



Storage Resiliency for the Data Center

The NetApp Approach: No Compromises

Shawn Kung, Network Appliance, Inc.

November 2006 | WP-7004-1006

Enterprise data centers require highly resilient storage systems that ensure high levels of application availability and data integrity. This paper presents the NetApp approach to preventing, curing, and recovering from storage system failures and reviews how NetApp storage resiliency technologies overcome the challenges of data availability and integrity, without compromising performance, cost, simplicity, or scalability.

Table of Contents

1. Introduction	3
2. Preventing and Curing Disk Malfunctions	3
3. Recovering from Storage System Failures.....	4
4. Conclusion	6
About Network Appliance	6

1. Introduction

Enterprise data centers today must provide high levels of application availability and consistent data integrity to support business-critical applications around the world. Data center managers continually wrestle with the challenges of avoiding unplanned downtime to ensure application data is available and avoiding data corruption to ensure that application data is correct and up to date. Compromises to either data availability or integrity can have disastrous consequences for a company's bottom line and reputation. While regional disasters and site failures get the most attention by virtue of causing the most pain, the most common causes of unplanned outages are local errors due to operational failures followed by component or system faults.¹ To achieve 99.999% application availability requires a highly reliable storage environment that prevents downtime and data corruption whatever the cause.

Two industry trends—storage consolidation and the widespread adoption of larger-capacity storage—make high availability a more urgent priority for storage and IT managers. With consolidation, even higher availability is required as more data and applications are at risk. At the same time, increased adoption of SATA storage with larger disk capacities increases the risk and probability of failures. Increased global competition and productivity growth make additional demands: in the quest for high application availability and consistent data integrity, performance and cost must not be compromised.

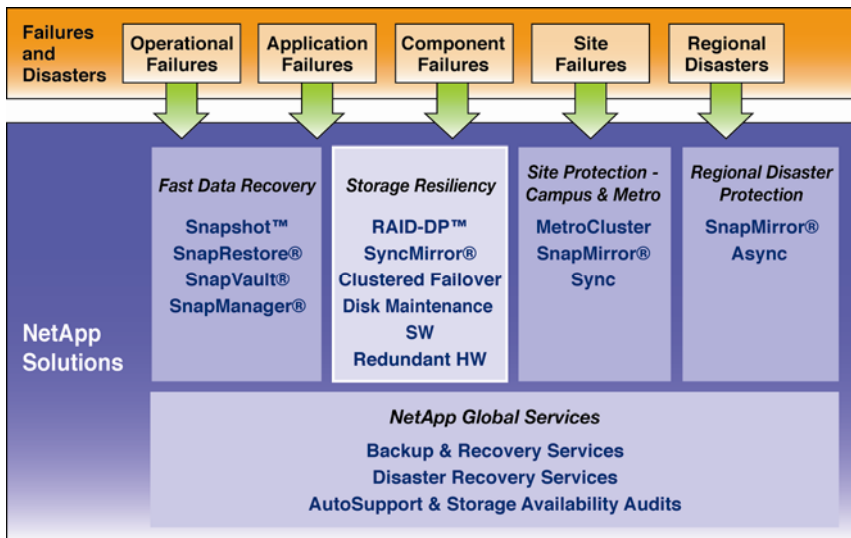


Figure 1) Storage resiliency is an important component of NetApp data availability and disaster recovery solutions

NetApp offers an exceptional and comprehensive portfolio of storage resiliency technologies that help support exceptionally high levels of application availability. To protect against business interruption, storage resiliency is built into every aspect of the NetApp storage architecture. True storage resiliency has two aspects: (1) preventing errors and system failures from happening by means of early detection and self-healing processes and (2) recovering quickly and unobtrusively from errors and system failures when they do happen. Innovations throughout our unified architecture make possible unique storage resiliency technologies that have powerful capabilities without compromising performance, cost-effectiveness, simplicity, or scalability.

2. Preventing and Curing Disk Malfunctions

Disk malfunctions are a reality of life in the data center despite the remarkable dependability of modern drives. Drives can fail suddenly, or they can slowly degrade. Firmware bugs can cause unrecognized data

¹ Gartner Data Center Survey, April 15, 2005

corruption by causing committed writes to be dropped. The resiliency features in Data ONTAP® proactively identify and fix disk issues before they can do harm. In addition to many standard disk maintenance features, such as Checksums, Background Media Scans, Proactive RAID Scrubs, and Rapid RAID Recovery, Data ONTAP includes special technology that not only predicts nascent data integrity and availability problems, but also automatically repairs disk drive errors to prevent unplanned outages and data loss. The innovative features that follow exemplify the unique NetApp approach to proactive disk maintenance.

2.1 Maintenance Center

Maintenance Center is a set of storage resiliency tools that work together to predict and fix drive defects before they can cause problems. Maintenance Center software monitors numerous real-time data points for early indicators of potential drive issues. Suspect drives are immediately removed from use, while other tools run proactive diagnostic tests and heal fixable errors. Healed disks are returned to use, and only those disks that cannot be fixed need to be replaced. When the potential for data loss is detected, Rapid RAID Recovery, a feature of Data ONTAP, proactively copies data to a spare disk before the questionable disk actually fails. This method provides faster reconstruct times and lowers the risk of data loss with minimal impact on performance. By reducing the number of drive failures and nonfunctioning drives that must be replaced, Maintenance Center helps to lower support costs and improve productivity.

2.2 Lost Write Protection

Although it is rare, disks can malfunction during a write operation in such a way that the write fails to reach the intended location: either it is dropped while being written to the physical media or it is lost by being written to a random location. In either case, the disk is unable to detect the failure and signals a successful write status. This event, called a lost write, is particularly insidious because it causes silent data corruption. NetApp Lost Write Protection technology detects and fixes lost writes as data is being written to disk. Only NetApp, with our innovative Data ONTAP 7G technology, can identify this failure as it is happening. Other storage vendors must rely on read-after-write methods, which can significantly degrade performance.

2.3 Momentary Offline Recovery

Sometimes a disk drive which is degrading over time can slide into an unresponsive state. As the disk retries reads and writes, I/O response times increase and raise the risk of timeouts and application downtime. Conventional solutions fail out the unresponsive drive and reconstruct it, a potentially time-consuming and CPU-intensive activity. NetApp offers the unique ability to momentarily offline a drive, resolve the unresponsiveness, and put the drive back in service without any reconstruction taking place and without compromising performance. This innovative disk offline feature—available only from NetApp—ensures the consistent high performance that enterprise applications demand from the storage subsystem.

3. Recovering from Storage System Failures

We provide the highest level of hardware quality. Each component has not only passed a highly rigorous qualification process but is constantly screened during the manufacturing assembly phase. Each new system undergoes a comprehensive set of manufacturing tests for the highest level of quality assurance before being shipped to customers. Despite having one of the lowest defect rates in the industry, NetApp realizes that however low the probability of a single component failure, simple calculations tell us that failures can occur given a large number of hardware components.

3.1 RAID-DP: RAID 6 Without Compromise

The risk of data loss as a result of multiple drive failures is growing as organizations increasingly choose large-capacity Serial ATA (SATA) disk drives, which have higher failure rates than FC drives. The probability of encountering unrecoverable errors during drive reconstruction is also increasing with the adoption of larger capacity SATA and Fibre Channel (FC) drives because many more bits of data are involved in a disk

reconstruction than in the past, and unrecoverable error correction technologies have not kept pace with newer drive capacities. NetApp protects you from these types of failures with our unique implementation of double-parity RAID 6, which we call RAID-DP. With this resiliency technology of Data ONTAP, aggregates and volumes can withstand up to two failed disks in a RAID group. They can also withstand the increasingly common event of a single disk failure followed by an uncorrectable bit-read error from a second disk during reconstruct. RAID-DP dramatically increases data availability without sacrificing cost, performance, or capacity utilization. Only NetApp provides 100% protection against all form of double disk failure cost-effectively and with minimal impact to performance.

3.2 SyncMirror: Synchronous Replication for High Availability

Mission-critical environments require maximum fault tolerance and random read performance. SyncMirror, composed of RAID 1 mirroring and RAID-DP technologies, offers the ultimate storage resiliency by maintaining data availability in spite of triple disk, enclosure, and storage loop failures. Dual RAID mirrors also maximize performance for random read workloads, especially important for mission-critical OLTP database applications. Only SyncMirror can ensure data availability and integrity in the event of mirrored disk failure. SyncMirror works seamlessly in combination with other NetApp technologies to provide enhanced protection for your data.

3.3 Clustered Failover

If a storage controller (e.g., motherboard) fails, NetApp controller failover software automatically initiates a failover that transmits the data service to the partner controller. Clustered Failover (CFO) software manages the takeover and giveback procedures to ensure continuous data availability. The takeover and giveback procedures are simple, fast, and often transparent to end users and applications.

Our unique Multipath HA Storage configuration option for active/active cluster configurations provides quad paths from each controller to each storage loop for maximum availability and performance consistency.

MetroCluster software, in conjunction with SyncMirror and CFO, provides a unified high availability and disaster recovery solution. By permitting the location of the second pair of storage controllers and arrays in a different building or site, MetroCluster software offers protection against site disasters and hardware outages within a campus or metro environment.

3.4 Fully Redundant, Fault-Tolerant Systems

Every NetApp storage system provides industry-standard hardware redundancy. All system components are 100% fully redundant and fault tolerant to avoid single points of failure (SPOFs) to help exceptionally high levels of application availability.

3.4.1 Storage Controller

Each system offers active/active storage controller configurations to avoid controller SPOF. If one controller fails, it will automatically fail over to its partner controller to maintain application availability. Dual redundant Infiniband cluster interconnect paths provide fault-tolerant connectivity between each controller.

3.4.2 NVRAM

Each controller uses nonvolatile random access memory (NVRAM) as a write cache for improved performance. During each write, data is instantly mirrored to its partner's NVRAM. In the event of failure, the mirrored NVRAM can automatically flush writes to disk to avoid data loss. In addition, each NVRAM has battery backup to ensure that data stored in the write cache will not be lost in the event of a power failure.

3.4.3 Power/Cooling

Each controller and disk enclosure offers redundant power supply units and cooling fans. In the event of a power supply or fan failure, the redundant power supply unit will take over the power and cooling load to ensure system availability.

3.4.4 Cabling

All cabling is fully redundant, including all disk enclosure cables and Infiniband cluster interconnect cables. Quad pathing is provided for active/active systems in a Multipath HA Storage configuration for enhanced availability and performance consistency. In the event of a cable fault, alternate paths ensure continuous data access.

3.4.5 I/O Modules

In addition to redundant power, fans, and cabling, disk enclosures also have fully redundant I/O modules. In the event of an I/O module failure (due to chip failure or inadvertent user pull), the redundant module takes over to maintain data availability. The ESH2 I/O controller module in the disk shelf employs an embedded switch architecture for point-to-point fault isolation of FC disk drives. This second-generation embedded switched hub maximizes fault tolerance for highly available data center environments.

3.4.6 Disk Drives

All disk drives are protected by RAID software (RAID 4, RAID-DP, SyncMirror). In the rare instance that a drive fails, it can be easily hot-replaced. Disk shelves with empty slots allow hot additions of disk drives in any of the empty drive bays. In addition, hot disk spares are available for drive reconstruction (or Rapid RAID Recovery) to ensure prompt recovery and minimal performance overhead.

4. Conclusion

Enterprise data centers running mission-critical applications require highly resilient storage systems that support high levels of application availability and consistent data integrity. They need to avoid unplanned system downtime to ensure application data is always available, and they need to avoid data corruption to ensure that application data is correct and up to date. Comprehensive NetApp storage resiliency offerings, including Data ONTAP resiliency software and advanced, fully redundant hardware systems, address the two key data management challenges enterprises face today. The NetApp approach to preventing, curing, and recovering from storage system failures overcome the constant challenge to maintain maximum data availability and integrity, without compromising performance, cost, simplicity, or scalability.

For a more comprehensive understanding of how NetApp can help you implement a resilient storage infrastructure—without compromise—please refer to the companion publication, “A Comprehensive Approach to Application Availability.”

About Network Appliance

Network Appliance is a world leader in unified storage solutions for today’s data-intensive enterprises. Since its inception in 1992, Network Appliance has delivered technology, product, and partner firsts that simplify data management. Information about Network Appliance™ solutions and services is available at www.netapp.com.

