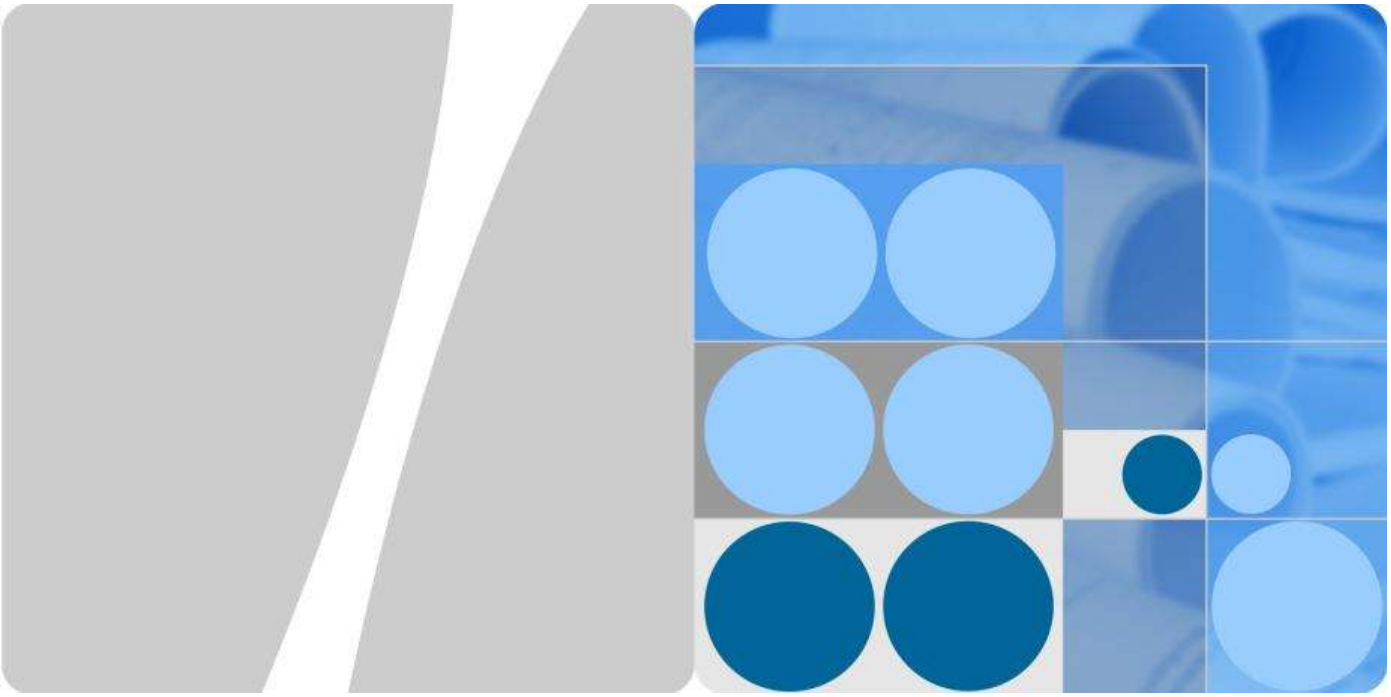


OceanStor 高端存储系统

RAID2.0+技术白皮书

文档版本 01

发布日期 2013-8-12

华为技术有限公司

商标声明


和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或默示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为技术有限公司

地址： 深圳市龙岗区坂田华为总部办公楼 邮编：518129

网址： http://www.huawei.com

客户服务邮箱： support@huawei.com

客户服务电话： 4008302118

目 录

修订记录/Change History

日期	修订版本	描述	作者
20130802	V1.0		秦烜/204091
20130830	V1.1		秦烜/204091
20131014	V1.2	更名	秦烜/204091
20131112	V1.3	刷新 FAQ	秦烜/204091

1 RAID2.0+概述

1.1 RAID 技术演变

RAID（Redundant Array of Independent Disk，独立冗余磁盘阵列）技术诞生于 1987 年，最初由美国加州大学的伯克利分校提出，其基本思想是把多个独立的物理硬盘通过相关的算法组合成一个虚拟的逻辑硬盘，从而提供更大容量、更高性能，或更高的数据容错功能。

作为一种成熟、可靠的磁盘系统数据保护标准，RAID 技术自诞生以来一直作为存储系统的基础技术而存在，但是近年来随着数据存储需求的快速增长，高性能应用的不断涌现，传统 RAID 逐渐暴露出越来越多的问题。

IDC 预测，未来 5 年内存储市场将继续保持年平均增长 10%以上的良好态势，全球存储总容量可能达到 16840PB。为了满足数据增长的需求，磁盘设备制造商不断地提升技术来增加磁盘单位存储密度，如今，4TB 的高容量磁盘和 900GB 的高性能 SAS 磁盘在企业和消费市场已经非常普遍，当这些高容量磁盘由于出现磁盘故障而需要进行数据重构时，传统 RAID 的弱点便会立即凸显。

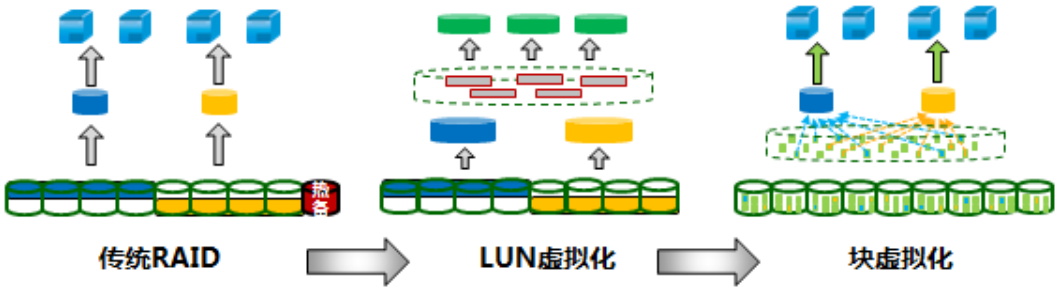
以 7.2K RPM 4TB 磁盘为例，在传统的 RAID5（8D+1P）中，其重构时间在 40 个小时左右。重构的进程会占用系统的资源，导致应用系统整体性能下降，当用户为了保证应用的及时响应而限制重构的优先级时，重构的时间还将进一步延长。此外，在漫长的数据重构过程中，繁重的读写操作可能引起 RAID 组中其他磁盘也出现故障或错误，导致故障概率大幅提升，极大地增加数据丢失的风险。

另一方面，传统 RAID 受限于硬盘数量，在数据容量剧增的年代无法满足企业对资源统一灵活调配的需求，同时，随着硬盘容量的增大，以硬盘为单位对数据进行管理也显得越来越力不从心。

为了解决传统 RAID 的上述问题，同时顺应虚拟化技术的发展趋势，众多存储厂商纷纷提出了传统 RAID 技术的替代方案：

- **LUN 虚拟化**：以 EMC 和 HDS 为代表的存储厂商，在传统 RAID 基础之上将 RAID 组进行更细粒度地切分，再将切分的单元进行组合，构建主机可访问的空间。
- **块虚拟化**：以华为和 HP 3PAR 为代表的存储厂商，将存储池中的硬盘划分成一个个小粒度的数据块，基于数据块来构建 RAID 组，使得数据均匀地分布到存储池的所有硬盘上，然后以数据块为单元来进行资源管理。

图 1-1 RAID 技术发展



1.2 华为 RAID2.0+简介

OceanStor 高端存储系统是华为技术有限公司（以下简称华为）根据存储产品应用现状和存储技术未来发展趋势，针对企业大中型数据中心，推出的聚焦于大中型企业核心业务的新一代（虚拟化、混合云、精简 IT 和低碳等）存储系统。

OceanStor 高端存储系统采用创新的 Smart Matrix 全交换式硬件架构，并结合专用的 XVE(eXtreme Virtual Engine)存储操作系统，来满足大型数据中心对存储系统的各种需求。

RAID2.0+技术是华为针对传统 RAID 的缺点，设计的一种满足存储技术虚拟化架构发展趋势的全新的 RAID 技术，其变传统固定管理模式为两层虚拟化管理模式，在底层块级虚拟化（**Virtual for Disk**）硬盘管理的基础之上，通过一系列 Smart 效率提升软件，实现了上层虚拟化（**Virtual for Pool**）的高效资源管理。

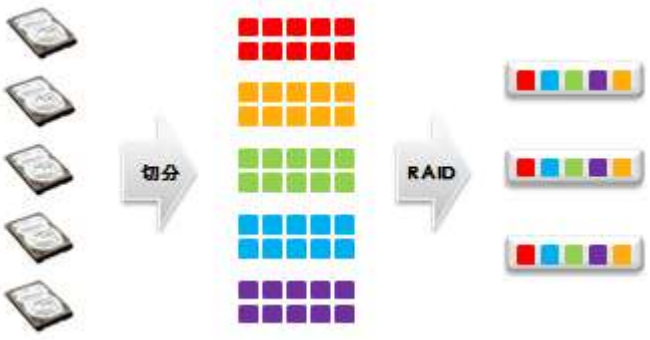
2 RAID2.0+热备策略

2.1 RAID2.0+基本原理

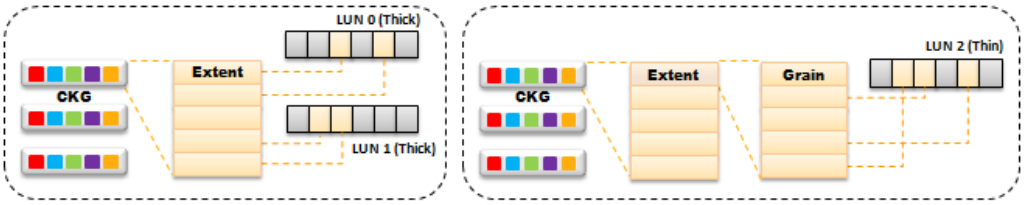
华为 RAID2.0+采用底层硬盘管理和上层资源管理两层虚拟化管理模式，在系统内部，每个硬盘空间被划分成一个个小粒度的数据块，基于数据块来构建 RAID 组，使得数据均匀地分布到存储池的所有硬盘上，同时，以数据块为单元来进行资源管理，大大提高了资源管理的效率。

1) OceanStor 高端存储系统支持三种不同类型（SSD、SAS 和 NL-SAS）的硬盘，这些硬盘组成一个个的硬盘域（Disk Domain），在硬盘域中，同种类型的硬盘按照一定的规则被划分为一个个的 Disk Group（DG）；

2) 在 DG 中，每个硬盘被切分成固定大小的数据块（Chunk，也叫 CK），OceanStor 高端存储系统通过随机算法，将不同硬盘的 Chunk（CK）按照 RAID 算法组成 Chunk Group（CKG）；

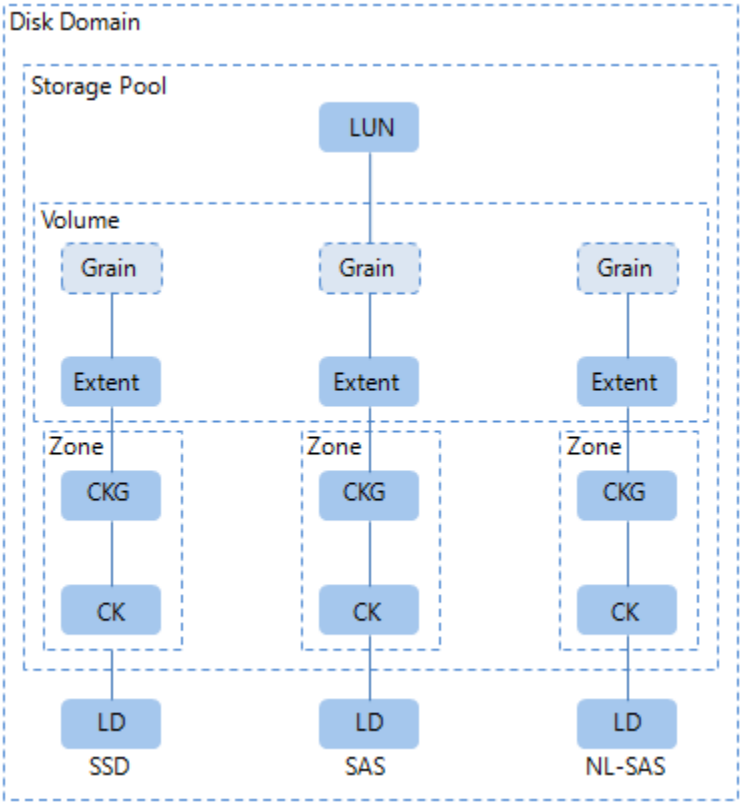


3) CKG 被划分为固定大小的逻辑存储空间（Extent），Extent 是构成 Thick LUN（也叫 FAT LUN）的基本单位；对于 Thin LUN，会在 Extent 上再进行更细粒度的划分（Grain），并以 Grain 为单位映射到 Thin LUN。



2.2 RAID2.0+实现框架

OceanStor 高端存储系统 RAID2.0+的实现框架如下图所示：



- OceanStor 高端存储系统的硬盘域（Disk Domain）由一个或多个层级的硬盘组成，不同层级支持不同类型的硬盘：构成高性能层的 SSD 硬盘，构成性能层的 SAS 硬盘和构成容量层的 NL-SAS 硬盘。
- 各存储层的硬盘被划分为 **64MB** 固定大小的 Chunk（CK）。
- 每一个存储层的 Chunk（CK）按照用户设置的“RAID 策略”来组成 Chunk Group（CKG），用户可以为存储池（Storage Pool）中的每一个存储层分别设置“RAID 策略”。
- OceanStor 高端存储系统会将 Chunk Group（CKG）切分为更小的 Extent。Extent 作为数据迁移的最小粒度和构成 Thick LUN 的基本单位，在创建存储池（Storage Pool）时可以在“高级”选项中进行设置，默认 **4MB**。

  若干 Extent 组成了卷（Volume），卷（Volume）对外体现为主机访问的 LUN（这里的 LUN 为 Thick LUN）。在处理用户的读写请求以及进行数据迁移时，LUN 向存储系统申请空间、释放空间、迁移数据都是以 Extent 为单位进行的。例如：用户在创建 LUN 时，可以指定容量从某一个存储层中获得，此时 LUN 由指定的某一个存储层上的 Extent 组成。在用户的业务开始运行后，存储系统会根据用户设定的迁移策略，对访问频繁的数据以及较少被访问的数据在存储层之间进行迁移（此功能需要购买 SmartTier License）。此时，LUN 上的数据就会以 Extent 为单位分布到存储池的各个存储层上。

- 在用户创建 Thin LUN 时，OceanStor 高端存储系统还会在 Extent 的基础上再进行更细粒度的划分（Grain），并以 Grain 为单位映射到 Thin LUN，从而实现对存储容量的精细化管理。

2.3 RAID2.0+逻辑对象

本章节主要针对 RAID2.0+的主要软件逻辑对象和重点概念进行阐述。

Disk Domain

Disk Domain 即**硬盘域**，是一堆硬盘的组合（可以是整个系统所有硬盘），这些硬盘整合并预留热备容量后统一向存储池提供存储资源。

- OceanStor 高端存储系统可以一个或多个硬盘域
- 一个硬盘域上可以创建多个存储池（Storage Pool）
- 一个硬盘域的硬盘可以选择 SSD、SAS、NL-SAS 中的一种或者多种
- 不同硬盘域之间是完全隔离的，包括故障域、性能和存储资源等

Storage Pool & Tier

Storage Pool 即**存储池**，是存放存储空间资源的容器，所有应用服务器使用的存储空间都来自于存储池。一个存储池基于指定的一个硬盘域创建，可以从该硬盘域上动态的分配 Chunk（CK）资源，并按照每个存储层级（Tier）的“RAID 策略”组成 Chunk Group（CKG）向应用提供具有 RAID 保护的存储资源。

Tier 即**存储层级**，存储池中性能类似的存储介质集合，用于管理不同性能的存储介质，以便为不同性能要求的应用提供不同存储空间。存储池根据硬盘类型可划分为多个 Tier，OceanStor 高端存储系统支持的存储层级和硬盘类型如下表所示：

存储层级	层级名称	支持硬盘类型	应用
Tier0	高性能层	SSD	性能和价格较高，适合存放访问频率很高的数据
Tier1	性能层	SAS	性能较高，价格适中，适合存放访问频率中等的数据
Tier2	容量层	NL-SAS	性能较低，价格最低且单盘容量大，适合存放大容量的数据以及访问频率较低的数据

- 创建存储池可以指定该存储池从硬盘域上划分的存储层级（Tier）类型以及该类型的“RAID 策略”和“容量”。
- OceanStor 高端存储系统支持 RAID5、RAID6 和 RAID10，支持的 RAID 策略和配置如下表所示：

RAID 级别	RAID 策略
RAID5	4D+1P，8D+1P
RAID6	4D+2P，8D+2P
RAID10	系统自动选择 2D +2D 或 4D+4D

- 容量层由大容量的 NL-SAS 盘组成，RAID 策略建议使用双重校验方式的 **RAID6**。

Disk Group（DG）

Disk Group（DG）即**硬盘组**，由硬盘域内相同类型的多个硬盘组成的集合，硬盘类型包括 SSD、SAS 和 NL-SAS 三种。OceanStor 高端存储系统会在每个硬盘域内根据每种类型的硬盘数量自动划分为一个或多个 Disk Group（DG）。

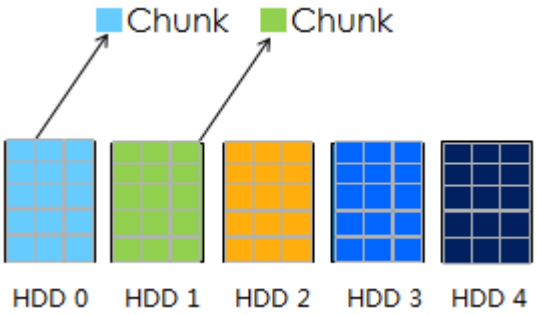
- 一个 Disk Group（DG）只包含一种硬盘类型
- 任意一个 CKG 的多个 CK 来自于同一个 Disk Group（DG）的不同硬盘
- Disk Group（DG）属于系统内部对象，主要作用为故障隔离，由 OceanStor 高端存储系统自动完成配置，对外不体现

Logical Drive（LD）

Logical Drive（LD）即**逻辑磁盘**，是被 OceanStor 高端存储系统所管理的硬盘，和物理硬盘一一对应。

Chunk（CK）

Chunk 简称 **CK**，是存储池内的硬盘空间切分成若干固定大小的物理空间，每块物理空间的大小为 **64MB**，是组成 RAID 的基本单位。



Chunk Group（CKG）

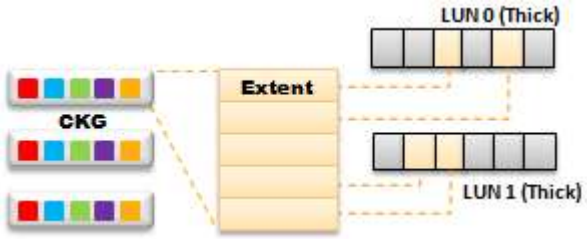
Chunk Group 简称 **CKG**，是由来自于同一个 DG 内不同硬盘的 CK 按照 RAID 算法组成的逻辑存储单元，是存储池从硬盘域上分配资源的最小单位。

- 一个 CKG 中的 CK 均来自于同一个 DG 中的硬盘
- CKG 具有 RAID 属性（RAID 属性实际配置在 Tier 上）

- CK 和 CKG 均属于系统内部对象，由 OceanStor 高端存储系统自动完成配置，对外不体现。

Extent

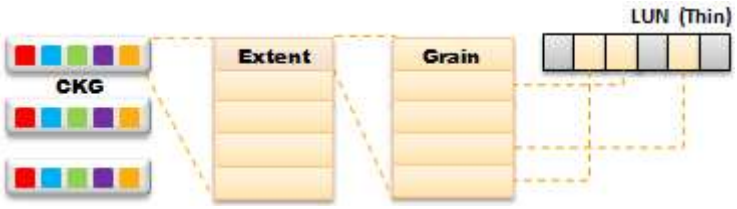
Extent 是在 CKG 基础上划分的固定大小的逻辑存储空间，大小可调，范围为 **512KB～64MB**，默认为 **4MB**，是热点数据统计和迁移的最小单元（**数据迁移粒度**），也是存储池中申请空间、释放空间的最小单位。



- 一个 Extent 归属于一个 Volume 或一个 LUN
- Extent 大小在创建存储池时可以进行设置，创建之后不可更改
- 不同存储池的 Extent 大小可以不同，但同一存储池中的 Extent 大小是统一的

Grain

在 Thin LUN 模式下，Extent 按照 **64KB** 的固定大小被进一步划分为更细粒度的块，这些块称之为 **Grain**。Thin LUN 以 Grain 为粒度进行空间分配，Grain 内的 LBA 是连续的。



- Thin LUN 以 Grain 为单位映射到 LUN，对于 Thick LUN，没有该对象

Volume & LUN

Volume 即**卷**，是系统内部管理对象，一个 Volume 对象用于组织同一个 LUN 的所有 Extent、Grain 逻辑存储单元，可动态申请释放 Extent 来增加或者减少 Volume 实际占用的空间。

LUN 是可以直接映射给主机读写的存储单元，是 Volume 对象的对外体现。

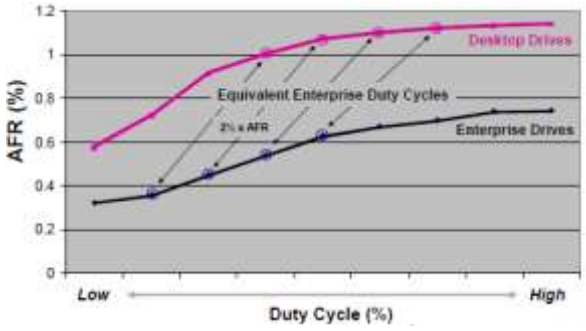
3 RAID2.0+技术亮点

RAID2.0+通过两层虚拟化管理模式，克服了传统 RAID 的一些固有缺点，大大提升了存储系统的可靠性和资源管理的效率，借助于 RAID2.0+的创新技术，OceanStor 高端存储系统真正实现了高端存储的安全可信、弹性高效。本章节主要从以上两个维度来介绍 RAID2.0+的技术亮点。

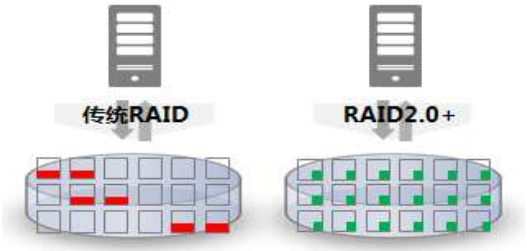
3.1 安全可信

3.1.1 自动负载均衡，降低整体故障率

传统 RAID 存储系统中一般会有多个 RAID 组，每个 RAID 组中包含几块到十几块硬盘。由于每个 RAID 组的业务繁忙程度不同，导致硬盘的工作压力不均，部分硬盘存在热点，根据 SNIA 的统计数据，热点盘的故障率会明显增高。如下图所示，其中 Duty-Cycle 即忙闲度，指的是硬盘工作时间占总上电时间的比例，AFR 为硬盘年故障率，不难看出，Duty-Cycle 高时的硬盘年故障率几乎是低时的 1.5～2 倍。



RAID2.0+技术通过块虚拟化实现了数据在存储池中硬盘上的自动均衡分布，避免了硬盘的冷热不均，从而降低了存储系统整体的故障率。



3.1.2 快速精简重构，改善双盘失效率

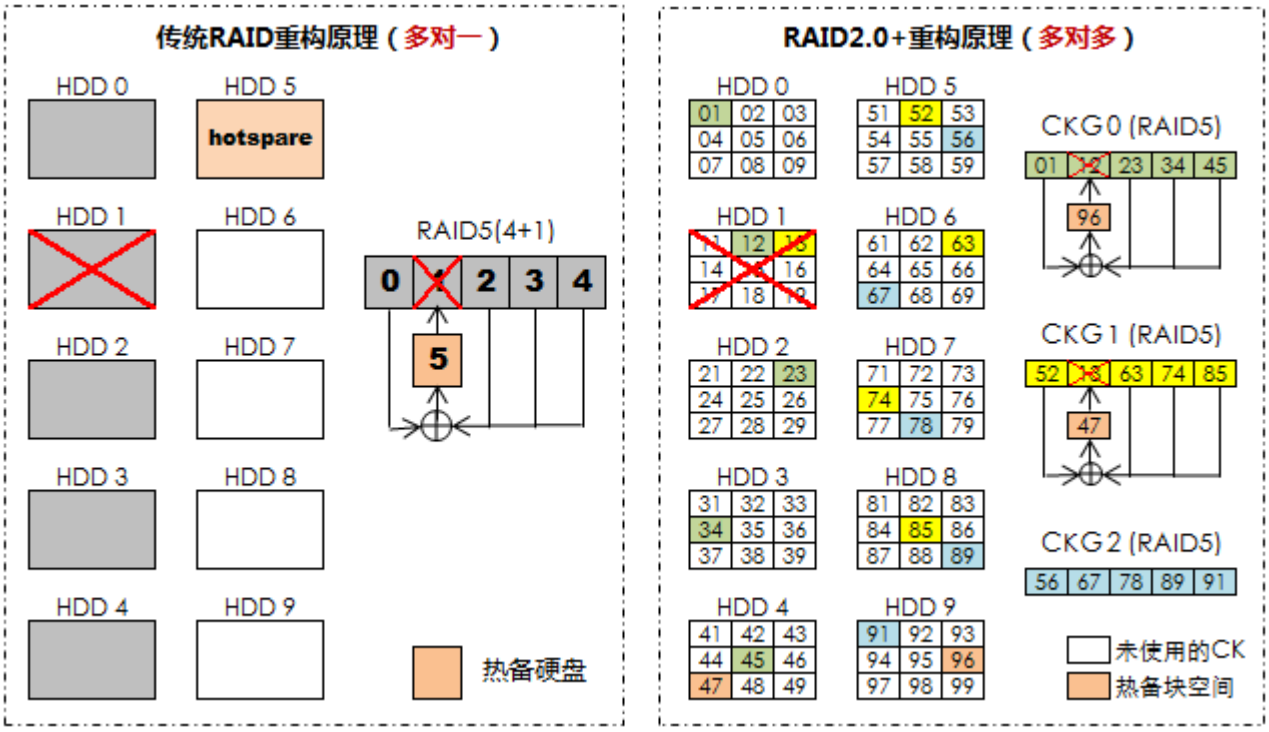
纵观近 10 年硬盘的发展，其容量的增长远远快于性能的进步，现在 4TB 的高容量磁盘在当前的企业和消费市场已经非常普遍，而 5TB 的高容量磁盘也将在 2014 年 Q2 会出现，即便是专门针对企业市场的高性能 SAS 磁盘，也已经达到了 1.2TB 的容量。

容量的增长使得传统 RAID 不得不面临一个严重的问题：10 年前重构一块硬盘可能只需要几十分钟，而如今重构一块硬盘需要十几甚至几十个小时。越来越长的重构时间使得企业的存储系统在出现硬盘故障时长时间处于非容错的降级状态，存在极大的数据丢失风险，存储系统在重构过程中由于业务和重构的双重压力导致数据丢失的案例也屡见不鲜。

基于底层块级虚拟化的 RAID2.0+技术由于克服了传统 RAID 重构的目标盘（热备盘）性能瓶颈，使得重构数据流的写带宽不再成为重构速度的瓶颈，从而大大提升了重构速度，降低了双盘失效的概率，提升了存储系统的可靠性。

下图是传统 RAID 和 RAID2.0+两种技术重构原理的对比：

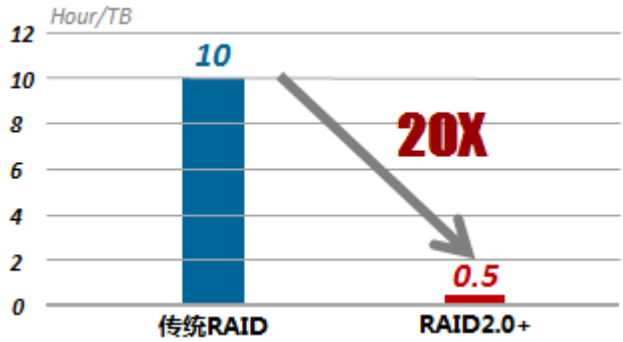
- 左图传统 RAID 中，HDD0～HDD4 五块硬盘创建 RAID5，HDD5 为热备盘，当 HDD1 故障后，HDD0、HDD2、HDD3、HDD4 通过异或算法将重构的数据写入 HDD5 中；
- 在右图的 RAID2.0+示意图中，当 HDD1 故障后，故障盘 HDD1 中的数据按照 CK 的粒度进行重构，只重构已分配使用的 CK（图中 HDD1 的CK12和CK13），存储池中所有的硬盘都参与重构过程，重构的数据分布在多块硬盘中（图中的 HDD4 和 HDD9）



重构速率的提升还得益于 RAID2.0+技术对故障的处理更加精细有效，RAID2.0+在原有坏道修复和全盘失效重构两级故障修复之间增加了数据块的故障修复，能够基于块（CK）的粒度只重构已分配并使用了的空间，通过对实际使用空间的有效识别，当硬

盘出现故障时，RAID2.0+能够通过精简重构进一步缩短重构时间，降低数据丢失的风险。

由于 RAID2.0+技术在重构方面的巨大优势，使得 OceanStor 高端存储系统在重构方面与传统阵列相比具有明显的优势，下面是采用传统 RAID 的存储系统与采用 RAID2.0+的 OceanStor 高端存储系统在采用 NL-SAS 高容量磁盘环境中重构 1TB 数据所需时间的对比：



3.1.3 故障自检自愈，保证系统可靠性

OceanStor 高端存储系统针对硬盘采用了多重故障容错设计，具有硬盘在线诊断、DHA（Disk Health Analyzer，硬盘故障诊断与预警）、坏道后台扫描、坏道修复等多种可靠性保障，RAID2.0+技术会根据热备策略自动在硬盘域中预留一定数量的热备空间，用户无需进行设置，当系统自动检测到硬盘上某个区域不可修复的介质错误或整个硬盘发生故障时，系统会自动进行重构，将受影响的数据块数据快速重构到其他硬盘的热备空间中，实现系统的快速自愈合。

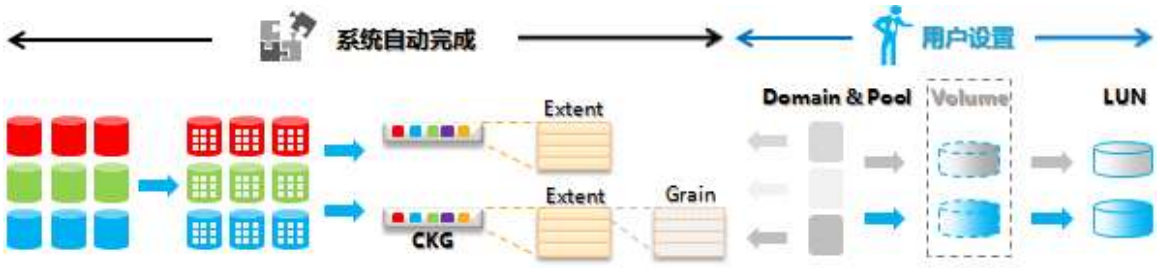
传统 RAID	RAID2.0+
需要手动配置单独的全局或局部热备磁盘	分布式的热备空间，无需单独配置
多对一的重构，重构数据流串行写入单一的热备磁盘	多对多的重构，重构数据流并行写入多块磁盘
存在热点，重构时间长	负载均衡，重构时间短

3.2 弹性高效

3.2.1 虚拟池化设计，降低存储规划管理难度

目前主流的高端存储系统拥有成百上千块不同类型的硬盘已经非常普遍，如果使用传统 RAID 技术，对于管理员来说，意味着不仅需要管理数量众多的 RAID 组，而且需要针对每一个应用，对每一个 RAID 组进行周密的性能、容量规划，在当今这样一个变化迅速的时代，要作到准确预估 IT 系统生命周期内业务的发展趋势以及与之对应的数据增长量级几乎是一项不可能实现的目标，这使得管理员不得不经常面临存储资源分配不均等一系列管理问题，大大增加了管理的复杂度。

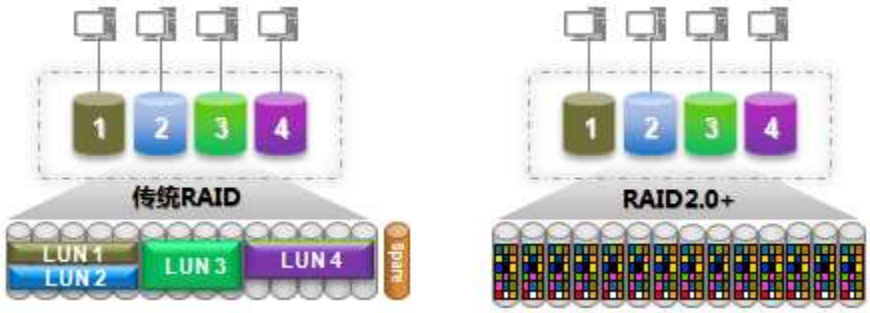
使用 RAID2.0+技术的 OceanStor 高端存储系统，采用了领先的虚拟化技术，对存储资源进行池化设计，管理员只需要维护少量的存储资源池，所有的 RAID 配置在创建存储池时自动配置完成，同时，系统会自动根据制定的策略来智能管理和调度系统资源，大大降低了规划和管理的难度。



3.2.2 增加 LUN 所跨硬盘数，大幅提升单 LUN 性能

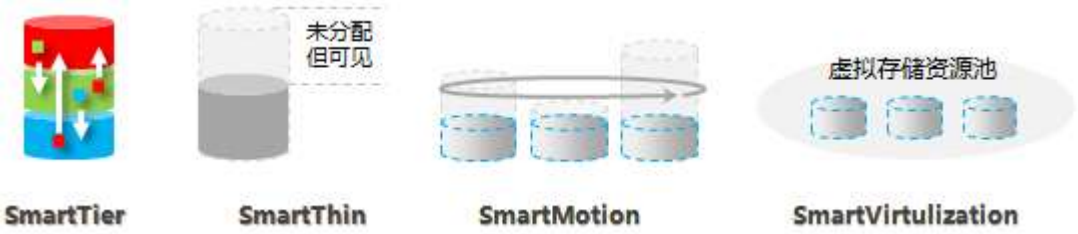
进入 21 世纪之后，服务器计算能力的不断发展和越来越多的主机应用（数据库、虚拟机等）对存储的性能、容量、灵活性都提出了更高的要求，传统 RAID 组受到硬盘数的限制，容量小、性能差且难以扩展，已经越来越无法满足业务的需求。当主机对一个 LUN 进行密集访问时，只能访问到有限的几个磁盘，容易造成磁盘访问瓶颈，导致磁盘热点。

RAID2.0+技术支持由几十甚至上百块硬盘组成一个大的存储资源池，LUN 基于存储池创建，不再受限于 RAID 组磁盘数量，宽条带化技术能够让单个 LUN 上的数据分布到很多不同的磁盘上，避免了磁盘热点，使得单 LUN 性能和容量都得到了大幅提升。如果当前存储的容量无法满足要求时，只需要简单向硬盘域中增加硬盘就可以完成存储池和 LUN 的动态扩容，提升了磁盘的容量利用率。



3.2.3 空间动态分布，灵活适应业务变化

RAID2.0+基于业界领先的块虚拟化技术实现，卷上的数据和业务负荷会自动均匀分布到存储池所有的物理硬盘上，借助于智能的 Smart 系列效率提升套件，OceanStor 高端存储系统能自动根据业务所需的性能、容量、冷热数据等因素在后台进行智能调配，灵活地适应企业业务的快速变化。



4 附录 A：RAID2.0+ FAQ

Q1、OceanStor 高端存储系统的所有硬盘都位于一个存储池吗？

A1：不一定，OceanStor 高端存储系统基于硬盘域（Disk Domain）和存储池（Storage Pool）来对所有硬盘进行管理，用户可以创建一个或多个硬盘域，不同硬盘域之间的资源、性能和故障是完全隔离的，在硬盘域之上，可以创建一个或多个存储池，每个存储池由各种硬盘提供存储空间资源。

Q2、OceanStor 高端存储系统在硬盘扩容和硬盘故障时，数据是如何变化的？

A2：在新硬盘添加到存储池中时，OceanStor 高端存储系统会自动将部分数据根据硬盘的空间使用情况移动到新增的空间中去，做到容量均衡，保证同一个存储池中的每个硬盘空间利用率大致相当。

当硬盘故障时，与故障硬盘相关的 CKG 会自动进行重构，重构的数据会自动均衡写入其他正常硬盘的热备空间中。热备空间不需要用户指定，由系统根据硬盘的使用情况自动选择。具体重构过程可参考 3.1.2 小节。

Q3、RAID2.0+技术与传统 RAID 相比，可靠性高体现在哪些方面？

A3：RAID2.0+的可靠性提升主要体现在以下几个方面：

- **负荷分担**：RAID2.0+使得硬盘更加均衡地工作，避免了传统 RAID 可能出现的硬盘“过劳死”问题。详见 3.1.1 小节描述。
- **稳健重构**：RAID2.0+技术使得发生重构时有更多的硬盘来分担重构负荷，减少了每块硬盘承担的重构工作量，大大降低了重构期间再发生硬盘故障的风险。
- **快速重构**：RAID2.0+大大减少了重构的时间窗，使得系统能在尽可能短的时间内恢复到容错状态，从而提升系统的可靠性。详见 3.1.2 小节描述。
- **精简重构**：RAID2.0+能够通过元数据感知已分配空间中哪些是已使用的，因此在重构时仅重构已使用空间，减少了重构数据量，进一步缩短了重构时间，降低了重构风险。
- **自检自愈**：RAID2.0+采用分布式的热备空间，当系统检测到故障时，只要硬盘中有空闲的空间（CK），即可自动启动重构，在提升可靠性的同时大大降低了管理成本。详见 3.1.3 小节描述。

✧ **失效数据量**：传统 RAID 失效后，影响的是 RAID 组上的所有数据；而 RAID2.0+发生多盘失效后，只有和多块失效硬盘都相关联的数据才会失效，大部分的数据仍然可以访问，失效数据量与传统 RAID 相比按数量级减少。

下表是基于 Markov 模型，综合考虑**数据丢失概率**和**丢失数据量**得出的两种技术的**数据丢失风险**：

系统配置及参数	RAID2.0+配置	传统 RAID 配置	数据丢失风险（传统 RAID/RAID2.0+）
40*600GB SAS 盘（硬盘失效率 1%）	RAID5(4+1)/(40 盘/DG)	8 组 RAID5(4+1)	**16.09**
40*2TB SATA 盘（硬盘失效率 2%）	RAID6(8+2)/(40 盘/DG)	4 组 RAID6(8+2)	**69.29**
40*600GB SAS 盘（硬盘失效率 1%）	RAID10(10 盘)/(40 盘/DG)	4 组 RAID10(10 盘)	**39.15**

Q4、OceanStor 高端存储系统的硬盘利用率很低吗？

A4：OceanStor 高端存储系统采用了基于 RAID2.0+的两层虚拟化软件架构，为了便于对数据进行灵活高效的管理，预留了一定的容量，然而这一点常常被错误理解为硬盘利用率不高，而忽略了针对对用户有实际意义的存储效率。正是借助于该独特创新的架构，华为存储实现了业界最高的存储效率。比如，借助于以数据块为单元的资源管理，实现了全局负载均衡和快速重构，重构速率提升 20 倍；上层虚拟化通过 Smart 系列软件实现了系统资源的智能分配，大幅提升资源利用率。

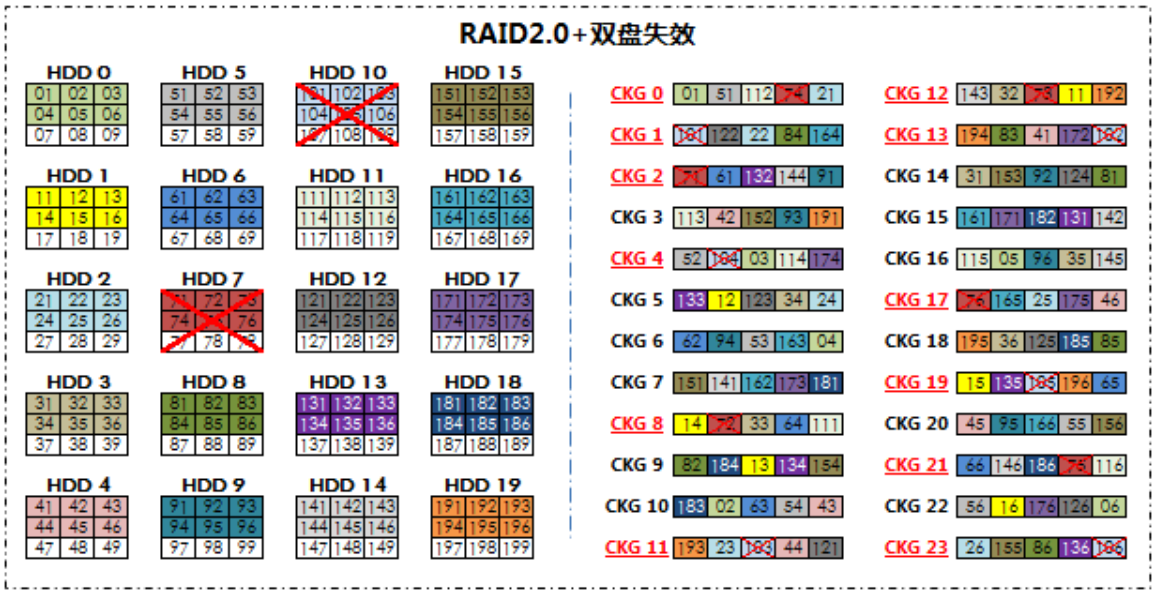
以 SAS 硬盘采用 RAID5（8D+1P）的 RAID 策略为例，其容量利用率约为 83.42%，仅比传统 RAID 的 88.89%低 5%。

Q5、OceanStor 高端存储系统在双盘失效时数据是否会丢失？

A5：RAID 技术是构成存储数据保护的基础，讨论双盘失效的问题，其本质仍然在于 RAID 的容错能力：对于 RAID5 来讲，其可容错的故障数为 1（对于传统 RAID 来说，单位为硬盘，对于 RAID2.0+来说，单位为块）；对于 RAID6 来讲，可容错的故障数为 2。因此，若采用 RAID6 等双校验的保护类型，无论是传统 RAID 还是基于块虚拟化的 RAID2.0+，在双盘失效时数据都是不会丢失。

若采用 RAID5，对于传统 RAID 来说，双盘失效一定会导致数据丢失，而采用 RAID2.0+技术的 OceanStor 高端存储系统，只要双盘失效时每个 CKG 中不会同时出现两个失效的块（CK），那么数据是不会丢失的。

如下图是一个由 20 块硬盘组成的存储池，以 LUN 的形式对上层主机提供存储空间，RAID 策略为 RAID5（4D+1P）。



当 HDD 7 硬盘和 HDD 10 硬盘同时故障时，受影响的只是与 HDD 7 和 HDD 10 硬盘相关联的 CK，分别为 CK71～CK76，CK101～CK106（CK77～CK79 和 CK107～CK109 为空闲 CK，无数据不受影响）；对应到 CKG 中，分别为 CKG 0、CKG 1、CKG 2、CKG 4、CKG 8、CKG 11、CKG 12、CKG 13、CKG 17、CKG 19、CKG 21 和 CKG 23（红色下划线标识），由于 CKG 采用 RAID5（4D+1P）的保护策略，而每个受影响的 CKG 中都只有 1 个 CK 失效，因此整个 CKG 的数据仍然可用，从主机层面看，对应的 LUN 仍然是可以正常访问的，业务也不会中断。

5 附录 B：RAID2.0+周边资源

1、 **RAID2.0+特性多媒体**，直观的动画让您轻松了解 RAID2.0+的优势和原理！

http://3ms.huawei.com/mm/video/videoMaintain.do?method=showVideoDetail&f_id=1421965

2、 **RAID2.0+全景图及重点概念**

http://3ms.huawei.com/mm/docMaintain/mmMaintain.do?method=showMMDetail&f_id=STR13072905120079

6 缩略语表/Acronyms and Abbreviations

表6-1 缩略语清单

英文缩写	英文全称	中文全称
RAID	Redundant Array of Independent Disk	独立冗余磁盘阵列
RPM	Revolutions Per Minute	每分钟转速
LUN	Logical Unit Number	逻辑单元号
RAID	Redundant Array of Independent Disks	独立磁盘冗余阵列
XVE	eXtreme Virtual Engine	XVE
CK	Chunk	Chunk
CKG	Chunk Group	Chunk Group
DG	Disk Group	硬盘组
LD	Logical Drive	逻辑磁盘
SNIA	Storage Networking Industry Association	存储网络工业协会
AFR	Annual Failure Rate	年故障率
DHA	Disk Health Analyzer	硬盘故障诊断与预警