bm.com

e-business

iSeries SAN

GP07

ITSO iSeries Technical Forum

Nick Harris

niharris@us.ibm.com

Redbooks

International Technical Support Organization

F03GP07.PRZ

Trademarks and Disclaimers

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.
The following terms are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both:

400	BRMS	Host Integration Series	JustMail	Payment Manager	Stylized e
ADSTAR	Client Series	Host on Demand	MQSeries	Payment Server	SystemView
Advanced Function Printing	ClusterProven	Host Publisher	MQSeries Integrator	PCOM	VisualAge for Java
AFP	CODE/400	HTTP Server for AS/400	Net.Commerce	PowerPC	VisualAge for RPG
AIX	DataGuide	IBM	Net.Data	PowerPC AS	WebSphere
AnyNet	DB2	IBM Logo	Netfinity	Print Service Facility	WebSphere Advanced Edition
Application Development	DB2 Extenders	IBM Network Station	NetView	pSeries	WebSphere Commerce Suite
APPN	DB2 UDB for AS/400	Information Warehouse	NUMA-Q	PSF	WebSphere Development Tools for AS/400
AS/400	DB2 Universal	Integrated Language Environment	OfficeVision	S/390	WebSphere Standard Edition
AS/400e	e-business logo	Intelligent Printer Data Stream	OS/2	SanFrancisco	Workpad
AT	e(logo) Server	IPDS	Operating System/400	Screen Publisher	xSeries
BrioQuery	Enterprise Storage Server	iSeries	OS/400	SmoothStart	

other countries, or both.
Tivoli and NetView are trademarks of Tivoli Systems Inc. in the United States, other countries, or both.
C-bus is a trademark of Corollary, Inc. in the United States, other countries, or both.
Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.
PC Direct is a trademark of Ziff Communications Company in the United States, other countries, or both and is used by IBM Corporation under license.
ActionMedia, LANDesk, MMX, Pentium and ProShare are trademarks of Intel Corporation in the United States, other countries, or both.
IBM's VisualAge products and services are not associated with or sponsored by Visual Edge Software, Ltd.
Linux is a registered trademark of Linus Torvalds.
UNIX is a registered trademark of The Open Group in the United States and other countries.
SET and the SET Logo are trademarks owned by SET Secure Electronic Transaction LLC.
Other company, product and service names may be trademarks or service marks of others.

Information is provided "AS IS" without warranty of any kind.

All customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer.

Information in this presentation concerning non-IBM products was obtained from a supplier of these products, published announcement material, or other publicly available sources and does not constitute an endorsement of such products by IBM. Sources for non-IBM list prices and performance numbers are taken from publicly available information, including vendor announcements and vendor worldwide homepages. IBM has not tested these products and cannot confirm the accuracy of performance, capability, or any other claims related to non-IBM products. Questions on the capability of non-IBM products should be addressed to the supplier of those products.

All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. Contact your local IBM office or IBM authorized reseller for the full text of the specific Statement of Direction.

Some information in this presentation addresses anticipated future capabilities. Such information is not intended as a definitive statement of a commitment to specific levels of performance, function or delivery schedules with respect to any future products. Such commitments are only made in IBM product announcements. The information is presented here to communicate IBM's current investment and development activities as a good faith effort to help with our customers' future planning.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput or performance improvements equivalent to the ratios stated here.

Photographs shown are of engineering prototypes. Changes may be incorporated in production models.

Objectives

The objective of this presentation is to describe the storage support, features and scenarios available at V5R1 and V5R2 of OS/400.

This presentation also includes updates for the January 2003 announcement.

iSeries offers SAN support. We will position iSeries and SAN. We will also cover what this SAN support and implementation.

Agenda

- iSeries Storage V5R1
- iSeries Storage V5R2
  - January 2003 announcement
- More about iSeries Storage
- HA or DR - you get what you pay for.....
- Independent ASPs
- Positioning iSeries storage methods
- iSeries external storage attachment
- iSeries Storage in a SAN
- iSeries External storage migration
- iSeries external storage configurations
- External disk overview
- Storage Area Networks overview
- Sources of information
- Additional Information

iSeries Storage V5R1

V5R1 Storage Announcement Highlights

Storage Area Network (SAN) enablement for iSeries

- FC adapters
- Point-to-point and Arbitrated Loop support

Switched disk capability through 1 GB High Speed Link (HSL)

Dynamic disk add for Windows Servers

Disk add can take some time

- PTF to put on is MF27545, improves 'add' LUN performance

Notes V5R1 Storage announcements

With V5R1 the iSeries was first enabled for Storage Area Network (SAN).

- There are two Fibre Channel (FC) adapters #2765 Tape adapter and #2766 Disk Apapter.
- The FC connnection methods are Point-to-point and Arbitrated Loop.

The iSeries disk storage was enhanced with Independent Auxiliary Storage Pools (iASPs). These new storage areas are able to hold IFS objects. These new storage spaces are further enhanced by switchable capability enabled with iSeries Cluster Services and through the 1 GB High Speed Link (HSL).

OS/400 Windows Integrate added to its SAN like disk functions with V5R1 and dynamic disk add support for Windows Servers. This allows you to define an additional 'windows disk' using the Operations Navigator GUI. When linked this disk will automatically appear in windows without having to restart the Windows Server.

While the actual amount of time will vary based on a number of things, including how the LUNs are configured on the ESS and how busy the box is, there was a V5R1 PTF put out in late December '0. This PTF significantly improves the amount of time to add the Shark LUNs to the system ASPs.

The PTF to look for is MF27545. With this PTF, there is no need to do the Initialization and format prior to the add to ASP.

Fiber Channel adapters at V5R1

New #2765 FC Tape adapter

FC distance 10 kilometers

Arbitrated Loop

One target per loop

Hubs support

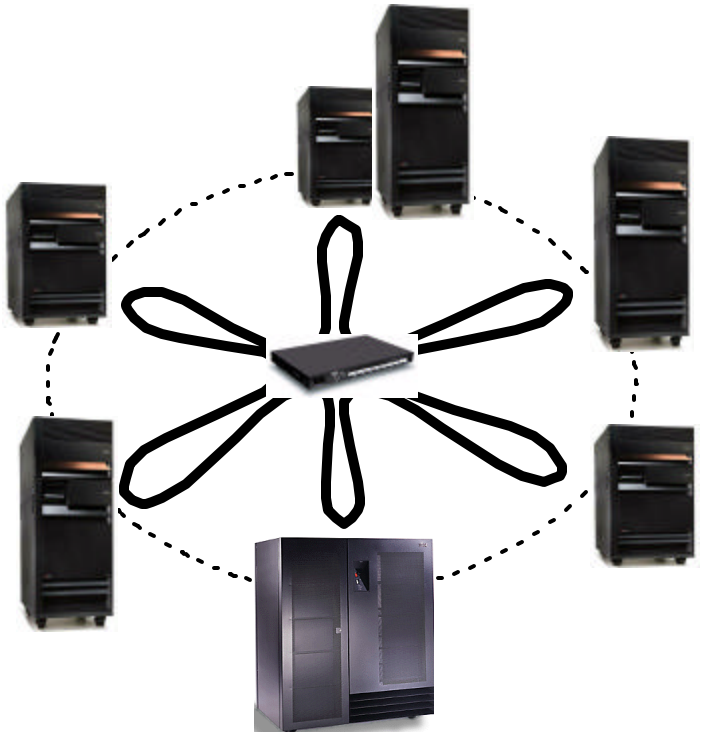
New #2766 FC Disk adapter

FC distance 10 kilometers

Point-to-point or Arbitrated Loop

5 initators per loop

Hub support



Notes Fiber Channel adapters at V5R1

The #2765 PCI Fibre Channel Tape Controller provides fibre channel attachment capability for external tape devices. The #2765 supports arbitrated loop topologies. With the use of managed hubs the distance tapes can be removed from the systemare greatly increased to up to 10km. One tape attachment is permitted for each #2765.

Each #2765 is shipped with a wrap connector (PN#05N6767). The following options are available to attach SC-type fibre cables :

- #0371 is a two meter LC-SC Adapter kit that can be used to connect the #2765 to a 50µm (50 micron) cable.
- #0372 is a two meter LC-SC Adapter kit that can be used to connect the #2765 to a 62.5µm (62.5 micron) cable.

Fibre channel attachment for tape drives offers tremendous performance capabilities and long distance options. It is also easier for you to share these valuable resources with multiple systems.

The #2766 PCI Fibre Channel DASD Controller provides fibre channel attachment capability for external disk devices. The #2766 supports point-to-point and arbitrated loop topologies. With the use of managed hubs the ESS containing disk can be up to 10km away from the server by using managed hubs. Up to five iSeries (initiators) can be connected in an arbitrated loop.

Each #2766 is shipped with a wrap connector (PN#05N6767). Just as for the #2765, the following options are available to attach SC-type fibre cables :

- #0371 is a two meter LC-SC Adapter kit that can be used to connect the #2766 to a 50µm (50 micron) cable.
- #0372 is a two meter LC-SC Adapter kit that can be used to connect the #2766 to a 62.5µm (62.5 micron) cable.

Some iSeries customers may find SAN-attached DASD devices to be appealing for their environment. If consolidating large amounts of DASD from different platforms is important, you should consider SAN. Note, however, a complex commercial business environment usually requires good, predictable response time to maintain user productivity and satisfaction. Carefully consider the performance implications of sharing resources in this environment, as the sharing may introduce more variable performance. For these critical workloads, dedicated direct attach DASD resource can ensure more predictable performance.

Fibre Channel Adapters require OS/400 V5R1 or later.

IMPORTANT NOTE: IBM withdrew the #6501SCSI disk adapter from marketing on July 31, 2001. The only attachment to IBM ESS is now through #2766 Fibre Channel Adapter on iSeries and OS/400 V5R1.

iSeries Storage at V5R2

April iSeries Storage announcements

2Gb/sec SAN support

- new cable distances for 2Gb switches, 150m and 300m

Switched fabric support

- Recommendation 2 ISL (inter switch links), max 3

New HSL-2 support

FC adapter for Linux partition

- #0612 - Linux Dir Attch-2766

Switched disk support

- most native objects

Notes iSeries Storage announcements

With the iSeries announcements in April 2002 comes new, exciting FC enhancementst.

V5R2 provides 2Gb/sec SAN support. This allows the iSeries to connect to 2Gb switched fabric.

- The 2 Gb support has new cabling distances to attached to local devices or switches. The 62.5 micron cable distance is150m and the 50 micron cable distance is 300m. These are compared to the 1Gb distances of 175m for xx micron and 500m for yy micron FC cables.

Switched fabric support has been added to the connection method of iSeries FC.

- The new supported switches are 2109-F08/F16. The current switches are still supported as managed hubs or switches if they allow that type of fabric connection.
- A standard FC recommendation is that there should not be more than 2 (max 3) Inter-Switch Links (ISL) between the server and the device connected. This is a SAN recommendation, not related to iSeries support.

iSeries internal disk support is further enhanced with new HSL-2 support. This offers a faster connection between the iSeries system unit and expansion units containing I/O adapters and disk.

Before V5R2 a Linux partition on iSeries had to use network storage spaces or native SCSI attached disk. With V5R2 there will be support for FC attached disk via the #2766 FC disk adapter.

- this support will be available soon

Switched disk support has been enhanced at V5R2 by allowing most native OS/400 objects to be supported. This will allow many OS/400 applications to used iASPs and swtched disk support.

January 2003 announcements

- No new FC adapters
- New Models all support 2765 and 2766
  - Max of 2 #2765/#2766s (any combination) allowed in 270, 800, 810 CEC.
  - Maximum of 4 #2765/#2766s (any combination) allowed per model 820, 825 CEC.
- New Towers and Expansion units support 2765 and 2766
  - Maximum of 3 #2765/#2766s (any combination) allowed in #0595/#5095 tower.
  - A maximum of 6 #2765/#2766 (any combination) allowed in the #5074,#0578/#5078, #5094, #0588/#5088 in the top and bottom enclosures of #5079/5294 and in the top unit of a #8079/#8093/#8094.
- New #2844 IOP supports 2765 and the 2766

Fiber Channel adapters at V5R2

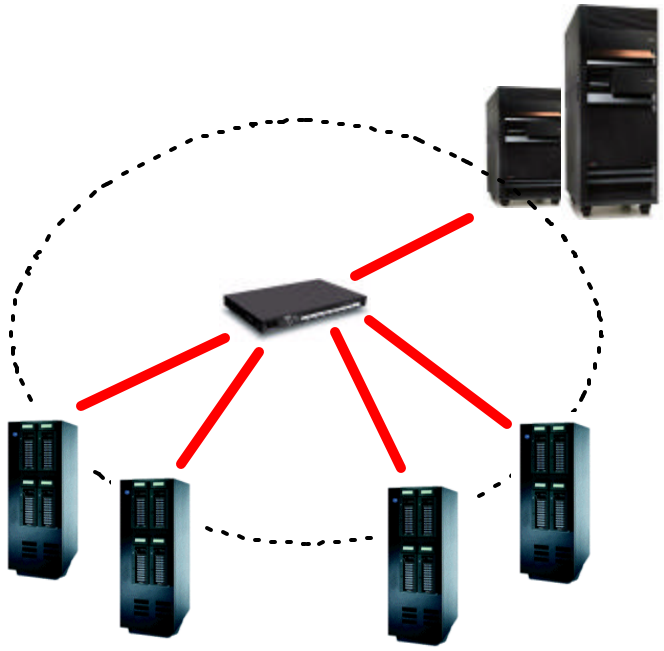
#2765 FC Tape adapter

FC distance 10 kilometers

Arbitrated Loop (last leg to tape)

Switched Fabric

Multiple targets (16 tape devices)



#2766 FC Disk adapter

FC distance 10 kilometers

Point-to-point, Arbitrated Loop

Switched Fabric

Still 32 LUNs across 1-32 sharks

Notes What's new with V5R2

The physical hardware of the #2765 FC Tape adapter remains unchanged at V5R2, but from a fabric prespective it now supports 2Gb/sec. This adapter was always rated at 2Gb/sec, with V5R1 this was limited to 1Gb/sec.

FC distance 10 kilometers, the total distance between switches or managed hubs across dark fibre remains at 10km.

All iSeries FC attached tape have the last segment of attachment to the tape drive setup as Arbitrated Loop.

New with V5R2 is Switched Fabric support. This enables iSeries to attach to tape drives across a switch SAN. This new supoort offers the benefits of redundancy and intelligent routing through the fabric. It is recommended for performance that there are not more than 2 Intersystem links between the server and the tape drive.

Each #2765 can access multiple targets (16 tape devices) with V5R2. For counting purposes most tapes are two devices, one for the tape and one for the loader. Whilst an individual #2765 can access 16 devices, it is doubtful it would achieve throughput to drive all 16 at rated speeds. A realistic number would be 2-3 tapes drives per adapter.

The physical hardware of the #2766 FC Disk adapter remains unchanged at V5R2, but from a fabric prespective it now supports 2Gb/sec.

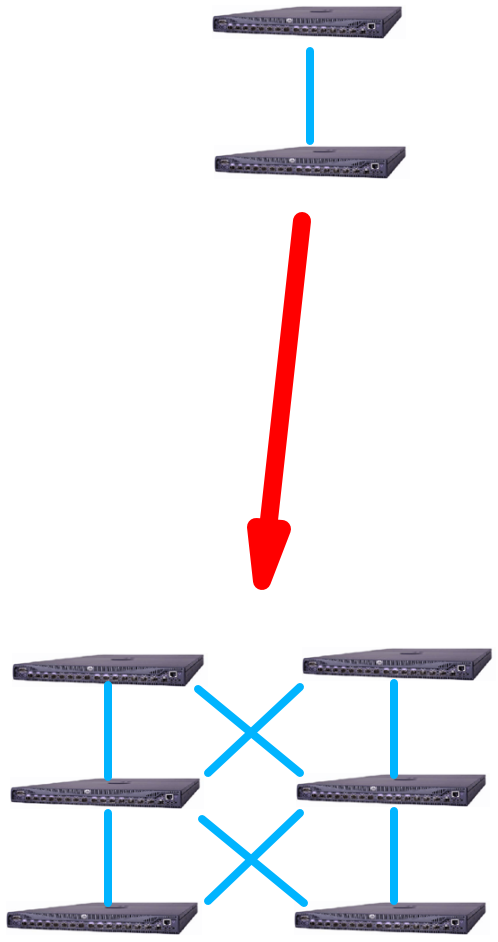
FC distance 10 kilometers, the total distance between switches or managed hubs across dark fibre remains at 10km.

The existing Point-to-point, Arbitrated Loop with Quickloop connection methods remain. But with V5R2 Switched Fabric connection is now supported. This new supoort offers the benefits of redundancy and intelligent routing through the fabric. It is recommended for performance that there are not more than 2 Intersystem links (max 3) between the server and the disk storage device.

With V5R2 there are still 32 LUNs per #2766 Disk adapter. But these LUNs can now be spread across 1-32 ESS. Realistically we can think of a reason for spreading LUNs across multiple ESS.

2Gb SAN

- New support with V5R2
- No support for V5R1
- New connectors
  - 150m - 62.5 micron
  - 300m - 50.0 micron
- Shorter distance
  - 1Gb = 175m and 500m
- Expectations.....



Notes 2Gb SAN

The new 2Gb SAN support offers significant benefits to SAN fabric users. These are:

- High Speed for core SAN backbones
-

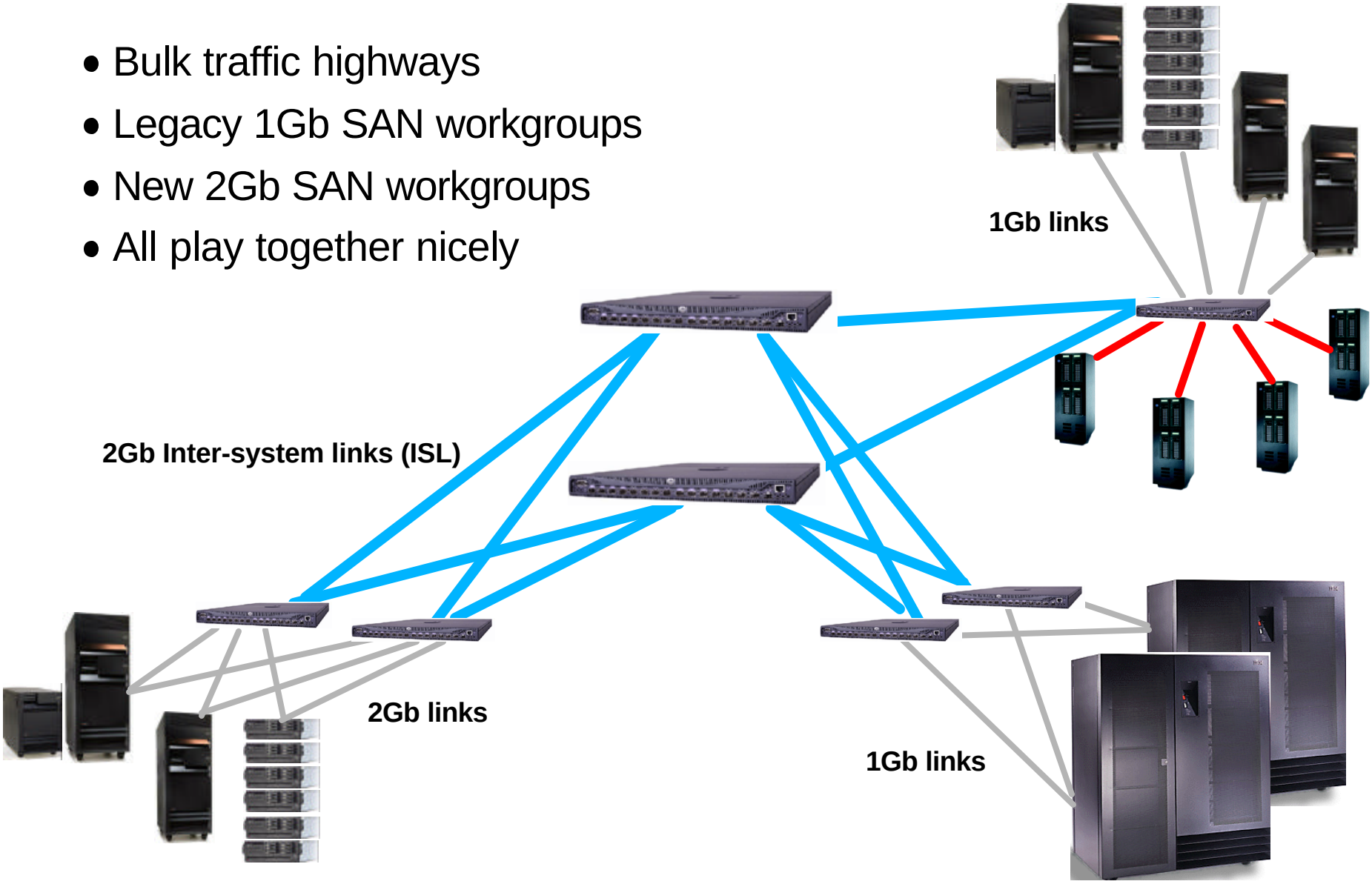
V5R1 does not support the new 2Gb/sec fabric speed. However, and existing adapters will automatically allow 2GB/sec links once V5R2 is installed.

There are two new connector for 2Gb links. These new connector have short cable lengths than their 1Gb counterparts. The dark fibre distance remains unchanged at 10km.

Any customer considering moving to 2Gb/sec fabric must take into consideration performance and the device capabilities within the fabric. Many fabrics will not see any end to end throughput improvements. This is for a numbe of reasons discussed in a later chart.

2Gb SAN Example

- Bulk traffic highways
- Legacy 1Gb SAN workgroups
- New 2Gb SAN workgroups
- All play together nicely



Notes 2Gb SAN example

The chart shows the possibilities of 2Gb SAN. The evolution of SAn is very similar to Token Ring and Ehternet LANs, as speed increases backbones are increased first followed by local workgroups.
In the graphic you can between the two large switches we have a 2Gb trunk service, This reduces the possible bottlenecks throughout the main SAN backbone.
Locally SAN links may remain at 1Gb until the technology adopts the new SAN technology or is replaced with new technology.

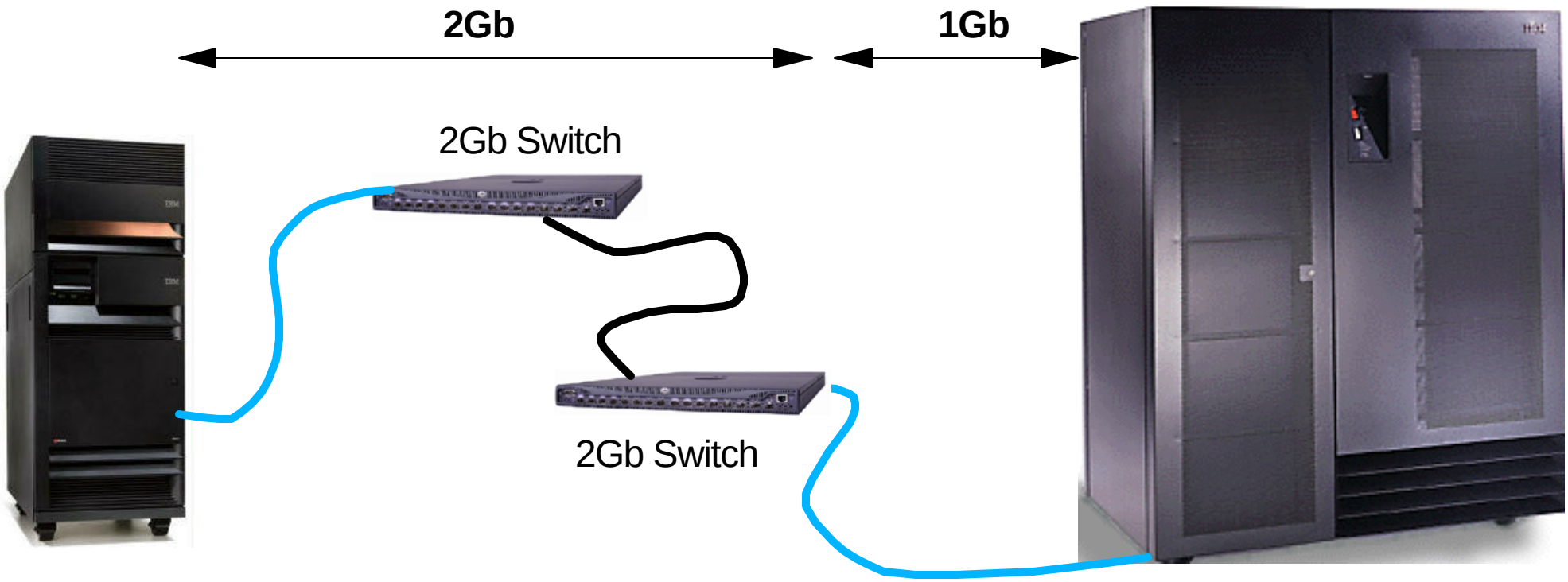
If you look at the route to the disk, there are not currently 2Gb links available from end-to-end. Assuming there is a spread of I/O requests and Host Bay Adaptors on the Storage sub-systemare not swamped the SAN provides a high speed redundant network.

During failures the network can become very slow, which may impact the performance of applications throughout the enterprise.

2Gb performance

Don't look for huge disk performance

- Unless bottlenecks exist



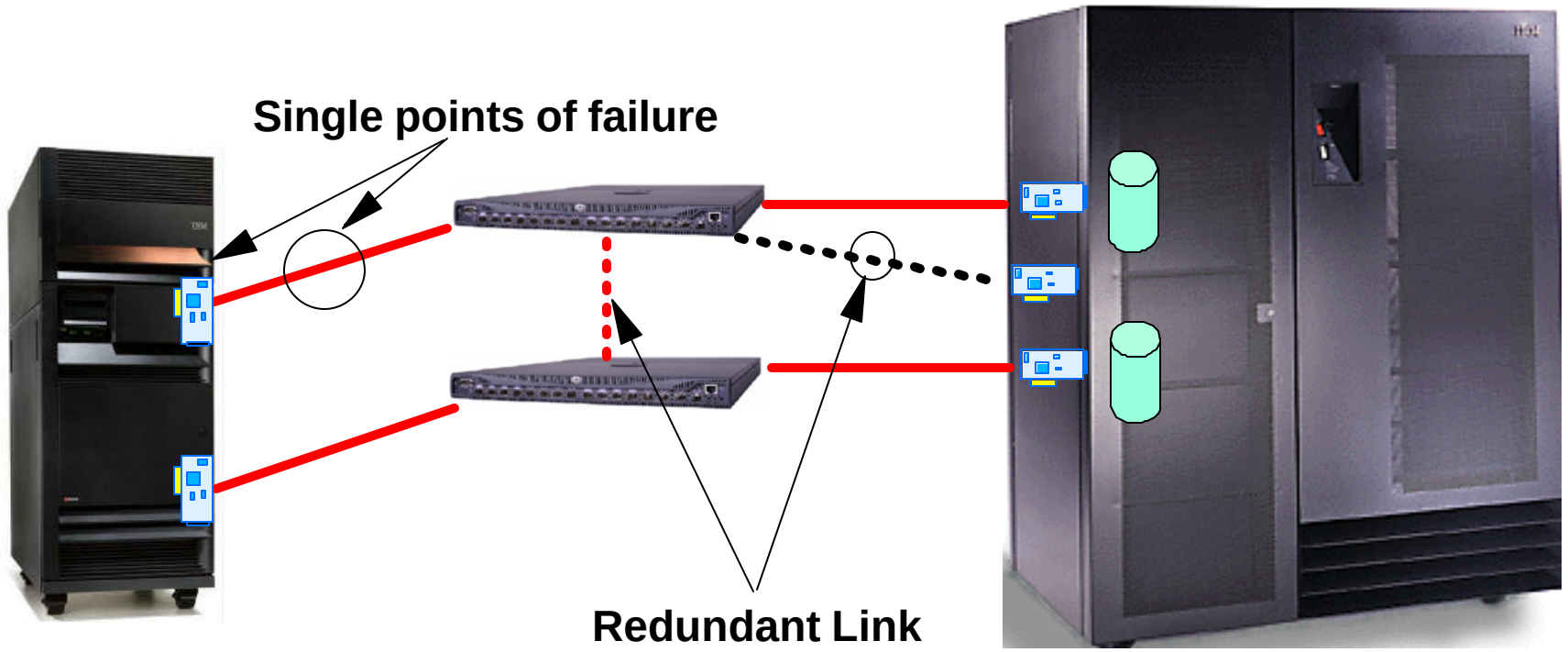
Notes 2Gb SAN disk

When considering 2Gb SAN for ESS attach don't expect huge performance improvements. Currently the ESS only supports 1Gb Host Bay Adapters. There may be benefits to be gained through the network by improving bandwidth.

Link redundancy

No dual path for disk with switch

Other failover alternatives



Notes Link redundancy

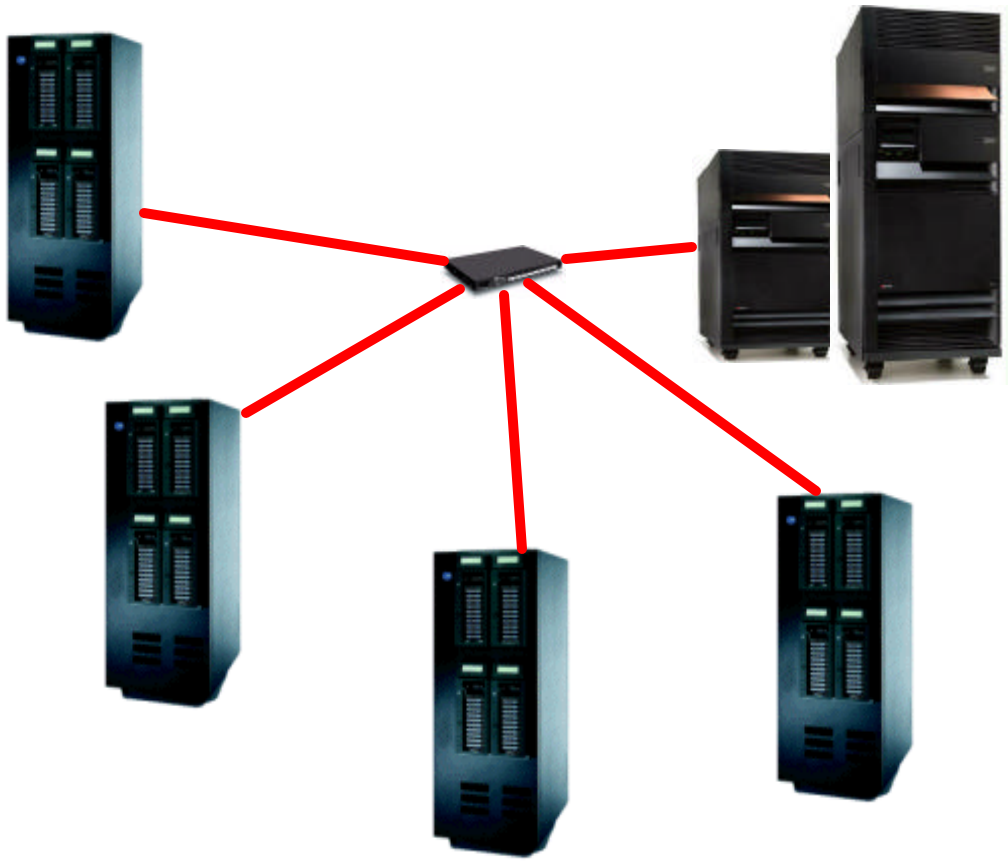
Whilst the iSeries can attach to multiple targets from a tape drive perspective, it cannot implement dual path to a LUN currently. Dual path to a LUN would provide protection from iSeries #2766 failure or failure of the first link segment to the switch or direct to the ESS.

There are other methods to protect oneself from adapter failure, this would be IOP mirroring. This function is enabled as a base function within OS/400. Two #2766s would be installed with their corresponding IOPs. These would then be mirrored to the ESS. This setup would also require the iSeries LUNs to be mirrored in the ESS.

iSeries reliability is extremely high and there are other single points of failure that are probably more vulnerable, e.g. the fabric and switches, even possibly the host bay adapters (during maintenance these can become unavailable). A potential solution to this would be to provide a zoned redundant path from the switch to the ESS as shown in the chart.

Tape and 2Gb/sec - great support

- Can support up to 16 devices
- Most tapes = 2 devices
  - tape
  - loader



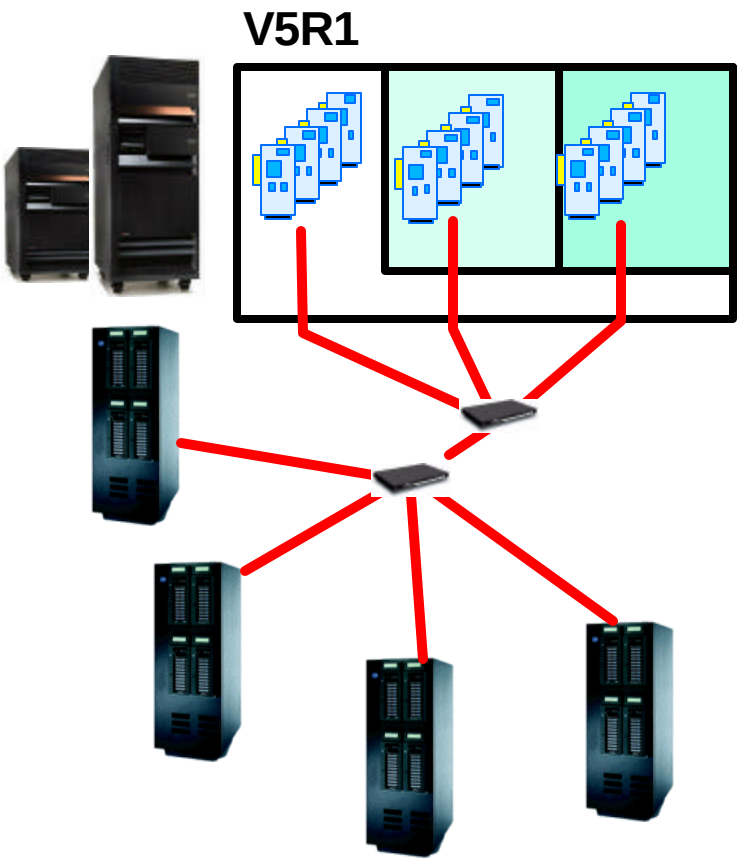
Notes Tape and 2Gb support

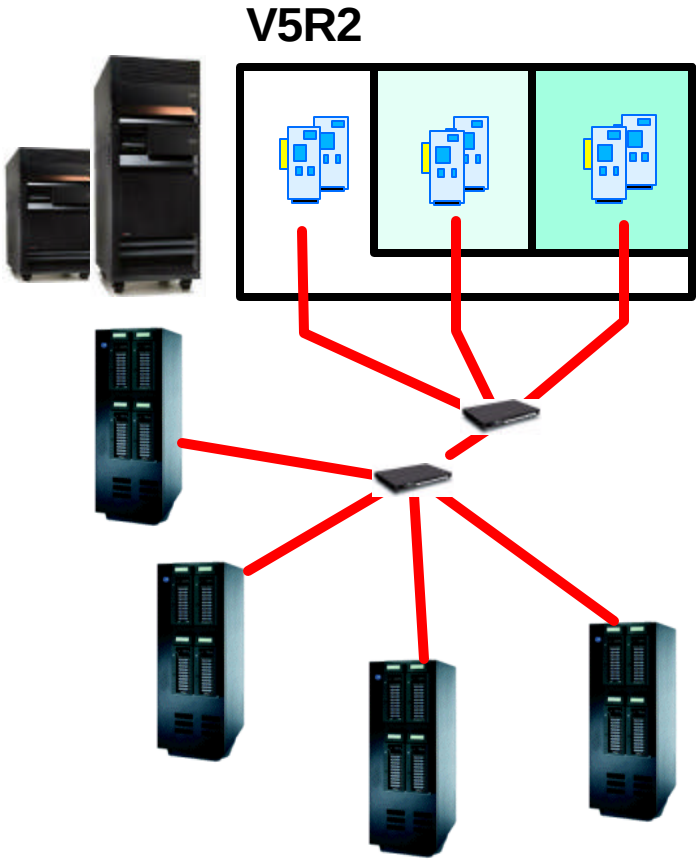
With V5R2 the most significant 2Gb support is with Tape attach #2765. The ability to throw tape drives over large distances and allow the attachment of many sources is a huge benefit to must customer IT shops.

The ability to create a single backup vaulting facility is now very affordable for most iSeries customers. Allowing other servers in the enterprise to also access the large style tape drives. Driving down the tape backup window across the enterprise.

Tape and LPAR

- In V5R1 separate adapter for each drive/each partition
- V5R2, only one adapter per partition for up to 16 devices
- But beware of performance





Notes Tape and LPAR

At V5R1 FC attached tapes offered a great solution for multiple partitions, but it was very expensive. If the LPAR'd system had many partitions and was attached toa number of high speed tape drives. Each partition had to have a #2765 for each tape drive it needed to access. This is a very expensive solution.

With V5R2 and multiple target support the need for a 1:1 ratio of #2766 to tape drive is removed, as up to 16 devices can be support from one adapter. 2-3 Tape drives per adapter should provide the best performance. But as with all things tape this is very dependant on processor, main storage, file size, etc..

Disk Migrate While Active

Concurrent data migration

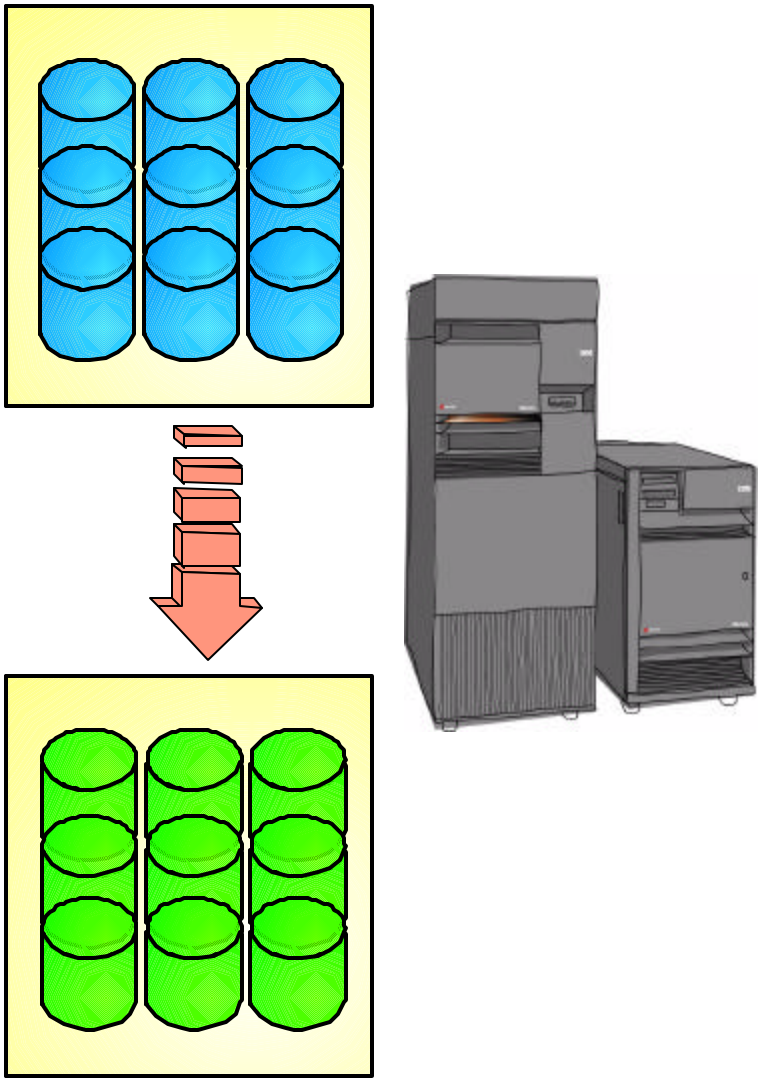
- Enabled through Start ASP Balance (STRASPBAL) command
- Data migrated to new disks during normal production

Significantly reduces downtime

- Restricted system state only required to remove disks from ASP configuration

Enables easier transition to new disk technology

- Plan ahead for migration project
- Install OS/400 V5R2
- Add new disk units
- Mark disk units to be removed
- Activate data migration



Notes Disk Migrate While Active

Disk migrate while active (DWMA) allows for concurrent data migration on a single server or LPAR partition. DMWA is enabled through Start ASP Balance (STRASPBAL) command. This is a standard command within OS/400. Use of this command allows data on disk to be migrated to new disks during normal production.

This significantly reduces server or partition downtime, as the server or partition does not need to be in restricted system state. The only downtime required is to remove disks from ASP configuration, and possibly physically remove the disks from the server.

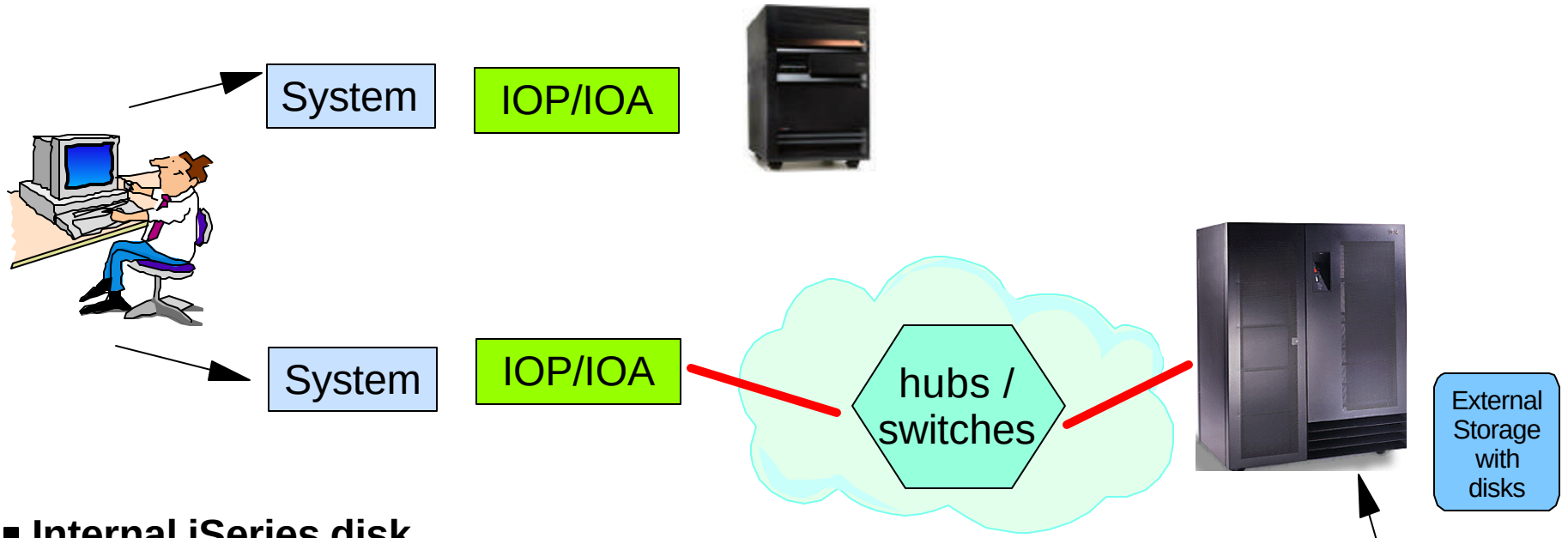
This faciiltity will enables easy transition to new disk technology, by allowing customers to plan ahead for a migration project. The basic steps for DMWA are as follows:

- Install OS/400 V5R2
- Add new disk units
- Mark disk units to be removed. This is done with SETASPBAL command. The marking starts the movement and prevents further allocations from being written to disk. Unless there is a possible ASP overflow condition, where the disk would allow allocations to occur.
- Activate data migration
- Once the data migration is complete, the disks are ready to be removed from the ASP. This should be a relatetively short running task, but depends on number of disks and unmoved data.

Not all data is migrated. There are certain objects that do not move. Journals, temp files, network storage spaces, IFS objects. These objects are moved during the remove from ASP function. Approximately 6% occupied space has remained in our testing. But on some systems this could be much larger. This would then mean that the remove from ASP would take longer.

Positioning iSeries Storage methods

So which way to go.....



- **Internal iSeries disk**
  - Use where lowest iSeries cost, highest performance is desired
  - Usually easier to set up, maintain, easier problem determination
- **External iSeries disk (ESS)**
  - Use where storage consolidation across multiple platforms desired
  - Provides exploitation of SAN infrastructure (requires SAN skills and h/w)
  - IBM has integrated iSeries with ESS Copy Services

Notes Which way to go

Internal iSeries disk

iSeries internal integrated disk solutions is the best solution where lowest iSeries cost, highest performance is desired. This is typical for low-end processors without expansion towers and the largest enterprise class servers.

The disk setup using internal disk is easier to set up, maintain, and problem determine with the Operational Navigator graphical user interface.

External iSeries disk (ESS)

Where storage consolidation across multiple platforms desired the ESS provides a great solution. This is provided the iSeries takes up a relatively small part of the overal disk requirements. This is also a good solution where the customer has already made an investment in an ESS solution.

Consolidated storage provides exploitation of SAN infrastructure, allow great distance to be reached for both disk and tape. But management of the SAN requires significant investment skills and h/w. Typically, iSeries IT shops do not employ database administrators to manage their storage. The iSeries does it itself.

If there is a requirement for distance DR, IBM has integrated iSeries with ESS Copy Services. This provides support for PPRC and Flashcopy.

High Availability or Disaster Recovery

For iSeries High Availability:

- For 24 x 7 availability:
  - Lead with iSeries Clustering and High Availability Business Partner software Solutions (MIMIX, Vision, DataMirror, etc.....)
- For disaster recovery across multiple heteregenous platforms:
  - ESS PPRC or iSeries Clustering
  - Journal enablement is universally required

Storage consolidation for smaller Windows NT/2000

- Use iSeries as "SAN" server for xSeries
  - NT/2000 consolidation through iSeries Integrated xSeries Server/Adapter

Disk Path

CPU, L1 + L2 Cache

Main Storage and Expert Cache

HSL Bus, IOP

CTLR

CTLR -104MB Write cache

All in managed by or in OS

All customer maintained

Hubs

SAN

Switches

80MB/s Bus

80MB/s Bus

80MB/s Bus

External storage

Host Bay

HBA

HBA

Host Bay

SSA Adapters

Processor

Processor

SSA Loop

NVR

Cache

Cache

NVR

10K RPM Disk Units

Notes Disk Path

The chart shows a high level view of the I/O path to disk on iSeries.
Initially the two routes to disk are the same. The split occurs at the IOP level.

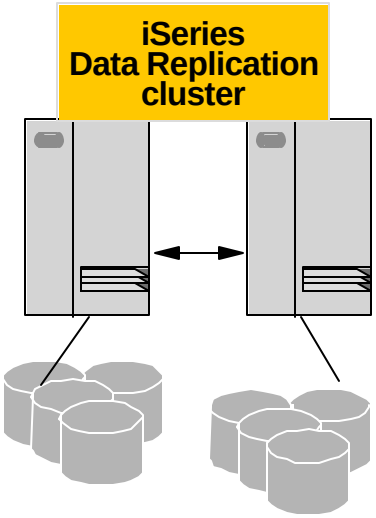
From the IOP the SCS controller or adapter (IOA) connects directly to the disks via one of three buses. The whole of the path to disk is known and managed by iSeries Storage Management and integrated with OS/400 and DB2/400.

The other path to ESS disk shows a more complex route. Once you leave the controller #2766 IOA, you are into the SAN fabric. This is not completely unknown to the iSeries storage management, OS/400 or DB2/400. Flowing through the SAN, cabling and switches you arrive at the ESS host bay adapter which is located in one of the four Host Bays. The IO then heads for one of the two internal cluster processors and its cache. Cache has no prioritization, IOs are handled on a first come first servered basis. Any writes are handed over to the Non-volitile memory of the other cluster processor. At this point the write is released to iSeries storage management as it is considered written. The IO is then 'destaged' to disk based on ESS internal algorithms.

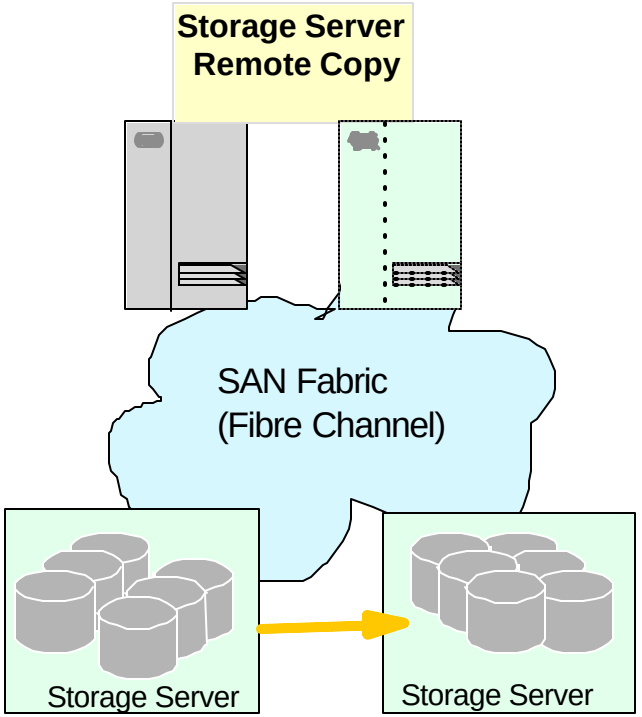
This second route has many components and is very complex. Significant investment is skills is required to manage this environment. Depending on the design and without many parts of the SAN and consolidated storage the IOs cannot get through. If this happens the servers are down.

HA or DR - you get what you pay for.....

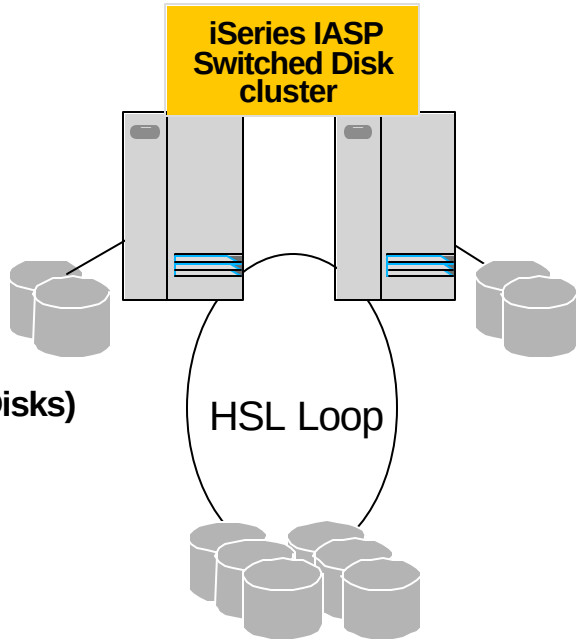
Availability Management Approaches



Replication Services (HABPs)
- Data center & Disaster center solution



Storage Server Copy Services
- Disaster center solution



Independent ASPs (Switchable Disks)
- Data center solution

Notes Availability management

Replication Services (HABPs)

- **Data center & Disaster center solution**
- Provides distance for disaster back-up
- Two active copies of data
- Scheduled and Unscheduled Downtime
- Can Eliminate Save Window
- Workload distribution

Independent ASPs (Switchable Disks)

- **Data center solution**
- Switching for OS/Hardware maintenance
- Switching for unscheduled outage
- Homogeneous distributed application consolidation
- Does not eliminate save window
- Workload switching

Storage Server Copy Services

- **Disaster center solution**
- Independent from OS/System
- Device level mirroring & recovery processes
- One active copy of data (second copy dormant)
- Does not eliminate save window
- 10KM distance (performance impacts unknown)
- Fits for certain DR scenarios

New stuff from IBM Storage Group

PPRC - Extended Distance or PPRC/XD

- Distances over 103km - Global.....
- 'Fuzzy' copy
- No iSeries experience yet
- Available from Storage now with Copy Services

Notes PPRC-XD

PPRC-XD is a new component of Copy Services available since mid-2003. This allows for PPRC services to operate over continential or global distances. Whilst this function is available for iSeries external disk, information on testing and operation is very limited at this time. PPRC-XD is meant for data migration not availability.

PPRC-XD works by maintaining a 'fuzzy copy' of data volumes. This is an asyncronous copy. With an async copy, the order of the writes is not guarenteed. Therefore, the data on the remote disk will not maintain its integrity. It is doubtful that OS/400 would IPL from a fuzzy copy. Other open systems would be able to mount 'some' volumes, but not all.

PPRC-XD is syncronized regularly and it is the sync copy that is used for recovery. This sync copy is roughly equivalent to a flashcopy. It is recommended that the sync copy is taken regularly and a lull in transactions. At present with our limited knowledge of PPRC-XD we feel this is not a practical implementation for OS/400.

iSeries availability - built-in

Faster restart of secondary system during failover or switchover

- Flexibility to start journaling in stand-by mode
- Journaling is pre-started on target system to for rapid recovery and improved performance

Configuration flexibility for enabling self-managing access path protection

- SMAPP default value for protection lowered to 70 minutes
  - V5R1 was 90 minutes, V4R5 was 120 minutes
- Selective enablement of access paths

Save-While-Active

- enhanced to allow save new member without waiting for unrelated commit cycle

Notes iSeries availability - built-in

Faster restart of secondary system during failover or switchover

- Flexibility to start journaling in stand-by mode. Journaling is pre-started on target system to for rapid recovery and improved performance

Configuration flexibility for enabling self-managing access path protection

- System Managed Access Path Protection (SMAPP) default value for protection has been lowered to 70 minutes at V5R1 this value was 90 minutes, and at V4R5 it was 120 minutes. This is a built-in function of OS/400 that speeds database recovery after a failure or disaster.
- Selective enablement of access paths

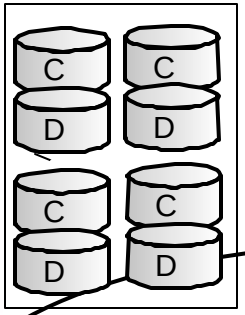
Save-While-Active

- Similar to Copy Services PPRC. Produces a T0 save for system libraries and objects.
- Can be used instead of Flashcopy and has more functions
- Can perform incremental saves.
- Enhanced to allow save new member without waiting for unrelated commit cycle
- Parallel save capabilities for decreaseing the save window of large databases

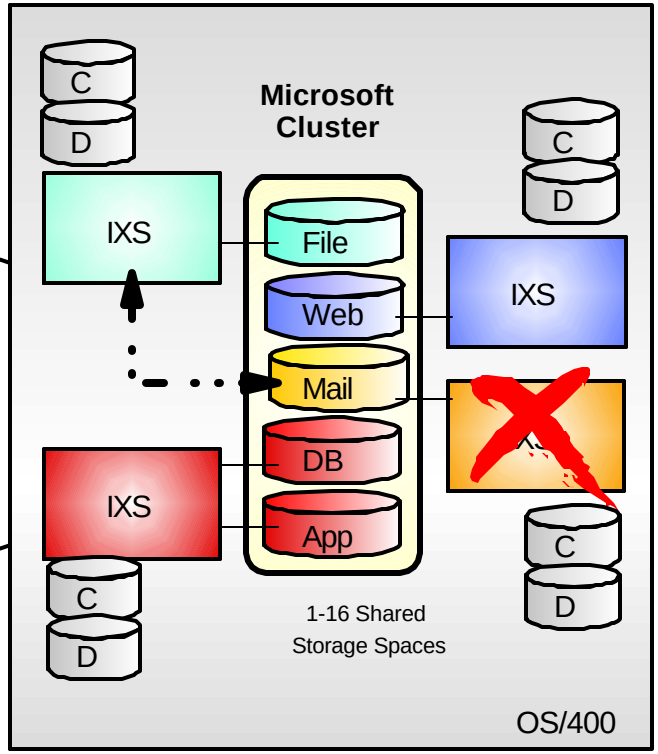
IASP Cluster & Microsoft Cluster Server

Production Server

Network storage spaces



IXS adapters



Notes iSeries and MS Clusters

With V5r2 support OS/400 Windows integration has been enhanced to include support for Microsoft Cluster Services (MSCS). This new support enables iSeries customers to take advantage of the OS/499 Windows intergration code benefits and now include MS clusters and applications that demand MSCS to become highly available.

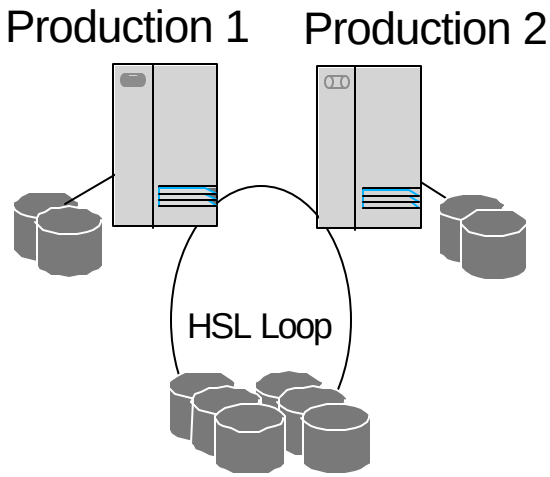
iSeries offers a SAN like storage consolidation structure for Windows (and Linux) Servers. The windows storage is in Network Storage spaces in the IFS on the iSeries.

In the example on the chart we have 4 IXS adapters running various PC based applications. Each of the IXS/Windows servers is clustered to another, providing failover cabability.

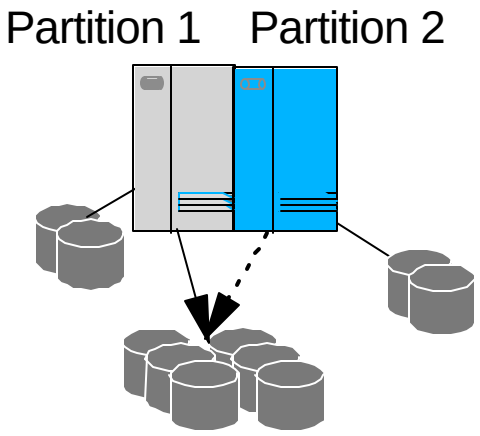
iSeries Windows Integration supports MS Window 2000 Server and Advanced Server.

Independent ASPs

IASP Operational Requirements



Dual System Configuration



LPAR Configuration

Basic Requirements

- Dual servers or LPARs
- V5R1 (IFS), V5R2 (DB)
- Chargeable feature of OS400, (Option 41)
- HSL Switchable Towers (Dual Systems)
- Journaling Active

Planning Considerations

- Some system objects not supported
  - Spool Files
  - User Profiles
  - Job Queues & Output Queues
  - Other system objects*

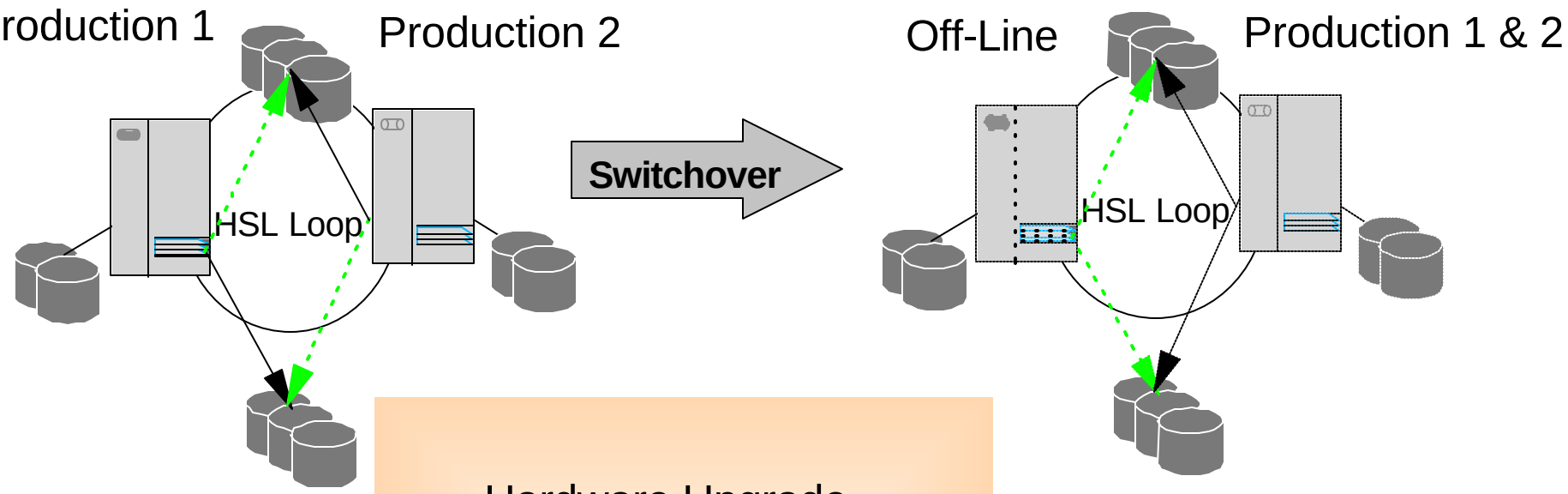
Benefits

- No system to system in flight data loss
- No lag time (primary/backup)
- Simple switchover operation for maintenance

Go to Info Center look under System Management (Clustering)
http://publib.boulder.ibm.com/html/as400/infocenter.html
Redbook publications **SG24-6802 & SG24-5194**

Notes

IASP Switchover Environments



Hardware Upgrade
OS Upgrade
PTFs
Workload Switch
Unscheduled Outage

Notes

Storage Migration

#6501 Migration at V5R2

Use Disk Migrate while Active

- Move to FC Lun or internal PCI drive
- System remains active
- Run STRASPBAL between LUNs

Use StorWatch Specialist

- Upgrade to 8xx with #2766
- Power down system
- Point and click with StorWatch

MUST remove #6501 before upgrade to i890

- No SPD support
- No migration tower support

Notes #6501 migration V5R2

Use Disk Migrate while Active
- Move to FC Lun or internal PCI drive
- System remains active
- Run STRASPBAL between LUNs

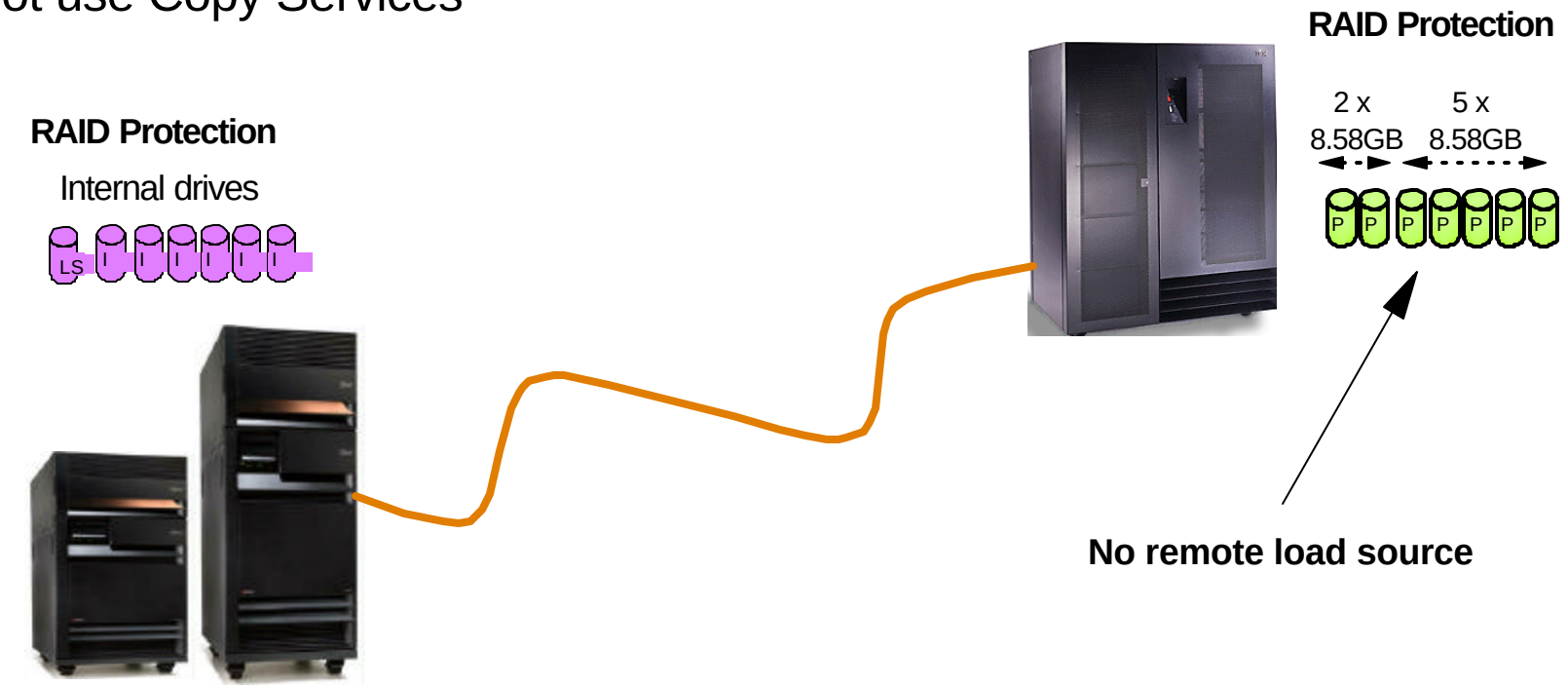
Use StorWatch Specialist
- Upgrade to 8xx with #2766
- Power down system
- Point and click with StorWatch

MUST remove #6501 before upgrade to i890
- No SPD support
- No migration tower support

iSeries SAN External Disk Configurations

iSeries with combination of disks

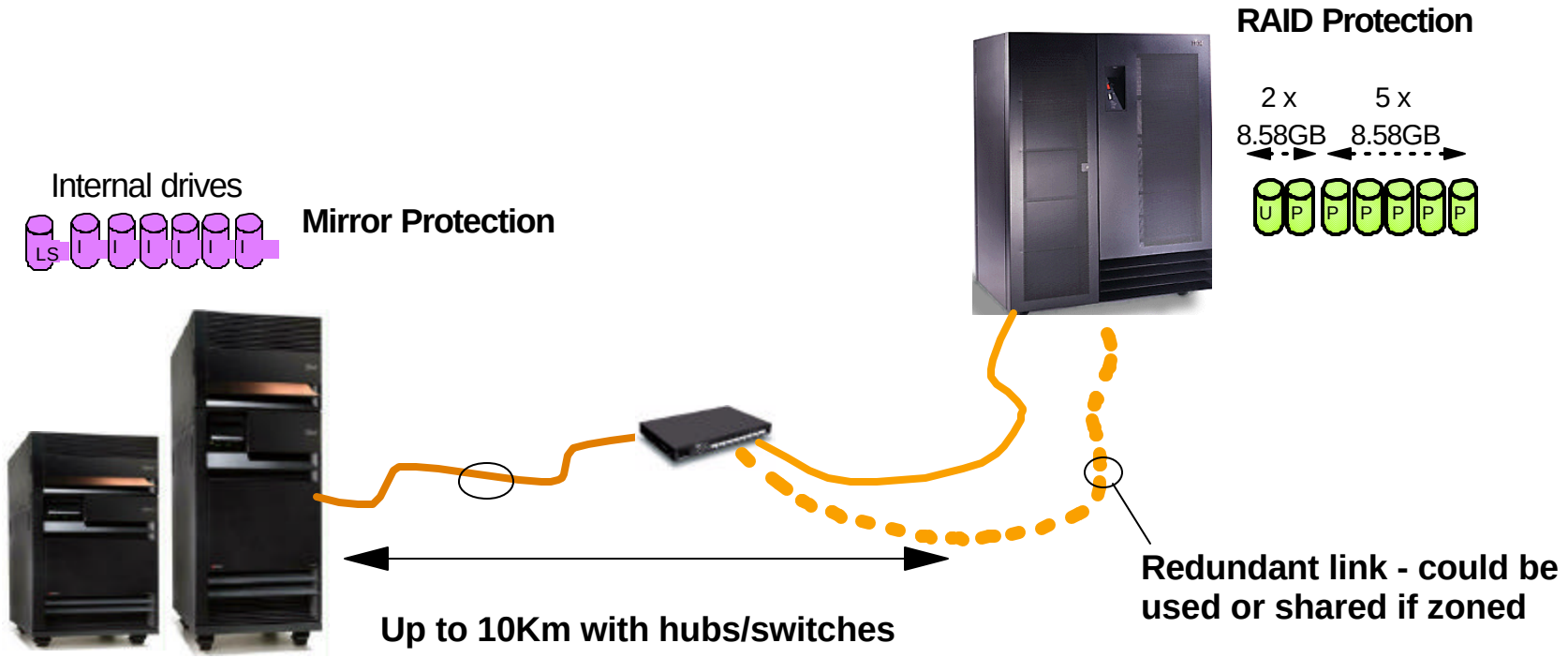
- Internal drives + ESS drives
- No RLS in ESS
- Limited concurrent maintenance
- Similar protection to internal RAID disks
- Cannot use Copy Services



Notes

iSeries with redundant path

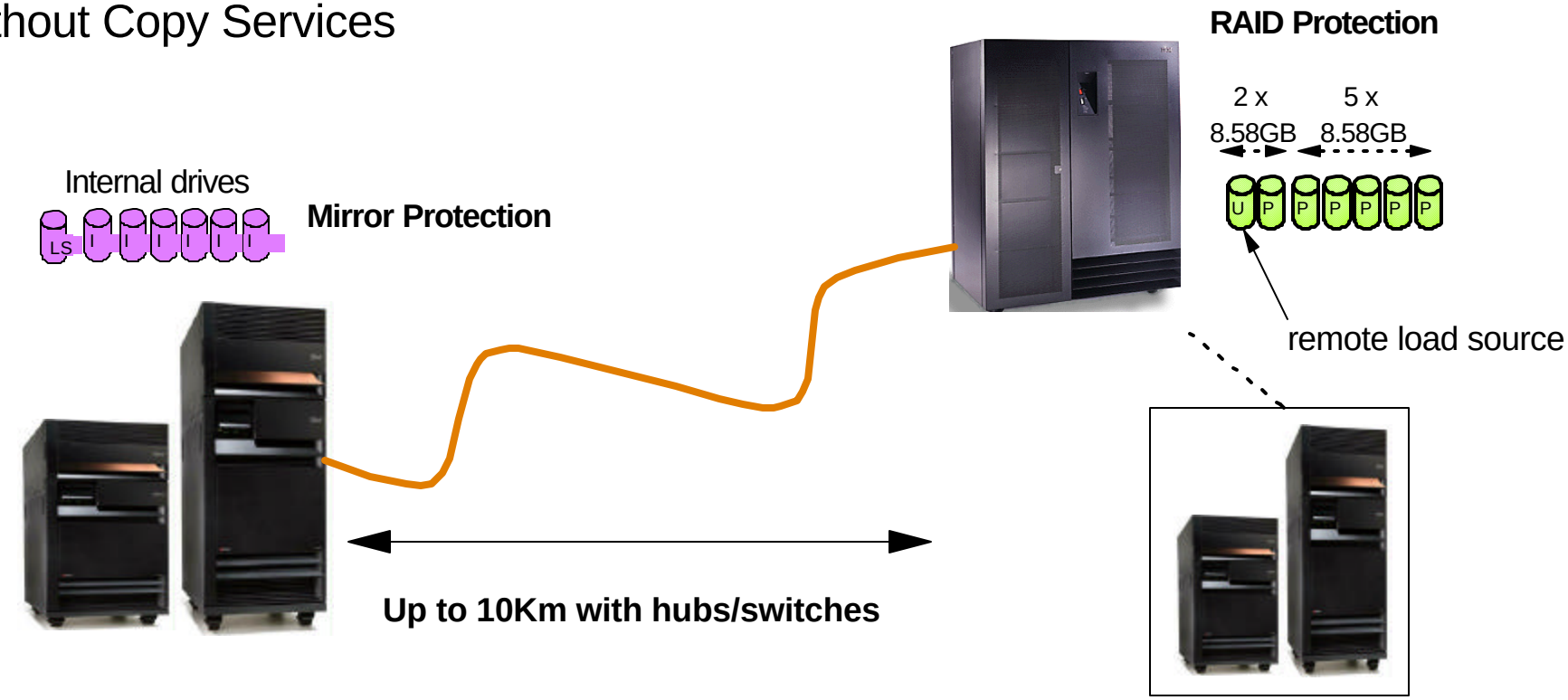
- Second path from Hub
- Must use separate host bays
- Will failover if host bay lost



Notes

iSeries mirrored to ESS

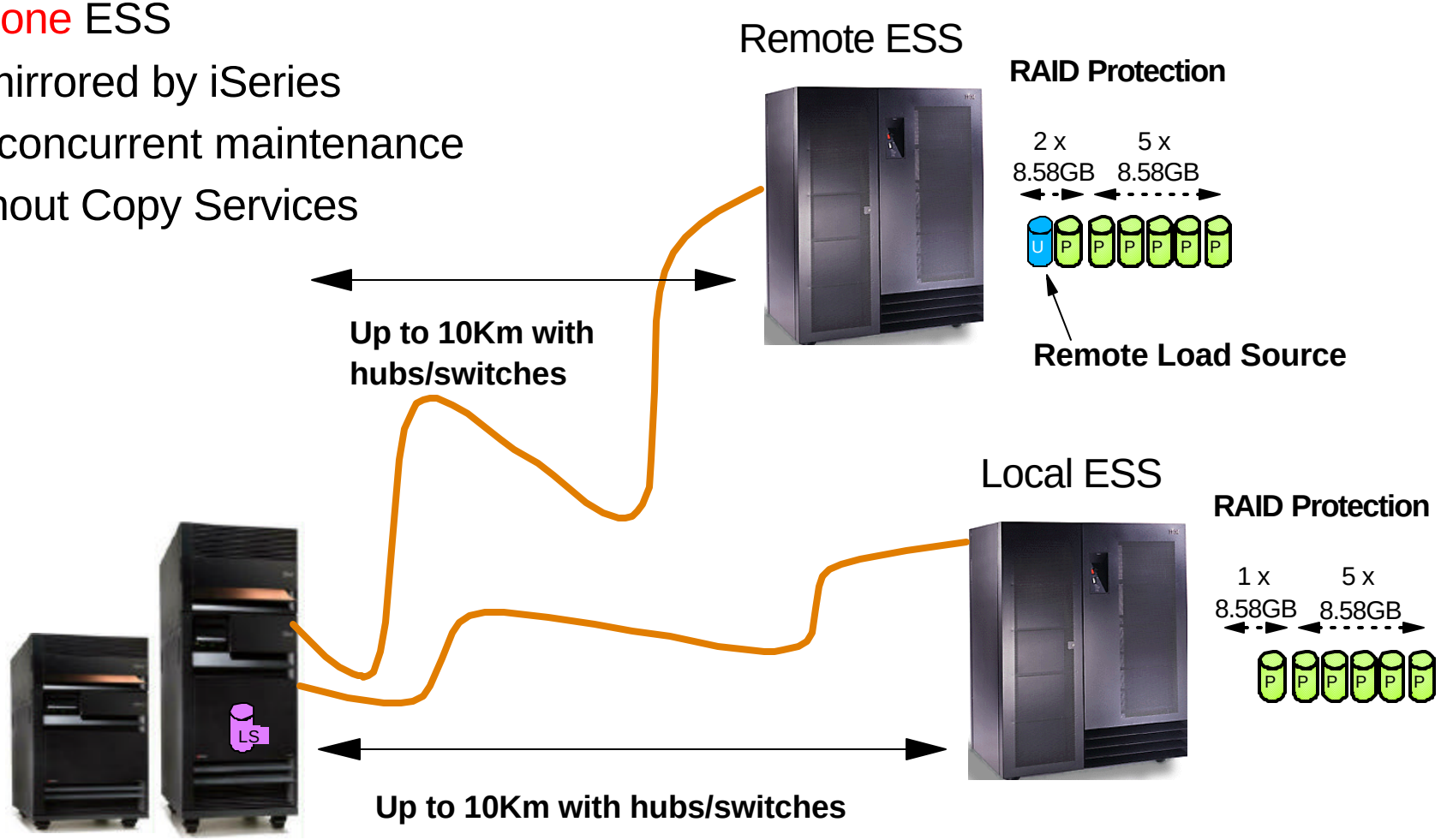
- All internal drives mirrored to ESS drives
- RLS in ESS
- All internal drives mirrored to ESS
- Limited concurrent maintenance
- ESS can be local or distant
- DR without Copy Services



Notes

iSeries to mirrored ESS

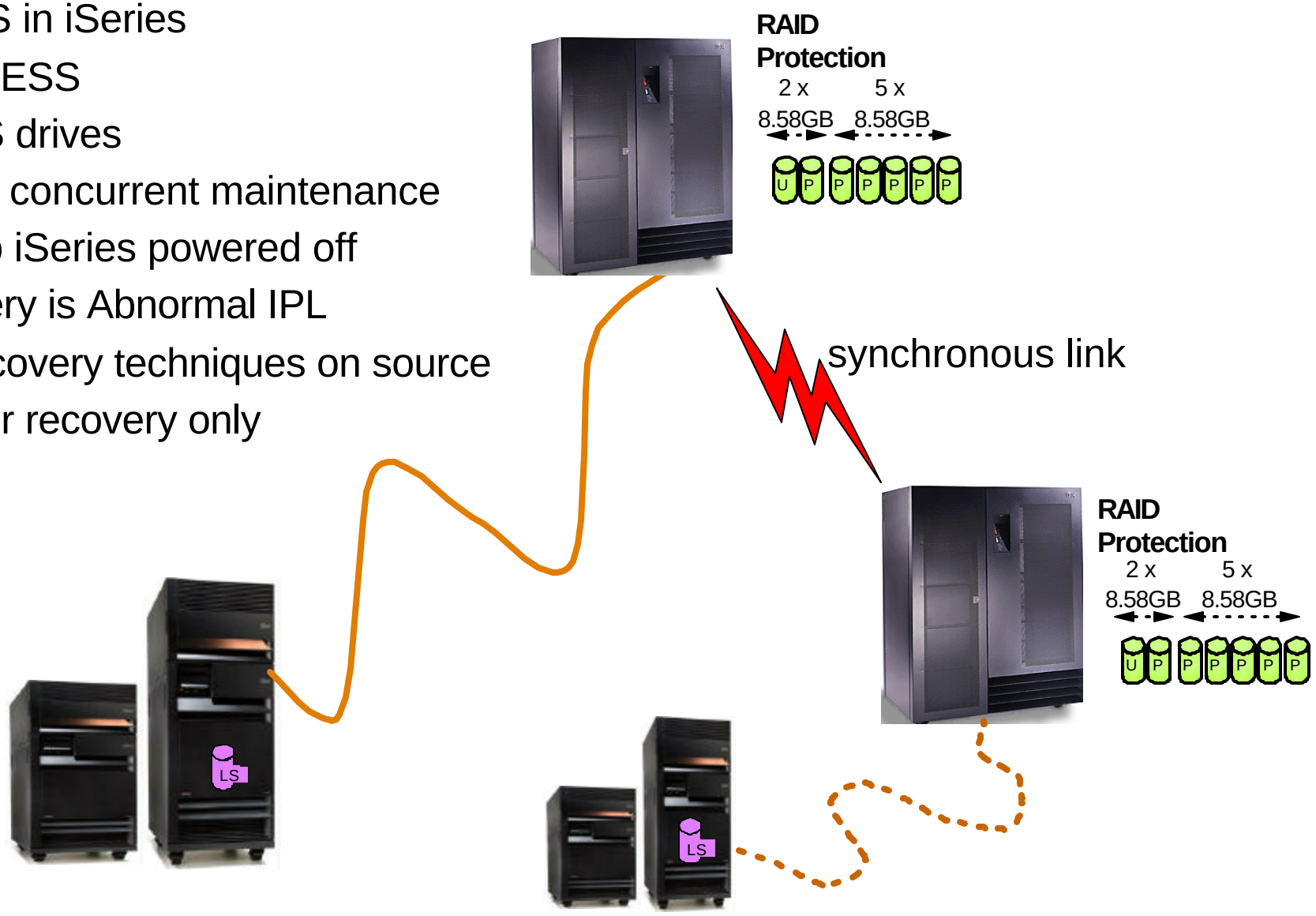
- All ESS drives
- Only LS in iSeries
- RLS in one ESS
- LUNs mirrored by iSeries
- Allows concurrent maintenance
- DR without Copy Services



Notes

iSeries to PPRC'd ESS

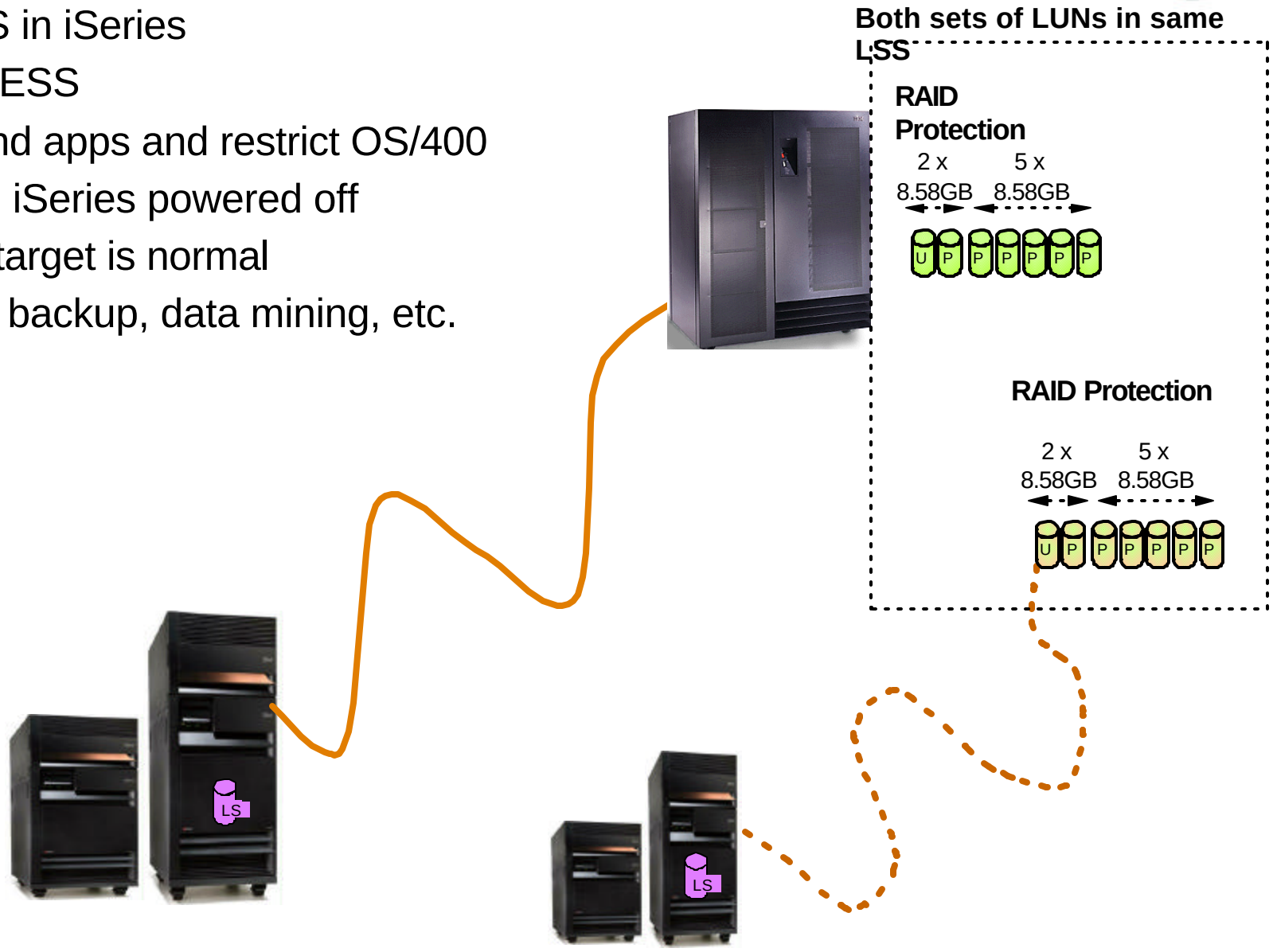
- Only LS in iSeries
- RLS in ESS
- All ESS drives
- Limited concurrent maintenance
- Backup iSeries powered off
- Recovery is Abnormal IPL
- Use recovery techniques on source
- Disaster recovery only



Notes

iSeries to ESS with FlashCopy

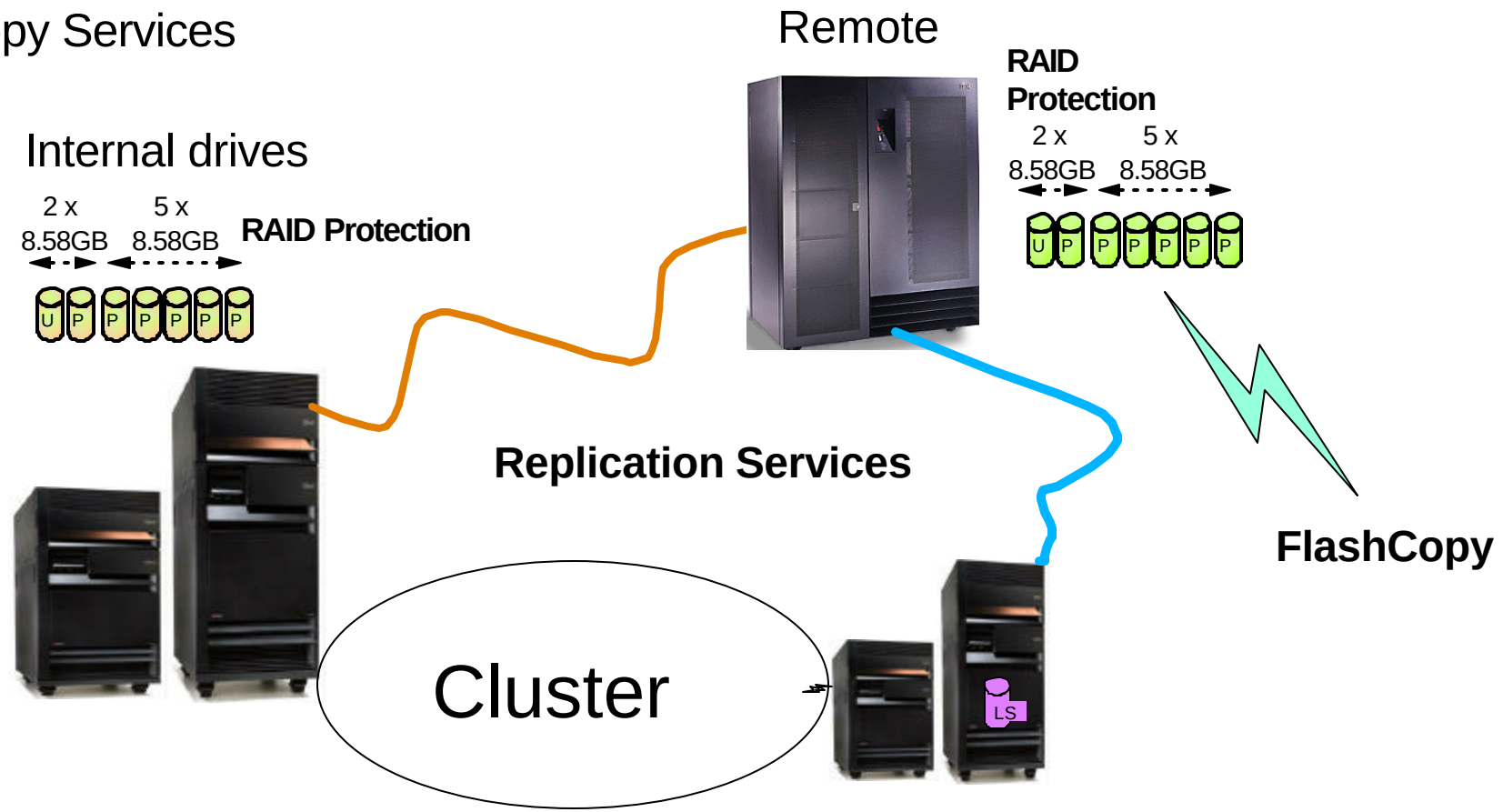
- Only LS in iSeries
- RLS in ESS
- Must end apps and restrict OS/400
- Backup iSeries powered off
- IPL on target is normal
- Off-line backup, data mining, etc.



Notes

The Best of Both Worlds

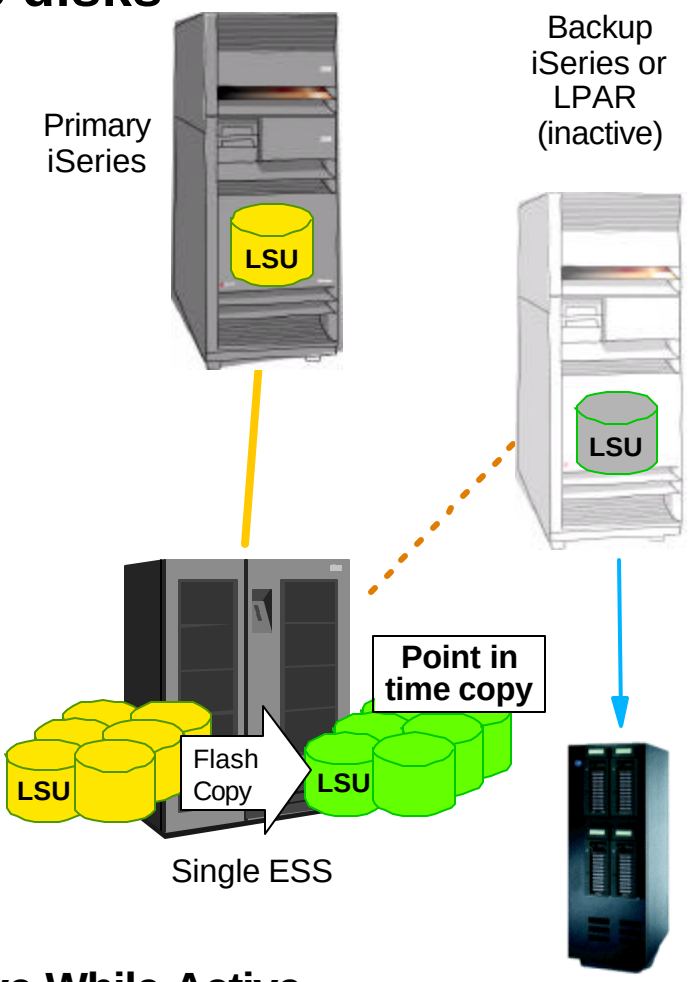
- Local is 'normal' iSeries with internal disk
- Remote is external disk iSeries
- Only LS in remote iSeries
- Replication svs still used
- No Copy Services



Notes

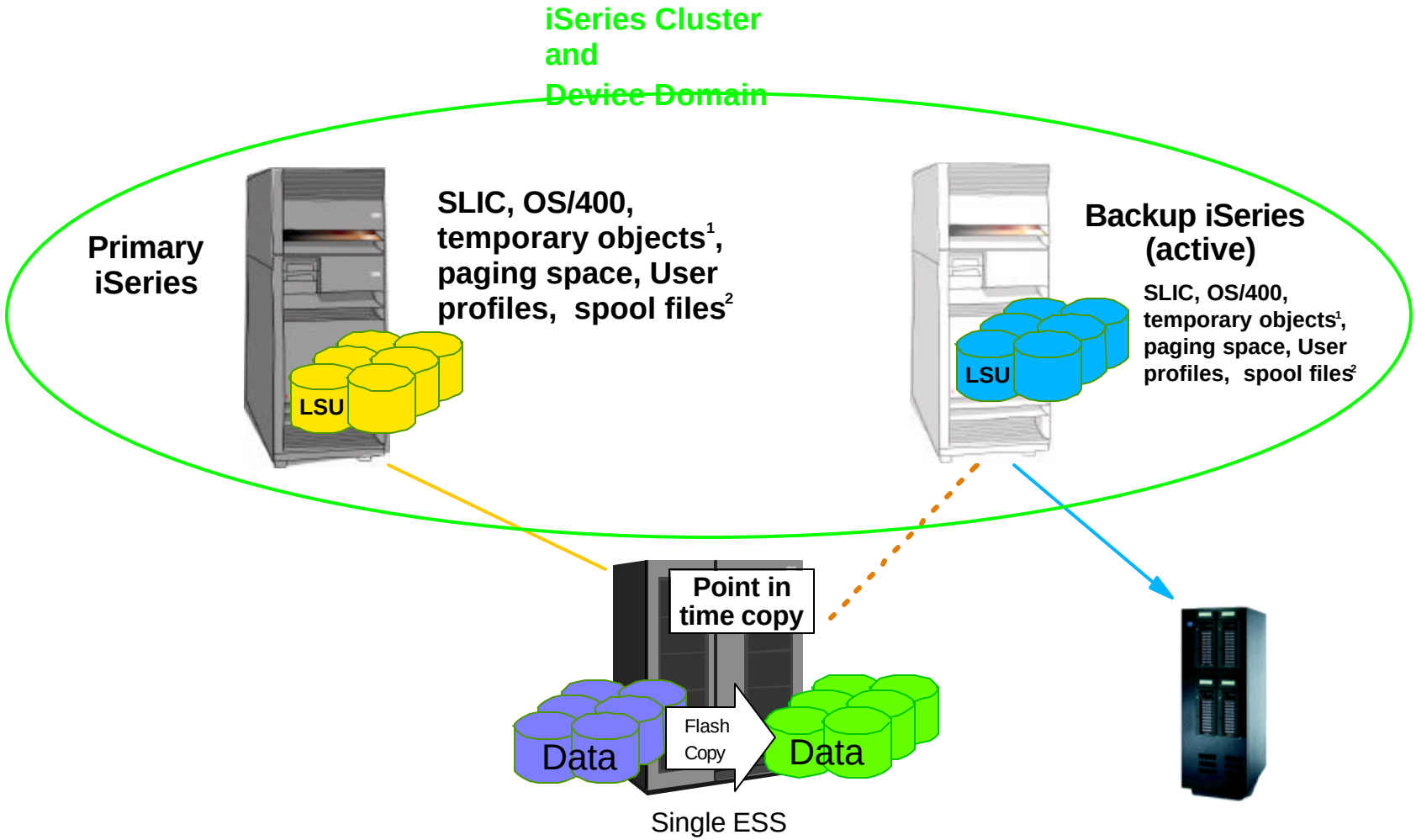
Current Flash Copy

- **Local, instant point-in-time copy of all iSeries disks**
  - **Addresses outages due to backups**
  - **Invoked from IBM TotalStorage ESS Specialist**
  - **'NOCOPY' option**
- **Benefits:**
  - **Fast checkpoint**
  - **Full system backups with minimal downtime**
  - **Multiple platforms supported**
- **Considerations:**
  - **Must quiesce (power off) primary server**
    - **Up to 2-hour outage or more on primary server**
  - **New complexities**
  - **Manual processes - no CLI**
  - **Recovery considerations**
  - **May impact disk performance**
  - **For nightly backups, may not be faster than OS/400 Save-While-Active**



Notes Current Flashcopy

IASP Flash Copy



1) Temporary space includes Qtemp libraries, index build space etc
2) There is a statement of direction to allow Spool files in an IASP in the future

Notes iASP Flashcopy

IASP Flash Copy

- Benefits over Remote Load Source process
  - No IPL to attach second copy of data to the backup server (An IPL must happen at the end of the process)
  - Saving can start after around 5 - 10 minutes after the flash
  - No issues with SLIC or OS/400. The second system has it's own copy
  - Saves will be done through OS/400. This will take advantage of the performance of our save techniques including parallel saving.
  - Replication on the production server (with HABP) to another system can continue to run without getting behind while the save is running
  - Multiple IASP's can be attached and saved in a single process. This is a huge benefit compared to current remote load source mirroring technique.
- Considerations:
  - Must quiesce (vary off IASP)
    - Around 10 minute outage on primary server
  - New complexities - customer must be in an IASP environment to use it
  - Manual processes to Flash- no CLI
  - Recovery considerations
  - Disk sizing for system ASP is very important - this requires the fastest disk on the system as this is where memory paging, index builds etc happen.
  - IPL the backup system after the save has taken place

Notes

iSeries SAN attached Tape

#2765 FC Tape adapter

- Announced with V5R1
- Only for M270 and M8xx
- Private Loop Direct Attach
- Higher performance
- Multi-host
- 10km distance
- Single device attach

- Announced with V5R2
- Fabric support (Switch)
- N-N connection
- 16 devices
- 100MB/sec

Notes #2765 FC Tape

SAN or Fibre Channel attachment support was first enabled with iSeries release V5R1. A new adapter, FC 2765, provides the Fibre Channel attachment. This IOA is only supported on the newer PCI based system hardware used in the model 270 and 8xx system released with V4R5 and later. Fibre Channel tape attachment provides the iSeries user with; higher performance, multiple host attachment, remote attachment (10 KM).

In V5R1, the iSeries is limited to Private Loop Direct Attach support of simple SAN connections. The V5R1 configurations are limited to single device and multiple iSeries. In V5R1 multiple device connection are made with multiple Arbitrated Loops built with multiple FC 2765 adapter, one per loop, per drive.

In V5R2, the iSeries has added the support for Fabric attachment. The V5R2 configurations can be assembled to provide essentially N to N connections between systems and tape devices. The V5R2 limitations are a maximum number of 16 devices and a performance bandwidth of 100 MB/s per adapter.

#2765 attachment

Fabric Devices

- 3534 Managed Hub
  - 2 hubs can be partnered for increased port count.
- 2109 Switch
  - V5R1 Quick Loop support only.
  - V5R2 Switch fabric support.
- 2103 Hub is not supported.
- 2108 SAN Data Gateway is not supported

Fibre Channel Tapes

- 3590 models Exx and Hxx including drives for the 3494
- 3584 LTO Library
- 3583 LTO Library

Fibre Cables... three types

- 9um single mode used for switch to switch up to 10 KM.
- 50um multi mode used for host/device to switch or switch to switch.
- 62.5um multi mode used for host/device to switch or switch to switch.

Notes

Supported AS/400 Releases

PCI systems (270 and 8xx) V5R1 and later.

- V5R1 supports Arbitrated Loop only
- V5R2 adds support for Fabric
- Point to Point connections are not supported for tape.

Maximum performance...

- 100 MB/s per 2765 with 830 and 840.
- 80 Mb/s per 2765 with 270 and 820.

Maximum devices...

- 2-IOA's per IOP
- 16 devices per IOA
  - A device is a tape drive or media changer
  - V5R1 > one target 2 to 16 LUNs
  - V5R2 > multiple targets and LUNs

Notes

V5R1 Configurations

Single iSeries - Single tape with FCAL

iSeries

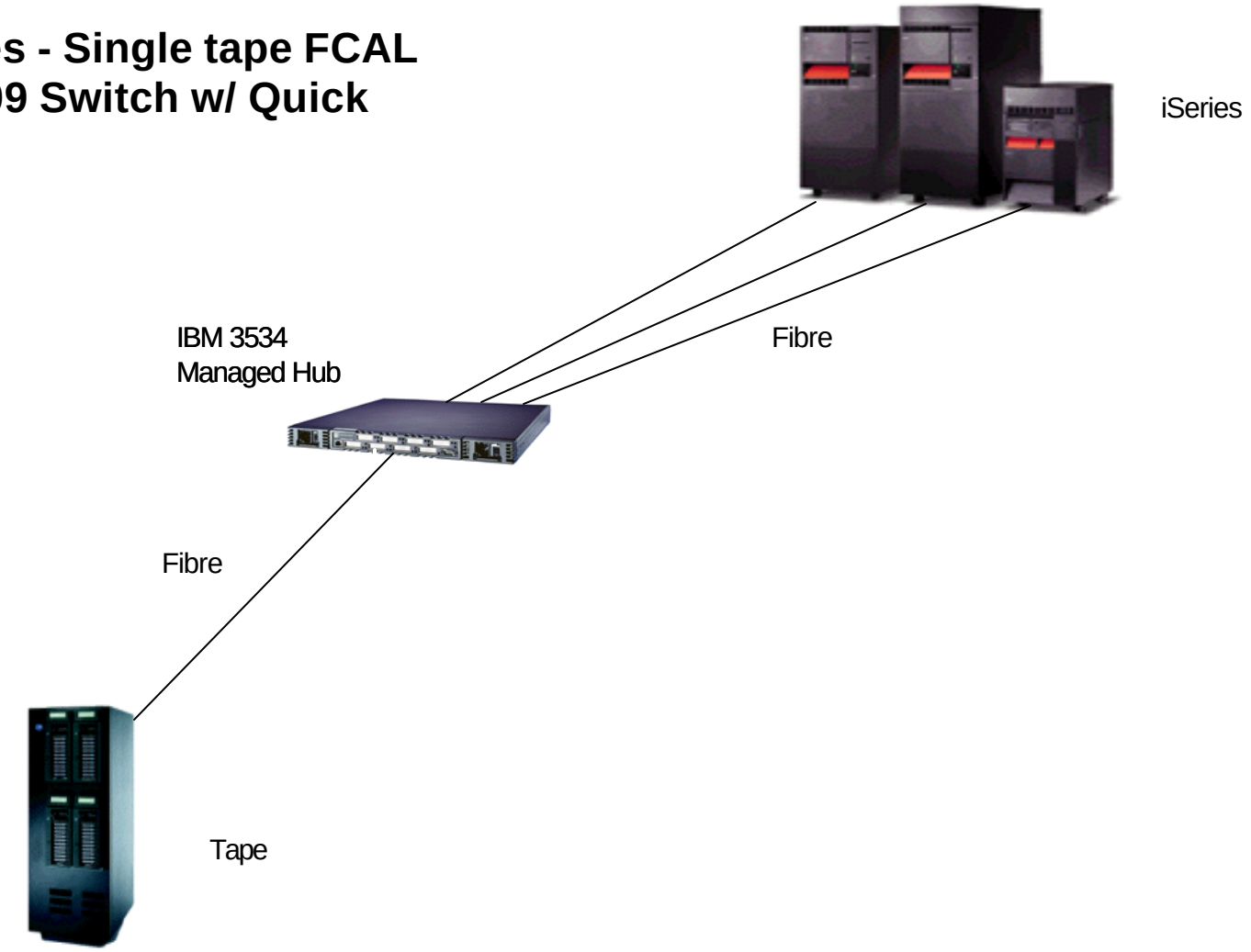
Fibre

Tape

Notes

V5R1 Configurations

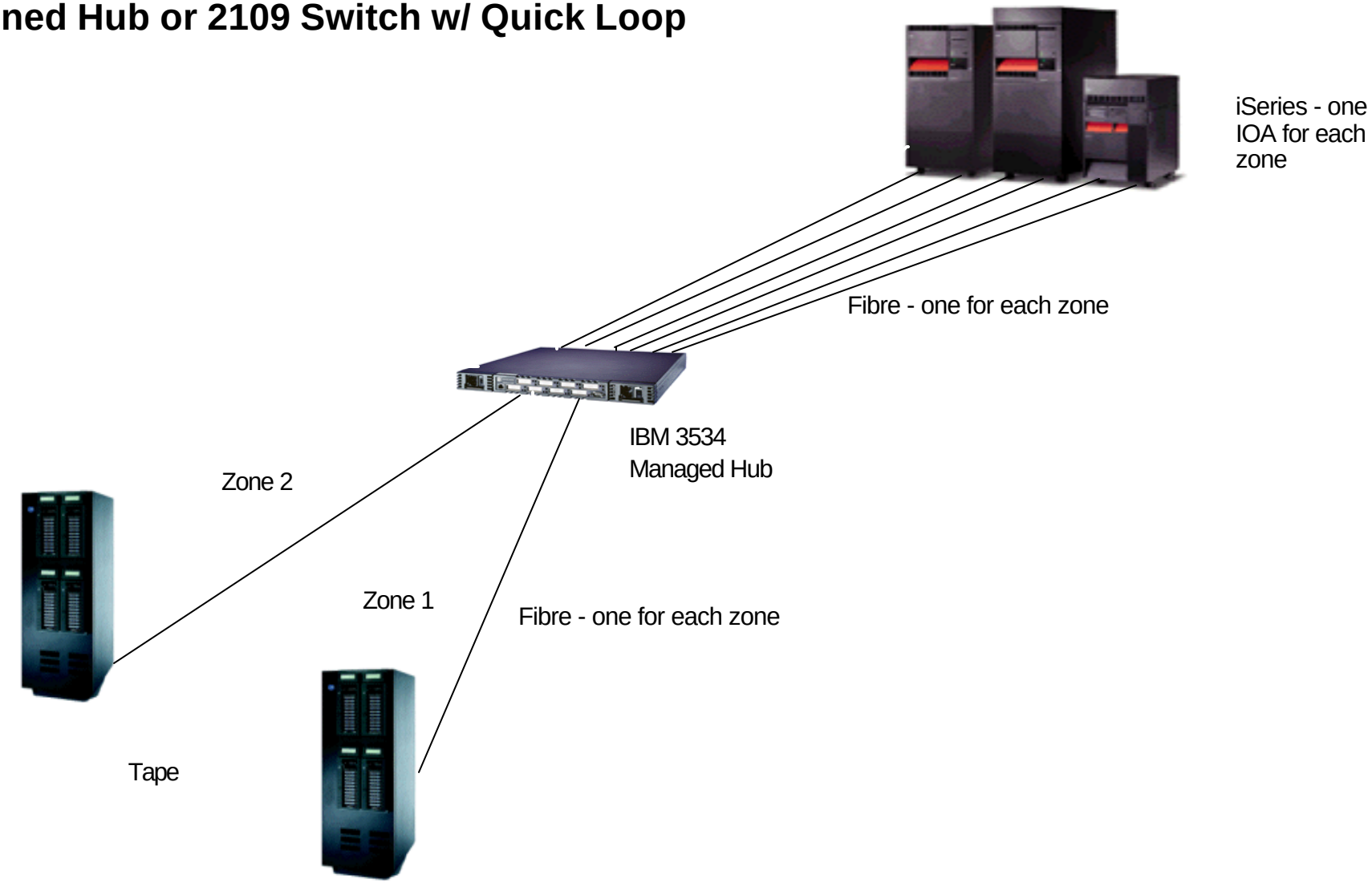
Multiple iSeries - Single tape FCAL via Hub or 2109 Switch w/ Quick Loop



Notes

V5R1 Configurations

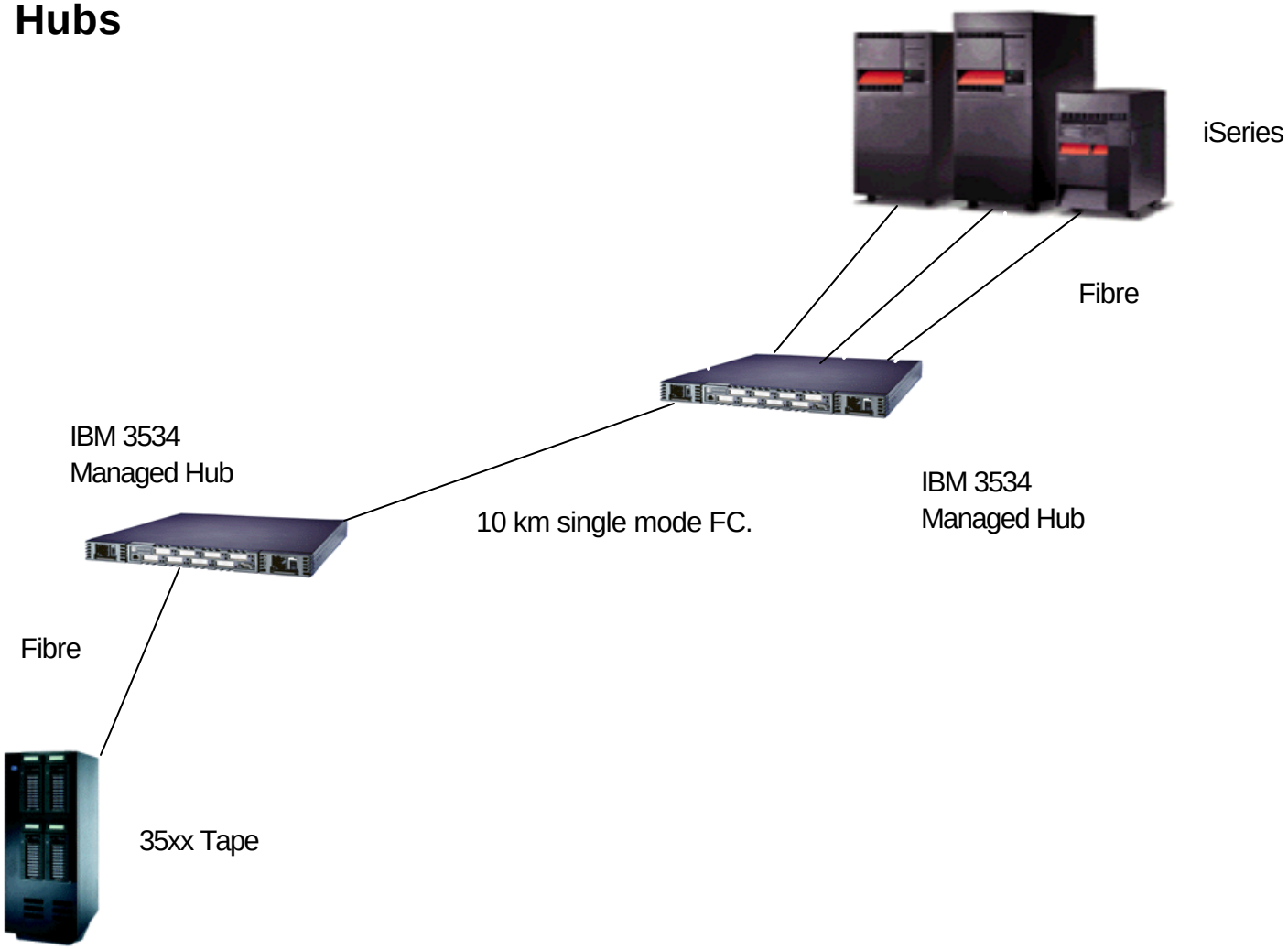
Multiple iSeries - Multiple FCAL, via Zoned Hub or 2109 Switch w/ Quick Loop



Notes

V5R1 Configurations

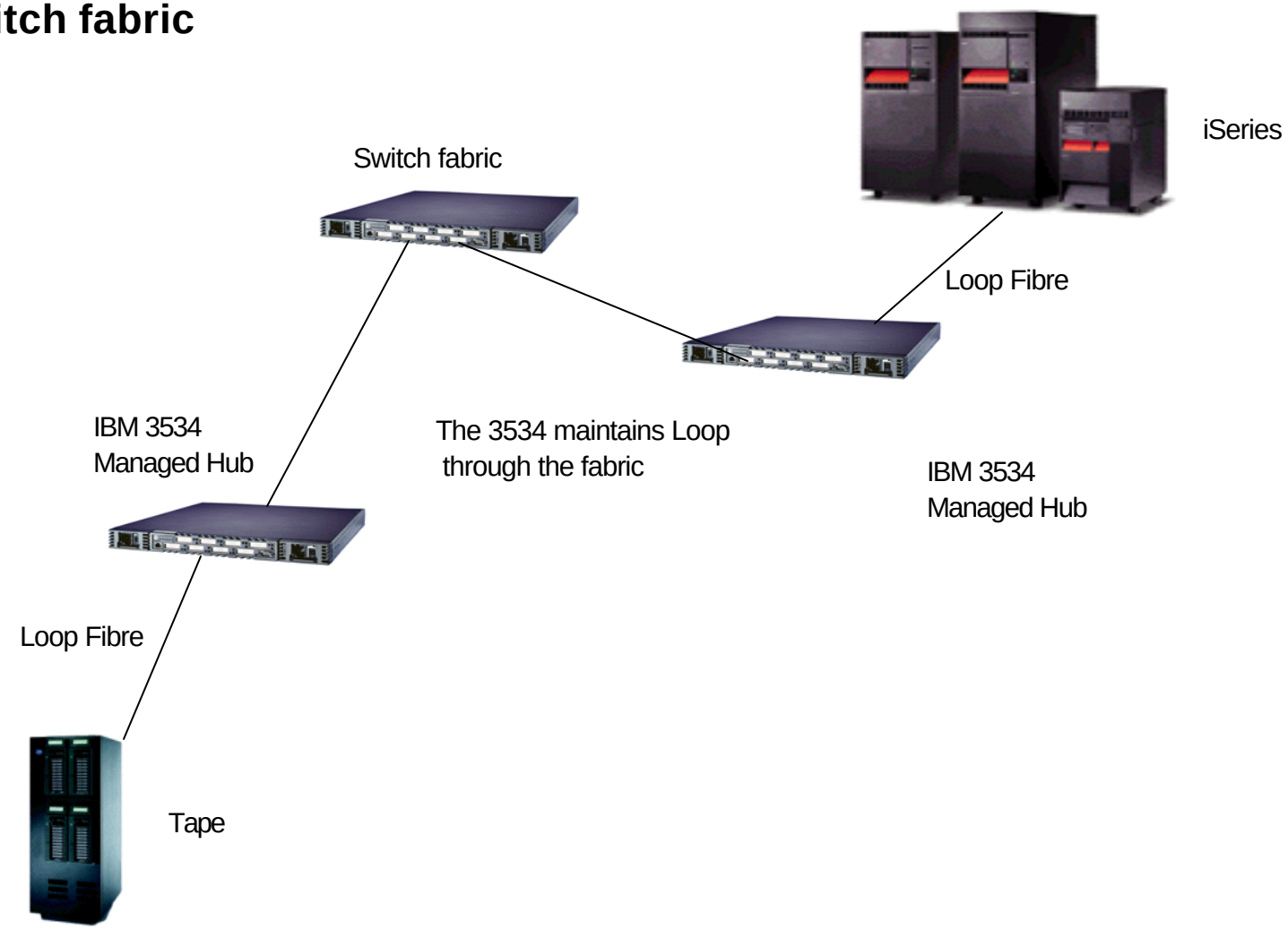
Multiple iSeries - tape FCAL long distance via partnered Hubs



Notes

V5R1 Configurations

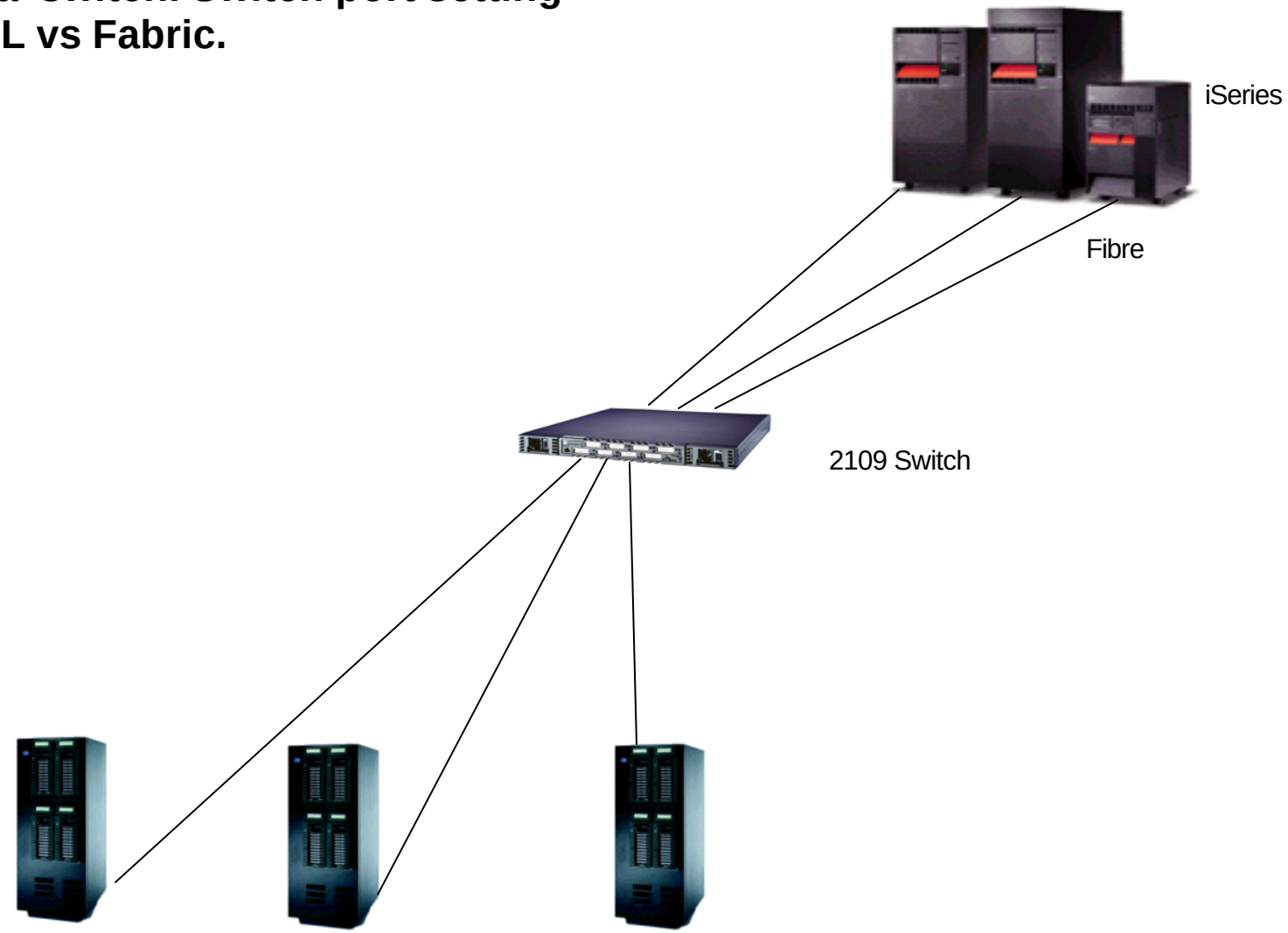
iSeries - connected with 3534 Quick Loop through switch fabric



Notes

V5R2 Configurations

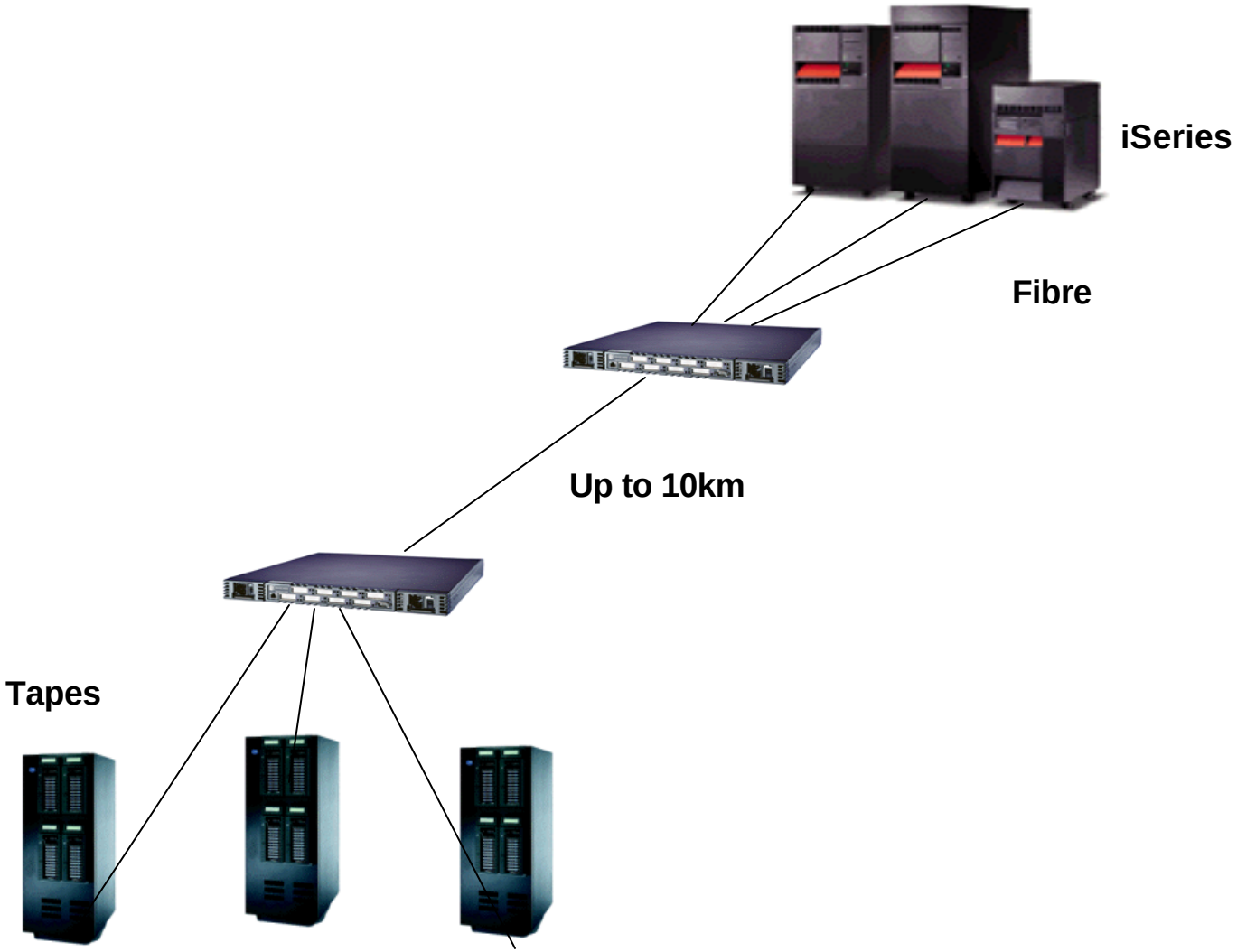
Multiple iSeries - Tape FCAL or Fabric n-to-n via a switch. Switch port setting controls AL vs Fabric.



Notes

V5R2 Configurations

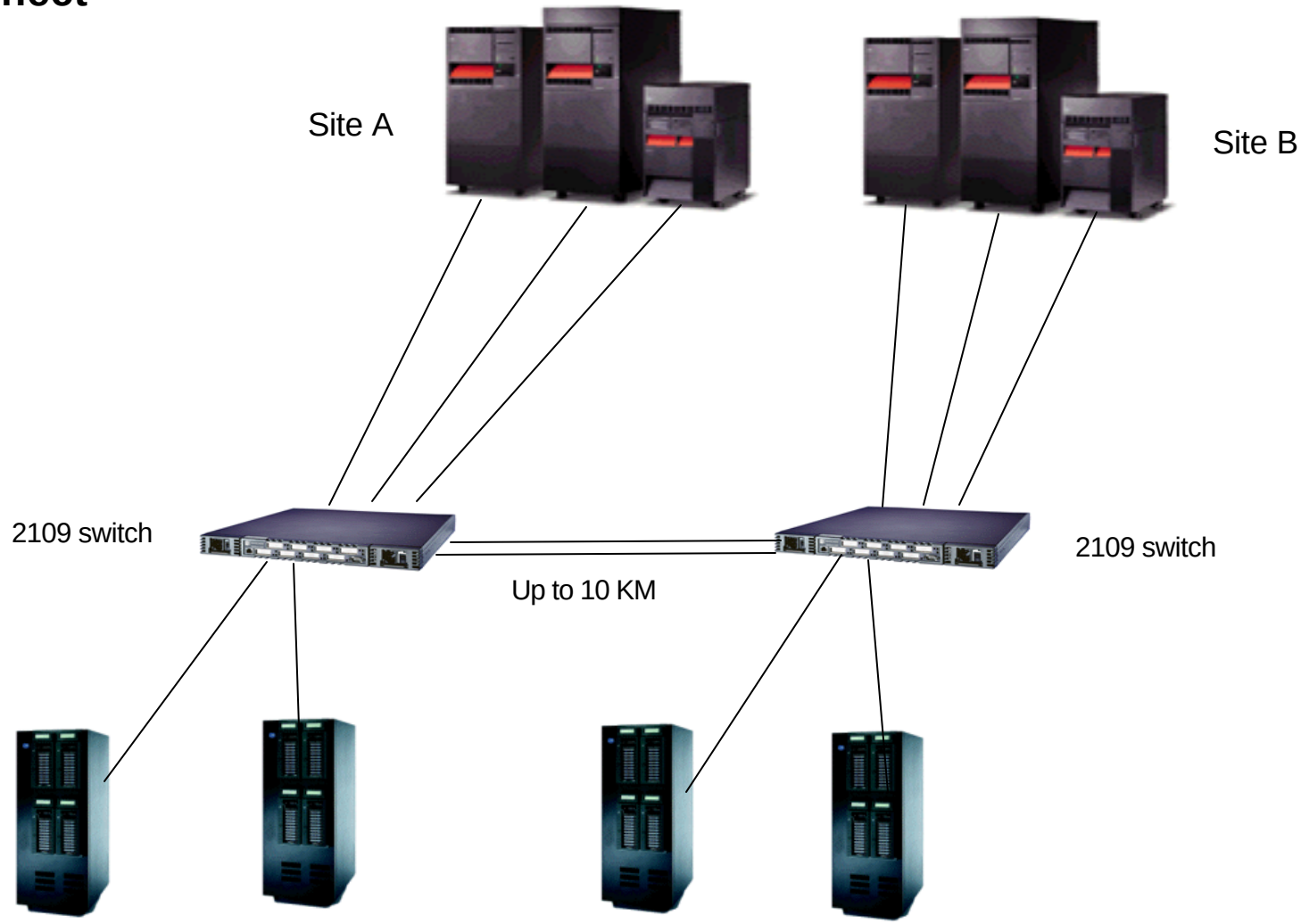
Multiple iSeries - Tape Fabric n-to-n via a switches with 10km link.



Notes

V5R2 Configurations

Multiple iSeries - Tape Fabric for a remote site interconnect



Notes

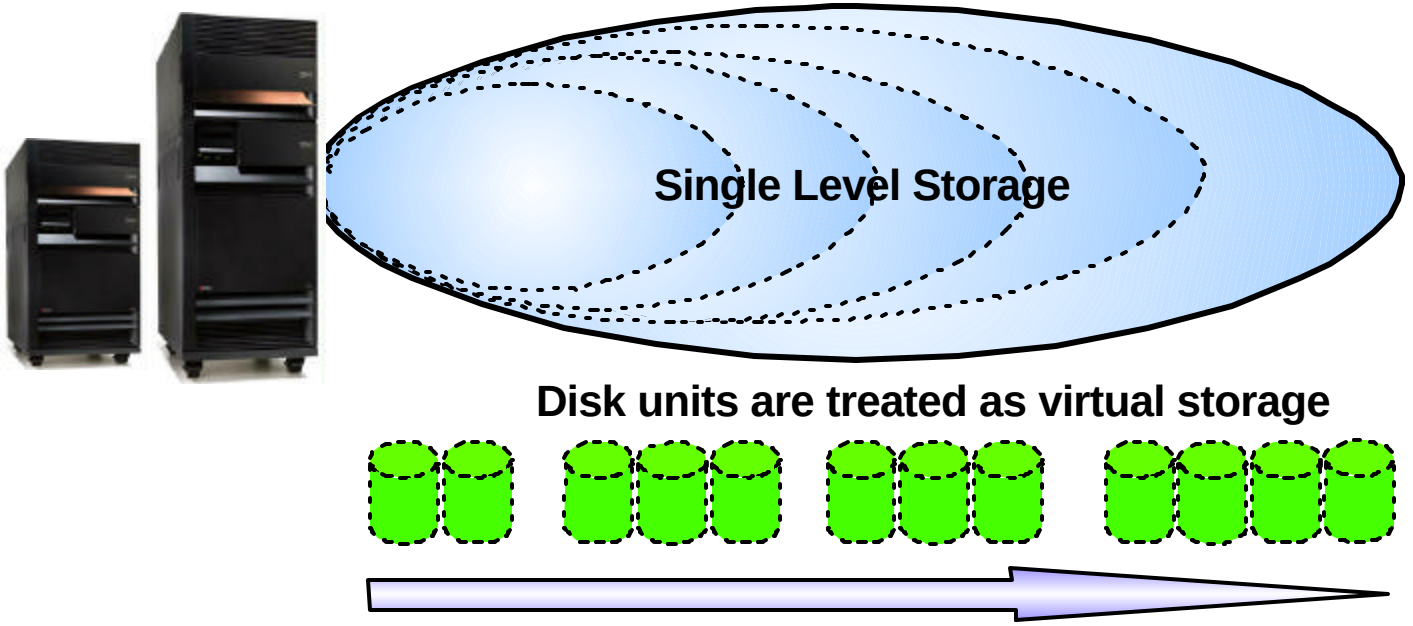
Sources of information

Where to find stuff.....

- www.redbooks.ibm.com - for both iSeries and storage redbooks
  - iSeries in a Storage Area Network - SG24-6220-02 - *updated*
- This presentation will on the IBM Redbooks website in 'Additional Materials' area.
- iSeries Information Center
  - Backup and recovery guide
- www.ibm.com\eserver\iseries\storage
- iTC iASP feasibilty study and workshop

Additional Information

iSeries Storage Evolution



S/38, AS/400, iSeries storage heritage

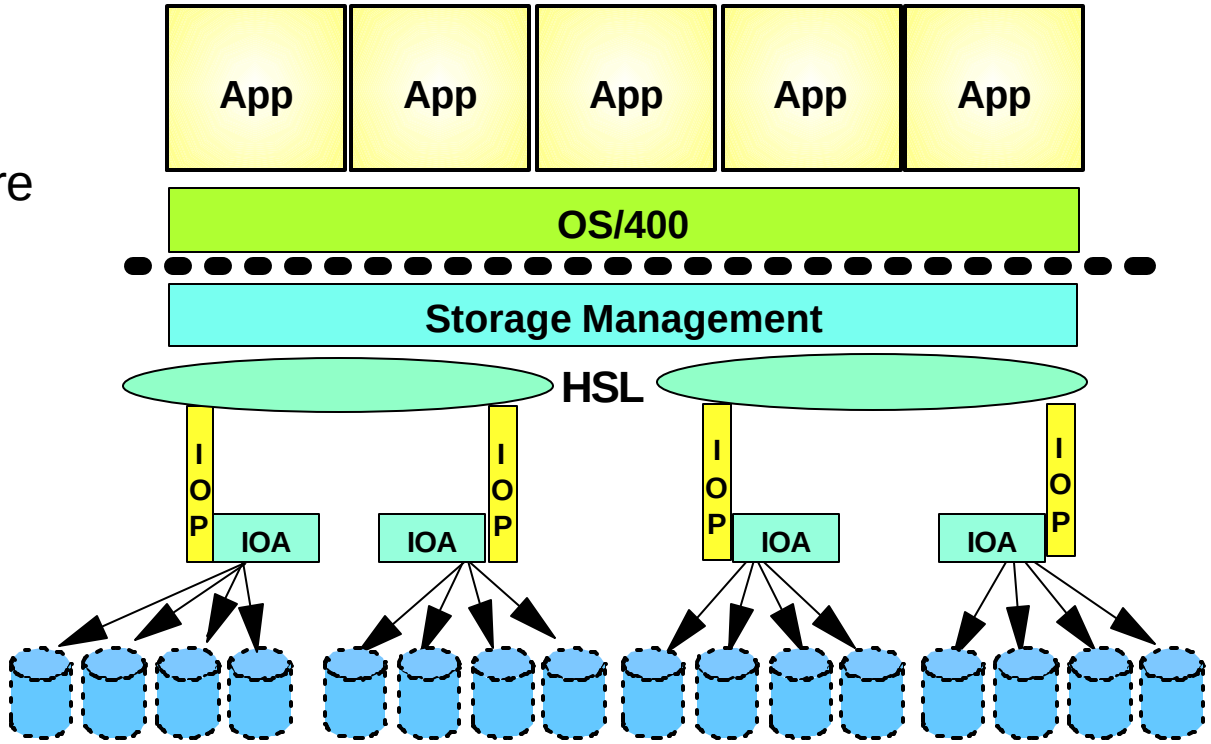
- Always multi-user, multi-application
- Technology Independent Machine Interface hides storage management complexities
- Non-disruptive growth, automated data stripping and load balancing
- Disk drives are essentially viewed as virtual memory store

iSeries Storage Mgmt Architecture

Integrated SAN like features built into the Operating System

Automated storage management

- Single view of storage to applications
- Hardware and software are independent
- Data stripped across all disk units
- Expert cache, parallel I/O
- Workload optimization
- Saves performed at object level to enable easy recovery



Single Level Store

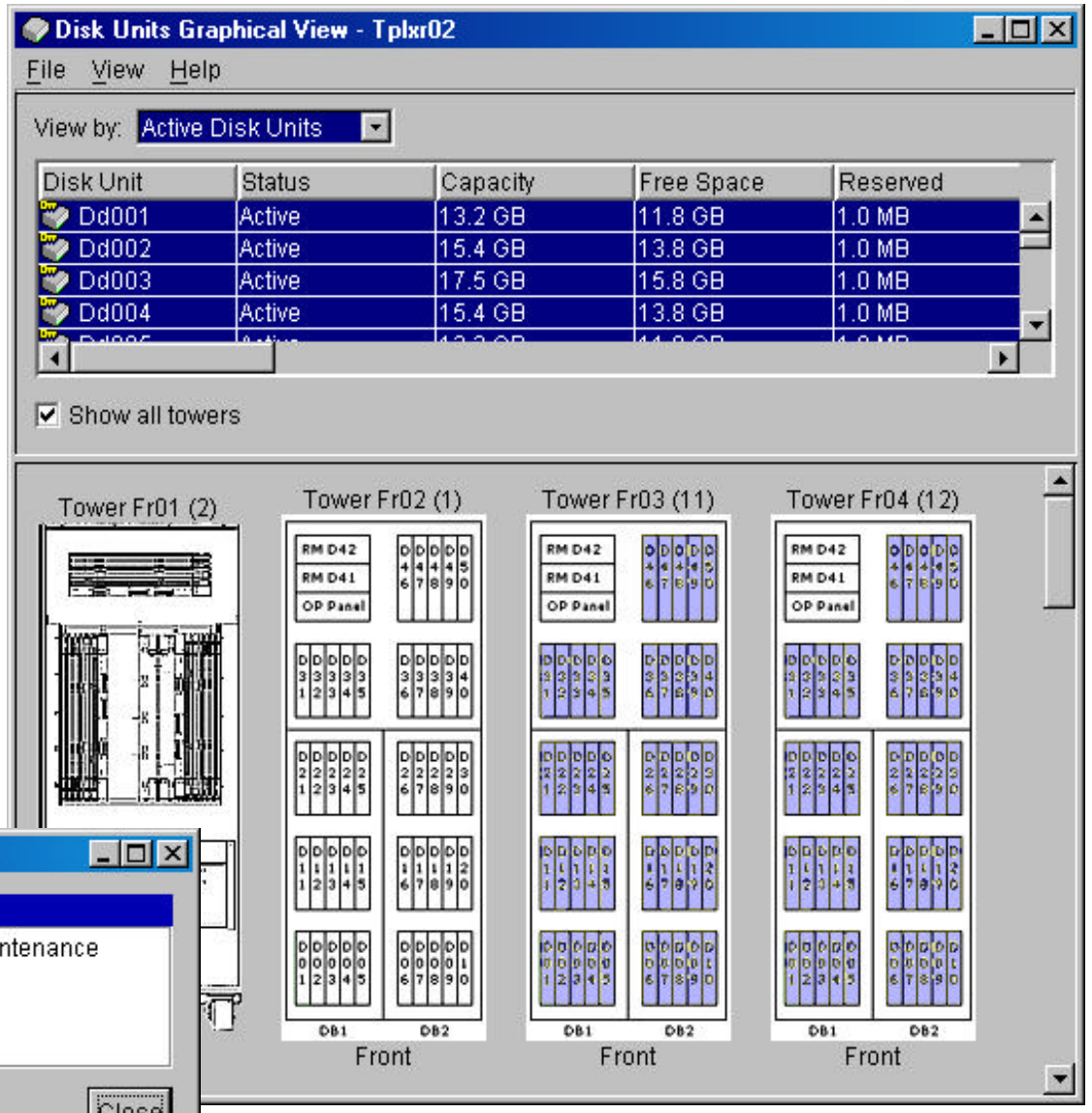
Self Configuring disk managment

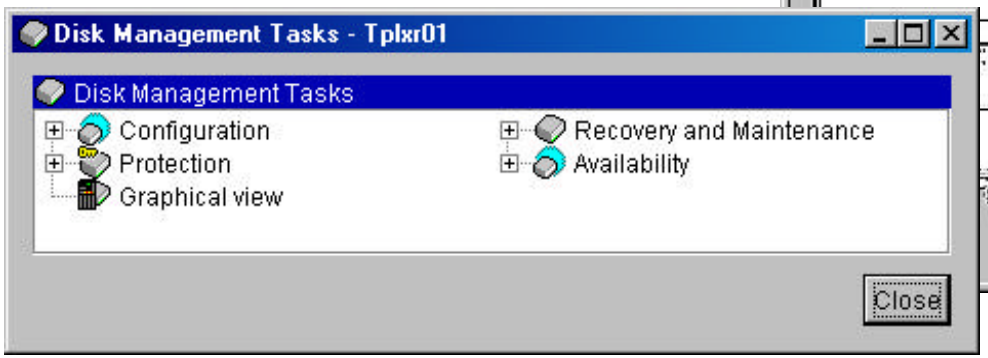
Create, manage and monitor storage virtualization on iSeries

- Secured access
- Configuration, protection, availability, recovery and maintenance

Self-guided wizards

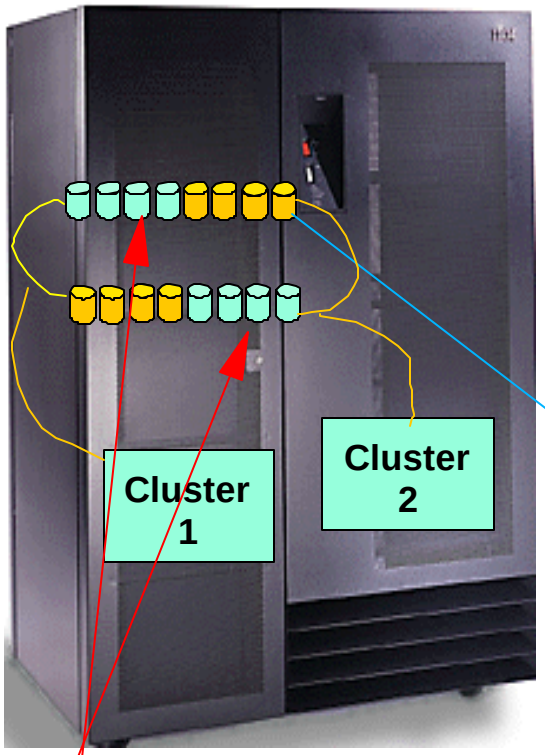
Graphical representation of complex management tasks



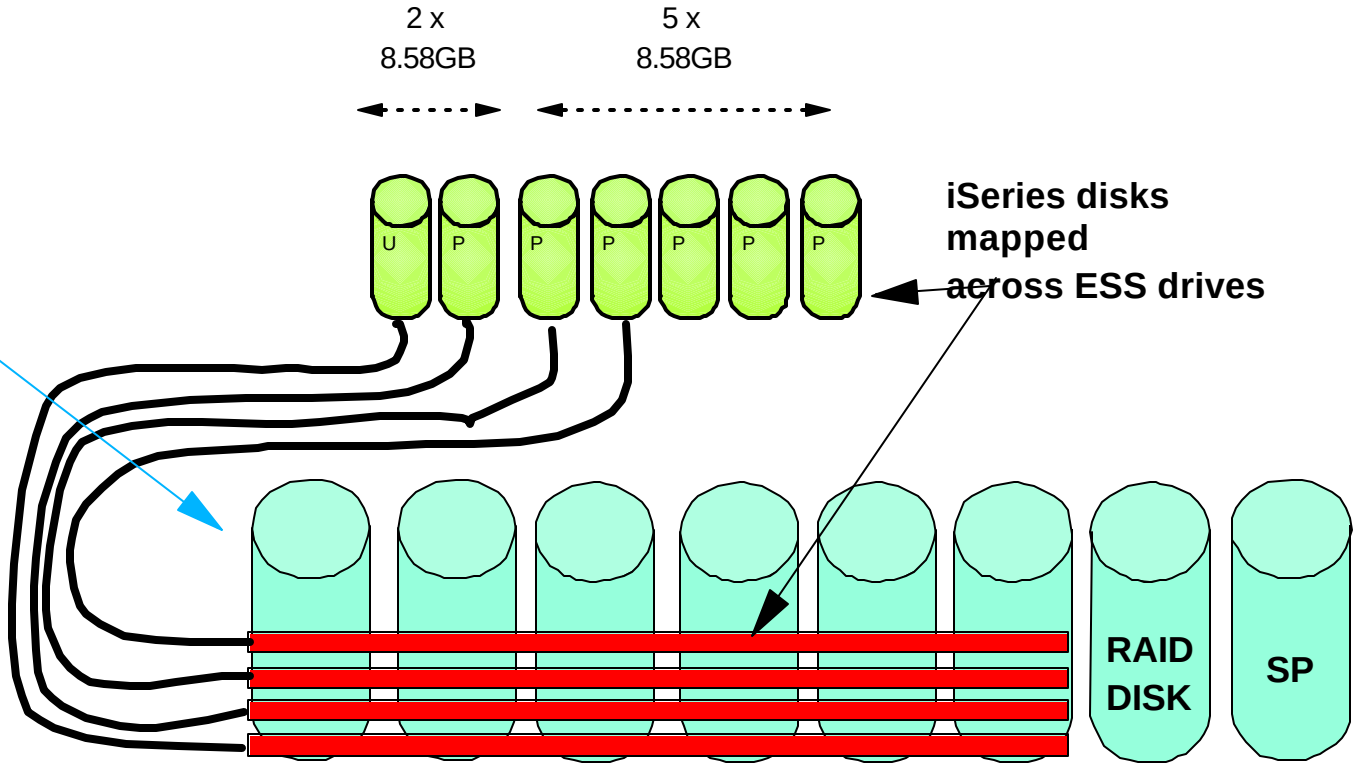


External disk

External storage basics

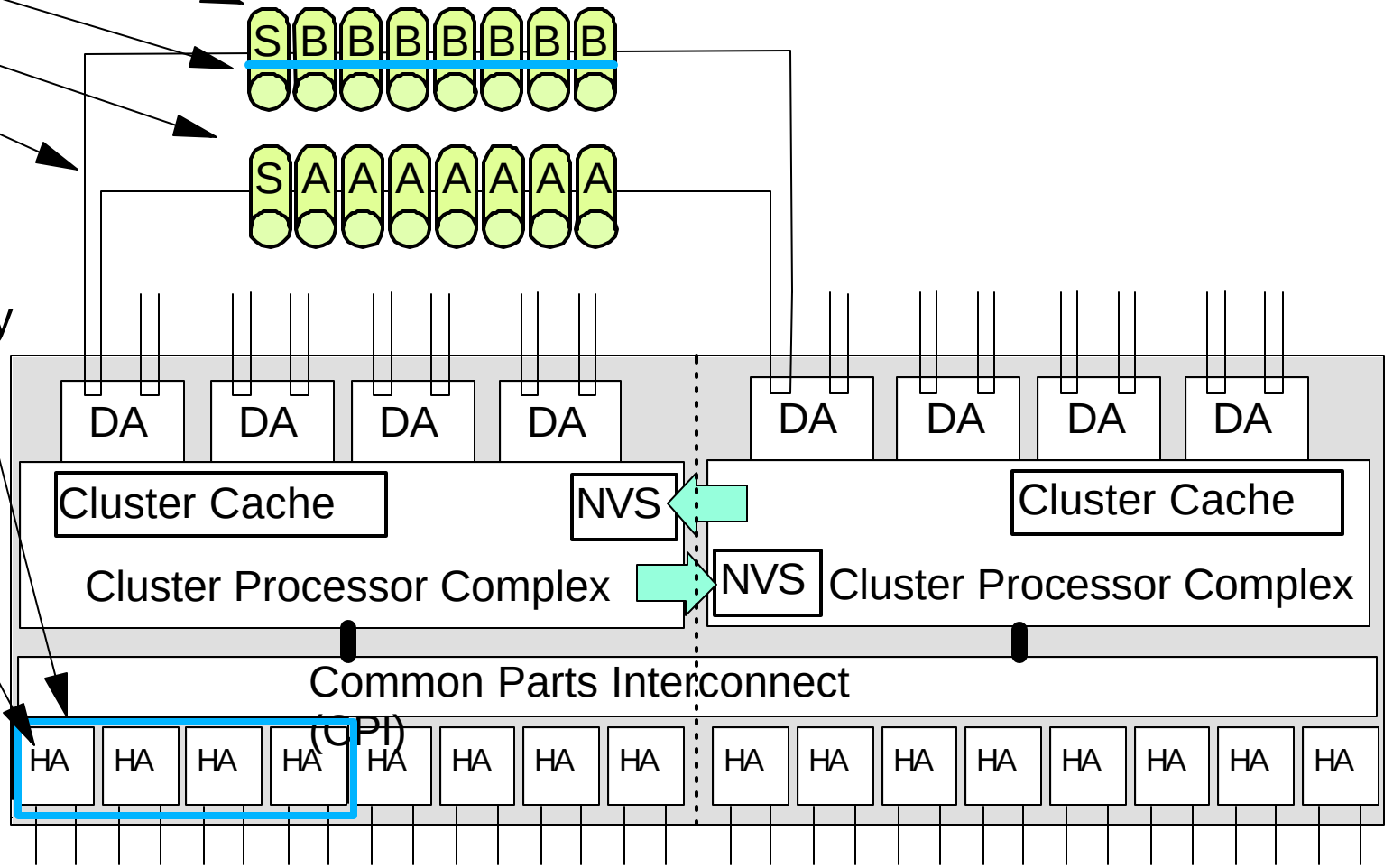


RAID Arrays on same SSA loop



ESS terminology

- DDM
- LUN
- Array
- SSA Loop
- Cluster
- Cluster cache
- NVS
- Host adapter bay
- Host adapter

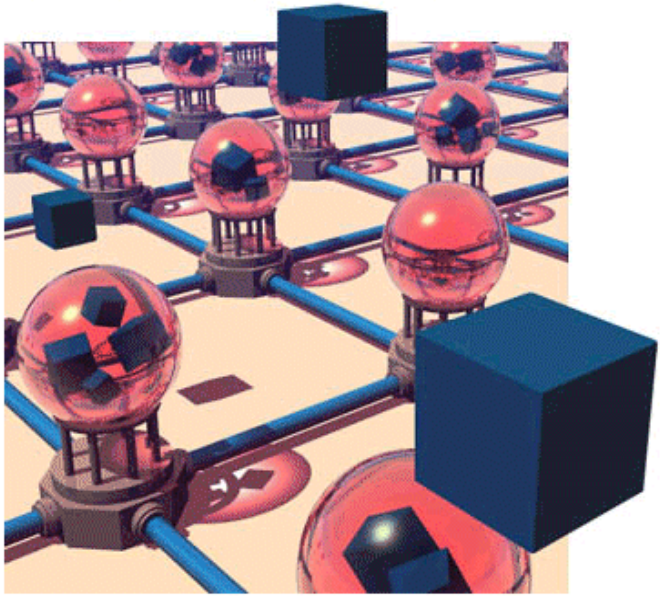


ESS StorWatch Specialist

StorWatch Solutions

- Introduction
- Status
- Problem Notification
- Communications
- Storage Allocation
- Users
- Licensed Internal Code
- Copy Services

Enterprise Storage Server Specialist



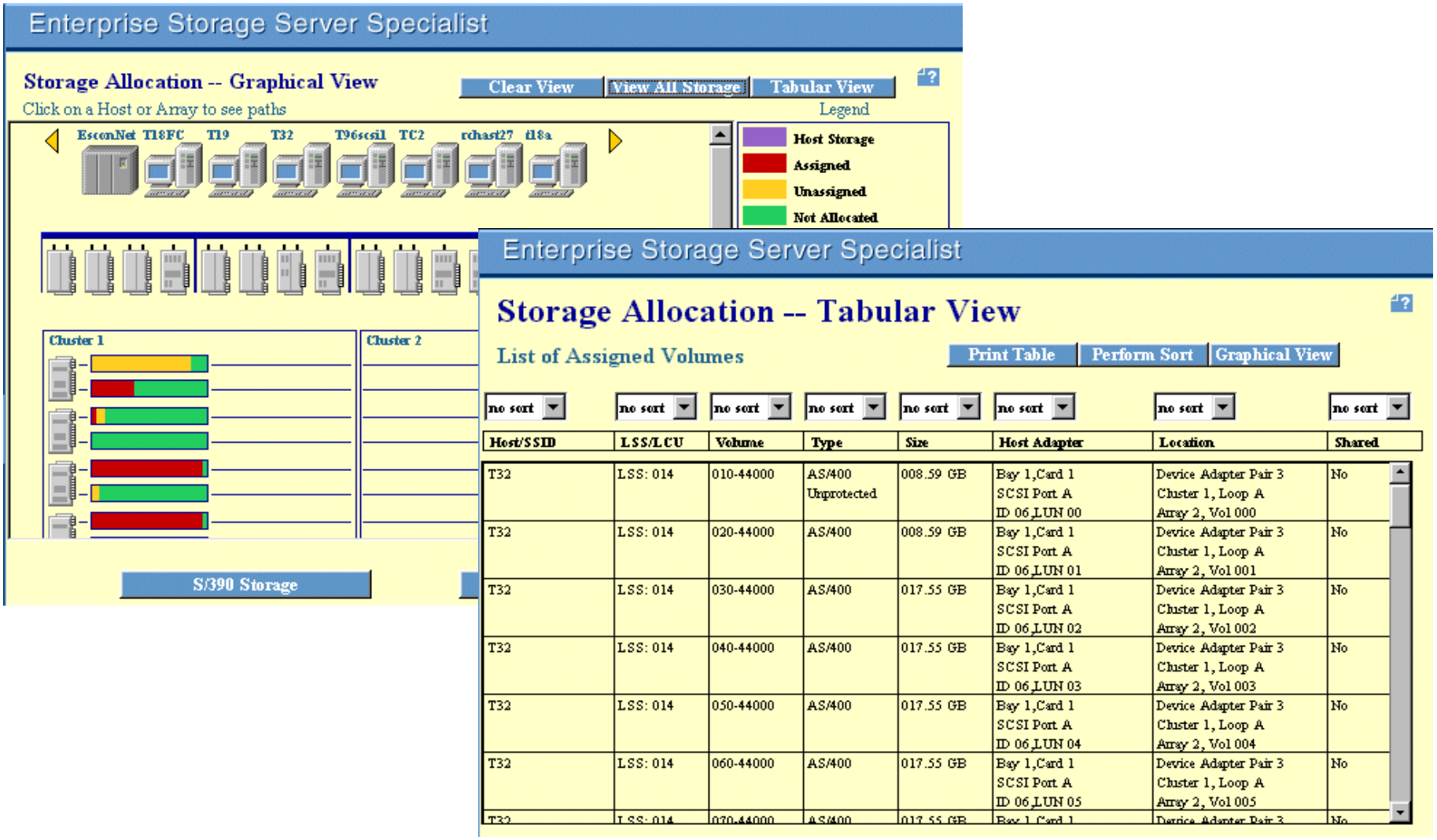
Welcome to

IBM StorWatch Enterprise Storage Server Specialist

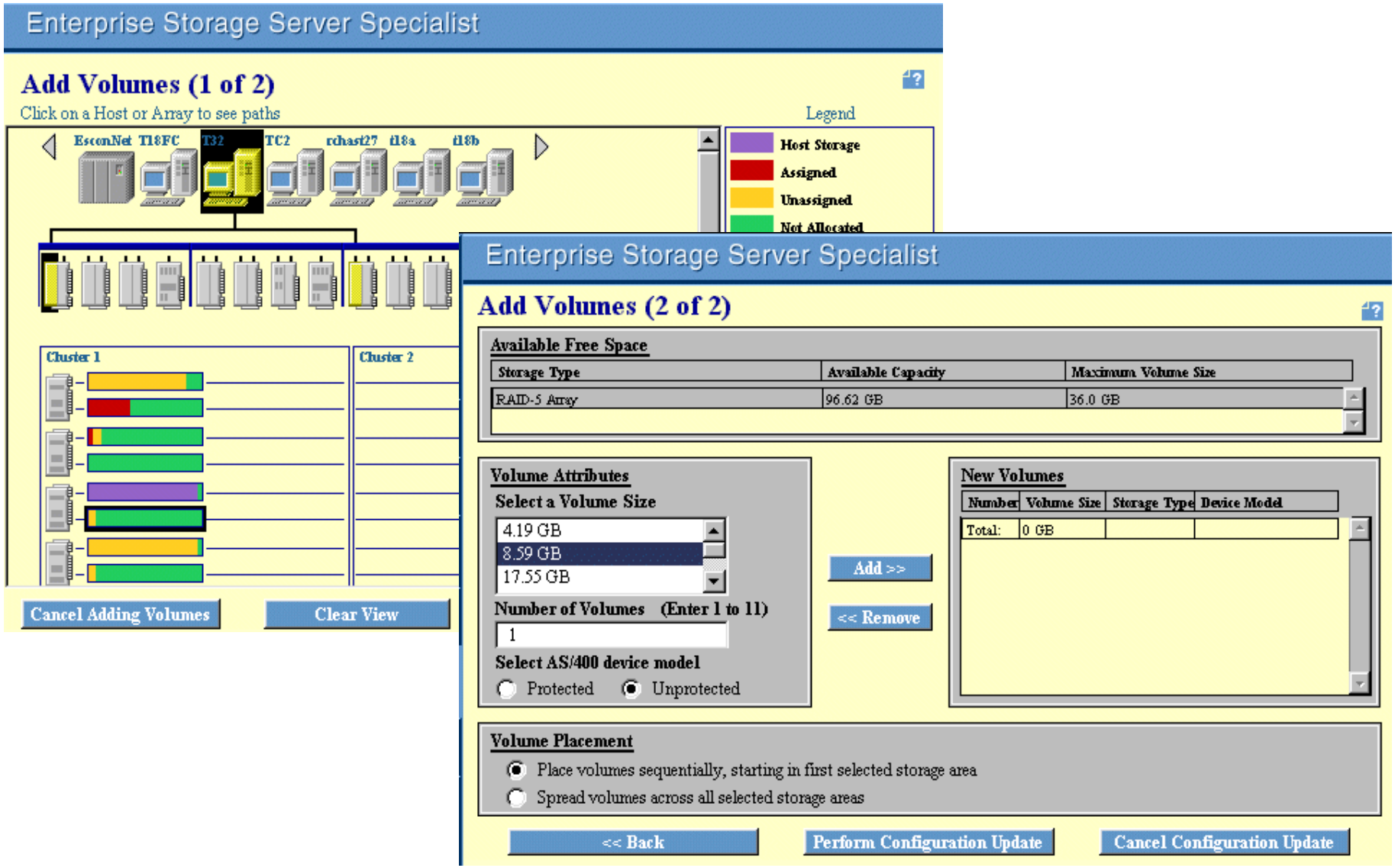
```
Machine Type:      2105
Machine Model:     E20
Serial Number:     075-14419
WWNN:        5005076300C00A40
```


View README

ESS StorWatch Configuration



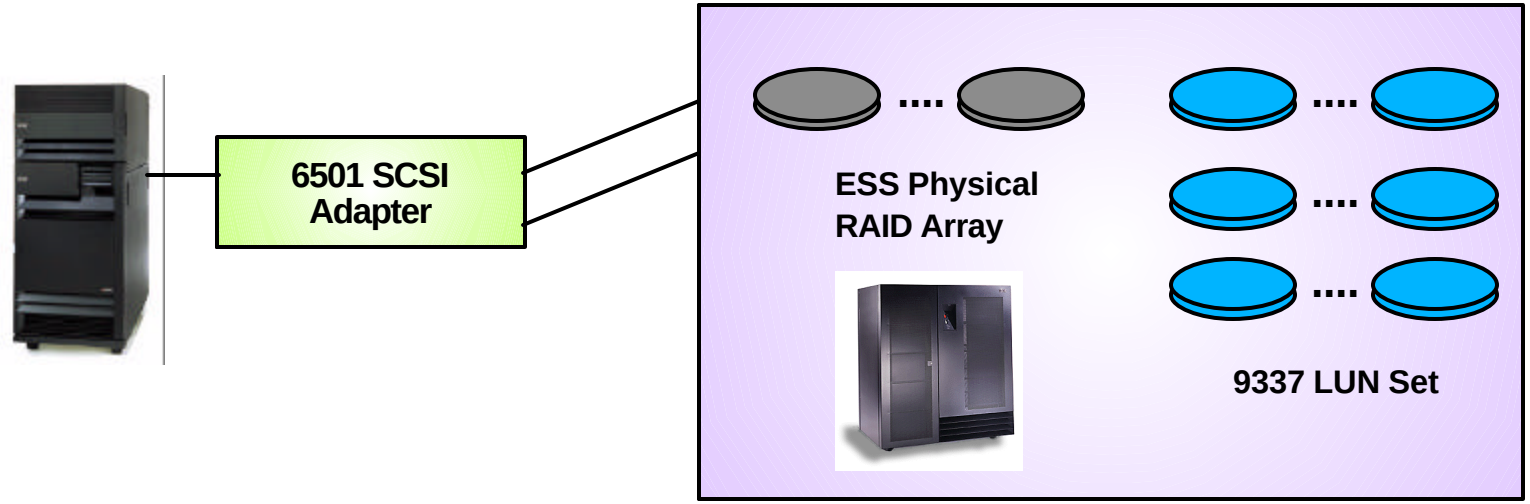
Configuring storage



iSeries external storage attach

SCSI External DASD Storage Attachment

ESS 9337 Emulation

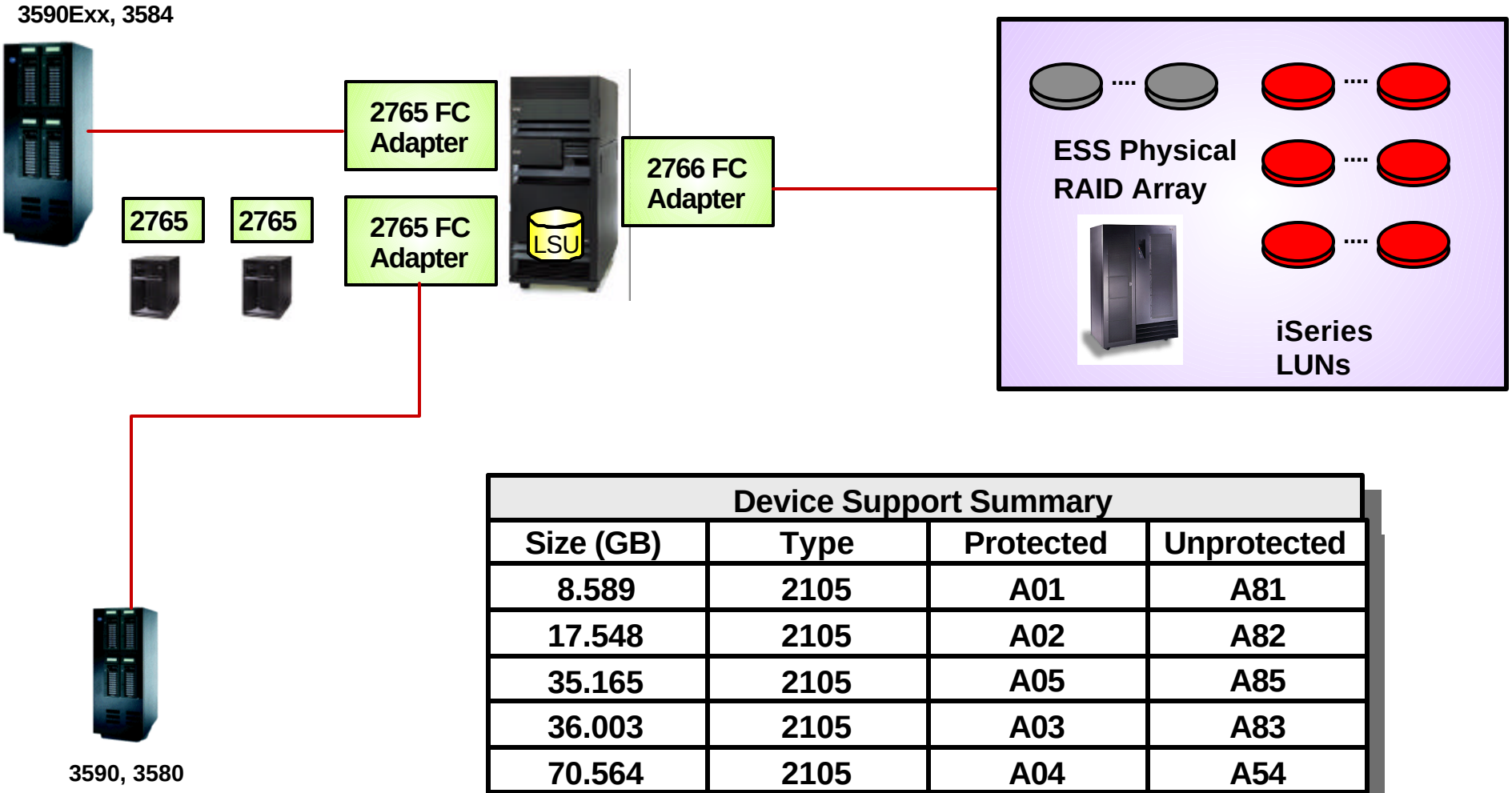


Device Support Summary			
Size (GB)	Type	Protected	Unprotected
4.190	9337	48C	48A
8.589	9337	59C	59A
17.548	9337	5AC	5AA
36.003	9337	5BC	5BA

Used for Load Source mirroring

Used for all other iSeries drives

FC External Storage Attachment

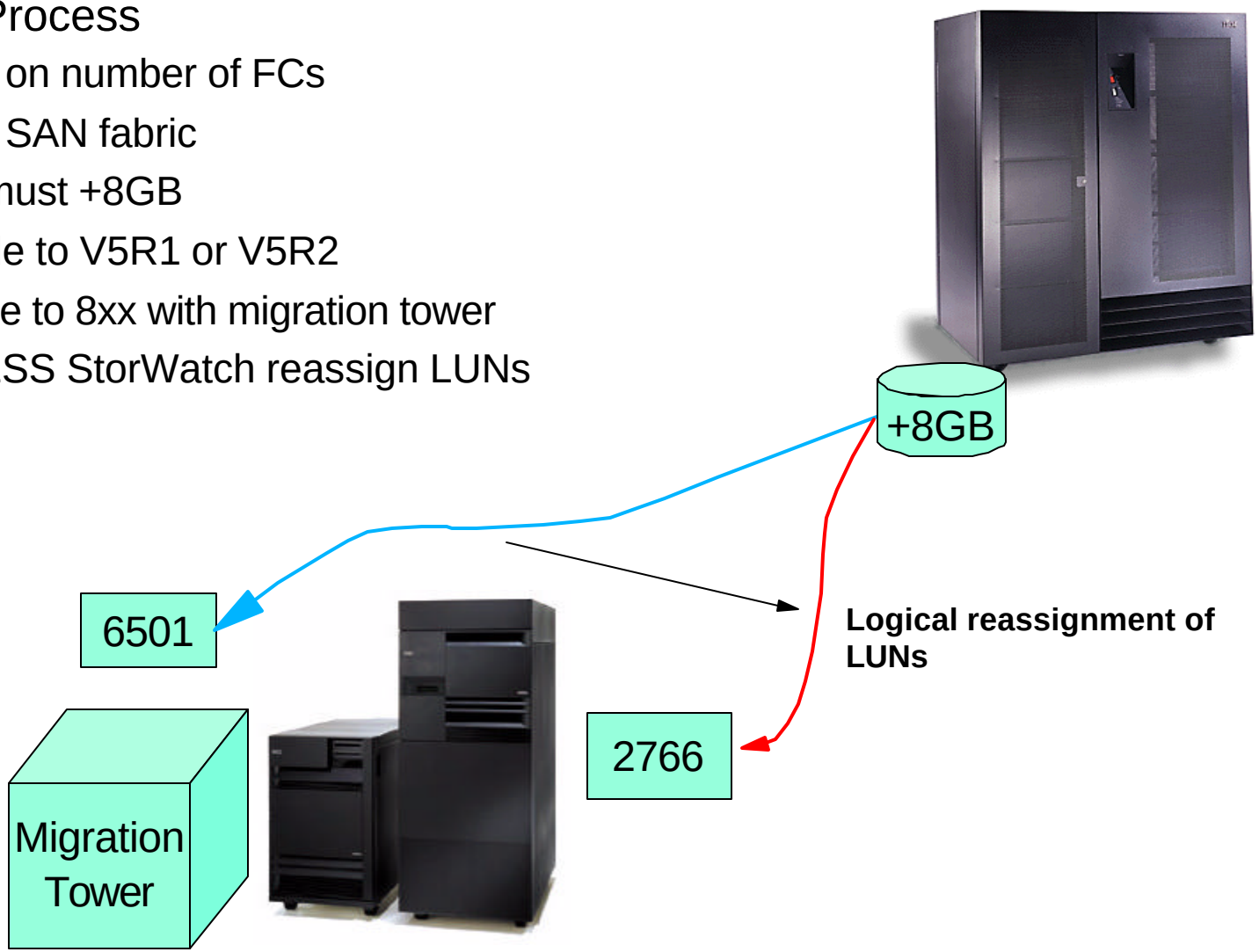


Device Support Summary			
Size (GB)	Type	Protected	Unprotected
8.589	2105	A01	A81
17.548	2105	A02	A82
35.165	2105	A05	A85
36.003	2105	A03	A83
70.564	2105	A04	A54

Disk migration - SCSI to FC

Simple Process

- Decide on number of FCs
- Design SAN fabric
- LUNs must +8GB
- Upgrade to V5R1 or V5R2
- Upgrade to 8xx with migration tower
- From ESS StorWatch reassign LUNs



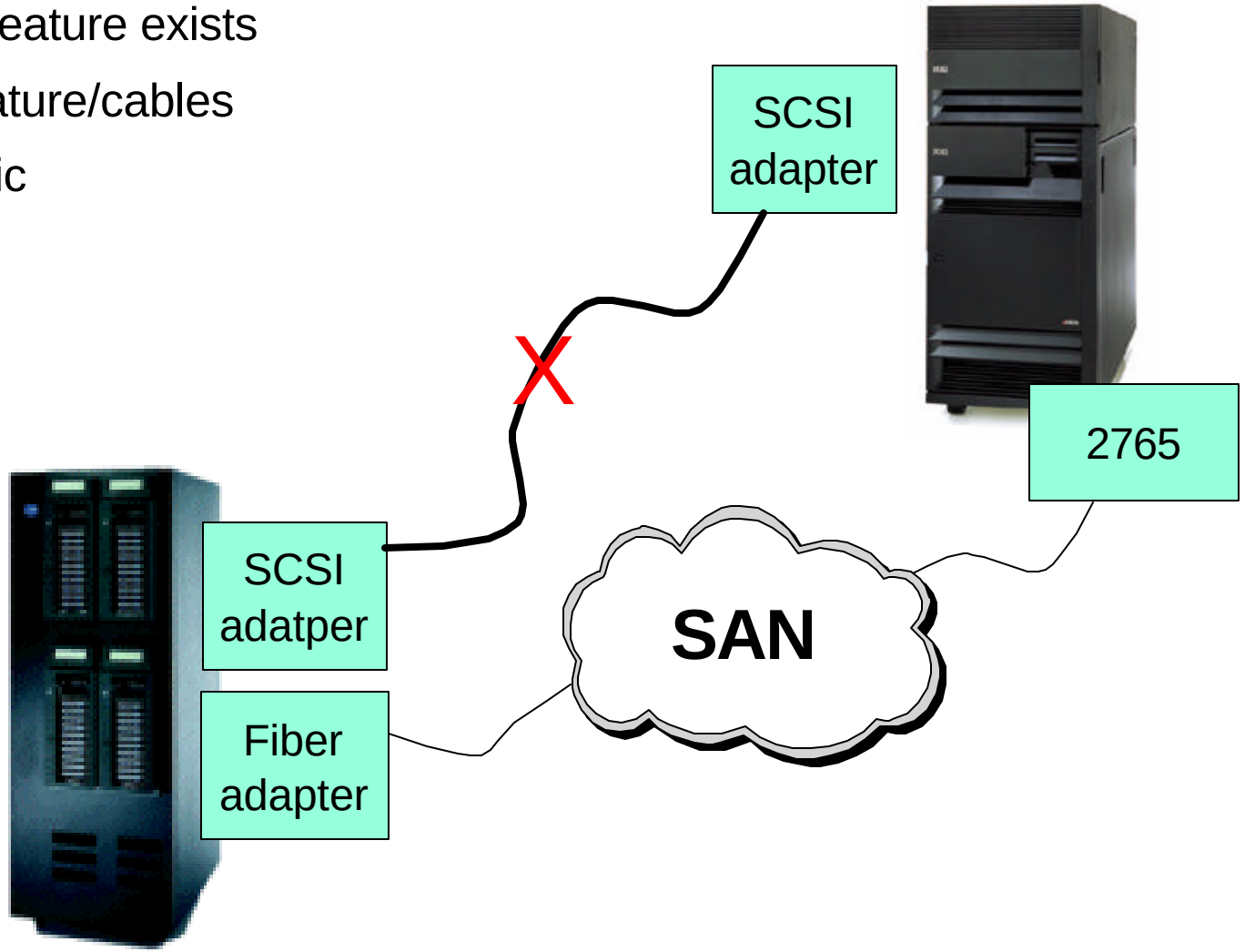
Tape migration - SCSI to FC

Again a simple process

Establish if tape FC feature exists

Swop attachment feature/cables

Attach FC cable/fabric



Remote load source mirror

New function available Sept 2000

PTFs available

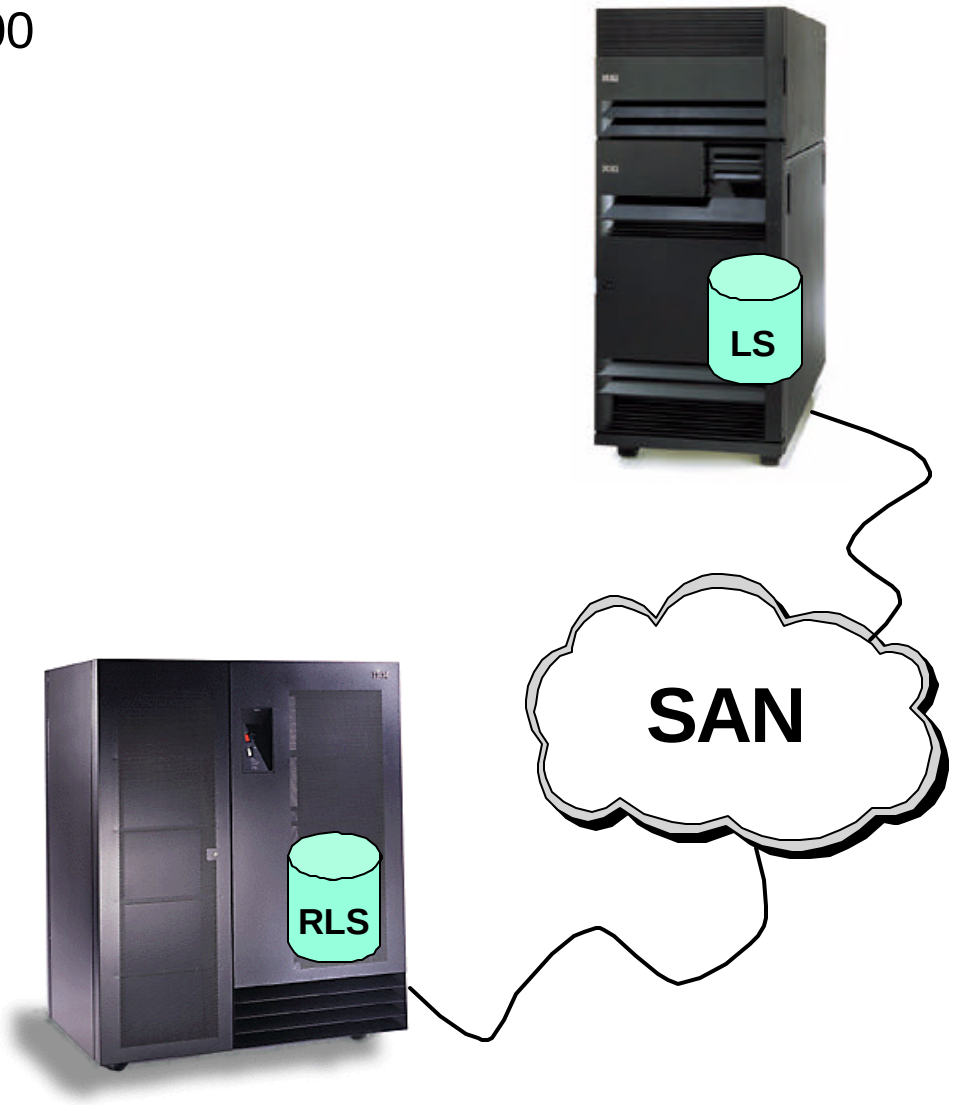
- V4R3 and later on SCSI
- V5R1 on FC

Unprotected LUN on ESS

Allows

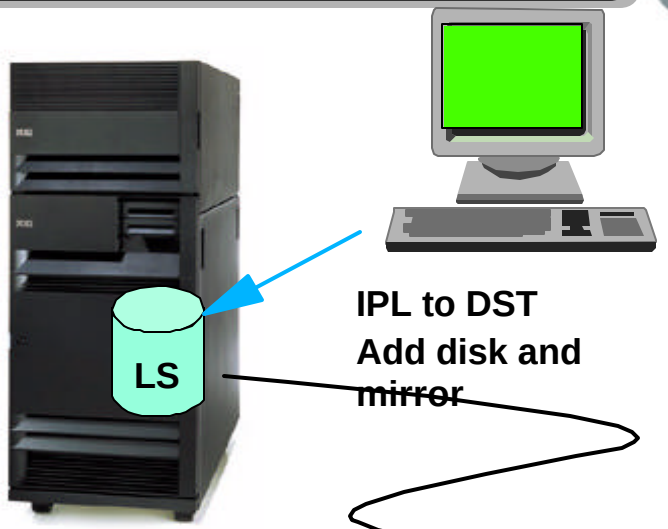
- remote ESS
- DR site recovery

Remeber single level storage

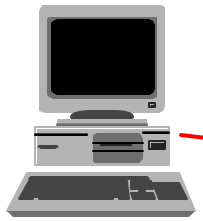


Implementing RLS mirror

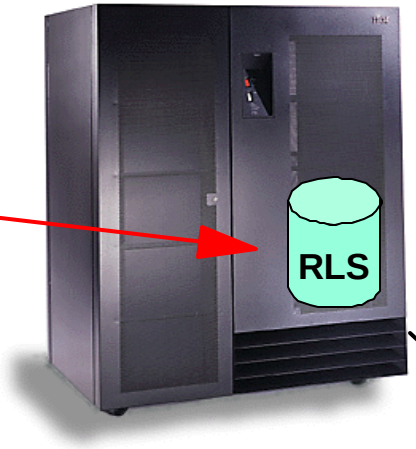
1. No internal disk other than LS
2. Create unprotected LUN on ESS (other LUNs protected)
3. Size equal to LS
4. Assign LUN to iSeries
5. IPL to DST
6. Disk Management
7. Start mirroring
8. Add disk to System ASP



IPL to DST
Add disk and mirror



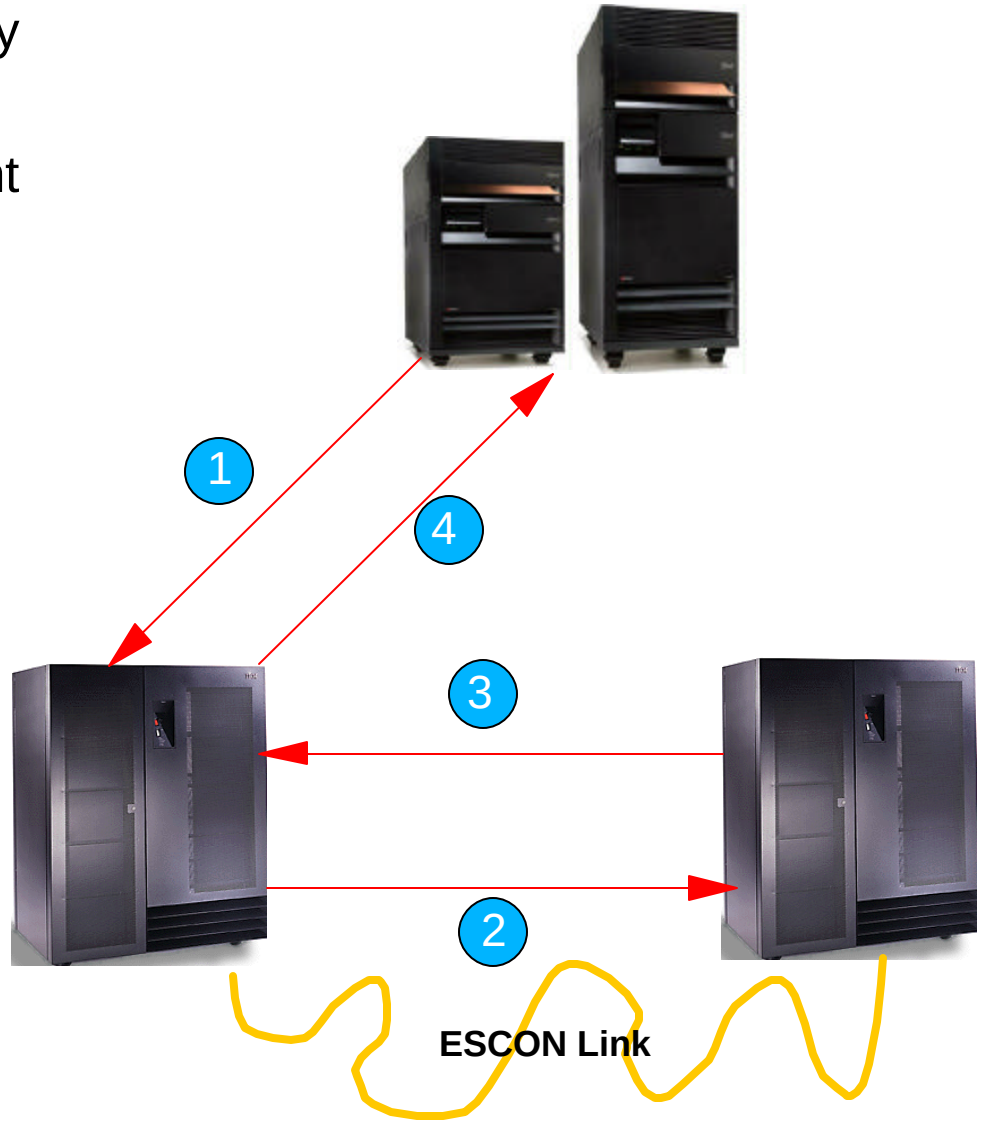
ESS StorWatch Specialist
Create and assign function



SAN

Point-to-Point Remote Copy (PPRC)

1. Write (channel end)
2. Synchronous write to secondary
3. Write hit at secondary
4. Asynchronous acknowledgment (Device End)
5. Requires ESCON link(s)



iSeries configuration guidelines

Card placement rules for iSeries

#2765

- e-Config will enforce rules
- Requires about half IOP resources
  - Imbedded IOP in CEC/#5075
  - #2842 PCI IOP
  - #2843/#9943 PCI IOP
- Not more than 2 #2765 per EDS boundary
- Need fiber connector
  - #0371 - 50 micron
  - #0372 - 62.5 micron
- System CEC/Expansion unit maximums (any combination)
  - 270 2 x #2765/#2766
  - 820 4 x #2765/#2766
  - 830/840 base 5 x 2765/#2766
  - 5074/0578/5078/bot 5079/top 8079 6 x #2765/#2766

Migration planning 1

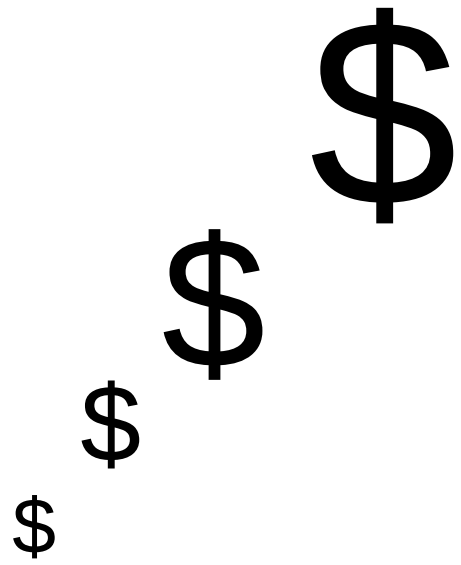
#6501 to #2766

- Requires V5R1
- iSeries/ESS planning
- Need FC feature on ESS
- Size the SAN fabric
  - Watch for bottlenecks
  - Install any SAN fabric requirement (hubs, cables)
- Upgrade or box swop to iSeries 270/8xx
- LUNs must be greater than 8GB
  - 4GB LUNs will require a reload
- Re-assign LUNs in storWatch
- Suspend remote LS mirror

Why SAN ?

The challenge IT managers are facing today ....

- Costs of Storage Management
  - Distributed UNIX/NT costs 1 person per 250 GB
  - Centralized UNIX/NT costs 1 person per 500 GB
  - OS/390 only costs 1 person per 1 to 7 Terabytes!
- Cost of Down-time
  - From thousand to millions of Dollars



SAN Business Values

Enhanced Connectivity

Storage Management Disciplines & Standardization

Backup/Recovery Performance Improvements

Disaster Tollerance

Resource Sharing

Reduced Data Duplication (NT/W2000)

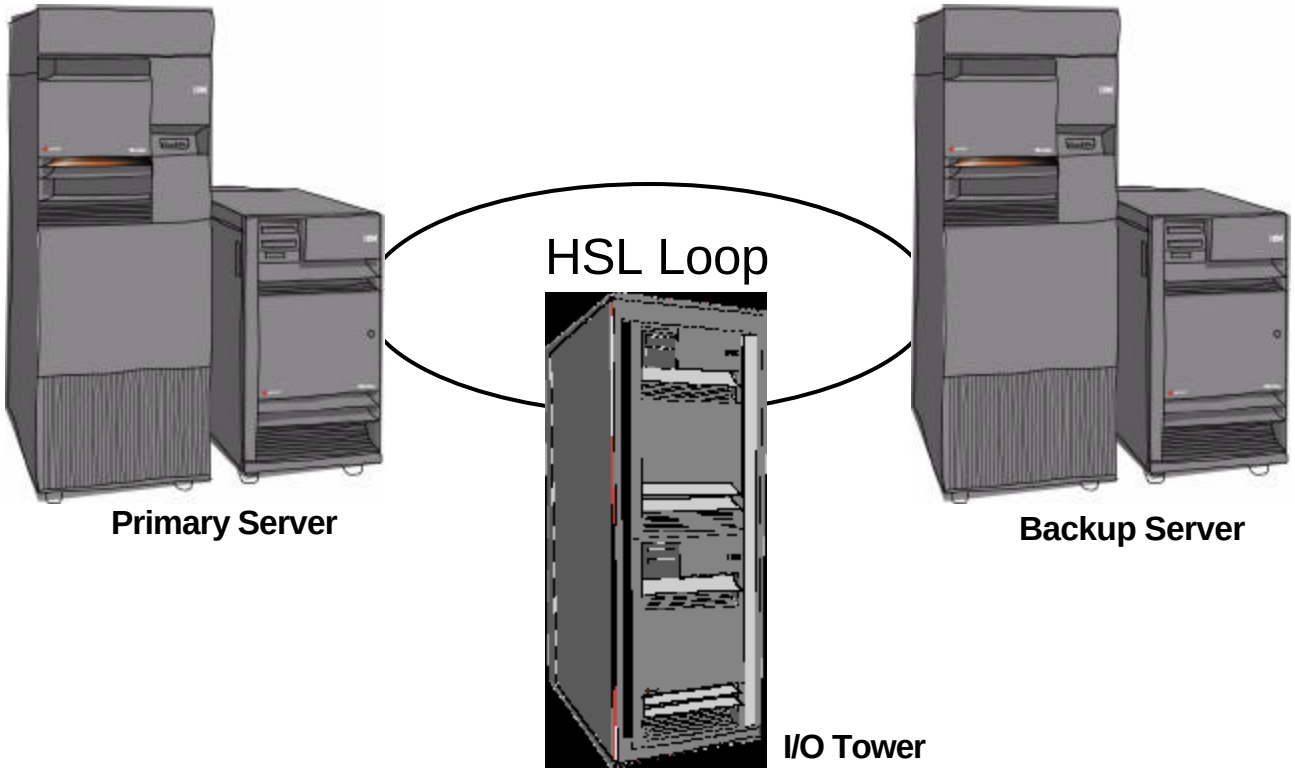
Data Center & Disaster Center Solution

Replication Data Resiliency - Part of the iSeries Cluster Architecture

- **Can provide the most resilient availability solution**
- **Scheduled Outages: Concurrent Saves & Maintenance Procedures**
- **Unscheduled Outages: Failover support for primary server failure**
- **Workload balancing (Queries, Batch, Web etc) on Backup server**
- **No IPL required for switchover**

IASPs: Data Center Solution

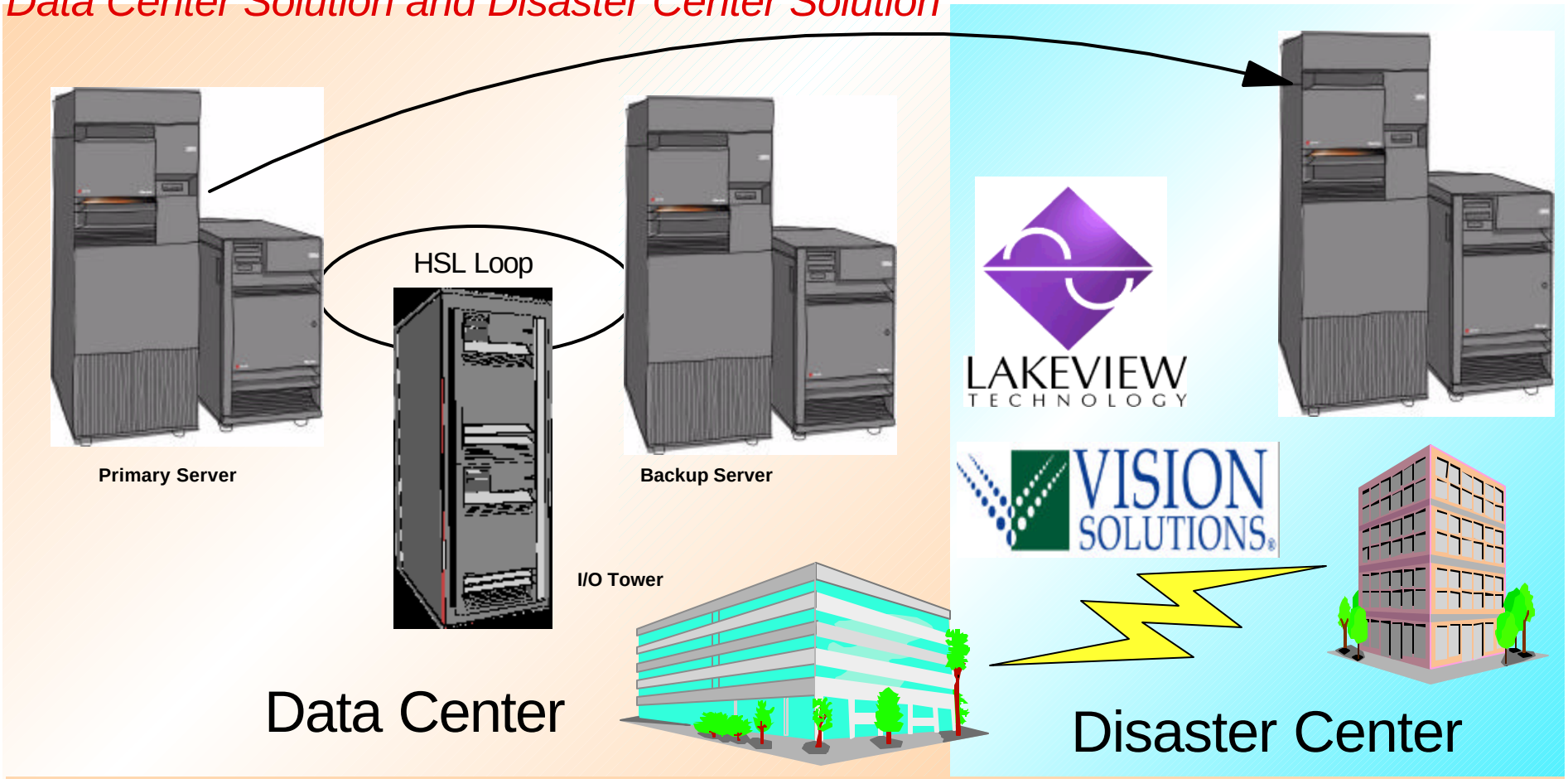
Device Resiliency/Switchable DASD Tower - Part of the iSeries Cluster Architecture



- **Integrated OS/400 Solution (HACMP on iSeries !)**
- **Scheduled switchover support for maintenance - hardware & OS**
- **Unscheduled failover for primary server failure**
- **Distributed homogeneous application consolidation**
- **No IPL required for switchover**

Replication & IASP Hybrid Solution

Data Center Solution and Disaster Center Solution



- **Provides Data Center and Disaster Center availability solution**
- **Data Center: Scheduled Outage support for maintenance**
- **Data Center: Unscheduled failover for primary server failure**
- **Disaster Center solution: Remote recovery, archives back-up/recovery**