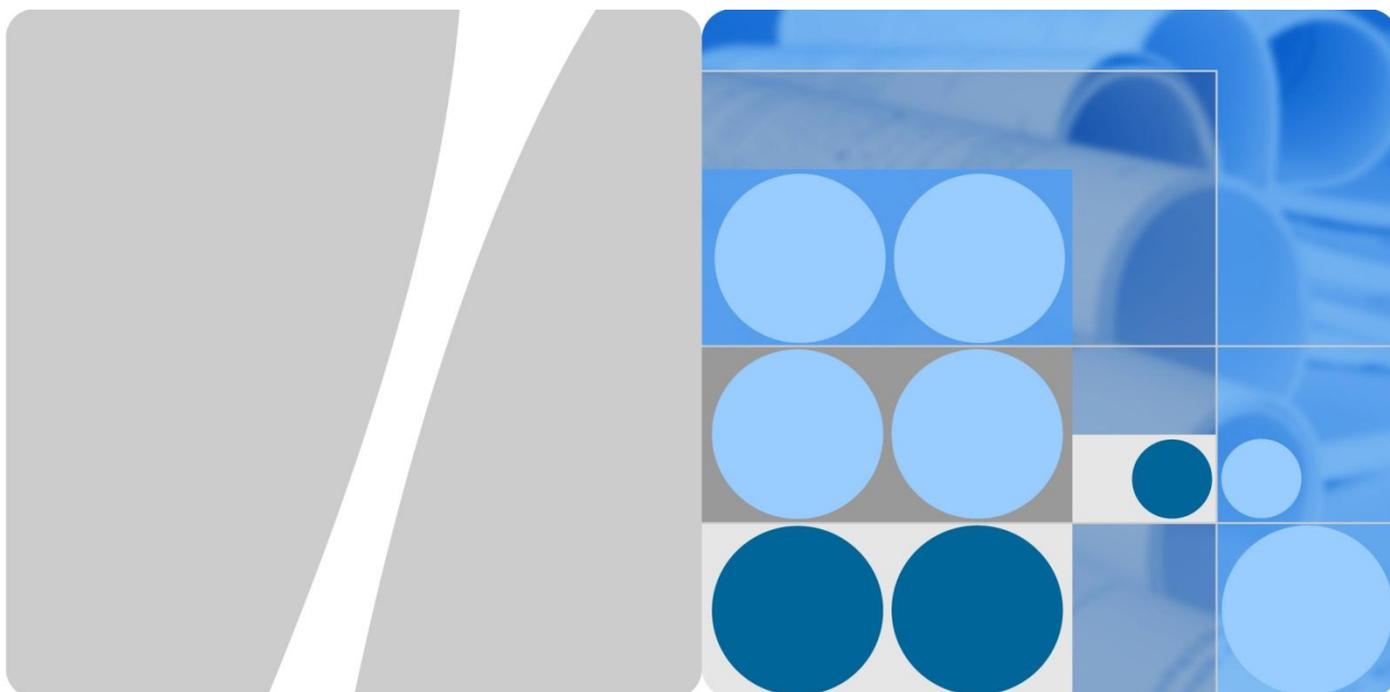


资料编码



# OceanStor VTL6000 产品白皮书

文档版本 V2.2  
发布日期 2012-5-5

华为技术有限公司



**版权所有 © 华为技术有限公司 2009。 保留一切权利。**

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本档内容的部分或全部，并不得以任何形式传播。

## 商标声明



**HUAWEI** 和其他华为商标均为华为技术有限公司的商标。

本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，本档内容会不定期进行更新。除非另有约定，本档仅作为使用指导，本档中的所有陈述、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](mailto:support@huawei.com)

客户服务电话： 0755-28560000 4008302118

客户服务传真： 0755-28560111

---

# 目 录

---

<b>1 执行摘要/Executive Summary .....</b>	<b>5</b>
<b>2 简介/Introduction .....</b>	<b>6</b>
<b>3 解决方案/Solution .....</b>	<b>7</b>
3.1 重复数据删除 .....	7
3.2 HA 集群 .....	10
3.3 节能降耗 .....	13
<b>4 推广/Experience .....</b>	<b>15</b>
4.1 典型应用和客户价值 .....	15
<b>5 结论/Conclusion .....</b>	<b>19</b>
<b>6 缩略语表/Acronyms and Abbreviations .....</b>	<b>20</b>

### 修订记录/Change History

日期	修订版本	描述	作者
2010-2-23	1.2	初次发布，V1.0 和 V1.1 已废弃。	汪野
2010-3-8	1.3	根据评审意见修改了 4.1.1 小节客户价值部分的内容： 将高性能大容量部分的内容分为高性能、大容量两个部分分别进行说明	汪野
2010-4-13	2.1	修改了 SIR 引擎配置描述： 配置 2 个 SIR 引擎时，它们为 HP 集群。此前的说法为它们各自独立，该说法有误。	汪野
2012-5-5	2.2	修改了部分参数	朱峰

# 1 执行摘要/Executive Summary

---

本文首先说明了中高端用户在选择 VTL 产品时面临的诸多难题，然后通过对 VTL6000 虚拟磁带库产品的关键技术特性、应用场景、解决方案、客户价值进行说明和分析，最终得出 VTL6000 可以很好地解决中高端用户面临的问题这一结论。

# 2 简介/Introduction

---

随着数据的爆炸式增长，传统磁带备份系统的备份恢复速率已无法满足用户的备份需求。随着磁盘技术的快速发展，磁盘介质的容量得到了很大的提高，单位容量磁盘存储的价格也大幅下跌，基于 SATA 磁盘的备份系统已经在用户 IT 环境中大量部署，并以其高备份恢复性能而逐渐为用户所青睐，VTL 就是其中具有代表性的产品。VTL 产品兼有磁盘设备的高性能、易维护和磁带设备先进成熟的介质管理等优点，一经推出，即获得了良好的发展和成长，市场前景极为广阔。

对那些中高端用户而言，其一次完全备份数据量即可达数十 TB，以 8 小时备份窗口计，其对备份系统性能需求至少为 1000MB/s。而且，以每天一次增量备份、每周一次完全备份、保存周期为三个月计算，备份设备存储容量需求高达几百 TB，要构建、运行和管理如此规模的磁盘备份系统，用户需要投入巨大的存储成本、能耗成本以及管理成本。此外，除生产系统外，中高端用户对备份系统的可用性同样有较高要求，一般的单引擎 VTL 产品显然无法满足其高可用需求。

可见，中高端用户选择 VTL 产品面临的几个主要问题就是：性能低下、容量不足、能耗过高以及不支持备份系统高可用。

# 3 解决方案/Solution

VTL6000 是一款面向中高端用户的虚拟磁带库产品，它支持重复数据删除、HA 集群以及磁盘休眠技术，以此作为解决中高端用户面临的诸多问题的切入点。

本章介绍 VTL6000 的重复数据删除、HA 集群以及节能降耗三个特性，其他如磁带缓冲、远程复制等 VTL 高级特性的介绍请参见《VTL3500 产品白皮书》，这里不赘述。

## 3.1 重复数据删除

### 3.1.1 重复数据删除简介

重复数据删除 (Deduplication) 技术，简单来说，就是一种消除重复数据的技术，它用软件或硬件的方式，对存储数据进行处理，以消除其中的重复数据，从而减小数据占用的存储空间。

备份系统中，数据在备份服务器的控制下，从备份客户端（源端）传输到备份设备（目标端），根据重复数据删除操作发生的位置，可以将其分为源端重复数据删除 (Deduplication at the Source) 和目标端重复数据删除 (Deduplication at the Target)。源端重复数据删除指的是：备份客户端将数据传输给备份设备之前，首先对这些数据进行处理，以删除其中的重复数据，然后将删除得到的没有重复的数据发送给备份设备。目标端重复数据删除指的是：数据从备份客户端发送时，并没有经过重复数据删除，只有在到达备份设备时，才由备份设备来执行重复数据删除。

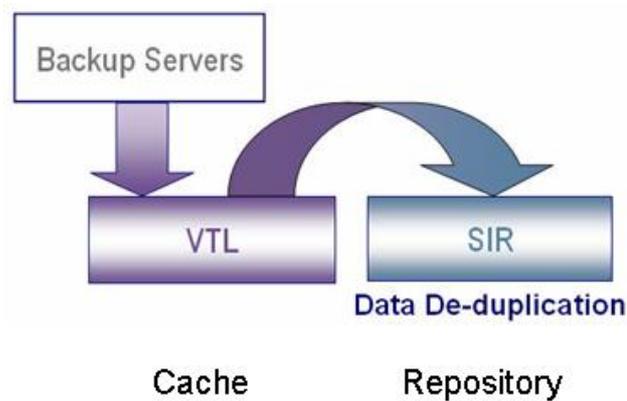
在目标端重复数据删除中，根据重复数据删除操作发生的时间，又可以将目标端重复数据删除分为内处理重复数据删除 (Inline Deduplication) 和后处理重复数据删除 (Post processing Deduplication，业界又有 Offline Deduplication 的叫法，即 offline 重复数据删除，其意义相同)。内处理重复数据删除指的是：备份设备在接收备份数据的同时执行重复数据删除操作，即一边接收数据一边做重复删除操作，备份数据接收完成则重复数据删除操作也执行完成。后处理重复数据删除指的是：备份结束后备份设备才开始执行重复数据删除操作，即备份设备首先接收完所有的备份数据，然后在某一时刻才开始对备份数据执行重复数据删除。

无论哪种重复数据删除技术，其终极目的都是消除重复数据，因而都会涉及将新数据和已有数据进行比较以判断新数据是否重复的过程。识别重复数据的方法主要有两类，一类是基于索引的比较方法，一类是基于内容的比较方法，前者通过比较数据的索引以识别重复数据，后者则直接比较数据本身以识别重复数据。数据的索引指的是，系统将数据划分为定长或不定长的数据块，然后以每个数据块为输入，使用一定的算法计算出一个唯一的值，该值即为该数据块的索引。基于索引的比较方法中，由于索引存储空间远小于其对应数据块的存储空间，因而索引的比较操作可以直接在内存中执行，其效率显著高于基于内容的比较方法。当前，基于索引的比较方法已经在众多重复数据删除技术中得到应用。

### 3.1.2 VTL6000 重复数据删除原理

VTL6000 的重复数据删除技术为 SIR (Single Instance Repository)，它是一种 offline 重复数据删除技术，其数据比较方法为索引比较方法。

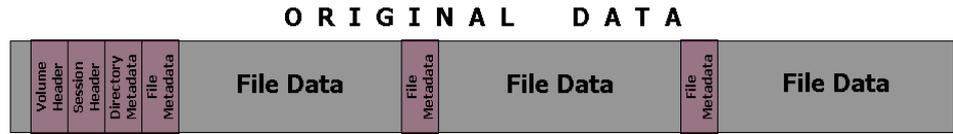
图3-1 SIR 技术原理



VTL6000 中的 VTL 软件包含两个模块：VTL 模块和 SIR 模块。VTL6000 的存储空间逻辑上被划分为两部分，分别为 VTL 模块和 SIR 模块所用，称之为 VTL 存储空间（又称为 Cache）和 SIR 存储空间（又称为 Repository）。VTL6000 接收到备份数据后，首先将其存放于 VTL 存储空间，此后的某个时刻（如某时间点、备份结束、达到存储水位等），SIR 模块会读取这些数据，并将其和 SIR 存储空间中的已有不同数据块进行比较：SIR 模块将原始备份数据划分为大小为若干 KB 的数据块，然后使用 SHA-1 算法为每个数据块计算出一个哈希值（又称为索引/Index），通过比较新数据块和已有数据块的哈希值以确定新数据块是否重复，重复的数据块将被丢弃，而仅保留其数据块指针，全新的数据块才会被存放到 SIR 存储空间，如图 3-1 所示。

VTL6000 中，运行 VTL 软件模块应用的实体物理机以及运行于其上的 VTL 软件模块应用统称为 VTL 引擎，运行 SIR 软件模块应用的实体物理机以及运行于其上的 SIR 软件模块应用统称为 SIR 引擎。VTL 引擎必须配置，以支持对外提供虚拟磁带库服务；SIR 引擎可选配置，用于提供重复数据删除功能。VTL 引擎和 SIR 引擎可以为同一个引擎，也可以为不同的引擎，当选配单独的 SIR 引擎时，VTL6000 最多可以支持 3 个 (2+1) SIR 集群配置。

图3-2 VTL6000 中虚拟磁带原始数据



SIR 引擎执行重复数据删除之前，VTL 存储空间中备份数据以虚拟磁带的形式存放，如图 3-2 所示。重复删除执行后，虚拟磁带中的数据被指针替代（此时的虚拟磁带称为 VIT，Virtual Index Tape），该指针指向 SIR 存储空间中的单一实例数据块，所释放的 VTL 存储空间用于存储新的备份数据，如图 3-3 所示。

图3-3 VTL6000 中重复数据删除后的数据分布

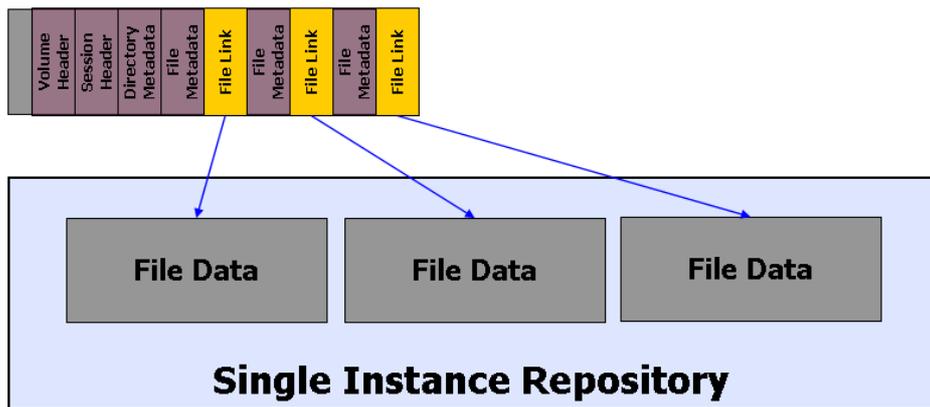
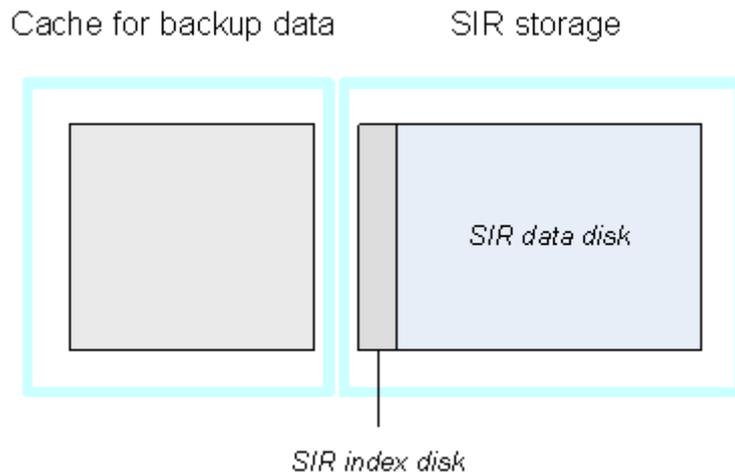


图3-4 VTL6000 存储空间分配情况



前面已经说明，VTL6000 存储空间逻辑上被划分 VTL 存储空间和 SIR 存储空间两部分，即 Cache 和 Repository 两部分，其中，SIR 存储空间又被划分为 SIR data disk 和 SIR index disk 两部分，如图 3-4 所示。SIR data disk 中即存放重复数据删

除后的不同数据块，新备份到 Cache 中的数据即和 SIR data disk 中的数据块进行比较以确定其是否重复。SIR index disk 中则存放 SIR data disk 中所有数据块的索引，即 SHA-1 哈希值，可见，SIR data disk 越大，则 SIR index disk 越大。SIR 模块运行过程中，即将 SIR index disk 中的所有索引（索引表）读取到 SIR 引擎的内存中，快速完成索引表查找，可见，SIR index disk 越大，则 SIR 引擎需要配置的内存容量也越大。

---

更多重复数据删除技术的详细内容请参见《技术白皮书 2012 重复数据删除》。

---

## 3.2 HA 集群

### 3.2.1 HA 集群简介

HA (High Availability) 集群即高可用集群，指的是一旦集群中的某个节点出现故障，导致其不能继续正常工作时，集群中的其他节点会接替故障节点继续工作。

HA 集群中的节点分为活动节点和备份节点。活动节点就是正在执行任务的节点，备份节点为活动节点的备份。一旦活动节点发生故障，则备份节点就开始接替活动节点来执行任务。

HA 集群的实现是基于资源切换的。所谓“资源”，是指 HA 集群中某个节点发生故障后，备份节点所要接管的所有东西的集合，如 IP 地址、主机名、磁盘卷、应用程序上下文等。备份节点接管了故障节点的这些资源后，才能够接替故障节点继续正常工作，从而将对客户端的影响降到最小。

资源的监控和接管依赖于 HA 软件。各种操作系统几乎都自带 HA 软件，众多厂商也各自开发了 HA 软件，无论哪种 HA 软件，其作用都是让备份节点监控活动节点的状态，一旦侦测到故障，则强行接管活动节点的资源并继续提供服务。

注：本小节参考了《大话存储》群集部分的内容。

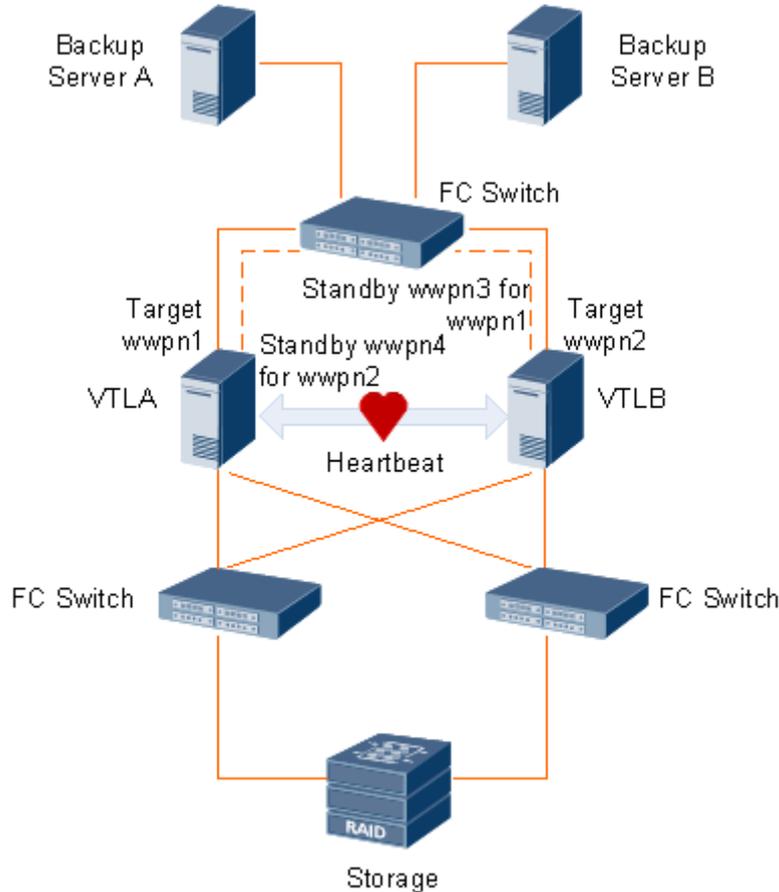
### 3.2.2 VTL6000 高可用特性

VTL6000 中 VTL 软件模块和 SIR 软件模块均支持 HA 特性，即 VTL6000 高可用特性包括 VTL 引擎 HA 配置以及 SIR 引擎 HA 配置，确保 VTL 引擎高可用以及 SIR 引擎高可用。

#### VTL 引擎 HA

VTL6000 支持 2 个 VTL 引擎 HA 配置，两个 VTL 引擎可以为单向活动——备份节点（单向 Failover），也可以互为活动——备份节点（双向 Failover）。

图3-5 VTL 引擎 HA 配置示意图



以图 3-5 为例，当 VTLA 和 VTLB 均正常运行时，二者各自对外提供虚拟磁带库服务。单向 Failover 配置中，假设 VTLA 为活动节点，VTLB 为备份节点，此时，VTLB 可以监控并接管 VTLA：当 VTLA 出现故障时，VTLB 可接替 VTLA 继续对外提供虚拟磁带库服务；当 VTLB 出现故障时，VTLA 不可接替之。双向 Failover 配置中，VTLA 和 VTLB 互相监控对方，并在对方出现故障时接管对方以继续对外提供虚拟磁带库服务。

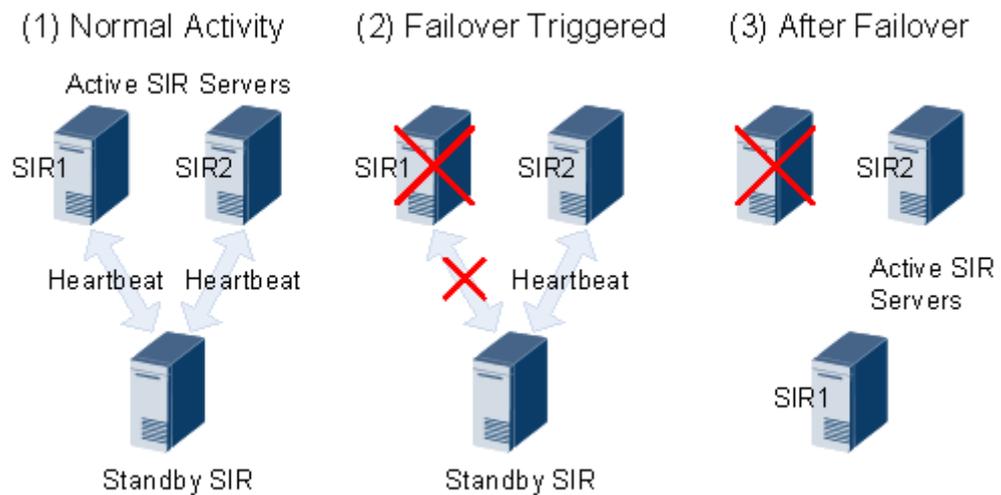
如图 3-5 所示，VTLA 指的是 VTLA 软件应用和 VTLA 物理机的组合，VTLB 同理。所谓“接管”指的是：例如 VTLA 被 VTLB 接管后，原来运行于 VTLA 物理机上的 VTLA 软件应用，现在运行于 VTLB 物理机上。读者不禁要问，接管完成后，运行于 VTLB 物理机上的 VTLA 软件应用又是如何访问其原始存储（VTL 存储空间），并继续为其原始备份服务器提供虚拟磁带库服务的呢？答案是共享存储和 Standby 备份主机接口。共享存储指的是，VTLA 物理机和 VTLB 物理机都有物理通道连接 VTL6000 的存储单元，当 VTLA 软件应用运行于 VTLB 物理机上时，它仍然可以通过 VTLB 物理机上连接存储单元的物理通道访问其原始存储。Standby 备份主机接口指的是，以 FC 通道为例，VTLA 通过 FC 端口 Target wwpn1 对外提供虚拟磁带库服务，相应地，在 VTLB 上就必然有一个 FC 端口 wwpn3 作为 wwpn1 的备份端口，实际配置时，这两个 FC 端口都有 FC 通道连接到 VTLA 的备份服务器，当 VTLA 软件应用运行于 VTLB 物理机上时，它仍然可以通过 Standby wwpn3 继续为其原始备份服务器提供虚拟磁带库服务。

以双向 Failover 为例，VTLA 和 VTLB 的软件应用通过心跳网络监控对方的状态，一旦发现对方的软件、硬件以及连接存储单元的通道出现故障（影响到其继续对外提供正常的虚拟磁带库服务），则自动触发 Failover，即由监控节点接管故障节点，接管过程持续约 4 分钟时间。当故障节点的故障得到排除，并可以正常工作时，则自动（或通过手动的方式）触发 Failback，即当前接管节点自动（或通过手动的方式）放弃接管。

## SIR 引擎 HA

VTL6000 支持 3 个 SIR 引擎 HA 配置，其中 2 个 SIR 引擎为活动节点，另外 1 个 SIR 引擎为备份节点。

图3-6 SIR 引擎 HA 配置示意图



以图 3-6 为例，SIR 引擎的 HA 配置和 VTL 引擎的 HA 配置稍有不同，它拥有 1 个专职备份的 SIR 引擎，称为 Standby SIR 引擎，其上的 SIR 软件应用通过心跳网络监控活动 SIR 引擎的状态，当 SIR1 出现故障时（SIR2 出现故障时情形类似），如软件、硬件以及连接存储（SIR 存储空间）的物理通道出现故障（影响到其继续提供正常的重复数据删除服务），则自动触发 Failover，此时 Standby SIR 通过 IPMI 指令将 SIR1 下电，然后接管 SIR1。这里的接管指的是，SIR1 软件应用将运行于 Standby SIR 物理机上。此后，当故障 SIR 引擎修复时，其上电后将自动成为 Standby SIR 引擎。

SIR 引擎 HA 配置中，各 SIR 引擎都有物理通道以访问所有的 SIR 存储空间，包括访问其各自的存储单元以及其他 SIR 引擎的存储单元，从而确保各 SIR 软件应用无论运行于哪一个 SIR 物理机上，它们都能访问其各自的存储单元。另外，各 SIR 引擎都有物理通道和 VTL 引擎互联（SIR 引擎读取 VTL 存储空间中的原始备份数据执行重复数据删除，VTL 引擎读取 SIR 存储空间中的数据块执行恢复）。总体来看，各个 SIR 引擎的物理连接是完全相同的，从而支持活动节点和备份节点的任意切换。

顺便一提，SIR 软件模块也支持 HP（High Performance）特性，VTL6000 的 3 个（2+1）SIR 引擎 HA 配置中，实际工作时，总是 2 个活动 SIR 引擎在执行重复数

据删除，它们组成的是一个 HP 集群，可以提供更高的重复数据删除性能和效率。当然，VTL6000 也支持配置 1 个或 2 个 SIR 引擎，此时不提供 SIR 引擎 HA 特性。配置 2 个 SIR 引擎时，它们组成 HP 集群。

## 3.3 节能降耗

据国际能源机构(IEA)2008 年的统计，从 1973 年到 2006 年，全球能源消耗上升 73%。从 70 年代以来，能源价格一路上涨，给工业、制造业带来的成本压力一直上升，全社会比以往更重视节能。在全社会大力倡导绿色节能的大背景下，存储产品的节能效应也日益走进广大用户的视野。当前，很多机构和企业已经将节能特性作为采购存储产品的一个必要考虑条件，节能特性在一定程度上直接决定了产品的成败。

### 3.3.1 重复数据删除节能

重复数据删除技术可以大幅降低用户存储投资和管理成本，此外，重复数据删除还可以显著降低存储系统能耗。在需要存储的数据量相同的情况下，借助重复数据删除技术，用户可以用更小的存储容量存放更多的数据，在存储密度相同的情况下，更小的存储容量就意味着更少的存储设备，更少的存储设备也就意味着更低的能耗。下面将使用一个具体的例子，对 VTL6000 中重复数据删除技术的节能效应进行说明。

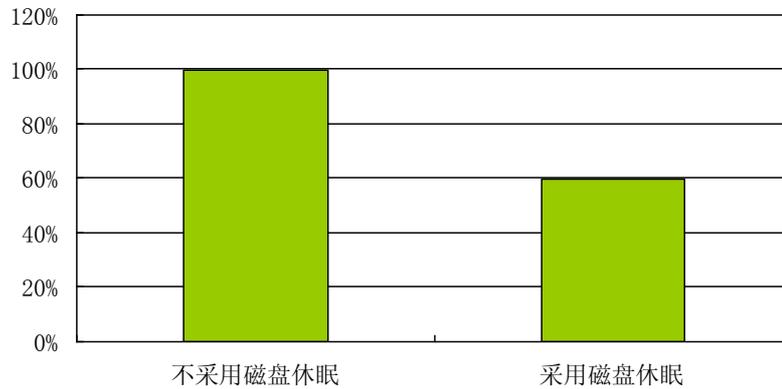
用户环境一次完全备份的数据总量为 10TB，每日增量数据为 50GB，用户要求的备份策略为：每天一次增量备份，每周一次完全备份，备份数据保留周期为 6 个月。以每月 30 天计算，备份系统共需存储 25 个全备数据集和 155 个增备数据集。对于不支持重复数据删除技术的存储设备，共需提供 362.75TB 存储容量 ( $10*25+(25*24*0.05*7)/2+155*0.05$ )。对于支持重复数据删除技术的 VTL6000，共需提供 40TB 存储容量 (20TB 为 VTL 存储空间，20TB 为 SIR 存储空间，SIR 存储空间中的实际存储数据量为 18.95TB( $10+179*0.05$ ))。如果采用 2TB 的 SATA 磁盘存储这些数据，则前者需要提供约 200 块磁盘，VTL6000 则仅需要提供约 30 块磁盘 (考虑到 RAID6 和热备所需磁盘)，两者数量之比约为 7:1。此外，相比 VTL6000，不支持重复数据删除技术的存储设备还要配置更多磁盘框。此种情形下，VTL6000 存储单元消耗的电能约为不支持重复数据删除技术存储设备的十分之一。考虑到 VTL6000 存储单元功耗约为 VTL6000 整机功耗的 40%，则通过重复数据删除技术，VTL6000 可节省约 36% 的整机功耗。

### 3.3.2 磁盘休眠节能

除了重复数据删除之外，磁盘休眠也可以显著降低存储系统能耗。磁盘休眠的工作原理是使长时间无访问请求的磁盘处于休眠甚至下电状态，以达到绿色节能和延长磁盘寿命的目的。在使用磁盘休眠技术的存储设备中，没有读写操作的磁盘处于休眠状态，有读写操作的磁盘才处于正常运行状态。对于那些处于休眠状态的磁盘，一旦有对它们的读写操作，它们就会被唤醒从而进入正常运行状态；任何处于正常运行状态的磁盘，在读写操作结束后，它们会在相应的条件下转入不同级别的休眠状态。因此，磁盘休眠技术非常适用于访问频率低、数据即时可得性要求低的大容量近线存储设备和分级存储设备，尤其适用在以数据恢复为目的的数据备份和数据归档设备，比如 VTL。对支持磁盘休眠技术的 VTL 而言，在

每天若干小时的备份窗口时间内，极端情况下 VTL 中的磁盘全部处于正常运行状态；在当天剩余的绝大部分时间中，VTL 中的所有磁盘都将处于休眠状态。

图3-7 VTL6000 存储单元磁盘休眠效果



VTL6000 支持磁盘休眠，具体表现为：VTL6000 存储单元中的磁盘被划分为若干个 RAID 组，用户可根据需要为每个 RAID 组设置不同的磁盘休眠策略，如定时休眠、某时间段休眠等，已设置磁盘休眠 RAID 组中的磁盘在没有读写访问时可依据策略休眠甚至下电，此后，备份软件（如 NetBackup）一旦向已休眠 RAID 组中的 LUN 发出读写请求，即触发这些 RAID 组中已下电的磁盘自动上电，RAID 组即可重新支持正常的读写访问。通过磁盘休眠技术，VTL6000 存储单元可节约 40% 的能耗。考虑到 VTL6000 存储单元功耗占约为 VTL6000 整机功耗的 40%，则通过磁盘休眠技术，VTL6000 可节省约 16% 的整机功耗。

# 4 推广/Experience

VTL6000 支持 VTL/SIR 引擎集群配置，备份性能高达 2500MB/s，存储裸容量接近 384TB，且支持重复数据删除和磁盘休眠，能够很好地满足中高端用户对 VTL 产品高性能、大容量、节能降耗以及高可用的需求。

本章针对 VTL 备份系统、远程备份系统两个典型应用场景，对 VTL6000 解决方案和客户价值进行了分析和说明，其他如分级备份、远程容灾等应用场景、解决方案以及客户价值的介绍请参见《VTL3500 产品白皮书》，这里不赘述。

## 4.1 典型应用和客户价值

### 4.1.1 VTL 备份系统

#### 应用场景

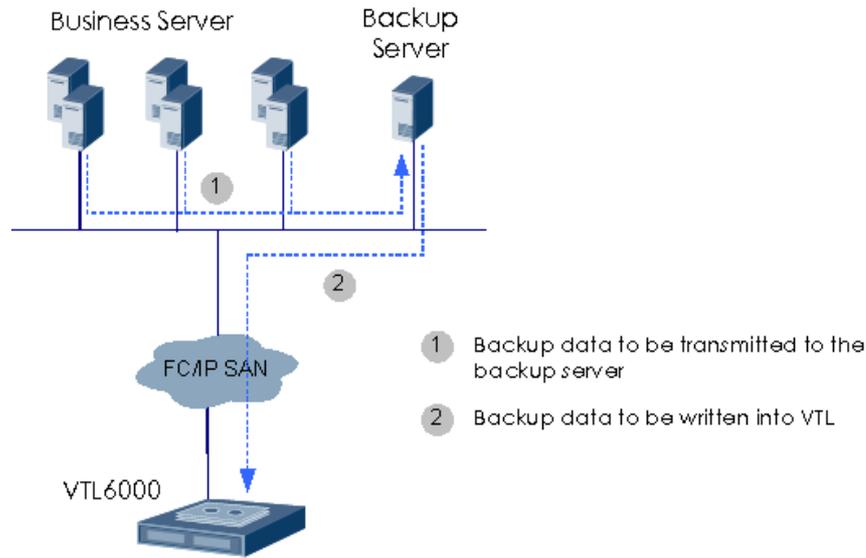
VTL 备份系统应用场景为如下多种情形的任意一种：

- 新建备份系统。此前用户应用环境并无备份系统，或者备份系统需要重新改造，致使客户需要重新选定备份设备。
- 使用 VTL6000 替换用户备份环境中已有的物理带库。此前客户已有基于物理带库的备份系统，由于性能、可靠性、管理等原因，用户需要选择一款新的备份设备替换原有物理带库。

#### 解决方案

使用 VTL6000 作为备份设备，应用组网如图 4-1 所示。

图4-1 VTL 备份系统



其中，VTL6000 通过 FC SAN 或者 IP SAN 连接备份服务器。

## 客户价值

VTL 备份系统客户价值如下：

- 高性能，满足用户对备份窗口的需求。单 VTL 引擎配置时，VTL6000 实测备份性能可达 1250MB/s，8 小时内可支持备份 36TB 数据；双引擎配置时，VTL6000 实测备份性能高达 2500MB/s，8 小时内可支持备份 72TB 数据。对于一次备份数据量不超过 72TB 的用户，VTL6000 可以很好地满足其对备份窗口的要求。
- 大容量，满足用户对存储容量的需求。双 VTL 引擎配置时，VTL6000 最大支持 384TB 裸容量（可用容量为 320TB），提供重复数据删除功能时，VTL6000 最大可提供 152TB（其中 VTL 存储空间可用容量为 160TB，SIR index disk 可用容量为 8TB，SIR data disk 可用容量为 152TB）存储容量用于存放重复数据删除后的数据，以 20:1 重复数据删除比计算，VTL6000 可存放 3.0PB 备份数据，满足中高端用户对备份容量的需求。
- 重复数据删除，极大节约系统能耗和存储投资。VTL6000 支持重复数据删除和磁盘休眠，可显著减少备份系统需要配置的磁盘存储，极大节省系统能耗和存储投资。重复数据删除比例以及节能效应和备份数据类型、备份策略密切相关，以 3.3 小节内容为例，相比不支持重复数据删除和磁盘休眠的 VTL 产品，VTL6000 可以为用户节省 50% 以上的备份设备整机能耗。
- HA 集群配置，满足用户高可用需求。双 VTL 引擎 HA 集群配置下，VTL6000 的两个 VTL 引擎一般都独立工作，当任何一个 VTL 引擎出现故障导致其无法支持备份业务时，另一个引擎将自动接管该故障引擎，接管时间约 4 分钟，即 4 分钟后，VTL6000 又恢复正常工作，满足用户对备份系统的高可用需求。3 个（2+1）SIR 引擎 HA 集群配置下，当 VTL6000 任何一个 SIR 引擎出现故障导致其无法支持重复数据删除业务时，Standby SIR 引擎将自动接管该

故障引擎，接管时间约 4 分钟，即 4 分钟后，VTL6000 又恢复正常工作，满足用户对备份系统的高可用需求。

## 4.1.2 远程备份系统

### 应用场景

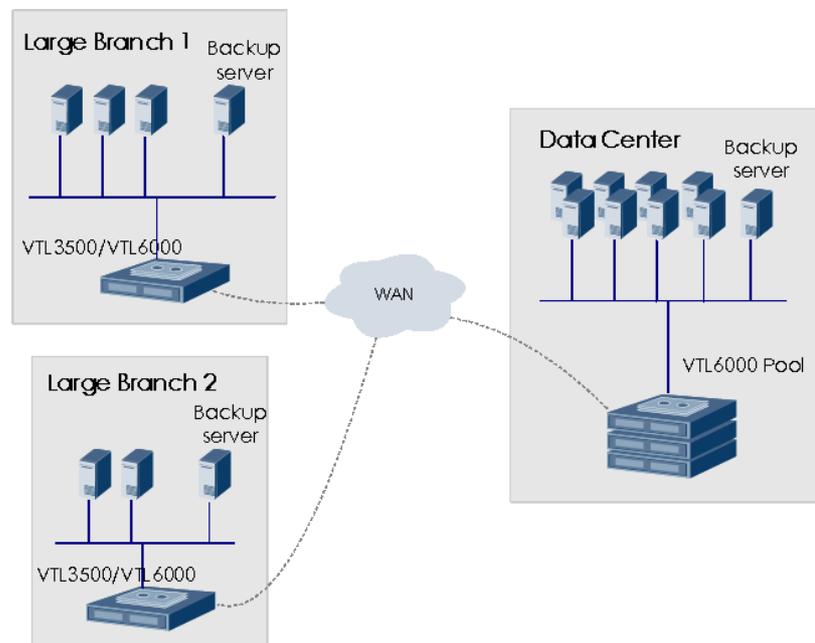
远程备份系统应用场景为如下多种情形的任意一种：

- 用户网络带宽资源不充足。除数据中心之外，用户应用环境中还有多个远程分支节点的数据需要备份。此前客户应用环境并无备份系统，或者已有备份系统，但用户分支节点和数据中心之间的实际可用带宽资源不充足，导致分支节点数据备份性能低下，无法满足用户对分支节点备份窗口的要求。

### 解决方案

解决方案如下，组网图如图 4-2 所示。

图4-2 远程备份系统



- 在用户分支节点部署支持重复数据删除功能的 VTL3500（参见 VTL3500 产品描述）或者 VTL6000 All In One（参见 VTL6000 产品描述），在用户数据中心部署提供重复数据删除功能的 VTL6000 Multi-Nodes（参见 VTL6000 产品描述）。
- 分支节点的数据直接备份到其本地 VTL3500 或 VTL6000，然后通过基于重复数据删除的远程复制功能复制到数据中心 VTL6000；数据中心的数据则直接备份到其本地 VTL6000。

## 客户价值

远程备份系统客户价值如下：

- 本地 VTL 备份，满足用户对分支节点备份窗口的要求。分支节点数据直接备份至本地 VTL3500/VTL6000，此为 VTL 备份系统应用场景，其固有客户价值见 4.1.1 小节内容。借助 VTL 的高性能特性（VTL3500 备份速率为 400MB/s，VTL6000 All In One 备份速率为 1250MB/s），即使分支节点备份数据量高达 10TB，备份时间也不会超过 3 小时，满足用户对分支节点备份窗口的要求。
- 基于重复数据删除的远程复制，显著降低用户网络带宽需求，节约网络带宽投资。分支节点的备份数据首先在本地 VTL 中完成重复数据删除后，再通过 WAN 复制到数据中心 VTL，复制过程中仅传输重复数据删除后，且在数据中心不存在的那些数据块，相比直接传输未经重复数据删除的备份数据，此举可大幅节约复制带宽，显著降低用户网络带宽需求，节约网络带宽投资。
- 全局重复数据删除，进一步降低用户存储投资。VTL6000 支持全局重复数据删除功能，即复制过程中仅传输重复数据删除后的、在数据中心不存在的那些数据块，因此各个分支节点 VTL 之间的，以及各个分支节点 VTL 和数据中心 VTL 之间的重复数据均可得到消除，相比 VTL 备份系统应用场景（见 4.1.1 小节内容）中的重复数据删除功能，全局重复数据删除的重复数据消除效率更高，进一步降低用户存储投资。

# 5 结论/Conclusion

---

VTL6000 支持重复数据删除和磁盘休眠，且支持 VTL/SIR 引擎 HA 集群配置，可以很好地解决中高端用户面临的备份系统性能低下、容量不足、能耗过高以及不支持高可用等诸多难题：

- 对于一次备份数据量不超过 72TB 的用户，VTL6000 可以很好地满足其对备份窗口的要求。
- 以 20:1 重复数据删除比计算，VTL6000 可支持存放 3.0PB 备份数据，满足用户对备份容量的需求。
- 通过重复数据删除和磁盘休眠技术，VTL6000 可以为用户节省 50% 以上的能耗成本。
- VTL6000 支持 VTL 引擎和 SIR 引擎 HA 集群配置，故障切换时间约 4 分钟，满足用户对备份系统的高可用需求。

# 6 缩略语表/Acronyms and Abbreviations

表6-1 VTL6000 产品缩略语清单

英文缩写	英文全称	中文全称
VTL	Virtual Tape Library	虚拟磁带库
SATA	Serial Advanced Technology Attachment	串行 ATA
HA	High Availability	高可用性
SIR	Single Instance Repository	单一实例库
Cache	Cache	缓存
FC	Fibre Channel	光纤通道
IPMI	Intelligent Power Management Interface	智能电源管理接口
LUN	Logical Unit Number	逻辑单元号
RAID	Redundant Array of Independent Disks	独立磁盘冗余阵列