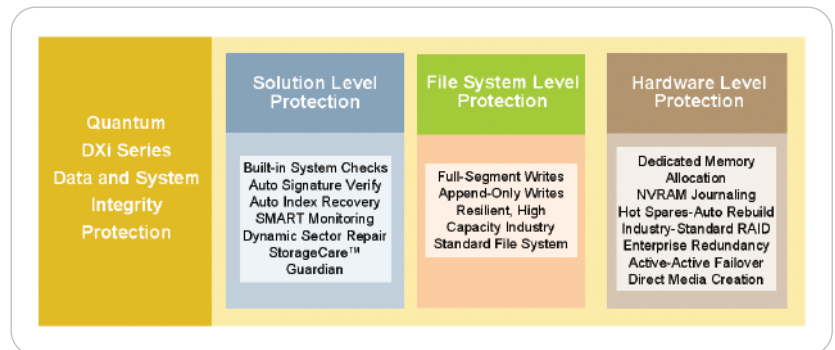


DXi-Series Disk Backup Systems with Data De-duplication: Providing Comprehensive Data and System Integrity

Quantum's DXi-Series disk backup solutions use data de-duplication technology to increase the capacity of disk to retain backup data by 10 to 50 times and to make WAN-based replication a practical DR protection tool. In addition to taking advantage of the inherent reliability of disk-based read and write operations, the DXi-Series solutions provide multiple levels of automated processes which monitor, protect, and maintain the integrity of the system and the data stored within it. This paper describes the processes that operate at the solution, file system, and system hardware levels.



VARIABLE-LENGTH DATA DE-DUPLICATION BACKGROUND

In the DXi-Series de-duplication system, data is segmented into variable length blocks, a digital signature is generated for each segment, the blocks are written to the system's RAID array, and the signatures are stored in an index. As data is written to a DXi-Series system, each unique block only needs to be stored one time. If a block is sent to the system again, a pointer to the original instance of the data segment is stored instead of storing the block again. Since the pointer uses a small fraction of the space used by the original block, data de-duplication technology dramatically increases the capacity of the disk to retain data. The same logic is also applied to replication of de-duplicated backup data between DXi-Series systems. The replication source checks the target before sending blocks to see if the target already holds copies of them, and it only sends new blocks during replication. The effect is that replication of de-duplicated backup sets uses much less bandwidth than conventional backup.

SOLUTION LEVEL INTEGRITY PROCESSES:

Built-in system integrity checks. These routines protect the DXi-Series solutions by continually checking the state of system's hardware and software conditions using a built-in, automated test process during normal operations. If anomalies are detected, the system suspends de-duplication operations and completes full data and index verification/correction in background.

Automated signature verification. During all read or access operations, the system checks validity of the index signatures in real time. If a fault were to occur, the system would invoke correction activities, recovering the index or re-acquiring the segment.

Auto-index recovery. If any faults are identified in the signature index, the index is automatically corrected by regenerating the index signatures from the original blocks.

Auto-data re-acquisition. If faults were to occur in the data blocks, an automated process is initiated to correct blocks by re-acquiring the segments during subsequent backup events.

SMART predictive monitoring. DXi-Series systems continually monitor system-level trend data using the industry-standard System Monitoring and Reporting Technology (SMART) protocol, including recoverable error rates and re-tries. When specific thresholds are reached, the system issues alerts recommending preventative maintenance.

DXi-Series Disk Backup Systems with Data De-duplication: Providing Comprehensive Data and System Integrity

Dynamic sector repair. In the event of a read failure from any drive sector, the system automatically re-directs the request to a redundant copy of the data maintained on disk. Then it isolates the faulty sector and creates a new sector to maintain full internal redundancy.

StorageCare™ Guardian monitoring and alerting. Continually monitors hardware and software status of DXi-Series systems in any location and links data to Quantum's service team and/or to the end user. Provides alerting functions, log download, and remote corrective action capability.

FILE SYSTEM LEVEL INTEGRITY PROCESSES:

Full-segment write policies. Each segment is fully written to a buffer location at ingest prior to beginning of de-duplication and is only overwritten after de-duplication has successfully been completed.

Append-only write policies. New data is only appended to existing data pool. Data blocks are never deleted or made available for overwrites until the data status and all references are verified.

Resilient, high capacity file system. DXi-Series solutions leverage the file system technology used in StorNext® data management software. StorNext FS is a highly resilient journaled file system that scales to millions of files, and it is the standard primary file system used throughout the world in very high capacity, high availability applications, including broadcast, rich media, and science and engineering.

HARDWARE SYSTEM LEVEL INTEGRITY PROCESSES:

Dedicated, internal memory allocation. Isolates all data de-duplication processes and indexes from other operations to ensure integrity.

Write journaling in NVRAM. All data written to disk is journaled in non-volatile RAM so that it can be maintained and correctly recorded even in the event of a power failure or accidental drive removal/insertion.

Hot spare and automated rebuild. For DXi-Series systems with hot spares, RAID rebuild occurs automatically as a background process if a disk were to fail. All drives are hot swappable.

Proven, industry-standard RAID. The DXi-Series RAID controllers leverage commercial array technology with a history of tens of thousands of installations supporting primary datasets worldwide.

Enterprise-level redundancy. The DXi3500 and DXi5500 appliances, both of which were designed for operation in distributed environments and midrange data centers, include redundant, hot swappable components (fans, power supplies) as a standard feature. The DXi7500, designed for disk backup in Enterprise data centers, offers a fully redundant architecture that eliminates all single points of failure. Main system controllers, RAID controllers, and switches all provide redundant components with active-active failover.

Direct creation of removable media. The DXi7500 offers the added protection feature of direct tape creation, automatically generating physical tape in a connected library to provide both redundancy and long term retention. Tape creation can run in parallel with the backup operation, data is moved over a dedicated SAN without using the media server's processors, and barcodes on physical media match those created in the backup application's media register.