EMC²
where information lives®

White Paper

EMC Solutions for Backup to Disk Accelerating NAS Backup with EMC Celerra, MPFS, and EMC NetWorker

Applied Technology

Abstract

This white paper describes the benefits of using MPFS with EMC® NetWorker® to accelerate EMC Celerra® LAN backups and restores. The paper describes upgrading existing NAS backup infrastructure from legacy NFS access to EMC's more efficient MPFS product. This paper also presents comparative performance data from EMC internal testing.

May 2008

Copyright © 2008 EMC Corporation. All rights reserved.

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED "AS IS." EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com

All other trademarks used herein are the property of their respective owners.

Part Number H4363

Table of Contents

Executive summary

With limited resources and increasing demands, today's enterprises need to manage IT costs and reduce the risk of business disruption, therefore backing up data is an essential storage requirement in IT operations. Regular backups are necessary to restore data and resume critical business operations in the event that data is lost, destroyed, or corrupted. Enterprises face a daunting array of problems when they consider the best way to protect valuable data, including:

- Budget constraints and strained IT resources
- Risk of business disruption from loss of access to application data
- Long, time-consuming backup windows
- Slow restore times interfere with accessing application data

EMC® Celerra® and MPFS technology allow customers to experience optimal performance, shorter backup windows, and faster restore times, maximizing the investment in hardware infrastructure. MPFS helps EMC NetWorker® achieve significant backup and restore performance, which lowers the amount of primary business application downtime.

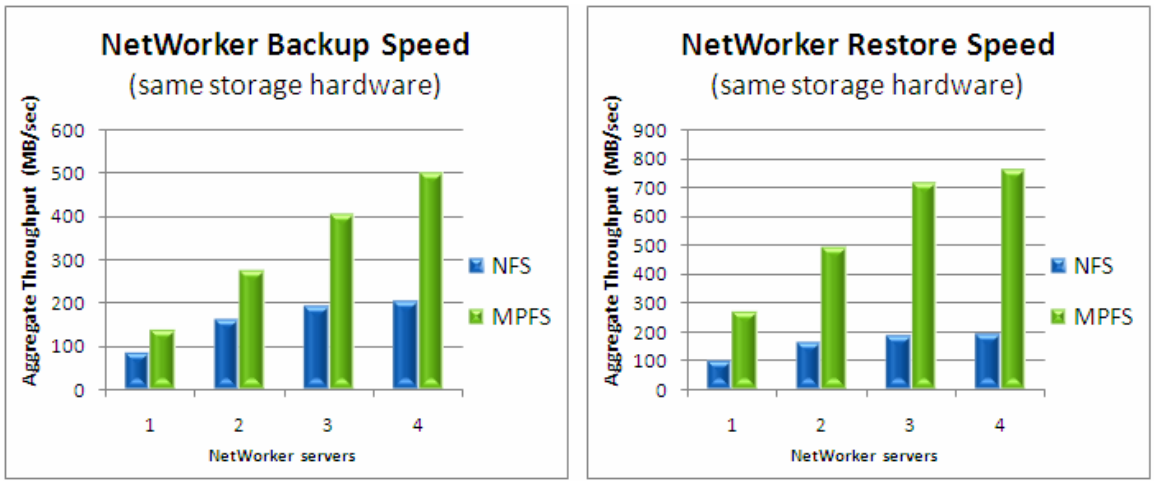

Figure 1 NetWorker restore and backup speeds using MPFS vs. legacy NFS

Introduction

Data backup and recovery are essential IT operations that have special storage requirements. Enterprises must quickly restore data and resume critical business operations in the event that data is lost, destroyed, or corrupted. In the last five years, the value of data has increased dramatically within all kinds of organizations, and the importance of protecting information has grown more rapidly. Disk-based backup is becoming popular for a number of reasons, including excellent performance, reliability, and the decreasing cost of large capacity disks. It is no longer financially and strategically practical to use tape-only solutions.

EMC leverages our industry-leading MPFS (Multi-Path File System) technology to enable EMC NetWorker to centralize, automate, and accelerate data backup and recovery across an IT environment with outstanding aggregate performance. MPFS uses a standard NAS file-sharing environment without application changes on existing infrastructures. Figure 1 shows greatly improved NetWorker backup and restore is achieved under MPFS.

Audience

The intended audience for this document is IT planners, storage architects, and storage and backup administrators who are involved in IT business and backup strategies or seek a context for the implementation and configuration of their NAS backup solutions.

Terminology

- **Backup window**: The period of time that a system is available to perform a backup procedure. Backup procedures can have detrimental effects to system and network performance, sometimes requiring the primary use of the system to be suspended. These effects can be mitigated by arranging a backup window with the users or owners of the system(s).
- **CIFS**: Common Internet File System.
- **DM**: Data Mover. The EMC Celerra NAS blade.
- **EMC NetWorker**: A client-server-based software application used for backup and recovery of data.
- **Fallthrough**: The event when MPFS temporarily reverts back to the NFS or CIFS protocol to provide continuous data availability, reliability, and protection while block I/O path congestion or unavailability is resolved.
- **FC**: Fibre Channel. A high-bandwidth data-transfer mechanism that uses optical cables to connect devices and is designed for high-performance storage systems in a SAN.
- **FMP**: File Mapping Protocol. On Celerra systems, the file mapping protocol is used to exchange file-layout information between a MPFS client and the MPFS server (the Celerra Network Server).
- **iSCSI**: Internet Small Computer System Interface. A network protocol standard that uses SCSI protocol over TCP/IP networks.
- **MPFS**: Multi-Path File System. A Celerra Network Server feature that allows heterogeneous clients with MPFS client software to concurrently access shared data stored on a Symmetrix® or CLARiiON® array directly over Fibre Channel or SCSI in a SAN.
- **NAS backup**: In the context of this white paper, a solution where data is written to hard disk instead of tape. Backups are stored on disk in the same way that a backup would be stored on a tape drive.
- **NFS**: Network File System.
- **ROI**: Return on Investment. The ratio of money gained or lost on an investment relative to the amount of money invested.
- **RTO**: Recovery Time Objective. The boundary of time and service level within which a business process must be accomplished to avoid unacceptable consequences associated with a break in continuity.
- **SLA**: Service Level Agreement. The part of a service contract where the level of service is formally defined.
- **SP**: Storage Processor. An integral component of the storage array (CLARiiON) that controls the transfer of data from the NAS file or application server to the disk drives.
- **TCO**: Total Cost of Ownership. The financial estimate that is designed to help managers assess direct and indirect costs commonly related to software or hardware.

The benefits of MPFS to NAS backups

Testing has shown that MPFS outperforms legacy NFS in several categories. The benefits of MPFS with EMC NetWorker on Celerra NAS backups include:

- Significantly faster backup speeds produce shorter backup windows
- Greatly improved restore speeds produce quicker recoveries and decreased downtime
- Reduced TCO due to fewer backup servers and fewer software licenses
- Stronger SLA with shorter backup maintenance windows
- Increased ROI by extending the life of existing infrastructure

Faster NetWorker backup speeds

Testing shows that MPFS improves performance by an average of 51 percent and scales much better than NFS. Customers should consider using MPFS for data backup. Figure 2 shows that the NFS server quickly becomes the bottleneck for NetWorker backups. Adding more NetWorker servers showed only modest gains in overall performance. MPFS eliminates this chokepoint because the MPFS clients write their file content directly to the Celerra via the SAN.

NetWorker servers were scaled from one to four and NetWorker clients were scaled from four to 16. In our testing we realized linear performance gains with four MPFS-enabled NetWorker servers in this configuration which approaches the maximum Fibre Channel performance of the Celerra, in this case an NS40FC. An investigation was performed for the single MPFS-enabled NetWorker server performance and we found that four to eight clients per NetWorker server provided optimal backup throughput per backup server.

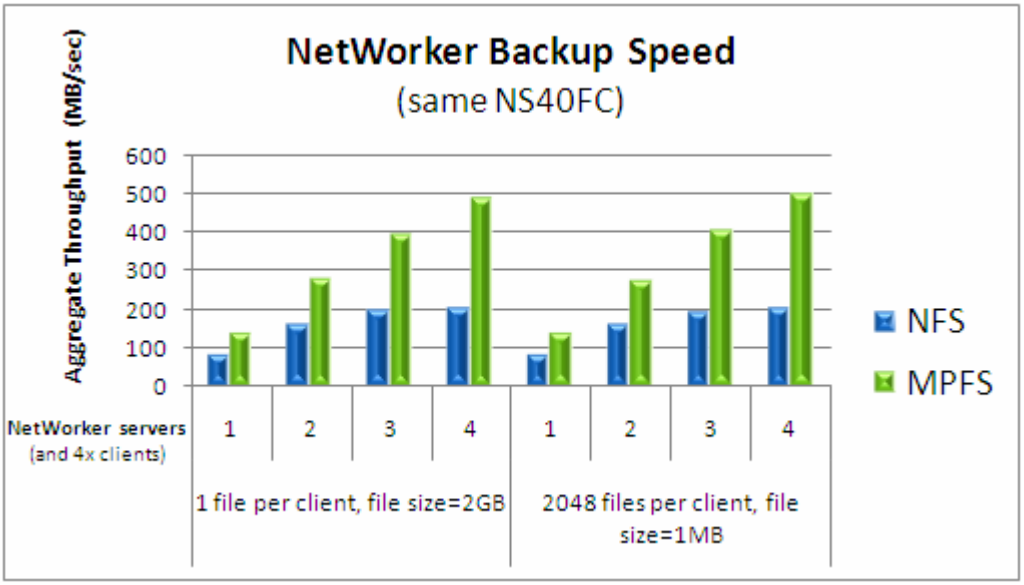

Figure 2 Backup duration comparison for MPFS and NFS

Improved NetWorker restore speeds

MPFS also improves restore performance by two to three times over NFS, which results in quicker recoveries and decreased application downtime. NetWorker servers were scaled from one to four, and NetWorker clients were scaled from four to 16, just as in the backup test. Our results showed restore operation on 16 NetWorker clients using four MPFS-enabled NetWorker servers took almost half the time of four NetWorker clients using one NFS-enabled NetWorker server.

Figure 3 compares NFS and MPFS restoration speeds. Note that a single NFS server only doubles throughput when scaled to four backup servers and MPFS improves throughput eight times NFS with the same number of servers.

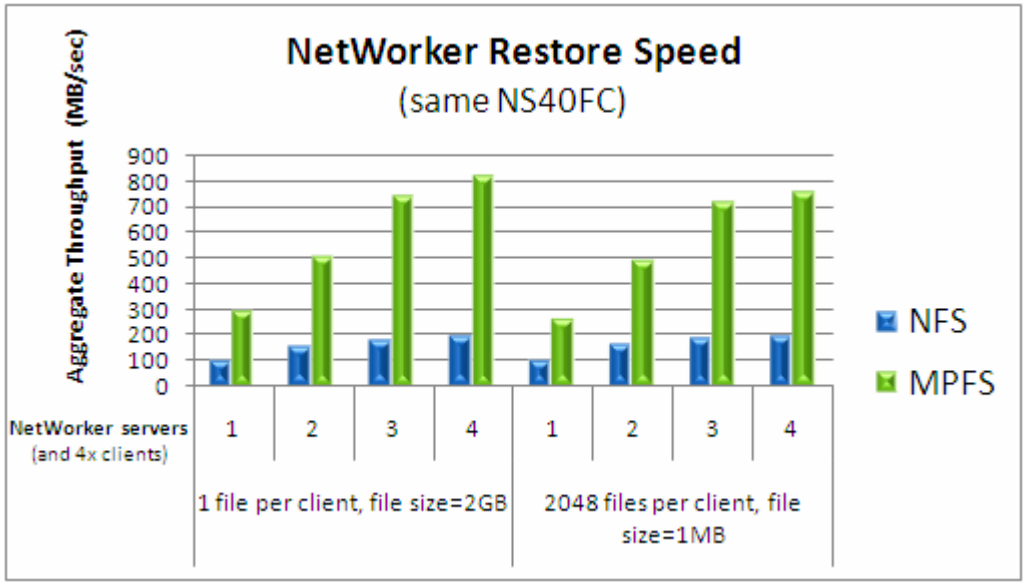

Figure 3 NetWorker restore speeds using MPFS versus legacy NFS

Reduced TCO due to fewer backup servers

MPFS improves total cost of ownership by reducing hardware, software, and operating costs. Due to performance gains by using MPFS with NetWorker, fewer servers are required. This gives financial benefits by decreasing the number of NetWorker server licenses, reducing the number of server OS licenses, and decreasing the enterprises overall power consumption. This last point is extremely popular as more companies are attempting to adopt greener technologies into their infrastructure.

Our competitors would have you purchase completely new systems in order to achieve similar performance!

Stronger SLA with shorter backup maintenance windows

As organizations grow, they often find that the backup windows lengthen to the point that they adversely affect server uptime. Employing MPFS shortens backup windows and increases the amount of peak performance operations. These shorter maintenance windows also allow for more constant and manageable server loads.

Increased ROI by extending the life of existing infrastructure

As a customer's enterprise grows, so do the demands of their data management solution. The adoption of MPFS technology to an existing Celerra platform can extend the life of older hardware by increasing efficiency without purchasing large, expensive components. It can delay the purchase of hardware upgrades to take advantage of technology improvements and improved prices. Even when customers are ready to buy new hardware, MPFS continues to stretch IT dollars over legacy NFS installations.

Is it possible to upgrade to MPFS in your environment?

MPFS requires an existing Celerra platform. Most environments can reap the benefits of MPFS enablement with its faster backups and restores, and lower hardware and software costs.

EMC has outlined some simple guidelines that are necessary to upgrade popular Celerra products to MPFS.

Celerra NS40FC: simple upgrade, low cost

Customers who use the Celerra NS40FC can enable MPFS by completing the following upgrades:

- Install a Fibre Channel host bus adapter (FC HBA) on your backup server.
- Direct-connect the backup server to NS40FC via direct connect or Fibre Channel SAN.

Celerra NSX or NS gateway with any supported CLARiiON or Symmetrix: simple upgrade, low cost

Customers who use the NSX or NS gateway can enable MPFS by completing the following upgrades:

- Install a Fibre Channel host bus adapter (FC HBA) on the backup server.
- Connect the backup server to SAN.

Celerra NS40 for MPFS: simple upgrade, lowest cost

Customers who use the Celerra NS04 for MPFS can enable MPFS by completing the following upgrades:

- Configure iSCSI on the backup server.
- Add Gigabit Ethernet NICs (as needed) to the backup server.

Tested system configuration

Two testing scenarios were created for this NAS backup testing. In the first scenario we configured four NetWorker servers to write to an NFS share. We then used the same servers to write to an MPFS share. The server and client parameters followed current EMC best practices guidelines.

Network architecture

Figure 4 shows the overall architecture of the Celerra NAS Backup Solution with EMC NetWorker and how the VLANs are configured to back up the data. The solution includes:

- Four servers as the source destination for the client data connected to the local area network. The servers are running the Enterprise version of EMC NetWorker 7.3 on an Intel 3.0 GHz server with 4 GB RAM.
  These servers are configured to access five VLANs. The first interface is dedicated to the Celerra Data Mover on VLAN 1. The second, third, fourth, and fifth interfaces were dedicated to VLAN 2, 3, 4, and 5 and each NIC was connected to one NetWorker client.
- EMC Celerra NS40FC acting as the target destination for the backup data connected to the Local Area Network.
- Sixteen Linux clients used for backup and restore operations. The NetWorker clients were installed with an EMC NetWorker client for Linux and configured to access one EMC NetWorker server through one 1 Gb NIC. A group of four backup clients, one on each VLAN (VLANs 2, 3, 4 and 5), connects to the corresponding interface on the NetWorker server.

Every backup server was connected to one of its four clients through a separate VLAN. The tested configuration used a Cisco 6509 switch that is capable of supporting Gigabit Ethernet connections at full line speed, with no backplane oversubscription or ISL trunks.

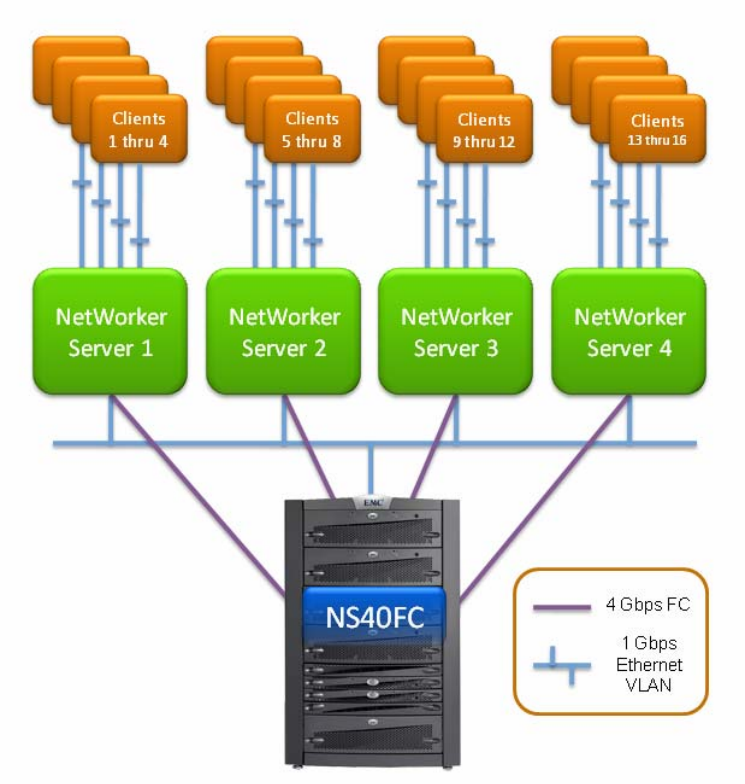

Figure 4 MPFS test environment for EMC NetWorker

Hardware resources

The following hardware was used in this tested storage configuration:

Table 1 Hardware resources

Hardware	Quantity	Configuration
EMC Celerra NS40FC	1	• Two X-blades, one active and one standby • One Control Station • Four GbE ports combined together using LACP • Four 4 Gb/s front-end ports and two 4 Gb/s back-end ports • 4 GB RAM, 512 MB Read Cache, 2500 Write Cache per SP • 120 spindles on each array • 146 GB, 15k rpm FC HDDs • First five spindles reserved for Celerra Control LUNs and FLARE®
1U Intel server	20	• Two 3.0 GHz Xeon processors • 4 GB RAM • Two onboard Intel 82541GI/PI GbE NICs • One Intel PRO Quad Port GbE card
Gigabit Ethernet switch	1	Cisco Catalyst 6509
QLogic QLA2462-E (FC HBAs)	4	Two 4 Gb ports per card

Software resources

The following software was used in this tested storage configuration:

Table 2 Software resources

Software	Quantity	Configuration
DART release 5.5.30.4	1	Installed on NS40FC
Red Hat Enterprise Linux 4 Update 3 Kernel version 2.6	20	Installed on the 1U Intel servers
MPFS Linux client 4.3.31.0	4	Installed on MPFS enabled NetWorker servers **Note**: The command **mpfsctl max-readahead 32768** was used for maximum restore performance.
NetWorker Client 7.3	16	Installed on backup clients
NetWorker Server 7.3	4	Installed on NetWorker Server
NetWorker Node 7.3	4	Installed on NetWorker Server
NetWorker Management Console 3.2	1	Installed on NetWorker Server
NetWorker License Manager 7.3	1	Installed on NetWorker Server
QLogic driver 8.01.07.15	4	Installed on HBAs
QLogic Firmware 4.00.26	4	Installed on HBAs
CLARiiON FLARE 3.24.040.5.011	1	Installed on NS40FC SPs
Navisphere Analyzer	1	Installed on NS40FC SPs
Access Logix	1	Installed on NS40FC SPs
Navisphere® Manager	1	Installed on NS40FC SPs

RAID layout as tested

Figure 5 shows the RAID layout on an NS40FC CLARiiON array as it was tested for both MPFS and NFS. The array has 120 spindles. Not counting the Celerra Control LUNs (CCL), a total of 22 RAID groups were created and one RAID 5 4+1 LUN was bound per RAID group. Figure 6 on page 12 shows a single striped volume was created across 22 LUNs.

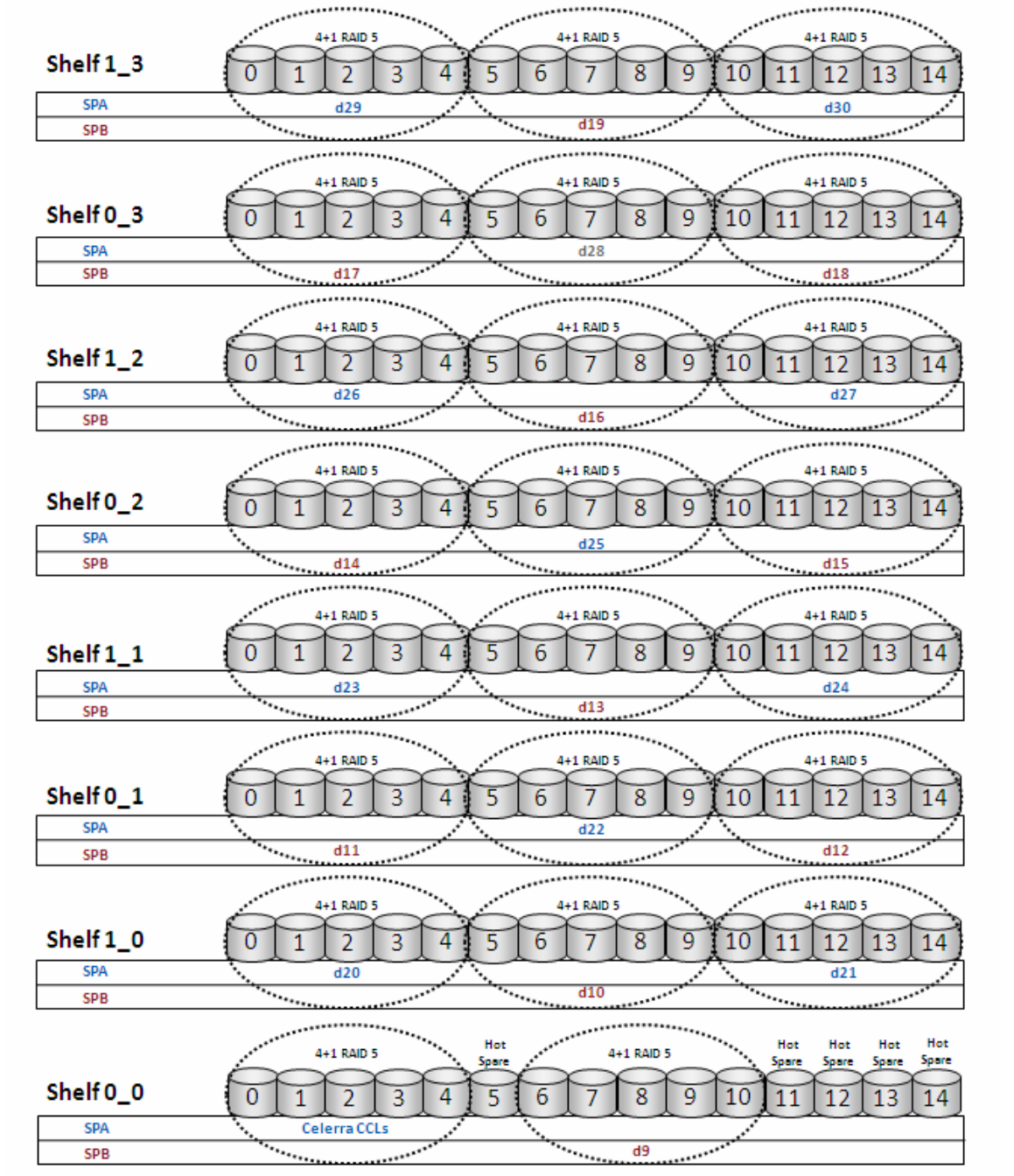

Figure 5 RAID layout in the tested configuration

A Celerra volume stripe size of 256 KB was used. A metavolume was created on the striped volume, and the file system was created on that metavolume.

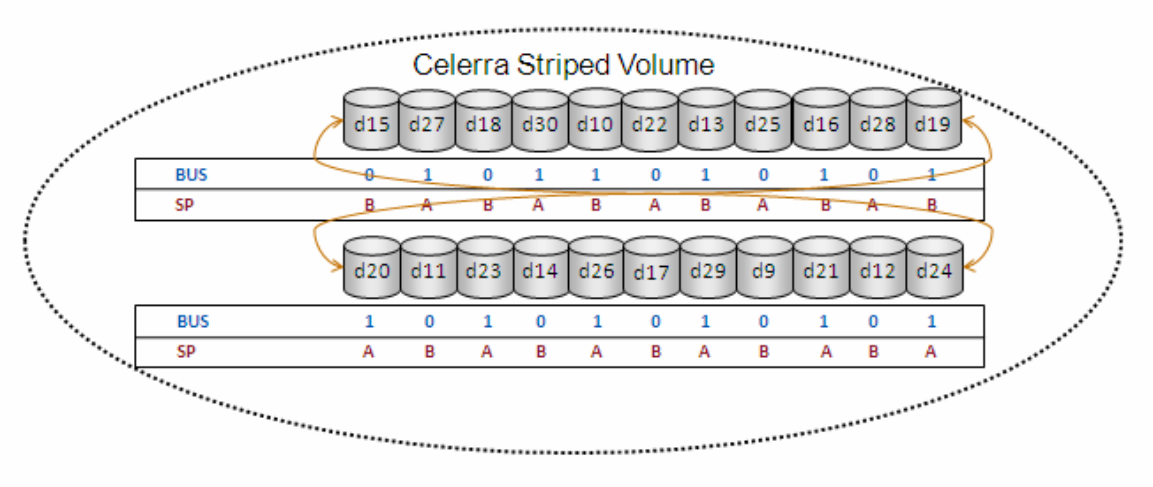

Figure 6 Celerra striped volume creation

Testing strategy

Backup and restore tests were performed using both MPFS and legacy NFS environments. Table 3 shows the various dimensions that were tested. The NetWorker servers were scaled from one to four. Each client backed up or restored 2 GB of data, either as a single 2 GB file or as 2048 one-megabyte files.

Each NetWorker server serviced four clients. For each NetWorker server that was added to the configuration, four clients were added (that were served by that NetWorker server), so that each server always handled four clients at a time.

The source data was warmed in the cache for large files to avoid the client's local disk from being the bottleneck. The clients were load balanced across different VLANs and connections to the NetWorker server(s) to avoid network bottlenecks. Four NetWorker ADV_file type devices were created on the target Celerra file system. The save and restore commands were used to initiate the backup and restore operations from a management node.

Table 3 Testing strategy for MPFS and NFS environments

Test run	Test environment	File size	NetWorker systems
Backup	• MPFS • NFS	• 1 MB files (small files) • 2 GB files (large files)	• 1 server, 4 clients • 2 servers, 8 clients • 3 servers, 12 clients • 4 servers, 16 clients
Restore	• MPFS • NFS	• 1 MB files (small files) • 2 GB files (large files)	• 1 server, 4 clients • 2 servers, 8 clients • 3 servers, 12 clients • 4 servers, 16 clients

Conclusion

EMC offers customers a wide variety of benefits with Celerra NAS backups using MPFS with EMC NetWorker including:

- Significantly faster backup speeds produce shorter backup windows
- Greatly improved restore speeds produce quicker recoveries and decreased downtime
- Reduced TCO due to fewer backup servers and fewer software licenses
- Stronger SLA with shorter backup maintenance windows
- Increased ROI by extending the life of existing infrastructure

References

For more information on MPFS and NAS backups, please see the following documents:

- *EMC Celerra MPFS for Linux Clients Version 5.0.15.1 Release Notes*
- *EMC Celerra MPFS over FC and iSCSI v5.0 Linux Clients Product Guide*
- *EMC Celerra MPFS for NS40 Integrated Configurations Quick Start Guide*
- *EMC Celerra MPFS for Gateway Configurations Quick Start Guide*
- *EMC Celerra MPFS for Windows Clients Version 5.0.78.1 Release Notes*
- *EMC Celerra MPFS over FC and iSCSI v5.0 Windows Clients Product Guide*
- *EMC Solutions for Backup to Disk EMC Celerra LAN Backup to Disk with EMC NetWorker – Reference Architecture*
- *EMC Solutions for Backup to Disk - Solution Overview*
- *EMC Solutions for Backup to Disk EMC Celerra LAN Backup to Disk with EMC NetWorker - Best Practices Planning* white paper
- *EMC Solutions for Backup to Disk EMC Celerra LAN Backup to Disk with CommVault Galaxy - Reference Architecture*
- *EMC Solutions for Backup to Disk EMC Celerra LAN Backup to Disk with IBM Tivoli Storage Manager - Reference Architecture*
- *EMC Solutions for Backup to Disk EMC Celerra LAN Backup to Disk with Symantec Backup Exec - Reference Architecture*
- *EMC Solutions for Backup to Disk EMC Celerra LAN Backup to Disk with Veritas NetBackup - Reference Architecture*
- *EMC Celerra Backup-to-Disk with Symantec Backup Exec for Windows Servers - Best Practices Planning* white paper
- *EMC Solutions for Backup to Disk EMC Celerra LAN Backup to Disk with CommVault Galaxy - Best Practices Planning* white paper
- *EMC Solutions for Backup to Disk EMC Celerra LAN Backup to Disk with IBM Tivoli Storage Manager - Best Practices Planning* white paper
- *EMC Solutions for Backup to Disk EMC Celerra LAN Backup to Disk with Veritas NetBackup - Best Practices Planning* white paper