

戴尔 PowerMax：数据缩减

线内压缩和重复数据消除

2022 年 7 月

H19254

白皮书

摘要

PowerMax 存储平台采用多种数据缩减技术，如线内压缩和重复数据消除。此外还包括模式检测和高效数据放置，可在性能和效率之间实现出色的平衡。

Dell Technologies

版权

本出版物的内容按“原样”提供。Dell Inc. 对本出版物中的信息不作任何形式的陈述或担保，并明确拒绝对适销性或针对特定用途的适用性进行任何暗示担保。

需具备适用的软件许可证才能使用、复制和分发本出版物中说明的任何软件。

版权所有 © 2022 Dell Inc. 或其子公司。保留所有权利。Dell Technologies、Dell、EMC、Dell EMC 和其他商标为 Dell Inc. 或其子公司的商标。Intel、Intel 徽标、Intel Inside 徽标和 Xeon 均为 Intel Corporation 在美国和/或其他国家/地区的商标或注册商标。其他商标可能是其各自所有者的商标。

中国印刷，2022 年 7 月 H19254。

Dell Inc. 确认本文档在发布之日内容的准确性。该信息如有更改，恕不另行通知。

目录

执行摘要.....	4
数据缩减.....	5
系统资源使用情况.....	9
管理和监视	9
支持的数据服务	17
总结.....	18
参考资料.....	19

执行摘要

概览

戴尔 PowerMax 将线内压缩、线内重复数据消除和模式检测相结合，实现数据缩减，从而提升系统效率。使用这些数据缩减技术，用户可以实现巨大的容量节省。数据缩减可压缩数据并消除冗余数据拷贝。本白皮书介绍了数据缩减在 PowerMax 系统中的作用，并描述了如何使用戴尔管理应用程序（如 Unisphere for PowerMax、Solutions Enabler 和 Mainframe Enablers 软件）进行报告。

修订记录

日期	描述
2022 年 7 月	初版

我们非常重视您的反馈

Dell Technologies 和本文档的作者欢迎您对本文档提供反馈。通过[电子邮件](#)联系 Dell Technologies 团队。

作者： Robert Tasker

提醒： 有关此主题的其他文档的链接，请参阅 [PowerMax 和 VMAX 信息中心](#)。

数据缩减

概览

数据缩减结合了线内压缩、线内重复数据消除、模式检测、高效数据放置和机器学习 (ML)。此组合会创建一个系统，用户可以写入比总物理容量更多的主机数据，同时实现企业级存储系统预期的性能。默认情况下，此功能处于开启状态，您可以在存储组级别启用或禁用它。此外，PowerMax 2500 和 8500 系统中提供的所有数据服务均受支持。该支持也适用于 CKD 模拟，但不包括 CKD 的重复数据消除。

压缩可减小数据大小，而重复数据消除 (dedupe) 则将数据存储为单个实例。模式检测包括一个非零分配功能，该功能将连续的零字符串从压缩数据中排除。使用集成在系统中的硬件辅助来执行压缩、重复数据消除和模式检测，以降低执行这些功能的开销。机器学习识别存储在磁盘上的繁忙数据，并确保其保持未缩减状态以获得出色性能。高效的数据放置使用一种称为**压缩**的功能，该功能可战略性地存储数据，以最大限度地减少浪费的空间，并减少对垃圾数据收集或碎片整理 (defrag) 功能的需求。

基于活跃度缩减

基于活跃度缩减 (ABR) 可降低通过解压缩频繁访问的数据而产生的性能成本。该功能允许将多达 20% 的繁忙数据存储在未压缩的系统上。这种能力对系统有利，因为它可最大限度地减少因不断解压缩频繁访问的数据而导致的性能延迟。要确定哪些数据比较繁忙，系统使用机器学习算法处理 IO 统计信息。执行此任务可维护平衡、优化的环境，以实现数据缩减节省和性能提升。

压缩

压缩可将传入写入工作负载减少到尽可能小的大小，以消耗极少的容量。数据在通过使用 GZIP 压缩算法的数据缩减硬件时被压缩。通过数据缩减硬件时，数据划分为四个部分，这些部分并行压缩，以最大限度地提高效率。这四个部分的总和是存储在磁盘上的数据的最终缩减大小。此功能可提供对缩减数据的精细访问。仅处理包含部分读取或写入请求的已请求数据的部分，因为每个部分均可单独处理。

重复数据消除

重复数据消除是一种容量节省方法，可识别完全相同的数据拷贝并存储每个拷贝的单个实例。为了提供有效的容量节省，需要几个重复数据消除组件。

- **哈希 ID**：哈希 ID 是传入数据的唯一标识符，用于确定是否需要重复数据消除关系。系统使用 SHA-256 算法生成哈希 ID。
- **哈希 ID 表**：哈希表是在系统控制器之间分布系统内存的一种分配方案。这些表为重复数据消除过程使用的哈希 ID 编制目录。表中的条目用于确定是否存在重复数据消除关系，或者是否需要新条目，以及数据是否可以存储在磁盘上。
- **重复数据消除管理对象 (DMO)**：DMO 是系统内存中的 64 字节对象，仅在存在重复数据消除关系时才存在。这些对象存储和管理前端设备与在阵列中使用后端容量的

去重数据之间的指针。DMO 管理前端设备和磁盘上存储的数据之间的去重数据的指针。当存在重复数据消除关系时，这还可以管理将哈希 ID 存储在哪个哈希表中。

使用与压缩相同的数据缩减硬件执行重复数据消除，并且在硬件处理数据时生成唯一的哈希 ID。然后，将哈希 ID 与哈希 ID 表进行比较，以查找相同的 ID。找到匹配项时，数据不会存储在磁盘上，并且会创建重复数据消除共享。在前端卷和哈希 ID 表中的唯一 ID 之间设置指针。指针将磁盘上存储的数据的单个实例链接到卷，以便将来访问数据。DMO 管理数据、访问数据的前端卷和哈希 ID 表之间的指针。当哈希 ID 表中没有匹配项时，将添加一个新条目，以便将来进行哈希 ID 比较。

重复数据消除 算法

PowerMax 系统使用在数据缩减硬件中实施的 SHA-256 哈希算法来查找重复数据。然后，将数据存储为单个实例，供多个源共享。此过程可提高数据效率，同时保持数据完整性。

SHA-256 算法为每个 32 KB 数据块生成 32 字节的代码。考虑一个使用 1 PB 写入数据的系统，每天更新 5% 的数据。在 100 万年的运行中，发生哈希冲突的可能性为 20%。由于每个 128 KB 磁道作为 4 个 32 KB 块来处理，因此需要在同一个 128 KB 磁道的所有四个数据块上发生哈希冲突，才算是发生实际的哈希冲突。全部 4 个块都发生冲突的几率非常小，只在理论上才有可能发生（在万亿年的运行中，发生这种情况的几率不到 1%）。此外，当在重复数据消除比较阶段找到匹配项时，将执行逐字节比较。进行此比较是为了在更新表和设置指针以允许访问数据之前确认存在匹配项。

压缩

使用称为压缩的过程执行数据放置。压缩会直观地将缩减或未缩减数据放置在磁盘上尽可能合适的位置。在磁盘上存储数据的操作使用写入对象。在系统中配置的驱动器上，每个对象都是 6 MB 的连续后端数据设备容量。写入对象在 1 K 边界上对齐，并在一次使用中按顺序消耗。对于所有受支持的 RAID 类型，写入对象都分布在整个条带上，以优化写入。对于 FBA 和 CKD 模拟，每个对象都支持缩减或未缩减数据。

- **FBA 写入对象：**未缩减的写入对象包含 48 个 FBA 磁道。缩减的写入对象包含 1000 个缩减的磁道。写入对象的缩减条目范围为 1 KB 到 96 KB。
- **CKD 写入对象：**未缩减的写入对象包含 108 个 CKD 磁道。缩减的写入对象包含 1000 个缩减的磁道。写入对象的缩减条目范围为 1 KB 到 52 KB。

扩展数据压缩

PowerMax 2500 和 8500 系统包含一项称为扩展数据压缩 (EDC) 的额外功能，可压缩已压缩的数据以进一步节省容量。通过识别在延长时间内未被访问过的数据来完成此任务。使数据成为 EDC 候选的因素如下：

- 数据属于已启用数据缩减的存储组。
- 数据已有 30 天未被访问。
- 数据尚未通过 EDC 压缩。

符合 EDC 条件的数据使用 Def9_128_SW 算法进行压缩，以进一步减少用于存储数据的容量。此进程是系统中的自动化后台进程。在存储组级别实现的压缩比包括额外的节省。EDC 仅适用于 PowerMax 存储阵列。

CKD 压缩

基于活跃度缩减 (ABR) 可降低通过解压缩频繁访问的数据而产生的性能成本。该功能允许将多达 20% 的繁忙数据存储在未压缩的系统上。这一结果对系统有利，因为它消除了因不断解压缩频繁访问的数据而对性能产生的负面影响。要确定数据的繁忙级别，系统使用机器学习算法处理从传入 I/O 到前端设备收集的统计信息。此操作能够维持系统资源之间的平衡，从而提供一个出色的环境，实现数据缩减节省和性能提升。

压缩可将传入写入工作负载减少到尽可能小的大小，以尽可能减少消耗的容量。数据在通过内置于使用 GZIP 压缩算法的系统的存储硬件时被压缩。通过数据缩减硬件时，数据划分为四个部分，这些部分并行压缩，以最大限度地提高硬件的效率。这四个部分的总和是存储在磁盘上的数据的最终缩减大小。当存在部分读取或写入请求时，此结果可为缩减的数据提供精细访问权限。仅处理包含已请求数据的部分，因为每个部分均可单独处理。

使用称为压缩的数据放置过程执行数据放置。压缩会直观地将缩减或未缩减数据放置在磁盘上尽可能合适的位置。在磁盘上存储数据的操作使用写入对象。在系统中配置的驱动器上，每个对象都是 6 MB 的连续后端数据设备容量。写入对象在 1 K 边界上对齐，并在一次使用中按顺序消耗。对于所有受支持的 RAID 类型，写入对象都分布在整个条带上，以优化写入。每个对象都支持缩减或未缩减数据。未缩减的写入对象包含 108 个 CKD 磁道。缩减的写入对象包含 1000 个缩减的磁道。写入对象的缩减条目范围为 1 KB 到 52 KB。

数据缩减 I/O 流

所有 I/O 都通过高速缓存，然后由系统处理。在系统接收数据后且将其放入磁盘之前，执行数据缩减操作。使用线内流程需要在应用数据缩减的 I/O 流中进行额外检查。系统使用这些检查来确定传入数据是否需要通过数据缩减硬件。已启用数据缩减的存储组的传入数据将遵循数据缩减流。然而，由于具备基于活跃度缩减 (ABR) 功能，已启用数据缩减的

存储组的活动数据会跳过数据缩减流以实现性能优化。由于 ABR 而未压缩的数据可能会在以后压缩并移动到压缩池。已禁用数据缩减的存储组的数据会忽略数据缩减流，并在未缩减的情况下写入系统。

需要考虑几个不同的 I/O 类型：读取、写入和写入更新。

- **读取**：访问已填入阵列的数据的请求。
- **写入**：会占用磁盘空间的传入 I/O。
- **写入更新**：会更改分配给阵列磁盘空间的数据的传入 I/O。

下图描绘了 I/O 将遵循的路径，它由数据集或相关存储组的特征决定。

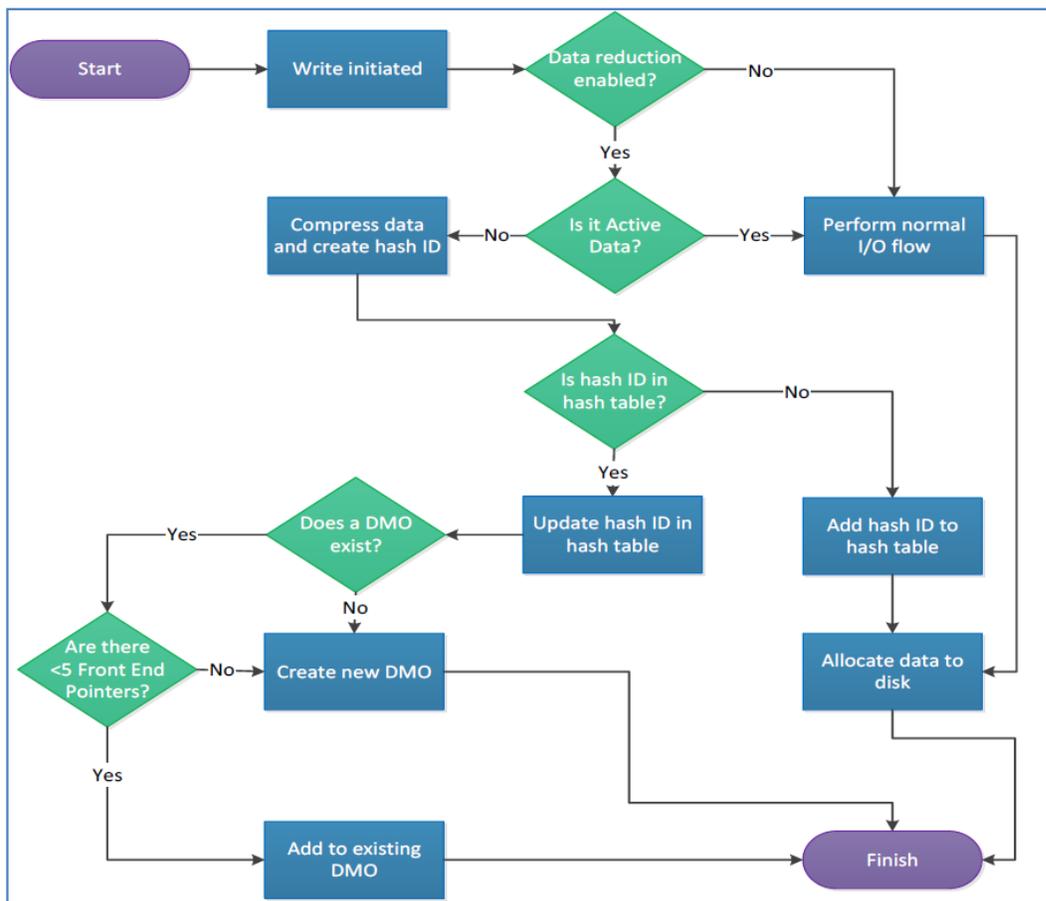


图 1. PowerMax 企业级存储系统的数据缩减 I/O 流

系统资源使用情况

概览

容量和内存是每个 PowerMax 系统中配置的两个主要资源。容量分为物理容量、有效容量和调配容量。内存被归类为系统资源。

内存资源支持用于调配容量和物理容量的元数据结构。可用的有效容量与物理容量、可用系统资源量以及写入系统的数据的可缩减性相关。写入的数据如果高度可缩减，则消耗的物理容量更少，从而会增加有效容量。反之亦然：写入的数据如果不可缩减，则会减少可用的有效容量。[数据缩减](#)部分中描述的信息（容量、系统资源）在用于 PowerMax 2500 和 8500 系统、Unisphere for PowerMax、Solutions Enabler 和 Mainframe Enablers 软件的管理应用程序中可用。Unisphere for PowerMax 是一个用户界面 (UI)，以图形、图表和列表形式提供数据。Solutions Enabler 是一个标准的命令行界面，提供相同的数据，但不是以图表和图形的形式。Mainframe Enablers 是一套组件，用于监视和管理大型机环境中的 Dell Storage 系统。本白皮书接下来的部分所示的图像描述了 Unisphere for PowerMax 管理 PowerMax 2500 或 8500 系统的过程。

物理容量

物理容量是基于安装的磁盘和应用的 RAID 保护在系统中配置的磁盘空间量。在未使用数据缩减的配置中，物理容量是可用于主机数据的总容量。例如，显示 100 TB 物理容量的系统表示它可以容纳 100 TB 不使用数据缩减的主机数据。

有效容量

有效容量是使用数据缩减时可用的空间量。初始安装时的总容量取决于系统中配置的内存量，并且基于 4:1 的默认数据缩减节省 (CKD 模拟为 3:1)。

例如，具有 100 TB 物理容量的同一系统将显示 400 TB 的有效容量。值 400 TB 是有效容量的起始点，随着数据写入系统并应用数据缩减，该值将会发生变化。

调配容量

调配容量是以创建的设备形式表示的可用容量，并呈现给打算消耗系统中物理容量或有效容量的主机和应用程序。

管理和监视

概览

Unisphere for PowerMax 是一个用户界面，用于管理和监视系统的容量和资源使用情况。在系统级别，有关容量使用情况、数据缩减和系统资源的信息将显示在容量控制面板中。在容量控制面板中，用户可以转到显示有效容量和调配容量、快照容量、数据缩减以及系统资源信息的屏幕。

容量控制面板

在 Unisphere for PowerMax 中，有多个显示屏提供与容量使用情况相关的信息。

主控制面板显示一个交互式图形，其中显示了有效容量使用情况和随时间推移的数据缩减情况。此显示屏显示有效容量使用情况历史记录，以及与有效容量相关的数据缩减比率。此信息可用于监视和跟踪相对于所示数据缩减比率的有效容量使用趋势。PowerMax 2500 或 8500 系统可在同一存储资源池中配置 FBA 和 CKD 模拟，但历史图表特定于所选模拟视图。

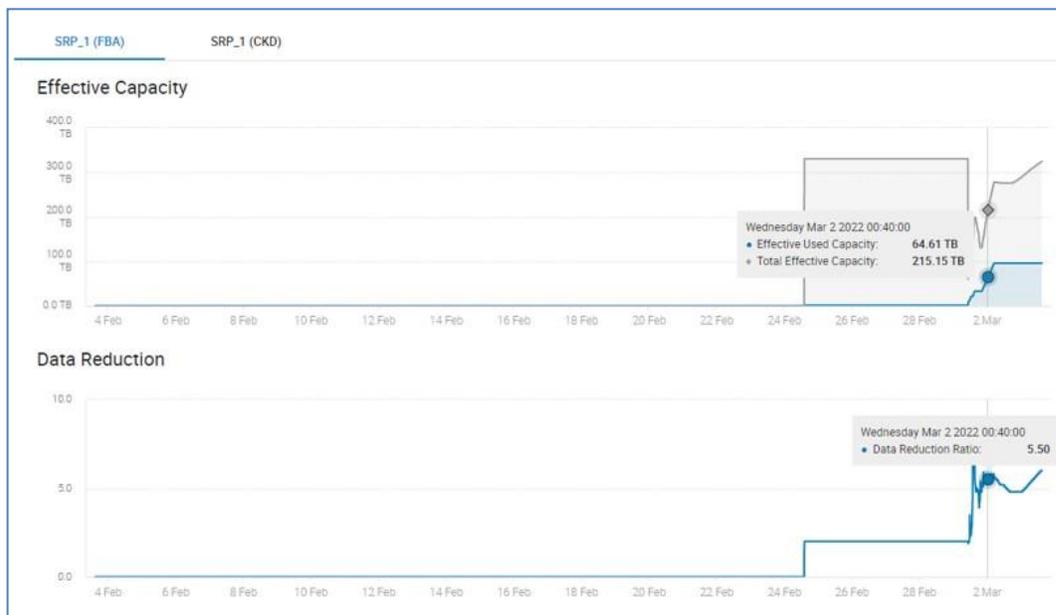


图 2. 显示 FBA 模拟的有效容量和数据缩减的容量控制面板历史图表

主控制面板还以条形图形式提供调配容量、有效容量、快照使用情况和数据缩减的数据。每个部分均可展开为更详细的显示屏，显示每个项目的更精细数据。

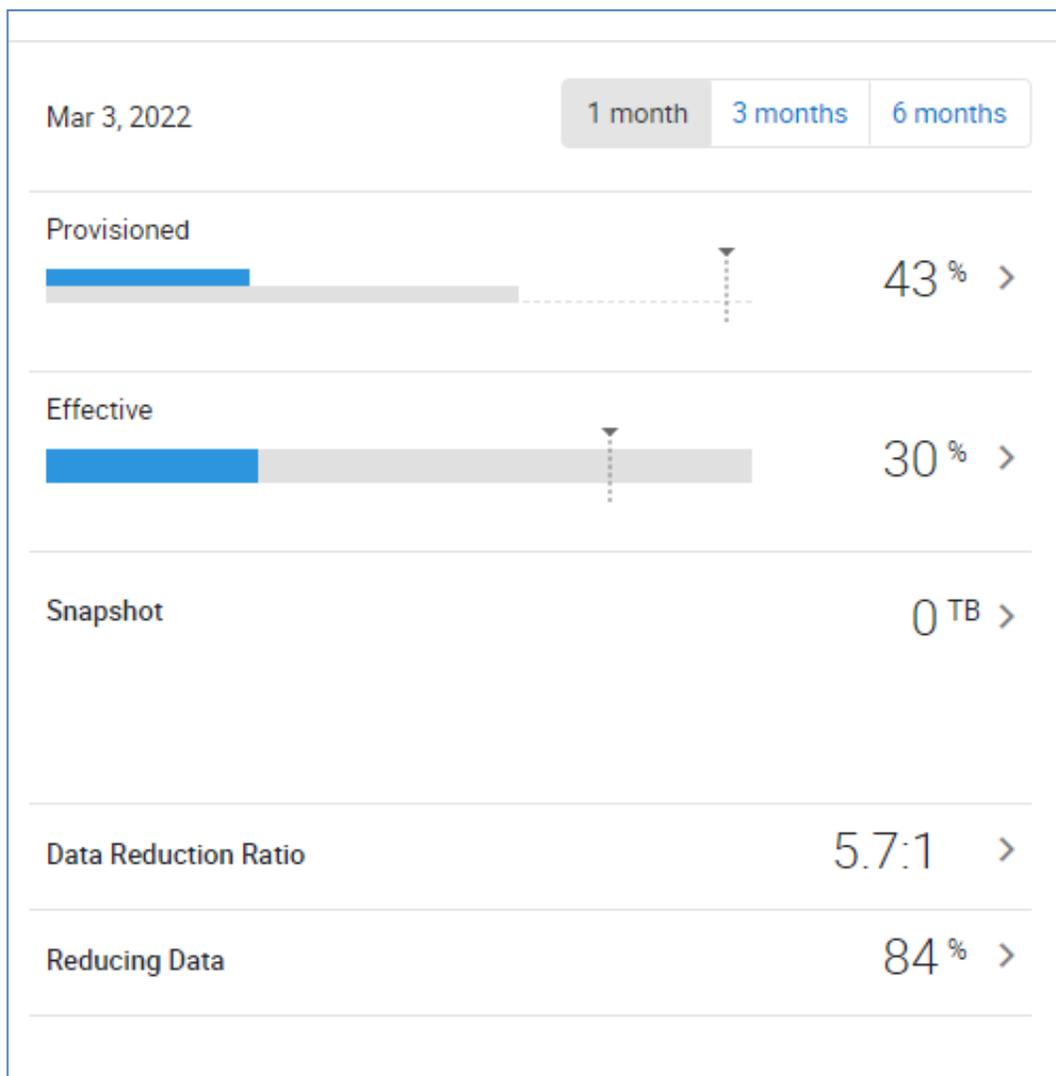


图 3. 调配容量、有效容量、快照使用情况和数据缩减的容量控制面板条形图

调配

调配容量是以设备形式调配的容量，该容量作为可用容量呈现给主机和应用程序。使用“SRP 容量”和“系统资源”这两个指标跟踪调配容量。

- **SRP 容量**显示以调配容量值和可用有效容量值（以 TB 为单位）的调配容量。有效 SRP 容量最初基于 4:1 的默认数据缩减节省（CKD 模拟为 3:1）。该容量是使用系统中配置的物理容量计算得出。调配容量会随着用户创建设备而增加。显示的百分比值是订阅百分比，根据调配容量和有效容量进行计算。

- **系统资源**代表以元数据形式支持调配容量的可用内存量。仅当向系统添加高速缓存时，显示的总值才会更改。已用容量就代表这一点。可用容量表示系统可以支持的额外调配容量。当用户创建设备时，已用容量将会增加。当两个值之间存在差异时，将使用内存资源来支持数据缩减或其他使用内存的功能，例如拍摄现有设备的快照。

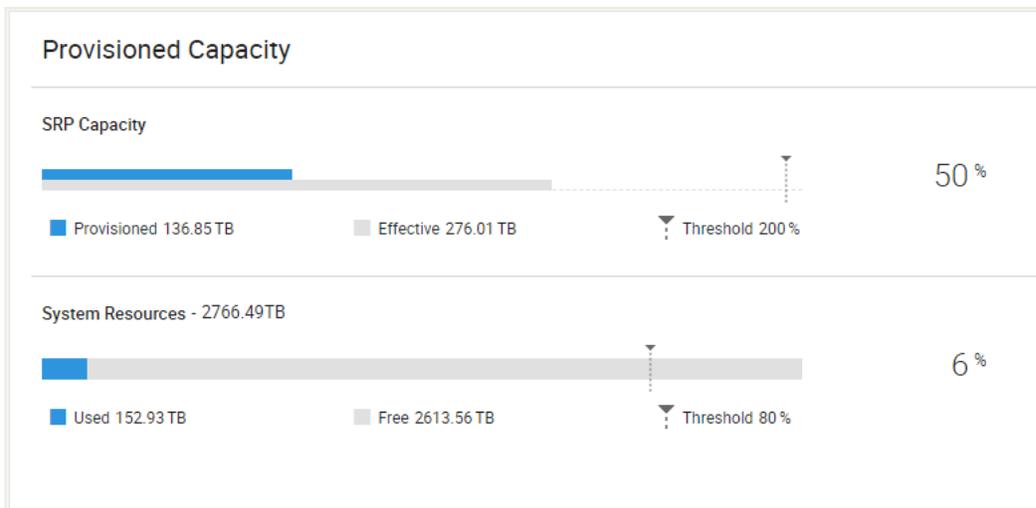


图 4. 调配容量显示屏中的系统资源使用情况

有效

有效容量表示用户可使用的容量，基于对使用“数据缩减”所带来的节省的预期。有效容量显示屏提供了可用物理容量和有效资源的详细视图。这显示在三个部分：“物理容量”、“有效容量资源”和“有效容量使用情况”。

- **物理容量**显示系统中配置的硬盘中可用的物理容量。所示的容量是应用格式化和 RAID 保护后的值。所示的值是系统在未使用数据缩减时可支持的主机数据容量。
- **有效容量资源**根据当前系统资源使用情况指示可实现的值。所示的有效容量资源值将相对于当前数据缩减节省以及物理和有效容量使用情况进行调整。
- **有效容量使用情况**显示根据系统资源使用情况和当前数据缩减节省可用的当前有效容量。圆形图中显示的值是当前可用的有效容量。呈现在右侧的值将使用情况细分为三个类别：快照已用容量、用户已用容量和可用容量。

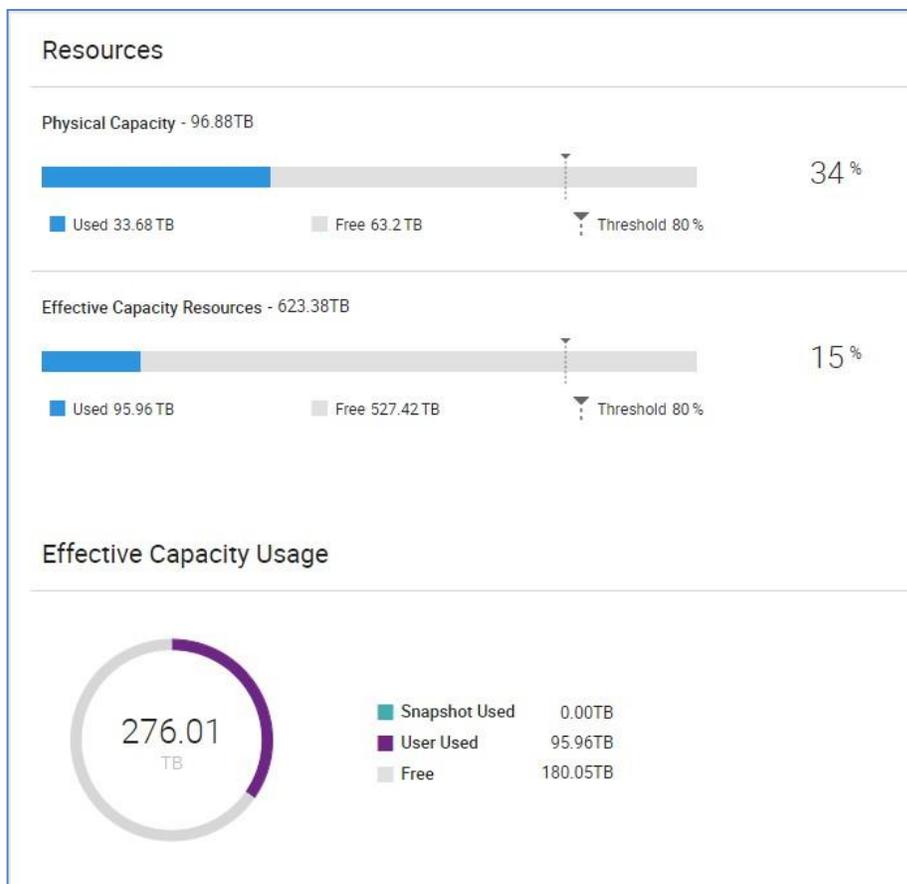


图 5. 显示有效容量显示屏的资源使用情况

快照

由于受到共享分配和数据缩减等功能的效率影响，后端快照容量可能显著低于每个快照的增量。

将鼠标悬停在容量控制面板上的快照条形图上，以获得概要详细信息。快照值定义为：

- **已用**：用于快照更改数据的有效容量
- **可用**：基于已用和剩余元数据量的剩余快照更改数据量
- **总计**：已用 + 可用
- **阈值**：针对快照更改数据发出警报的阈值

单击条形图以转至有效容量控制面板和快照容量控制面板。

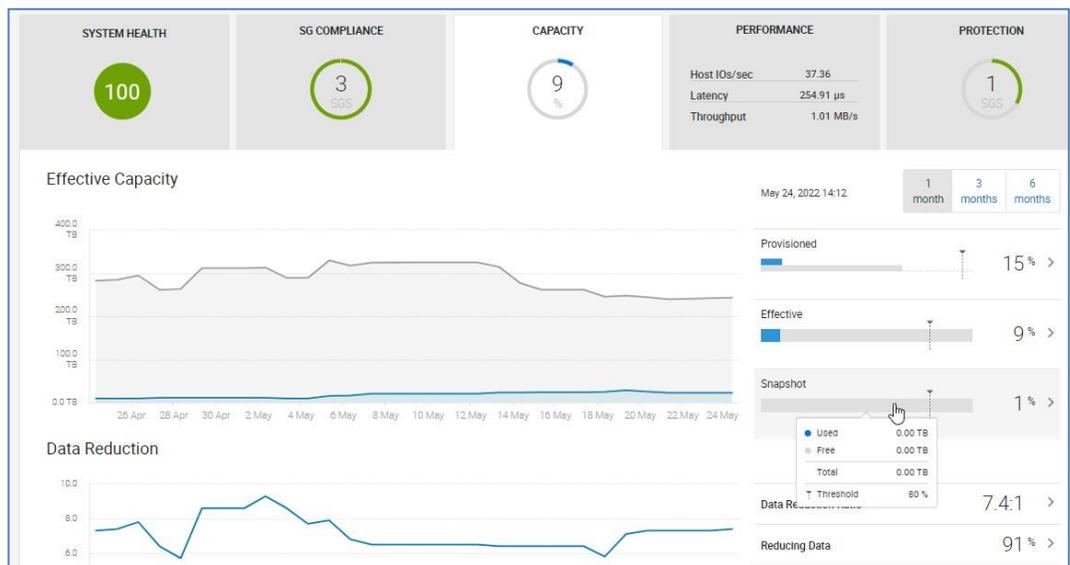


图 6. 容量控制面板

在有效容量控制面板上：

- **快照已用**：快照更改数据所使用的有效容量

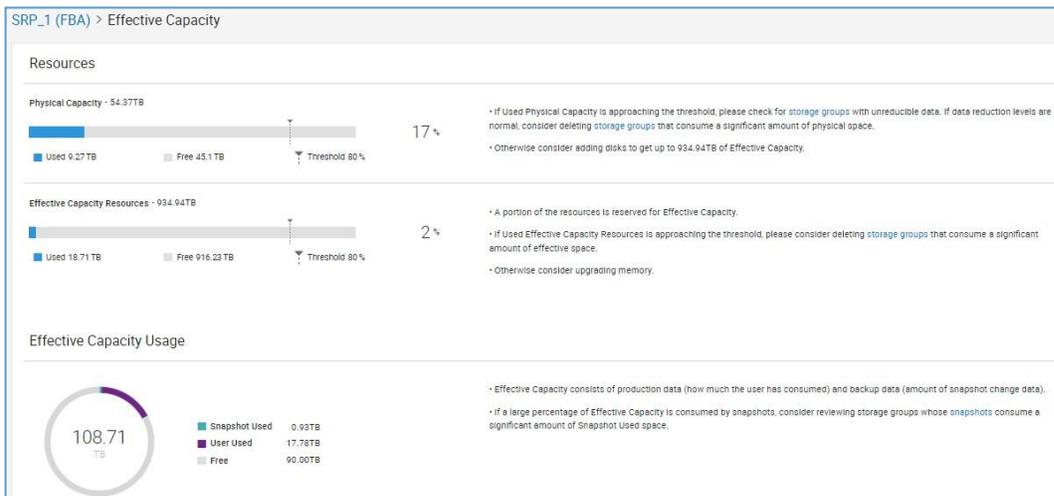


图 7. 有效容量控制面板

“快照容量” 页面中的值定义为：

- **快照有效已用**：SRP 的快照数据使用的有效容量百分比
- **快照物理已用**：SRP 的快照数据消耗的可用容量百分比
- **快照资源**：为整个系统消耗的快照元数据百分比
- **快照资源已用**：快照使用的作为快照元数据容量一部分的容量

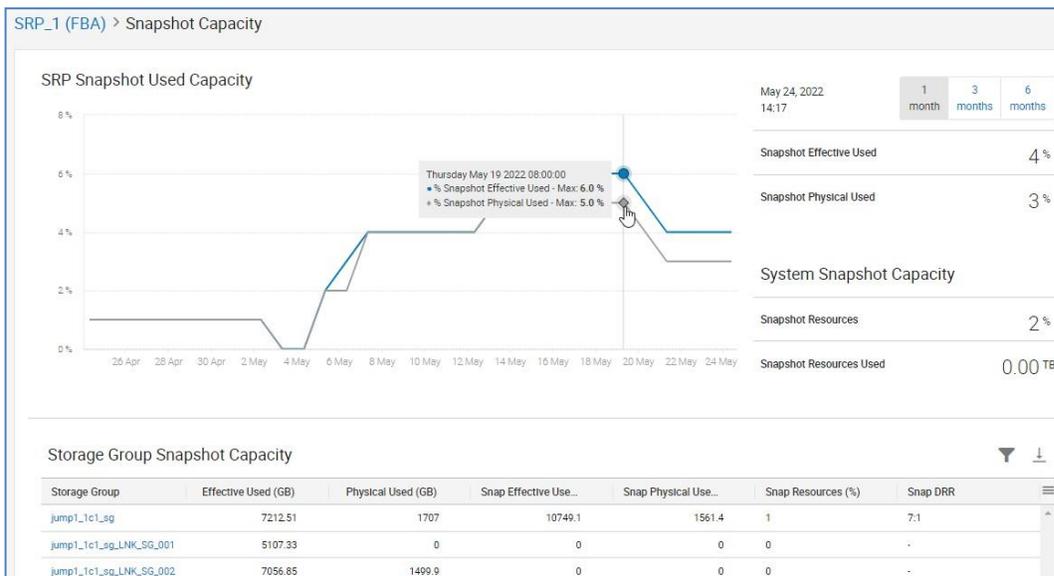


图 8. 快照容量使用情况

数据缩减比率

数据缩减显示屏为用户提供了单个位置来查看数据缩减效率。有三个部分：“数据缩减比率”、“历史交互式图形”和所有存储组的表。显示的数据缩减比率仅考虑已启用的数据和写入系统的缩减数据。

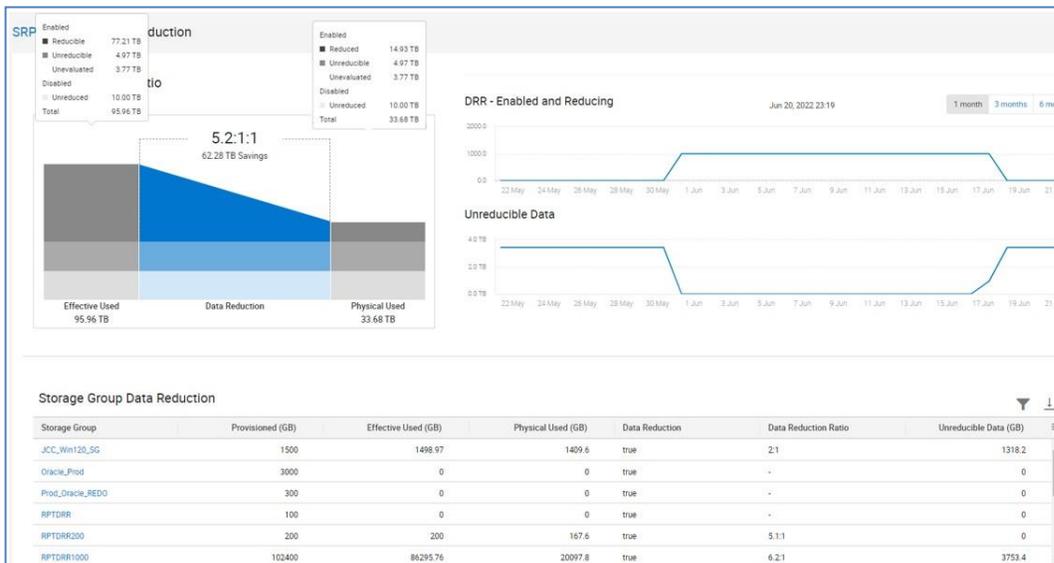


图 9. 数据缩减比率

数据缩减比率显示为一个图形，其中显示有效已用容量、数据缩减比率和实际已用容量。

“物理已用”是指正在使用的实际物理容量。数据缩减以比率表示节省。

有效已用表示在应用数据缩减时实现任何节省之前写入系统的数据。显示的所有值均表示由主机或应用程序写入的完整大小。写入的数据分为两类：“已启用”和“已禁用”。

- **已启用**表示要考虑的数据已启用数据缩减，并受数据缩减过程和基于活跃度缩减功能的约束。启用数据缩减时，数据可归入三个额外类别：“可缩减”、“不可缩减”和“未评估”。
 - **可缩减**数据是数据缩减过程已确定为数据（可缩减以使用比写入系统更少的物理容量）的写入数据量。
 - **不可缩减**数据是无法缩减的数据。
 - **未评估**数据是尚未由数据缩减过程评估的数据。尚未确定数据是可缩减，还是不可缩减。
- **已禁用**表示写入系统的数据不受任何数据缩减节省的约束。标识为已禁用的所有数据将显示为未缩减。

物理已用表示存储在磁盘上之后写入系统的数据。这会考虑所有已启用和已禁用的数据，以及所有缩减和未缩减的数据。有两个类别表示存储在磁盘上的数据：“已启用”和“已禁用”。

- **已启用**表示已经历数据缩减过程的数据。此数据有三个子类别：“已缩减”、“不可缩减”和“未评估”。
 - **已缩减**数据已经通过数据缩减过程发送。该过程既包括通过数据缩减硬件，也包括存储在磁盘上。与主机或应用程序写入的数据相比，存储在磁盘上的已缩减数据占用的磁盘空间更少。
 - **不可缩减**表示数据已通过数据缩减过程（包括数据缩减硬件）发送，但无法缩减。物理已用部分中考虑的一些不可缩减数据可能会导致由于重复数据消除而共享的数据，从而实现数据缩减节省。
 - **未评估**数据是数据缩减过程尚未评估的数据。因此，尚未确定数据是可缩减还是不可缩减。
- **已禁用**表示写入系统的数据不受任何数据缩减节省的约束。标识为已禁用的所有数据将显示为未缩减。

交互式图形图表 **DRR 已启用和缩减和不可缩减数据**提供历史数据。这显示了不可缩减数据对数据缩减比率的影响。此图允许用户跟踪和监视数据缩减比率中可能由不可缩减数据引起的更改。

存储组列表提供特定于系统中每个存储组的容量使用情况和数据缩减信息。使用交互式图形跟踪对数据缩减比率的更改时，可以使用存储组列表来识别具有大量影响数据缩减比率的不可缩减数据的存储组。

计算效率比率：可在数据缩减图的弹出窗口中找到计算数据缩减比率所需的数据。

- **数据缩减比率**：使用“有效使用”中的“已启用”和“可缩减”以及“物理已用”中的“已启用”和“已缩减”来计算数据缩减比率。

$$\text{已启用可缩减} \div \text{已启用已缩减}$$

- **总体数据缩减比率**：使用“有效已用”和“物理已用”的总值计算总体系统数据缩减比率

$$\text{有效已用总量} \div \text{物理已用总量}$$

支持的数据服务

概览

FBA 存储支持数据缩减。在同一存储资源池中支持混合 FBA/CKD 系统，但 CKD 模拟的数据缩减将仅利用压缩和基于活跃度缩减。支持 PowerMax 和 VMAX 全闪存系统中提供的所有其他数据服务。这些服务包括本地复制 (SnapVX)、远程复制 (SRDF)、D@RE 和 VMware vSphere Virtual Volumes (vVol)。

本地复制 (SnapVX)

SnapVX 快照可保护应用程序，而无需使用目标卷来捕获称为增量的更改数据。使用指向相关时间点映像的指针，在存储后端自动维护快照增量。资源共享和重复数据消除会自动利用此设计来提供高速缓存、容量和性能优势。

总结

成为快照增量时，已压缩的源数据仍保持为已压缩。未压缩的源数据在变为快照增量或变为不太活跃时可能会被压缩。通过链接目标的读取活动可能会阻止压缩未压缩增量，或者可能导致将压缩的快照增量解压缩。快照增量可用于重复数据消除。

在链接目标上启用数据缩减只会影响链接目标拥有的数据。链接目标和克隆上的数据可用于重复数据消除。

远程复制 (SRDF)

支持 SRDF 的压缩，称为 SRDF 压缩。SRDF 压缩是一项旨在降低带宽消耗的功能，同时使用远程复制将数据发送到已连接的系统以及从系统中发出数据。SRDF 压缩和数据缩减都使用相同的硬件；但是，它们有不同的用途。使用数据缩减压缩的数据在 SRDF 链路中发送之前解压缩。如果 SRDF 压缩和线内压缩适用，则解压缩数据，使用 SRDF 压缩功能进行压缩，然后发送到远程站点。

静态数据加密 (D@RE)

D@RE 提供基于硬件的阵列上后端加密，数据缩减作为线内过程执行。数据通过数据缩减硬件，然后通过加密硬件发送。因此，在加密数据之前，先对数据进行压缩和/或重复数据消除。在已启用 D@RE 的系统上，在磁盘上加密的数据已经压缩、去重或同时执行两者。

虚拟卷

将数据分配到 vVol 时支持数据缩减，并遵循与所有其他数据相同的 I/O 路径。IO 路径可在图 1 中看到。由于 vVol 没有存储组，因此在 vVol 存储容器中的存储资源级别启用数据缩减。

总结

总结

物理存储容量的使用是存储行业的存储管理员普遍关注的问题。不断增长的数据量催生了提高物理容量使用效率的需求。戴尔 PowerMax 2500 和 8500 数据存储系统将此效率提升到更高水平。数据缩减可提供卓越的容量节省，同时提供出色性能。这一结果减少了数据中心占地面积，降低了总体 TCO。除了节省之外，使用数据缩减非常简单，只需单击即可启用或禁用。系统会处理所有工作。

参考资料

Dell Technologies 文档

以下 Dell Technologies 文档提供了与本文档相关的其他信息。对这些文档的访问取决于您的登录凭据。如果您无法访问某个文档，请与 Dell Technologies 代表联系。

- [PowerMax 和 VMAX 信息中心](#)