

倾情奉献



# 软件定义存储

傻瓜书

Wiley

控制存储成本

消除存储瓶颈

解决存储管理挑战



Chris Saul

IBM 限量版第二版



# 软件定义存储

IBM 限量版第 2 版

作者：Chris Saul

## 傻瓜书

Wiley

# 软件定义存储傻瓜书, IBM 限量版第二版

出版商:

**John Wiley & Sons, Inc.**

111 River St.

Hoboken, NJ 07030-5774

[www.wiley.com](http://www.wiley.com)

版权所有 © 2019 by John Wiley & Sons, Inc.

除依据《1976年美国版权法案》第 107 或 108 条规定允许的情况外, 未经出版商事先书面许可, 不得以任何方式(包括电子、机械、复印、录制、扫描或其它任何方式)对本出版物中的任何章节进行翻印、传播或将其存储在任意检索系统中。如需申请许可, 请与出版商联系, 地址: Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, 电话: (201) 748-6011, 传真: (201) 748-6008。在线申请网址: <http://www.wiley.com/go/permissions>。

**商标:** Wiley、For Dummies、Dummies Man 徽标、The Dummies Way、Dummies.com、Making Everything Easier 及相关商业外观是 John Wiley & Sons, Inc. 和/或其关联公司在美国和其他国家/地区的商标或注册商标, 未经书面许可, 不得使用。IBM 和 IBM 徽标是 International Business Machines Corporation 的注册商标。所有其他商标均为各自所有者的财产。John Wiley & Sons, Inc. 与本书提及的任何产品或供应商无任何关系。

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保证, 包括但不限于对特定用途的适用性。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。本书的销售前提是, 本书出版商不借此提供任何有关法律、会计或其他专业服务。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, 读者应该意识到, 本书所列网站可能在本书写成后的时间里发生改变或消失。

关于我们提供的其他产品和服务的信息, 或者如何为您的企业或组织定制 *For Dummies* (傻瓜书) 系列书籍, 请联系我们在美国的业务发展部, 电话: 877-409-4177, 电子邮件: [info@dummies.biz](mailto:info@dummies.biz), 网址: [www.wiley.com/go/custompub](http://www.wiley.com/go/custompub)。关于如何为产品或服务申请 *For Dummies* 品牌许可, 请联系: [BrandedRights&Licenses@Wiley.com](mailto:BrandedRights&Licenses@Wiley.com)。

ISBN: 978-1-119-55407-3 (pbk); ISBN: 978-1-119-55406-6 (ebk)

美国印刷

10 9 8 7 6 5 4 3 2 1

## 出版商鸣谢

谨此感谢帮助本书成功出版的人士:

**项目编辑:** Carrie A. Burchfield

**策划编辑:** Steve Hayes

**编辑经理:** Rev Mengle

**业务开发代表:** Sue Blessing

# 目录

引言 .....	1
关于本书 .....	1
本书中使用的图标.....	2
参考资料 .....	2
<b>第 1 章: 什么是软件定义存储? .....</b>	<b>3</b>
SDS 的定义 .....	3
SDS 主要类型讨论.....	5
面向存储管理的 SDS.....	5
可提供现代数据保护和服务的 SDS.....	6
作为基本存储基础架构运行的 SDS.....	7
IBM Spectrum Storage 简介 .....	9
<b>第 2 章: 面向存储管理的 SDS .....</b>	<b>11</b>
IBM Spectrum Control.....	11
IBM Storage Insights.....	13
IBM Spectrum Discover .....	14
IBM Spectrum Connect.....	16
<b>第 3 章: 面向数据保护和服务的 SDS .....</b>	<b>19</b>
IBM Spectrum Archive.....	19
IBM Spectrum Protect .....	21
IBM Spectrum Protect Plus.....	22
IBM Spectrum Copy Data Management (CDM).....	24
<b>第 4 章: 面向基础架构的 SDS .....</b>	<b>27</b>
IBM Spectrum Virtualize.....	27
IBM Spectrum Scale .....	29
IBM Cloud Object Storage .....	32
IBM Spectrum Accelerate.....	33
IBM Spectrum NAS.....	36

**第 5 章: IBM Spectrum Storage 系列解决方案的 10 个用例 39**

充分利用 AI.....	39
支持敏捷性.....	40
通过现代化提升业务弹性.....	41
实现数据价值的倍增.....	42
结合采用闪存.....	43
支持混合云.....	43
借助混合多云实现数据保护的现代化.....	43
简化大数据分析.....	44

# 引言

*数据驱动型企业*是指尽最大努力从其数据资产中获取价值的组织。他们使用数据来改善和加速决策，提升客户支持和体验，创造新的收入流，并以多种不同方式获取并保持竞争优势。数据驱动型企业采用多种不同的信息技术 (IT) 工具和方法从数据中发掘更多价值。云 IT 模型可帮助企业降低成本并提供更大的业务敏捷性，这使得云计算在最近成为最强大的功能之一。各种类型和规模的企业都已开始利用私有云、公有云、混合云以及现有多云解决方案和架构带来的优势。

软件定义存储 (SDS) 是另一套不断扩展的 IT 工具，此类工具正在数据驱动型企业中迅速流行。“*软件定义的存储*”这一术语目前使用得非常频繁，其中涵盖了许多其定义到目前为止尚不清楚且存在争议的要素。我提供了一个有用的定义，而且会在之后的章节中探讨其属性，但请牢记一个重要的概念 - SDS 是一个工具集，可帮助您解决数据存储、移动、操作、保护和分析方面的难题，以便更好地实现您的业务目标。

最重要的是，SDS 是构建数据驱动型多云企业的基础。

## 关于本书

《软件定义存储傻瓜书》IBM 限量版第二版介绍了数据驱动型组织如何利用这种快速发展的技术实现业务和 IT 目标，同时还介绍了市场领先的 SDS 解决方案系列 - IBM Spectrum Storage 系列解决方案。此外，本书还就 IBM 推出的这款功能强大的软件如何帮助企业成为市场领导者提供了大量的相关信息和观点。

## 本书中使用的图标



切记

在本书中，会偶尔使用一些特殊图标来引起读者对重要信息的注意。预计可能出现的图标如下所述。

此图标用于指示可能需要您给与完全关注的信息。



技术内容

在出现此图标时，肯定不会提供人类基因组图或 IBM Watson 蓝图，不过却会针对行话下方的术语提供一些有用说明。



提示

感谢您阅读本书，也希望您能喜欢本书，请珍惜作者的成果。此图标用于指示一些非常有用的信息块。

## 本书之外的参考资料

尽管本书所含信息非常丰富，但其内容毕竟有限。如果您在阅读完本书后发现自己想要了解更多信息，敬请访问 [ibm.com/spectrumstorage](http://ibm.com/spectrumstorage)。在该网站，您将会找到更多与 SDS 和 IBM Spectrum Storage 相关的信息。

- » 软件定义存储的定义
- » 纵观 SDS 的类型
- » IBM Spectrum Storage 简介

## 第 1 章

# 什么是软件定义存储？

本章针对软件定义存储 (SDS) 给出了一个简单、有用的定义，同时还探讨了该技术的主要功能类型，介绍了市场领先的 SDS 解决方案系列 - IBM Spectrum Storage 系列解决方案。

## SDS 的定义

在最传统的 IT 架构中，计算机程序或应用会以各种方式对数据进行处理，进而将原始数据转化为业务价值。但是原始数据往往并不会存储在应用中，而是存储在其他位置，即存储设备中。传统 IT 架构中的应用从存储设备请求或查询所需的数据，而存储设备则通过提供所请求的数据进行响应。

不过想象一下，如果将软件放置应用和存储设备之间的数据路径中，情况将会怎样？从应用的角度来看，什么都没有改变，即在查询之后返回数据。但是从存储角度来看，一切都发生了改变。这就是 SDS 的基础所在。



通过在应用和存储之间部署一层或多层软件，那么存储设备的类型实际上可能就会变得无关紧要，并且设备本身可以互换。此类软件能够维持每个数据单元的位置，并且无论数据的物理位置如何，都可以对其进行检索。将物理存储设备与存储软件分离或“解耦”可实现诸多益处和优势。这种布置还赋予了 SDS 最流行的定义之一，即不依赖于任何特定底层存储设备的存储管理软件。

不过 SDS 已经实现了长足发展。考虑 SDS 的另一种方法是将其想象为在应用及其数据之间的数据路径中的战略点处插入的一个智能层。随着智能软件开始部署在相对被动的物理存储介质（如闪存、磁盘或磁带等）与 IT 基础架构的其余部分之间，使得存储领域突然出现了广阔的可能性。

假设您想通过复制数据并将其发送到其他系统或位置的方式来保护数据。当然，您也希望在数据进出时以及挂起时对其进行加密。想象一下：如果您能够监控数据的访问模式，然后仅将最常访问的一小部分存储在速度最快、成本最昂贵的存储中，然后将其余部分的数据发送到成本仅相当于其他存储介质很小一部分的磁带上，将会实现多少的成本节省？您是否希望实现真正的智能？想象一下，某个存储系统可以学习预测所需的数据并能够从真正廉价的磁带系统中提前检索数据，这样您就不必长时间地等待 IT 人员通过盒式磁带去装载数据。



切记

这就是智能化。这也是目前 SDS 的现状，同时也是我所说的 SDS 的定义 - 在应用主机和物理存储设备之间部署的智能与功能层。

该定义补充了许多其他更为传统的 SDS 定义。一旦完成了实施，SDS 将能够响应用对数据的请求，而不是成为保存实际比特与字节的硬件。相反，SDS 能够管理后端存储设备，根据请求从中检索数据，移动数据以优化业务和 IT 目标，并对数据进行加密和监控...SDS 可完成的此类活动有很多，而且其数量还在不断增加。因此，在该定义下，几乎所有的存储硬件（我们称之为“商用存储”）都可以在智能 SDS 后面使用。

基于 SDS 的 IT 战略催生了下一代的存储基础架构。与硬件分离的软件可在不造成任何不利影响的情况下实现最佳配置，而且价格低廉。这种方法对于随着时间的推移而进行的硬件升级也非常有用，使得企业可以使用不同代次的硬件，而无需再次购买软件。



切记

另外，SDS 只是代码。只要能满足 SDS 适当的处理和存储要求，您就可以根据需求下载和部署 SDS。这适应了 SDS 的另一个通用定义，即：与存储相关的、可以与底层存储硬件分开升级和/或下载及实施的产品。

就 SDS 的定义而言，真正的问题在于：一旦您开始想象的话，可能性似乎是无限的，甚至是混乱的。SDS 几乎可能提供我们想要的任何数据服务（它也正在朝着这一点迅速发展），或执行我们想得到的任何技巧性策略。为了简化这些观点并确保 SDS 相关对话的一致性，我们对各种类型的 SDS 产品和工具进行了一些有用的划分。

## SDS 主要类型讨论

考虑一下 SDS 的作用。SDS 工具和解决方案的作用之一是监控、评估、管理和控制存储过程及其所有组件。也许最常见的作用是，SDS 执行与数据保护、存储、移动及提供数据访问相关的特定任务和服务（此类任务和服务的范围非常广泛而且在不断增多）。SDS 产品和解决方案可以为底层的基础性存储功能本身提供基本的组织特征和功能。通过这些简单的观察结果，我们可以将 SDS 分为三个基本类别 - 管理类、数据保护和服务类以及基础架构类。

### 面向存储管理的 SDS

您知道您的组织目前所用的存储容量是多少吗？最近的增长速度如何？您有多少未使用容量储备？有哪些访问模式？这些模式是否会在月底或每个工作日的早上达到峰值？坦率地说，您的每个数据集及其所有副本都存放在何处？如果这些数据集位于不同的存储之中，您是否可以节省资金或提高应用性能、员工生产力或客户满意度？

如果您不知道所有这些问题的答案，则表示您的组织需要使用 SDS 来提供更好的管理服务。这些问题仅仅只是存储管理的冰山一角。我们不妨考虑以下问题：

- » 您是否可以几乎随时随地找到所需的任何文件吗？
- » 您是否可以针对所有存储活动生成足够详细的报告，进而使其能够在与其他公司高管举行与重大存储投资相关的会议时使用？
- » 在出现问题时，您的存储基础架构中的某些功能将无法正常运行；在此情况下，您能否快速、轻松地识别出问题，进而向存储供应商的客户支持团队提供所需信息？
- » 您是否对通过基于行业标准和不断演变的实际经验优化存储配置来实现 IT 和业务目标充满信心？

这些问题尚未涉及您在虚拟化业务应用或尝试利用容器等新技术时可能会遇到的系统集成和兼容性问题，而此类集成和兼容性问题就引出了更多问题：您存储在云端之外的数据情况又是如何？您是否能将本地发生的事件与本地存储系统中发生的事件整合到一起？

上述这些问题仍旧没有触及 SDS 管理工具和解决方案的功能。此类 SDS 十分重要，但许多 IT 专业人员在考虑使用 SDS 时并未将存储管理和控制考虑在内。相反，他们经常想到的是数据服务和数据保护。

## 可提供现代数据保护和服务的 SDS

就您希望完成的事项以及您的数据而言，相关的列表几乎每天都在“变长”。但是，保护数据免受意外或恶意破坏、丢失或攻击几乎是每个人的首要任务。SDS 提供了诸多强大且有效的方法，可为您提供数据保护和相关服务。事实上，数据保护已经变得非常重要，以至于在我们讨论到 SDS 时，需要也值得将其划分为单独的一类。

数据保护战略和解决方案实质上会影响 IT 环境的所有其他方面，包括如何管理数据增长、选择哪种解决方案来确保业务及其应用始终保持运行且可供客户使用。举例来说，快速数据增长的管理与数据保护解决方案的规模化性能及其在降低不断上升的存储成本方面的效用有很大关系。

此外，数据保护不只是简单的备份和恢复。数据副本也会引起很多麻烦。在许多数据中心中，仅仅是数据副本的存储就占到了总存储容量的一半以上。当应用开发和测试团队想要最新数据集的副本时，情况将会怎样？有人知道最新的副本位于何处吗？IT 部门如何确保用户拥有对数据副本的访问权限？向团队提供已清除了所有敏感信息的副本需要多长时间（一周还是数分钟）？

您的数据保护解决方案如何降低存储需求？是否可以轻松地添加磁带和云存储资源，以实现经济高效的数据归档？并不是说简单地致本地云提供商，然后就能够以所需的性能和成本来启用所需的所有服务，这是不可行的。实际上，有效地集成公有云资源可能会很复杂，也可能会有风险，这取决于您的 SDS 数据保护和数据服务解决方案的质量。

## 作为基本存储基础架构运行的 SDS

简单来说，所有数据都可以看成比特和字节。但是，只要您将足够的比特和字节集聚到一起，或者将所有数据（例如图像或视频的像素）关联起来，那么数据就会展现出可用于描述性或组织性之目的的特性。之后，我们开始将数据称为结构化数据或非结构化数据、块数据、文件或对象。结构化数据是数据库中流转的信息，并以表格的形式进行管理和操作。总体而言，在整个信息领域中，大约有 80% 的数据被描述为非结构化数据，例如文档、图像、音频和视频文件、电子邮件、通过物联网 (IoT) 传输的汽车遥测数据、数字化之后的 DNA 数据等。为了让这些数据易于管理，我们将其组织成文件并使用被称为“文件系统”的特殊软件。最后，某些类型的文件会携带有关自身或元数据的额外信息，因此可以在不依赖传统分层组织结构的情况下对其进行管理和搜索。这些被称为对象。每种类型的数据（块数据、文件或对象）都需要自己专用的存储软件，不过一些产品也可以有效地处理所有这三种类型的数据。

此类软件位于离存储硬件和存储介质最近的位置，有助于形成所谓的基本存储基础架构。他们通常直接集成到存储设备中并作为控制器功能的一部分，但并非必须如此。这些基础架构级别的解决方案可以像合法的 SDS 一样运行良好 - 仅作为不依赖任何特定硬件的软件，可下载、可升级和可部署。

无论是与特定设备深度集成还是完全不耦合，这种基础架构 SDS 在过去十年左右的时间里已变得异常强大。它可以监控访问模式并将数据集移动到最佳存储介质中。它还可以压缩数据或识别冗余信息，而且仅存储数据一次，进而让存储效率变得更高。这种类型的 SDS 可以加密和复制数据，将其发送到云中，或者在集群和网格中一起管理许多存储设备。它甚至可以令磁带表现得像磁盘一样。

如果您将 SDS 划分到以下三种基本类型（管理类、数据保护类或存储基础架构类）之一时，就更容易理解 SDS。但是，并非您购买的任何可提供管理服务的 SDS 产品都能够与任何数据保护解决方案或作为基本存储基础架构运行的 SDS 相兼容 - 除非这些产品经过精心设计并可以很好地协同工作。

# IBM Spectrum Storage 简介

IBM 不断投资于 SDS 的开发和创新，即：将以前与 SDS 相关的独立解决方案整合到一个具有统一用户体验的产品系列中，。该 SDS 产品系列被称为 IBM Spectrum Storage 系列。

IBM Spectrum Storage 是一个领先的 SDS 解决方案系列。该系列提供了广泛的 SDS 功能，可以满足您的存储目标。图 1-1 所示为 IBM Spectrum Storage 系列当前所包含的产品，各个产品按 SDS 的功能类别进行组织。有关最新信息，敬请访问 IBM Spectrum Storage 网站：[ibm.com/spectrumstorage](http://ibm.com/spectrumstorage)。

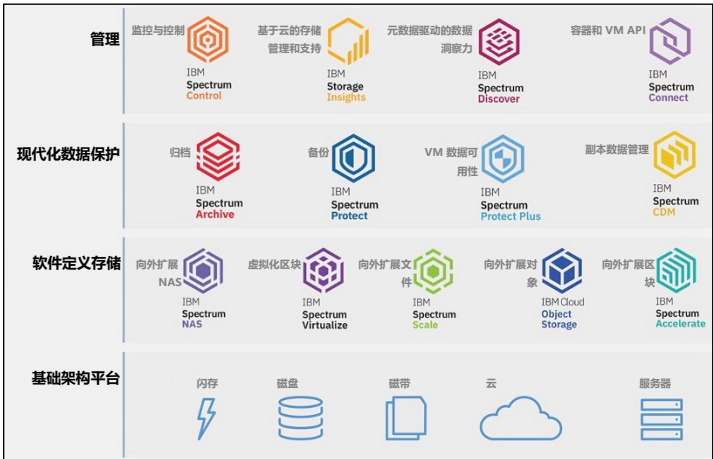


图 1-1: IBM Spectrum Storage 系列产品。

IBM Spectrum Storage 可提供全面的 SDS 应用组合，可帮助您管理所有类型的数据，无论其位于何处。您可以在本地和多云环境中统一存储，利用该系列产品的丰富功能更轻松、更有效地实施和管理重要的业务工具，例如分析和人工智能 (AI)，同时提高业务敏捷性并降低成本。IBM Spectrum Storage 系列解决方案所提供的许可选项，与购买和部署来自不同供应商的单独产品相比，具有许多优点和优势。

IBM Spectrum Storage 系列解决方案可通过多种不同的方式进行部署。其中一些产品已经集成到 IBM 存储系统（如 IBM FlashSystem 或 Storwize 系列）之中，因此您可以在存储系统中内置大量的 SDS 功能。不过，它们也可以作为纯软件解决方案在数据中心服务器上实施。此外，IBM Spectrum Storage 系列的许多产品都可以在云端运行，因此您无需进行管理和维护即可获得它们的功能。

在接下来的三章中，我将为您介绍 IBM Spectrum Storage 系列当前所涵盖的产品，并讨论每款产品的作用、为什么要部署特定解决方案，以及企业从中可获得哪些益处和优势。

## 本章提要

- » 监控、分析和管理您的存储环境
- » 利用 AI 加速支持并优化存储运营
- » 深入了解非结构化数据
- » 支持和简化多云部署

# 第 2 章

## 面向存储管理的 SDS

IBM Spectrum Storage 具有许多组件。在本章中，我将介绍该系列中四款可提供各种存储管理功能的产品。

### IBM Spectrum Control

您的数据存储环境很有可能是混合环境。您有来自不同供应商的存储系统，每个存储系统的部署时长不一，而且都有其自己的用户界面和管理软件。在复杂的存储生态系统中，您必须监控可能数百甚至数千个磁盘、闪存驱动器，甚至是磁带库的性能、容量和其他特性。完成所有必须做的事情似乎是不可能完成的任务，但事实上，对于 IBM Spectrum Control 这款可提供存储管理服务的软件定义存储 (SDS) 产品来说，只需要运行一天即可完成这些事情。IBM Spectrum Control 是一个综合性的端到端数据和存储管理解决方案，可监控和分析多供应商存储环境并实现其自动化。



它可帮助您管理存储系统、文件系统和数据库的性能、可用性和容量利用率。

IBM Spectrum Control 提供了单个控制点，可帮助管理员跨多站点存储环境（包括 IBM 的存储环境和非 IBM 存储环境）管理存储基础架构的各个方面 - 从应用主机到存储网络结构，再到物理驱动器。您可以使用 IBM Spectrum Control 来执行日常存储管理任务，例如容量配置、存储层优化、性能调优以及数据复制控制和调度等。它可以提供诊断功能，查明受可用性或性能问题影响的资源，然后及时生成警报，以启用事件操作。

IBM Spectrum Control 可以帮助您节省 IT 人员的时间并提高他们的生产效率，进而降低总体存储管理成本，具体包括：



提示

- » 简化库存控制、资产管理和报告流程
- » 通过单个集成式基于 Web 的管理控制台简化多个存储设备的管理
- » 提供各种工具，使得 IT 人员能够从单个位置执行例行管理任务，例如聚合、设备分组以及基于策略的操作等

由于这些功能，IBM Spectrum Control 能够帮助您：

- » 基于性能指标设置阈值，使 IT 人员能够主动管理性能
- » 减少服务台人员所收到的呼叫量并改善诊断效果
- » 允许管理员监控指标，例如 I/O 速率和缓存利用率等，以改善容量管理和规划
- » 提供了一个中央用户界面，用于配置和管理 SAN 上所有受支持的设备



提示

IBM Spectrum Control 是一款内部部署软件。不过 SDS 的一个特性在于软件可以在内部环境或云端运行。为了简化实施，IBM Storage Insights 提供了类似的功能，不过是通过简单的订阅计划从安全的 IBM Cloud 中提供。

## IBM Storage Insights

IBM 认为，人工智能 (AI) 可以为数据存储和管理带来巨大的益处。每个存储管理员都知道，大型存储系统的管理需要在监控、分析、决策和调整方面花费大量的时间。至少可以这么说，一旦出现问题，复杂存储基础架构的故障排除，以及最高效解决方案的实施都极具挑战性。

为了解决这些挑战并减少人工投入及失误，IBM 的存储系统配备了 IBM Storage Insights，它是一款久经企业验证的、经 AI 增强且基于云的系统洞察力平台，可帮助组织更好地了解存储容量和性能方面的趋势，加快需要支持时解决问题的速度。Storage Insights 能够分析根本原因并在应用受到影响之前预测性能和容量问题，进而帮助您监控基础架构，优化存储并解决问题。

Storage Insights 是管理类别中真正的基于云的 SDS 解决方案，可为您提供有关存储运行状况、性能、容量趋势、成本和利用率的更深入的洞察力。

- » 监控整个存储环境中的存储运行状况、性能和容量。
- » 查看多年累积的 70 多项指标，以了解相关趋势，同时将其与最佳实践进行对比，以在应用受到影响之前识别异常。
- » 查看多年的历史数据，同时提升对数据增长率及所有可用容量的可视性，使您在容量规划方面无需猜测。

- » 使用分析驱动型数据管理功能将数据移动到最具成本效益的存储层，进而降低成本。
- » 识别并回收已调配但未使用的存储，进而延后未来的存储购买时间。



技术内容



切记

Storage Insights 可以在单个视图中监控由 IBM Spectrum Virtualize 控制的所有 IBM 块存储和外部存储的运行状况、容量和性能。Storage Insights Pro 还能够监控 IBM Spectrum Scale 和 IBM Cloud Object Storage 的存储。

作为一种基于云的服务，Storage Insights 只需几分钟即可完成部署，并且 IBM 会通过新功能自动对其进行更新，无需您执行任何操作。Storage Insights 可以提供前瞻性的最佳实践，而且有助于识别潜在问题，并在需要支持时快速解决问题。Storage Insights 可帮助 IBM 存储产品的客户通过全球领先的云提供商之一加快解决问题的速度，提升用户体验，提高系统可用性，并确保服务满意度。

## IBM Spectrum Discover

从非结构化数据资产中挖掘最大的业务价值并非易事。管理庞大的非结构化数据以及存储这些数据所需的系统是一项重大挑战。存储管理员经常会发现与其数据相关的信息（也称为元数据）无法提供实现高效优化存储所需的、一个能够呈现存储消耗、数据所有权、数据内容和数据质量的适当视图。对于可能花费大量时间来搜索包含有机密数据或敏感数据的文件和对象的数据科学家、业务分析师和知识工作者来说，基本的系统级元数据也是不够的。这就是 IBM 推出 IBM Spectrum Discover 的原因所在。

企业需要可提供卓越数据可视性的元数据管理解决方案。如果组织能够清晰地了解他们的非结构化数据，他们就能更轻松地优化存储系统，确保他们的非结构化数据符合政府法规和行业标准，并充分利用非结构化数据的价值，进而获得竞争优势和关键数据洞察力。

IBM Spectrum Discover 是一款高级元数据管理解决方案，可针对 EB 规模的非结构化存储提供一流的数据洞察力。它可以轻松连接到 IBM Spectrum Storage 系列产品中的另外两款产品 - IBM Cloud Object Storage 和 IBM Spectrum Scale (我将在第 4 章中介绍此产品)，快速摄入、合并数十亿个文件和对象的元数据并创建索引，进而提供一个丰富的元数据层，以便存储管理员、数据管理员和数据科学家高效地管理和分类海量的非结构化数据，并从中获取一流的洞察力。如此一来，便可提高存储的经济性，改善数据治理，并加速大规模分析，以创造竞争优势并加快关键研究的速度。



切记

IBM Spectrum Discover 可提供以下功能：

- » 通过事件通知与基于策略的工作流实现 PB 规模元数据摄入及元数据索引的自动化
- » 基于广泛的系统元数据和定制化元数据的存储使用量精细视图
- » 可快速、高效地搜索数 PB 的数据，进而生成高度相关的结果集用于大规模分析
- » 能够快速从拟删除的数据或拟移动至成本更低、使用频次更小的存储层的数据中区分任务关键业务数据
- » 基于策略的定制化标记功能可帮助组织对非结构化数据进行高效分类，并使其符合业务需求



技术内容

IBM Spectrum Discover 能够自动从源存储系统捕获系统元数据，基于用户定义的策略构建定制化元数据，而且支持更高级的用例，比如允许使用 Action Agent 应用程序接口 (API) 对文件标题和内容进行深入的内容检查并从中提取元数据。如此一来，便可形成一个由单个集中式解决方案管理的丰富的文件和对象元数据层。



提示

IBM Spectrum Discover 作为 VMware 虚拟设备提供，可以轻松部署，还可与许多 IBM 和非 IBM 的数据管理工具相集成。

## IBM Spectrum Connect

IBM 的相关研究证实，就目前而言，超过四分之三的企业已经部署了某种类型的云计算功能，并且大多数组织已开始采用多云战略。但是多云并不容易实现。新增的每个云都会带来连接性和兼容性问题，更不用说随之而来的诸多管理挑战。在整个 IBM 存储环境中支持和简化多云部署是 IBM Spectrum Connect 的重点所在。

当今的组织都需要在多云环境中实现轻松、快速的存储集成。在单个解决方案中，IBM Spectrum Connect 可帮助您充分利用现有的 IBM 存储功能，而且能够让存储团队及其他利益相关者在容器化环境、VMware 环境和 Microsoft PowerShell 环境中配备、监控、自动化和编排 IBM 块存储。

IBM Spectrum Connect



切记

- » 能够从单个位置管理 IBM 存储系统的 API 对话，提供单一视图，用于实现多云平台与 IBM 存储设备之间的编排
- » 在诸如 IBM Cloud Private 之类的解决方案中使用 Kubernetes 支持容器编排
- » 按服务级别协议 (SLA) 或按应用工作负载定义策略，进而简化容器的存储配置
- » 提供更好的存储可视性，以改善容器化环境中的故障排除流程
- » 允许定义易于使用的存储类（例如按 SLA 或按工作负载），从而简化自助服务以及存储配置的自动化

- » 支持 IBM FlashSystem 系列产品、IBM Storwize 系列产品、IBM DS8000 系列产品、IBM SAN Volume Controller 和 IBM XIV，以及 IBM Spectrum Virtualize、IBM Spectrum Accelerate 和 IBM 与 Cisco 共同推出的 VersaStack 聚合基础架构解决方案
- » 充分利用现有的 IBM Storage 功能，同时免费向 IBM 存储系统的客户提供



技术内容

IBM Spectrum Connect 可将私有云管理功能与多个支持的 API 结合在一起。举例来说，借助 IBM Storage Provider for VMware VASA (vStorage APIs for Storage Awareness)，VMware vSphere vCenter 能够识别和使用各种 IBM 存储系统的功能，从而让 VM 数据存储的分配、维护和监控变得更轻松。

## 本章提要

- » 使用经济高效的磁带机和库
- » 支持高级数据保护
- » 专注于运营恢复和数据复用
- » 加强您的副本数据安全态势

# 第 3 章

## 面向数据保护和服务的 SDS

本章介绍了 IBM Spectrum Storage 系列中可提供广泛的现代数据保护、归档和复制管理服务的几款产品，这些服务对业务和 IT 的成功而言至关重要。

### IBM Spectrum Archive

自从计算出现以来，数字化信息一直都是存储在长条状的磁带上。数十年来，随着硬盘驱动器以及最近的闪存固态存储等新存储技术的出现，许多行业分析师都预测磁带存储将会不复存在。但事实并非如此，而且在 IBM Spectrum Archive 等软件定义存储 (SDS) 解决方案的推动下，磁带存储仍然是许多组织所用存储架构的基石，而且对于云和分析工作负载必不可少的新兴业务解决方案而言，磁带存储也是一种重要工具。

磁带存储之所以能够继续为许多企业提供巨大的价值，其中一个关键原因在于：业务所生成的绝大多数（可能多达 80%）数据在一两个月后就会很少或永远不会被访问。但是由于各种各样的原因，某些数据仍然必须存储，一些情况下甚至需要存储数十年，这其中既有政府法规要求的原因，也是因为企业希望留在随后从这些数据中发掘宝贵业务洞察力的机会。如此一来，这些数据存储可能会变得非常庞大且成本高昂。即使是采用低成本的磁盘存储，单位 GB 存档数据每月也要花费几美分的成本，而使用磁带的话，单位 GB 数据每月的成本只有十分之几便士。由于这些原因，磁带存储 - 尤其是 IBM Spectrum Archive - 总体上对于许多企业而言仍然非常有价值。

IBM Spectrum Archive 基于 IBM 线性磁带文件系统 (LTFS) 技术而构建, 能够以直接、直观、图形化的方式访问存储在 IBM 磁带机及磁带库中使用的 Linear Tape-Open 磁带和 IBM 企业级盒式磁带上的数据。如此一来, 无需额外的磁带管理和软件即可访问存储在磁带上的数据。

IBM Spectrum Archive 有助于简化闪存/磁盘和磁带之间的数据移动, 从而能够降低成本, 无需专门的磁带应用。



切记

IBM Spectrum Archive 为组织提供了一种简便方法, 使其能够在分层式存储基础架构中使用经济高效的磁带机和磁带库。通过使用磁带库 (而非闪存或磁盘存储) 来存储需要长期保留的数据, 组织可以提高效率、降低成本。

IBM Spectrum Archive 可帮助组织

- » 通过磁带构建运营存储层, 而不是将静态、无变化的数据存储在本高昂的磁盘存储上
- » 长期归档数字资产, 如此便可在未来多年确保资产的可参考性和货币化
- » 在运营存储上构建数据副本, 进而改善分层式存储基础架构的效率和成本高效性



提示

IBM Spectrum Archive 提供了三种软件解决方案, 它们可用于管理 LTFS 格式的数字文件, 这三种解决方案分别是: Single Drive Edition 和 Library Edition (这两个版本均可免费下载), 以及最完整的版本 - IBM Spectrum Archive Enterprise Edition。



# IBM Spectrum Protect

出于各种原因，各种类型和规模的组织都开始向多云架构迁移，而数据保护可能是其中的重中之重。借助公有云资源，组织可以大幅简化数据保护解决方案。资源只需根据需要购买即可，基本上没有资本支出。数据可以自动复制到任何位置的多个云存储站点，不仅能够降低数据丢失风险，还能够提高分支机构和远程位置的应用性能。不过，连接到云端、复制和移动海量数据集、最大程度地提高网络性能、利用最具成本效益的云和本地存储层、调度和监控及跟踪备份（以及许多其他任务）可能会导致一些复杂且艰巨的挑战。这些原因以及许多其他原因正是企业选择 IBM Spectrum Protect 的缘故所在。

IBM Spectrum Protect 提供了广泛的功能，包括备份、快照、归档、恢复、空间管理、裸机系统恢复、灾难恢复 (DR) 和数据复用等。它可以通过单个控制点保护各种规模的系统上的数据。您可以下载并部署 IBM Spectrum Protect 来保护从小型到大型各种规模的企业。它能够为云环境、虚拟化环境和软件定义环境提供高级数据保护，还可以用于恢复无数的应用、数据库和文件系统，包括单个文件。

通过将 IBM Spectrum Protect 作为 SDS 解决方案来实施，您可以轻松、安全且经济高效地实现云端备份。借助容器存储池，您可以充分利用对象存储，而无需在主流的云环境（如 IBM Cloud、Amazon S3 和 Microsoft Azure）上使用其他硬件或网关。这些容器存储池可提供内联数据去重和压缩功能，使您可以高效地使用空间和带宽，还可以提供加密功能，确保数据安全。

IBM Spectrum Protect 可以帮助用户大幅降低备份基础架构的成本。一般来说，它可以帮助用户降低存储媒介、备份服务器、数据中心占地面积、耗电及冷却方面的成本。结合使用 IBM Spectrum Protect 的高性能数据去重、压缩和永久增量快照功能，可以帮助您减少备份存储需求。该软件内部完全支持其高效功能，因此，在去重、加密、网络加速或云访问方面，完全不需要其他基于硬件的设备。为了提供最大的成本灵活性，IBM Spectrum Protect 还为备份数据提供了广泛的存储选项，包括闪存、磁盘、磁带和对象存储等。



提示

IBM Spectrum Protect 所提供的指导性蓝图和配置脚本可实现大中小各种环境的自动部署并集成最佳实践，进而帮助您缩短部署时间、减少猜测。经简化的解决方案捆绑包（例如 IBM Spectrum Protect Suite）提供了许多易于管理的许可选项，包括快照以及面向虚拟环境、数据库、电子邮件和企业资源计划的主流备份代理。IBM Spectrum Protect 支持数据量高达数 PB 的大型环境，不过小型组织也可以从中受益。对于托管服务器数量不足 50 台或备份数据量不足 100 TB 的企业而言，可以使用 IBM Spectrum Protect 入门级解决方案产品。

## IBM Spectrum Protect Plus

当今的数据中心正在从手动任务转向自动化、简单化和敏捷化。IBM Spectrum Storage 系列 SDS 解决方案正在通过 IBM Spectrum Protect Plus 等产品引领着这种转变（IBM Spectrum Protect Plus 代表着 IBM 数据保护产品组合中的下一波潮流）。IBM Spectrum Protect Plus 采用了现代化的数据保护方法：专注于数据的运营恢复和复用（而非仅仅专注于备份解决方案）。

IBM Spectrum Protect Plus 能够保护各种虚拟环境，尤其是 VMware ESXi 和 Microsoft Hyper-V 环境。IBM Spectrum Protect Plus 还支持文件和数据库恢复，以及灾难恢复所需的多站点数据复制。IBM Spectrum Protect Plus 的数据副本可经过调整用于多个数据拷贝用例，包括应用开发/测试环境的配备、分析与报告支持等。该软件不到一个小时即可上线并运行，而其他备选的方案，不仅需要耗费数周才能完成部署，而且需要成本高昂的专业服务。

IBM Spectrum Protect Plus 由两个组件构成，分别是 IBM Spectrum Protect Server 和 vSnap 存储库。内置的数据去重和压缩功能都包含在 vSnap 存储库中，而且在与永久增量快照结合使用时，IBM Spectrum Protect Plus 还可以大大减少存储需求。快照存储为本地 VM 映像，通过载入就可用于恢复，进而提供对受保护数据的近乎即时的访问。若要长期保留数据，可以将 IBM Spectrum Protect Plus 快照卸载到 IBM Spectrum Protect，同时它能够充分利用所提供的各种存储介质选项，包括磁盘、磁带、云或对象存储等。

IBM Spectrum Protect Plus 能够创建并维护所有虚拟机、应用和文件副本的全球性目录，方便管理员了解受保护的对象；更重要的是，一旦发生数据丢失，强大的搜索功能和该全球性目录可帮助管理员快速地精确识别想要使用/访问的副本，无需手动浏览数百个对象及多个时间点副本。管理员可以恢复整个 VM 或快速恢复单个文件。



切记

IBM Spectrum Protect Plus 的关键在于其“选择 + 点击 + 完成”的操作简便性。配置可以像选择恢复服务级别一样简单：从预定义的黄金、白银或青铜 SLA 中选择一种，或自定义创建自己的 SLA；然后将该服务级别应用到一台或多台虚拟机；最后即可完成。如果您遗漏了任何一台虚拟机，仪表板视图可帮助您快速识别它们，确保被覆盖在内。该解决方案具有鲜明、现代化的观感，其中配有一个仪表板，可让您一目了然整个保护环境。您可以在单个位置查看有关保护状态、SLA 合规性及存储利用率的信息。



提示

您不必成为备份专家即可使用 IBM Spectrum Protect Plus。该软件易于部署，并且可以通过安装预构建的虚拟设备在数分钟内完成使用准备。您既可以将其用作实现敏捷、简单的 VM 和数据库保护的独立解决方案，也可以将其集成到 IBM Spectrum Protect 环境中，获得全面的数据保护解决方案。

# IBM Spectrum Copy Data Management (CDM)

现代企业已经无法继续维持常规的数据副本创建和管理方法。在许多情况下，传统方法会导致整个 IT 环境中构建并存储大量的数据副本，包括主要实例及各种备份、快照、副本和数据库转储文件等。

IBM Spectrum CDM 可帮助您改变这一状况。借助 IBM Spectrum CDM 的 SDS 功能，企业可以构建高效的存储环境，进而显著提高整体数据经济性。IBM Spectrum CDM 提供了一个领先的数据和存储服务套件，该套件可以简化副本管理，并加速 21 世纪企业的敏捷性和竞争力。

IBM Spectrum CDM 可以帮助企业提高运营和存储效率，继而降低成本和复杂性。它所提供的服务套件可帮助企业跟踪、构建、刷新和管理数据副本的使用。这些通用服务还可用于数据中心中许多不同用例，包括灾难恢复、应用开发和测试、分析、归档等。借助这些适于多个用例的通用服务，管理员可以对整个数据环境进行编目和优化管理，进而降低成本和复杂性。

IBM Spectrum CDM 能够随时随地为数据使用者提供副本，使您无需创建不必要的副本，也不会导致不使用的副本浪费宝贵的存储资源。该解决方案在带外运行，即在生产数据路径之外运行，因此不会造成性能问题，能够创建所有存储、应用和 VMware 环境的目录。借助该目录，IT 管理员可以管理、编排并分析数据，以释放数据价值并满足业务需求。数据使用者可以通过自助服务门户来构建所需副本，进而实现业务敏捷性。

IBM Spectrum CDM 可帮助 IT 管理员编排原本复杂的数据复用场景并实现其自动化。通过该解决方案的“使用数据” workflow，管理员可以快速交付数据副本，并消除需要应用所有者向 IT 部门提交凭单并等待响应的手动流程。诸如灾难恢复之类的业务操作可以实现自动化，进而可以每天对其进行测试和验证。IBM Spectrum CDM 的强大功能使得这些业务运营中所用的资源可以在隔离的环境中调动，只需按下按钮即可快速提升到生产环境，或者在测试后撤除并清理。编排可提高可重复性和可审计性，从而使企业可以将单个数据副本用于多种目的，减少数据散乱情况并降低成本。

IBM Spectrum CDM 还支持多云架构。它不仅能够帮助您将数据迁移到云端，还有助于建立活跃的应用环境，使您可以充分利用极具弹性的云端计算基础架构。您可以可靠地启动和关闭工作负载。通过这种方式，您便可根据需求使用适当的基础架构并支付相应费用，最大程度地实现云的经济优势。



提示

如今，大多数 CIO 和其他业务领导者都非常关注发生数据泄露事件（会导致敏感数据丢失的）的可能性。副本数据是最容易导致组织受到攻击的向量之一，因为通常不会对其进行跟踪，有时也不会擦除其中的敏感数据，而且也不总是会对其进行应有的安全保护。这就是 IBM Spectrum CDM 可以发挥作用的地方。通过部署 IBM Spectrum CDM，您可以使用可靠的数据屏蔽功能，以此来增强您的副本数据安全性。此功能使用基于策略和基于角色的筛选功能，以防止敏感数据遭到泄露。

## 本章提要

- » 通过虚拟化“魔力”实现存储转型
- » 创建单个高性能的全局文件和对象存储解决方案
- » 部署内部对象存储和基于云的对象存储
- » 充分利用高度并行网格架构的优势
- » 建立高度可扩展的网络连接存储

# 第 4 章

## 面向基础架构的 SDS

在本章中，我将介绍 IBM Spectrum Storage 系列产品中可提供基本存储基础架构功能的产品。

### IBM Spectrum Virtualize

在传统的企业数据中心的存储容量通常被隔离或限制在由各个应用配备和拥有的不同系统孤岛中。由于无法根据需要将存储资源移动到其他应用和工作负载，导致管理点重复、容量利用率低下。

十五年来，IBM Spectrum Virtualize 一直致力于通过虚拟化“魔力”将不灵活的孤岛式存储系统转变为敏捷、强大的存储资源。从本质上讲，存储虚拟化只是通用 SDS 范式的一种特殊体现：在应用主机及其数据之间的数据路径中插入软件。IBM Spectrum Virtualize 通过其虚拟化“魔力”为其支持的所有存储系统提供了单个管理点，从而实现更好的利用率和性能，同时将其市场领先的服务和功能集扩展到所有这些存储系统。

IBM Spectrum Virtualize 的功能包括：一系列数据服务，这些服务可扩展到超过 440 种 IBM 与非 IBM 异构存储系统；自动化数据移动；同步及异步拷贝服务（拷贝到内部系统或公有云）；高可用性配置；存储分层及数据降维技术等等。

IBM Spectrum Virtualize 软件可以交付一整套高级存储功能，包括：

- » IBM Easy Tier AI 驱动型自动分层
- » 加密功能，有助于提高现有存储上的数据安全性
- » IBM FlashCopy 及远程镜像，有助于实现本地复制和远程复制
- » 支持使用云存储来补充内部存储
- » IBM HyperSwap 高可用性配置，可在不造成业务中断的情况下实现应用和数据在数据中心之间的移动
- » 支持虚拟化和容器化服务器环境，包括 VMware、Microsoft Hyper-V、IBM PowerVM、Kubernetes 和 Docker

IBM Spectrum Storage 系列中的基础性产品与 VMware vCenter 等虚拟化工具进行了集成，有助于自动配备存储、轻松部署新存储技术，进而提升敏捷性。它还让支持的存储能够与 Kubernetes 和 Docker 容器环境一起部署。



切记

为了进一步推动您的 IT 转型，IBM Spectrum Virtualize for Public Cloud 还提供了多种方式，以帮助您在内部存储与公有云之间构建混合云解决方案。它支持基于存储的实时数据复制及灾难恢复，也支持本地存储与 IBM Cloud 之间的数据迁移。由于 IBM Spectrum Virtualize for Public Cloud 具有 SDS 特性，因此无论采用何种类型的存储，云服务提供商均可采用与内部环境相同的方式来管理存储。



技术内容

容器是一种开源技术，能够将软件与其在任何环境中运行所需的所有组件一同进行打包。容器通常可实现虚拟机的通用性，同时占地面积小、成本低。因此，就灵活地向私有云和公有云交付工作负载而言，容器化是一个关键的支持性技术。结合 IBM Spectrum Connect（参见第 2 章），IBM Spectrum Virtualize 可以让在其管理之下的所有存储系统成为容器环境中的有效组件，进而提升灵活性、简化部署并进一步降低成本。

IBM Spectrum Virtualize 可提供强大的数据降维功能，其中包括数据块去重功能，它能够最大程度地减少所存储数据副本的数量，还包括硬件加速数据压缩技术，它能够针对各种应用工作负载范例提供统一的高性能结果。



切记

IBM Spectrum Virtualize 可作为软件在 Intel x86 服务器上运行，同时也是 IBM Storwize 系列产品、IBM FlashSystem 9100 和 V9000、IBM SAN Volume Controller 以及来自 IBM 和 Cisco 的 VersaStack 聚合基础架构解决方案的核心。

## IBM Spectrum Scale

如今，您的信息生态系统会涵盖不同的数据类型和存储介质。您不仅拥有结构化数据，还拥有使用并生成这些数据的数据库和事务处理应用。您的大部分数据都是非结构化数据；您需要文件系统来管理这些数据。此外，您也可以使用对象存储。

IBM Spectrum Scale 是一款 SDS 基础架构解决方案，旨在针对业务活动可能生成的所有类型的数据（包括结构化数据、非结构化数据和对象）提供高性能、高功能的数据管理。



技术内容

文件和对象中都包含有与其自身相关的信息（我们称之为元数据），这些信息与文件和对象中所包含的数据相关联，但是对象的特征在于其扩展的元数据。每个对象都分配有一个唯一标识符，通过该标识符，服务器或最终用户无需知道数据的物理位置就可以检索对象。



对于云计算环境中的数据存储自动化和简化而言，这种方法非常有用。

IBM Spectrum Scale 是一个功能齐全的文件数据管理工具集，包括高级存储虚拟化、集成式高可用性、自动分层存储管理和高性能配置，可高效管理数量巨大的文件数据。它能够通过各种访问协议支持广泛的应用工作负载，而且经实践证明，非常适于要求苛刻的大型环境。

该款解决方案与其他存储解决方案（必须借助 Spark 或 Hadoop 之类的工具实施单独附加系统来处理文件、数据对象或大数据分析）有所不同，所有数据不管其访问方式如何，都以通用方式存储在 IBM Spectrum Scale 中。

在成为核心文件系统的一部分之后，基于文件和基于对象的数据都能够以基本上相同的方式进行访问和管理。可以适当地在所有数据上运用广泛的数据存储服务 and 功能，包括快照、信息生命周期管理、存储分层、异步或同步数据复制，以及一些独特的数据保护战略。

IBM Spectrum Scale 几乎可通过所有的标准存储协议（包括 OpenStack Swift、Amazon S3、CIFS、NFS、HDFS 和 POSIX）提供对文件和对象数据的原生、高性能、可扩展访问。借助 IBM Spectrum Scale，多个系统和应用可以共享公用存储池，从而使您可以透明地管理基础架构，同时确保不会中断应用。IBM Spectrum Scale 与平台无关，因此无论是在 IBM Power Systems 和 x86 机器上，还是在来自 IBM 和其他供应商的存储系统上均可运行。这种灵活性有助于降低成本并提高能效。



切记

IBM Spectrum Scale 提供了诸多独特的功能。借助这些功能，您可以：

- » **加速文件和对象存储性能。** IBM Spectrum Scale 可提供对数据和共享存储的并行访问，进而提升高性能工作负载的可扩展性。数据和元数据会在高级的分布式锁管理器的控制下，从 IBM Spectrum Scale 管理的存储节点流至并行的所有存储，该管理器通过每次仅允许一个用户进行更改的方式来防止文件损坏。

- » **简化数据管理。** IBM Spectrum Scale 包含有多个集成式工具，可帮助您控制成本，或管理高达数 PB 数据、数十亿个文件的存储增长。这些是在单个易于管理的整体资源池或名称空间中实现的，此类资源池或名称空间非常易于管理，也可以根据需要快速扩展，只需要添加更多的向外扩展资源即可，如此一来，便可避免每次需要更多容量时都要添加全新文件系统而导致的“文件管理器蔓延”问题。IBM Spectrum Scale 能够自动将文件数据分布在多个存储设备上，进而最有效地利用所有可用存储并提供高性能。它具有一个易于使用的通用管理界面，并且可以通过 IBM Spectrum Scale 系统或集群中的任何节点来管理操作，从而实现用户配额、快照和存储管理等标准文件系统管理功能。
  
- » **促进全球协作。** 借助活动文件管理 (AFM) 分布式缓存和路由技术，IBM Spectrum Scale 可实现对世界各地数据的低延迟访问。AFM 可跨越地理距离对 IBM Spectrum Scale 全局名称空间进行扩展，进而能够从任何位置提供快速读写性能以及自动命名空间管理。当数据在某个位置进行编写或修改时，所有其他位置均能够以最小的延迟得到相同的数据。AFM 可利用 IBM Spectrum Scale 固有的可扩展性，提供一款与位置无关的高性能解决方案，该款解决方案可屏蔽故障并隐藏广域延迟和中断。这些功能可加速全球范围内团队的项目进度并提高效率。
  
- » **支持多云架构。** 多云解决方案有助于防止系统故障，它能够利用来自各种提供商的功能或定价模型来优化解决方案，而且可将存储资源放置在不同的位置，以最大程度降低全球最终用户的延迟。IBM Spectrum Scale 提供了许多功能，可帮助您构建和管理多云环境。透明云分层支持在闪存、磁盘、磁带和云存储层之间基于策略的无中断智能迁移，从而允许您通过统一视图管理和访问内部数据和云数据。借助云数据共享功能，您可以将数据从云存储导入到 IBM Spectrum Scale 文件系统，或将数据从 IBM Spectrum Scale 导出到云端，以供其他系统访问。



提示

IBM Spectrum Scale 可以加载到任何适用的硬件上。它之所以如此命名，是因为它使用大规模并行网格架构来获得更高的性能和更高的系统弹性。这意味着您只需要简单地添加更多的服务器和存储，IBM Spectrum Scale 便会自动将其合并到单个存储资源中，并将管理和数据都分布在所有可用的组件或节点上，以实现性能、效率和数据保护的最大化。作为真正的 SDS，它还可以在存储系统 (IBM Elastic Storage Server) 或 IBM Cloud 和 Amazon Web Services 上使用。

## IBM Cloud Object Storage

您每次将照片发布到 Facebook、在 Spotify 上听歌或在 Box 上共享文件时，您都是在使用对象存储。对象不同于传统的数据文件，因为它们包含有额外的元数据，这些元数据使得对象存储系统可以从根本上消除分层组织，例如文件夹和目录等。这项强大的技术旨在满足快速增长的存储需求，与传统存储技术相比，它能以高度可扩展、更可靠高效、极具成本效益的方式存储大量的非结构化数据。

对于在线交易处理、实时浏览会话期间的客户趋势分析或实时欺诈检测和防范等高性能应用而言，对象存储解决方案并不合适。相反，对象存储特别适合与全球性企业、遗传学、天文学、宏观经济学和物联网 (IoT) 相关的大量数据集。对于所有这些用例而言，对象存储都非常高效，因为它易于扩展和访问，并且可以降低存储成本。

IBM Cloud Object Storage 既可以提供内部部署的对象存储解决方案，也可以提供基于云的对象存储解决方案。它能够交付提供连续可用性、保护和数据资产访问所需的功能。借助该款产品，您可以在需要的时间和地点利用数据，以改善业务流程、决策以及对法规/法律要求的响应能力。

IBM Cloud Object Storage 有助于降低存储成本，同时还能够可靠地支持传统和新兴的移动、社交、分析和 AI 应用工作负载。实际上，与传统存储相比，IBM Cloud Object Storage 可以节省 90% 的存储相关总成本，这一点已经得到了行业分析师的确认。



技术内容

IBM Cloud Object Storage 技术采用了一种创新的方法，能够以成本高效的方式存储大量非结构化数据，同时确保安全性、可用性和可靠性。它所采用的独特信息离散算法能够将数据分离成不可识别的“切片”，而这些“切片”通过本地或全球存储节点的网络连接分布。借助 IBM Cloud Object Storage 技术，企业能够从根本上确保数据传输与存储的安全性。没有完整的数据副本存在于任何一个存储节点中，只要节点的子集可用，便可完全检索数据。这种方法既消除了基于 RAID 存储的高昂开销，也消除了传统数据中心中镜像、复制和灾难恢复等功能的管理复杂性。

默认情况下，存储在 IBM Cloud Object Storage 中的所有对象都是加密的。静态数据通过高级加密标准 (AES) 256 位加密和安全哈希算法 (SHA)-256 哈希进行加密，而动态数据则通过内置的运营商机级传输安全层 (TLS)/安全套接字层实现安全保护。

## IBM Spectrum Accelerate

IBM Spectrum Accelerate 能够让您通过与传统系统不同的存储架构来实现块存储。常规存储使用被称为*控制器*的计算机来管理“磁盘盒”或存储机柜。此类存储通常会配备两个或多个控制器；如果某个控制器出现故障，则系统可以继续运行，不过性能可能会有所降低。相反，IBM Spectrum Accelerate 使用的是网格存储架构，在这种架构中，相同的构建块或节点基本上都同时拥有控制器和存储。

在这些高度并行的网格存储系统中，添加另一个构建块或节点可增加存储量、CPU 能力、内存缓存和联网功能，进而能够让您获得更多的容量和性能。节点彼此“交谈”并分担存储管理职责；此外，它们能够自动在所有节点中的所有存储资源上散布或“划分”传入的数据。本质上来说，IBM Spectrum Accelerate 解决方案中的所有节点都可以协同工作，进而构建一个能够随着您添加更多构建块而自动增长的存储系统。

这种方式带来了一些非常可观的效果。举例来说，由于系统会执行自动数据分配和负载均衡，因此无需进行存储“调优”；如此一来，性能会始终保持一致、可预测而且没有热点。架构和内置机制（如自我修复）能够使系统提供高可用性。由于相同的原因，保护数据也不是问题 - 随着数据的到达，多个副本将会在整个系统中复制。如果某个节点发生故障，数据将会自动重新分配，并且可以从其他节点获得。此外，添加到网格系统中的节点越多，它自然就会变得越可靠、越具弹性，速度也更快。



提示

与 IBM Spectrum Virtualize（有关更多信息，请参见前文“IBM Spectrum Virtualize”一节）一样，IBM Spectrum Accelerate 也提供了一套久经验证的丰富存储服务功能，而且这些功能均为免费提供。它与市场上最受欢迎的应用侧虚拟化产品相集成。对于希望利用异构服务器硬件并轻松快速地为云计算环境向外扩展经济高效的存储的企业和服务提供商来说，IBM Spectrum Accelerate 可为其提供诸多优势：

- » **非常简单**：独特的网格架构可提供高级数据分配和自我修复功能，这些功能消除了对规划和复杂管理的需求，从而只需要比传统存储系统所需更少的人员配备和培训，即可更有效地管理企业级存储容量。
- » **远程复制**：镜像功能通常用于让灾难恢复站点填满当前数据；如此一来，如果主数据中心发生故障，远程站点便可接管，确保将故障对业务的影响降到最小。IBM Spectrum Accelerate 可提供异步和同步镜像功能，为需要持续运行的应用提供高可用性。
- » **多租户和服务质量 (QoS)**：在云计算领域，一个存储设备可以满足多个用户或“租户”的需求。在这种情况下管理资源并确保所有租户都能获得其所期望的存储容量和服务级别，就需要额外的 SDS 功能。

- » **快照：**数据保护是任何企业级存储解决方案的重中之重。协调快照、采用对容量资源影响最小的方式存储快照，以及在系统恢复期间管理快照的使用都属于有价值的存储服务。
- » **监控：**SDS 提供的一个关键益处在于详细地或“粒度化”地自动收集与存储系统运行相关的信息。此类信息会报告给 IT 经理，而 IBM Spectrum Accelerate 可使用此类信息来执行其他自动化存储服务，例如在网格节点发生故障后重建额外的数据副本。
- » **安全：**IBM Spectrum Accelerate 可提供轻量级目录访问协议 (LDAP)、基于角色的访问以及静态数据加密。
- » **规模化管理：**IBM Spectrum Accelerate 配有一个基于多项专利的创新型用户界面。对于混合云解决方案而言，重要的一点是，它采用了 IBM Hyper-Scale 技术，可帮助您扩展所需的存储系统，通过单个视图进行管理，并灵活地将数据卷从一个系统迁移到另一个系统，无论您的云使用模式为何。



技术内容

IBM Spectrum Accelerate 可在 VMware vSphere 上作为虚拟机运行，使您能够通过商用硬件构建基于服务器的 SAN。它能够通过底层物理磁盘对虚拟节点进行高效分组，并将数据均匀分布在各个节点上，从而创建单一虚拟阵列。该款解决方案能够以成本高效的方式使用标准数据中心网络，从而实现节点间以及与应用主机的互联性。IBM Spectrum Accelerate 可支持满足最低要求的任何硬件配置和组件，并且不需要明确的硬件认证。节点的扩展是线性扩展而且不会导致业务中断。



提示

如果您想要构建其功能可轻松扩展、经济高效且可提供灵活性的云存储解决方案，IBM Spectrum Accelerate 绝对是不二之选。IBM Spectrum Accelerate 隶属于两个 IBM Storage 产品系列（IBM FlashSystem A9000/A9000R 和 IBM XIV Storage Systems），可提供私有云、公有云和混合云组合式解决方案。

# IBM Spectrum NAS

文件系统最初是为了向非结构化数据赋予结构而开发的。文件系统能够以文件、文件夹和目录的形式提供分层组织，进而使构成图像、文档、电子邮件等的比特和字节变得更易存储和查找。

网络连接存储 (NAS) 是一种专用的文件存储，它支持多个用户和异构客户端设备从集中式存储中检索数据。局域网 (LAN) 上的用户可通过标准以太网连接访问共享存储。NAS 的最大特点在于其易于访问性、大容量和相当低的成本。NAS 设备所提供的基础架构能够将存储整合到单个位置，而且支持归档和备份等任务。

不过，NAS 的主要缺点在于，每个 NAS 系统作为一个独立的网络节点驻留在 LAN 上，由其自己的唯一互联网协议 (IP) 地址所定义。本质上来说，要增加更多的存储、性能和功能，必须先部署一个额外的 NAS 系统，然后在需要时再部署另外一个，依此类推。这种模型不易于增长/扩展；在当今非结构化数据量爆炸式增长的时代，必须确保存储的可扩展性。

这就是 IBM Spectrum NAS 可以发挥作用的地方。它为文件系统提供了一款高度可扩展的 SDS 解决方案。它可以加载到您所选的廉价服务器（如“商用”x86 计算机）上，将其转换为 NAS 存储设备或“节点”的组/网格/集群，这些设备或节点协同工作，进而提供一个 NAS 解决方案。IBM Spectrum NAS 可用于管理节点集群，其中每个节点都拥有自己的 CPU、内存、存储和缓存，所有节点共用单个文件系统。每个节点都会增加集群的访问点、缓存、存储容量和性能。IBM Spectrum NAS 的架构从本质上来说非常均衡，因为每个节点在存储集群中的作用完全相同，也就是说没有专用的元数据或特殊节点。用户和应用可通过任一节点访问集群，不仅确保了负载的均匀分布，还有助于消除瓶颈。该架构结合数据保护机制（如纠删码和快照）使用的话，可以避免单点故障，进而提高可靠性。

IBM Spectrum NAS 采用软件定义的方法，因此支持组织实现从固定的、刚性的、以软件为中心的方法转变为灵活且实用的方法。随着存储或带宽需求的提升，可向集群中添加更多节点，分担相应的计算和存储负载。通过这种方式，可确保目前的部署在未来依然能够发挥作用。

集群中的节点数量越多，集群的容量和性能就越高。节点数更多意味着，当出现节点故障时，能够使用更多的选项来重新均衡负载、调整自愈流程。应用可通过 IBM Spectrum NAS 虚拟文件系统访问数据，因此允许多个客户端与相同的存储数据进行交互。IBM Spectrum NAS 允许多个单独的域和文件系统跨越单个存储集群，因此可简化管理，降低成本，同时相比更传统的 NAS 方法，存储资源的使用效率更高。



提示

IBM Spectrum NAS 专为实现轻松部署与管理而设计。每个节点可以是“裸机”，也可以是虚拟机。它配有一个向导式指南，使您可以在 30 分钟内完成设置，而且支持无中断的网络升级。由于每个节点均采用相同的软件，因此可进一步简化安装和维护。



- » 将 AI 融入到存储中
- » IBM Spectrum Storage 系列产品的创新使用

## 第 5 章

# IBM Spectrum Storage 系列解决方案的 10 个用例

本章介绍了各种规模和类型的企业利用 IBM Spectrum Storage 系列解决方案和技术实现其存储目标、解决实际业务挑战所采用的 10 种方式（实际上是 8 种）；有关本章所提及不同组件的详细信息，请参见第 2、3、4 章。

## 充分利用 AI

据行业分析师预测，在未来几年内，大多数企业将会实施某种形式的人工智能 (AI)，以期提高生产效率、管理风险并推动 IT 基础架构内的成本降低。人工智能有望改变我们生活的许多方面。如果您所在企业还未开始的话，我相信您的业务组织或 IT 小组很快将会开始研究您的企业可能从 AI 中获得的收益和优势。在您了解 AI 是什么以及它现在可以做什么时，您可能会问的第一个问题就是实际上如何部署 AI？

然后，您可能会想知道您在 IT 基础架构中使用 AI 的目的，以及它能够带来的益处。

IBM Spectrum Storage 提供了一个简单而强大的解决方案：总体而言，IBM Spectrum Storage 解决方案将 AI 融入其中。举例来说，IBM Storage Insights 就利用了 AI 技术，您只需要部署一款 IBM 存储解决方案，即可实现 AI 的收益。它能够从多个来源收集数据 - 这些来源既有来自您的 IT 基础架构内部的来源，也有来自全球数以千计的其他 IBM 客户所构成的 IT 环境的来源，然后不断处理大量信息，从中寻找趋势、最佳实践、重复出现的问题、潜在问题，该过程要远远早于人工方法。然后，您和您的 IBM 支持团队可以评估建议并规划如何实施（如果您选择的话）。

同样地，IBM Spectrum Virtualize 可以使用 AI 技术来评估工作负载随时间推移的变化趋势，并决定在存储层之间移动哪些数据。在这些情况以及其他许多情况下，IBM Spectrum Storage 都可以将 AI 功能的优势添加到存储中，甚至可以添加到您可能已经在数据中心部署的非 IBM 存储中。

## 支持敏捷性

*业务敏捷性*是指您的公司能够尽快地添加、增强、更改、消除或修复您所售产品，以适应快速变化的市场趋势的能力。这就是“容器”这种新技术能够发挥作用的地方。企业对容器技术的兴趣很高，并且正在快速增加。容器按照灵活、轻巧、便携的原则来设计，将会被用于运行从传统数据中心和云数据中心到汽车、游轮、机场航站楼、物联网 (IoT) 网关等各种环境中的所有应用。容器能够提升面向业务及其他用例（如科学研究）的应用的敏捷性。

IBM Spectrum Connect 提供了 IBM Spectrum Storage 系列产品中的技术，使您能够使用容器来改善业务和 IT 敏捷性。IBM Spectrum Connect 充分利用了现有的 IBM 存储功能，同时免费向 IBM 的客户提供。它能够与 Docker、Kubernetes、IBM Storage Enabler for Containers 和 IBM Cloud Private 等工具和解决方案一同使用，以简化容器的使用。



切记

此外，它还增强了容器化环境的可视性，进而提高利用率和故障排除能力。

业务环境并非容器能够提升敏捷性的唯一场景。我们以科学研究场景为例。举例来说，在基因组学中，应用的开发生命周期较短，保存期限更短。环境在不断演变，一种工具很快就会被新的变体或其他工具完全取代。借助容器，可以在独立于特定操作系统或其他应用的情况下完整封装研究应用。用户可以在同一主机上运行不同的软件版本，而不必担心冲突。此外，新的软件包和应用可轻松地按需推出计算节点，从而有可能减轻管理员的应用管理负担。

## 通过现代化提升业务弹性

大多数业务应用都被视为任务关键型应用。对于这些应用来说，组织需要“永续”的高可用性 IT 基础架构解决方案，以满足其特定的业务需求、法规要求和灾难恢复 (DR) 战略。即便是一个小时的宕机，也会产生数千甚至数百万美元的成本，而且可能会损害品牌价值，因此关键业务数据的恢复点目标 (RPO) 及恢复时间目标 (RTO) 必须在预算允许的前提下尽可能地接近于“零”。

基于公有云的业务连续性解决方案有助于最大程度地减少资本超支，同时维持宝贵数据的可访问性。如此一来，企业便可将生产数据存储在内网系统，以确保安全性、改善系统性能，同时在几乎不投入资本开支的情况下，即可在云端配备 RTO/RPO 几乎为零的冗余基础架构。市场领先的 IBM Spectrum Storage 解决方案可提供各种高效方式，帮助您实施基于云的业务连续性解决方案。

举例来说，可以使用 IBM Spectrum Virtualize 和 IBM Spectrum Virtualize for Public Cloud 实施创新型业务弹性解决方案，以支持内部系统和公有云之间的数据镜像。借助云端的数据副本，您可以选择从云端快速恢复或将数据从云端复制到另一个数据中心。

同样地，针对超高可用性的要求，IBM Spectrum Virtualize 和 IBM Spectrum Accelerate 中的 IBM HyperSwap 支持具有自动故障转移功能的主动/主动、潜在多站点配置。

通过部署这些 IBM Spectrum Storage 系列产品，您可以借助功能强大的业务连续性解决方案，确保满足所有的业务服务级别协议 (SLA)。

## 实现数据价值的倍增

业务运营会产生各种数据。每次交易、每次购物，甚至是潜在客户浏览的每个网页，都会生成信息。在过去，业务活动产生的数据通常会被丢弃，即便是最好的情况，也是仅仅对其进行归档。现在，高管们已经认识到，数据在最初使用之后还可以提供巨大的业务价值和竞争优势。数据可以多次复用，而且通过每次复用，它可以为企业带来越来越多的价值。

您可能不会这么认为，但为了实现数据保护和备份而制作副本，进而实现灾难恢复，的确是数据复用的形式。举例来说，由业务应用生成的现有数据集可被开发团队重新用于测试目的。通过发掘数据资产以获取业务洞察力，已变得极具价值。此外，出于展示监管合规性等各种原因，数据副本还可用于满足报告需求。

由于它们的深度集成和兼容性，IBM Spectrum Storage 系列产品可以轻松地组合在一起，进而构建出可同时解决多个挑战和需求的解决方案。多云、数据复用和现代数据保护功能可能是最为常见的 IT 需求组合。几乎每个企业都希望同时满足这三个需求。结合使用 IBM Spectrum Virtualize、IBM Spectrum Copy Data Management (CDM) 和 IBM Spectrum Protect Plus 可满足所有这些需求，甚至还可以满足更多需求。通过结合使用这些组件，可提供强大的数据复用、保护、高效及多云功能。

## 结合采用闪存

在使用 SDS 之前，闪存存储的部署在某些情况下极具挑战。举例来说，采用手动方式将数据集移入或移出闪存的流程速度会很慢且劳动强度大。这种做法完全抹杀了闪存本身在降低成本方面的价值。不过，如果您将 IBM Spectrum Virtualize 所提供的智能、自动存储分层功能添加到其中时，数据将根据活动级别或您选择的其他策略快速移动到最有利的存储介质上。如此一来，就会大幅提升闪存的价值。您的应用会获得它们所需的性能 - 不过也仅仅是在需要时，而且在不需要使用高性能闪存且磁盘或磁带即可满足要求时，预算就会节省下来。随着文件存储时间的增加，IBM Spectrum Scale 中基于策略的分层功能可提供类似的功能。在更大规模的数据移动方面，借助 IBM Spectrum Virtualize 和 IBM Spectrum Accelerate 中的无中断迁移功能，您可以将整个存储系统迁移到新环境。

## 支持混合云

除了其网格架构固有的优势外，IBM Spectrum Accelerate 还提供了其他诸多功能，因此可用作构建混合云存储的一种有效工具。举例来说，通过使用 IBM Hyper-Scale 系列中的各种工具，您可以自动将数据复制到异地的 IBM Spectrum Accelerate 或 IBM XIV 存储实例。借助 Hyper-Scale 工具集，您几乎可以永久地扩展存储资源、构建灾难恢复解决方案并支持混合云。

使用 IBM Spectrum Accelerate 构建的混合云存储解决方案具有许多优势，关于这一点，我将在第 4 章中予以介绍。

## 借助混合多云实现数据保护的现代化

数据保护解决方案需要解决您的所有工作负载，包括物理文件服务器、虚拟化环境以及物理/虚拟数据库及应用。更进一步来说，为了确保在本地和异地保护所有数据，您可以在混合多云环境中利用 IBM Spectrum Protect 和 IBM Spectrum Protect Plus 的优势。

IBM Spectrum Protect 和 IBM Spectrum Protect Plus 都可以利用内部部署的 IBM Cloud Object Storage 和公有云来卸载数据，以实现长期数据保留。IBM Spectrum Protect Plus 还可以将 IBM Spectrum Protect 用作 Amazon 云目标存储库，以进行数据卸载。



提示

对于需要确保更高级别数据安全性的企业而言，IBM Spectrum Protect Plus 还支持 IBM Cloud Object Storage 免疫存储功能，例如支持数据保留的存储段。在指定的保留期内，存储在支持数据保留的存储段中的对象是无法删除或更改的。您可以在存储段级别预定义保留期，也可以使用保留策略为每个对象设置保留期。

## 简化大数据分析

如果您参与过企业 IT 管理或决策的话，就会知道：您的企业迟早会使用某种形式的大数据分析。当然，SDS 在大数据分析中发挥着重要作用，这意味着一个值得加以强调的用例。

IBM Spectrum Storage 软件可以协同工作，提供完整的分析管道解决方案。我们以无人驾驶汽车为例，它会生成大量数据。当数据从车辆中流出时，IBM Cloud Object Storage 可提供理想的存储库，使您可以按照法规要求长时间地存储大量信息。IBM Spectrum Discover 会在提取数据时对数据进行索引处理，然后数据分析师可以使用 IBM Spectrum Discover 来查找所需的特定数据。这些数据会被复制到 IBM Spectrum Scale，该产品可提供支持 AI 工作负载所需的高性能。最后，分析结果可以存储在 IBM Cloud Object Storage 中进行长期保留。

如要解决需要分析的数据量不断增长的问题，以及现有的多个分析框架的相关问题，您可以选择 IBM Spectrum Storage SDS 解决方案，这些解决方案不仅灵活，而且其容量和性能可以轻松扩展，同时支持最新的开源工具和商用工具，比如 Spark、Hadoop 和新的 AI/机器学习/深度学习框架。

# 使用 SDS 系统满足您的业务需求

对于面对日益增长的数据挑战的企业而言，传统的存储系统已成为成本高昂的瓶颈。本书介绍了软件定义存储 (SDS) 如何让组织借助可执行基本存储功能的智能软件来大幅降低其存储成本，同时提高性能、可靠性和可扩展性。

## 本书介绍了...

- 提高灵活性
- 简化管理流程
- 促进全球协作
- 解决性能瓶颈
- 部署一站式 SDS 解决方案
- 发现 SDS 用例



Chris Saul 是 IBM Spectrum Storage 的营销经理。他协助推出了多款 IBM 存储产品，包括 2015 年推出的 IBM Spectrum Storage、2010 年推出的 Storwize，以及 2003 年推出的 SVC。他从 1981 年开始就一直在 IBM 工作，担任过与 IBM 存储软件和存储系统相关的各种职位。

如需获取相关视频、分步骤演示照片、指南文章或购买本书，敬请访问

**Dummies.com®!**

ISBN: 978-1-119-55407-3

部件号: TSM03044USEN-

01

© 2019 IBM Corp.

## 傻瓜系列

Wiley



还可提供电子版



# WILEY 最终用户许可协议

访问 [www.wiley.com/go/eula](http://www.wiley.com/go/eula), 浏览 Wiley 电子书最终用户许可协议。