The Library of Congress

Storage Architectures for Digital Preservation

Module 4: Moving and Distributing Data

Requirements for Moving Data and Distributing Data

Move significantly large amounts of data M TB/hour ($M > 2$)

Heterogeneity i.e. move data across vendor arrays

Move data across distance

Data integrity checks after migration of data

Moving data should not impact performance of servers/arrays

Seamless migration of data – no host agents, no re-wiring etc.

SAN Switches provide performance, flexibility, and intelligence to move and distribute data

Trends in SAN-based Data Movement

Storage Array → SCSI Read → **Cisco MDS 9000 SAN Director** → SCSI Write → Storage Array

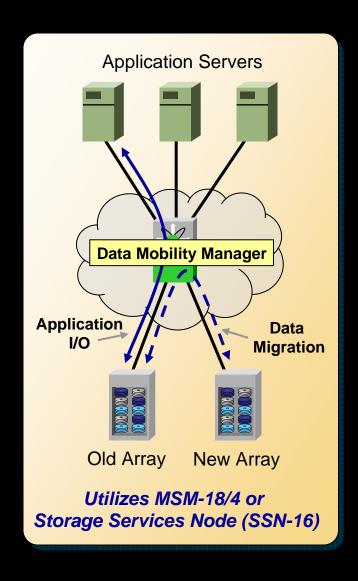| Requirement	SAN-based Data Movement Capabilities
Move Large Amounts of Data	• Purpose-built line cards with SCSI intelligence that reside in SAN Directors that can move up to 4TB/hour of data between storage arrays
Move Data Across Distance	• SAN-like performance over distance with SCSI Acceleration • Encapsulation of FC in IP protocols for distance connectivity • Encryption and compression
Heterogeneity	• Because SAN Directors operate at the SCSI level and have visibility into storage arrays from different vendors , it can provide heterogeneous data movement
Data Integrity Checks	• Because SAN Directors have high performance line cards with SCSI intelligence, they can perform data integrity checks to ensure data has not changed during the migration process
Seamless	• SAN Directors have the intelligence to migrate data between arrays without any host agents, no re-wiring, and no re-configuration.

An Example of SAN-based Data Movement: Cisco MDS Data Mobility Manager (DMM)

Application Servers

Data Mobility Manager

Application I/O

Data Migration

Old Array   New Array

Utilizes MSM-18/4 or Storage Services Node (SSN-16)

- Online migration of heterogeneous arrays

- Move up to 4 TB of data in 1 hour

- Simultaneous migration of multiple LUNs

- Unequal size LUN Migration

- Rate adjusted migration

- Verification of migrated data (integrity)

- Secure erasure of old data

- Dual fabric support for HA

- Requires no SAN re-configuration or rewiring

- CLI and wizard-based management with Cisco Fabric Manager