

OceanStor N8000 集群 NAS 存储系统技术白皮书

Issue V2.1
Date 2012-07-31

华为技术有限公司

版权所有 © 华为技术有限公司 2012。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



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。

华为技术有限公司为客户提供全方位的技术支持，用户可与就近的华为办事处联系，也可直接与公司总部联系。

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>

客户服务电话： 0755-28560000 4008302118

客户服务传真： 0755-28560111

客户服务邮箱： Support@huawei.com

目 录

1 概 述	6
2 N8000 集群 NAS 存储系统架构	7
2.1 产品逻辑结构.....	7
2.2 集群 NAS 软件.....	7
2.3 集群 NAS 引擎.....	8
2.4 交换设备配置.....	9
2.5 存储单元配置.....	9
2.6 用户管理子系统.....	9
3 组网以及环境要求	10
3.1 集群 NAS 存储系统.....	10
3.2 NAS/SAN 一体化典型组网.....	11
3.3 数据备份组网.....	11
3.4 系统运行环境.....	13
4 软件特性	14
4.1 集群文件系统.....	14
4.1.1 技术背景	14
4.1.2 技术原理	14
4.1.3 技术特点	19
4.2 快照技术.....	19
4.2.1 技术背景	19
4.2.2 技术原理	19
4.2.3 技术特点	21
4.3 动态分级存储技术.....	22
4.3.1 技术背景	22
4.3.2 技术原理	22
4.3.3 技术特点	26
4.4 远程复制技术.....	26
4.4.1 技术背景	26
4.4.2 技术原理	26
4.4.3 技术特点	28
4.5 重复数据技术.....	29
4.5.1 技术背景	29
4.5.2 技术原理	30
4.5.3 技术特点	36
4.6 配额管理技术.....	37

4.6.1 技术背景	37
4.6.2 技术原理	37
4.6.3 技术特点	39

图目录

图 1 产品逻辑结构图.....	7
图 2 FileStore 软件及应用架构示意图.....	8
图 3 N8000 集群 NAS 存储系统关系图.....	10
图 4 SAN、NAS 一体化典型组网	11
图 5 使用 NBU 客户端备份方式组网	12
图 6 使用 NDMP 备份组网	12
图 7 数据通道全互联技术.....	15
图 8 通信通道全互联技术.....	15
图 9 全局存储空间管理技术.....	16
图 10 高效的集群锁技术	16
图 11 对称式元数据处理技术	17
图 12 全局名字空间技术	18
图 13 集群多节点并发访问技术	18
图 14 主文件系统及其快照文件集	20
图 15 新生成快照	20
图 16 主文件系统数据更新	20
图 17 对快照进行数据更新	21
图 18 N8000 多卷文件系统示意图.....	23
图 19 N8000 远程复制技术实现步骤.....	27
图 20 N8000 远程复制技术实现机制.....	28
图 21 重复数据删除技术示意图	29
图 22 定长分块技术与变长分块技术对比.....	31
图 23 文件相似性检测技术——相似性组.....	32
图 24 重复数据删除集群架构	33
图 25 分块 Cache 预读加速.....	33
图 26 分块 Cache 预读加速.....	34
图 27 元文件与数据文件管理	35



图 28 重复数据删除比率计算公式	35
图 29 配额管理逻辑架构	38
图 30 配额管理执行方式	39



1 概述

传统NAS系统存在扩展性差、可靠性差等缺陷。为克服传统NAS系统的此类缺陷，集群NAS技术应运而生。集群NAS系统具有高扩展、高性能、高可用等特点。本文档介绍华为N8000集群NAS存储系统软件和技术的相关细节。

2 N8000 集群 NAS 存储系统架构

2.1 产品逻辑结构

N8000集群NAS存储系统（以下简称N8000或N8000系统）由集群NAS引擎（由集群NAS引擎硬件+集群NAS软件FileStore组成）、交换设备、存储单元以及管理子系统组成，其中集群NAS软件包含定制化的操作系统，不需要额外选择操作系统平台。

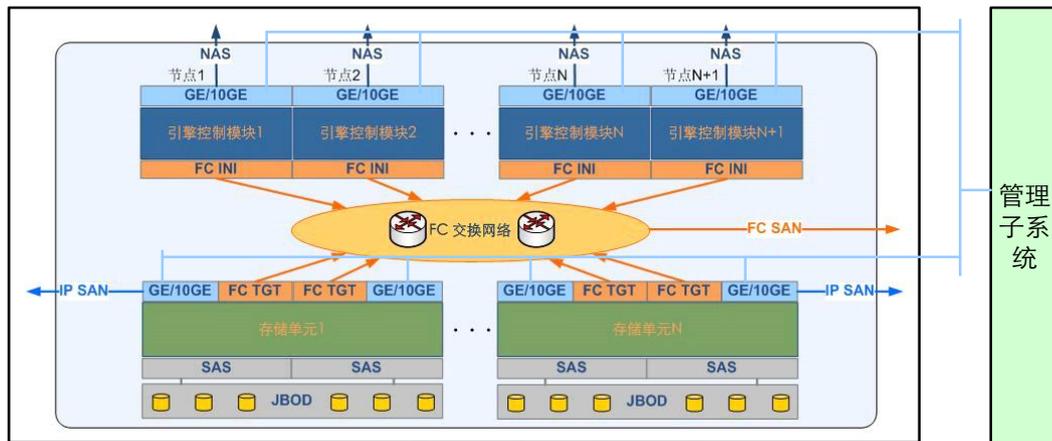


图 1 产品逻辑结构图

2.2 集群 NAS 软件

集群NAS软件FileStore作为整个N8000集群NAS存储系统的核心部件，其作用除了管理存储并提供NAS服务外，还包括对集群系统自身的管理以及相关的故障处理。

N8000的集群NAS软件基于商用的CFS开发，通过在专用64位存储系统上安装CFS软件，并配置NFS和samba等服务以对外提供NAS访问接口。

N8000集群NAS软件的核心部件CFS提供了大量的增值功能，如快照、动态存储分级等功能。FileStore软件及应用架构示意图如下图所示。FileStore软件由核心层的CFS、CVM软件和用户层的各类NAS服务应用组成。FileStore软件部署于集群NAS引擎节点（Node）上，集群NAS引擎节点可通过集群方式扩展，提供全Active的集群架构。所有节点可访问同一个共享存储空间，对外可通过以太网向上层应用提供NAS业务。

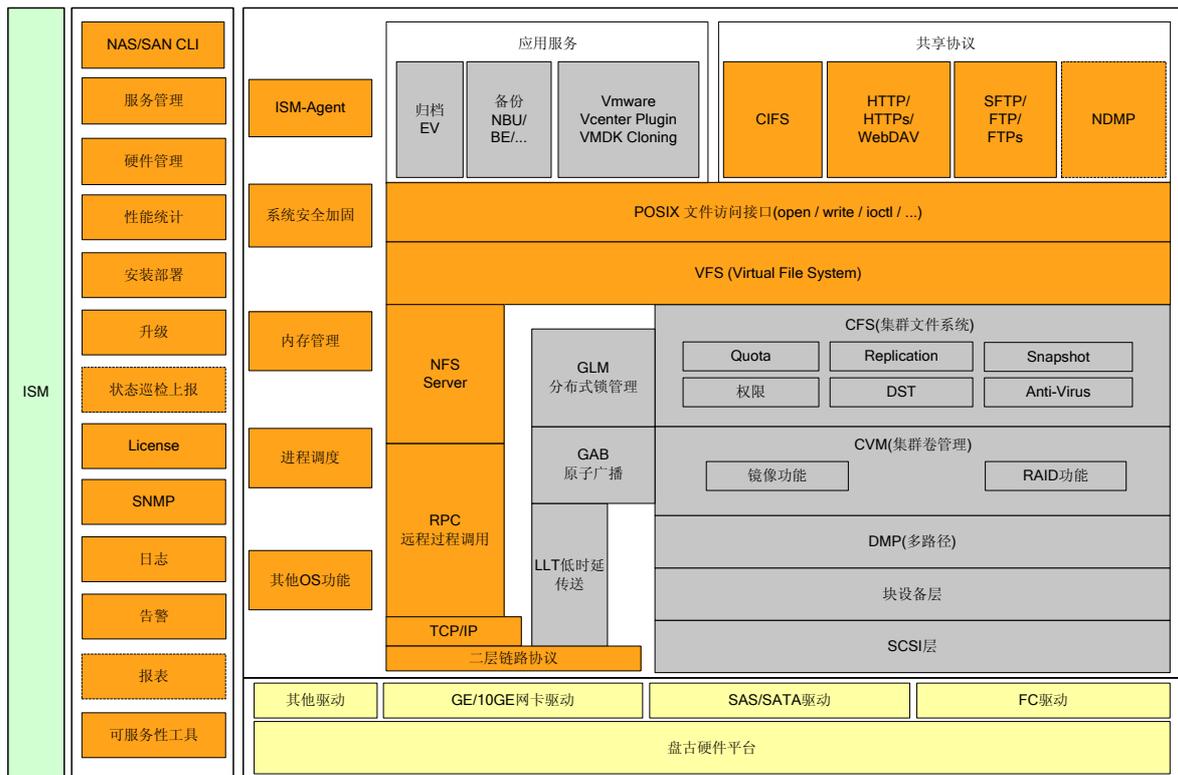


图 2 FileStore 软件及应用架构示意图

2.3 集群 NAS 引擎

NAS引擎是集群软件运行的硬件平台，通过与FileStore软件的配合，将后端存储提供的卷以NAS方式提供给客户使用，需要较好的I/O处理能力、网络带宽和可靠性。

2.4 交换设备配置

在N8000集群NAS软件中，集群节点间的通讯通过专用的心跳网络来实现，同时，连接多套存储单元时需要光纤交换设备。心跳网络对交换机的带宽要求低，光纤交换设备要求支持8Gb接口。当选择的节点数超过2个时，需要额外配置2台以太网交换机，作为集群NAS的心跳链路使用。

2.5 存储单元配置

支持S2600T、S5500T、S5600T、S5800T、S6800T存储系统。上述所有类型的存储单元，均可向下细分为两个组件，即控制框和硬盘框，其中控制框包含了控制器及控制软件，硬盘框用于磁盘容量扩展

在N8000集群NAS存储系统中，将控制框也称之为存储单元控制框，将硬盘框也称之为存储单元硬盘框。存储单元控制框前可增加型号的描述区分各种型号，如S5x00T存储单元控制框。

2.6 管理子系统

N8000集群NAS存储系统提供GUI和CLI两种管理方式。GUI支持NAS引擎和存储单元统一管理，CLI仅支持NAS引擎和存储单元独立管理。另外有些高级功能及部分系统配置功能仅适用CLI支持，面向高级管理维护人员。

3 组网以及环境要求

3.1 集群 NAS 存储系统

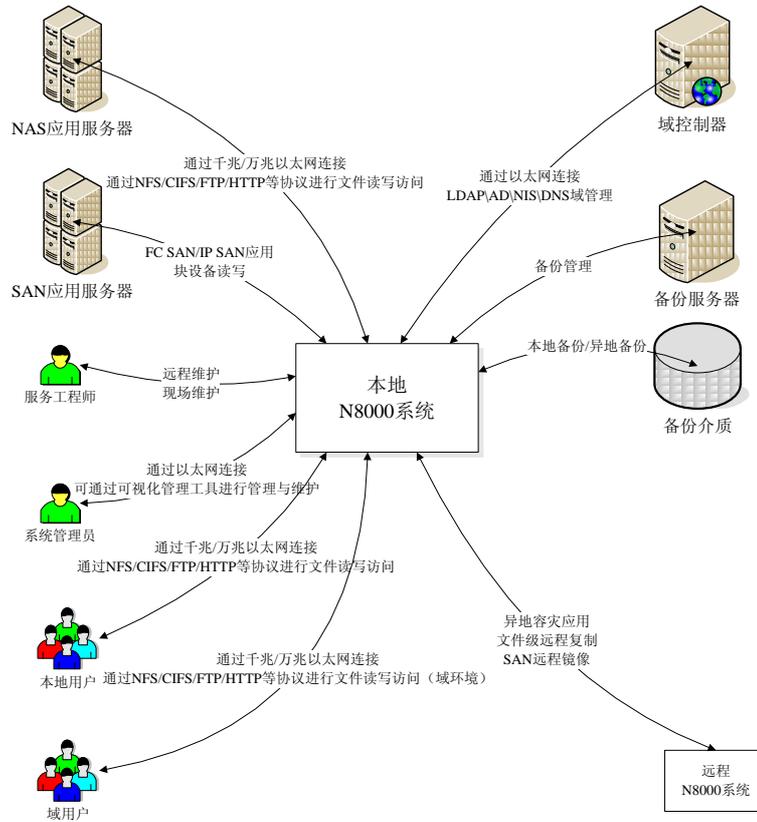


图 3 N8000集群NAS存储系统关系图

3.2 NAS/SAN 一体化典型组网

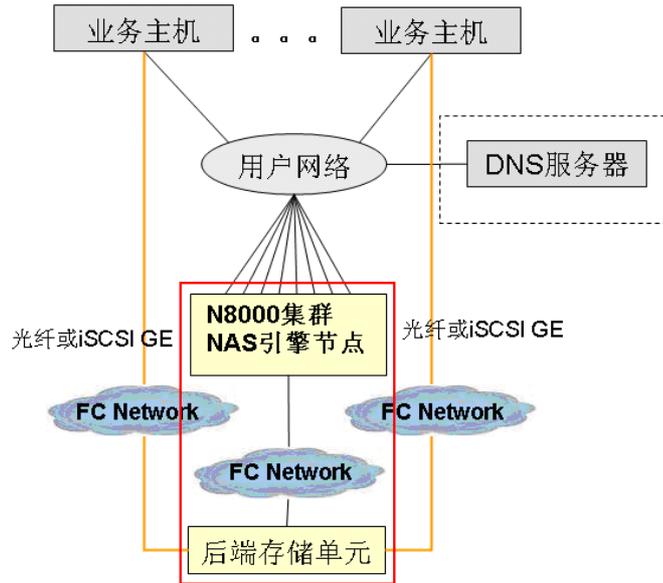


图 4 SAN、NAS一体化典型组网

组网说明：

N8000系统的NAS、SAN一体化体现为可直接提供FC或iSCSI接口，提供FC-SAN和IP-SAN的组网方式供用户业务使用。

在N8000系统中，仅使用了存储单元的部分端口，在引擎节点数量比较少或使用交换形态的情况下，可以根据需要配置10Gbps iSCSI扣卡，从而为用户提供万兆IP SAN；若用户有SAN的需求，则可以将部分磁盘空间（以LUN为单位）通过空闲的FC端口或iSCSI/10Gbps iSCSI端口提供给用户使用，达到NAS/SAN一体化的效果，在这种组网下，用户的具体网络环境可能为无域的环境、AD域环境、NIS域环境或LDAP环境等。

3.3 数据备份组网

N8000集群NAS存储系统支持两种备份方式，一是通过集成的NBU客户端进行备份；二是通过NDMP协议进行备份。两种备份方式的组网如图示：

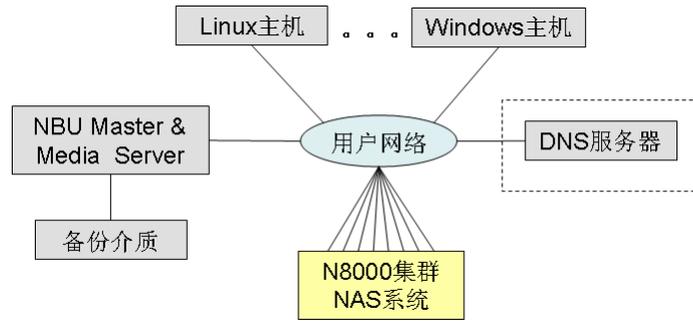


图 5 使用NBU客户端备份方式组网

组网说明:

使用NBU客户端进行备份时，需要环境中包含NBU服务器(包括Master Server和Media Server)，如果使用DNS，则要求DNS服务器能对NBU服务器进行正确解析，网络连接好后在N8000系统上配置NBU服务器相关的信息，然后在NBU Master Server上进行备份策略等配置后，备份即可按照设置的策略正常进行，数据流从N8000引擎经Media Server后备份到后端介质，无需再进行干预。

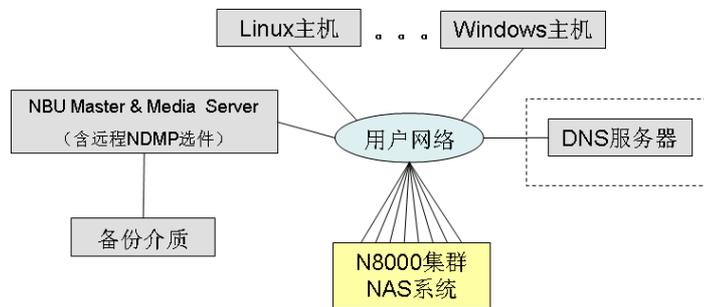


图 6 使用NDMP备份组网

组网说明:

使用NDMP协议进行备份时，当前版本支持通过LAN进行备份的方式，即备份数据通过IP网络进行传输，备份采用的协议为标准的NDMP协议。在采用基于IP网络的NDMP备份时，同样需要有备份服务器（如NBU），由备份服务器连接备份介质（磁带库设备），根据备份软件的不同，可能需要相应的选件支持，如使用NBU时需要备份服务器上安装有远程NDMP选件，在进行备份时，NDMP控制信息由备份服务器下发到N8000系统，N8000系统根据控制命令进行备份，将备份数据通过IP网络传送到备份服务器的相关模块（如NBU的Media Server）再由该模块写入备份介质。

3.4 系统运行环境

N8000集群NAS存储系统通过SAN共享技术将存储阵列提供的卷虚拟成可同时供多个集群节点使用的逻辑卷，在此基础上，通过NFS共享、CIFS共享、FTP等方式为用户提供文件服务。

对最终用户来说，N8000集群NAS存储系统可看做一个文件服务器，用户通过该文件服务器存取文件。用户IT的环境可能是比较复杂的，如AD域、NIS域、LDAP等环境，N8000系统也提供了这些域环境的支持，用户只需要进行相应的配置即可将N8000在用户的域环境中运行起来供应用主机访问。

4 软件特性

4.1 集群文件系统

4.1.1 技术背景

N8000系统的集群架构，采用全互联的组网机制，对称式集群设计，实现集群节点的全局命名空间，从而允许集群任何节点并发访问任何文件系统的任何文件；并且支持文件内的细粒度的全局锁，提供从多个节点并发访问相同文件的不同区域，实现高并发读写，最终达到高性能访问系统。

4.1.2 技术原理

➤ 数据通道全互联技术原理

N8000系统分为引擎单元和存储单元，存储单元处理数据块的存储，引擎单元提供集群文件系统的逻辑功能。引擎单元和存储单元通过直连全互联，或者交换网络实现全互联，从任何一个引擎节点都能够访问到任何一个存储单元，从而保证数据通道的畅通无阻，并且提供冗余路径实现高可靠性，如下图所示。

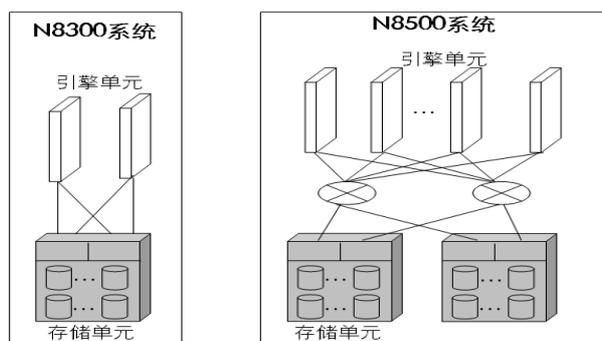


图 7 数据通道全互联技术

对于N8300系统的典型配置，2个引擎节点连接到存储单元，每个节点都有路径连接到阵列的双控端口中，从而实现全互联，并且提高冗余链路；对于N8500系统的典型配置，引擎节点通过冗余的交换机连接到存储单元，每个存储单元也分别连接到冗余交换机，从而形成全冗余的组网。

➤ 通信通道全互联技术原理

N8000系统的引擎节点之间需要相互通信，提供集群节点的统一视图，保证元数据/数据处理的一致性。节点之间采用双通道连图，冗余互联，保证通信通道的高效可靠。

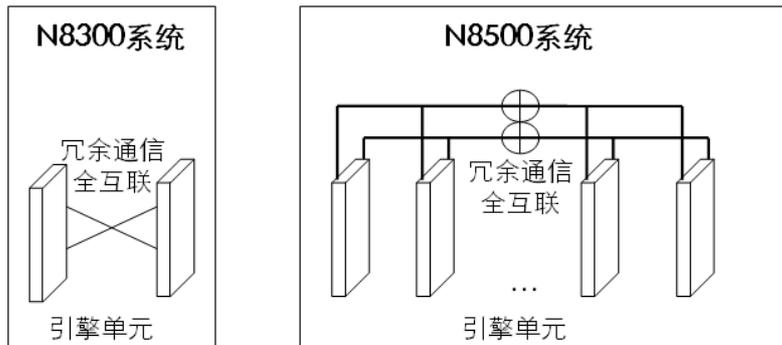


图 8 通信通道全互联技术

N8300系统的2引擎节点典型配置下，每个节点提供2条链路到另外一个节点，从而即使一条链路故障时，也能保证通信链路的可靠性。N8500系统的多引擎节点，则是连接到2个交换机上面，从而提供冗余通信平面，保证可靠性。

➤ 全局存储空间管理技术原理

N8000系统在数据通道和通信通道实现全互联的基础上，采用集群卷管理技术，提供了全局的存储空间管理。也就是说，在集群中的任何节点能够以相同的卷设备名字访问同一存储空间，保证块设备访问的一致性，如图-3所示。

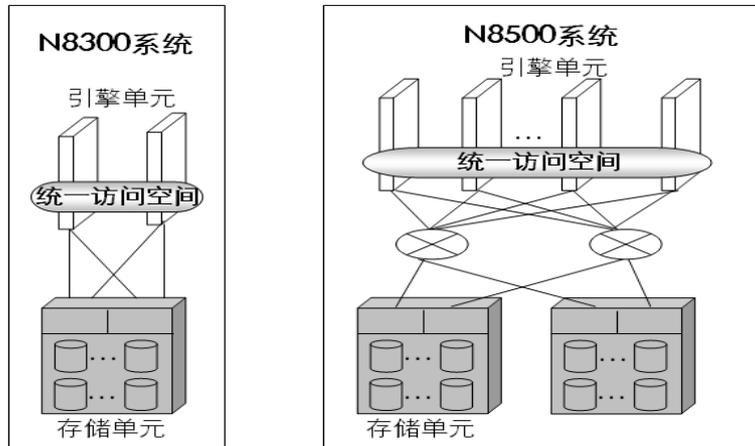


图 9 全局存储空间管理技术

➤ 高效的集群锁技术原理

N8000系统需要在节点之间传递元数据/数据，而为了保证数据的一致性，提供了GLM(Global lock manager)集群锁技术，如图-4所示。

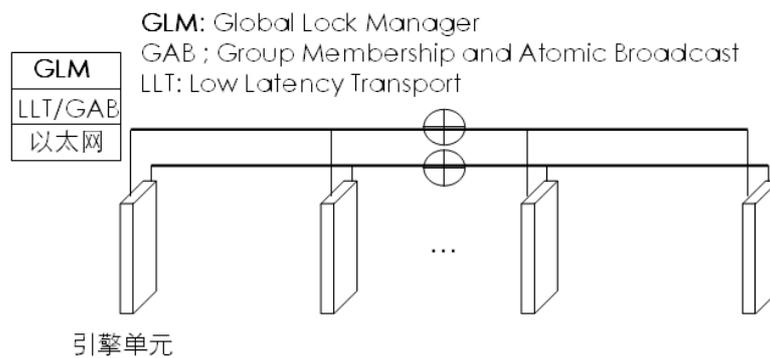


图 10 高效的集群锁技术

N8000系统的GLM集群技术是构建在LLT(Low Latency Transport)/GAB(Group Membership and Atomic Broadcast)底层通信基础上，该通信层直接构建在以太网二层协议上，而不是TCP/IP层上。考虑到集群通信主要传送的锁信息，通常帧都比较小，而TCP/IP包会带来额外的大帧头开销；同时集群通信主要在私有网络内传递，对TCP/IP在广域网的超时重传、滑动窗口没有太多需求。

所以，LLT针对集群锁通信需求，直接利用以太网帧传送包，从而达到提高利用率的目的，并且设计自己的流量控制、超时重传、以及重复帧处理，实现高效传输。

同时，针对集群多引擎节点通信的原子要求，设计了GAB模块，它保证通信的原子性；也就是说，保证在集群多引擎节点中，消息要么被所有引擎节点受到并确认，要么所有引擎节点都丢弃该消息，从而不会出现部分确认消息、部分丢弃消息的中间状态。

➤ **对称式元数据处理技术原理**

N8000系统中，元数据分为了不同的类型，比如：超级块、日志区、分配单元(AU, Allocation Unit)、INODE。这些元数据具有不同的属性，分别对不同的数据管理负责，比如，超级块对整个集群文件系统的管理负责，而inode则是针对单独的那个文件负责。

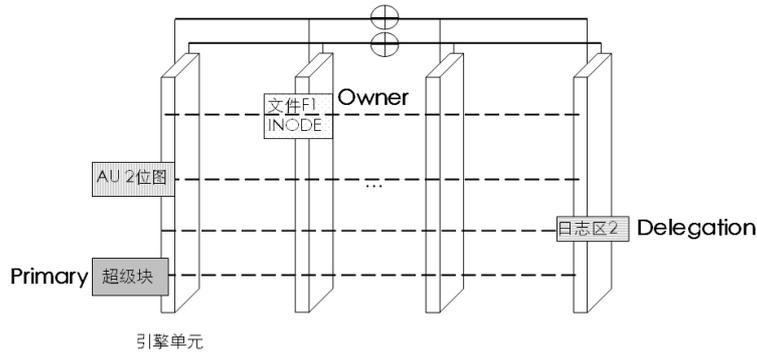


图 11 对称式元数据处理技术

在元数据分类的基础上，N8000系统的每个引擎节点就可以担当不同元数据处理的责任主体。如图-11所示，第一个引擎节点负责超级块的更新修改，其他节点需要对元数据的处理，都把请求发送给他；第二个引擎节点负责文件F1的INODE元数据更新修改，其他节点需要操作F1文件，都需要把请求发送给该引擎节点。

同时，在任何引擎节点故障后，都可以把该节点负责的元数据处理转移到其他任何引擎节点上，从而实现对称式的元数据处理，实现元数据处理的负载均衡，提高元数据处理的效率。

➤ **全局名字空间技术原理**

N8000系统采用对称式的设计架构，每个节点都可以参与处理元数据，从而很容易形成全局名字空间，让系统的每个处理节点都能够读写相同的文件系统，实现节点之间的并发访问，如图-7所示。

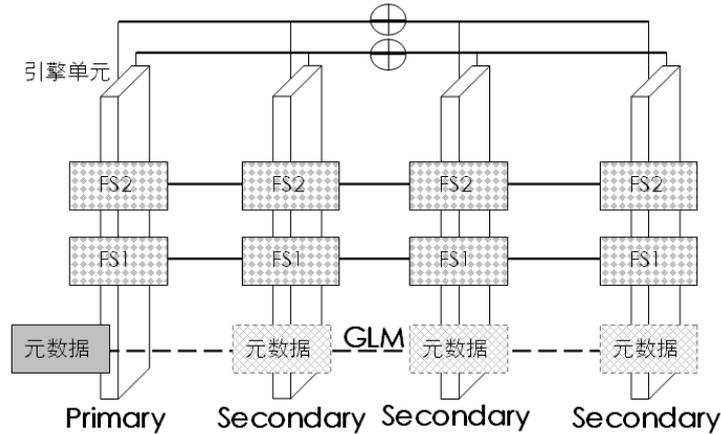


图 12 全局名字空间技术

根据元数据的分类、分片，让每个节点都能参与元数据的处理，减少节点之间转发元数据/数据请求的通信量，实现负载均衡。

➤ 集群多节点并发访问技术原理

N8000系统提供了全局的命名空间，从而实现了多个节点可以并发访问同一个文件系统。除此之外，N8000系统还提供全局锁机制，保证不同节点访问相同文件系统相同文件并发访问的一致性，通过细粒度的锁设计，保证系统的并发性，如图-8所示。

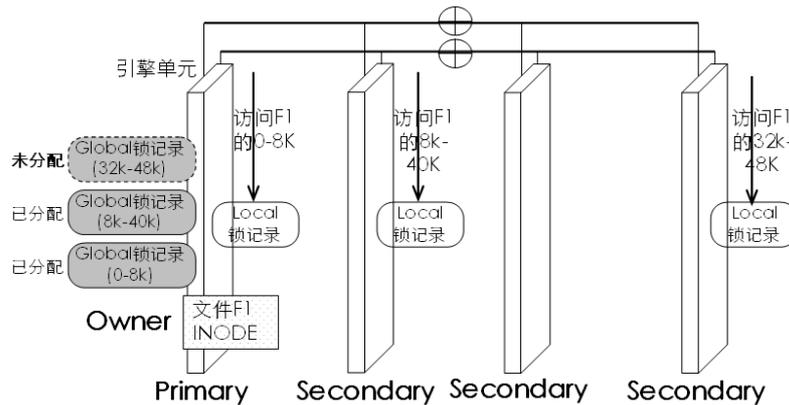


图 133 集群多节点并发访问技术

当不同节点访问文件F1的不同区域时，通过该文件的所有者(Owner)进行仲裁，如果没有其他节点访问对应区域，则允许该节点读写数据；如果发现有其他节点访问该区域，则出现冲突，该所有者就会让新请求等待，直到读写该区域的请求完成，然后通知等待该区域的节点竞争锁。通过此机制，实现了集群多节点范围内的一致性所处理；同时，通过细粒度的区域锁设计，保证高并发。

4.1.3 技术特点

N8000集群存储系统采用全Active、全共享设计。文件系统为所有引擎节点共享而不是与特定节点绑定，所有引擎节点可以并发访问同一文件系统，不存在负荷热点问题。N8000所有引擎节点互为备份，某一节点故障时，故障节点原来承担的应用可以切到其它在线节点。N8000所有节点提供业务，互为备用，具有更高的系统可靠性和资源利用率。

4.2 快照技术

4.2.1 技术背景

N8000集群NAS存储系统中的文件系统快照，支持生成源文件系统在某个时间点上的一致性映像，在不中断正常业务的前提下，快速得到一份与源文件系统一致的数据副本。副本生成之后立即可用，并且对副本的读写操作不再影响源文件系统中的数据。因此通过文件系统快照技术就可以解决如在线备份、数据分析、应用测试等难题。用户可以通过多种方法使用存储检查点。例如，它们可用于：

- 创建可以备份到磁带的文件系统的稳定映像。
- 提供源文件系统的磁盘备份，在意外删除情况下，最终用户可以恢复自己的文件。
- 在安装修补程序之前，创建应用程序的二进制副本，以允许出现问题时进行还原。
- 创建文件系统的磁盘备份，可用作传统的基于磁带备份的补充，从而提供更快的备份和恢复功能。

4.2.2 技术原理

N8000的文件系统快照采用了写前拷贝（copy-on-write）技术方式来实现，其实现原理如下：

- 1、在文件系统未进行任何快照前，数据写入动作与未创建快照的文件系统一样：对数据的任何修改操作，会直接写入文件系统中对原有数据进行覆盖，原有数据不会被保留。
- 2、为文件系统创建快照时，首先会冻结该文件系统，创建一个没有任何实际数据的快照文件集，快照文件集仅包含指向主文件系统的一组指针。由于不需要拷贝数据，因此整个过程耗时极少，从一秒以内到数秒不等，与文件系统中的文件数量有关。N8000中的快照文件集与主文件系统共同使用文件系统空间，无需单独为快照分配特定的空间。

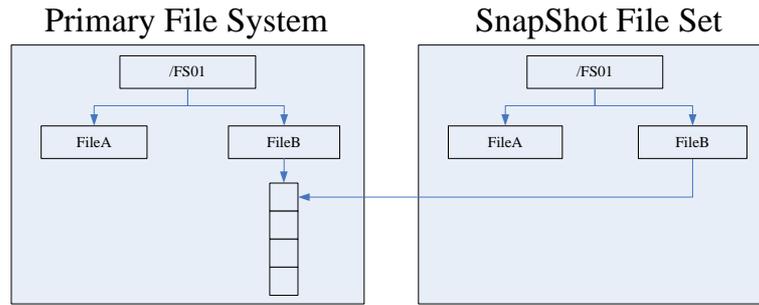


图 14 主文件系统及其快照文件集

- 3、刚创建的快照不包含任何实际数据，只包含指向主文件系统的一组指针，当用户访问快照时，实际上访问的是主文件系统中的数据。当主文件系统中的数据发生变化后，首先旧数据被拷贝到快照文件集中，然后新数据才会写到主文件系统中（即copy-on write）。

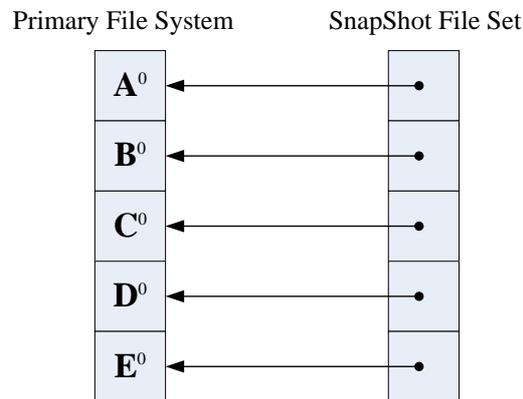


图 15 新生成快照

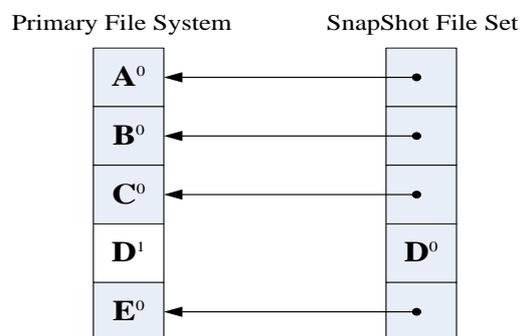


图 16 主文件系统数据更新

- 4、随着主文件系统不断被更新，原有的数据块会逐渐的拷贝到快照文件集中，但如果需要更新的块已经被快照文件集保存过一次，这时就不会发生拷贝操作，只更新主文件系统即可，

因为快照文件集所映像的只是生成快照时刻的主文件系统映像。

- 5、当用户需要恢复出快照点时刻的数据时，可通过快照数据的回滚快速实现，通过回滚，文件系统可将数据恢复到快照点时刻，从而避免了快照点后因为人为的错误或者病毒的入侵等引起的源文件系统损坏造成的数据丢失。需要说明的是，快照的回滚是不可逆的，回滚只能将数据恢复到某一特定的时间点，但该时间点到故障点之间的数据将会丢失。如果仅仅是特定的几个文件被损坏、误修改、误删除，则无需进行整个文件系统的回滚，直接从特定时间的快照文件集中将这些文件恢复到主文件系统中即可。
- 6、当快照被删除后，指向主文件系统的指针和写入时拷贝数据都将被清空，快照创建时间点的的数据变为不可用。
- 7、从上面的过程可以看到，当快照创建后，无论后续的数据怎样的读写和改变，只要快照不被删除，我们均能够及时的获取到快照创建时间点的的数据。

4.2.3 技术特点

➤ 可写快照

N8000中的文件系统快照可以为用户提供读写业务，用户可通过创建快照方便的重定义数据用途，实现诸如测试、备份、归档等各类操作。对快照的修改操作，需要说明的是，如果对某一快照执行了写操作，那么该快照将不再是快照创建时间点的一致性映像，不能使用该快照进行文件系统回滚。

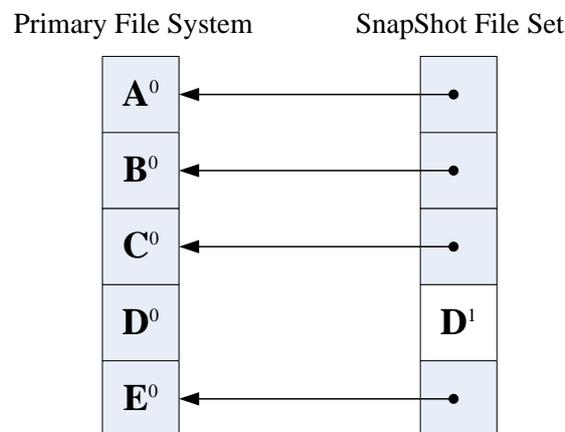


图 17 对快照进行数据更新

➤ 快速数据恢复

对于传统的离线备份，备份数据无法直接在线读取，必须经过较长时间的数据恢复过程才能够获得原数据在备份时间点的可用副本，从而实现数据的还原。N8000的文件系统快照可以直接从快照中恢复出快照时间点的数据，从而实现了方便的数据回滚。

➤ 循环定时快照，实现持续数据保护

N8000对同一文件系统支持多个时间点的快照，用户可以定制策略定时自动进行快照的创建操作。当多个时间点的快照采用循环的方式沿时间轴向前推进自动操作时，就非常方便且低成本的实现了持续数据保护。

➤ 快照对性能的影响

在文件系统中使用快照并非对系统性能没有影响。在文件系统中创建快照后，读操作的性能不会受到影响，但却增加了写操作的复杂度。当原文件系统数据更新较大时，写操作增加，系统性能有一定的降低。

4.3 动态分级存储技术

4.3.1 技术背景

N8000动态分级存储特性的主要价值是降低平均联机存储硬件成本，而又不会相应地增加管理的复杂性。重要的文件放置在高可用性、高性能（相应地，成本也非常高）的存储设备上；不太重要的文件放置在成本较低的、性能和可用性规格较低的设备上。

N8000动态分级存储特性可自动将文件重定位到正确类型的存储设备上，而不更改它们在文件系统命名空间中的逻辑位置。由于物理重定位的文件保留在相同的逻辑位置上，因此使用动态分级存储特性不需要对应用程序配置和操作步骤进行更改。对于应用程序和用户而言，动态分级存储中的文件重定位是透明的。N8000动态分级存储包括两部分：支持多卷文件系统，以及在由文件系统管理的存储内基于策略自动迁移文件。

4.3.2 技术原理

➤ 多卷文件系统的属性

N8000可以构建多个文件系统，并让这些文件系统的逻辑命名空间分布在多个存储卷上。构建文件系统所使用的各个卷被称作该文件系统的卷集。卷集中的各个卷是从单个存储单元中的 LUN 配置的。事实上，占用多个卷的文件系统对于应用程序是透明的。该文件系统中的所有文件都是同

一命名空间的一部分，在访问和处理这些文件时就如同它们全部占用单个卷一样。

基于文件系统的卷存储层相对于基于硬件的多层存储方法有一个重要的优点。与基于硬件的解决方案（它的各个层都基于由磁盘阵列提供的 LUN）不同的是，N8000卷可以是任何需要的容量和配置，如果有 I/O 性能或容错需要，甚至可以跨接多个N8000存储单元。

N8000动态分级存储特性所使用的文件系统必定是多卷文件系统。

➤ 使用放置级别以管理文件位置

多卷文件系统的管理员可以通过定义文件放置策略（可控制初始文件位置和重定位现有文件的环境）来控制卷集内的文件位置。文件放置策略由一些规则组成，这些规则将文件的位置限制为由管理员定义的、文件系统卷集中各个卷的子集。这些子集被称为放置级别。通常，用存储层来标识放置级别。根据策略规则，将在指定的放置级别中创建和扩展一些文件，并当这些文件满足某些与命名、活动、访问率以及大小相关的限定条件时，将它们重定位到其他放置级别。

例如，下图中表示一个 N8000 多卷文件系统，该系统的卷集由两个分别名为 tier1、tier2 的放置级别组成。这样的卷集可用于包含重要文件（tier1），以及大量非活动的文件（tier2）的文件系统。

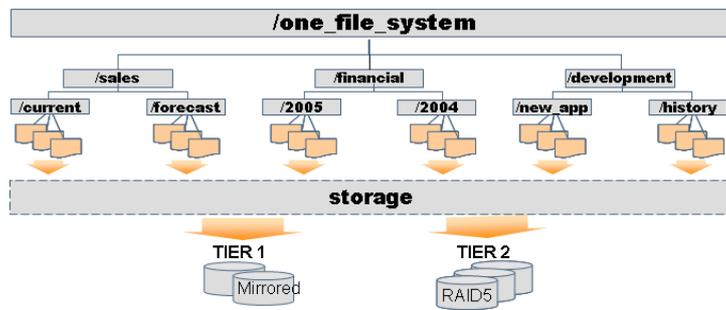


图 18 N8000多卷文件系统示意图

➤ 文件放置策略

根据文件系统的活动文件放置策略，可将文件动态放置在文件系统卷集的各个卷之间。文件放置策略由用来管理文件的初始位置和随后的重定位的各个规则组成。这些规则定义了三方面的内容，一是文件写入文件系统时的初始位置，用户可以自定义文件写入时放置于哪一个卷集中；二是文件如何从存放热点数据的卷集重定位到存放非活动文件的卷集，N8000系统中，该规则设定为指定时间长度内，未进行访问的文件则需进行重定位；三是文件如何从非活动文件的卷集重定位至存放热点数据的卷集，N8000系统中，该规则设定为若文件平均访问热度超过指定值，则将这些文件进行重定位。

策略规则将初始分配目标和重定位目标指定为按优先顺序排列的放置级别的列表。如果空闲空间允许，则在列表中的第一个放置级别中分配文件；如果在第一个放置级别中没有可用的空闲空间，则在列表中的第二个放置级别中分配文件，依此类推。

➤ 执行文件放置策略

执行文件放置策略时将实施文件重定位，可按需手动进行重定位操作，也可通过一个时间表策略定期进行，此操作与初始分配类似。如果有可用的空间，则将文件重定位到规则中列出的第一个放置级别中，如果第一个放置级别中没有可用的空间，则将文件重定位到第二个放置级别中。文件重定位基于一些必要的限制条件，例如：在一段指定的文件是否未被访问、文件的访问热度。若文件放置策略未能执行，那么，即使某些文件已经满足重定位的条件，重定位操作也不会实施。只有当文件放置策略执行时，文件系统才会对这些限制条件的满足情况进行判断，并且对符合条件的文件实施重定位。

➤ 计算文件的访问热度

动态分级存储特性的重要应用是将非活动文件从高性能存储区自动重定位到低成本存储区，或是将热点文件从低成本存储区自动重定位到高性能存储区。N8000系统通过平均访问热度来判定文件是否为热点文件。

文件的平均访问热度是在指定数量的 24 小时时段内文件的读取或写入请求数除以时段数。与 I/O 热度不同，访问热度与文件大小无关；一个在 2 天的时间段内有 20 次 I/O 请求的大文件与一个在 2 天的时间段内被访问过 20 次的小文件具有相同的平均访问热度。

➤ 数据的自动迁移

N8000的DST策略制定中，可基于文件访问频率策略，自动将文件从一级存储迁移至二级存储，或将文件自动从二级存储回迁至一级存储，对业务主机透明。

➤ 卷的可访问性：分级存储的重要假定

动态分级存储特性背后存在一个通常未表述出来的重要假定，即：系统正常工作时并不要求所有存储设备总是可访问的。企业通过将数据存储在不容错功能较低的设备上，默认表示这些数据对企业运营的重要性不及存储在容错能力较高的设备上的数据。极端情况下，如果企业将数据存储在没有容错能力的设备上，则默认表示即使没有这些数据企业也能够正常运营，至少在安装替换存储设备并从备份副本中恢复这些数据所需的时间段内是这样的。

如果常规数据中心内的某个存储设备不可用，则无法装入该存储设备上的文件系统，而且使用

该文件系统的应用程序也无法运行。不过，对于多卷文件系统，要使文件系统正常工作，只需要最关键的数据必须是可访问的；包含非关键数据的存储设备不必是可访问的。

与任何商业文件系统类似，VxFS 多卷文件系统必须可以访问其所有元数据才能正常工作。但是 VxFS 文件系统不需访问任何特殊用户文件的数据就可以装入和运行。例如，如果由于包含某些文件的数据的卷出现故障，导致这些数据不可访问，则需要这些文件的应用程序无法运行，但是文件系统可以像往常一样工作。由此可以得出一个结论：应将多卷文件系统的元数据放置在其卷集中容错能力最好的卷上，而用户文件的数据可以放置在与该文件的商业价值相称的卷上。

➤ 元数据放置

VxFS 为文件系统卷集中的每个卷都提供了一个 `dataonly` 标记，从而可以对元数据的放置进行管理控制。如果已为某个卷设置 `dataonly` 标记，则 VxFS 仅在其上存储用户文件数据，而不存储元数据。未设置 `dataonly` 标记的卷可以存储元数据和用户数据。

VxFS 文件系统卷集中的第一个卷必定可以存储元数据（名为 `metadataok` 的属性）。默认情况下，VxFS 在文件系统创建时或在向卷集中添加卷时，设置所有其他卷的 `dataonly` 标记。

通过将卷指定为 `dataonly`，管理员可以阻止 VxFS 在该卷上存储任何文件系统元数据。仅应使文件系统的容错功能最好的卷有资格存储元数据。

N8000动态分级存储特性默认将分级存储文件系统的第一个放置级别定义为高性能存储区，该存储区所使用的卷默认拥有`metadataok`属性；而第二个放置级别定义为低成本存储区，该存储区所使用的卷默认设置`dataonly`标记。因此文件系统元数据将全部存放于第一个放置级别中。

➤ 可部分访问的文件系统

即使多卷文件系统某些甚至全部 `dataonly` 卷不可访问，仍可以装入和访问VxFS 多卷文件系统。尽管仍可以对不可访问的卷进行纯粹的元数据操作（如所有权或权限更改），但存储在这些卷上的文件中的数据是无法访问的。不过，其数据位于可用卷上的文件是可访问的，VxFS 可以处理所有文件系统元数据。

新文件的空间分配操作会因没有可用于元数据的空闲空间而失败，即使在 `dataonly`卷上有足够的可用空间也无济于事。因此，管理员应在 `metadataok` 卷上分配足够的空间以满足预期需求，并应持续监视这些卷上的空闲空间。

指定 `dataonly` 仅在文件系统级别上有意义。它是与文件放置级别完全不同的属性。由于在进行初始分配和重定位决定时，VxFS 会将放置级别中的所有卷视为等同的，因此给定放置级别中的所

有卷应具有相似的容错能力和 I/O 性能，并应具有相同的 dataonly 或 metadataok 属性。

4.3.3 技术特点

文件迁移策略可以定制，数据在各级存储间自动迁移。数据迁移过程对用户和管理员透明，无需人工干预。由于多卷文件系统的使用，文件迁移完成后，不会在原址保存文件相关链接信息，没有额外的空间开销。

4.4 远程复制技术

4.4.1 技术背景

N8000 集群 NAS 存储系统中，远程复制技术建立在 VxFS (Veritas File System) 文件系统基础上，利用了文件系统的两项功能，文件改变日志 (FCL) 和存储检查点 (Storage Checkpoint)，可基于用户定制的周期性时间策略在不同的 N8000 系统间进行数据复制。我们通常将复制的数据源称为源系统，将复制的目的地称为目标系统。

➤ 文件改变日志 (FCL) 技术

该技术能够跟踪文件系统中数据更改的时间并记录更改类型，可记录多种文件系统更改操作，例如文件的创建、附加数据，覆盖数据、截短数据、扩展属性修改、打孔、链接、取消链接、重命名等。远程复制利用这项技术可迅速识别给定周期时间范围内，哪些文件中发生了数据改变，但该项技术无法识别被改变的数据块。

➤ 存储检查点技术

该技术可在任意时间点快速创建文件系统持久映像。映像生成时只包括指向文件系统的一组指针，当文件系统的的变化后，旧数据被拷贝到映像中，然后新数据才会写到主文件系统中，即写入时复制方式。每一个存储检查点映像均可以独立进行读写操作，因此，远程复制的整个过程可基于文件系统存储检查点提供，这样即使在复制过程中，有持续的写操作存在，复制过程也能够顺利进行，且用户业务不会受到任何影响。

4.4.2 技术原理

N8000 远程复制是基于给定周期的异步复制，对于每一个配置了远程复制服务的文件系统，会

为其保留至少一个文件系统存储检查点即时映像，该映像是最近一次执行复制任务时所创建，可为下一次复制任务提供参照。

N8000的远程复制可分为三个关键步骤，文件改变识别、增量数据生成、增量数据同步。文件改变识别过程可在源文件系统中识别当前周期内被改变的文件；增量数据生成过程可获取文件系统被改变的扩展区（extent），并生成待同步的增量数据；增量数据同步过程可将增量数据同步到目标文件系统。

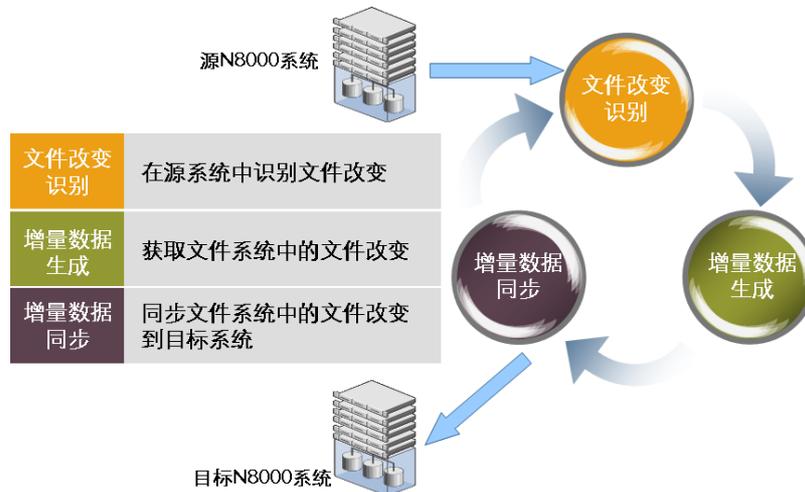


图 19 N8000远程复制技术实现步骤

在每一个复制周期时间点到来时，首先会为需要复制的源文件系统创建一个存储检查点即时映像。该周期的复制操作，将使用这个存储检查点映像作为复制源。另外，每个远程复制的文件系统均会保留上一个复制周期时间点的存储检查点，用于数据对比。

存储检查点创建完成后，N8000远程复制通过文件改变日志技术，迅速识别出在这个周期内被改变的文件，生成一个记录了这些文件的列表。该过程中，系统无需对整个文件系统进行扫描，与传统的文件级远程复制技术相比，显著提升了处理性能。

随后，N8000远程复制使用文件系统存储检查点的元数据对比功能，在当前周期的存储检查点及上一周期的存储检查点间逐一比较这些被改变文件，识别出被改变的扩展区（extent）并进行记录。整个比对过程均通过元数据进行，无需访问文件数据，效率极高。另外，由于只需对比被改变的文件，因此，即使整个文件系统中存在海量文件，也无需担心数据对比过程会耗费较长时间。

最后，N8000远程复制将被改变的扩展区（extent）中的数据提取出来，生成待传输的数据文件，并利用业界广泛应用的rsync数据同步工具，通过SSH方式，将这些数据同步到目标系统中，还原到目标文件系统中。

当前周期的复制过程完成后，源系统中，较老的存储检查点将被删除，而最新创建的存储检查点会被保留下来，以便下一个周期复制时进行数据对比。

值得一提的是，远程复制的目标文件系统可随时对用户仅提供只读业务。此时用户访问的实际上并不是真实的文件系统，而是目标文件系统在上一周期复制完毕后创建的存储检查点即时映像。这样既保证了用户能够随时访问到完整可用的文件，也不会因为用户的访问而影响到远程复制过程的顺利进行。

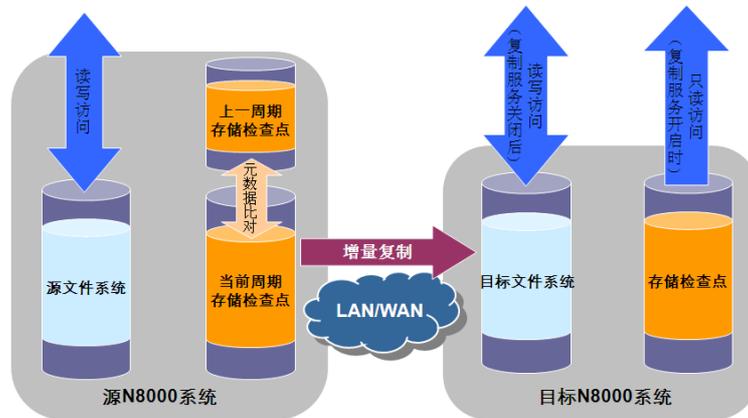


图 20 N8000远程复制技术实现机制

4.4.3 技术特点

➤ 可访问的复制数据

N8000的远程复制技术，当复制服务开启时，允许用户通过只读方式访问目标系统中的复制数据。当复制服务关闭后或用户业务已经切换到目标N8000系统，则允许用户对目标文件系统进行读写访问。

➤ 快速的业务恢复

N8000的远程复制中，源系统发生故障时，系统业务可迅速切换到目标系统，通过目标系统提供业务。由于目标系统实际上已经可以单独为用户提供只读数据访问。因此，可在极短的时间内切换到目标系统，提供只读业务。若需目标系统提供可读写业务，仅需进行少量的设置即可完成切换。

➤ 与集群技术紧密结合

N8000的远程复制服务与集群技术紧密结合，当承载复制业务的节点发生故障后，远程复制服务可自动切换到其他正常节点上，持续提供复制业务，保障了复制业务的持续性。

➤ 复制业务与数据业务分离

N8000的远程复制中，用户可配置专门的网络接口用于复制业务，可将复制业务与NFS、CIFS等数据业务完全分离。避免了复制业务与主营业务的相互影响与带宽争用。

➤ 复制过程的高效性

N8000的远程复制技术在整个复制流程中对复制效率进行了优化。利用文件改变日志技术与存储检查点的元数据比对技术提升了增量数据识别的效率。另外，由于只需要传输发生改变的数据块，大大降低了需要传输的数据量，提升了复制效率。

4.5 重复数据技术

4.5.1 技术背景

重复数据删除（De-duplication）是一种数据缩减技术，旨在减少存储系统中使用的存储容量。它通过删除存储系统中重复的数据，只保留其中的一份，从而消除冗余数据。重复数据删除技术可以很大程度上减少对物理存储空间的消耗，从而满足日益增长的数据存储需求。

在磁盘存储环境中，重复数据删除技术搜索重复数据对象（例如，数据块、文件）并删除这些重复数据。检测到重复数据时，不进行数据保留而是修改“数据指针”，以便存储系统可引用已存储在磁盘上的数据对象的精确拷贝。

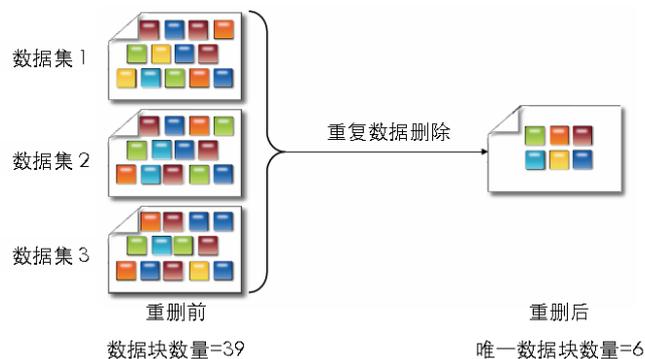


图 21 重复数据删除技术示意图

重复数据删除技术可以带来许多实际的利益，主要包括以下方面：

- 满足ROI（投资回报率，Return On Investment）/TCO（总持有成本，Total Cost of Ownership）需求。

- 可以有效提高存储效率，控制数据的急剧增长。

- 节省存储总成本与管理成本。

- 节省设备空间占用、电力供应、冷却等运维成本。

4.5.2 技术原理

➤ 基本处理流程

N8000中的重复数据删除是针对在线数据的重删，采用了文件变长分块、后处理方式的重删技术，主要处理流程如下：

1. 新文件写入时，将其划分成变长数据块，对这些数据块计算出哈希值；
2. 对新文件数据块的哈希值进行抽样，组成该文件的抽样数据；
3. 将新文件的抽样数据与现有文件的抽样数据进行比较，找到该文件的相似性组；
4. 将新文件的哈希值与其相似性组中已有的数据块哈希值进行比较，相同的块不再进行保存，新块压缩后保存，随后生成该文件的元文件，并将新的数据块哈希值保存到该相似性组的元数据库中；
5. 用户访问文件时，如果用户访问的是非重删文件，则由文件系统直接处理读写操作，如果访问的是重删文件，则根据重删后的元文件进行解析、完成读写等操作。

➤ 基于策略重删技术

N8000中的重复数据删除允许配置自定义的重删策略。用户可指定重删操作的周期，控制重删操作的启动时间，此外，用户还可以自行设定允许被重删的数据类型，精确定义重删操作的处理对象。

重复数据删除操作在系统后台运行，对业务完全透明。与传统后处理式在线数据重删技术不同的是，N8000中的重删可动态调整重删操作对系统资源的占用，优先保障在线业务的性能，对系统在线业务的影响降至最低。

➤ 文件变长分块技术

尽管使用定长分块方式算法简单，运算效率高，但此方法对数据的插入和删除非常敏感，实际应用中数据重复度非常低，重删效果十分有限。定长分块由于分块方式的固定，只要新文件中的某一部分数据大小发生变化，那么这部分数据之后的数据分块都将随之发生变化。因此，差异很小的两个文件很可能拥有几乎完全不同的数据块，这种情况在文件系统保存增量备份数据时尤其严重。

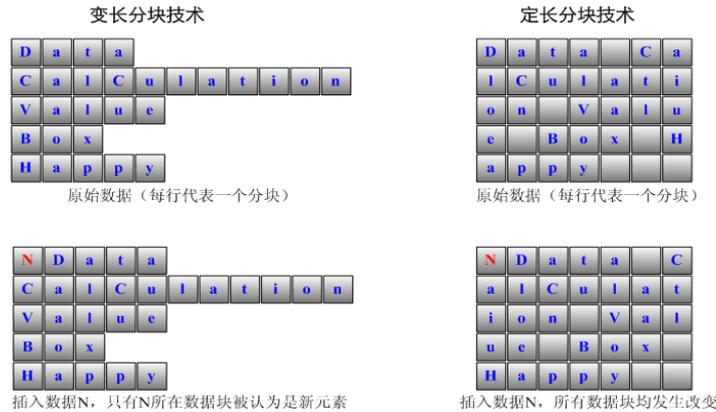


图22 定长分块技术与变长分块技术对比

N8000中的重复数据删除技术在传统的重复数据算法上进行了一系列的优化。在保证重删效果的同时提升了重删效率。

首先，采用了滑动窗口技术进行文件变长分块，这种技术基于文件内容进行数据切分，将文件分割成长度大小不等的数块。该算法对文件内容变化不敏感，插入或删除数据只会影响到较少的数块，其余数块不受影响。

其次，采用了Adler32算法计算滑动窗口的校验和时，这种算法只需要将滑动窗口滑动前的校验和，滑入字节值和滑出字节值进行复杂度极低的计算，即可获得滑动后窗口内部字节数组的校验和，提高了重删操作的运算效率。

另外，N8000引入了动态跳跃和双因子动态分块对重删操作进行了进一步的性能优化。动态跳跃规定了数块大小的最小值，每一个可变数块从开始到规定的最小值区间内的数据无需进行哈希值计算，降低了计算开销，提高了重删操作的运行效率。双因子动态分块采用两个因子来判定是否在当前位置分块，增加了数据重复的概率，提高数据重删率。

➤ 文件相似性检测原理

N8000中的重复数据删除使用获专利的文件相似性检测技术来进行重复数据的识别。这种检测技术通过一定的算法对文件数块的哈希值进行抽样，将这些抽样哈希值与当前系统中已有数块的哈希值对比确定文件的相似性。根据文件的相似性，可将重删后的文件分成不同的相似性组。每个相似性组中，各文件的抽样哈希值构成了该组的相似性索引。同一相似性组中的所有文件分块哈希值保存在该组的元数据库中，供新文件重删时进行哈希比对。

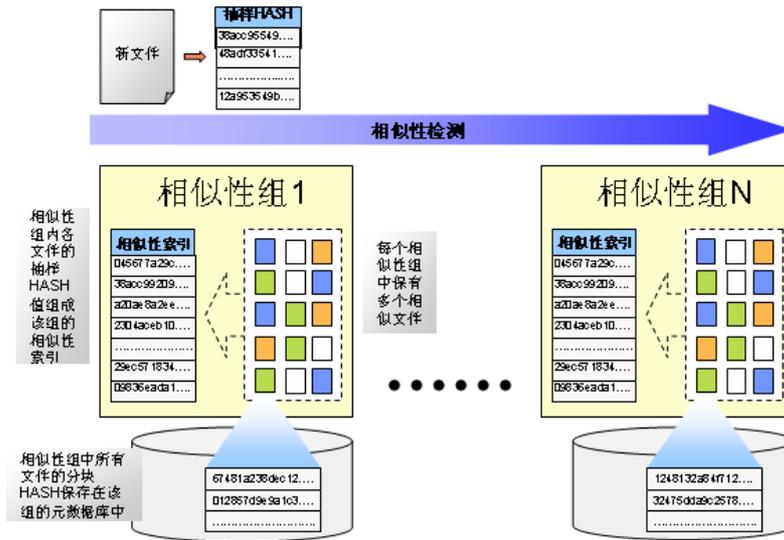


图23 文件相似性检测技术——相似性组

有文件需要进行重删时，先将该文件的抽样哈希值与各相似性组中文件的相似性索引进行比较，若该文件与某一相似性组的相似度超过一定数值，则确定该文件归属于该组，最后再将该文件中各个分块的哈希值与该组元数据库中哈希值相比较，存储唯一的数据块，并刷新相应的元数据。

这种技术降低了识别重复数据过程中的数据查询次数，与传统的在线数据重删技术相比，大大提升了重复数据删除的性能。

➤ 文件分块压缩技术

N8000中的重复数据删除为了进一步提高数据重删率，采用LZS无损数据压缩技术对需要存储的新数据块再进行压缩，通过压缩可以把分块内的冗余数据如前置零和空格删除掉。该压缩算法平均压缩率为2:1，可在重复块消冗的基础上将重删率提升一倍。

压缩的粒度为文件变长分块，粒度较小，因此压缩和解压的性能较好，对业务系统的性能影响较低。

➤ 高可用性集群重删架构

N8000集群NAS存储系统中的重复数据删除特性与系统的集群应用紧密结合。通过采用集群重删工作模式，重删整体性能及可靠性大大提高。

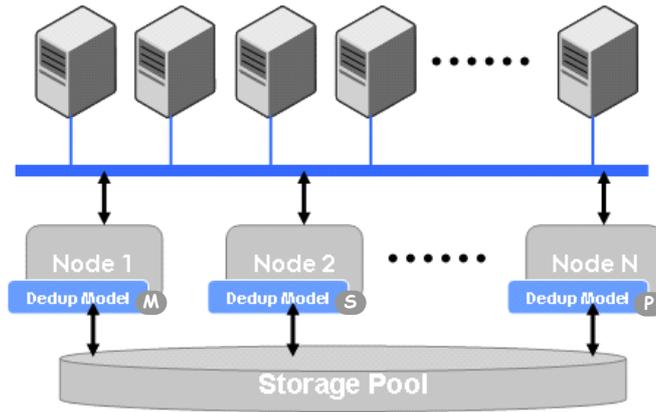


图24 重复数据删除集群架构

如上图所示，重复数据删除特性模块部署在N8000系统中每个集群节点上。重删引擎采用N+1集群工作模式，各节点的重删引擎分别工作于主、从、被状态。主节点在整个集群中唯一，负责对重删集群进行管理。从节点负责执行主节点下发的重删任务。备节点在整个集群中唯一，其上的重删引擎处于待机监控状态，一旦集群中任一节点发生故障，其将会自动代替故障节点角色进行工作，并上报告警提示用户处理。

➤ 元数据管理分离架构

N8000中的重复数据删除采用元数据管理（MDS）与重删引擎（DDE）相分离的架构。在多节点集群重删的模式下，该架构能使单个重删节点上的重复数据查询工作均分到集群中的所有工作节点上并行处理。从而减少单个节点上的访问压力，提高重删系统整体性能。

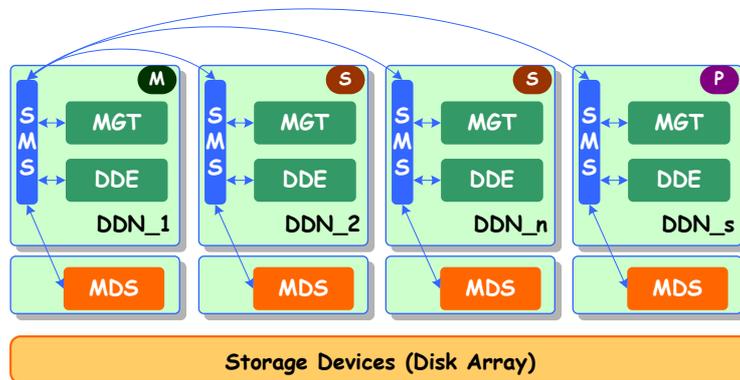


图25 分块Cache预读加速

➤ 分块信息查询加速技术

N8000中的重复数据删除引入分块Cache和布隆过滤器技术，用于提高重复块查询性能。在备份

归档场景，其每次备份数据和前一次相比，数据连续重复概率很高。分块Cache技术在查询当前分块的时候，即通过预读机制，将后续紧接的大量分块信息读取到内存当中，当需要查询下一分块时，能够直接在内存当中命中，减少查询过程中对磁盘I/O的访问。

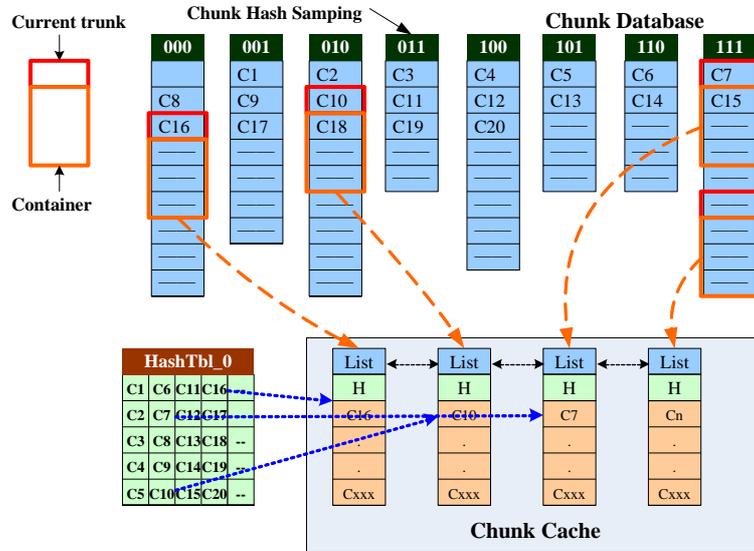


图26 分块Cache预读加速

布隆过滤器利用位数组很简洁地表示一个集合，并能判断一个元素是否属于这个集合。采用布隆过滤器优化查询效率，将已存在的块数据映射到布隆过滤器的位数组中，通过查询位数组，快速准确的发现新块，最大程度的减少数据库查询次数，提高处理效率。

分块Cache和布隆过滤器技术的引入，可以使备份归档场景下，重删对磁盘阵列的重复块查询I/O次数减少99.9%，极大提高该场景下的重删查询性能，进而提高了重删系统整体性能。

➤ 重删数据的保存

N8000中的重复数据删除将重删后的数据按照相似性组分组保存。包含了两种类型的数据对象，元文件、数据文件。元文件中包含了原文件的各项文件属性等内容，并记录了原文件各数据块的存放位置；数据文件则保存文件内容数据块。

下图表示了一个相似性组中，元文件与数据文件的保存方式。

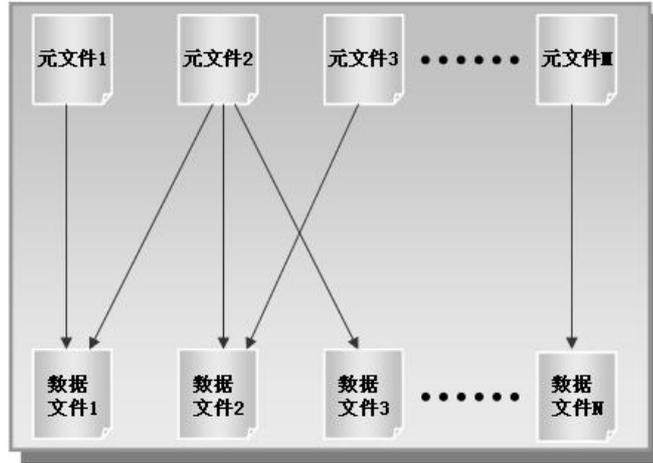


图27 元文件与数据文件管理

➤ 重删数据读写方式

重删后的数据读写方式与普通数据的读写方式不同，当文件系统收到文件读写请求时，会判断该文件是否属于重删数据。若否，则按照文件系统原有的处理方式处理读写数据流；若是，则通过重删后的元数据对文件进行检索解析，找到需要进行读写操作的数据块，最终完成数据读写。

值得注意的是，重删数据的元数据读操作通常都会在缓存中进行，通过专用的元数据管理模块及缓存管理模块，可以使得读操作更具效率；而重删数据写时，系统不会将该文件所有的数据块全部解析出来，只定位出此次写操作影响的数据块，进行小范围的数据操作。这些措施在最大程度上保障了开启重删特性的N8000系统上的业务读写性能。

➤ 文件类型对重删比的影响

我们通常使用重复数据删除比率（简称重删比）来衡量重复数据删除的效果。若用重删前的总数据量表示在传统存储系统中存储数据所需的空间大小，用重删后的总数据量表示在带重删的存储系统中存储相同数据所需的空间大小。这两个数值间的差异率即是重复数据删除比率。下图显示了重复数据删除比率的计算公式。

$$\text{重复数据删除比率} = \frac{\text{重删前的总数据量}}{\text{重删后的总数据量}}$$

图28 重复数据删除比率计算公式

通过重复数据删除特性对文件系统的数据进行处理时，数据集内重复数据段的数量会因为数据性质的不同而存在很大的差异，这通常取决于数据文件的类型以及创建这些文件的应用程序。对

具体应用场景的分析有助于我们了解重复数据删除特性在这些场景中的应用效果与价值。

在某些特定的应用场景下，例如：来自特定数据库的一组备份映像，将数据写入文件系统中时，重复数据删除的优势往往非常明显，因为每个新的写入操作只会写入该操作引入的新数据段，而在传统的数据库备份应用中，两个备份之间的数据段级差异时常只有1%~2%，尽管高变化率也会存在。在这样的应用场景下，高效的重删比将令投资者满怀信心。

相反，在另一种应用场景下，例如：保存数万份照片的素材库，重复数据删除能够取得的效果则差强人意。因为不同的照片间能够找到的重复数据段数量极为有限。这将最终体现为低效的重删比。

因此，在应用重复数据删除技术时，需要针对具体应用场景进行具体的分析。我们推荐在两种应用场景中开启重复数据删除特性，一是备份应用场景，此种场景下，数据的重复率极高，重删效果十分明显；二是虚拟机应用场景，此种场景下，存储系统中会存放大量虚拟机文件及这些文件的拷贝，数据重复率高，重删效果明显。

4.5.3 技术特点

➤ 业务无感知

N8000的重复数据删除使用了多种技术优化性能，对系统在线业务的性能几乎不会带来任何影响。首先，N8000的重删可动态调整重删操作对系统资源的占用，优先保障在线业务的性能，可以智能调整重删操作对系统资源的占用。其次，N8000的重删采用了领先的文件变长分块技术，利用了滑动分块技术以及高效的Adler算法及动态跳跃和双因子算法，在文件分块运算效率上优于传统重复数据删除技术。再次，N8000的重删使用领先的文件相似性检测技术检测重复数据，这种技术通过相似性检测算法将存储系统中的文件分为多个相似性组，在组内进行数据比对，更有利于重复数据的识别，并且减少了数据比对次数。与传统重复数据删除技术相比，去重率更高，性能更优。另外，N8000的重删使用LZS无损压缩算法对需要存储的新块数据进行压缩存储，在分块消冗的基础上将重删率提升了一倍。

➤ 集群保护提供系统级的可靠性

传统重复数据删除产品中，通过单个物理设备提供数据业务，此时，重复数据删除软件以及承载该软件的物理设备都会成为故障单点，对系统可靠性带来了挑战。

N8000将重复数据删除与Active-Active集群技术相结合，提供了系统级的可靠性。在多节点的集

群环境中，只要系统中任意一个节点仍然在正常运行，那么重删数据的读写都能够顺利执行，保障了客户业务的连续性。

➤ 最大限度降低TCO

N8000的重复数据删除特性通过消除数据存储空间中的冗余数据，用户可从存储空间效率上获益。这将直接体现为初期存储采购成本的降低，并且，重复数据删除特性可有效控制数据增长，也延缓了后继的存储扩容需求。另外，存储空间需求的降低使得用户无需管理大量的存储设备，带来了空间、电力、制冷及维护管理等运维成本的降低。最大限度降低TCO。

4.6 配额管理技术

4.6.1 技术背景

磁盘配额是一种基于用户和分区的文件存储管理。通过磁盘配额管理，管理员就可以对本地用户或登录到本地计算机中的远程用户所能使用的磁盘资源进行合理的分配，每一个用户只能使用管理员分配到的磁盘资源。

磁盘配额使得管理员可以方便合理地为用户分配存储资源，可以限制指定账户能够使用的磁盘资源,这样可以避免因某个用户的过度使用磁盘资源造成其他用户无法正常工作甚至影响系统运行避免由于磁盘资源使用的失控可能造成的系统崩溃，提高了系统的安全性。

4.6.2 技术原理

类unix操作系统的用户管理常常分为用户和用户组。对于文件的来说，创建文件的用户和他所属的组拥有该文件，在文件对象用userID和groupID这两个属性来表示其拥有者。据此，我们可以统计和控制特定用户或用户组在特定的文件系统中拥有的文件数和文件占用空间。

在特定的文件系统上应用配额管理主要流程如下，首先，在挂载文件系统时使用quota选项；其次，从应用层使能quota特性；最后，为某个用户或用户组设置配额的门限值。

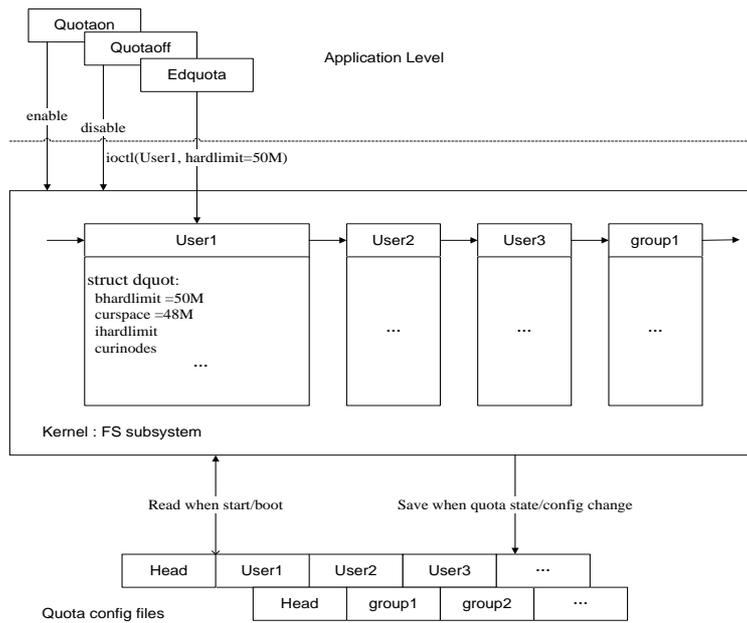


图29 配额管理逻辑架构

内核用dquot结构表示user或group的配额管理项，hardlimit为用户从应用层下发的门限值，curinode和curspace分别代表运行时user或group所拥有的文件数目和文件所占字节空间。此外，内核还维护二个基于文件系统的配额配置文件，分别保存着当前用户和用户组的配额状态，这样下次启动时系统可以获取上次的配额状态。

配额的实现附着在内核文件子系统上面，从应用层下来的文件读写创建请求，会先经过quota模块，与对应用户配额中的门限值相比较，根据门限值允许或禁止此次文件操作。如果允许，请求才能继续进行，并更新此用户的配额运行状态。不合法则错误返回。

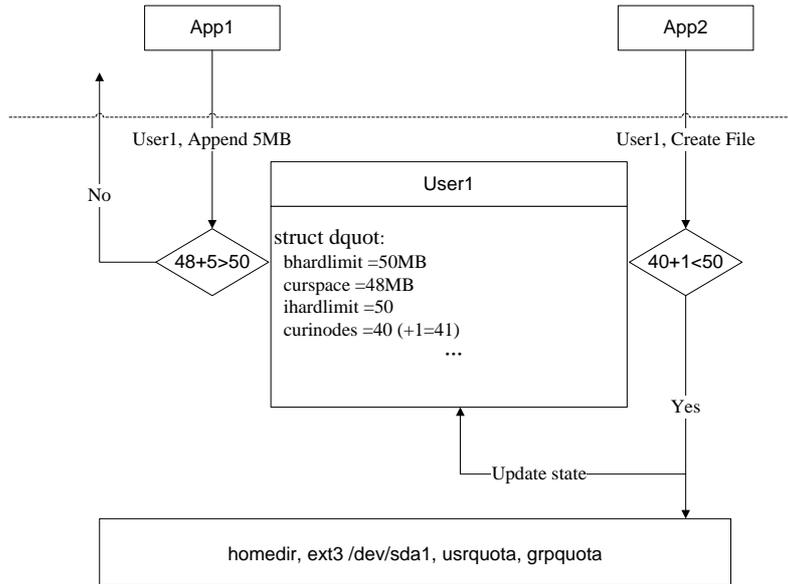


图30 配额管理执行方式

4.6.3 技术特点

N8000系统中的配额管理从多个角度，不同类别对用户使用的系统资源作灵活的限制。

➤ 配额的文件系统

N8000系统中的配额管理功能分为两大类，一是对文件系统的配额管理，二是对CIFS homedir的配额管理。

homedir共享中的配额是指homedir用户可操作的存储空间，也是这个home文件夹的可用大小。用户配额的指定必须不大于文件系统的总大小，否则指定配额操作报错。

N8000提供两种方式来设定用户配额：其一是手动创建homedir时，指定用户配额；其二是直接指定所有homedir的用户配额。

➤ 配额的作用对象

N8000系统可以设置传统的用户配额和用户组配额。

➤ 配额的限制资源

N8000配额系统限制对文件系统两种主要资源的使用：文件数量和文件系统空间。针对每种资源，可以向各个用户和组指派配额，以限制对这些资源的使用。

➤ 配额的限制策略

N8000系统可以为设置以下这两种限制：

- 硬配额：在任何情况都不能超出的绝对限制。
- 软配额：软配额的值必须低于硬配额，用户实际使用的资源可以超出软配额的限制，但仅局限于有限的一段时间内。默认限制为七天。