NetApp
Go further, faster

Technical Report

Storage Subsystem Resiliency Guide

Jay White, NetApp
September 2010 | TR-3437

ABSTRACT

This document provides technical recommendations and best practices as they relate to data availability and resiliency in the NetApp® storage subsystem. The topics addressed in this document are important to understand when planning and architecting a NetApp storage environment that will meet customer needs and expectations.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

The predecessor to this document approached storage resiliency with a focus on key features and options that allowed you to configure your storage system for maximum storage resiliency within the scope of a predefined set of data availability tiers. This document looks to build further upon that foundation. It is not always possible or even necessary to configure systems for maximum resiliency depending on the purpose and requirements of a given storage configuration. Furthermore, the end objective of any storage configuration is not necessarily to make sure of storage resiliency but rather to make sure of data availability. How resilient is a system that experiences a failure that incurs such a performance impact to the system as a whole that applications depending upon the storage system stop functioning, even if the system is technically still responding for foreground I/O?

As a result of situations such as the one just described, simply focusing on resiliency is not enough. Resiliency must be approached with data availability in mind and how it affects the system as a whole.

1.1 DATA AVAILABILITY

A core measurement of a NetApp storage system is data availability. For the purposes of this document, data availability is assessed based on three factors:

- **Performance:** Performance can be broken down to two key perspectives from a data availability point of view. The first is that customers will have specific performance requirements that are necessary to meet in order to satisfy applications that are dependent on the storage system data being readily available. A data availability outage from this perspective means the system might still be responding to foreground I/O, but it might have fallen below the requirements of the dependent applications' ability to function. The second is that if a system's performance suffers to the extent that the system stops responding completely to foreground I/O, then a data availability outage situation has been encountered.
- **Resiliency:** Resiliency from the point of view of data availability is the system's ability to suffer a single failure or multiple failures, while continuing to respond to foreground I/O in the degraded state. There are multiple options and features that contribute to a system's ability to suffer failures, which are discussed throughout this document.
- **Recoverability:** Recoverability defines the system's ability to both automatically recover from failures and continue to respond to foreground I/O while conducting recovery operations on the storage system.

These three factors are further applied to the three layers of data availability:

- **Storage subsystem:** The storage subsystem layer addresses all hardware components and software features that relate to the storage system's internals. Primarily this can be considered to be from the HBA down through the attached storage arrays from a physical perspective, or around the storage and RAID software layers that are a part of Data ONTAP®: in short, the system's ability to communicate internally from the controller to the attached storage arrays.
- **System:** The system layer addresses the ability of a storage system to suffer failures. This is primarily focused on controller-level failures that affect the ability of a system to continue external communication. This applies to single controller and HA pair configurations and the components that contribute to external controller communication such as network interfaces.
- **Site:** The site layer addresses the ability of a group of collocated storage systems to suffer failures. This is primarily focused on the features related to distributed storage system architecture that allow for an entire storage system failure, which would likely be related to a site-level incident, such as a natural disaster or act of terrorism.

To further quantify best practices and recommendations, this information must be applied to a set of defined tiers of data availability. This is necessary as, for example, you are not able to accomplish maximum resiliency and recoverability without sacrificing performance to a certain extent.

Depending on the requirements of each tier of data availability, the balance between the three factors outlined above change, which results in different best practices and recommendations between the defined tiers of data availability. The tiers of data availability and the recommendations for each are covered in section 6, "Data Availability Tiers," in this document.

1.2 SCOPE

The primary scope of this document is to address the storage subsystem layer of data availability while taking into account the three factors (performance, resiliency, and recoverability) outlined in section 1.1, “Data Availability.”

2 RELIABILITY

The most common measure of reliability that is publically stated in the industry today is mean time between failures (MTBF). The problem is that MTBF is not nearly as accurate a measure of reliability as average return rate (ARR) or average failure rate (AFR), both of which are tracked by companies but in several cases not made publically available.

NetApp does track ARR and AFR for critical storage components and can provide data to customers that are under NDA with NetApp if a request is filed with the appropriate product management contact. NetApp is not required to provide this data, and it is completely within the discretion of the product management team as to how much data, if any, is provided upon request.

2.1 MEASURING RELIABILITY

There are generally three reliability measures available for hardware components today. They are mean time between failures, average return rate, and average failure rate. These measures are discussed in detail below, but to summarize the key points to take away from this section:

- The expected operating life of an enterprise drive is five years. NetApp highly recommends replacing drives that are older than five years. This also aligns with the standard five-year warranty for drives.
- The more drives you have in your configuration, the more likely you are to encounter drive failures within the time that the drives are in use.
- MTBF is the least accurate measure of reliability.
- AFR is the best measure of reliability but takes time to establish an accurate data set.

This section primarily focuses on drives, but the same methods and information are applied to the other devices present in the storage subsystem and beyond.

MEAN TIME BETWEEN FAILURES

MTBF is the least accurate measure of reliability. MTBF is commonly misunderstood to be the useful life of a hardware device. Since hardware manufacturers can’t reasonably test devices for their entire expected life before release, they test many devices in an attempt to extract what the failure rate should be during the expected life of the device. The formula most commonly used is as follows:

test duration * # of drives tested / # of drives that failed during testing = MTBF

Figure 1) MTBF formula.

The most commonly referenced MTBF values for storage subsystem devices are for drives. SATA, SAS, and FC drives have different MTBF values as follows:

- SAS and FC drives are 1.6 million hours
- SATA drives are 1.2 million hours

The standard drive warranty is five years (43,800 hours), which is far short of 1.6 million or even 1.2 million hours. Again, MTBF is a measure not of the usable life of the drive but rather of the error rate within the useful drive life.

Purely based on MTBF, the math suggests that for SATA drives (1.2 million hours MTBF) ~0.73% of your deployed drives should fail each year. For FC and SAS drives (1.6 million hours MTBF) ~0.55% of your deployed drives should fail each year.

SATA 1,200,000 hours MTBF / 8,760 hours per year = 136.9863 years

1 failure / 136.9863 years = 0.00730 * 100 = 0.73%

FC and SAS 1,600,000 hours MTBF / 8,760 hours per year = 182.6484 years

1 failure / 182.6484 years = 0.00547 * 100 = 0.55%

Figure 2) Calculating drive failure rate based on MTBF.

To apply this further, let's consider the following two example configurations:

- 300 SAS drives with an expected use of five years
- 36 SAS drives with an expected use of five years

Math can now be applied to determine how many failures would be expected to occur over the operating life of these configurations:

300 SAS drives * 0.55% (0.0055) = 1.65 failures per year * 5 years = 8.25 failures within 5 years

36 SAS drives * 0.55% (0.0055) = 0.198 failures per year * 5 years = 0.99 failures within 5 years

Figure 3) Expected failures within operating life based on MTBF and number of drives.

The primary point to take from all this is that the more drives you have, the more likely it is that one of those drives will fail within the time they are in use. Based on the standard five-year warranty that is applied to enterprise drives today, it is safe to state that the expected reliable life of a drive is five years, after which the likelihood that a drive will fail increases significantly the longer the drive stays in use.

AVERAGE RETURN RATE

The ARR of a device is a better measure of reliability than MTBF as it is based on the actual return rate of a device from systems that are in service and actually using the device. Unfortunately this is still not the best measure of reliability as it does not distinguish between devices that have been returned for reasons that are not associated with failures. Some examples of returns unrelated to actual failures include drives that are returned due to false positives, as a precautionary measure, or because of a mistaken shipment. Although not the best method for determining reliability, it is useful for companies to track this to understand if there are issues with operational efficiency, usability, or other business-related reasons.

AVERAGE FAILURE RATE

This is the most accurate measure of device reliability as it is based on devices that have been returned and verified to have failed. Unfortunately it takes time to establish AFR as it is based on an average over time. As a result AFR becomes more accurate as time progresses. Devices can fail for a multitude of reasons, of which some are discussed later in this document.

The purpose of this document is not to address what the ARR or AFR is for the various devices shipped by NetApp but rather to explain and put in context the measures that are either publically available or available upon request.

2.2 SYSTEM RELIABILITY

Many ask what the MTBF is for a controller or storage shelf. There are several reasons why MTBF is not published for collections of devices:

- The MTBF calculation is based on the usage of a single device or group of integrated devices. Controllers and storage shelves contain several components that are optional (expansion cards, shelf modules, and more) or can be different on a case-by-case basis.
- An MTBF value has to take into account all components in use, but with controllers and storage shelves not all components are critical. For example, if an LED fails on a storage shelf, the shelf will continue to operate in its primary role of providing access to drives.

- As stated in section 2.1, “Measuring Reliability,” MTBF is the least accurate measure of reliability. Adding additional devices further dilutes an already abstracted calculation and result.

Between storage shelves, shelf modules, and drives, the least reliable components of the storage subsystem are generally considered to be the drives. This does not mean that storage shelves and shelf modules are more reliable than drives. The logic behind this is as follows:

- There are many more drives present in a storage shelf than any other device. For example, a DS4243 has two to four PSUs, two IOM3 shelf modules, one shelf enclosure, and 24 drives.
- Drives contain just as much electronics and sophistication as the other components with the added factor that they contain moving parts.

As a result of this thinking, when storage subsystem reliability is discussed, it normally revolves around drives.

2.3 RELIABILITY BEST PRACTICES

Some key best practices to follow when attempting to maximize storage subsystem component reliability are as follows:

- Remove failed hardware components quickly to make sure that failures don’t propagate to healthy components in the system.
- Replace or retire hardware components that have exceeded their warranty period.
- Follow safe practices when handling hardware components to protect against physical damage and electrostatic discharge (ESD) damage.
- Understand that failures are a fact of life with technology and make sure spares for critical components are readily available. This means following best practices for hot and cold spares and understanding the parts turnaround for your site.
- The use of cold spares does not replace the need for hot spares. The longer components sit on a shelf, the more likely it is they might suffer physical damage or simply just not work. A drive installed and working in a storage system (hot spare) is in a state that results in a high reliability that it is ready to fulfill its role as a drive replacement.

3 ERRORS AND FAILURES

This section provides additional details around some key errors and failures that can occur in the storage subsystem. This is not meant to be inclusive of every possible error for failure that can occur; rather the focus is on conditions that affect system resiliency operations such as RAID reconstruction. Single points of failure (SPOFs) are also discussed as they affect system resiliency.

NetApp highly recommends removing failed hardware components from an active system as soon as possible to reduce the risk of the failure propagating to additional components within the system.

3.1 SINGLE POINTS OF FAILURE

Some potential SPOFs are eliminated by native system configurations. For example, every NetApp storage shelf uses more than a single shelf module, power supply unit, and drive. Other SPOFs might exist depending on what the system configuration selected is:

- **Controller:** NetApp does support single controller configurations, in which case the controller itself is a SPOF. Using an HA pair storage configuration that includes two controllers eliminates the controller as a SPOF.
- **Host bus adapter (HBA):** This includes onboard ports and separate HBAs, which are referred to as port groups. A port group is any set of interconnected ports. For example, onboard ports A and B might use a different ASIC than ports C and D, but they both depend upon the system motherboard to be able to function. A single quad-port HBA generally has two ASICs as well, but the HBA itself is a SPOF. As a result it is generally recommended to connect your storage loops (FC-AL) and stacks (SAS) to more than one port group. For example, this could be two HBAs or a combination of onboard ports and one or more HBAs. NetApp always recommends at a minimum that connections are split across ASICs.

- **Cables:** There are many types of cables used to connect a storage system together. Some cables are more resilient to physical damage than others; for example, optical cables are much more susceptible to physical damage than Ethernet cables. To avoid cables as a SPOF in your storage configuration, NetApp recommends (and in many cases now requires) the use of multipath high-availability (MPHA) cabling. MPHA provides secondary path connections to all storage shelves attached to the system.
- **Shelf enclosure:** Although complete shelf enclosure failures are very rare, they are a possibility. An approach used in the field to protect against this situation has been to make sure that no more than two drives from any RAID group are located on a single shelf (assuming RAID-DP®). This approach is not recommended as a shelf resiliency solution. The recommended method for protecting against shelf failures is to use local SyncMirror® or other mirroring methods to make sure data is quickly made available in a failure situation. Mirroring solutions also account for multiple failure situations.

3.2 DRIVES

Errors and failures associated with drives are very complex. As a result, many misconceptions exist around the types of failures that occur and how they are resolved.

Under some circumstances the perception could be that NetApp storage systems aggressively fail drives for reasons that are not always perceived to be critical. For example, after a single bad block is detected, NetApp could fail a drive, which might seem extreme. The term *bad block* is very generic. In reality the drive is returning an error code that is associated with an unsuccessful drive operation, and that error can indicate that there is a serious problem. Depending on the significance of the error returned from the drive, it could indicate that other blocks on the drive are also likely compromised. In this situation it is safer to fail the drive and remove it from the active file system so data is not compromised further.

The following five conditions will generally result in a drive being failed by the system and corrective action being initiated:

- The drive itself returns a fatal error.
- The storage layer of Data ONTAP reports that the drive is inaccessible.
- The drive returns a recommendation to Data ONTAP that the drive should be failed.
- The storage and RAID layer of Data ONTAP recommends a drive should be failed based on various error thresholds being exceeded by the drive.
- Lost write protection (LWP) occurs.

4 CORRECTIVE ACTIONS AND PREVENTIVE FEATURES

When issues are encountered, Data ONTAP will check the current RAID state and error condition. This results in one of three possible resulting actions:

- Initiate a RAID reconstruction
- Initiate a Rapid RAID Recovery (could also result in the use of Maintenance Center)
- Ignore the error

RAID reconstruction and Rapid RAID Recovery are discussed in more detail later. Errors are only potentially ignored for RAID groups that are already in a degraded state. This is because Data ONTAP already understands there are issues present and likely is in the process of resolving the degraded state.

Errors not associated with drive failures, normally detected by preventive actions such as RAID scrubs, can result in one of the following two actions:

- Rewrite the suspect data block to a new block (data block repair)
- Rewrite parity data for the block (parity repair)

Knowing that Data ONTAP conducts data block repair and parity repair is sufficient enough for the scope of this document as these operations are not specific to drive failures but rather are issues with individual data blocks in the file system. The key point is that Data ONTAP is taking several steps to make sure of data integrity, and those steps do not always result in drives being failed.

4.1 RAID RECONSTRUCTIONS

When a drive is failed and a RAID reconstruction initiated, several factors will determine how long the reconstruction process will take and how the system's performance will be affected as a result. Some of the factors that will contribute to system performance while operating in a degraded mode are as follows:

- System workload profile (random/sequential and read/write mixes)
- Current CPU and I/O bandwidth utilization
- RAID group size
- Storage shelf and shelf module technology in use
- Type of drives (SSD, SATA, FC, or SAS)
- RAID option settings
- Drive path assignments
- Distribution of drives across stacks/loops
- Single or double drive failure and reconstruction

Because of these factors, it is very difficult to accurately predict the impact to a storage system.

Once a drive is failed, all I/O that would normally be directed at the drive is redirected to the replacement drive. Reconstruction traffic will affect all drives in the degraded RAID group as reads will be occurring on all data drives in the RAID group. Additional bandwidth is needed on stacks/loops containing the degraded RAID group and the replacement drive. RAID reconstruction I/O will compete with foreground I/O within the confines of current system utilization and RAID options settings. This is discussed in greater depth in section 5.2, "RAID Options," in this document.

SINGLE DRIVE RECONSTRUCTION

A single drive reconstruction occurring in a RAID-DP RAID group results in data being reconstructed much like any single parity drive RAID group (double parity information is not needed). A reconstruction involves reads from all remaining drives in the RAID group and the parity drive.

A single reconstruction effectively doubles the I/O occurring on the stack/loop, as for each foreground I/O directed toward the RAID group, the data needs to be reconstructed on demand for the failed drive. This traffic is in addition to the reconstruction traffic associated with parity calculations and writes to the replacement drive.

DOUBLE DRIVE RECONSTRUCTION

A double drive reconstruction occurring in a RAID-DP group results in data being reconstructed from both single parity and double parity data. This type of reconstruction involves reads from all remaining data drives in the RAID group in addition to the single and double parity drives. Stack bandwidth requirements for foreground I/O triple in this case. Data ONTAP is intelligent enough not to require multiple reads in order to conduct both parity and double parity data reconstruction calculations; a single read operation is sufficient to do both calculations.

A double reconstruction effectively triples the I/O occurring on the stack/loop, as for each foreground I/O directed toward the RAID group the data needs to be reconstructed on demand for the two failed drives. This traffic is in addition to the reconstruction traffic associated with parity calculations and writes to the replacement drive.

4.2 RAPID RAID RECOVERY

A Rapid RAID Recovery is similar to a RAID reconstruction but without the need to reconstruct data from parity as the drive is still accessible. Some blocks on the drive might need to be reconstructed from parity data, but the majority of the drive will be copied at a block level to the replacement drive. As this is a block-level copy, all blocks are copied regardless of how full the drive might be.

A Rapid RAID Recovery does increase the I/O occurring on the stack/loop due to the read and write traffic occurring between the failing drive and the replacement drive. However, the impact to the remaining drives in the RAID group is far less than with reconstruction as parity calculations are not needed for most if not all of the data on the failing drive.

4.3 MAINTENANCE CENTER

When enabled on a system, Maintenance Center works in conjunction with Rapid RAID Recovery to assess the condition of failed drives prior to them being returned to NetApp. When a drive enters Maintenance Center, a Rapid RAID Recovery is initiated, failing the drive out of the RAID group. The failed drive is then assessed by Data ONTAP by running drive diagnostics. If the drive is deemed to be functional, it is returned to the systems spare pool. If the drive is not functional, it remains failed and needs to be replaced.

Maintenance Center requires a minimum of two hot spares available on the system, and Rapid RAID Recovery (`raid.disk.copy.auto.enable`) must be enabled. NetApp recommends setting the option `raid.min_spare_count` to 2 in order to allow the system to notify the administrator when Maintenance Center requirements are not met.

4.4 LOST WRITE PROTECTION

Lost write protection is a feature of Data ONTAP that occurs on each WAFL® read. Data is checked against block checksum information (WAFL context) and RAID parity data. If an issue is detected, there are two possible outcomes:

- The drive containing the data is failed.
- The aggregate containing the data is marked inconsistent.

If an aggregate is marked inconsistent, it will require the use of WAFL iron to be able to return the aggregate to a consistent state. If a drive is failed, it is subject to the same corrective actions as any failed drive in the system.

It is a rare occurrence for lost write protection to find an issue. Its primary purpose is to detect what are generally the most complex or edge case problems that might occur and determine the best course of action to take in order to protect data integrity.

4.5 BACKGROUND MEDIA SCANS

Background media scans are a drive diagnostic feature that is run continuously on all RAID group drives. This type of scrub (media scrub) is used to detect media errors. The purpose is not to make sure the data block is integral from the point of view of the file system but rather to make sure blocks on the drive are accessible.

System performance is affected by less than 4% on average. This is primarily because the drive internals are conducting the actual scans, which do not require CPU or I/O bandwidth from the system.

4.6 RAID PARITY SCRUBS

RAID parity scrubs are used to make sure of the integrity of data at rest. For very active data sets, the benefit of RAID parity scrubs is limited as the data is being read often, and thus Data ONTAP is making sure of data integrity through other means. The most common data that is at rest is archive data. RAID parity scrubs offer the best return when used with this type of data.

This process traverses data at rest and triggers reads on that data. As a result of triggering the read, the data is checked against parity to determine that it is correct. If a block is found to be incorrect, the block is marked as bad, and the data is recreated from parity and written to a new block. RAID scrubs minimally affect foreground I/O, and data suggests this impact is less than 10% on average. For large archival data sets, NetApp recommends increasing the frequency of RAID parity scrubs.

RAID parity scrubs are enabled by default, and by default a scrub will run for 360 minutes (six hours). The performance impact is set to low by default, which results in the process only using idle system resources.

5 ADDITIONAL CONSIDERATIONS

In addition to specific resiliency features and corrective actions, there are considerations that are important to understand when configuring your storage system. This section focuses on specific configuration factors such as RAID group size, RAID options, and best practices around mixed configurations.

5.1 RAID GROUPS

RAID group configuration can greatly affect a storage system's resiliency. NetApp highly recommends using RAID-DP for all storage configurations, as it offers the best resiliency features and enables nondisruptive background firmware updates for drives. The best practices and points discussed in this section assume the use of RAID-DP.

It is tempting to always create the largest RAID groups in an aggregate to minimize parity tax and maximize performance, but the implications of doing such are:

- **Larger failure domains:** The more drives you have in a RAID group, the more likely it is that one or more of those drives will fail in the course of the operational lifetime of the storage system. Drive reliability is a primary factor in attempting to understand the risk of encountering multiple drive failures (MDFs) within a single RAID group. Ultimately any calculation is a guess as there is no guarantee that drives will fail all at the same time, in the same RAID group, or fail at all (within the five-year warranty period).
- **Increased drive reconstruction times:** The more data drives that are present in the RAID group, the greater the calculation overhead for reconstructing data from parity. Each data drive contributes a data point that needs to be considered in the parity calculations. The more data points, the larger the parity calculation is, and as a result the reconstruction times increase. This increase is small (0% to 4%) in RAID groups of size 12 to 20, which is one of the factors behind the RAID group sizing policy outlined in TR-3838, "Storage Subsystem Configuration Guide" (available on Field Portal). Significant increases in RAID reconstruction times have been noted in RAID groups of size 21 to 28. At RAID group size 28, an increase of 2% to -28% has been noted in RAID reconstruction times for nonidle systems.

SATA VERSUS FC/SAS

Many consider SATA drives (1.2 million hours MTBF) to be less reliable that FC and SAS drives (1.6 million hours MTBF). NetApp's AFR and ARR data suggests that enterprise SATA drives are as reliable in actual deployments as FC and SAS drives, which leads to the question "Then why is the maximum SATA RAID group size less than SSD, FC, or SAS?"

Although the reliability might be similar, the capacity and speed differences cannot be overlooked. SATA drives are of larger capacity and slower than FC/SAS drives, which means they take a significantly longer time to reconstruct than FC/SAS drives. In Data ONTAP 8.0.1 the maximum SATA RAID group size has been increased from 16 to 20. This change was decided after analyzing field data and based on the better-than-expected reliability results (among other factors). After RAID group size 20, however (size 21 to 28), an inflection point was seen in the risk of multiple drive failures due to perpetual drive reconstructions.

Perpetual drive reconstructions occur when the reconstruction time for drive is so long that it greatly increases the probability of encountering another drive failure before completing the current reconstruction activity. This is normally only a risk with large (>1TB) SATA drives. The risk is currently only increasing as we see larger and larger SATA drives coming to market (3TB, 4TB, and larger).

SOLID STATE DRIVE

Given the small capacity points of SSD and the significantly better drive-level performance, the use of larger SSD-based RAID groups is less risky. Data shows that RAID reconstruction of a 100GB SSD is <20 minutes on a loaded system. Given this fast reconstruct time, it is reasonable to expect sufficient system resiliency even at the largest RAID group size of 28 (assuming RAID-DP).

5.2 RAID OPTIONS

Within the scope of data availability, it is important to understand how you can tune a storage configuration in order to make sure data availability requirements are met. For example, in an archival storage configuration, it is better to tune the system to allow reconstruction I/O to compete efficiently with foreground I/O, whereas in an Exchange configuration, it might be necessary to make sure the foreground I/O competes for system resource more effectively than reconstruction I/O.

RAID options are the primary user-configurable method for telling Data ONTAP how foreground I/O and corrective I/O (RAID reconstruction I/O and Rapid RAID Recovery I/O) should compete for system resources.

The option `raid.reconstruct.perf_impact` can be set to low, medium, or high. By default it is set to medium. Changing this option results in the following behavior:

- **Low:** This allows corrective I/O and foreground I/O to compete with 0% of system resources during peak controller performance. This effectively guarantees foreground I/O can consume 100% of system resources when capable of doing so. Corrective I/O will only use idle system resources.
- **Medium:** This allows corrective I/O and foreground I/O to compete with 40% of system resources during peak controller performance. This effectively guarantees foreground I/O can consume 60% of system resources without interference from corrective I/O.
- **High:** This allows corrective I/O and foreground I/O to compete with 90% of system resources during peak controller performance. This effectively guarantees foreground I/O can consume 10% of system resources without interference from corrective I/O.

For the purposes of corrective I/O the term system resources refers to:

- CPU
- I/O bandwidth
- Drive utilization

There is no limit on the amount of idle CPU and I/O bandwidth that corrective I/O can consume, hence 0% means that only background processing will be allocated when the system is under load. This also means that this option makes very little difference on idle systems as there is no foreground I/O with which to compete.

The percentages listed are not guaranteeing that corrective I/O will consume that much, rather only that foreground I/O and corrective I/O will compete for system resources within those percentages. A setting of high does not mean you will see a 90% impact to foreground I/O, as both foreground and corrective I/0 are still occurring within that percentage. Additionally, corrective actions might be able to compete with foreground I/O up to those percentages, but that does not mean the corrective actions are demanding that percentage of system resources.

The hidden option `raid.reconstruct.threads` can be set to a value from 1 through 8. The default value is 4. This option should not be changed without first receiving guidance from the NetApp Product Management or Engineering teams. This option controls how efficiently corrective I/O can compete with foreground I/O for system resources within the confines of the `raid.reconstruct.perf_impact` option. Increasing the threads allows corrective I/O to compete more efficiently (using more threads), whereas decreasing the threads will make corrective I/O compete less efficiently for system resources.

Another factor that is important to consider is if the application(s) using the storage are single threaded or multithreaded. A single-threaded application will be less efficient at driving foreground I/O than a multithreaded application. As a result, corrective I/O can also be advantaged or disadvantaged. The default setting for the `raid.reconstruct.threads` option assumes a multithreaded application is driving foreground I/O on the system. If your application is single threaded and sensitive to latency, this could warrant adjusting the `raid.reconstruct.threads` option to a lower value.

5.3 SPARES POLICY

Spares recommendations vary by configuration and situation. In the past NetApp has based spares recommendations purely on the number of drives attached to a system. This is certainly an important factor but not the only consideration. NetApp storage systems are deployed in a wide breadth of configurations. This warrants defining more than a single approach to determining the appropriate number of spares to maintain in your storage configuration.

Depending on the requirements of your storage configuration, you can choose to tune your spares policy toward:

- **Minimum spares:** In configurations where drive capacity utilization is a key concern, the desire might be to use only the minimum number of spares. This option allows you to survive the most basic failures. If multiple failures occur, it might be necessary to manually intervene to make sure of continued data integrity.

- **Balanced spares:** The configuration is the middle ground between minimum and maximum. This assumes you will not encounter the worst case scenario and will provide sufficient spares to handle most failure scenarios.
- **Maximum spares:** This option makes sure enough spares are on hand to handle a failure situation that would demand the maximum number of spares that could be consumed by a system at a single time. Using the term maximum is not stating the system might not operate with more than this recommended number of spares. You can always add additional hot spares within spindle limits as you deem appropriate.

Selecting any one of these approaches is considered to be the best practice recommendation within the scope of your system requirements. The majority of storage architects will likely choose the balanced approach, although customers who are extremely sensitive to data integrity might warrant taking a maximum spares approach. Given that entry platforms use small numbers of disks, a minimum spares approach would be reasonable for those configurations.

For RAID-DP configurations, consult Table 1 for the recommended number of spares.

Table 1) Determining recommended spares.

Recommended Spares		
Minimum	**Balanced**	**Maximum**
Two per Controller	**Four per Controller**	**Six per Controller**
Special Considerations		
Entry Platforms	Entry-level platforms using only internal drives can be reduced to using a minimum of one hot spare.	
RAID Groups	Systems containing only a single RAID group do not warrant maintaining more than two hot spares for the system.	
Maintenance Center	Maintenance Center requires a minimum of two spares to be present in the system.	
>48-Hour Lead Time	For remotely located systems there is an increased chance they might encounter multiple failures and completed reconstructions before manual intervention can occur. Spares recommendations should be doubled for these systems.	
>1,200 Drives	For systems using greater than 1,200 drives an additional two hot spares should be added to the recommendations for all three approaches.	
<300 Drives	For systems using less than 300 drives you can reduce spares recommendations for a balanced and maximum approach by two.	

Additional notes regarding hot spares:

- Spares recommendations are for each drive type installed in the system. See section 5.4, “Mixed Configurations,” for more information.
- Larger capacity drives can serve as spares for smaller capacity drives (they will be downsized).
- Slower drives replacing faster drives of the same type will affect RAID group and aggregate performance. For example, if a 10k rpm SAS drive (DS2246) replaces a 15k rpm SAS drive (DS4243), this results in a nonoptimal configuration.
- Although FC and SAS drives are equivalent from a performance perspective, the resiliency features of the storage shelves in which they are offered are very different. By default Data ONTAP will use FC and SAS drives interchangeably. This can be prevented by setting the RAID option `raid.disk.type.enable` to on. See section 5.4, “Mixed Configurations,” for more information.

HOT AND COLD SPARES

NetApp does not discourage administrators from keeping cold spares on hand. NetApp recommends removing a failed drive from a system as soon as possible, and keeping cold spares on hand can speed the replacement process for those failed drives. However, cold spares are not a replacement for keeping hot spares installed in a system.

Hot spares are also present to replace failed drives, but in a different way. Cold spares can replace a failed part (speeding the return/replace process), but hot spares serve a different purpose. That is to respond in real time to drive failures by providing a target drive for RAID reconstruction or Rapid RAID Recovery actions. It is hard to imagine an administrator running into a lab to plug in a cold spare when a drive fails. Cold spares are also at greater risk of being "dead on replacement" as drives are subjected to the increased possibility of physical damage when not installed in a system. For example, handling damage from electrostatic discharge is a form of physical damage that can occur when retrieving a drive to install in a system.

Given the different purpose of cold spares versus hot spares, you should never consider cold spares as a substitute for maintaining hot spares in your storage configuration.

ENFORCING MINIMUM SPARES

The RAID option `raid.min_spare_count` can be used to specify the minimum number of spares that should be available in the system. This is effective for Maintenance Center users as when set to the value 2 it effectively notifies the administrator if the system falls out of Maintenance Center compliance. NetApp recommends setting this value to the resulting number of spares that you should be maintaining for your system (based on this spares policy) so the system will notify you when you have fallen below the recommended number of spares.

5.4 MIXED CONFIGURATIONS

The ability to create mixed configurations with NetApp storage solutions is a significant benefit for many customers. The purpose of this section is not to generically dissuade the use of mixed configurations but to identify that as technology changes or is introduced, there is a need to assess and reassess mixed configurations to make sure resiliency and/or performance has not been unintentionally compromised. This is not to say that simply by creating a mixed configuration you have compromised resiliency, as there are several mixed configurations supported today that offer the same resiliency level as the equivalent segregated configurations.

SHELF TECHNOLOGY

As NetApp transitions from the DS14 storage shelf family to the SAS storage shelf family, it is common to see both shelf technologies attached to the same system. The SAS storage shelf family has new and unique resiliency features that are not available with the DS14 storage shelf family. For example, Alternate Control Path (ACP) is a feature only available with the SAS storage shelf family.

NetApp recommends segregating logical system configuration between DS14 and SAS storage shelf technologies.

FC AND SAS EQUIVALENCY

The option `raid.disktype.enable` is off by default. This means that for the purposes of aggregate creation and spares selection, Data ONTAP will treat FC and SAS drives the same. For example, a SAS drive can serve as a replacement for a failed FC drive and vice versa. Although from a performance perspective FC and SAS drives are equivalent, they are different from a resiliency perspective, as the storage shelves in which those drives are available are very different. FC is available in DS14mk2 and Ds14mk4 shelves using ESH2 and ESH4 shelf modules. SAS is available in DS4243 using IOM3 shelf modules and DS2246 using IOM6 shelf modules. The SAS shelf family has improved resiliency features over the DS14 family. For example, if a DS14-based drive replaces a drive that was part of a RAID group contained completely within a SAS shelf, the resiliency level of that RAID group has effectively dropped when considered as a whole.

NetApp recommends setting the option `raid.disktype.enable` to on in order to enforce the separation of FC and SAS drives.

DRIVE SPEED

With the introduction of the DS2246 storage shelf, we see the availability of 10k rpm 2.5" SAS drives at a time when DS4243 is offering 15k 3.5" SAS drives. 10k rpm FC drives are still around in customer configurations even though they are no longer available from NetApp. Mixing 10k rpm drives with 15k rpm drives in the same aggregate effectively throttles all drives down to 10k rpm. This results in longer times for corrective actions such as RAID reconstructions.

NetApp recommends that you do not mix 10k rpm and 15 rpm drives within the same aggregate.

5.5 MY AUTOSUPPORT SYSTEM RISKS

For customers who use My AutoSupport™, there is a newly expanded Health Summary section that includes a feature known as System Risk Details (SRD). The SRD section proactively identifies risks in deployed NetApp storage configurations that can negatively affect system performance, availability, and resiliency. Each risk entry contains information about the specific risk to the system, potential negative effects, and links to risk mitigation plans. By addressing identified risks proactively, you can significantly reduce the possibility of unplanned downtime for your NetApp storage system.

NetApp recommends using the SRD section to increase system resiliency by addressing system risks before they lead to unplanned downtime. More information can be found on the NOW™ (NetApp on the Web) site on the main My AutoSupport page located at https://now.netapp.com/NOW/asuphome.

6 DATA AVAILABILITY TIERS

This section covers the resiliency requirements and recommendations as they relate to the data availability tiers defined within this document. The majority of this document can be considered general best practices that are application to all data availability tiers.

DEFAULT VALUES

It is important to note that default values do not always equate to best practices. The reality of default values is that they are normally subject to one of the following factors:

- They represent the average or middle ground, neither optimized nor not optimized.
- Their values might have been determined years ago and might have not been reassessed for currency with today's solutions and features.
- They can be determined to be generically applicable settings or recommendations for the majority of known configurations.

Given the breadth and depth of the storage solutions offered by NetApp, it is almost impossible to make sure all default values align with best practices. There is no single answer in many circumstances, which means due diligence must be exercised in order to make sure your storage configuration is optimized appropriately for your customers.

6.1 TIER 1: MISSION CRITICAL

These types of environments are enabling services that are in high demand and cost the customer significant loss of revenue when an outage occurs. Online transaction processing (OLTP), batch transaction processing, and virtualization/cloud environments are examples of environments that fit into this tier of data availability. This tier of data availability is tuned toward prioritizing I/O response to foreground (client application) traffic to make sure dependent applications remain functional. By prioritizing foreground I/O over corrective I/O in degraded situations, you increase the time necessary to complete corrective actions. This increases the risk of encountering additional failures in the system before completing a corrective action: for example, encountering an additional drive failure before an existing reconstruction operation can complete.

Table 2) Recommendations and best practices for mission-critical data availability.

Mission-Critical Recommendations	
Flash Cache	Use Flash Cache to improve system performance and minimize the impact to foreground I/O while in degraded mode situations.
SyncMirror	Use local SyncMirror to make sure of shelf-level resiliency and to improve performance in degraded mode situations.
Spares	Use a maximum hot spares approach to make sure sufficient disks are available for corrective actions. Set the RAID option raid.min_spares_count to the recommended number of spares to make sure the administrator will be notified when spare counts are reduced below recommendations.
Drive Type	Use performance drives (SAS, FC, or SSD) instead of capacity drives (SATA). Smaller capacity 15krpm or SSD drives result in shorter times for corrective actions. This is important when foreground I/O is prioritized over corrective I/O, which increases times for corrective actions. Performance drives help offset that performance delta.
Aggregate Fullness	Monitor aggregate “fullness” as performance degrades as disks get full (the drive heads need to travel farther to complete I/Os). Drive failures will further degrade foreground I/O performance when drives are nearing full data capacity.
Utilization Monitoring	Monitor CPU utilization, disk utilization, and loop/stack bandwidth. If your utilization is greater than 50%, you are at increased risk to see greater foreground I/O degradation in degraded mode situations. This can also increase the time it takes for corrective actions to complete.
I/O Prioritization	Prioritize foreground I/O over corrective I/O by adjusting the RAID option `raid.reconstruct.perf_impact` to low.
Scrubs	Use the default settings for RAID scrubs and media scrubs. Systems are assumed to be highly utilized, so increasing the duration of scrubs will likely provide a reduced benefit to data integrity while consuming additional system resources.
Maintenance Center	Maintenance Center is recommended to enable intelligent triage of suspect drives in the field. This also facilitates the RMA process for failed drives to make sure of system return to a normal operating state in a timely manner.
RAID Groups	Follow the RAID group sizing guidelines described in TR-3838, "Storage Subsystem Configuration Guide." Selecting RAID group sizes that are smaller will create smaller failure domains, which affect fewer drives when failures occur. This also helps to reduce completion time for correction actions.

6.2 TIER 2: BUSINESS CRITICAL

These types of environments are likely subject to compliance requirements, and although maintaining client access to the storage system is important, the loss of data would be severely detrimental to the customer. No customer likes to lose data, but these customers are under legal obligations and are subject to significant penalties when found to be in noncompliance. This could also be a configuration that is protecting a company's intellectual property. Medical records, software source code, and e-mail are examples of environments that fit into this tier of data availability. This tier of data availability is tuned toward prioritizing corrective I/O while balancing foreground I/O. By prioritizing corrective I/O over foreground I/O in degraded situations, you increase the impact to foreground I/O performance.

Table 3) Recommendations and best practices for business-critical data availability.

Business-Critical Recommendations	
Flash Cache	Use Flash Cache to improve system performance and minimize the impact to foreground I/O while in degraded mode situations.
SyncMirror	Use local SyncMirror to make sure of shelf-level resiliency and to improve performance in degraded mode situations.
Spares	Use a maximum hot spares approach to make sure sufficient disks are available for corrective actions. Set the RAID option raid.min_spares_count to the recommended number of spares to make sure the administrator will be notified when spare counts are reduced below recommendations.
Drive Type	Use performance drives (SAS, FC, or SSD) instead of capacity drives (SATA). Smaller capacity 15krpm or SSD drives result in shorter times for corrective actions. This is important when foreground I/O is prioritized over corrective I/O, which increases times for corrective actions. Performance drives help offset that performance delta.
Aggregate Fullness	Monitor aggregate "fullness" as performance degrades as disks get full (the drive heads need to travel farther to complete I/Os). Drive failures will further degrade foreground I/O performance when drives are nearing full data capacity.
Utilization Monitoring	Monitor CPU utilization, disk utilization, and loop/stack bandwidth. If your utilization is greater than 50%, you are at increased risk to see greater foreground I/O degradation in degraded mode situations. This can also increase the time it takes for corrective actions to complete.
I/O Prioritization	Use the default setting of medium for the RAID option `raid.reconstruct.perf_impact` to balance foreground I/O and corrective I/O.
Scrubs	Consider increasing the frequency of RAID scrubs to increase integrity of data at rest.
Maintenance Center	Maintenance Center is recommended to enable intelligent triage of suspect drives in the field. This also facilitates the RMA process for failed drives to make sure of system return to a normal operating state in a timely manner.
RAID Groups	Follow the RAID group sizing guidelines described in TR-3838, "Storage Subsystem Configuration Guide." Selecting RAID group sizes that are smaller will create smaller failure domains, which affect fewer drives when failures occur. This also helps to reduce completion time for correction actions.

6.3 TIER 3: REPOSITORY

Repository environments are used to store collaborative data or user data that is noncritical to business operations. Scientific and engineering compute data, workgroup collaboration, and user home directories are examples of environments that fit into this tier of data availability. This tier of data availability is the middle ground that balances foreground operations with correction actions (should they be needed). Defaults are normally appropriate for these configurations.

Table 4) Recommendations and best practices for repository data availability.

Repository Recommendations	
Flash Cache	Use Flash Cache to improve system performance and minimize the impact to foreground I/O while in degraded mode situations.
SyncMirror	Use local SyncMirror to make sure of shelf-level resiliency and to improve performance in degraded mode situations.
Spares	Use a balanced hot spares approach to allow more disks to be used to add to the system capacity. Set the RAID option raid.min_spares_count to the recommended number of spares to make sure the administrator will be notified when spare counts are reduced below recommendations.
Drive Type	Consider using SATA drives (backed by Flash Cache) for these types of configurations.
Aggregate Fullness	Monitor aggregate "fullness" as performance degrades as disks get full (the drive heads need to travel farther to complete I/Os). Drive failures will further degrade foreground I/O performance when drives are nearing full data capacity.
Utilization Monitoring	Monitor CPU utilization, disk utilization, and loop/stack bandwidth. If your utilization is greater than 50%, you are at increased risk to see greater foreground I/O degradation in degraded mode situations. This can also increase the time it takes for corrective actions to complete.
I/O Prioritization	Use the default setting of medium for the RAID option `raid.reconstruct.perf_impact` to balance foreground I/O and corrective I/O.
Scrubs	Consider increasing the frequency of RAID scrubs to increase integrity of data at rest.
Maintenance Center	Maintenance Center is recommended to enable intelligent triage of suspect drives in the field. This also facilitates the RMA process for failed drives to make sure of system return to a normal operating state in a timely manner.
RAID Groups	Follow the RAID group sizing guidelines described in TR-3838, "Storage Subsystem Configuration Guide." Implementing larger RAID groups will help maximize usable capacity of the storage configuration (within spindle maximums).

6.4 TIER 4: ARCHIVAL

This type of environment is subject to a large initial ingest of data (write), which then becomes seldom accessed. System utilization on average is not expected to be very significant. As the data is seldom accessed, it is important to fully leverage subsystem features that exercise that data to make sure of continued integrity. Given the priority is maintaining data integrity, these configurations are tuned toward prioritizing corrective I/O and minimizing completion time for corrective actions. Backup and recovery, archiving, near-line, and reference data are examples of environments that fit into this tier of data availability.

Table 5) Recommendations and best practices for archival data availability.

Archival Recommendations	
Spares	Use a maximum hot spares approach to make sure sufficient disks are available for corrective actions. Set the RAID option raid.min_spares_count to the recommended number of spares to make sure the administrator will be notified when spare counts are reduced below recommendations.
Drive Type	Consider using SATA drives (backed by Flash Cache) for these types of configurations.
Aggregate Fullness	Monitor aggregate “fullness” as performance degrades as disks get full (the drive heads need to travel farther to complete I/Os). Drive failures will further degrade foreground I/O performance when drives are nearing full data capacity.
Utilization Monitoring	Monitor CPU utilization, disk utilization, and loop/stack bandwidth. If your utilization is greater than 50%, you are at increased risk to see greater foreground I/O degradation in degraded mode situations. This can also increase the time it takes for corrective actions to complete.
I/O Prioritization	Use the default setting of medium for the RAID option `raid.reconstruct.perf_impact` to balance foreground I/O and corrective I/O.
Scrubs	Consider increasing the RAID scrub duration (`raid.scrub.duration`) to help make sure of the integrity of data at rest. Consider increasing the media scrub rate (`raid.media_scrub.rate`) to increase drive-level block integrity.
Maintenance Center	Maintenance Center is recommended to enable intelligent triage of suspect drives in the field. This also facilitates the RMA process for failed drives to make sure of system return to a normal operating state in a timely manner.
RAID Groups	Follow the RAID group sizing guidelines described in TR-3838, "Storage Subsystem Configuration Guide." Implementing larger RAID groups will help maximize usable capacity of the storage configuration (within spindle maximums).

6.5 TIER 5: MULTIPURPOSE

One of the many strengths of NetApp storage is the ability to host multiple tiers of data availability within the context of a single system (HA pair). This tier is simply calling out this capability, which might result in conflicting configuration recommendations. In a Data ONTAP 7G multipurpose environment, aggregates and volumes are likely to be competing for the same system resources. For truly global options it is recommended that you tune those toward the most sensitive data availability tier being hosted in the storage configuration. With the introduction of Data ONTAP 8.0 Cluster-Mode, the ability to segregate how the logical configuration of a system utilizes system resources is much improved.

For example, the RAID option `raid.reconstruct.perf_impact` could be set to either low or high if your storage configuration is hosting both mission-critical and archival data. As mission-critical data is more sensitive to system configuration than archival data, the recommendation would be to set the option to low as per the application data availability tier (tier 1) recommendations.

Table 6) Recommendations and best practices for multipurpose data availability.

Multipurpose Recommendations	
Prioritize Recommendations	Prioritize configuration recommendations for the most sensitive tier of data availability when conflicting recommendations are present.
FlexShare®	Consider using FlexShare to prioritize system resources between data volumes.
Physical Segregation	Segregate the physical shelf and drive layout for multiple data availability tiers. For example, if you have both SAS and SATA (DS4243) attached to the same system, you could be using the SAS drives to host mission-critical data while using the SATA drives to host archival data. Although you can mix DS4243 SAS shelves with DS4243 SATA shelves in the same stack, it is recommended to separate the shelves into separate stacks so physical failures affecting one tier of data availability will not directly affect both tiers of storage being hosted (in this example).

NetApp provides no representations or warranties regarding the accuracy, reliability, or serviceability of any information or recommendations provided in this publication, or with respect to any results that may be obtained by the use of the information or observance of any recommendations provided herein. The information in this document is distributed AS IS, and the use of this information or the implementation of any recommendations or techniques herein is a customer's responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. This document and the information contained herein may be used solely in connection with the NetApp products discussed in this document.

www.netapp.com