

HS4000 系列存储产品技术白皮书

北京智云创新科技股份有限公司

2024.2

目录

1	产品概述.....	3
2	系统架构.....	3
2.1	专业的存储管理系统.....	3
2.2	硬件架构.....	4
3	产品功能.....	4
3.1	Omni Cache——读性能提升技术.....	4
3.2	OmniTiering——写/读性能提高 SSD/HDD 技术.....	6
3.3	OmniVOL——灵活卷管理技术.....	7
3.4	RAID PX——多重校验的高可靠 Raid 技术.....	12
3.5	OmniMirror——异地数据容灾技术.....	16
3.6	OmniDedup——重复数据删除技术.....	19
3.7	维护管理.....	23
4	HS4000 系列产品技术规格.....	25

1 产品概述

信创 NAS 存储产品 HS4000 隶属于“智云町”品牌，是北京智云创新科技股份有限公司推出的**高性能企业级信创机架 NAS/IP SAN 存储系统**，凭借**国产海光 5380 16 核 CPU** 架构，以及 SSD 和 HDD 盘混合分层构建高性能的 SSD 混合存储系统，将存储性能、成本和容量构建最佳匹配的**高性能国产信创企业级存储**。

产品定位

HS4000 系列产品满足特定行业高性能环境下：测绘三维建模、高校小规模 HPC、影视渲染、设计院 3D 设计渲染、大数据 AI 规模数据等行业中，解决海量数据存储并发共享。

2 系统架构

2.1 专业的存储管理系统

HS4000 系列存储产品配备智云创新的 COS V3 存储系统，配合**国产海光 5380 16 核 CPU** 的硬件可以让您在极短时间内搭建**高稳定性、高性价比**的网络存储系统，如 NAS 服务器、iSCSI 存储设备等。

软件特色功能：

- 灵活的多协议存储架构，同时支持 iSCSI /CIFS/NFS/HTTP/FTP 多协议访问；
- 闪存与大容量磁盘的完美融合，帮助用户在性能和成本之间实现最优均衡；
- 基于 NAS 的 SSD 和机械磁盘的数据分层；
- 基于 AI 的在线分析数据调优功能；
- 灵活卷、重复数据删除、数据压缩等功能提高空间利用率；

- 通过数据快照、卷克隆、卷镜像、RAID 多重校验等数据保护机制，有效保证用户数据安全；

2.2 硬件架构

HS4000 系列产品采用基于**国产海光 5380 16 核 CPU**的硬件架构，实现了系统的性能和高扩展性，可以满足**信创应用环境**和不断增加的容量需求。

- **主控**

单控制器，采用 1 颗**国产海光 5380 16 核**处理器。最大支持 128GB 内存；6 个 PCIE 插槽，1+1 冗余电源、风扇；

- **主控柜规格**

HS4000 系列支持 2U12 盘、4U24 盘、4U36 盘三种槽位的数据盘规格。

- **最大扩展能力**

最大扩展到 320 块磁盘，支持 24、60、84 磁盘位扩展柜，每个磁盘扩展柜的控制器、风扇、电源及硬盘采用全模块化设计。无任何延时，不会出现掉盘等可靠性问题，一旦控制器任何部件出现故障，只要抽出控制器就可以维护了。

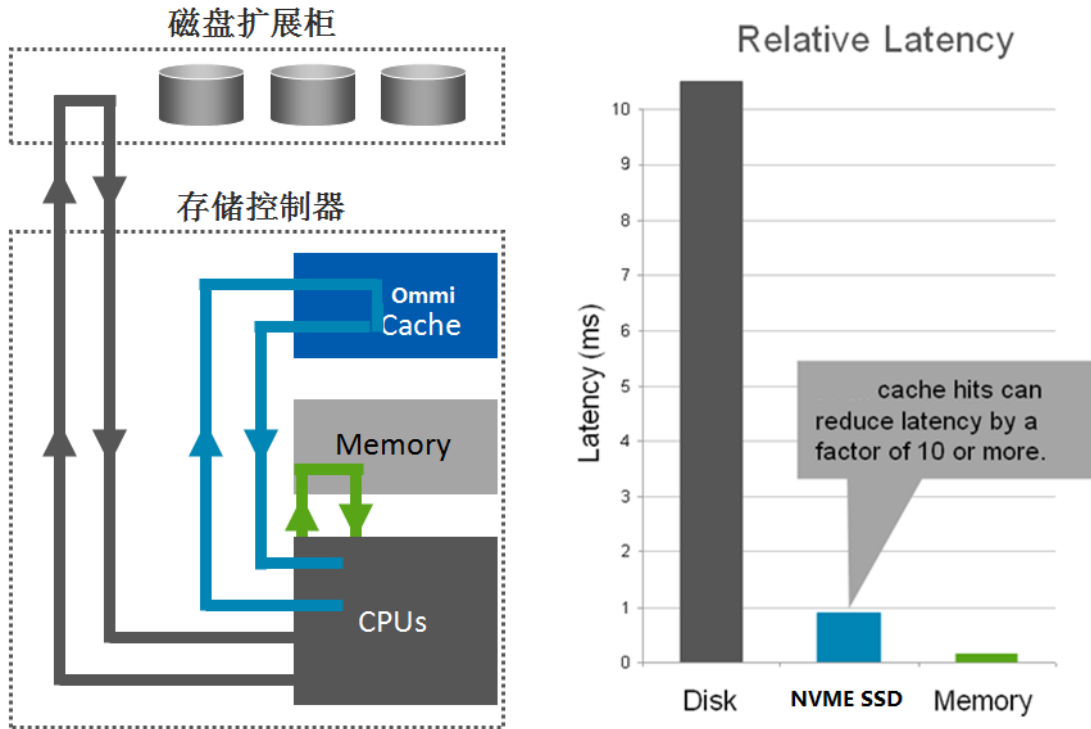
3 产品功能

3.1 Omni Cache——读性能提升技术

智云的 Omni Cache 性能加速模块通过存储控制器 PCIE 内置，Cache 容量根据存储控制器不同型号支持 512GB~16TB，该缓存模块与存储控制器之间采用 NVME PCIE 进行高速通信，其性能远远高于 SSD 硬盘通过接口与存储控制器进行通信。Omni Cache 技术可以

自动分析热点数据，将经常被访问的数据自动加载到 Cache 中，之后用户再访问时存储控制器直接从 Omni Cache 中将数据返回给用户，而不需要再从硬盘中读取这些热点数据，大大的提高数据读取性能。

Omni Cache（下图中性能提高即 Omni Cache 扩展缓存）数据处理流程如下图所示：



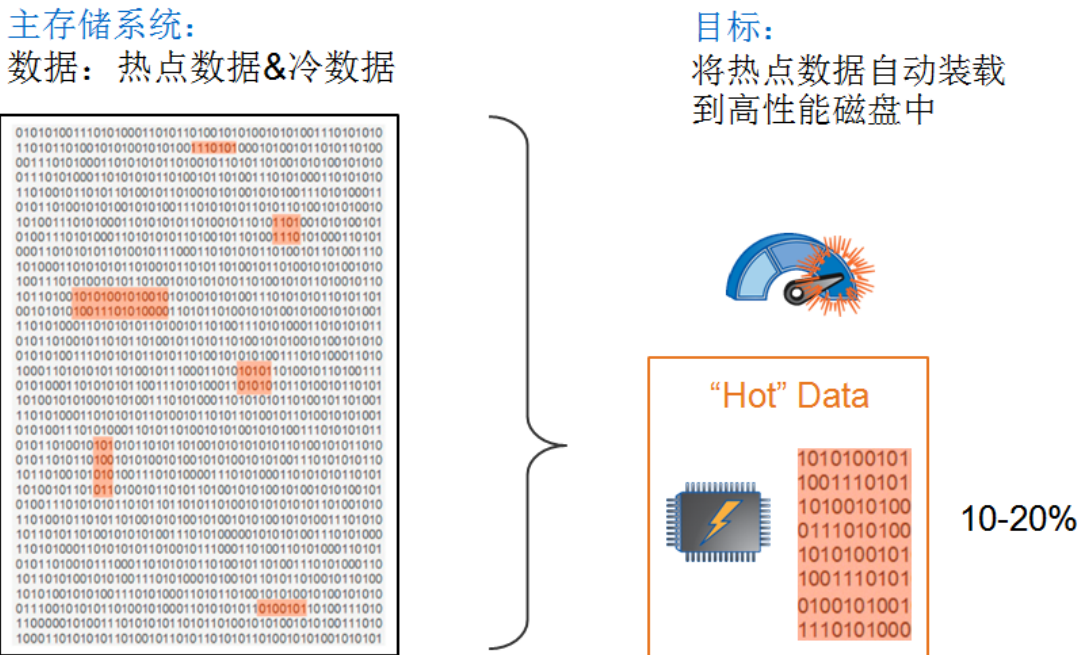
图：Omni Cache 架构图：Omni Cache 响应延时对比

智云的 OmniCache 大缓存技术实现了存储系统的 TB 级缓存，并且 OmniCache 还支持缓存级的重复数据删除，因此利用 OmniCache 技术，能极大提升 VMware 主机/桌面虚拟化环境的数据随机读取的缓存命中率，减少后端磁盘的数据读取压力，这不仅明显提升数据读取的性能，还能减少对后端磁盘数量的需求。特别对于桌面虚拟化环境下大量用户同时登录虚拟桌面系统所造成的启动和登录风暴问题，利用 OmniCache 就能很好的解决这个问题。

3.2 OmniTiering——写/读性能提高 SSD/HDD 技术

智云的 OmniTiering 技术,采用 SSD Pool 闪存池是特定的磁盘空间资源池,能够高速缓存写入和读取;大量并发写入数据首先由控制器写入到高性能 SSD 固态硬盘中,然后自动复制到 SATA、SAS 磁盘中,当数据读取时数据被自动复制到 SSD 固态硬盘中,存储同时留在磁盘上,用户可由 SSD 盘获取到高性能读取保证。一个闪存池通过故障转移和 RAID 保护具有缓存数据的持久性,智云的所有的数据保护和管理技术与 SSD Pool 一同工作。

Flash Pool 技术工作原理示意图:



当数据写入时,首先由存储控制器将数据写入到 SSD Pool 高性能缓存池中,之后存储系统自动将缓存池中的数据写入到 SATA、SAS 磁盘中。

解决影视、渲染、测绘用户拥有大规模数据集群高 i/o 业务并发问题,可以采用 SSD Pool 技术提高数据集群性能,同时采用磁盘混合配置方法降低系统建设成本。

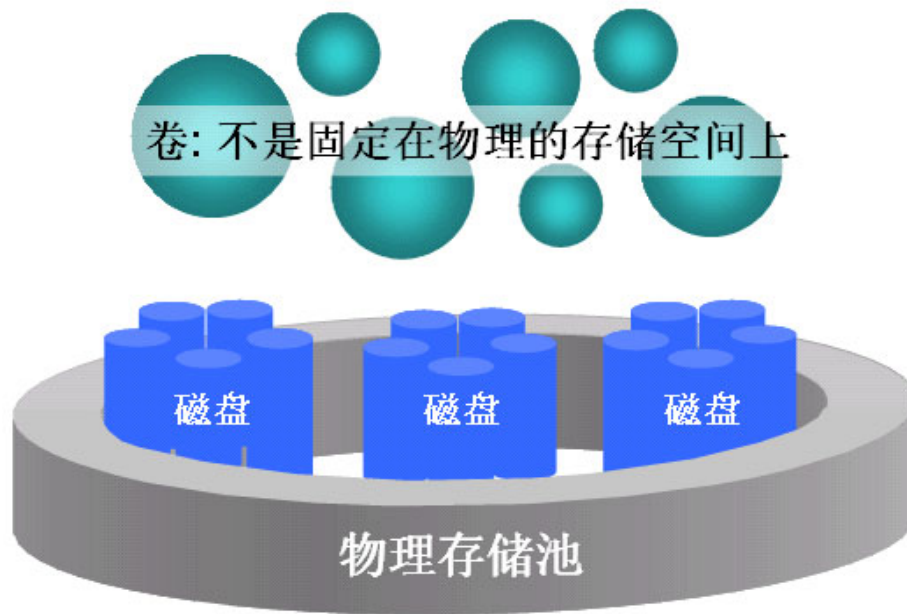
3.3 OmniVOL——灵活卷管理技术

OmniVol 是一个划时代的新技术，所有的卷被视为逻辑的数据容器，同时可以在不破坏底层物理存储结构的前提下被独立的设置、调整大小、管理和移动。

HS4000 存储系统 **OmniVol** 技术提供了真正的存储虚拟化解决方案，能够缩减开销和资本费用，减少业务中断并降低风险，同时还具有很高的灵活性，可以快速方便地适应企业不断变化的需求。**OmniVol** 技术可以自动集中存储资源，便于用户在一个大型磁盘池中创建多个灵活的卷。有了这一灵活性，就可以简化操作，最大限度地提高利用率和效率，并可以快速、无缝地进行修改。利用 HS4000 存储系统，用户可以随时随地根据需要以最低成本增加存储容量，并且无需中断业务运作。

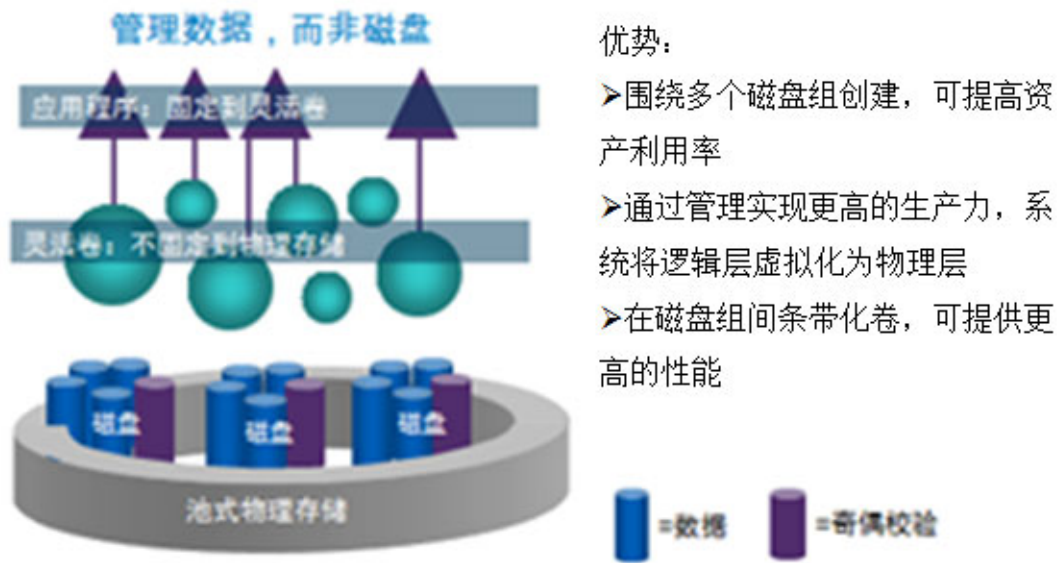
如下图所示，一个存储池被定义为给卷分配空间的许多磁盘的池。从管理者的角度来看，卷仍保持着对主要项的数据管理。但由于是对管理者透明，HS4000 的卷现在面向的是逻辑项而不是传统的物理存储了。

因此 HS4000 的卷不再受限于它们存在的磁盘的限制了。用户可以在一个存储的“存储池”中根据需要定义一个相应空间的卷，不再由其创建卷所在的磁盘个数硬性规定。一个卷也可以不需要停机任意收缩或者扩展。同时可以在存储池所拥有的所有物理磁盘上实现条带化。对于那些对性能敏感的应用来说，HS4000 的卷比那些由磁盘个数限定大小的卷拥有更好的性能。



➤ 性能的提高

在 HS4000 中，RAID 组被整合起来创建一个存储池。由于卷仍然是存储管理的基本单位，它将跨越存在于一个 HS4000 存储系统中组成存储池的所有磁盘，同时在这个大的存储池上也可以存在多个卷。这将使该卷充分利用所有磁盘的并行性能，满足在系统中某些比较繁忙的卷对于性能的需求。OmniVol 是灵活的，因为底层的存储的物理结构不需被预分区。



优势：

- 围绕多个磁盘组创建，可提高资产利用率
- 通过管理实现更高的生产力，系统将逻辑层虚拟化为物理层
- 在磁盘组间条带化卷，可提供更高的性能

图：OmniVol 比传统卷拥有更好的性能

➢ OmniVol 的容量池化

在 HS4000 中引入了一个新的存储管理的磁化概念——存储池。存储池的概念不同于以前用户们在使用 iSCSI 和 Fibre Channel 中所熟悉的“空间预留(space reservations)”。担保扩展了管理员的权限，使其在卷或者文件创建前可以决定预分配的策略。使其充分贯彻所谓的“自动精简配置 (thin provisioning) ”的概念。

担保，在卷一级设置，用以决定在存储池上给一个 OMNI volume 预分配多少空间。当用户在一个存储池上创建一个 **OmniVol**，将指定其容量，同时也可以指定担保的类型。

有三种担保的类型：

- 卷 (Volume)：卷担保类型确保在存储池上为 OMNI volume 分配的总空间总是可用的，这也是 OMNI volume 的默认设置。
- 文件 (File)：在文件担保类型中，存储池确保为可重写的 LUN 或者文件保留的空间总是可用。

- None: 对于一个无预留空间担保类型的 volume 来说, 不管给这给卷中的 LUN 设置了多少预留空间, 当容纳它的存储池没有足够的可用空间时, 对预留空间的 LUN 的写操作都将失败。

➤ 灵活的容量规划

对于一个 **OmniVol** 的大小本质上是没有限制的, 同时 volumes 可以动态的调整大小。管理员可以将 OMNI volumes 作为一个强大的工具为不同的用户、组和项目分配和提供存储资源。举例来说, 假设一个数据库比原先预计增长快很多时, 管理员可以在系统运行时随时重新配置相关的 volumes。重新分配存储资源的过程不需要任何宕机过程, 而且它是对用户透明的。

当需要额外的物理空间时, 管理员可以通过添加指定磁盘给存储池以提高其容量, 新的磁盘将成为存储池的一部分, 同时其容量和 I/O 带宽将对存在于该存储池中的所有的 volumes 生效。

所有存在于同一个存储池上的 OMNI volumes 的总容量也可以超过存储池的实际物理总容量。增加某一个 **OmniVol** 容量时也不需要改变存在相同存储池其它的 **OmniVol** 的容量或者存储池本身的容量。

➤ **OmniVol** 的技术限定

- HS4000 中最大的存储池个数 (包括存储池和传统卷) = 1000。
- HS4000 中最大的卷的个数 (包括传统卷和灵活卷) = 5000。
- 最大的存储池/传统卷的大小 = 1000TB。

➤ **OmniVol** 的技术优势

- 灵活的尺寸：HS4000 存储系统的灵活卷的最小颗粒度是 4KB，可以精确的适合各种大小的存储应用的需求。其它的系统需要卷或者文件系统的大小要基于底层物理或者逻辑磁盘的整数倍而定，显然会浪费数百兆甚至数 G 以至更多的物理容量。

- 灵活地改变尺寸（扩展和收缩）：HS4000 存储系统 **OmniVol** 技术允许 OFS 文件系统在线且安全的进行扩展和收缩，以使文件系统精确的适合用户的数据需求。其它的存储提供商基本能够提供不停机的容量扩展机制，但它们却不提供类似 HS4000 存储系统 **OmniVol** 的无缝和简单的容量扩展，并且不能提供在线且安全的文件系统收缩的功能。

- 空闲空间池：由于 HS4000 存储系统 **OmniVol** 可以像扩展一样简单的实现容量的收缩，已分配给 **OmniVol** 的未使用空间可以非常简单且安全的从 **OmniVol** 移出并重新分配给其它即将被填满需要更多空间的 **OmniVol**。通过 HS4000 存储系统 **OmniVol** 技术，未使用或者未分配的空间在存储池容器层被视为一个池（pool），可以被任意的 **OmniVol** 进行配置使用或者从 **OmniVol** 移出至池中。

- 瘦容量提供（Thin capacity provisioning）：HS4000 存储系统 **OmniVol** 可以实现瘦容量提供，即 **OmniVol** 仅仅占用已经被写入实际数据的物理存储空间，其他在 **OmniVol** 中已定义但未被使用的容量并不占用实际的物理磁盘空间。该功能在其它存储系统中是无法实现的，无论采用的是文件、LUN 或整个文件系统。

- 磁盘聚合提高性能：HS4000 存储系统 **OmniVol** 技术使用一个新的智能数据容器对象——存储池。通过使用存储池，**OmniVol** 可以充分利用一个含有大量磁盘的条带化的全部 I/O 对各种大小和种类的数据对象进行访问。

简单化的存储系统管理：通过使用存储池，HS4000 存储系统 **OmniVol** 实际上已经消除了了在存储产品上令数据管理员头痛的所有低级别的条带化的管理工作。存储池将注意力集

中于数据的管理,而不是磁盘的管理,将他们从一味的磁盘驱动器的配置的噩梦中解脱出来。

3.4 RAID PX——多重校验的高可靠 Raid 技术

传统的单一奇偶校验 RAID 技术, 能为单个故障磁盘驱动器提供保护。需要警告的是, 在重建故障磁盘时必须确保不能发生其他磁盘故障, 而且不能在读操作过程中出现不可修复的误码。如果在重建过程中出现上面任何一种事件, 那么保留在 RAID 阵列或卷中的部分或所有数据就可能会丢失。随着现代更大的磁盘存储介质的出现, 由于磁盘容量显著增加, 而误码率仍是原来的水平, 这就使得出现不可修复的误码的可能性相当高。因此, 传统的单一奇偶校验 RAID 保护数据的方法已经超出了其保护极限。更高一级的 RAID 数据保护已经发展到 RAID 双奇偶校验, 也称作 RAID PX, 它可以应用于整个网络应用数据存储产品线。

智云存储 HS4000 存储系统推出了双奇偶校验 RAID, 叫做 RAID PX。本节给出 RAID PX 的概览和 RAID PX 如何在各种磁盘故障情况下动态提高数据的容错性。其他所涉及的关键领域包括 RAID PX 的是免费的、无特殊硬件要求, 以及能够非常容易的将现存的 RAID 5 卷升级到 RAID PX。这里将借助对双磁盘故障的恢复情形来说明 RAID PX 如何做到允许卷既能不间断地提供数据, 又能同时对两块故障磁盘上损失的数据进行重建。

在最基本的层面上, RAID PX 为卷中的每个 RAID 组增加了第二个奇偶校验磁盘。RAID 组是建立卷的基本结构。每一个传统的 HS4000 存储系统 RAID PX 组都有一些数据磁盘和一个奇偶校验磁盘, 而一个卷通常会包括一个或多个 RAID PX 组。而 RAID PX 卷上的奇偶校验磁盘则通过 RAID PX 组上的磁盘对行奇偶校验进行存储, 额外的 RAID PX 奇偶校验磁盘则通过 RAID PX 组上的磁盘存储对角线上的奇偶校验。通过 RAID PX 上的这两个奇偶校验条, 一个传统的水平奇偶校验和一个对角线奇偶校验, 即使同一 RAID 组上的两个磁盘发生故障时也能得到数据保护。

在 RAID PX 中，仍然使用了传统的 RAID PX 水平奇偶校验结构，并且这种结构已成为 RAID PX 结构的子集，换句话说，RAID PX 并没有改变 RAID PX 在 HS4000 存储系统存储上的工作方式。数据通过每一行的奇偶校验计算被写入水平行，并被视作双奇偶校验中的行，这个同样的过程仍会保留到 RAID PX 中。实际上，如果是单一磁盘故障，或是从坏数据块中的读取错误，或者出现误码，那么 RAID PX 的行奇偶校验方法就是进行数据恢复的唯一手段，无需采用 RAID PX。在这种情况下，RAID PX 的对角线奇偶校验组件仅仅是行奇偶校验组件周围的保护层。

RAID PX 水平行奇偶校验

下列图表说明了在传统 HS4000 存储系统 RAID PX 解决方案中使用水平行奇偶校验的方法，也是您了解 RAID PX 和双奇偶校验的第一步。

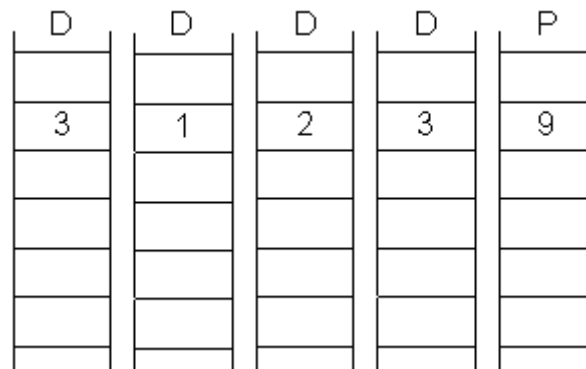


图 1: RAID P1 水平行奇偶校验

图表代表了传统的 RAID PX 组，使用包括 4 个数据磁盘（前 4 列，标注为“D”）和一列奇偶校验磁盘（最后一列，标注为“P”）的单向奇偶校验。上面图表中的行显示了实施传统 HS4000 存储系统 RAID PX 时使用的标准 4KB 块。上面图表中的第二行在每个 4KB 块和行中数据的奇偶校验计算中添加了一些样例数据，随后存储在奇偶校验磁盘上相应的块中。在这种情况下，计算奇偶校验的方式是向每个水平块中添加值，然后将这些值的和存储为奇偶校验值（ $3 + 1 + 2 + 3 = 9$ ）。实际上，奇偶校验是通过专门的 OR (XOR) 过程进

行计算，但其它情况也与此例非常相似。如果要从单一故障中重建数据，则产生奇偶校验的过程将与此相反。例如，当 RAID PX 恢复上面第一列的数据值 3 时，如果第一个磁盘发生故障，就会从奇偶校验中存储的值中减去剩余磁盘上的值 ($9 - 3 - 2 - 1 = 3$)。这个通过单一奇偶校验 RAID 重建的例子，应该有助于从概念上理解为什么能够对单一磁盘故障提供数据保护。

添加 RAID PX 双奇偶校验条

下表将一个以蓝色阴影块表示的对角线奇偶校验条和一个在第六列中以“DP”表示的第二个奇偶校验磁盘添加到前面的 RAID PX 组中。这表示 RAID PX 结构是 RAID PX 水平奇偶校验解决方案的超集。

D	D	D	D	P	P2
3	1	2	3	9	
		2			12
			2		
				7	

图 2：RAID PX 双奇偶校验条

在这个例子中对角线奇偶校验条通过相加的方法计算，并且存储在第二个奇偶校验磁盘上，而不是上面所说的实际使用的 XOR ($1 + 2 + 2 + 7 = 12$)。此时，需要注意的最重要的问题是，对角线奇偶校验条包括一个来自行奇偶校验的元素作为其对角线奇偶校验和的一部分。RAID PX 将把最初在 RAID PX 结构中的磁盘（包括数据和行奇偶校验磁盘）视作相同的磁盘。下一个图表为每个块加入了余下的数据，并且建立了相应的行和对角线奇偶校验条。

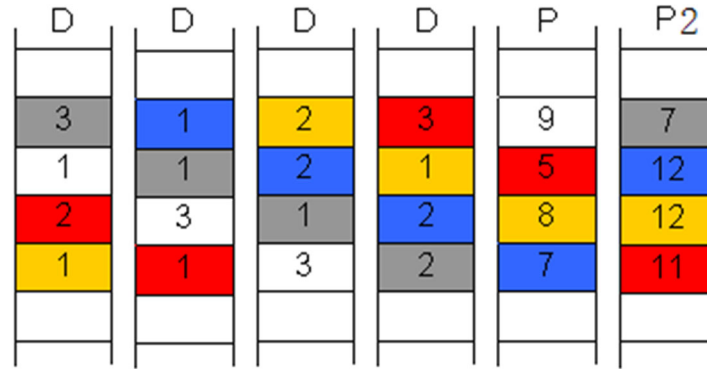


图 3: RAID PX

从上图可以明显看出, 在 RAID PX 中, 对角线奇偶校验条都在行奇偶校验结构的边缘。

在这个例子中, 对于 RAID PX 对双磁盘故障进行恢复的能力, 有两种重要情况可能不是很明显。第一种情况是, 每一个对角线条会错过唯一的一个磁盘, 但是每个对角线错过的磁盘都不一样。第二种情况的结果是, 在第二个对角线奇偶校验磁盘上会产生或存储一个没有得到奇偶校验的对角线条。在这个例子中, 被遗漏的对角线条是白色的块。在之后的重建例子中会明显的看出, 遗漏一个对角线条不会影响 RAID PX 在双磁盘故障中恢复所有数据的能力。

值得重视的是, 在本例中介绍的许多同样的 RAID PX 对角线奇偶校验都是真实的存储情况; 其中涉及到在 RAID 组中的许多磁盘上, 以及通过 RAID PX 组水平写入成数百万的数据行。并且, 通过上面较小的例子我们很容易地例证了 RAID PX, 而且不管 RAID 组中的有多少磁盘, 对于大尺寸的 RAID 组, 其恢复情况也与此相同。

由于 RAID PX 可以通过两种方式对双磁盘故障进行所有数据的恢复: 一种是使用数学原理并进行证明, 另一种则是简单地通过双磁盘故障并随后进行恢复。上述示例使用后面的方法来证明 RAID PX 的双奇偶校验保护的概念。

3.5 OmniMirror——异地数据容灾技术

技术原理

今天的全球性企业需要在发生自然或人为灾难、操作员出错或是技术和应用出现故障的情况下，保护数据并快速进行恢复。它们还需要一个有效的方式将数据发布到远程位置。如果没有有效的数据保护和发布策略，运作过程可能会被迫中断，导致数百万美元的收入损失。

智云的 OmniMirror 软件具有强大的功能，而且易于使用和管理，它满足了当前的全球性企业对灾难恢复和数据分布解决方案的需要。通过在 LAN 或 WAN 上高速复制数据，OmniMirror 软件将尽可能地为关键应用程序提供最高的数据可用性和最快的恢复速度。

OmniMirror 技术将数据镜像到一个或多个网络 NAS 存储上。OmniMirror 不断地更新镜像数据，以确保数据是最新的，并且能够用于进行灾难恢复、减少磁带备份、发布只读数据、在非生产性 NAS 存储上进行测试、执行联机数据迁移等等。如果您的企业分布在不同的地点，而所有地点都需要访问培训录像和 CAD 工具等相同的数据集，OmniMirror 可以将同一数据发布到所有地点。通过自动更新这些数据，并支持对镜像数据的本地访问方式，OmniMirror 可以大大提高员工的工作效率。

节省宝贵的网络带宽

智云的 OmniMirror 软件具有许多节省带宽的功能，可以降低数据复制和灾难恢复的基础设施成本。您可以先用磁带执行一次整卷数据传输，然后再用那些磁带在远程位置装入数据。之后，您只需要通过网络对新的数据块和更改的块进行增量更新。由于只需复制所有 NAS 存储数据的一小部分，OmniMirror 大大降低了对网络带宽的需求。此外，

OmniMirror 在数据传输过程中会设置检查点。如果系统出现故障，传输会从最近的检查点重新开始。OmniMirror 还会执行智能的重新同步过程，这样，在从镜像损坏的状态或不同步的状态进行恢复时，无需传输整卷数据。如果在应用程序测试期间修改了镜像副本上的数据，可以将生产系统上的新数据块和更改的数据块复制到镜像副本，使镜像副本快速实现与生产数据的重新同步。

灵活的配置

智云的 OmniMirror 可以方便地部署到任何具有足够的带宽来处理数据传输的网络基础设施中。它能够支持多种传输方式 (FC 和 IP)，这一点确保可以更好地利用现有设备，并且还因能够实现路径间的容错而提高了可用性。

OmniMirror 允许客户选择适当的同步级别 (同步、半同步和异步)，从而实现了较高的灾难恢复保护级别。例如，如果选择“同步”选项的话，远程站点的复制数据将始终是最新的，并能在发生故障后随时提供。这一方式简化了灾难恢复过程，降低了系统宕机时间。而“半同步”允许客户根据站点的需要，确定有多少 I/O 操作可以失去同步，或者是复制站点可以在多长时间内与数据源失去同步。而“异步”选项则允许您根据需要 (每分钟、每小时或每天) 任意安排传输过程。您可以设置最适合每一站点的频率。时间安排很容易修改，而且修改后可以立即生效。您还可以为源系统和镜像系统选择不同的 NAS 存储配置。源系统可以是具有 6TB 存储空间的集群 NAS 存储，它将 2TB 的关键数据镜像到另一种型号的 NAS 存储上。此外，得益于级联和多跃点镜像方式，智云的 OmniMirror 目标卷可以作为其他目标的源卷，而且每个“镜像对”都可以有自己的时间安排以满足各个站点的特定需求。如果需要远距离复制数据，级联镜像是非常理想的解决方案。例如，如果是从北京将数据复制到上海、南京和杭州，您可以将数据从北京复制到上海，然后使用低成本链接将数据从上海复制到南京和杭州。

易于使用，经济合算

用户在几分钟内就可以设置好 OmniMirror，并能通过易于使用的 DataFabric® Manager (DFM) 图形用户界面进行管理和操作。NAS 存储支持 OmniMirror 软件，您不需要再安装其他软件。通过内置的 SNMP 支持功能，可以方便地与 SNMP 框架进行集成。

技术特点

快速数据复制和容错—将主站点发生故障的宕机成本降至最低。

访问镜像数据—减少磁带备份，让您在灾难恢复上的投资物超所值。

卷或 Qtree 复制—镜像选定的数据集，大大降低了对网络基础设施的要求。

选择复制过程的同步级别—采用一个产品即可控制复制的频率（异步、同步或准同步）。

更有效地利用网络—降低了数据复制和灾难恢复的成本。

设置简便—几乎不需要增加 IT 资源；可以经常测试灾难恢复计划。

OmniMirror 可以用于：

- 灾难恢复—将数据镜像到远程位置以实现容错。
- 数据分布—使用级联或多跃点镜像方式，将数据集发送到世界各地。
- 远程数据访问—应用程序能够以只读模式访问镜像数据。
- 联机数据迁移—将数据迁移导致的宕机时间减到最少。
- 数据复制—使用镜像数据可以进行隔离测试。

- 同步的数据复制—保证各站点间关键数据的一致性。
- 负载均衡—通过访问镜像数据，在更多的客户端之间分摊负载。
- 实现一致的数据保护—使用 SnapLock™ 软件，能够在 Worm 卷之间实现完全一致的复制。

技术优势

- 智云的OmniMirror支持多种复制方式，用于适应不同的容灾环境，包括：同步，准同步，异步。异步方式的复制对于距离没有限制。
- OmniMirror支持多种拓扑方式的复制，包括：1对多；多对1；级联复制等。
- OmniMirror支持对多种数据的复制，包括：基于卷的复制，基于目录的复制，支持对NAS / iSCSI中的数据复制。
- OmniMirror只传输变化的数据块，可以大大节省网络带宽和减轻NAS存储的负载。

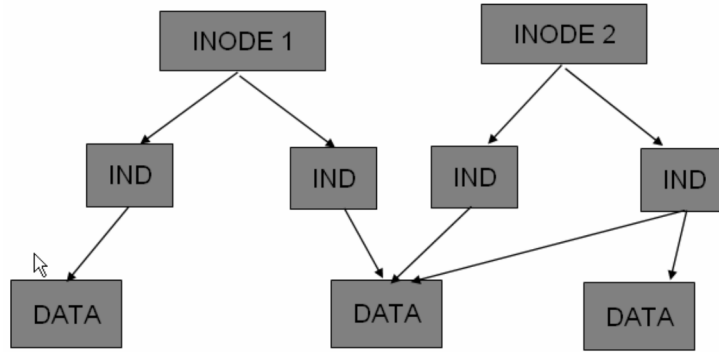
3.6 OmniDedup——重复数据删除技术

技术原理

HS4000 存储系统允许用户在任何存储环境中使用数据重复删除技术，包括生产应用环境、与备份归档环境。这将帮助用户以更少的存储空间来存储更多数据。

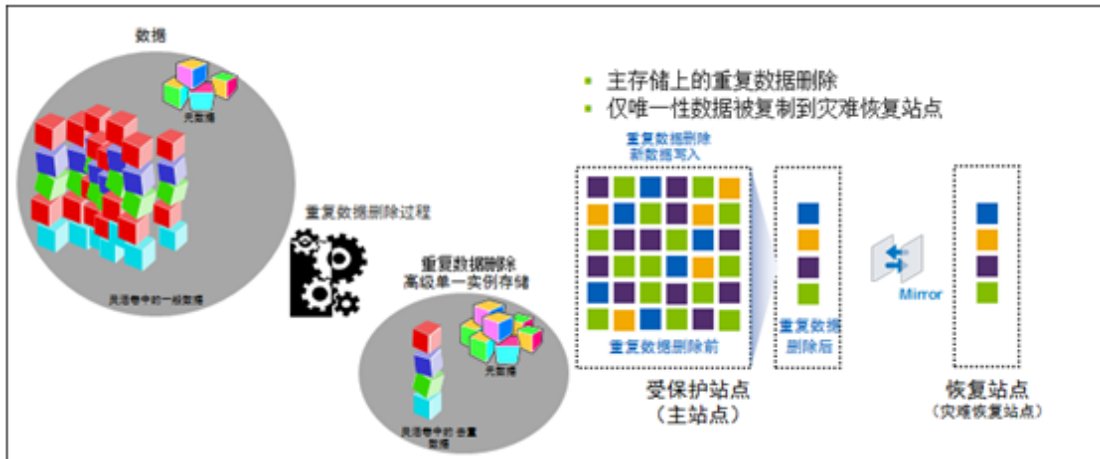
HS4000 存储系统的重复数据删除技术 **OmniDedup** 与应用无关，无论是文件存储、数据库、SAN 访问方式或者 NAS 访问方式，**OmniDedup** 都能够找出相同数据块并予以合并。

在存储系统中实现 **OmniDedup** 的关键是采用了 Dedup 卷。Dedup 卷是一个特殊的灵活卷，它允许卷中的包含相同内容的数据块共享物理存储空间。而 HS4000 的核心技术 OFS，在结构上支持这样的共享，从而能够最终实现存储空间的节省。如下图所示，OFS 的数据块索引结构支持多个索引指针指向同一个物理数据块，从而实现空间共享。



为跟踪这些指向自己的索引指针，每一个数据块还需要保持一个计数器，当指向自己的指针增加或者减少，计数器要相应地增减，以反应指针的变化。

OmniDedup 在 Dedup 卷内数据块级别实现重复删除，下图给出了其工作原理的逻辑示意



Dedup 的工作流程大致如下：当在一个数据卷上启用 Dedup 之后，存储系统仍然会以正常的方式保存新写入的数据块。系统为每个新数据块生成一个对应的指纹信息，用于与数据卷中其他数据块的指纹信息进行比对。如果发现两个数据块的指纹信息相同，则会对这两个数据块进行完全的字节比对，如果仍然相同，则存储系统就会丢弃重复的数据块并且回

收对应的存储空间。

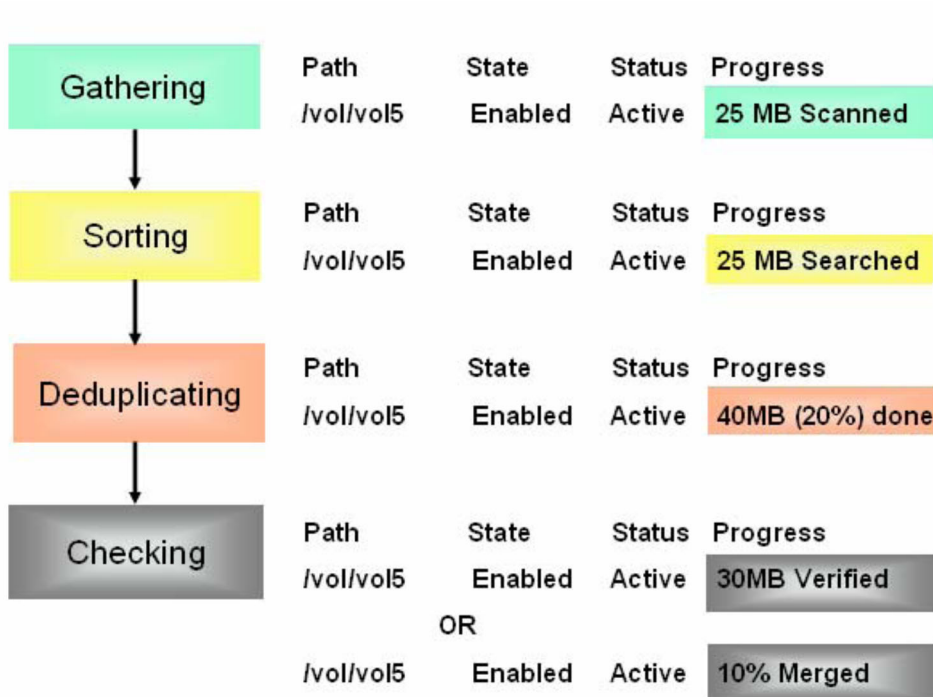
Dedup 进程以以下三种方式之一运行：

设定计划表，进程将按时启动运行；

通过命令行手动运行；

当累计有 20% 的新数据被写入到数据卷时，自动运行；

下图给出了 Dedup 进程的不同阶段的状态，



Gathering: 对数据卷中所有数据进行扫描并生成指纹信息，该过程仅对在已有数据的数据卷上激活的 Dedup 功能时才发生；如果在激活 Dedup 功能后写入的新数据，在其写入的同时系统指纹信息已经生成。

Sorting: 扫描指纹信息库并进行比对。

Deduplicating: 根据指纹信息和字节比对结果去除重复数据块，并修改数据块指针。

Checking: 检查指纹信息库，并去除其中的冗余信息（即：已合并的数据块的指纹信

息)。

技术特点

粒度细，在数据块级别；

作为后台进程操作，并可配置成自动运行，手动运行或计划表运行多种方式；

对应用透明，可用于数据中心的各个层面；

可以对卷中已有的数据进行重复删除操作；

US 全系列支持；

对于不同的应用环境，节省的空间比可能达到 20:1 甚至更高；

支持双控制器 cluster 结构；

技术优势

与存储协议无关

OmniDedup 对数据卷中的数据块进行比对操作并删除重复数据，与通过何种存储协议写入数据无关，无论是 FCP、NFS、CIFS、ISCSI 的数据均可以。因此从 Dedup 的操作层面看，对应用完全透明。

可直接用于生产环境

OmniDedup 技术是 MEKONG 系统的一个软件特性，只要一个 license 即可激活使用 (而且免费)。因此，可以方便地用于生产环境，当然也可以用于备份、归档等应用环境。特别是在服务器虚拟化等存储消耗型应用中，使用 **OmniDedup** 可以节省大量存储成本和空间。

性能损失最小

MeKong 存储系统的 **OmniDedup** 技术与 OFS 结构紧密集成，因此具有极高的效率。

不需要复杂的哈希算法和数据表查询，相反 **OmniDebup** 能够利用到 OFS 自身的数据结构特性来生成和比对指纹信息，重定向数据块指针，以及释放冗余数据块空间，因此将对性能的影响减至最小。

与 SnapMirror 结合，减少 WAN 流量

OmniDebup 可以与 **OmniMirror** 结合，将 **OmniDebup** 数据卷的数据 **OmniMirror** 到远端 HS4000。由于数据在源端已经 **OmniDebup** 过，通过网络传输的数据量将大大减少，从而进一步降低 WAN 带宽的需求，节省成本并且能够实现更好的 RPO。

3.7 维护管理

➤ 技术原理

每个 HS4000 存储系统设备上都提供内置的 GUI 工具，IT 管理员可以通过 Web 浏览器进入一个一致、易用的图形用户界面 (GUI)，进行日常管理。

管理员可以从任何地方，从任何计算平台远程设置并控制所有 HS4000 存储系统 存储设备，而无需中断关键业务操作。安装后，管理员可以有效地监视状态、满足增加存储容量的请求，并对文件系统配置进行更改。管理员还可以在不中断用户访问文件系统的情况下扩展卷，或在连接的 Windows 域中添加新的通用 InternOmni 文件系统 (Common InternOmni File System, CIFS) 共享。

HS4000 存储系统 GUI 是用于执行从安装到完成配置的管理任务的主界面。利用 HS4000 存储系统 View，管理员不必再学习特殊的语言或命令行界面来配置和维护 HS4000 存储系统 存储设备。只需一个 Web 浏览器，管理员就可以进入易用的图形用户界面，并且可以通过浏览各个屏幕来执行任务。界面上将始终显示一个浏览树，以便于在各个任务间切换。

OmniView 将 HS4000 的安装过程简化到只需执行几个简单的步骤。通过控制台或动态主机配置协议 (Dynamic Host Configuration Protocol, DHCP) 服务器将一个 IP 地址分配给 US 后, 管理员可以通过安装向导快速完成基本配置设置。

由于 **OmniView** 采用基于 Web 的方法, 因此无需学习和使用复杂难懂的命令行界面执行管理任务。通过一个图形用户界面, 管理员可以配置和监视任何系统资源或服务:

文件系统: NFS 与 CIFS 实施的磁盘、卷限额; **OmniView** 卷报告显示所有卷的当前配置状态。

数据保护: OMNISHot 和 **OmniMirror** 工具、网络数据管理协议 (NOmniwork Data Management Protocol, NDMP)

可用性: 集群容错软件

联网: 接口卡; HTTP、DNS (域名服务器) 和 NIS (网络信息系统) 服务器

管理: 安全功能、autosupport 和 **OmniView** 软件、SNMP 实用程序、资源监视工具

OmniView 的安全模式确保所有管理和配置命令都严格保密, 无论通过专用网络或公共 InternOmni 执行。安全模式需要 OMNICLI 软件, 该软件在上实施安全套接字层 (Secure SockOmni Layer, SSL 3.0) 服务, 用于进行用户验证和保密通过 HTTPS 处理的命令。

➤ 技术特点

完整的基于 Web 的管理工具集, 提供完善的存储设备配置管理功能, 监视功能。支持基于 SSL 的安全管理。

➤ 技术优势

在简单方便的图形化 Web 管理界面中提供易学易用的基本管理功能。低成本的管理工具。

4 HS4000 系列产品技术规格

系统硬件规格			
	HS4012	HS4024	HS4036
控制器数量	1	1	1
控制器外形规格	2U12 盘位	4U24 盘位	4U36 盘位
CPU	1*海光 5380/16 核 CPU	1*海光 5380/16 核 CPU	1*海光 5380/16 核 CPU
支持扩展盘位数	120	320	320
磁盘类型	SATA/SAS/SSD		
标配内存/最大内存	32GB/128GB		
千兆主机接口 (板载)	2		
万兆主机接口数 (选配)	2		
PCIe 扩展	6		
最大带宽	3GB/s	3GB/s	3GB/s
端口类型	1Gbps/10Gbps/25Gbps/100Gbps Ethernet、SAS2.0 (后端接口, 单端口 4*12Gbps)		
支持的磁盘扩展柜	ED3324 (4U; 24 个驱动器, 3.5 英寸 LFF)		
	ED3360 (4U; 60 个驱动器, 3.5 英寸 LFF)		
系统软件规格			
系统平台	OmniCOS v3 存储专用操作系统		
文件系统	OFS 事务型文件系统, 无单卷容量限制, 在线动态伸缩卷尺寸, 用户和目录级的配额管理		
RAID 保护	0,1,10,RAID P1(RAID5),RAID P2(RAID6),RAID P3(三重校验)		
支持的存储网络	iSCSI、NFS、CIFS/SMB、HTTP、FTP		
访问权限控制	ADS,NIS,LDAP,Kerberos,Home Directory		
网络协议支持	NTP,Telnet,SSH,SMTP,802.3AD,LACP,NDMP v2/v3		
系统监控	在线系统日志查看, 系统故障诊断, 网络/CPU/硬盘等状态实时监控, 声、光、E-mail 自动故障报警, SNMP v2/v3 协议		
管理界面	B/S 图形化操作界面(InnoView)标准中文和英文, 命令行管理 (CLI)		
数据管理	灵活卷(OmniVol), 重复数据删除(OmniDedup), 数据压缩(OmniCompress), 卷克隆(OmniClone), 虚拟化集成 (OmniVM), 数据自动分层 (OmniTiering)、热点数据二级缓存 (OmniCache)		
数据保护	数据快照 (OmniShot), 一键快照恢复 (OmniRestore), 数据加密		

许可软件	远程数据镜像(OmniMirror),企级网盘(OmniDrive), 文件备份		
最大 LUN 数量	1024		
支持的 SAN 主机数量	最多可支持 1024 个主机		
支持的操作系统	Windows 2000、Windows Server 2003、Windows Server 2008、Windows Server 2012、Windows XP、Linux®、Oracle Solaris、AIX、HP-UX、Mac® OS、VMware® ESXi、Citrix® XenServer、麒麟 Linux、统信 Linux®		
系统环境规格			
	HS4012	HS4024	HS4036
额定功率	450W	550W	800W
尺寸 (深*宽*高)	670*444*87.6mm	670*444*177mm	670*444*177mm
工作环境要求	10°C 至 40°C (50°F 至 104°F)；海拔高度不超过 3,000 米 (10,000 英尺)；相对湿度 20% 至 80%，非凝结		
存放条件	-40°C 至 70°C (-40°F 至 158°F)；在无空调的飞机货舱中，一般不高于 12,192 米 (40,000 英尺)，相对湿度 8% 至 80%，非凝结 (在原包装中)		