WINCHESTERSYSTEMS®
Purpose-Built Storage

Winchester Systems FlashDisk

Intelligent Drive Recovery

Protecting Your Data

Contents

Introduction

One of the basic capabilities of a storage system is to provide protection mechanisms against the failure of disk drives. With Winchester Systems FlashDisk **Intelligent Drive Recovery (IDR)** technology, storage ensures disk drive protection against data loss even when entire drives malfunction. In some cases, there may be faulty sectors of a drive that are not detected because they are rarely accessed by the host. Eventually, when users need data from those faulty sectors, especially in critical situations such as system volume rebuilding, users may lose that data without being able to recover it. However, if we can detect and repair those faulty sectors on drive media in advance, disastrous media errors can be avoided before they even happen.

Products covered by this document

FlashDisk VX and FX Series Storage Subsystems

What is a media error?

Media errors are defects, bad spots, or damaged areas/sectors of the disk surface that cannot reliably store data for subsequent retrieval. These may occur on any disk from any manufacturer. This is a known and accepted reality in the disk drive industry, particularly with high density storage usage.

Why do media errors occur?

Disk drive lifecycle has increased to the point where most enterprise-level disk drives boast MTBF (mean time between failures) specifications in excess of one million hours. Despite these improvements, they are still susceptible to media errors. Most drives in storage systems are equipped for high density capacity and well protected, serving host applications for long periods and sustaining large data volumes. As time goes by, even enterprise-level drives reach their mechanical fault threshold, and defects may appear on areas of the disk. With large capacity drives (2TB and more), the possibility or faulty sectors may also increase simply because of drive size.

Impact

Media defects affect drive ability to read data from a specific sector. They do not indicate general unreliability of a disk drive. For individual drives, data on those faulty areas may be irretrievably lost, but data on all other sectors can still be used even for several years. With RAID protection, we can leverage RAID parity mechanisms to recover the data on those faulty areas via other existing striped data. If the drive still has available space, we can even ignore media errors if we discover them and recover the affected data in advance. Of course, if the media error starts to occur on an

entire specific drive and the situation gets worst (meaning more errors found after each regular check), that may mean the drive is close to the end of its life cycle. To avoid disastrous whole drive failure and the risk of data loss, users should plan to replace the drive with a new one as soon as possible.

RAID 6 Protection

Why RAID 6?

Traditional RAID 3 and RAID 5 configurations were created to increase the reliability of data by using one extra disk drive to store parity and error correction information. The problem is that if one of the drives fails or media errors are found on it at the exact moment that a data recovery operation is being conducted (RAID rebuilding), the system cannot re-acquire the data that started the recovery operation and it becomes permanently lost because one of the disk drives is faulty.

RAID 6 systems store two parity and error correction information sets, which are arranged in such a way that even if one disk drive fails or media errors are found during the data recovery process, the system can continue operations and no data loss occurs, all thanks to the higher protection level compared to other typical RAID configurations. This is especially useful for high density disk drive usage.

RAID 5 vs. RAID 6 considerations

Due to RAID design, the RAID 5 configuration can get better performance because it only performs one drive parity calculation. It also offers more capacity because it only leverages one drive for parity data. But the more drives we have installed in a storage system, impact on performance and capacity becomes minor compared to the risk of data loss. That is why we recommend users leverage RAID 6 instead of RAID 5 if they have sufficient drives and capacity but want better RAID protection.

Drive count and RAID configuration policy

FlashDisk firmware defaults to the selection of all available drives for logical unit creation. Available drives affect the recommended RAID configuration based on the following policy:

- Eight or more drives: RAID 6
- Less than eight drives (three to seven): RAID 5
- Under three drives: defaults to only allow custom mode configuration

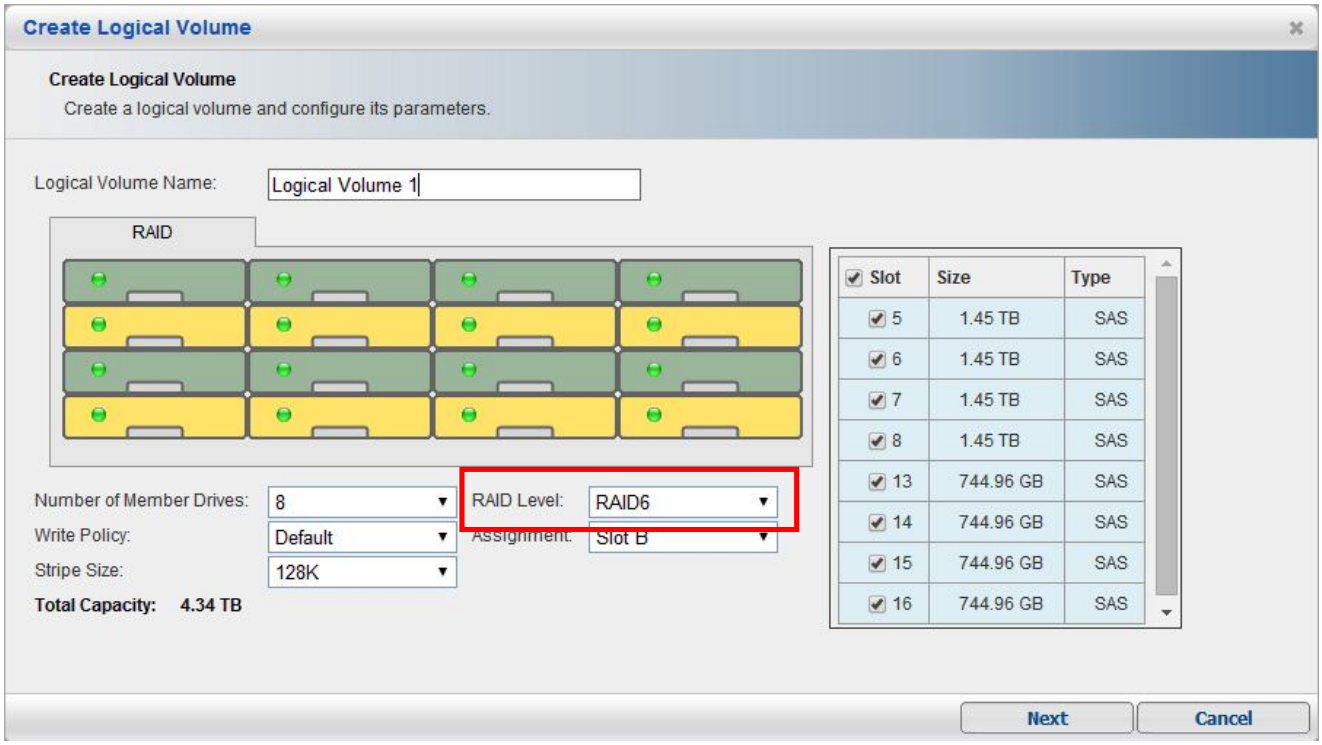

Figure 1: select RAID 6 as the default RAID configuration for logical unit creation

We strongly recommend users choose RAID 6 as a more reliable and secure configuration. If users still toggle the RAID configuration option from the default-selected RAID 6 to another option, a specific warning message will appear to inform them of this action. There are other available options in the configuration wizard: RAID 6 + local spare, RAID 5, and RAID 5 + local spare. Users can specify any supported RAID level in custom mode.

Enable S.M.A.R.T

FlashDisk systems enable S.M.A.R.T for all attached disk drives.

What is S.M.A.R.T?

S.M.A.R.T (Self-Monitoring, Analysis, and Reporting Technology) is a monitoring mechanism for computer disk drives to detect and report various indicators of reliability, aiming to anticipate and prevent failures.

S.M.A.R.T. options

The FlashDisk system enables S.M.A.R.T with the **detect-clone-replace** option

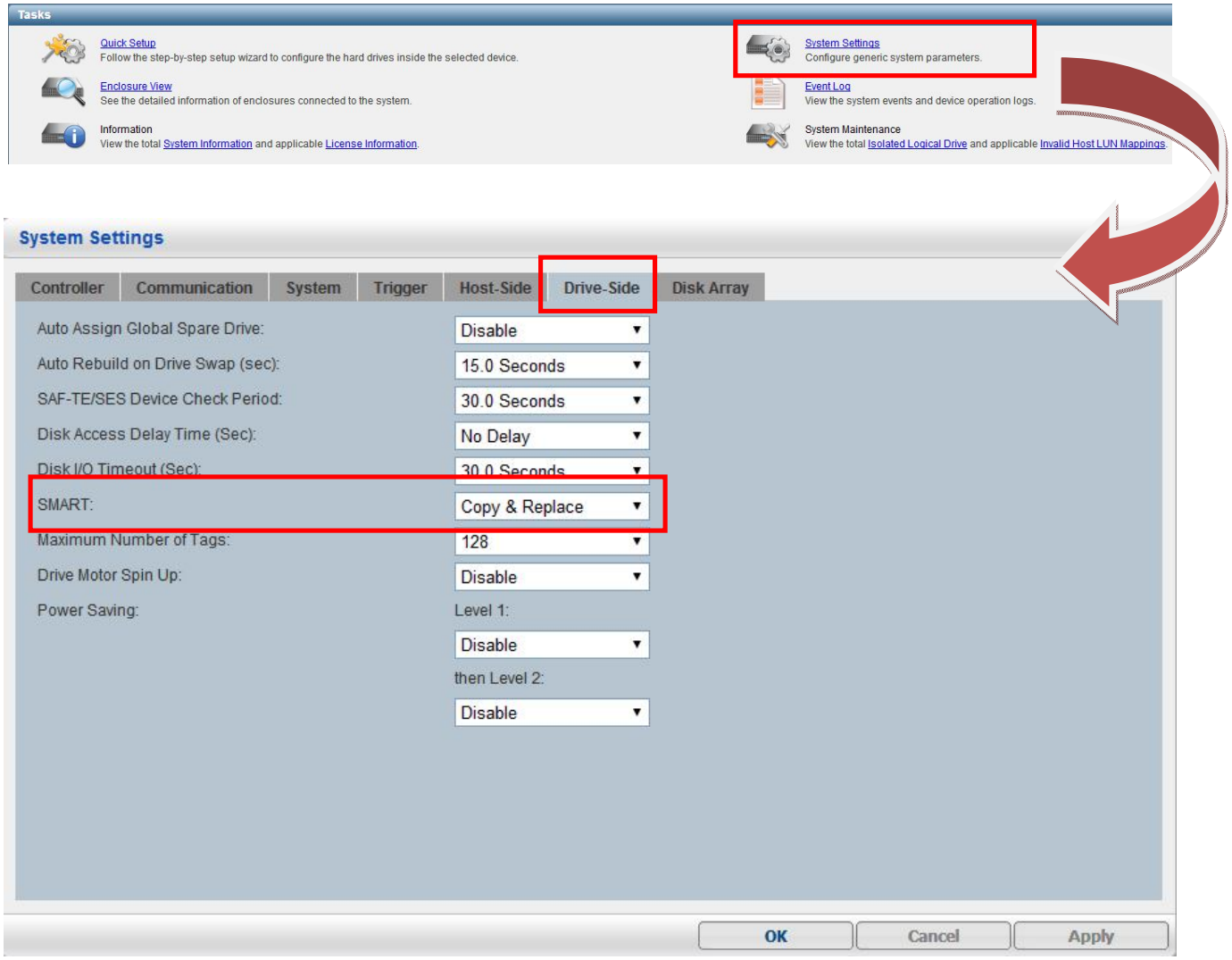

Figure 2: check if the storage system has enabled S.M.A.R.T. and related options

Detect-clone-replace option: if S.M.A.R.T reports a drive error, the system tries to clone the whole disk drive to an available spare and replace it. The system also informs the user with related event messages.

Media Scan

What is media scan?

The media scan feature sequentially checks a physical drive block by block for faulty areas. This surface scan operation detects and reallocates any sectors that are defective. It is run to reduce the possibility that a disk drive will experience soft media errors during operation.

How does media scan enhance data integrity?

Disk media errors occur over the course of disk drive life cycles, so drive manufacturers incorporate spare blocks into the media. The premise is that media errors can be replaced (reallocated or reassigned) with spares either automatically by the drive or using instructions from storage system firmware. When a block goes bad, it can be repaired (or reallocated) using a spare location that has been set aside for that purpose.

Not all areas of disk drives are accessed frequently, and sectors with errors that are not normally found and fixed can be detected by regular media scan operations.

When a disk drive encounters trouble reading data from a sector (so-called **disk soft error**), it automatically attempts recovery of the data through its various internal methods. Whether or not the drive is eventually successful at reading the sector, it reports the event to the storage system. If specific data sectors cannot be read and media errors are found during the media scan process, the storage system can try to fix them by reconstructing the data from parity via RAID protection, and attempt to write again on the same area.

If it still fails to write, the storage system posts an event and asks for the drive to reassign the sector to a spare physical location via a SCSI command. If the drive still has spare locations, the storage system directly rewrites the data to the new physical location and recovers data from any faulty sectors and areas. That is why media scan operations enhance data integrity after drive checks even following a repair process.

In the following example, one logical unit was created using a RAID 5 configuration with four disk drives (HDD1 to HDD4), and each drive has striped user data and RAID parity data for recovery. If media scan is started and one media error is found on a specific stripe in the first drive (HDD1) like in figure 3:

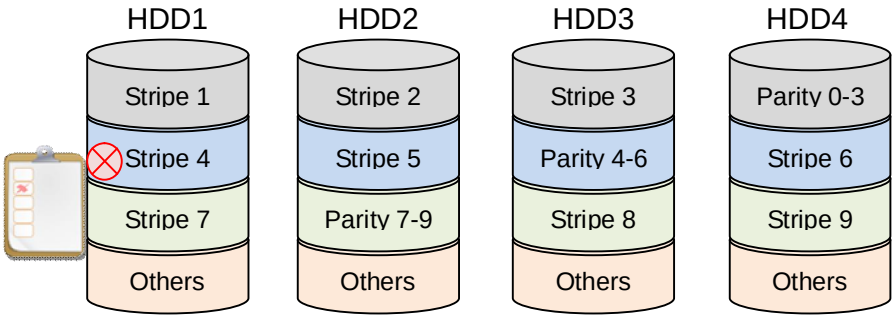

Figure 3: media scan started and found an error

the data in the faulty area can be regenerated from other user data and parity data by XOR calculation. The regenerated data is written to the same area first if possible. If the operation is successful, the media error is repaired and data is recovered.

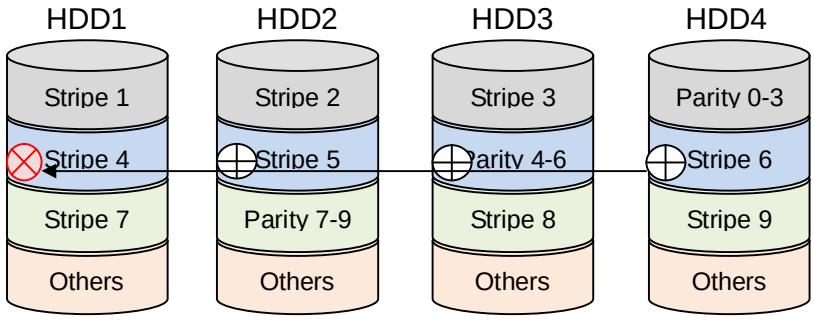

Figure 4: leverage RAID protection to regenerate and recover data for faulty area

If writing to the same location on the disk drive fails, the storage system tries to ask the disk drive to reassign capacity from spare space for the recovered data, and rewrite the data again. The data is redirected to the new location.

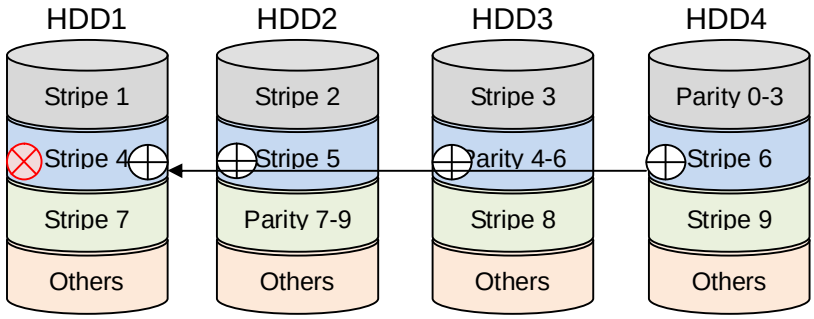

Figure 5: if writing to the same location fails, drives reassign space and rewrite data for recovery

Automatic Media Error Check

Media error handling is a high priority in RAID sets. Since media scan helps detect media errors in drives, storage systems should engage checks regularly to prevent loss of data availability before specific drives fail.

Why choose scheduled tasks instead of background continuous scan?

For repeated checks, we recommend to add scheduled tasking to check drive media periodically. Since some storage systems use green energy-saving design, they may toggle disk drives to idle mode to reduce power consumption. If the system engages in background continuous media scan operation, all disk drives never switch to idle mode even if there are no host I/O requests. This means continuous media scan in the background effectively cancels valuable green design features that help users save money.

Scheduled media scan tasks for periodical checks

Add an automatic media check task schedule after logical unit creation.

- Every basic logical unit is equipped to facilitate a media scan schedule. Users can remove or customize those scheduled tasks manually

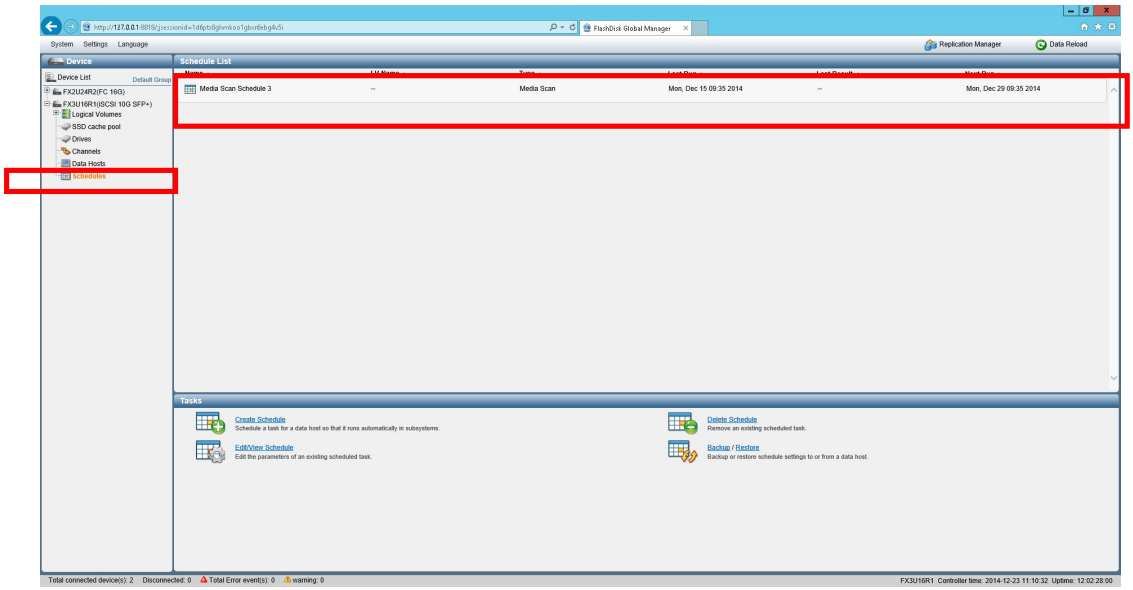

Figure 6: users can delete or customize tasks manually

- If users restore factory default settings, existing media scan schedules will be added back automatically after default restore is completed
- After firmware upgrade, all logical units retain existing media scan schedule settings. For legacy units without schedules, new firmware allows adding relevant settings
- There is no performance impact for background media scan progress because the check process is suspended if any host I/O request is incoming, and is only engaged/resumed when the system is idle

Policy of automatic media check schedule

Media scan priority: low

Start time: specific period after creation (or default restore)

Check period: every two weeks for RAID configuration (applies to RAID 1, RAID 3, RAID 5, and RAID 6)

Overdue schedule: continue after previous task is done

If the previous task is not finished and the next task is incoming, that means the storage system is constantly busy and cannot complete media scan in the configured period. The next task will be engaged after the previous task finishes, and the precise scheduled time may be missed. Users will be notified with an event message if the system is busy and the scan priority or check period is not suitable for their environment. Users can then change the scan policy by increasing the priority to normal or high to accelerate the process, but this may impact host I/O performance during scan operation. Users can also increase the scan period, but the check interval will also be increased, which is another tradeoff for protection – all based on user preference.

Caution:

1. Older models may not support IDR and advanced media scan without upgrading to the latest firmware. They still allow users to create media scan schedules manually.
2. The scan process can only be engaged if logical unit status is “good”. Should logical drives be in "rebuilding" or "drive adding" status, the scan process will pause and resume if the condition is resolved.
3. If a controller fails, in-progress scan operations resume when another controller takes over (only applies to models with redundant controllers).
4. Normal storage system reset may also pause in-progress scan operations, which resume after the system is fully back online.

Conclusions

Winchester Systems FlashDisk storage systems offer comprehensively and thoughtfully designed RAID products that augment media scan and error prevention with **Intelligent Drive Recovery (IDR)**. In addition to the standard data protection provided via generic RAID technology, partial and hard to detect media faults that occur in disk drives are handled and mitigated by Winchester Systems putting extra effort into relevant technologies – in the case of this white paper being **IDR**. These mechanisms help prevent possible data loss issues, as well as manage otherwise unrecoverable media errors to minimize impact and result in non-degraded sets. This is because small and otherwise easily-missed errors are handled as if the entire disk drive failed, so missing data can be regenerated with no negative effect on the overall RAID set.