



HP NonStop Servers- Storage Stuff

Derek Ginger
NonStop Enterprise Division
July, 12 2004

© 2004 Hewlett-Packard Development Company, L.P.
The information contained herein is subject to change without notice.



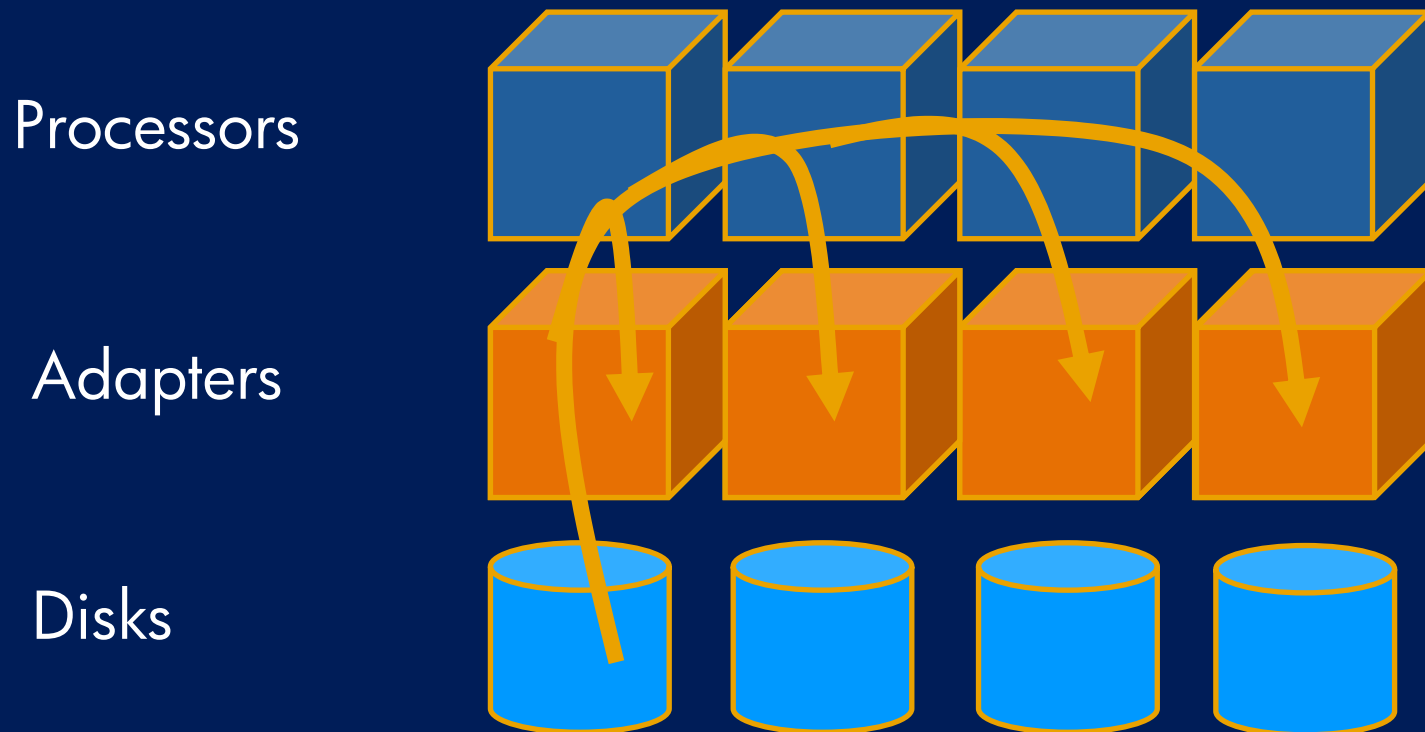
Agenda

- New features
- New I/O architecture
 - Modular I/O
 - Fibre Channel ServerNet adapter (FCSA)
 - Fibre Channel disks
 - Enterprise storage system (ESS)
- New data protection
- Tape drives

New features

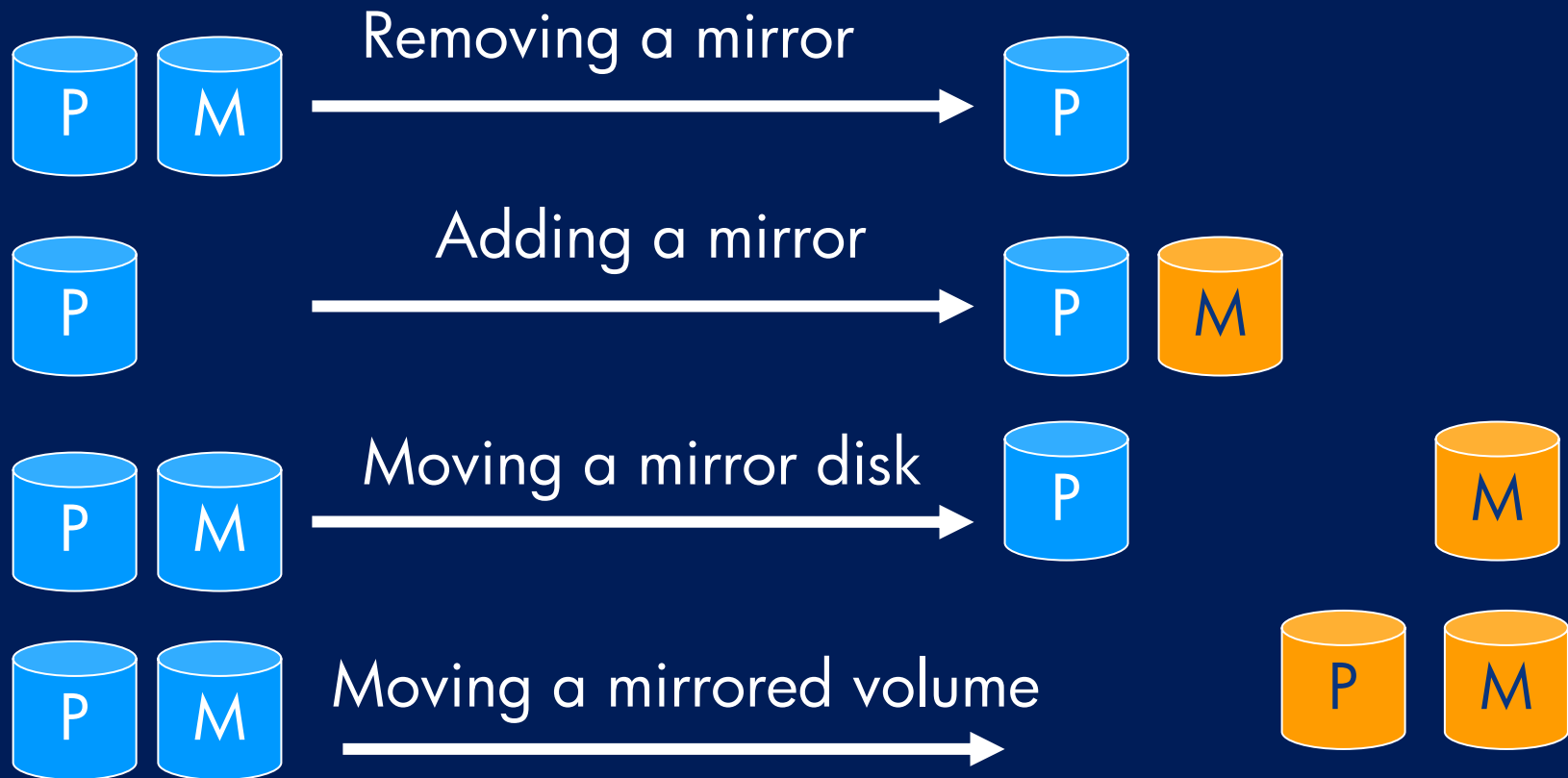


Online adapter firmware update (G06.18)



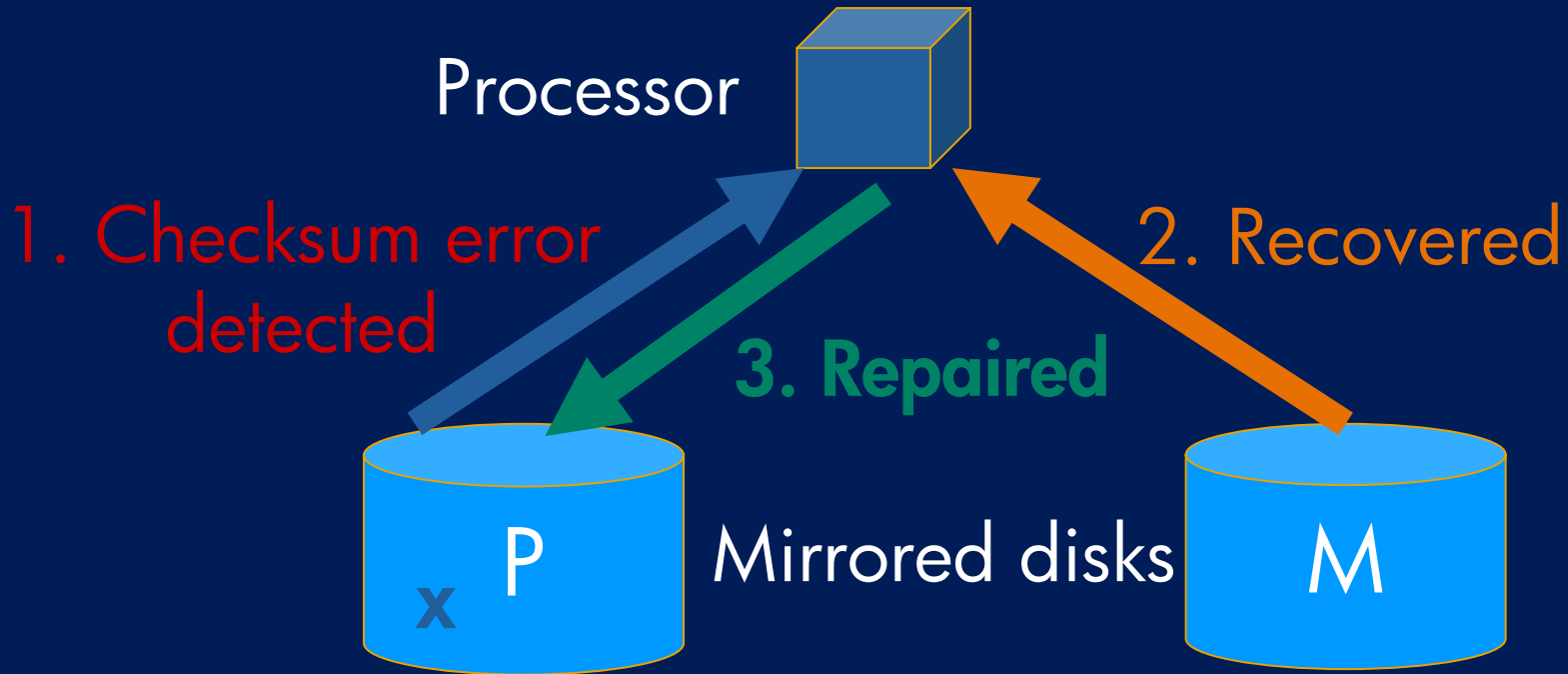
- All adapters firmware is re-downloaded one at a time
- Disks continue to run during update

Online disk re-mirroring (G06.21)



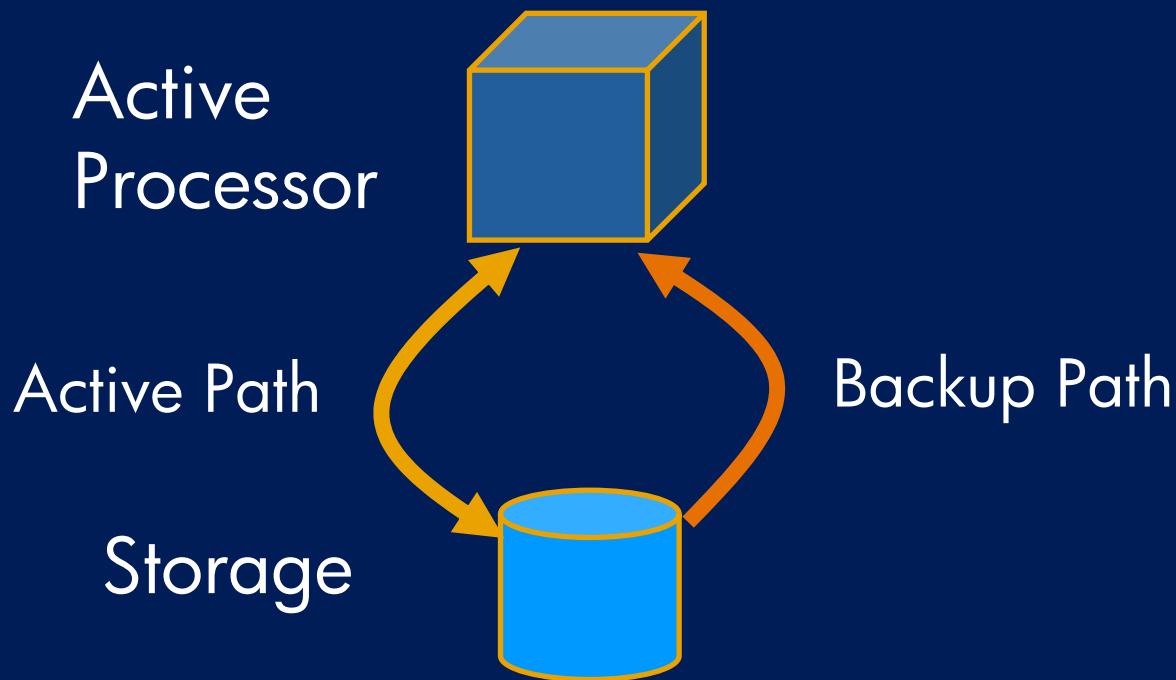
- Useful for online re-configuration and migration
- Utilizes disk Revive

Automatic checksum repair (G06.21)



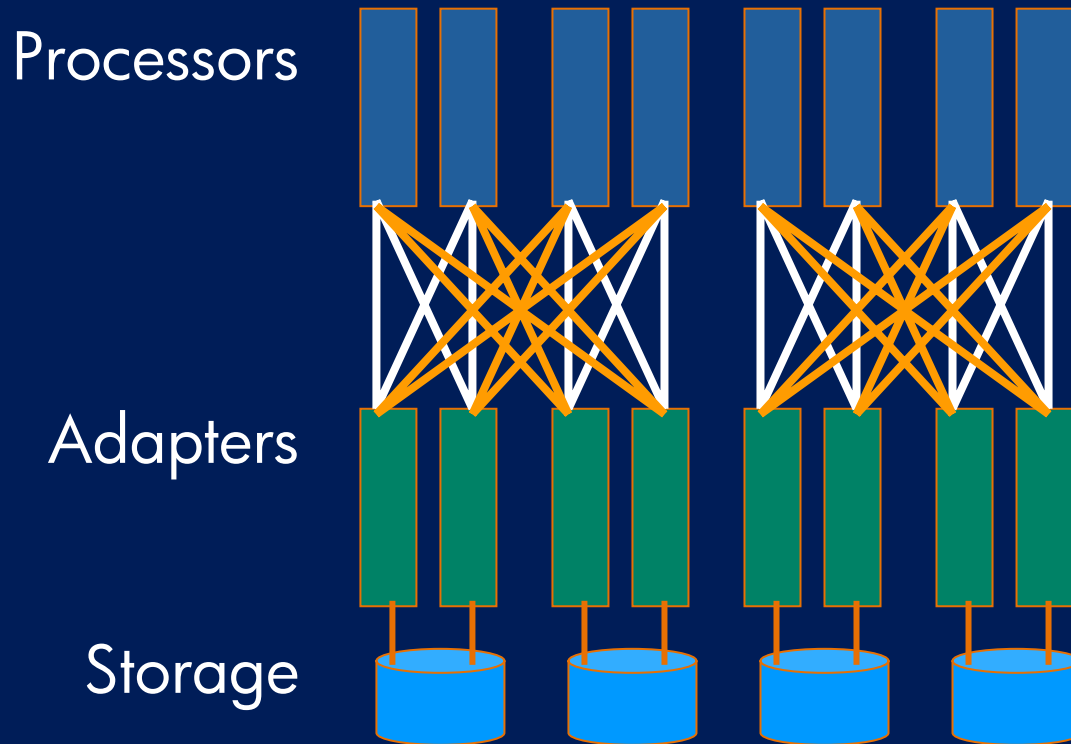
- When a processor detects disk checksum error:
 - Data and checksum is read and verified from the other disk
 - The error is repaired with data from the other disk
- Error repair without human intervention

Automatic backup path verification



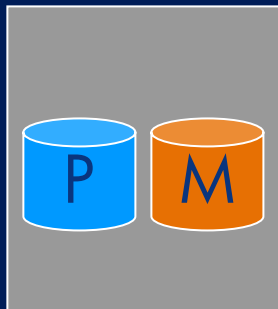
- The active path is used for data read and writes
- The active processor accesses the disk through the backup device path every few minutes
- Detects backup path latent failure

Flexible disk configuration



- Processors can be configured to access adapters in other enclosures
- Benefits:
 - Disk process can be load balanced over more processors
 - Protection against enclosure problems

Disks enclosure independence



Same Enclosure



With I/O Enclosure



Across
Processor Enclosures

- Protection against enclosure failure
- Enclosure interleaving with I/O Enclosure
- Near future: enclosure interleaving across processor enclosures
- System disks must be in the first processor enclosure

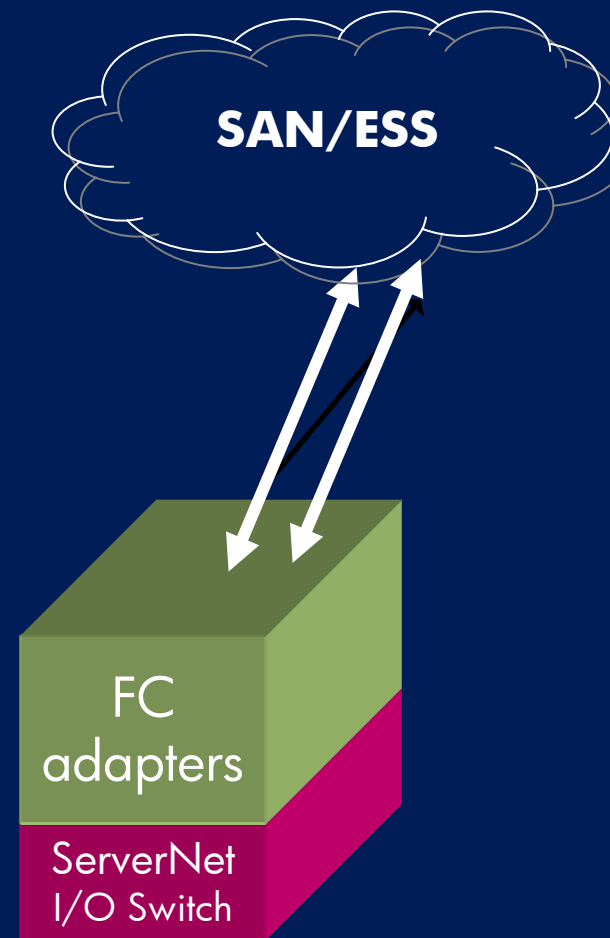
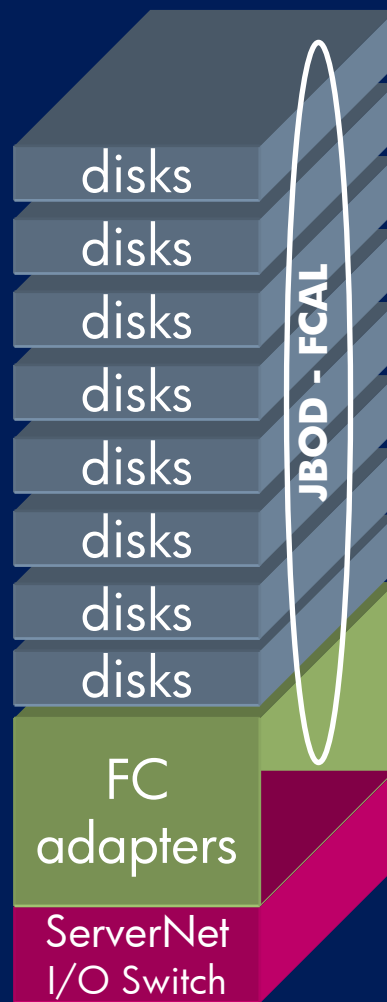
New I/O Architecture

New products

- Modular packaging
- Fibre Channel ServerNet adapter (FCSA)
- Fibre Channel disks
- Enterprise storage system (ESS)

Modular I/O for Storage

- Flexible and extensible configurations
- Can be connected to current S-series processor enclosures
- Migrates over to Itanium systems



Modular packaging

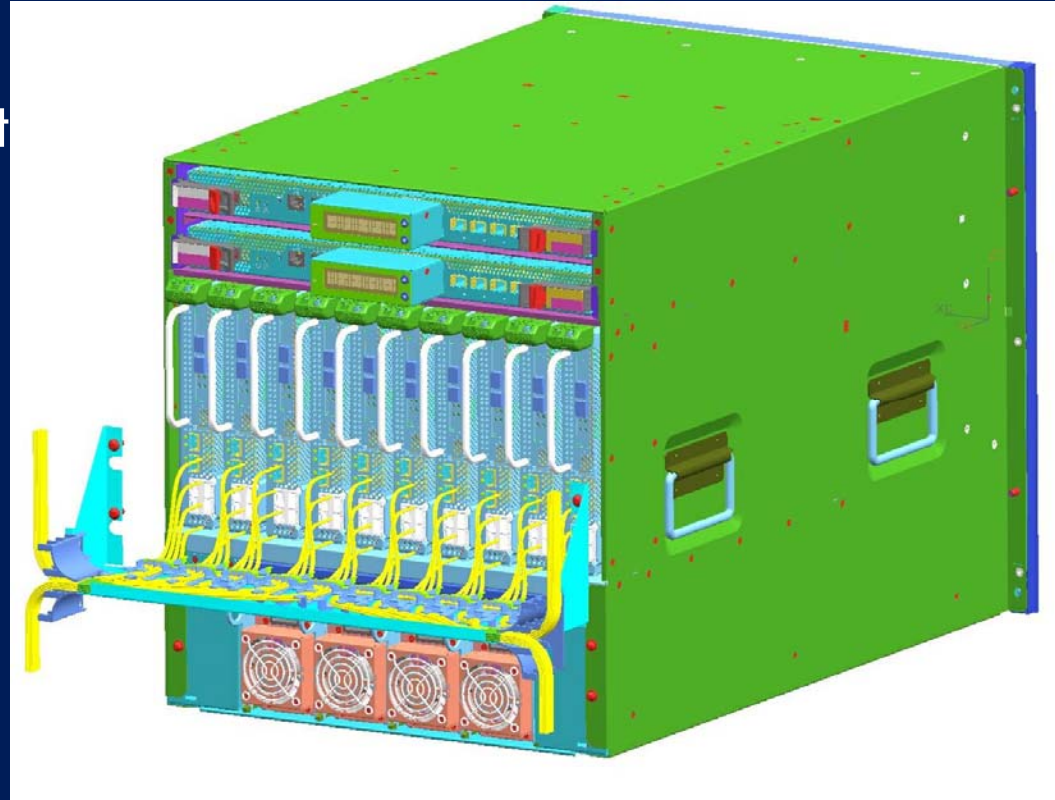


- Flexible and extensible configurations
- Allow independent technology update
- Granular part replacement
- Industry standard 19-inch rack
- Can be connected to current S-series processor enclosures
- With one I/O adapter module, up to 10 Fibre Channel disk modules



I/O adapter module

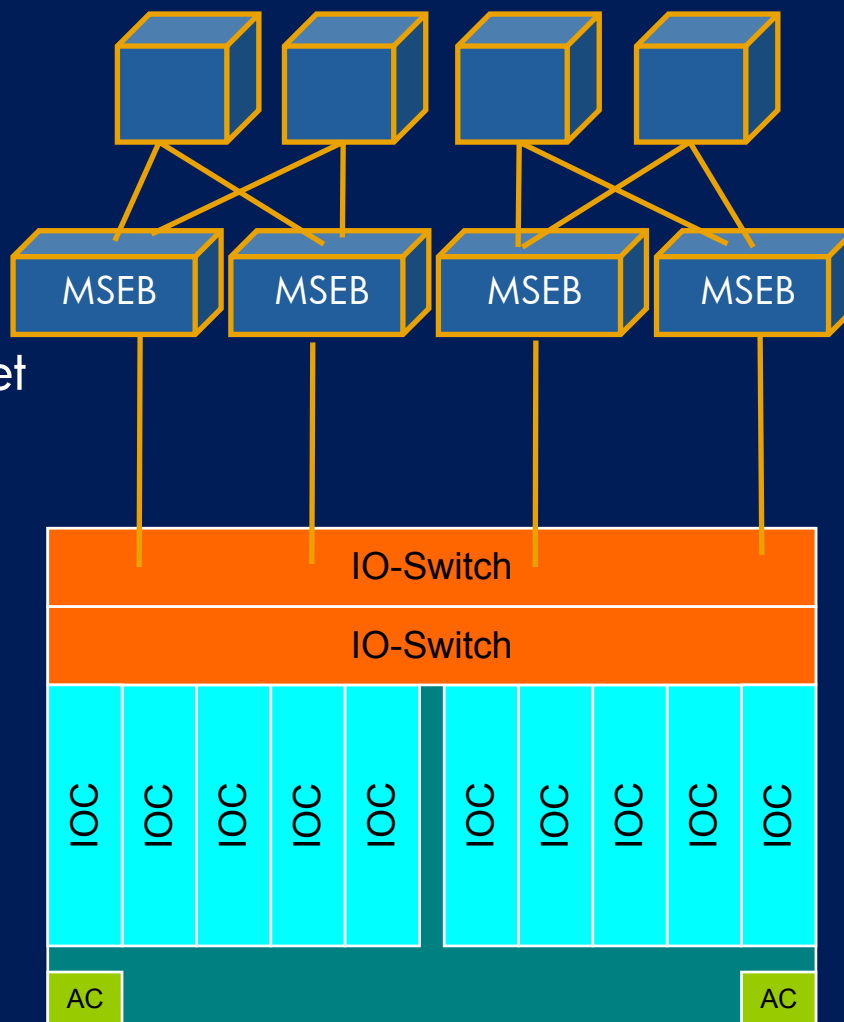
- Up to 10 I/O Adapters:
 - Fibre Channel ServerNet adapter
 - 4 port gigabit Ethernet ServerNet adapter
- 2 I/O Switches



Modular I/O System connection



NonStop Processors



Modular ServerNet expansion boards (MSEB)

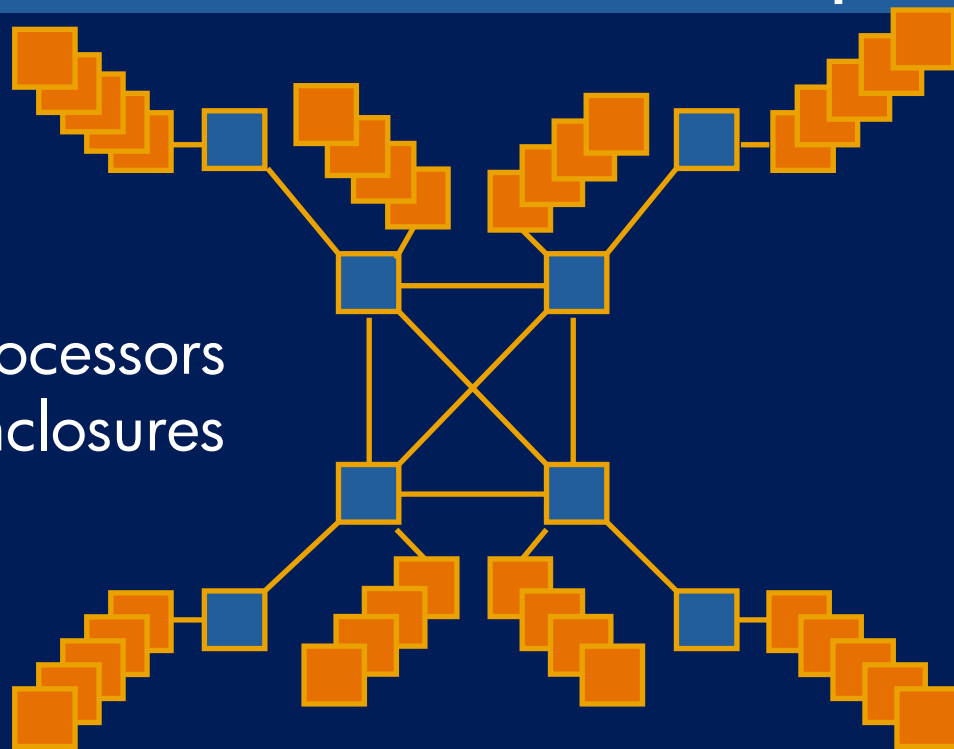
ServerNet links:

- 1.25 Gb/s per path, full duplex
- 125 meters

Modular I/O interconnect capacity

Modular
I/O
enclosures

Processors
enclosures



- Direct attached disks:
 - 5168 x 146 GB disks = 754,528 GB = 755 Tera Byte
- Enterprise storage provides additional storage
- Actual configuration should consider other factors such as application demand of processor capacity and memory

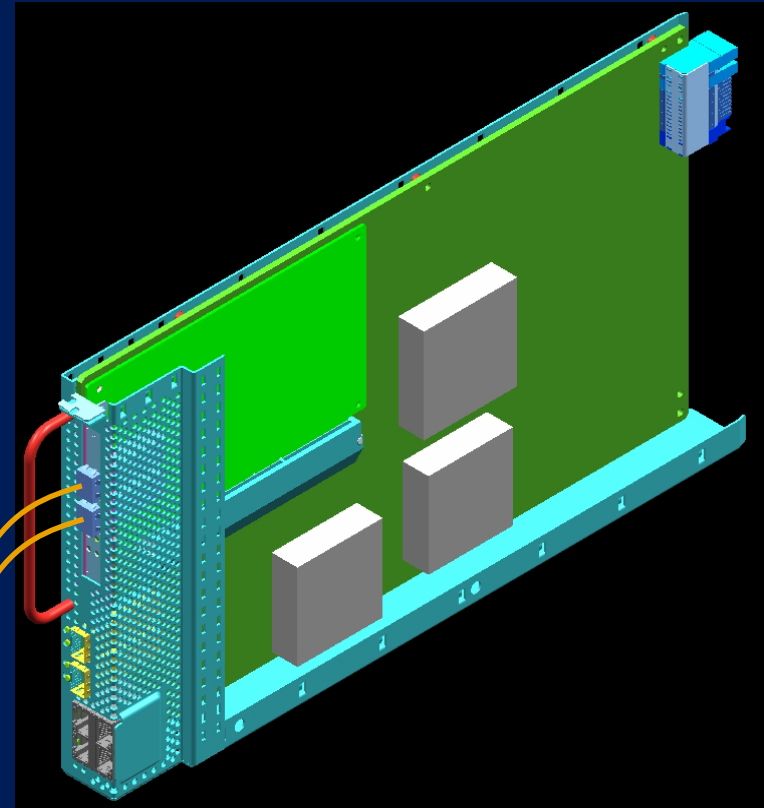
Fibre Channel

- An established standard
- The preferred interconnect for storage network
- Serial over optical fiber
- Data transmitted in frames protected by 32-bit CRC
- Fast data rates: 2 Gbit/s, future: 4 and 10 Gbit/s
- Long distances: 250 M, with switches: 10 KMs (for ESS)
- Large number of devices
- Topologies: arbitrated loop for FC disks,
 - point-to-point, and switched fabric (for ESS)



Fibre Channel ServerNet adapter

- Modular I/O adapter form factor
- High Performance, a powerful embedded RISC processor
- Very large number of concurrent I/Os
- 4 x 2.5 Gb/s ServerNet ports
- 2 x 128 MB on board buffer
- Industry standard Fibre Channel PCI card
- 2 x 2 Gb/s Fibre Channel ports



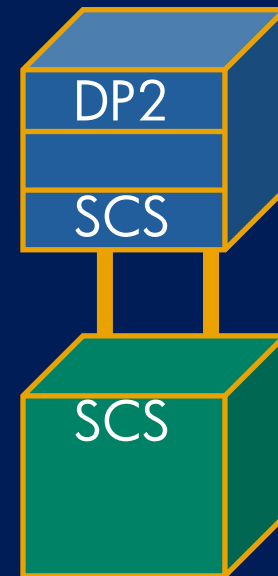
ServerNet connection services

- Multi-path ServerNet connection services
- Uses all available paths
- Automatic data transfer and path recovery
- Allow connection from all processors
- Common storage and networking code

Processor

ServerNet

Adapter



Fibre Channel disk module (FCDM)

Primary disks



Mirror disks



- Compact packaging: up to 14 disks per module
 - up to 140 disks per modular I/O enclosure
- Directly attached to NonStop servers

Fibre Channel disks

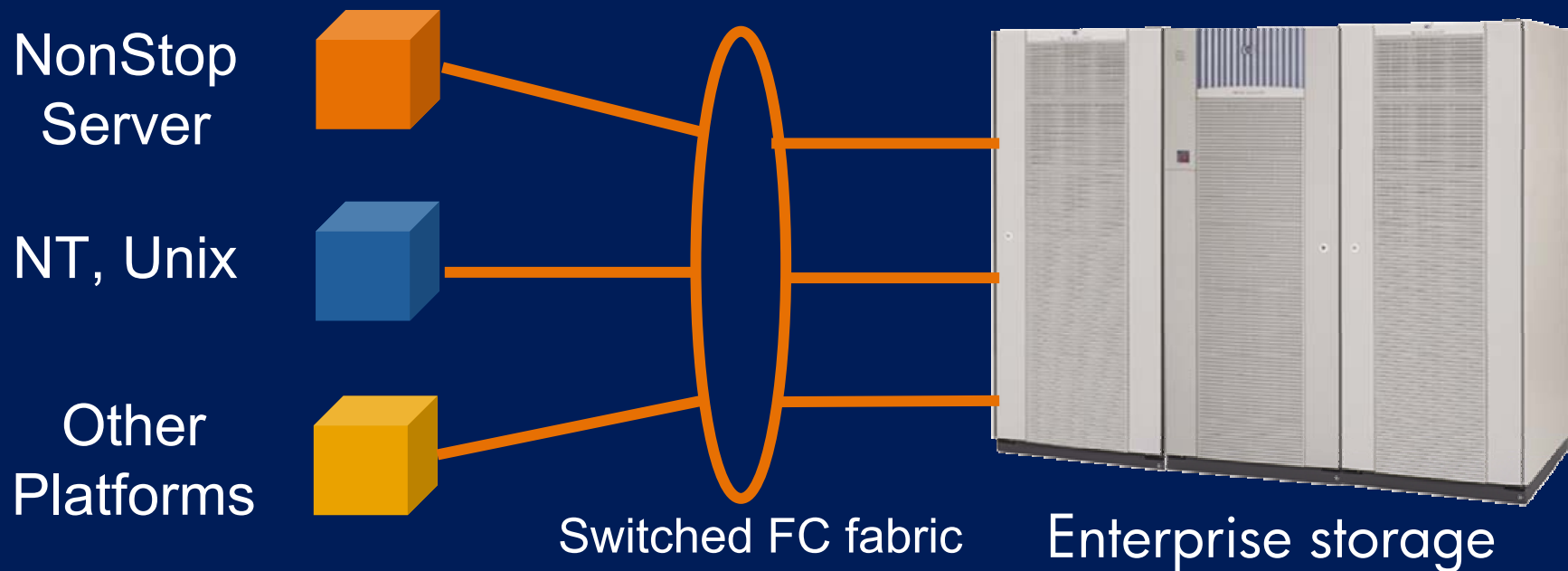


Adapters

Adapters

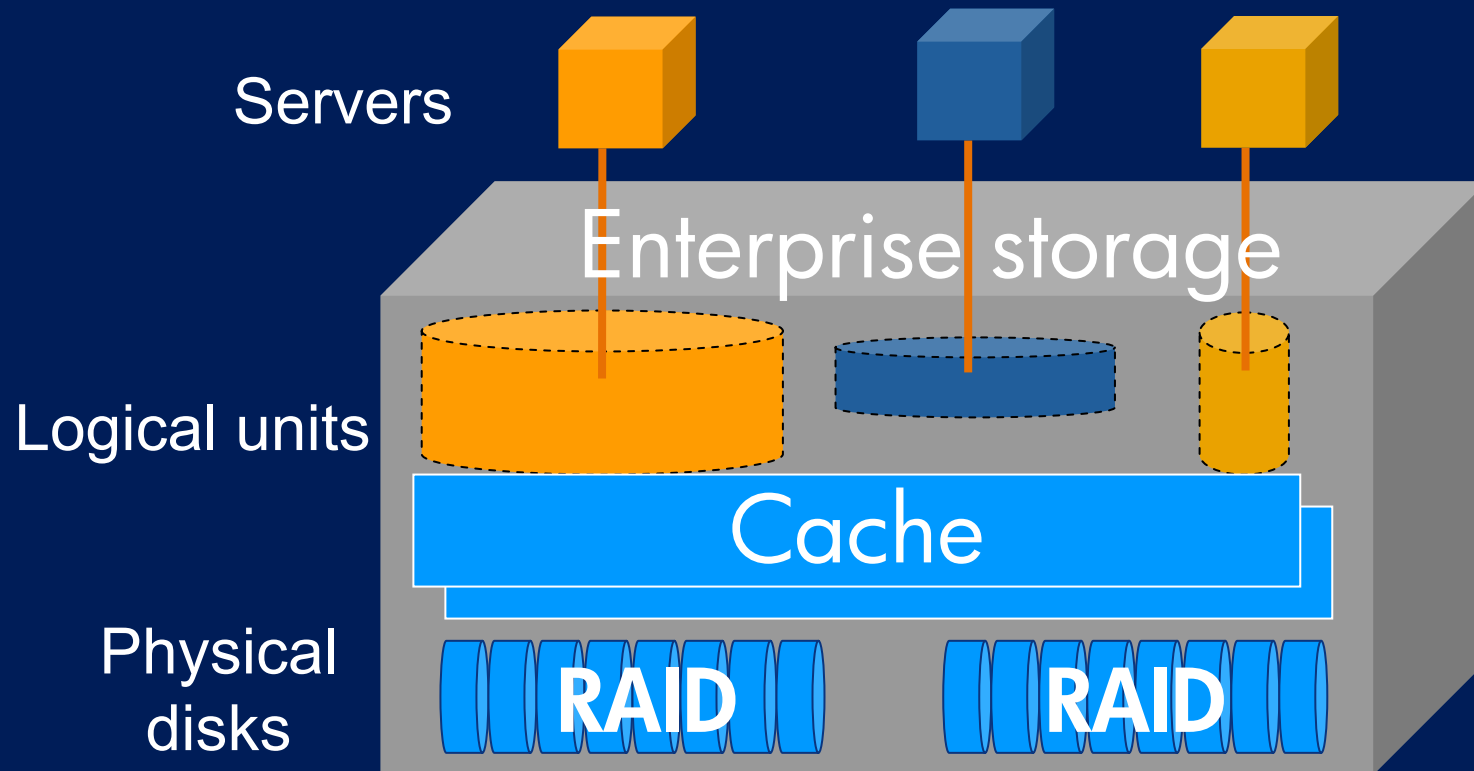
- Dual ported 2 Gbit/s Fibre Channel drives
- Dual Fibre Channel arbitrated loops
- 36 and 73 GB, 15 K RPM disks at FCS
- Future: 146 GB and larger

Enterprise storage



- Storage system with intelligent firmware
 - Highly available, terabytes of storage
 - Multi platform support
- No data interchange between platforms

Storage virtualization

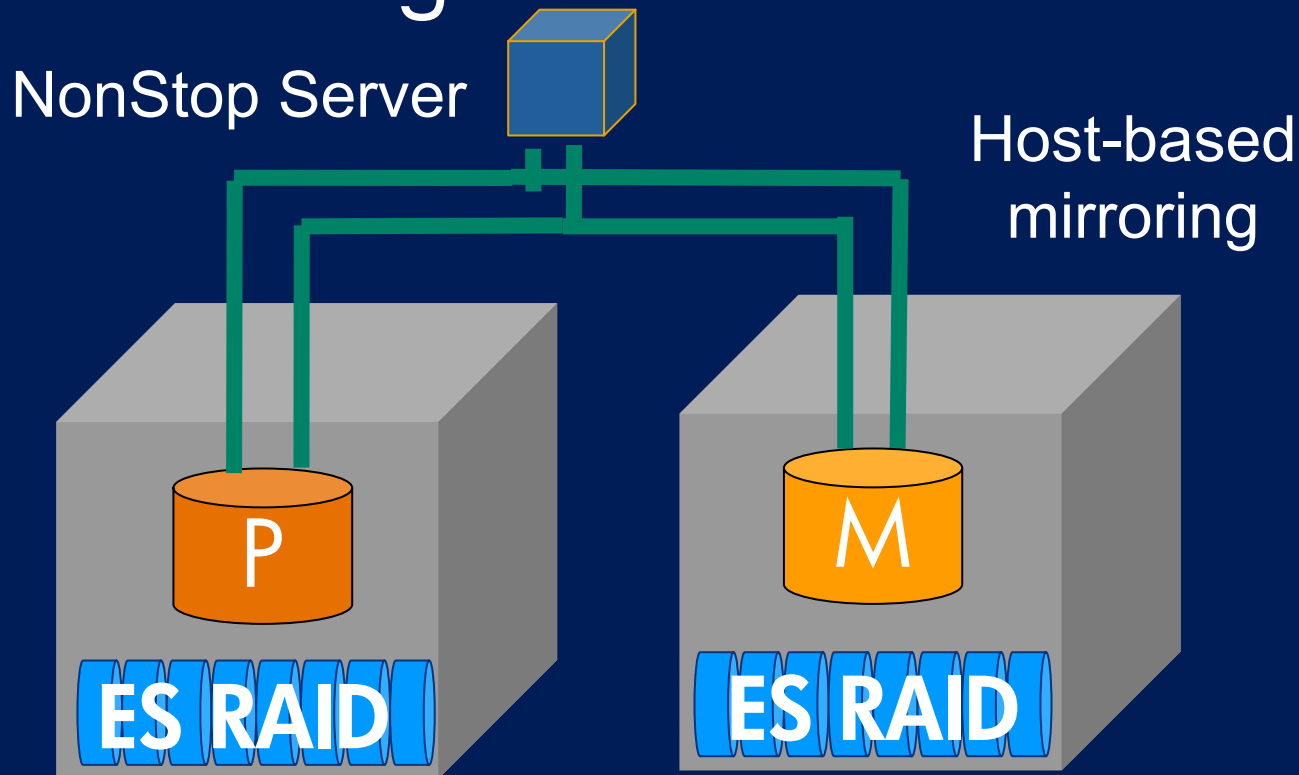


- Servers access logical units.
- Enterprise storage manages physical disks
- Logical units can be enlarged or relocated within the storage

Benefits of enterprise storage

- Management:
 - Unified and centralized
- Capacity:
 - Pooling of storage capacity
- Performance:
 - Many host adapter ports, large caches
 - and balancing of disk hot spots
- Flexibility:
 - Dynamic capacity, reallocation,
 - can be partitioned to several servers
- Availability:
 - Redundant components and RAID
 - (Redundant array of independent disks)

Redundant configurations



- ES RAID is for protection against physical disk problems
- Host-based mirroring is for protection against ES problems
- UPS (Uninterruptible Power Supply) is strongly recommended

HP XP



- Highly scalable up to 1024 disks, 149 tera byte
- High total performance: 500 K IOPS (cache), 10 GB/s
- 64 GB cache, 64 host ports, 8192 logical units
- Enhanced, tested, qualified, sold, supported by HP
- Single vendor for server and storage

Switched Fibre Channel fabrics

NonStop
Servers



Fibre Channel
switches

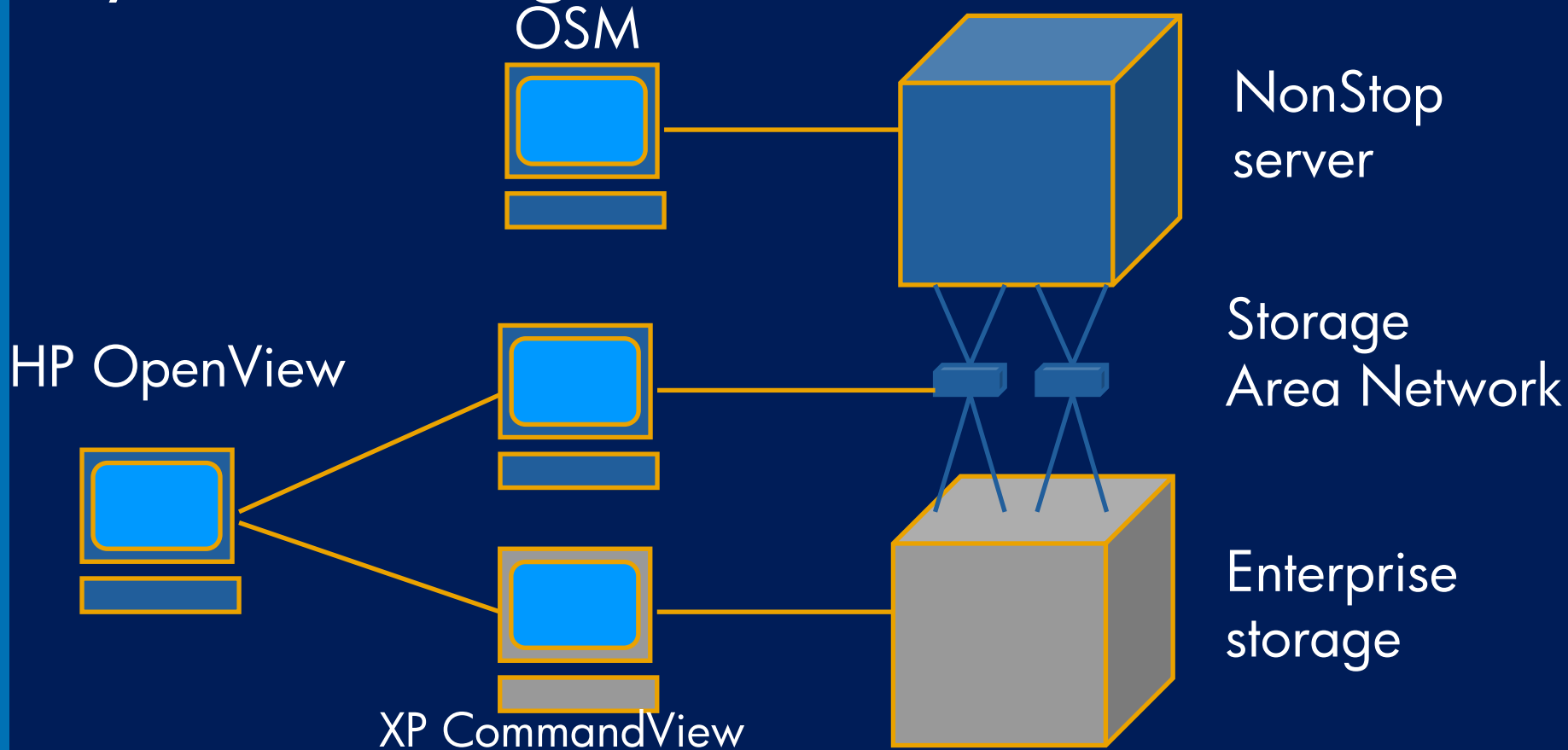


Enterprise
storage



- FC network allows any-to-any connection
- FC network can be dedicated, shared or zoned
- Redundant independent paths are required for fault tolerance

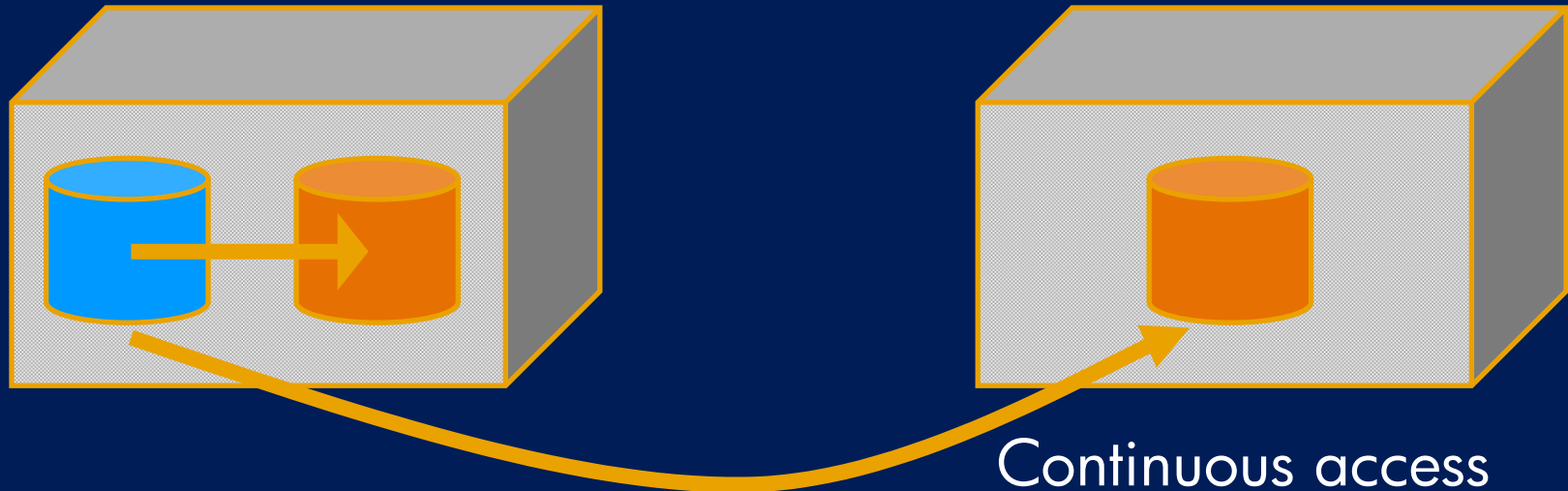
System management



- Specialized management systems for server, ESS and SAN

Disk copy

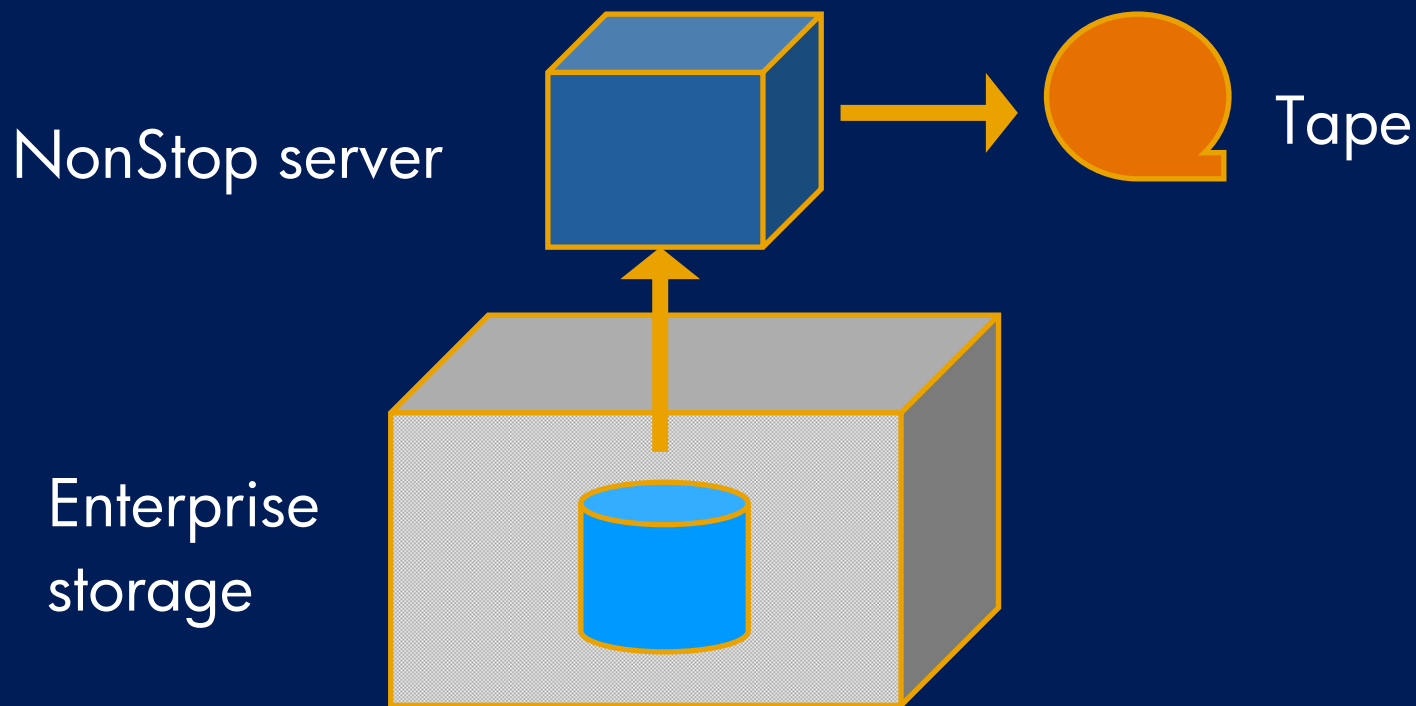
Business copy



- Disk must be quiesced momentarily to split pair
- Similar to removing a mirror disk
- Multi disk database must be quiesced manually to have a consistent backup

File backup to tape

- Standard Backup Restore



Storage options



Internal
Direct-attached storage



FC disks (2005)



Enterprise storage

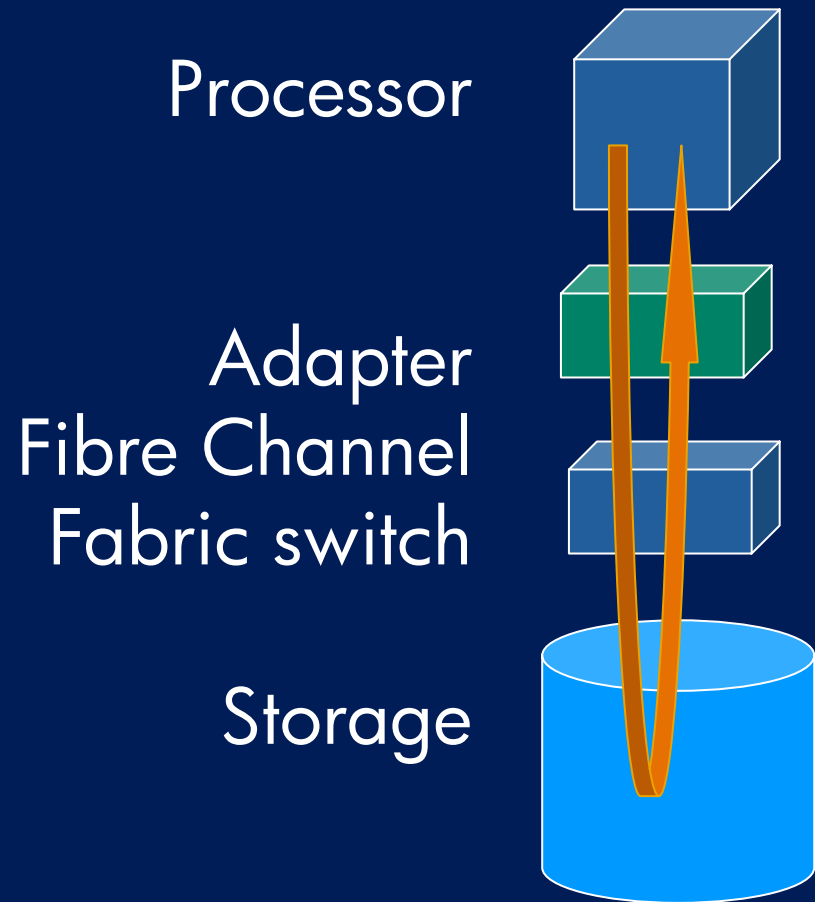
- Configure the storage to adapt to the enterprise
- Standard disk process serves all three types of storage
- The same application can run on all three types of storage

New Data Protection

New data protection

End-to-end checksum

- All storage is protected by checksum
- Disk checksum is calculated and checked in the processor



Checksum for new storage

- Industry standard disk sector size (512-byte)
- Better protection: 6 or 8-byte checksum
- All data (structured, unstructured files, audit trail and other structures on the disk)
- Transparent to applications (requires no changes)
- Internal SCSI disks will continue to use the current checksum





Online SQL partition migration

Internal SCSI disks

FC disks or ESS

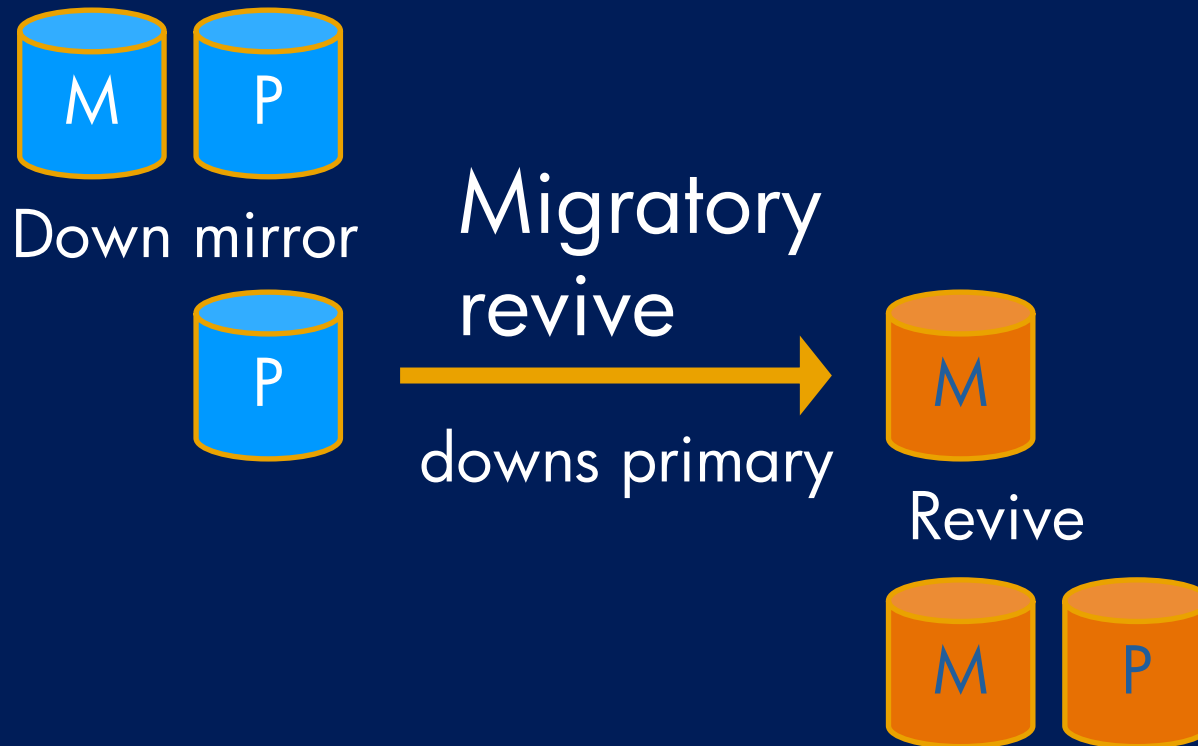


- Audited SQL partition
- Standard SQL command

Online disk migration

Internal SCSI disks

FC disks or ESS

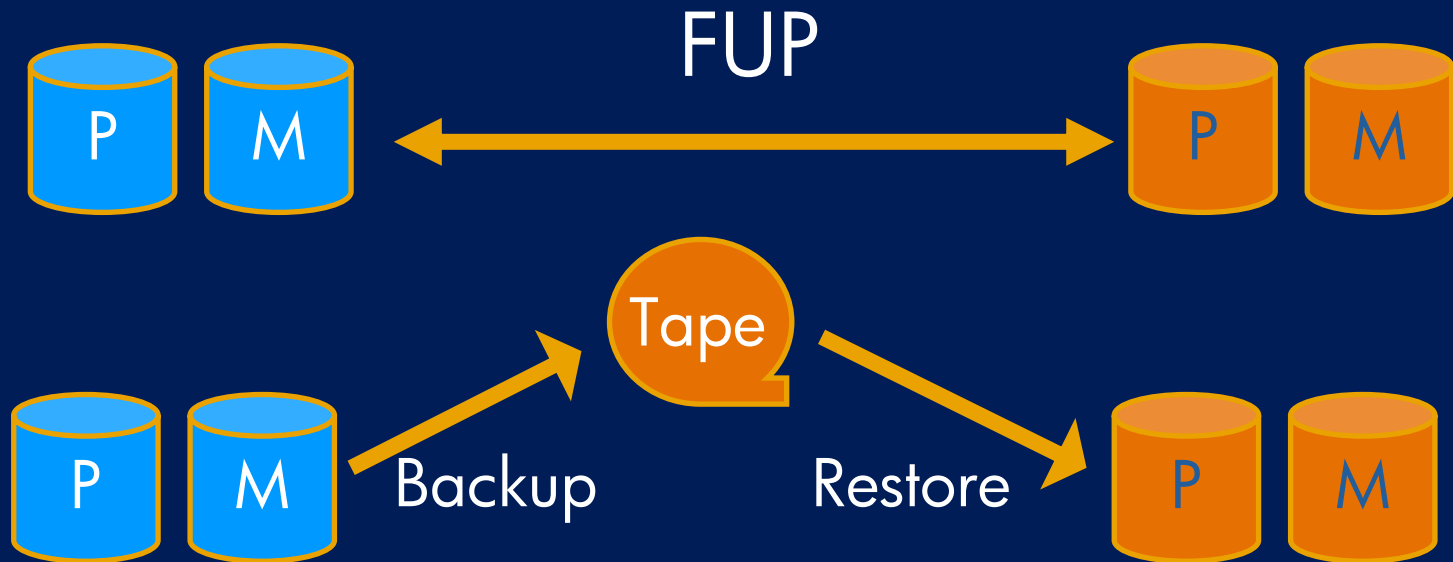


- Only from old to new checksum
- Similar to the direct disk online capacity upgrade

Offline data migration

Internal SCSI disks

FC disks or ESS



- Standard utilities:
 - FUP, File Mode Backup Restore, TMF Dump and Recover
 - not Volume Mode Backup
- Transfer files from old to new
 - and from new to old

There are two ugly parts of a computer system: Tape drives and Printers



- They involve humans
- Require supplies

Tape Storage and NonStop™ Servers

Tape Storage Products Update

- Low End: DAT, Table Top DLT 7000
- Mid Range: Super DL T, HP Ultrium LTO Gen 2
 - LTO is generally available
 - LTO is a new tape format – Requires migration planning
 - for customers moving from other tape platforms
- High End / Industrial Strength: STK “Eagle” 9840
- Additional tape storage products include tape libraries and
- long distance connectivity
 - CTL700/M tape library for mid-range and high end tape
 - SCSI extenders
 - Fibre Channel

New DAT Tape Drive for NonStop Servers

- HP StorageWorks DAT 72 tape drives
- Benefits
 - Utilizes the successful and popular DDS technology
 - Provides extraordinarily low cost of ownership
 - Provides backward compatibility with previous DDS generations
 - Enhanced and expanded user interface
- 28% lower list price and improved cost of ownership

**MORE STORAGE
FOR LESS**

Mid Range Tape Products

- HP Ultrium LTO Gen 2

- N1521A For use with CTL700/CTL700M Tape Library –
Replaces 5259 Super DLT

- N1524A 7186/7/8 cabinet mounted, with auto cartridge loader (10 cartridges)

- N1525A table top drive, with auto cartridge loader (10 cartridges)

- Generally available

- New tape format!

- Requires migration planning for customers moving from other tape platforms

Virtual Tape on NonStop Servers

- NED has partnered with two vendors to create a virtual tape offering
- Virtual Tape benefits to NonStop customers:
 - Improved throughput of tape operations, especially Restore
 - files restored from virtual tape are available in seconds as compared with hours in a physical tape environment
 - Improved flexibility in physical tape backups – file sets, retention, etc.
 - Physical backups take place “in the background” from the virtual tape environment
 - Improved media utilization; volume “stacking”
 - Enterprise wide, heterogeneous leveraging of tape storage resources
 - Total cost of solution is less than half of the IBM equivalent

Virtual Tape On NonStop Servers

Product Overview



- Virtual Tape is essentially tape emulated storage on disk
 - The following remain unchanged on NonStop servers
 - Physical SCSI connection from the NonStop to the tape device
 - Tape Process on the NonStop
 - Tape operations on the NonStop
 - Virtual tape looks like any other physical tape drive from an operations/end user perspective

Packaging

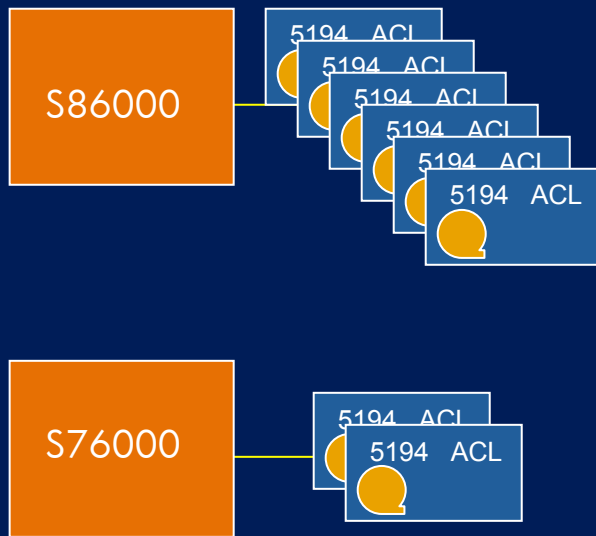
Virtual tape is:

- Hardware from other HP product divisions
 - ISS – ProLiant Server
 - NSS – StorageWorks MSA 1000 RAID Array
- Software developed by each supplier

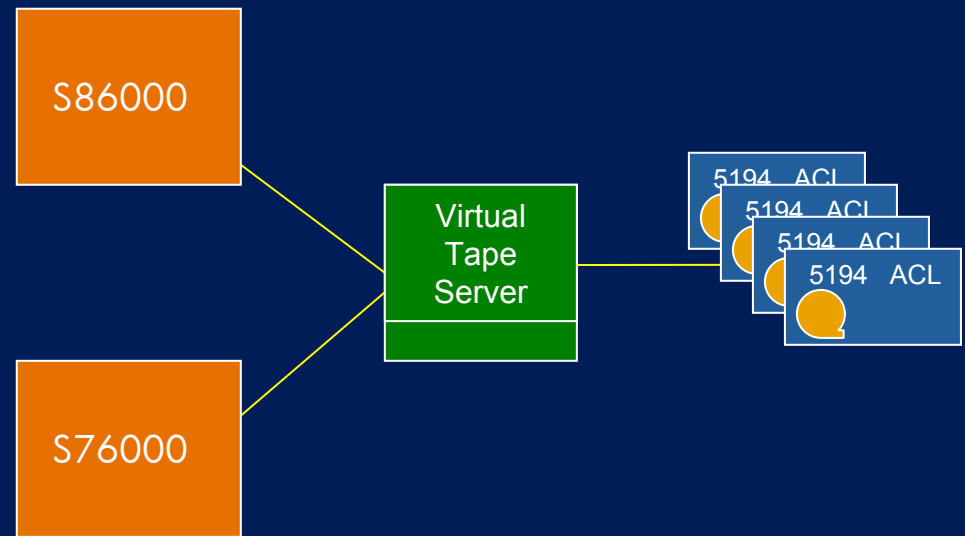
Virtual Tape Implementation on NonStop Servers



Physical Tape



Virtual Tape method



Disclaimer: Future product plans, dates, and functionality are subject to change without notice.

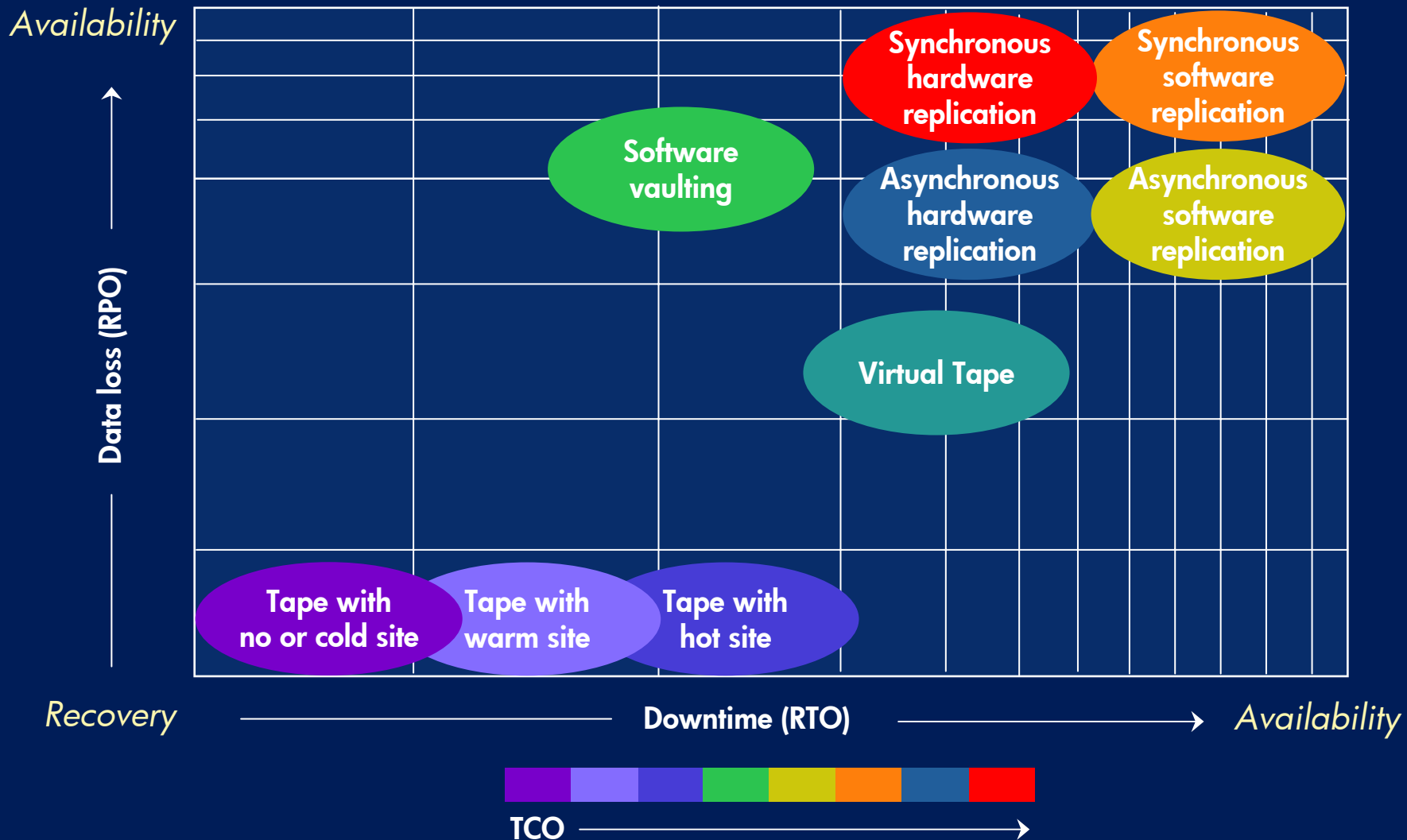
NonStop™ TMF and RDF

- ZLT – Zero Lost Transactions

Separating downtime and data loss

- RTO (recovery time objective)
 - How soon after an event does the business process need to be available?
 - Not all business processes need to be available at the same time.
- RPO (recovery point objective)
 - How much work in progress can be lost?
 - Not all work needs to be recovered to the same time.

It's a continuum



Integrated disaster tolerance and recovery products for the HP NonStop server



- NonStop TMF
 - The foundation for transaction integrity and data protection
- NonStop RDF
 - High-performance database replication

What can NonStop TMF do for an application?



- NonStop TMF software is the foundation for any 24 x 7 IT operation
 - Database reorganization, partition split/merge/move, and backup to disk or tape with zero application downtime
 - Near-real-time replication to a backup system
- Online database backup
 - One block split can ruin your whole day
 - Do you run tape retrieval and restore tests?
- Recovery from accidental or intentional database modification
 - Part or all of the database can be recovered to a point in time, including to just before that “drop table” command
 - Using third-party tools, accidentally deleted records can be reinserted

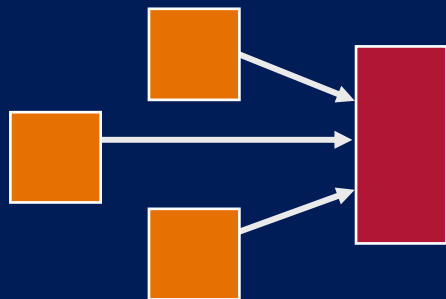
What is NonStop RDF software?

- High-speed, low-latency database replication software
- Peer-to-peer for NonStop servers only
- Focused on disaster tolerance
- Uses low-level system interfaces

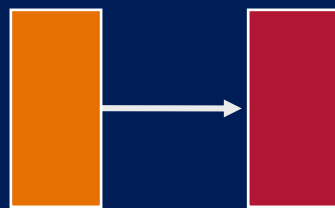


Some NonStop RDF topologies

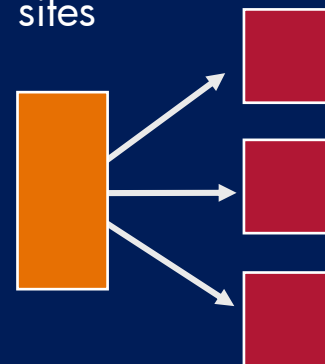
Centralized



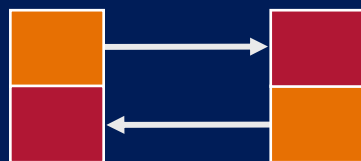
Simplex



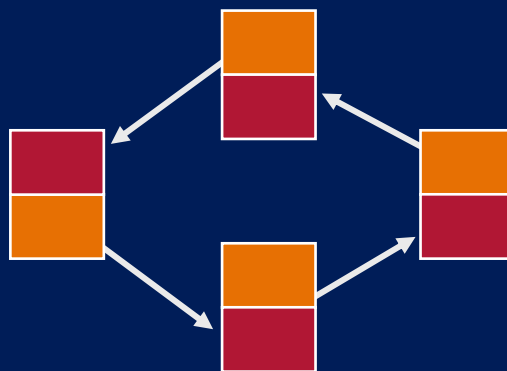
Multiple duplicate sites



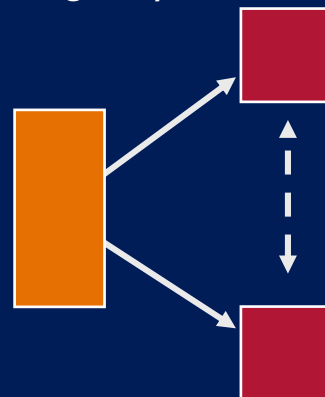
Reciprocal/
split workload



Ring



Triple contingency



Zero lost transactions (ZLT) overview

- All asynchronous replication products are vulnerable to losing some transactions if there is an unplanned outage of primary system
 - Some have lower latency than others
 - NonStop RDF has the lowest latency in the NonStop system environment because of its private audit trail reading and shipping logic
- Synchronous or lockstep replication
 - Guarantees no loss of transactions
 - Increases commit response times
 - Transaction cannot commit until data is safe in backup database
 - Perhaps valuable for low-volume applications

Disclaimer: Future product plans, dates, and functionality are subject to change without notice.

Goals for zero lost transactions

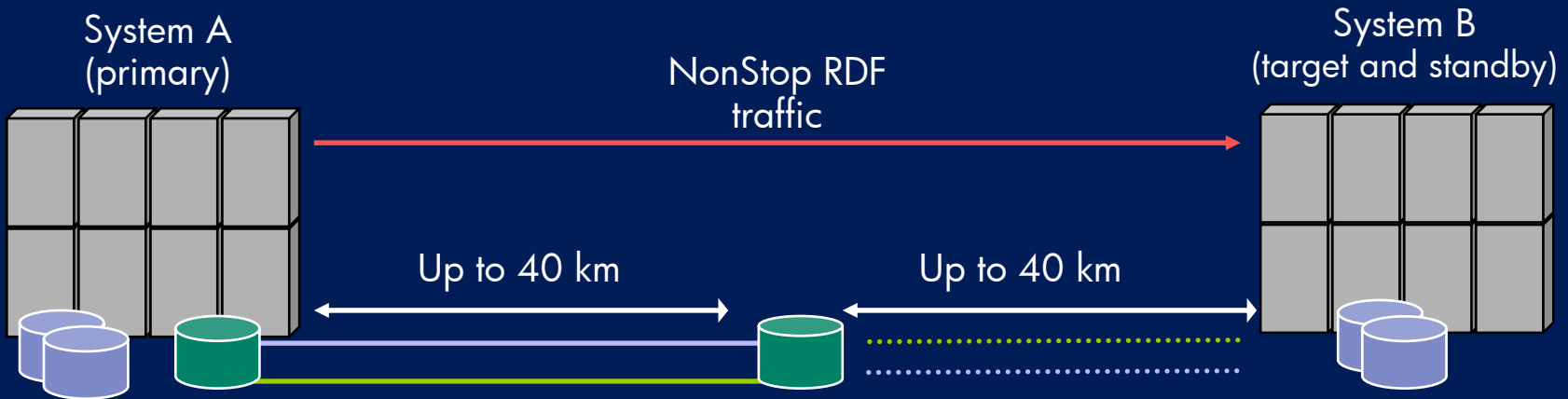
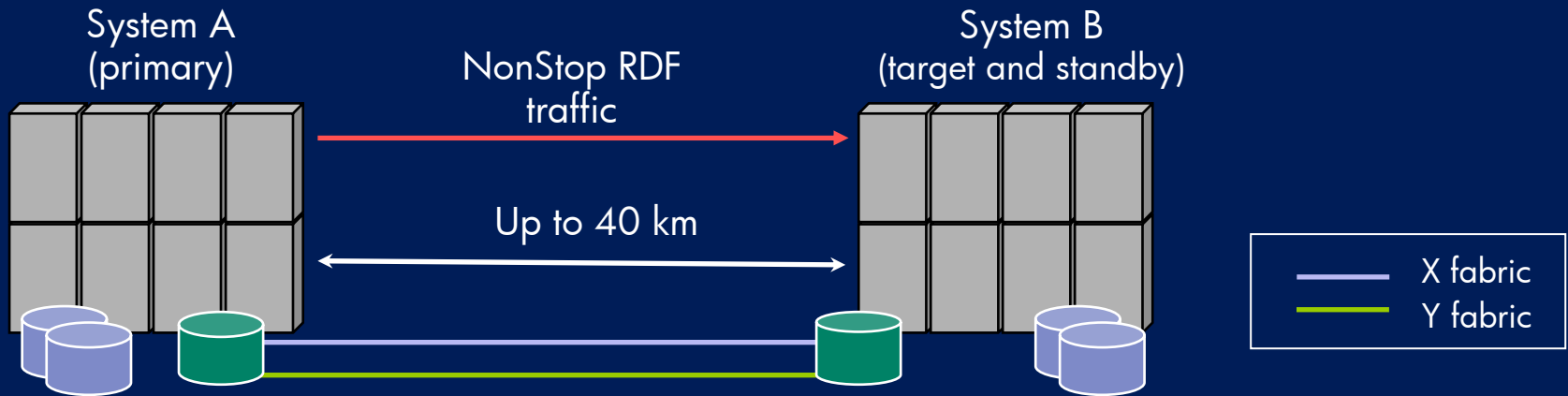
- Eliminate the loss of committed data after an unplanned outage
- Limit the increase in commit response
- Eliminate application changes
- Reduce NonStop RDF overhead on the primary system
 - Extractor overhead is consistently 1 percent of normal application costs

ZLT remote mirror

- For each audit trail in the NonStop RDF configuration
 - Use local and remote mirror
 - Remote mirror can be nomadic disk
 - Remote mirror can be up to 40 km from primary system
- NonStop RDF standby system
 - Can be a third system
 - Can be the backup system itself
- Remote mirror can be up to 40 km from standby system
 - Standby can therefore be up to 80 km from primary system
- If standby is not the target system, no distance limit for location of target

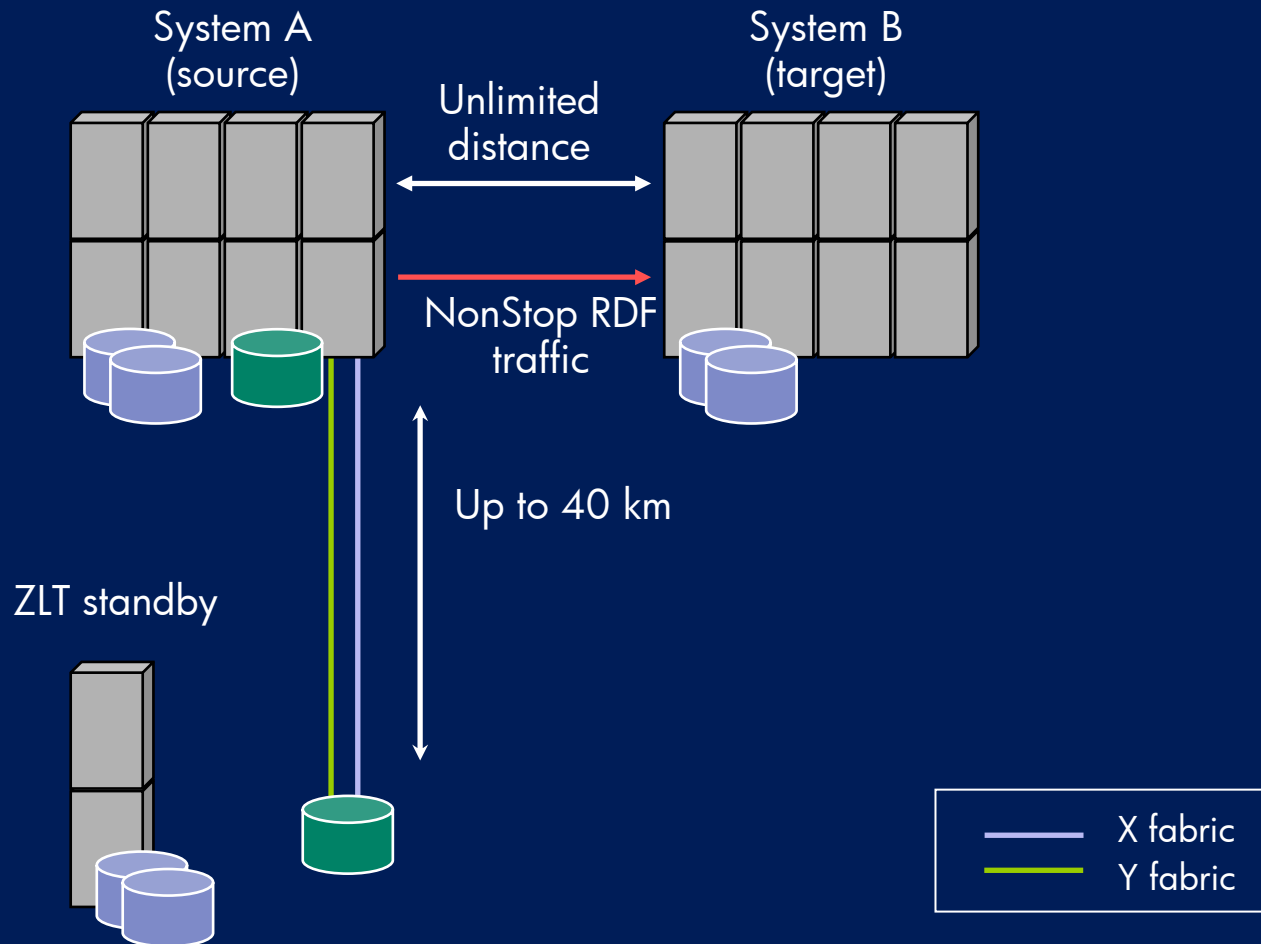
Disclaimer: Future product plans, dates, and functionality are subject to change without notice.

ZLT remote mirror with backup as standby

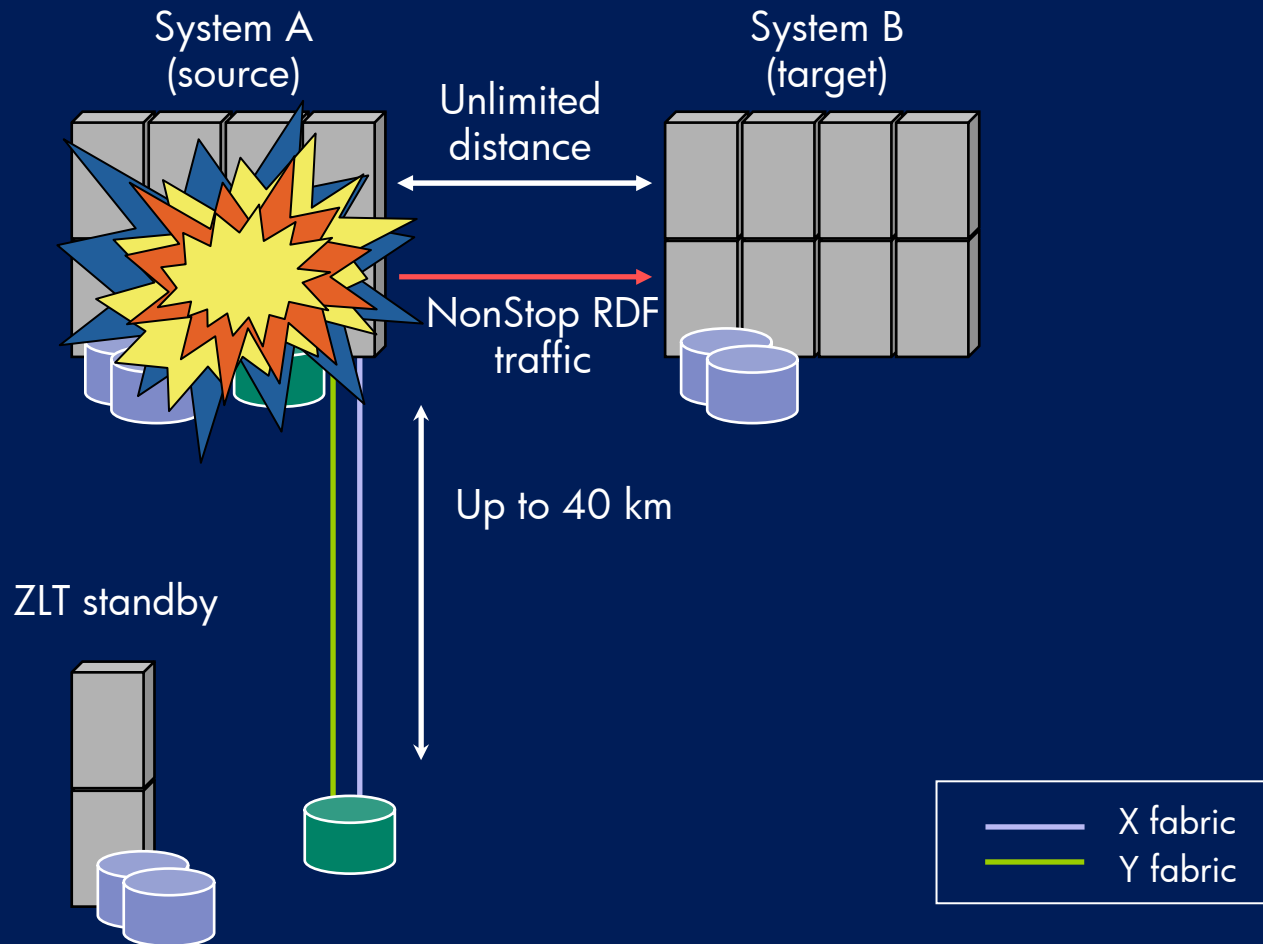


Disclaimer: Future product plans, dates, and functionality are subject to change without notice.

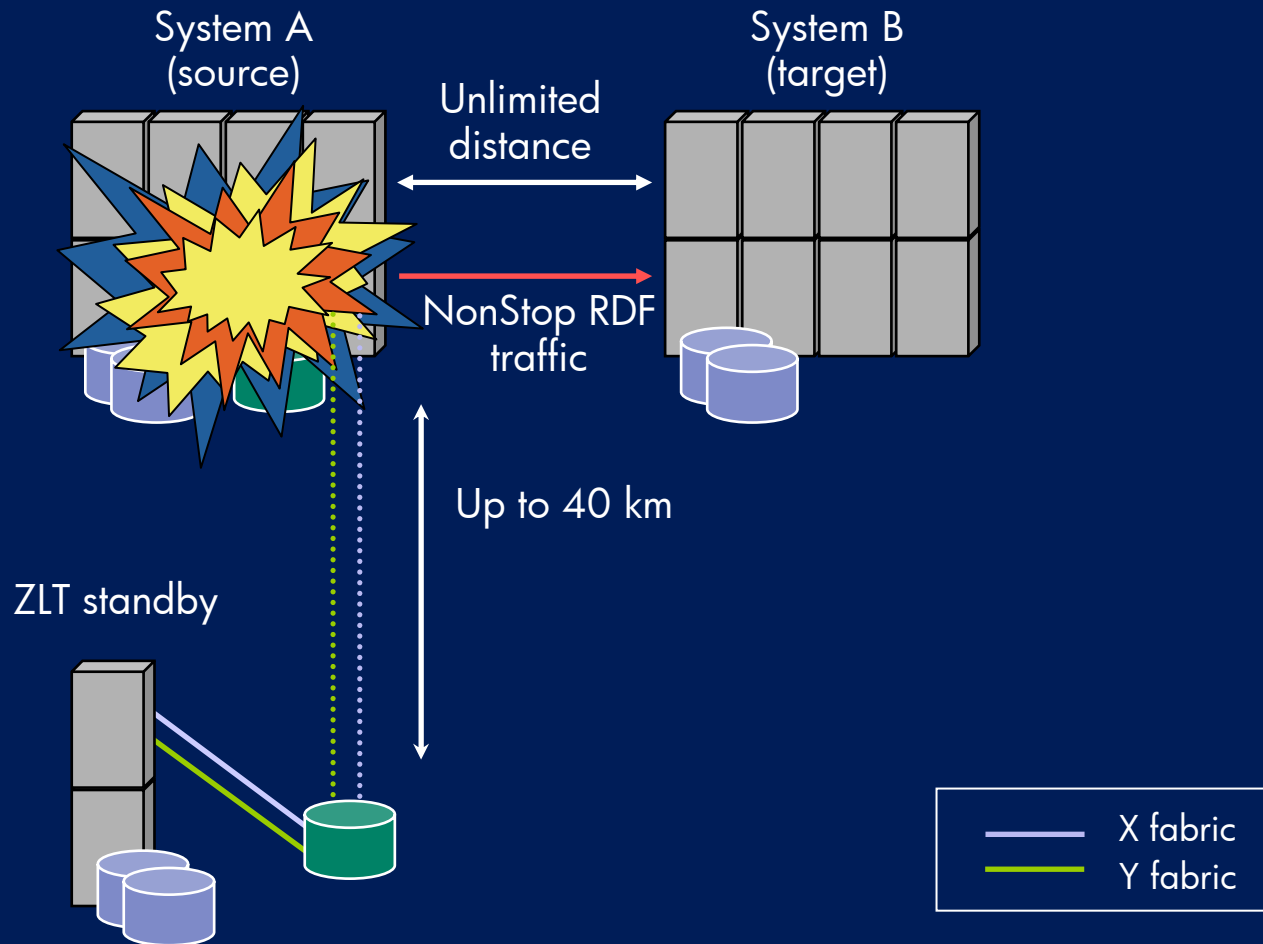
ZLT remote mirror with third system as standby Normal processing



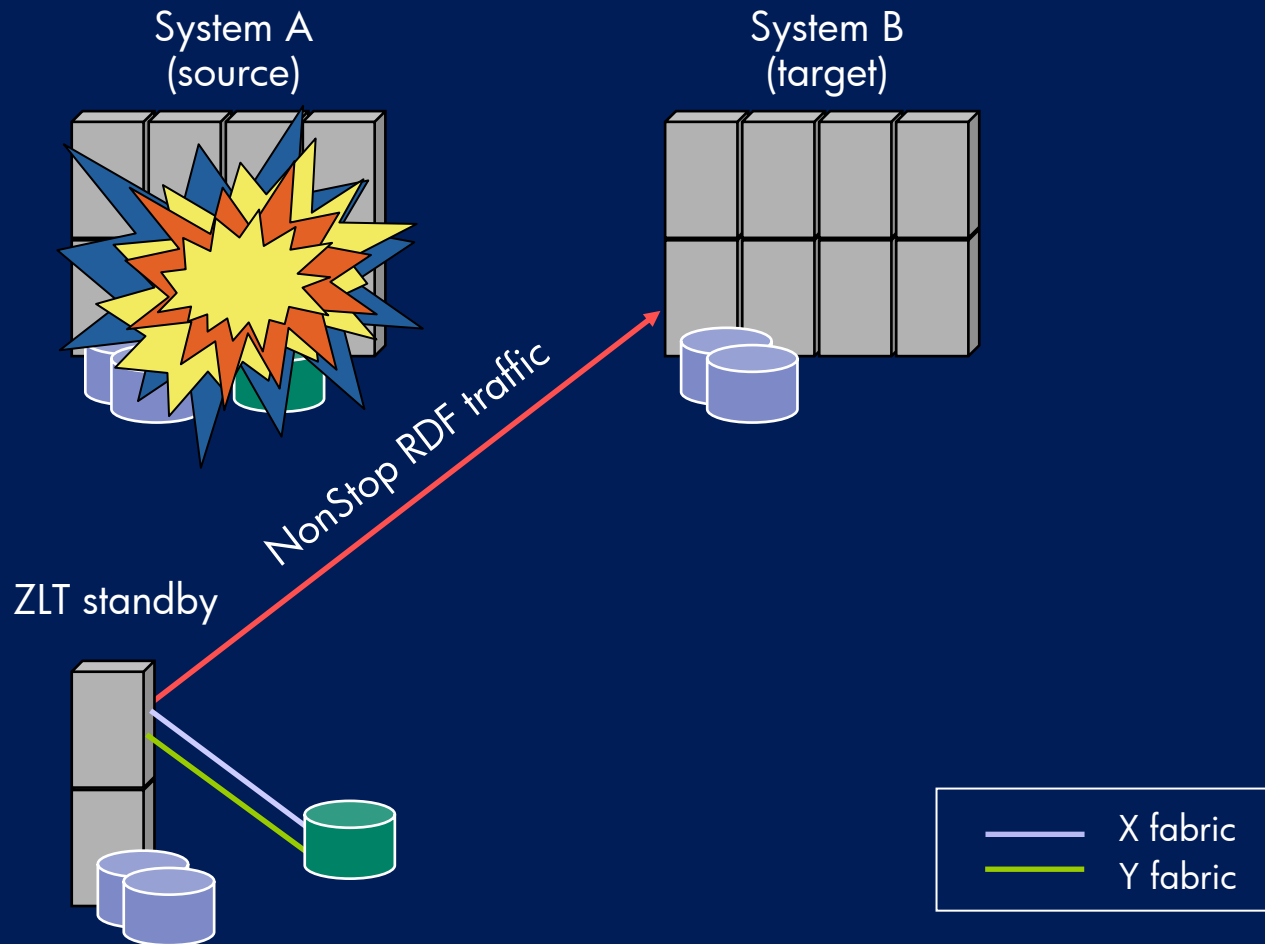
ZLT remote mirror with third system as standby Normal processing



ZLT remote mirror with third system as standby ZLT processing



ZLT remote mirror with third system as standby ZLT processing



ZLT remote mirror summary

- With access to the remote mirror, no transactions are lost
 - NonStop RDF reads all remaining audit prior to actual takeover processing
 - Low-cost method that delivers the same result as lockstep
- Has distance limitation of 40 to 80 km
- Performance considerations
 - Length of cable to remote mirror can affect application response times
 - Low to moderate audit generation rates may not affect applications
 - High audit generation rates can affect applications

Disclaimer: Future product plans, dates, and functionality are subject to change without notice.

Storage solutions for Adaptive Enterprise



- New features
 - New I/O architecture
 - New data protection
 - For the Adaptive Enterprise
-
- New solutions
 - New competitive edge
 - New business opportunity
 - For the Adaptive Enterprise

Don't be afraid to try something new

Remember that

- Amateurs built the ark
- Professionals built the Titanic



i n v e n t