，就像《爱丽丝漫游仙境》(Alice in Wonderland) 那样，本书将成为永恒的经典。

如果您不知道自己要看什么内容，您可以任意挑选一章进行阅读，但我还是建议您从第 1 章开始看起。不过，如果您对特定章节探讨的主题颇感兴趣，可以跳过其他内容直接前往该章节。请按适合自己的顺序阅读本书（但是我不建议前后颠倒或倒着读的读法）。

我保证，您绝不会云里雾里，像爱丽丝那样迷失方向掉进兔子洞。

第1章

闪存存储技术 什么是闪存存储技术以及 为什么要选择该技术

本章提要

- ▶ 了解闪存的工作原理
- ▶ 将闪存提升到新的高度
- ▶ 了解闪存的发展历史
- ▶ 告别 HSM、ILM 和 AST

本

章为您介绍闪存的工作原理，闪存技术如何发展成形乃至成为消费者和企业首选存储技术，以及闪存如何支持新的数据和存储管理策略。

闪存概念

相比传统的硬盘驱动器 (HDD) 将数据存储于带磁涂层的铝制旋转盘片上，闪存存储技术的工作原理截然不同。



闪存技术将信息存储在存储单元阵列中行和列的交汇处。每个存储单元都有两个晶体管，一个浮栅和一个控制栅，两者之间隔着一层薄薄的氧化层。每个单元仅存储一个比特的信息，通过对浮栅晶体管施加电压来确定该单元是代表 1 还是 0。每次进行数据存储时，浮栅晶体管上电荷充放电的过程类似于相机的闪光灯，闪存也因此而得名。多层单元 (MLC) 和三层单元 (TLC)（将在下一节进行介绍）可以通过向浮栅输送多层电压值来存储多个比特的数据。

SLC、MLC 和部分 TLC

闪存技术具有很多优势，但目前仍存在两大缺点：

- ✔ 像所有的电子擦除式可编程只读存储器 (EEPROM) 一样，存储单元内的那些小晶体管实现闪存有些勉为其难。经过多次电子擦除后，这些晶体管栅极就会逐渐断开，最终导致擦除失败。
- ✔ 由于在写入新数据之前您始终需要将闪存内的数据擦除，因此，在某些情况下闪存的速度不会很快。事实上，有些时候基于闪存的固态驱动器 (SSD) 的速度不如高性能 HDD，例如执行大型顺序写入操作时。不过，在随机读取等其他操作方面，SSD 的速度是 HDD 无法比拟的。



Thefreedictionary.com 将 EEPROM（电子擦除式可编程只读存储器）定义为：

一种断电后数据不会丢失的可重写存储芯片。**EEPROM** 以比特或字节为单位写入，也就是说，重新写入数据之前要擦除相应的字节。闪存是由 **EEPROM** 演变而来且在结构上与之大致相同，但其写入之前的擦除必须是一整个字节块。此外，**EEPROM** 通常出现在电路板上用来存储少量的指令和数据，而闪存模块却可以存储数千兆字节的数据，能够代替数码相机存储和硬盘。

目前正在开发的几项新技术将弥补这些缺陷，且可能会为闪存技术将来取代 HDD 助上一臂之力。

例如，单层单元 (SLC) 闪存一直以来都用在企业 SSD 中。SLC 闪存的耐用性比 MLC 闪存高十倍，但成本上要高出很多。

为使 MLC 闪存拥有 SLC 闪存一样的耐用性，同时保持较低的成本，一些存储公司开始提供采用精密的损耗平衡和坏块管理算法的 MLC 闪存。这些创新确保了将成本较低的 MLC 闪存应用到企业 SSD 的同时，不会对可靠性造成任何影响。

TLC 闪存可以提供超大的闪存容量和更低的成本，但在访问速度和可靠性上甚至不及 MLC 闪存。不过，随着数据存储模式算法的精密性越来越高以及耐用性的不断提升，TLC 闪存似乎也开始步入企业 SSD 应用行列。

闪存技术的发展

每一次技术创新都会在整个行业内掀起一场大革命。闪存存储这一突破性创新技术也是如此，它让整个存储行业发生了翻天覆地的变化。



突破性技术是指能够取代当前市场中的早期技术并改变价值网络的创新技术。

下面几节将重点介绍闪存技术的发展历程。

只读存储器

对于所有闪存产品而言（抱歉），闪存技术的起点并不高，归根结底还是属于只读存储器 (ROM)。在集成电路出现的早期（大约在 70 年代），计算机微代码可永久存储在 ROM 芯片中。顾名思义，ROM 芯片中的数据或编程指令会始终保持不变，且无法擦除或重写。如果不需要代码升级的话，ROM 还是很不错的，因为代码升级需要用包含升级代码的较新版本 ROM 芯片集替换原先的 ROM 芯片集。

擦除式可编程只读存储器

擦除式可编程只读存储器 (EPROM) 是由 ROM 演变而来第二代存储器。EPROM 克服了 ROM 永远只能“写入一次”的难题，这为计算和存储行业带来了重大突破。

EPROM 在其硅片上方留有一个透明窗，且该透明窗上覆盖有一层小贴纸，通常用来表示芯片上微代码的修订版本（或版本）。当您需要升级代码时，撕掉贴纸并将芯片放

入能放射紫外线的小盒（计算机工程师专用的微波炉）中即可。将硅片放在紫外线下照射约一个小时就可以擦除其中的数据，且可以使用专用的 EPROM 编程器进行重新编码。

EPROM 中昂贵的存储芯片可以重复使用，但操作还是过于繁琐。第三代存储器就是 EEPROM。对于 EEPROM，只要向芯片中的存储单元施加电压即可擦除该芯片上的数据。事实上，即使不将 EEPROM 从计算机系统的主板中移除，也能对其进行擦除和重新编程。

EEPROM 的用途不仅仅是存储微代码，它们还可像随机存取存储器 (RAM) 芯片那样用作固态存储。实际上，EEPROM 因能够对数据进行存储、擦除再存储而被归类为非易失性随机存取存储器 (NVRAM)。

随着需要大量存储容量的电子消费品（包括数码音乐播放器、相机和录相机）的流行，EEPROM 获得了长足发展。但在 90 年代早期，这些电子消费品唯一可采用的存储技术就是 1 英寸微盘驱动器，众所周知，该技术可靠性低且成本高昂。

Compact Flash

1994，Sandisk 率先推出 Compact Flash 模块。Compact Flash 是电子消费品市场标准封装模块领域的一大创新，该产品即插即用，取代了之前提到的微盘驱动器。Compact Flash 成功的关键在于其具有擦除和重新记录单个存储单元

的功能，不像早期的 **EEPROM** 必须先擦除其中的所有数据才能存储新的数据。

这开创了存储的新时代，而其他的一切，都已成为历史！随着电子消费品需求的飞速增长，闪存技术成本的不断下降引起了存储供应商的注意，这时许多 IT 专业人员不免产生这样的疑问，“闪存是否会彻底取代传统的硬盘驱动器？”

闪存存在企业中的应用

由于企业数据存储容量和性能方面的需求正以一个前所未有的速度飞速增长，存储行业一直奋力进行技术创新，以有效应对当今面临的企业存储挑战。

闪存存储影响企业数据中心的其中一个关键因素在于存储分层，而存储分层这一概念一直反反复复变化不定。在 80 年代初期，存储分层称为分层存储管理 (**HSM**)。到了 90 年代，**HSM** 演变为信息生命周期管理 (**ILM**)。现在，存储分层又有了新的“身份”，称为自动存储分层 (**AST**)。借助可实现整个分层概念变革的全新实施，闪存技术为存储分层注入了新的活力。

尽管名称上有所变化，但基本原理还是一样：存储分层采用过时且运作效率低下的方法将数据从高性能、成本高昂的介质移动到高容量、低成本的介质，以最大限度地提升

存储基础架构的利用率。但是，不管其目前是何称谓，存储分层都令 IT 组织对其有看似永不倦怠的热情。

问题在于目前大部分可用的存储分层解决方案都过于复杂，难以实施，且采用的方法都已过时（请参见图 1-1）。

- 监控多组磁盘块（区块）的 I/O 行为
- 将访问频率与各层平均值进行比较
- 将区块提升或降级至合适的层

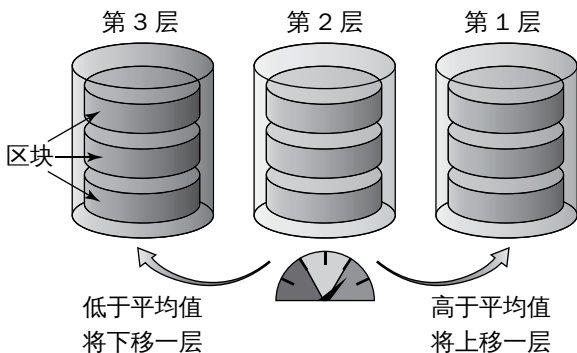


图 1-1：存储分层的工作原理。

存储分层的基本前提是在合适的时间以合理的成本将数据存储到适当的位置，以便为企业提供支持并最有效地利用存储基础架构。

以下三大基本设想是实现存储分层的基础前提：

- ✔ 数据的价值随着时间而降低；据估计，如果数据在 90 天内未被访问，将永远不会被访问。
- ✔ 估计在所有数据中，需要存储在高性能介质（成本相应较高）上的数据不超过 20%。
- ✔ 尽可能将数据尽快移动到大容量串行 ATA (SATA) 驱动器是改善对数据中心中存储基础架构的利用的可行策略。

那么，如今的自动分层解决方案是如何尝试满足这三点要求的？首先，我们来关注“自动”这个词；它看似不错且非常具有吸引力，但它究竟是如何发挥效用的？在 IT 部门实现全面“自动化”之前，大量的工作需要完成。存储架构师或管理员需要收集、分析并设计合理的工作流程以供系统自动运行，换言之，存储架构师的工作十分繁杂。

比方说，在构建合适的解决方案时，需要考虑以下问题：

- ✔ 我们的环境需要多少个存储层？根据驱动器类型、转速和 RAID（独立磁盘冗余阵列）级别，某些存储供应商提供多达九个存储层。
- ✔ 第 1 层应该为多大？第 2、3、4 等其他层又应该是多大？
- ✔ 如何确定哪些数据为“热”数据、“温”数据或“冷”数据（详见第 2 章）？有些自动存储分层实施需要清楚地

了解应用工作负载、其他软件以及不同存储层的具体规划和大小。

- ✓ 哪种类型的数据应进入第 1 层？如何对数据进行分类？
- ✓ 自动分层软件将数据迁移到其他层需要花费多长时间？在某些情况下，需要 3 天的时间来提升数据级别，12 天的时间来使数据降级。
- ✓ 何时将关键数据提升到第 1 层？谨记，数据迁移或重新定位会影响系统性能；需要的时间从几小时到几天不等，具体取决于供应商。
- ✓ 何时将冷数据移动至第 2 层和第 3 层？
- ✓ 数据迁移过程采取手动、自动还是按计划的方式？
- ✓ 数据迁移达到何种粒度？是要移动整个逻辑单元号 (LUN)，还是其中一个子 LUN？
- ✓ 如何知道我们的数据迁移策略、阈值或数据移动的时段是否合适？这需要持续监控或校准。
- ✓ 是否可以在第 1 层存储中使用重复数据删除和精简配置等数据效率功能？
- ✓ 是否需要使用不同的分层解决方案来支持 NAS 和 SAN？
- ✓ 存储分层会向我们的环境添加多少个新工具和管理端点？

✓ 存储分层的成本将是多少？我们需要购买多少个许可证？

能否设计出合适的解决方案取决于能否正确回答以上问题且收集和分析的数据正确无误。最后，即使您的数据和分析都正确，解决方案具体实施起来也会十分复杂。存储分层成功与否关键在于细致的前期工作和运行阶段的不断更新和维护。这一解决方案假设工作负载可预测，因此灵活性不高。

从统一存储入手就没那么复杂，效果也会好很多。统一存储是将相关类型的介质、阵列功能和所有协议结合于一体的单个平台。借助该解决方案，您可以使用软件来构建虚拟存储层 (VST) 策略，此策略利用闪存技术按需优化性能并减少不必要的存储成本。

VST 是存储基础架构的自行管理型数据驱动式服务层（请参见图 1-2）。它可以根据成本和性能要求实时评估工作负载优先级和优化 I/O 请求，无需复杂的数据分类和迁移。

同时，VST 可将热数据提升到闪存存储，不会产生其他存储分层方法提升热数据时产生的数据移动或迁移开销。一旦收到针对启用了 VST 的卷或 LUN 上某个数据块的读取请求，系统就会自动将该数据块列入提升范围（相比其他实施，4 KB 数据块的粒度非常细）。数据块向 VST 的提升并非数据迁移，因为该数据块在复制到 VST 后仍留在硬盘上。

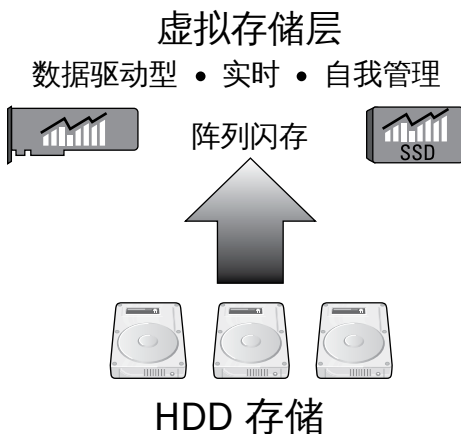


图 1-2：采用闪存技术的虚拟存储分层。

VST 采用了主要的存储效率技术，例如重复数据删除、卷克隆、精简配置、智能缓存和简化管理。您只需为卷或 LUN 选择所需的默认介质层，例如 SATA、光纤通道 (FC) 或串行连接 SCSI (SAS)。卷或 LUN 中的热数据会自动按需（应用驱动型）提升至基于闪存的介质。

VST 与早期存储分层方法的主要区别在于能够创造性地使用闪存技术。第 2 章将为您介绍 NetApp 的闪存技术解决方案以及这些解决方案如何相互配合，为企业提供一个完整的 VST 策略。

第2章

闪存存储技术 — 部署位置和部署方式

本章提要

- ▶ 选择您喜欢的类型：混合或纯闪存
- ▶ 用 SATA 和闪存代替 SAS
- ▶ 深入了解 NetApp Flash Pool
- ▶ 全面了解纯闪存存储

基

于闪存的存储可以提高许多应用（包括数据库、服务器和桌面虚拟化以及云基础架构）的 I/O 性能和效率。第 1 章介绍了什么是闪存技术以及它作为一项创新型企业存储技术如何支持虚拟存储分层等解决方案。在本章中，您将了解 NetApp 的混合和纯闪存阵列解决方案，以及 NetApp 如何将闪存应用到存储分层以最大限度地发挥闪存技术的价值。

混合解决方案与纯闪存解决方案

在存储阵列级别部署闪存技术有以下两种基本选项：

- ✓ 混合
- ✓ 纯闪存

混合存储阵列

混合存储阵列使用迁移或缓存技术将热数据存放到闪存中，实现闪存的性能与 HDD 的容量的完美结合。从性能角度来看，混合存储的主要目的是将尽可能多的随机 I/O 转移到闪存中，以增加 IOPS 并减少给用户和应用程序造成的平均延迟。

NetApp 为其 Data ONTAP 平台（Flash Cache、Flash Pool）和 SANtricity 平台（SSD 缓存）提供混合阵列选项。在这些混合解决方案中，采用相对较少的闪存（通常仅占总存储容量的 1-2%）就能够来非凡的效率和性能优势。

纯闪存阵列

纯闪存阵列可为业务关键型应用程序提供最高性能、快速响应和高度一致性。如果必须以稳定一致、通常低于 1 毫秒 (ms) 的低延迟完成所有 I/O 操作，纯闪存阵列将是最佳选择。它们特别适合性能要求较高且容量要求适中的情况。

NetApp 为 Data ONTAP 和 SANtricity 平台提供具有一系列性能和数据管理功能的纯闪存阵列解决方案。



选择混合阵列还是纯闪存阵列，要根据要部署阵列的应用程序工作负载的性能和效率要求来定。这些要求性能最高、最稳定一致且可预测的应用程序最适合部署纯闪存解决方案。请参见图 2-1。



来源：NetApp

图 2-1：选择混合阵列或纯闪存阵列的标准。

存储阵列闪存用作缓存

网络存储系通常很难满足企业在性能上的需求。扩大这些系统的容量十分容易，只需添加更多磁盘驱动器即可。但这种做法不会加快磁盘驱动器的速度 - 对于 SAS，转数为 10,000 或 15,000；对于 SATA，转数为 7,200。因此，人们通常使用大量小容量的磁盘驱动器来满足很多工作负载对

I/O 吞吐量的需求。这种方法既浪费存储容量、机架空间、电耗，又浪费金钱！



短行程技术时常用来提升 IOPS，其方法是限定磁盘驱动器磁头从磁盘驱动器上的任意点读取数据时必须移动的距离。这种方法缩短了平均寻道时间，但同时也减少了容量。

NetApp Flash Cache 可以提高随机读取密集型工作负载的性能，而无需添加更多高性能的磁盘驱动器。Flash Cache 是一种基于闪存的 PCI-e 扩展卡，适用于 NetApp FAS 存储控制器。Flash Cache 能够提升工作负载的处理速度，例如

- ✓ 业务处理
- ✓ IT 基础架构（文件服务）
- ✓ 协作（电子邮件和 Microsoft SharePoint）
- ✓ 工程应用（软件开发和电子设计自动化）

通过增加缓存层和提供专用过滤软件，该系统能够通过低延迟介质处理更多的操作。这有两个作用：

- ✓ 提高吞吐量，高于不使用 Flash Cache 的同类系统
- ✓ 减少磁盘子系统要处理的操作，并降低磁盘 IOPS 的延迟

表面上看，所有闪存都是大同小异。而实际上，它们在设计决策方面大不相同。NetApp 凭借 Flash Cache 另辟蹊径，集全力打造软件智能以最大限度地提升闪存效率，最终从企业闪存提供商中脱颖而出。

运行 Flash Cache 的 NetApp 系统能够实现超过 250,000 次的 IOPS，且延迟低于 1 毫秒。从磁盘驱动器访问数据时，延迟通常为 10 毫秒或更长时间。对比磁盘驱动器，Flash Cache 模块可将缓存命中的延迟时间缩短 90%（请参见图 2-2）。

Flash Cache 还证明了，可以用数量更少、成本更低的 7200 转数高容量 SATA HDD 来替代 15,000 转数高性能光纤通道和 SAS 硬盘驱动器，而不会影响性能。

NetApp 最新的 Flash Cache 模块 (Gen4) 包含 2 太字节 (TB) 的闪存。FAS 存储控制器上最多可安装八个模块，借此，每个控制器能实现高达 8 TB 的闪存容量。这些模块包含由板载现场可编程门阵列 (FPGA) 进行管理的 MLC 闪存（请参见第 1 章）。FPGA 控制系统主内存与位于 Flash Cache 板上的闪存之间的所有通信。FPGA 以提速为设计理念，采用多种技术来提升性能，例如：

- ✔ 智能交错多个写入队列中的写入进程，实现闪存擦除、写入和读取周期之间的平衡。由于每个闪存写入都会要求闪存擦除且擦除周期非常长，因此最好是并行运行尽可能多的周期以确保吞吐量。

- ✓ 支持多个内存接口，每个接口可连接多个存储库。当接口上的某个闪存库处于忙碌状态时（例如，处于擦除周期），FPGA 可向同一接口上的其他存储库发出命令。这可防止在请求过多而闪存库不够时出现延迟。

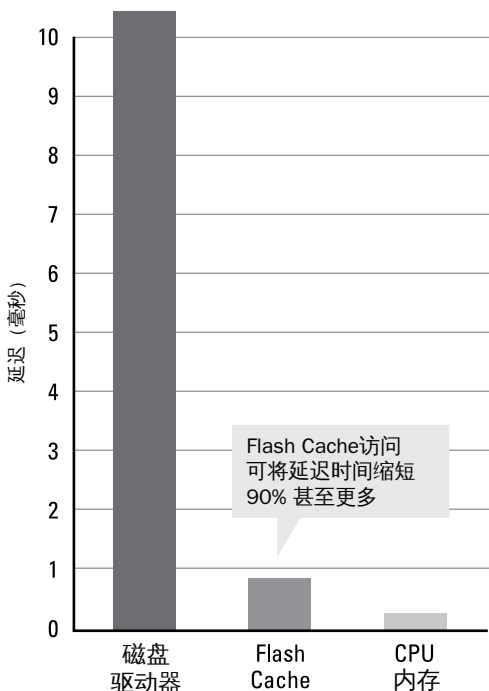


图 2-2：降低延迟。

- ✓ 整组读取多个闪存库中形成条带的闪存单元相比逐个读取这些闪存单元，可降低延迟超过 800%。

SSD 用作缓存

在“混合”存储池中，将 SSD 与 HDD 相结合，让网络存储系统既能够利用 SSD 的低延迟和高吞吐量优势，又能保持 HDD 的大存储容量。NetApp 在所有 NetApp FAS 和 V 系列存储系统中均提供了其 Flash Pool 解决方案。NetApp E 系列系统在其“SSD 缓存”中提供了类似的功能。

在单个池中结合使用 SSD 和 HDD 可让热数据在 HDD 和 SSD 之间进行自动分层，进而实现最佳性能（请参见图 2-3）。与 Flash Cache（详见本章前面的内容）一样，系统不会将热数据移动到 SSD，而是在 SSD 中创建一个数据副本。相比移动数据，复制数据有以下优点：

- ✓ 从 I/O 流程上讲，复制数据的速度要比移动快。（移动数据要求系统先复制数据，然后将其从原来的位置删除。）
- ✓ 热数据从 SSD 中删除后，则无需重写回 HDD。

然而，与 Flash Cache 不同的是，读取和写入都可以缓存到 Flash Pool SSD。如第 1 章中所述，在某些情况下，比如执行顺序写入操作时，HDD 实际上会发挥出基于闪存的 SSD 无可比拟的性能。正因如此，当在启用了 Flash Pool 的卷上执行写入操作时，系统会使用逻辑来判断是同时向 SSD 和 HDD 写入更快，还是只向 HDD 写入更快。相反，

如果是执行读取操作，逻辑会判断数据是应缓存到 SSD，还是应不加缓存从 HDD 直接读取。判断结果主要取决于数据模式是随机还是顺序。在任何情况下，Flash Pool 算法都是为提升速度而进行优化。

一级存储

```

10101001110101001110110101110
00101101010101001110101001110
11010111000101101010101001101
10110101010100110010110101010
100111010100111010100111010101
01110001011010101010011101010
01110110101110001011010101010
01101101101010101001100101101
10101001110101001110110101110
00101101010101001110101001110
11010111000101101010101001101
10110101010100110010110101010
10011101010011101010011101101
01110001011010101010011101010
01110110101110001011010101010
01101101101010101001100101101
10101001110101001110110101110
00101101010101001110101001110
11010111000101101010101001101
10110101010100110010110101010
  
```

目标：智能
放置“热”数据

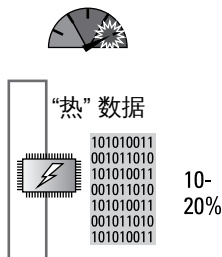


图 2-3：数据放置难题：确保热数据的速度快，冷数据的成本最低。

SSD 替代 HDD

越来越多的企业正想方设法提高用于其关键业务运营的应用程序的速度和响应能力，尤其是数据库系统。由于这些应用程序的性能与上市时间、营收、工作效率以及客户满意度密切相关，因此发挥其最大的潜能至关重要。

为达到极高的性能，企业可能会在其网络存储系统中过度配置容量，部署未占满的磁盘，以便阵列提供所需性能，但同时却会浪费各种昂贵的资源，包括硬件、容量、数据中心空间和电能。

为了避免过度配置并通过高性能应用程序获得最大的投资回报，许多公司现在寄望于纯闪存存储系统。但是，当购买者考虑这些新系统时，他们通常很难找到高度可靠、功能丰富且经济高效的企业级产品。

NetApp 纯闪存阵列提供超高性能、稳定一致的亚毫秒级响应能力和企业级功能，从而提高控制关键业务运营的延迟敏感型应用程序的速度、响应能力和价值。**NetApp EF** 系列闪存阵列是一个经过企业验证的平台，该平台采用的操作系统高效而简化，擅于处理注重极高性能、超低延迟和 IOPS 密度的 I/O 密集型工作负载。**NetApp** 纯闪存 **FAS** 系统兼具低延迟性能与 **Data ONTAP** 的丰富功能，其中包括重复数据删除、数据压缩、集成数据保护、多协议支持、横向扩展和无中断运行。

第3章

了解闪存存储技术的优势

本章提要

- ▶ 利用闪存提高存储性能
- ▶ 采用固态技术存储数据
- ▶ 利用 SSD 优化存储基础架构
- ▶ 衡量存储的实际成本

在 本章中，您将了解闪存存储技术如何为您的企业带来优势，以及多家不同行业的公司如何成功部署闪存存储技术并从中获益。

性能

闪存技术可提升整体 IT 性能。闪存存储能够在几微秒内执行读写操作，并提供高达数万次的 IOPS。相比之下，硬盘驱动器 (HDD) 的性能以毫秒计，且仅提供数百次的 IOPS（请参见表 3-1）。

表 3-1 闪存与 HDD 性能对比

	读/写	IOPS
闪存	微秒	数万次到数百万次
HDD	毫秒	数百次

且不同于 HDD，闪存没有移动部件，因此不会因硬件中产生的寻道时间或旋转延迟而带来性能影响。闪存没有移动部件还意味着损耗的减少，这样会提升机械可靠性（将在下一节进行论述）。



闪存存储可将响应时间缩短多达 90%，同时提高 I/O 吞吐量，使之比 HDD 高出多达 80%。

请参见本章末尾的案例研究“NetApp EF 系列闪存阵列提高 Oracle 数据库性能。”

可靠性

毫无疑问，可靠性是任何数据存储介质最为重要的一项属性。您必须保证以稳定可靠的方式对公司的重要记录和数据进行了存储，以确保其始终能够在需要时可用。

如前所述，闪存存储没有移动部件。单凭这一点就能显著提升可靠性并带来将近 200 万小时的平均无故障时间 (MTBF)，而 HDD 的平均无故障时间是 120 万小时左右。请参见边栏“磁盘驱动器的未来前景”了解更多有关 HDD 设计和可靠性的信息。

磁盘驱动器的未来前景

近来，关于硬盘驱动器的未来前景成为一大热门话题。HDD 问世已有超过半个世纪的时间。作为世界上第一个磁盘驱动器，IBM 350 磁盘存储单元由以下三大基本元件组成：铝制旋转盘片、磁涂层表面和可移动记录磁头。如今，尽管 HDD 的速度和容量都有所提升，但它们仍然是由这三大基本元件组成且工作原理也一样。

尽管 HDD 取得了惊人的成就，但有一天终究会被淘汰。精密电机和驱动器容易受机械磨损的影响，无法永久运行。一天 24 小时不停地旋转驱动器盘片却只是偶尔访问数据导致电能利用率低下。最后，既然热辅助磁记录 (HAMR) 技术预计至少能实现 100 TB 的 HDD 容量，那么从某种程度上讲 HDD 的容量、可靠性和经济效益将达到极限。

HDD 之所以长盛不衰是因为到目前为止都没有出现更好、能够取代它的产品。企业 SSD 采用闪存技术且在外观和风格上与 HDD 相似，但由于企业 SSD 完全由硅材料制成，因此不会像 HDD 那样受机械磨损或旋转延迟的影响。

企业 SSD 虽然是日渐普及，但其发货量与 HDD 相比仍是相形见绌。造成 SSD 销售滞后的主要原因有两个：成本和可靠性。

(转下页)

SSD 的成本以 GB 计算，价格比 HDD 贵将近 6 倍。从可靠性的角度看，闪存技术中的“闪存”意味着每执行一次电擦除（每次将数据存储到闪存都需要），存储单元就被削弱一次。

不过，闪存供应商正在采用精密的损耗平衡算法来解决 SSD 的可靠性问题，且生产规模经济将使成本逐渐降低。

如今，企业在混合存储阵列和小容量（专为实现高性能而设计）纯闪存存储阵列中应用闪存和 SSD 技术。

高效率

闪存技术对大幅提升数据中心效率的潜在影响显而易见。闪存是一项非常灵活的技术，有多种不同的外形规格，例如 PCI-e 闪存卡和 SSD。它可作为内存层部署在服务器、网络、存储控制器或存储阵列中。闪存能以缓存或独立持久的纯闪存阵列的形式存在（请参见图 3-1）。

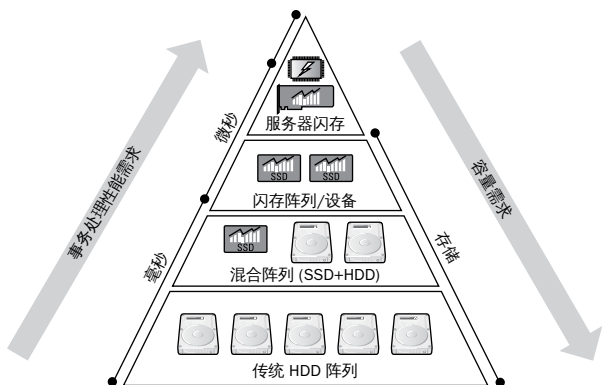


图 3-1：您可以灵活部署闪存技术以解决各种性能和容量需求。

闪存存储技术有助于 IT 部门提升运营效率。存储管理员可通过提升存储资源的利用率来达到使用 SSD 所有可用空间、优化存储以及减少 IT 整体支出的效果，而不必使用针对 HDD 的短行程技术（详见第 2 章）。（请参见案例研究“云服务提供商依靠 NetApp 存储效率技术节省超过 100 万美元。”）

案例研究：云服务提供商依靠 NetApp 存储效率节省超过 100 万美元

某家云服务提供商需要解决其存储基础架构的扩展性难题，与此同时要保持其全球超过 350,000 家客户可接受的服务级别。

挑战

灵活应对客户的需求，保持在托管通信和协作服务市场上的竞争优势。

解决方案

将客户迁移到基于 NetApp 无缝集成存储平台的 Microsoft Exchange Server 2010，并利用 NetApp 虚拟存储层 (VST) 提高性能。

优势

- ✓ 将总拥有成本降低 60%，节省超过 100 万美元的 IT 支出
- ✓ 将 Microsoft Exchange 2010 环境中的存储降低 75%
- ✓ 将存储管理时间减少 66%

概述

某家首屈一指的托管业务通信、协作、合规性和安全性以及基础架构解决方案提供商（全球拥有超过 350,000 个用户）必须通过一个经济有效且支持高可用性、高性能业务解决方案的基础架构来提供云服务。

支持该公司 Microsoft Exchange 2007 和 Microsoft SharePoint 2007 环境的存储平台已达到容量上限。该平台缺乏可扩展性和易管理性，因此需要对这些应用程序进行升级并将其客户迁移出去。

该公司与一家技术解决方案提供商兼 NetApp 星级合作伙伴密切合作，找到一个强大、无缝、统一的物理和虚拟资源管理平台。借助其主数据中心和二级数据中心中的 NetApp FAS3270 存储系统来支持其通信和协作服务，并充分利用其实验室环境中的 FAS3210 和 FAS3140 存储系统。

通过由其 FAS3270 存储系统中的 512 GB NetApp Flash Cache 卡（请参见第 2 章）驱动的 NetApp VST，NetApp 存储系统帮助云服务提供商优化性能并提升客户体验。该公司充分利用 VST 及其智能缓存和 NetApp 存储效率技术来为其 Microsoft Exchange 和 SharePoint 工作负载提供支持。NetApp VST 具有提高 IOPS、存储效率以及支持的 Microsoft Exchange 2010 邮箱数量的能力。该公司利用 NetApp 重复数据删除技术将其 Microsoft 环境的存储要求降低 40% 到 50%，使效率得到进一步提升。

NetApp 存储解决方案帮助该云服务提供商优化了 IT 预算、降低了风险并使之能够灵活扩展以应对未来的增长需求。该中端存储系统为混合工作负载提供了一个经济高效的环境，使之借助其内在扩展性保护公司的投资。FAS3200 系列存储系统通过实现额外的诊断和无中断恢复功能来进一步提升高可用性 NetApp 架构的性能。

成本

尽管传统 HDD 的每 GB 成本比 SSD 的每 GB 成本低，但这并非衡量存储成本的唯一因素。

衡量成本的其他重要因素还包括：

- ✓ **有效容量**：为优化 IOPS，许多 IT 存储管理员使用非占满的阵列来降低 HDD 上的寻道延迟。此做法会减少 HDD 的有效容量，同时大幅增加可用存储的每 GB 成本。
- ✓ **每 IOPS 成本**：如果廉价磁盘不能满足您要求最苛刻的应用程序和用户对吞吐量的要求，那么从生产效率和错失的机会上讲，廉价磁盘会让您花费更多的成本。HDD 每 IOPS 的一般成本大概为 1.25 美元，而 SSD 仅为每 IOPS 0.02 美元。
- ✓ **运营成本**：电耗、散热、数据中心空间以及存储托管和管理全都是必须考虑的重要运营成本。

闪存存储无需旋转盘片和移动读/写磁头便能提供高性能吞吐量。相反，闪存技术使用晶体管将数据存储在以行和列组成的阵列中（请参见第 1 章）。

性能越高，应用程序的运行速度就会更快，用户的效率也会更高。尽管某些人会因“软成本”微不足道而对这一因素不予考虑，但企业闪存 SSD IOPS 与 HDD IOPS 之间的差异还是很明显的。（请参见案例研究“NetApp FAS 和 Flash Cache 解决方案加速天然气勘探速度”）

最后，企业闪存 SSD 可帮助您提高数据中心空间的利用率并降低能耗成本。

企业可将高性能、容量相对较低的 SAS 和光纤通道 HDD 替换为低成本、高容量的 SATA HDD 和 SSD，从而减少其数据中心的物理占用空间。企业可用 24 个 SSD 轻松替换 1,000 个企业 HDD，前者能达到与后者相当的事务处理 IOPS 性能（请参见图 3-2）。

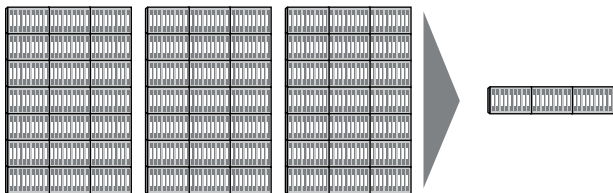


图 3-2：24 个 SSD 能达到与 1,000 个企业 HDD 相同的 IOPS 性能。



HDD 需要强大的电量来运转电机和旋转介质。相反，闪存存储消耗的电量非常少。相比 HDD，闪存将每 IOPS 的电耗减少 600%，且 HDD 的单位电耗约为 SSD 的 2.5 倍——通常 HDD 为 16 瓦，而 SSD 为 6 瓦。此外，SSD 产生的热量也比 HDD 少，从而降低您的散热运营成本。



借助对 SATA 技术的广泛支持，企业闪存 SSD 可在空间和电耗方面实现高达 60% 的节省，同时将每 TB 的成本降低多达 46%。

案例研究：NetApp FAS 和 Flash Cache 解决方案加速天然气勘探速度

一家大型能源勘探公司求助于 NetApp，希望 NetApp 帮助他们解决在容量和性能上面临的重大难题，以助其提升竞争优势。

挑战

应对每年 200% 到 300% 的增长，满足地质建模和地震解释应用环境下密集的性能要求。

解决方案

借助 NetApp FAS6000 和 FAS3000 系列存储系统以及 NetApp Flash Cache 智能缓存，提升系统性能和员工效率。

优势

- ✓ 在不额外增加人手的情况下，实现 700% 的数据增长
- ✓ 利用基于 SATA 驱动器的 Flash Cache，将每 TB 的成本降低 52%，节省 66% 的电耗以及 59% 的空间
- ✓ 近乎即时地访问大型数据卷

概述

美国一家从事天然气、原油和液化天然气勘探、开采和生产的独立能源公司采用最新技术来助力于实现资产价值最大化（超过 110 亿美元）并利用关键型决策工具提高近 4,450 名员工的工作效率。

随着公司及勘探数据增长到 13 PB 这一惊人的数字，该公司必须确保其 IT 基础架构能支持其不断变化、密集的运行要求。该公司成功的关键在于能够快速访问大量可靠的地震数据和解释数据。此外，该公司还投入了数百万美元的资金来保护其在特定区域的开采权。为了减少相关的财务风险，该公司必须使员工能够随时访问与最佳开采位置相关的高度精确的信息。

该公司部署了 NetApp 存储环境，进而能够方便快捷地进行扩展并具备几项重要的性能优势。位于主数据中心和二级数据中心的 FAS6280 和 FAS6080 存储系统负责提供生产环境，FAS3270 和 FAS3170 存储系统则负责支持其他运营领域。NetApp Flash Cache（请参见第 2 章）也是该解决方案的重要部分，其作用是助力该公司优化整个环境的性能。Flash Cache 通过对随机读取密集型工作负载进行智能缓存来加速对石油和天然气数据的访问。

借助 Flash Cache，该公司在共享地震数据工作环境中的缓存命中率达到将近 70%，将延迟时间缩短 90% 甚至更多。也就是说，有 70% 的数据存储到内存中，因此系统不必从 SATA 驱动器中检索数据并能在数秒而非数分钟内提供数据。

将开启和加载地质模型的时间从原来的 20 分钟缩短到现在的 5 分钟以内。地质学家现在可以在短短 2 分钟内进行过去需要 10 分钟才能完成的测试。提高系统性能也有助于地质团队加快计划（例如开发新技术）实施。

案例研究：NetApp EF 系列闪存阵列 提高 Oracle 数据库性能

面对实现极高业务增长的期望和稳固自身竞争优势的需求，投资公司为自己设定了一个目标，即提升其数据库基础架构的 I/O 性能。

挑战

提升业务关键型 Oracle 数据库的 I/O 性能。

解决方案

将四个 NetApp EF 系列闪存阵列分散部署至两个数据中心，为在 Oracle 上运行的投资管理和业务智能应用程序提供支持。

优势

- ✔ 数据库 I/O 的平均响应时间缩短至 0.86 毫秒，速度是之前的 12 倍
- ✔ 大规模分析作业的性能提高了 100%，目前的运行速度是之前的两倍
- ✔ 现在，将数据传输到相邻卫星系统所需的时间仅为之前的 30%
- ✔ 重新启动后重新创建数据库缓存所需的时间减少了 90%
- ✔ 备份时间减少超过 75%

第4章

闪存在未来存储项目中的十大（或五大）使用策略

本章提要

► 成功部署闪存

闪

存技术为存储专业人员带来了一系列新的存储选择。通过部署闪存技术，企业可以利用更高更快的吞吐量大幅提升任务关键型、I/O 密集型应用程序的性能。

存储专业人员面临的挑战就是：如何部署闪存技术才能在合适的时间以适当的成本提供正确的数据以满足您的需求？性能始终是关键。成本也是如此。企业必须继续从对新存储基础架构的投资中获取最大收益。

鉴于这一点，本章将为您介绍针对如何在未来项目中合理使用闪存的五大策略。

了解您的工作负载

闪存是优化随机读取密集型工作负载性能的最佳选择。虽然您可能会将闪存用于一个或多个任务关键型、I/O 密集型工作负载，例如文件服务、消息传送、虚拟基础架构或 OLTP 数据库，但您还必须了解这些工作负载的数据访问模式。您需要回答以下四个问题：

- ✔ 什么是读/写混合工作负载？
- ✔ I/O 一般是多少？
- ✔ 随机访问与顺序访问的数据量分别是多少？
- ✔ 应用程序的工作集有多大？

采用虚拟存储分层

大多数自动分层解决方案尝试根据数据访问模式和一组预定义策略在多个物理存储层之间持续移动数据。此过程相当复杂，需要从设计上对数据粒度和数据移动的时间安排进行权衡，以便提供一个可接受的性能级别。

另一种方案就是使用虚拟存储分层。作为一种自行管理型方案，虚拟存储分层可以提供最低粒度级别的热数据实时提升，而无需进行复杂的数据分类和移动。借助虚拟分层，您只需为卷或 LUN 选择所需的默认介质层，例如 SATA、光纤通道或 SAS，而该卷或 LUN 中的数据会自动按需提升至基于闪存的介质。

评估系统级别的性价比

评估闪存技术时，先在组件级别进行比较，然后以此来推断结果的方法看似很好。千万别被这一表象所迷惑。例如，如果系统要优化写入性能，闪存技术可能并不是最合适的存储介质。记住，闪存的主要优点通常是提升读取性能。在应用程序层中，最重要的不是由哪个子系统处理读取和写入，而是提供数据的速度和成本。

善用纯闪存

相比混合或纯 HDD 系统，纯闪存阵列的每 GB 成本更高，因此一定要掌握好利用这些系统以发挥最大成效的时机。纯闪存阵列最适合延迟敏感型应用程序，而对这些应用程序而言，提高响应能力意味着加快上市速度、做出更明智决策或获得更多收入。例如，事务处理工作负载 (OLTP)、实时分析 (OLAP)、决策系统和技术计算。纯闪存也非常适合用来改善对每桌面 IO 有极高要求的虚拟桌面基础架构中的最终用户体验。

搭配部署闪存和容量优化的驱动器

在磁盘阵列容量充足却还需要提升 I/O 的情况下，您应该考虑使用 NetApp Flash Cache（详见第 2 章）和 Flash Pool。测试表明，Flash Pool 能使采用 SATA 驱动器的系

统达到与采用 SAS 驱动器的系统一样的性能，同时还将存储容量提高 50%，每 TB 成本减少 46%，电量要求降低近 27%。

.....

[illegible]

.....

[illegible]

为任务关键型应用程序部署闪存存储

闪存技术为企业数据中心带来前所未有的速度、可靠性和效率。本书介绍闪存在存储基础架构中的部署方式和部署位置。

- 借助虚拟存储分层简化和优化存储 — 自动按需将“热”数据复制到缓存
- 利用闪存提升存储性能 — 在几微秒内执行读取和写入操作，并提供高达数万次甚至数百万次的 **IOPS**
- 降低运营成本 — 包括电耗、散热和数据中心空间成本



翻开书并找到以下问题的答案：

- 哪些闪存技术可以帮您更好地应对所面临的存储挑战
- 如何使用 **SATA** 和闪存来代替 **SAS** 硬盘驱动器
- 为何固态驱动器要比磁盘更快、更可靠
- 闪存如何提高用户的工作效率

Lawrence C. Miller 在信息安全和技术管理领域拥有超过 15 年的工作经验。他已编写了 35 本有关各种主题的“傻瓜系列”丛书，其中包括《傻瓜系列之 CISSP》(CISSP For Dummies)。

WILEY

ISBN 978-1-118-99423-8
本书不得转售