iSeries:
Planning for Continuous Availability

June 11, 2003

iSeries Advanced Technical Sales Support
Sue Baker - *smbaker@us.ibm.com*
Eric Hess   -   *texas@us.ibm.com*

Agenda

- Value of availability

- IT Challenges in the Business World and iSeries technology solutions
  - ► Outages
  - ► Storage
  - ► Systems management and consolidation
  - ► Backup and recovery
  - ► Clustering

- Summary

Value of availability

The value of availability

- Is "IT" mission critical to your business?
  - ► A competitive or image advantage?
  - ► A productivity advantage

- How much does it cost to be down?

- How often do you test your recovery plan?
  - ► Does it work?

- Does your business have a Business Continuity plan?
  - ► Is it a business plan, or an IT plan?
    - – Does it cover "IT" only, or the entire business?
    - – Who owns the Business Continuity plan?

- Frequency vs. duration

- Planning help:
  - ► IBM Cost of Downtime Calculator
    - – www-1.ibm.com/services/its/us/dtw10.html
  - ► www.availability.com
  - ► www.contingencyplanning.com

Notes: The value of availability

When disaster recovery emerged as a formal discipline and a commercial business in the 1980s, the focus was on protecting the data center—the heart of a company's heavily centralized IT structure. This model began to shift in the early 1990s to distributed computing and client/server technology. At the same time, information technology became embedded in the fabric of virtually every aspect of a business. Computing was no longer something done in the background. Instead, critical business data could be found across the enterprise—on desktop PCs and departmental local area networks, as well as in the data center.

This evolution continues today. Key business initiatives such as enterprise resource planning (ERP), supply chain management, customer relationship management and e-business have all made continuous, ubiquitous access to information crucial to an organization. This means business can no longer function without information technology: data, software, hardware, networks, call centers—even laptop computers.

A company that sells products on the Web, for example, or supports customers with an around-the-clock call center, must be operational 24 hours a day, 7 days a week—or customers will go elsewhere. An enterprise that uses e-business to acquire and distribute parts and products is not only dependent on its own technology but that of its suppliers. As a result, protecting critical business processes, with all their complex interdependencies, has become as important as safeguarding data itself.
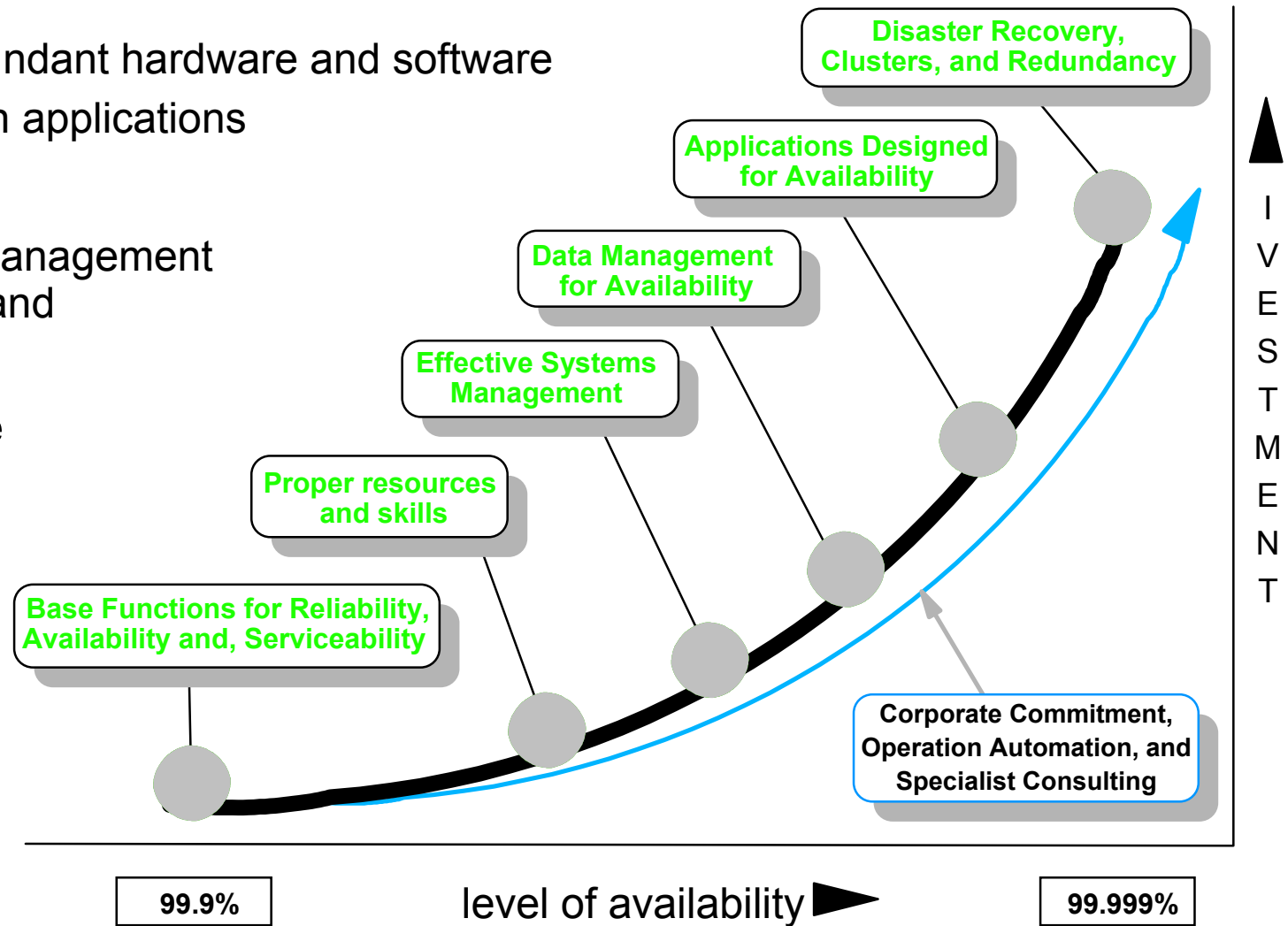
The goal for companies with no business tolerance for downtime is to achieve a state of business continuity, where critical systems and networks are continuously available, no matter what happens. This means thinking proactively: engineering availability, security and reliability into business processes from the outset—not retrofitting a disaster recovery plan to accommodate ongoing business continuity requirements.

From: Business Continuity: New risks,new imperatives and a new approach: www-1.ibm.com/services/continuity/recover1.nsf/files/Downloads/$file/buscont.pdf

The IBM Cost of Downtime Calculator is a Lotus 123 or Microsoft Excel spreadsheet (your choice) that will help you determine base outage costs for your business, and areas that need additional review. This tool maybe downloaded from:
www-1.ibm.com/services/its/us/dtw10.html

Availability is

- **A technology**
  - ► Resilient and redundant hardware and software
  - ► Intelligent, modern applications

- **A discipline**
  - ► Strong systems management disciplines, skills and education

- **A corporate culture**
  - ► Focus and commitment by both the business and IT organization

Disaster Recovery, Clusters, and Redundancy

Applications Designed for Availability

Data Management for Availability

Effective Systems Management

Proper resources and skills

Base Functions for Reliability, Availability and, Serviceability

Corporate Commitment, Operation Automation, and Specialist Consulting

INVESTMENT

| 99.9% | level of availability ▶ | 99.999% |

Notes: Availability is a corporate culture

Many senior executives and business managers consider business continuity the responsibility of the IT department. However, it is no longer sufficient or practical to vest the responsibility exclusively in one group. Web-based and distributed computing have made business processes too complex and decentralized. What's more, a company's reputation, customer base and, of course, revenue and profits are at stake. All executives, managers and employees must therefore participate in the development, implementation and ongoing support of continuity assessment and planning.

The same information technology driving new sources of competitive advantage has also created new expectations and vulnerabilities. On the Web, companies have the potential to deliver immediate satisfaction—or dissatisfaction—to millions of people. Within ERP and supply chain environments, organizations can reap the rewards of improved efficiencies, or feel the impact of a disruption anywhere within their integrated processes.

With serious business interruption now measured in minutes rather than hours, even success can bring about a business disaster. Web companies today worry more about their ability to handle unexpected peaks in customer traffic than about fires or floods—and for good reason. For example, an infrastructure that cannot accommodate a sudden 200 percent increase in Web site traffic generated by a successful advertising campaign can result in missed opportunities, reduced revenues, and a tarnished brand image.

Because electronic transactions and communications take place so quickly, the amount of work and business lost in an hour far exceeds the toll of previous decades. According to a report published by Strategic Research Corporation, a Santa Barbara, California, market research and consulting firm, the financial impact of a major system outage can be enormous: US$6.5 million per hour in the case of a brokerage operation; US$2.6 million per hour for a credit-card sales authorization system; or a mere US$14,500 per hour in automated teller machine (ATM) fees if an ATM system is off-line.
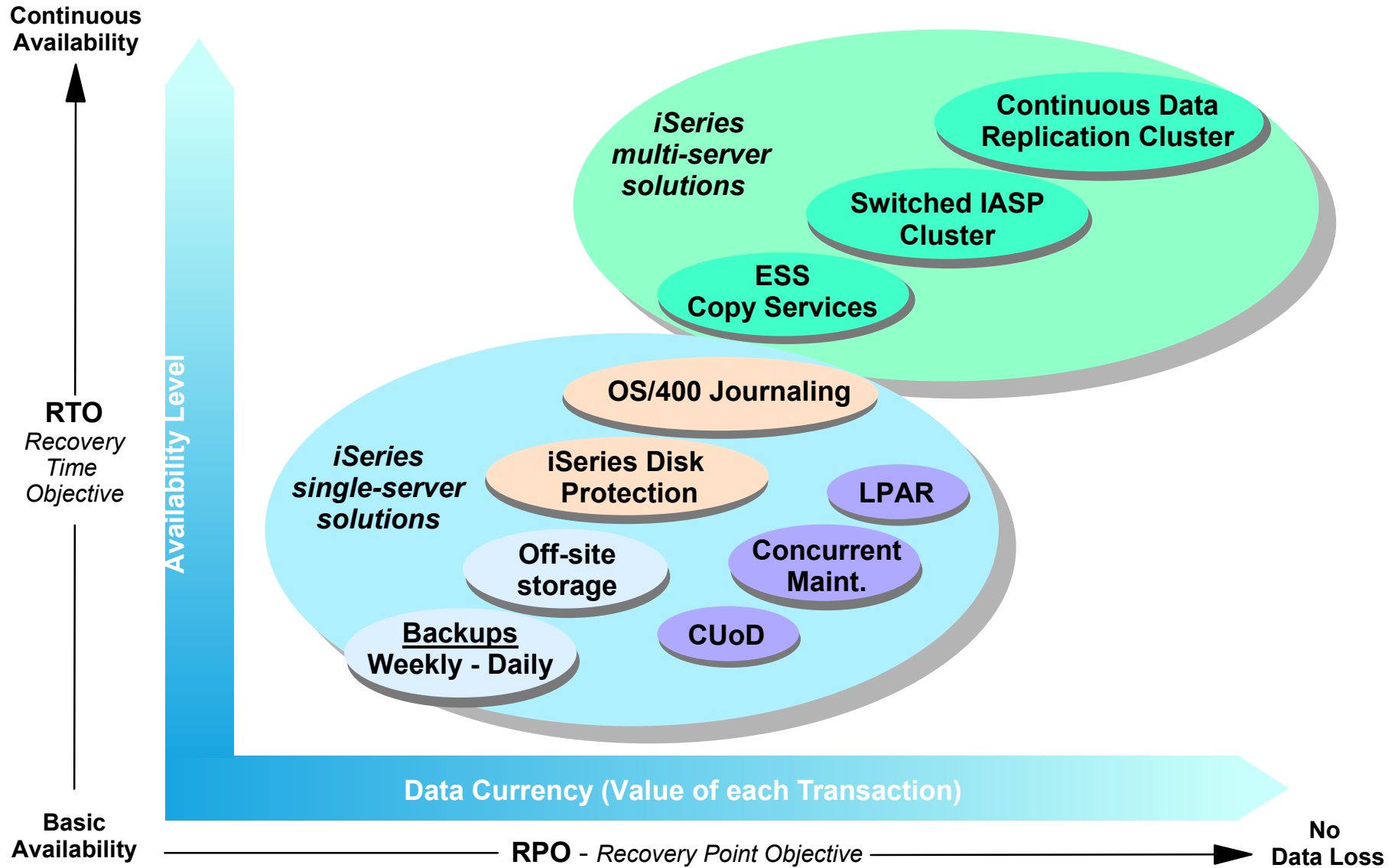
Even what was once considered a "minor" problem —a faulty hard drive or a software glitch—can cause the same level of loss as a power outage or a flooded data center if a critical business process is affected. For example, the New York-based research firm FIND/SVP calculates the average financial loss per hour of disk array downtime
at US$29,301 in the securities industry, US$26,761 for manufacturing, US$17,093 for banking and US$9,435 for transportation. More difficult to calculate are the intangible damages a company can suffer: lower morale and productivity, increased employee stress, delays in key project timelines, diverted resources, regulatory scrutiny and a tainted public image.

In this climate, executives responsible for company performance now find their personal reputations at risk. Routinely, companies that suffer online business disruptions for any reason make headlines the next day, with individuals singled out by the press. Moreover, corporate directors and officers can be liable for the consequences of business interruption or loss of business-critical information. Most large companies stipulate in their contracts that suppliers must deliver services or products under any circumstances. What's more, adequate protection of data may be required by law, particularly for a public company, financial institution, utility, health care organization or government agency.

Together, these factors make business continuity the shared responsibility of an organization's entire senior management, from the CEO to line-of-business executives in charge of crucial business processes. Although IT remains central to the business continuity formula, IT management alone cannot determine which processes are critical to the business and how much the company should pay to protect those resources.

From: Business Continuity: New risks, new imperatives and a new approach: www-1.ibm.com/services/continuity/recover1.nsf/files/Downloads/$file/buscont.pdf

iSeries strategies and technologies

Continuous Availability

RTO
Recovery Time Objective

Basic Availability

Availability Level

Continuous Data Replication Cluster

iSeries multi-server solutions

Switched IASP Cluster

ESS Copy Services

OS/400 Journaling

iSeries single-server solutions

iSeries Disk Protection

LPAR

Off-site storage

Concurrent Maint.

Backups Weekly - Daily

CUoD

Data Currency (Value of each Transaction)

RPO - *Recovery Point Objective*

No Data Loss

Outages

IT challenges in the business world - outage types
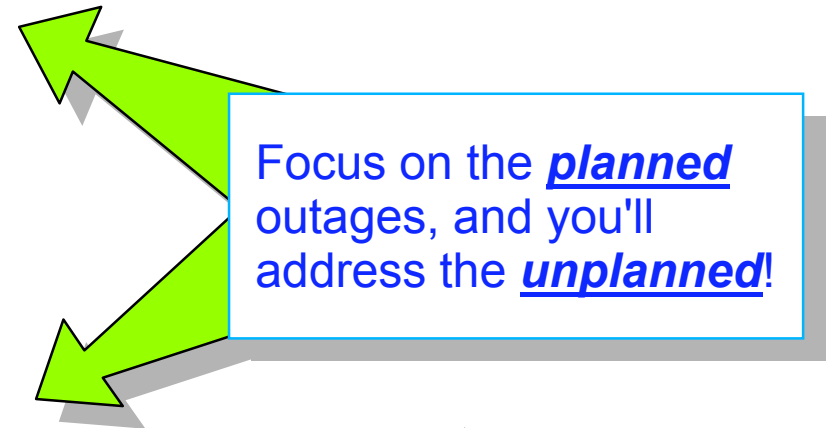
- **Planned**
  - ▶ Estimated to be >90% of all iSeries outages
    - – Backups (nightly, weekly, monthly saves)
    - – Server and application maintenance
    - – New software installation
    - – Upgrades (hardware or software)
    - – BI extracts/loads, batch jobs
    - – Building and utility maintenance

- **Unplanned**
  - ▶ Estimated to be <10% of all iSeries outages
    - – Power outage
    - – Human error or program failure
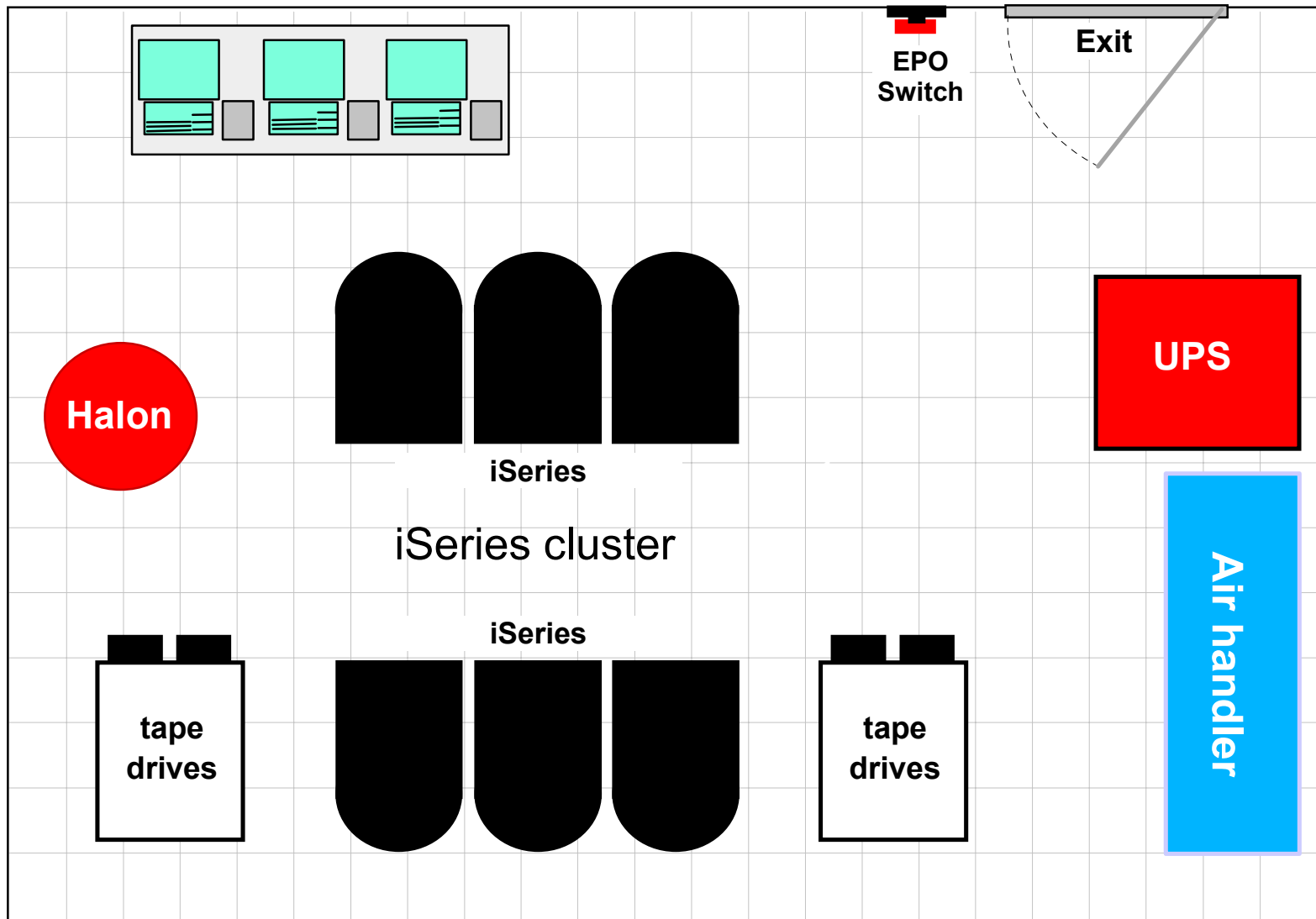   - – Data loss
    - – Hardware failure

- **Disasters**
  - ▶ Estimated to be <1% of all iSeries outages
    - – Site wide problem

Focus on the ***planned*** outages, and you'll address the ***unplanned***!

You may be able to cover ***all*** potential server outages by integrating local recovery techniques into your DR strategy
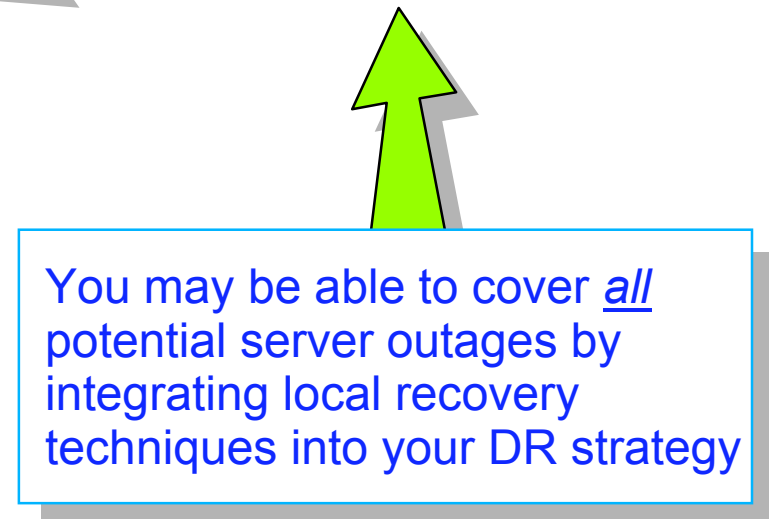
The odd outage

Or, what can go wrong will go wrong

iSeries technologies to avoid outages

Hardware Service

- Automatic Failure Notification

- Predictive Failure Analysis & Reporting

Power Subsystem

- Redundant power supplies
- Dual line cords
- Redundant cooling fans
- Dedicated UPS Monitoring Interface
- Concurrent maintenance

Memory

- "Chip kill" technology
- Error detection & correction
- Memory scrubbing
- Array repair at IPL using redundant array bits

Disk Subsystem

- RAID 5 protection
- Mirroring protection
  - Bus, IOP or disk level
- Concurrent maintenance
- Add disk concurrently
- Move data concurrently
- Predictive failure analysis & reporting

I/O Subsystem

- Hot pluggable PCI cards
- Dynamic hardware resource reallocation
- Redundant HSL loops
- IOP reset
- Virtual IP, route fault tolerance, load balancing

Notes: iSeries hardware technologies

	V5R1 and earlier	V5R2
Processor and memory	• ECC to correct single bit errors or chip kill • Memory scrub with redundant bit steering • Deconfigure failing component on IPL • Clustering	• Array repair at (re)IPL using redundant array bits (Array Gard) • Expanded ECC coverage (SP Flash, internal arrays on ASICs) • Retry of parity errors • Persistent boot time CPU/memory deconfiguration (Repeat Gard)
Power and service processor	• Redundant power supplies and cooling fans • UPS interface / battery backup • Generator • Dual utility power sources (dual power cords) • Clustering (with independent power sources)	• Fully redundant power path components
DASD	• RAID • Mirroring protection levels • Mirroring with manual hot spare • PFA & Service Director	
Other I/O	• Concurrent Maintenance • Redundant HSL links (loop) • Redundant I/O attachments, alternate paths • Virtual IP, Route Fault Tolerance, Load Balancing • Clustering	

Fault avoidance

   OS/400 and the iSeries attempt to correct problems before they cause an outage. Once example are the storage related system values QSTGLOWACN and QSTGLOWLMT. Used in combination, these system values help you prevent an outage caused by completely filling your DASD capacity. More information about these system values can be found in the information center.
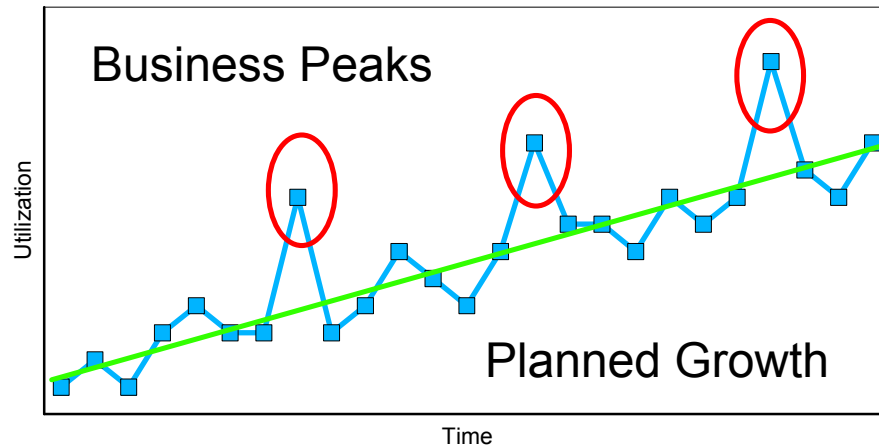
Service productivity via Electronic Service Agent

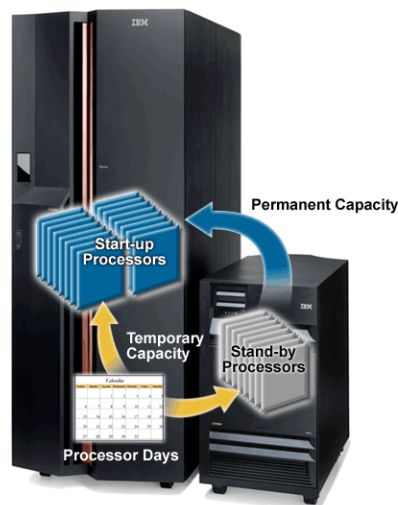   At this time, two functions are provided:

      Hardware problem reporting predicts and prevents hardware errors by early detection of potential problems, downloads fixes, and automatically calls IBM Service when necessary. **Note:** To activate the hardware problem reporting function of Electronic Service Agent requires that your server be under warranty, or that you purchase an IBM Maintenance Services Agreement.

      System inventory collection and transmission collects and electronically sends system information to IBM to be used for input for problem analysis and problem prevention functions, and to assist IBM in providing improved service.

Non-disruptive upgrades with Capacity Upgrade on Demand

Business Peaks

Utilization

Planned Growth

Time

Permanent Capacity

Start-up Processors

Temporary Capacity

Stand-by Processors
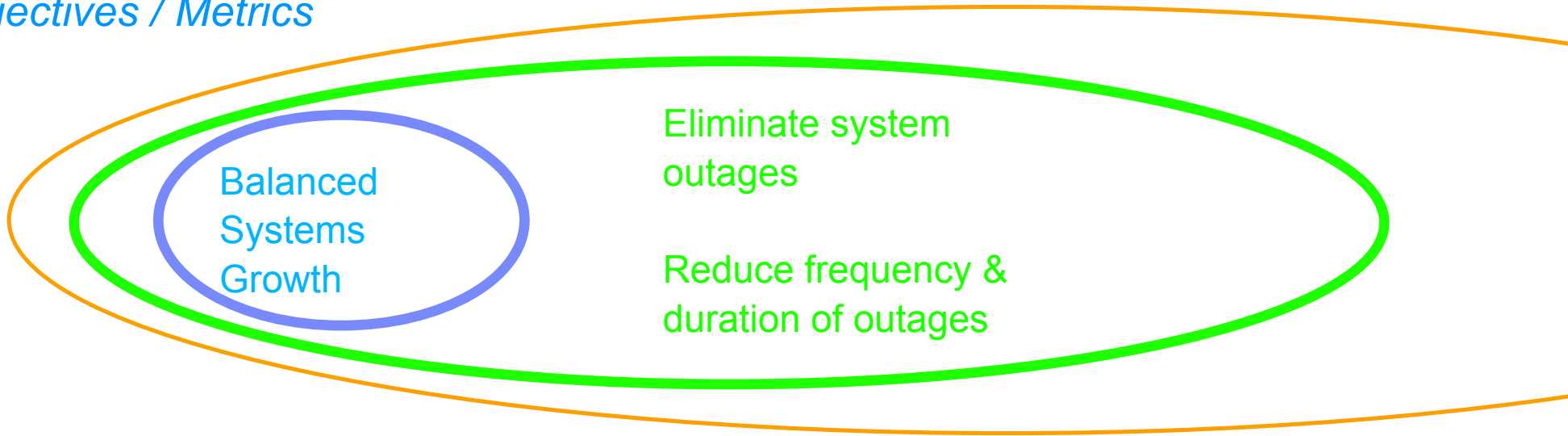
Calendar

Processor Days

- Adjust to unpredictable demand without interrupting operations or applications

- Permanent capacity
  - ► Planned growth
  - ► Pay when purchased

- Temporary capacity (NEW in 2003)
  - ► Business peaks
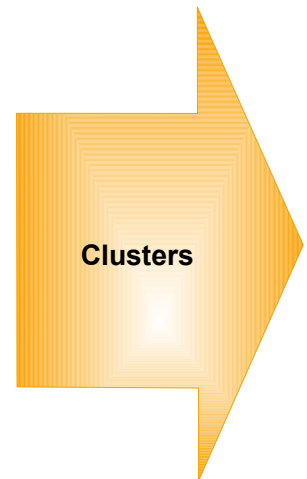 - ► Pay after activation
  - ► 14 no charge processor days

© 2003 IBM Corporation

See www.iseries.ibm.com/ondemand for more information.
Planning for Continuous Availability - 14

Scheduled downtime

Objectives / Metrics

Balanced Systems Growth

Eliminate system outages

Reduce frequency & duration of outages

Target Range for Single System

Save performance (backup window)	Save 25% of system in less than 4 hours	Save 25% of system in less than 2 hours
Save-While-Active	< 20 minutes interruption, no restrictions	No interruption, no restrictions
Normal IPL	< 30 minutes	< 15 minutes
Cumulative PTF install	< 60 minutes	< 30 minutes
Release install/upgrade	< 4 hours	< 2 hours
Dedicated system state	< 12 hours per year	< 4 hours per year
Hardware maintenance (scheduled)	< 8 hours per year	< 2 hours per year

Clusters

Availability and outages summary

- Analyze current risks

- Determine your availability objectives (RTO and RPO)
- Design your server to take advantage of
  - ► Hot spares
  - ► Concurrent maintenance
  - ► Electronic service agent
  - ► Capacity Upgrade on Demand
  - ► Dual power cords
- Use the latest techniques to minimize planned downtime
- Use iSeries clusters when availability requirements exceed a single server capabilities

Storage

Storage Solutions and Availability

The storage challenge

- **Storage redundancy**
  - ► Spare space
  - ► Duplication of data and objects

- **Operation complexities**
  - ► Multiple environments
  - ► Multiple storage technologies

- **Performance**
  - ► Arms, arms and more arms!

- **Increased availability**

- **RAID what?**

Occupied by Data objects

Occupied by Application objects

Unused Space
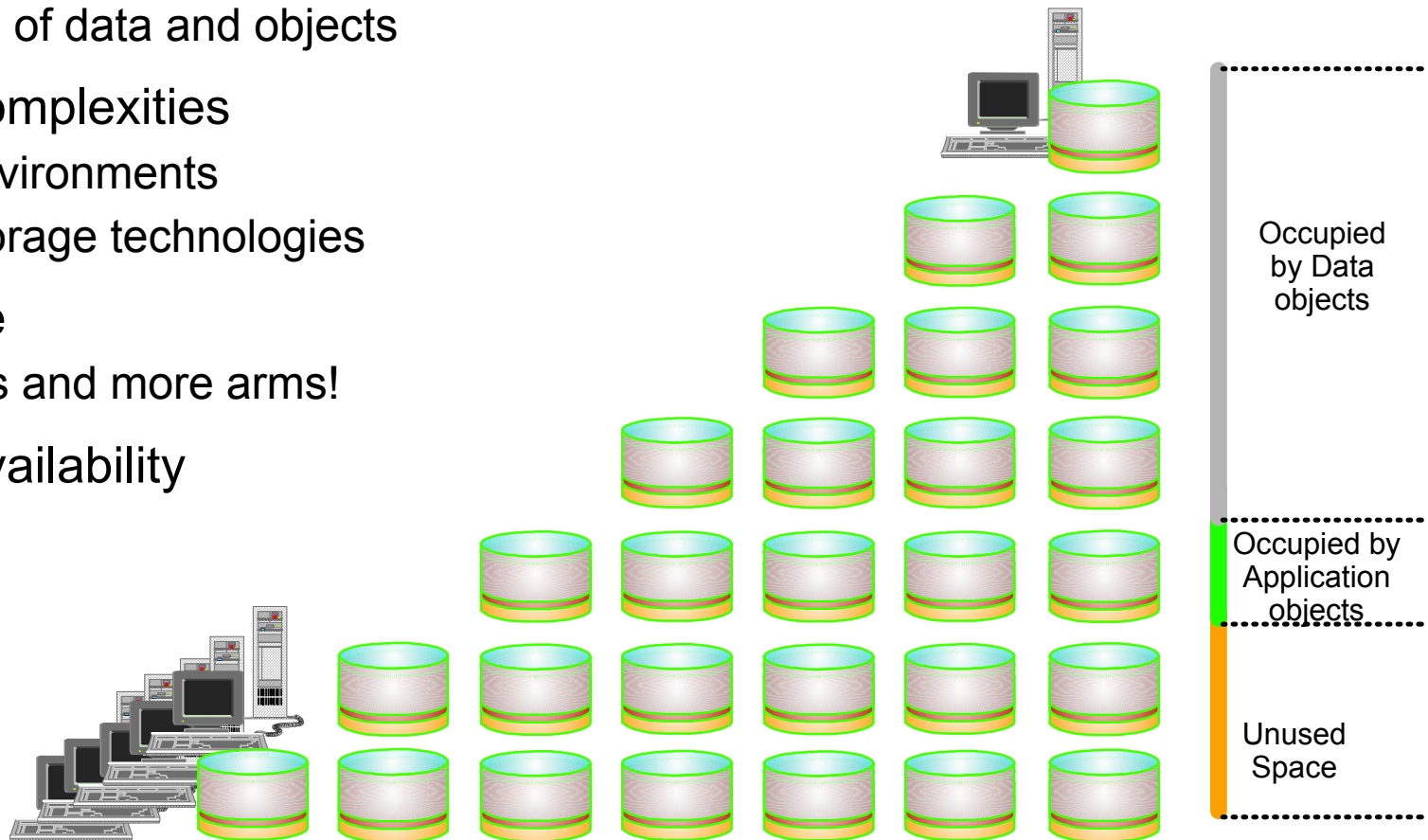
Storage Consolidation

IBM storage solutions for iSeries

- **Integrated iSeries storage**
  - ▶ Internal disk capacity to 144 TB
  - ▶ Solutions:
    - – Most cost effective - price performer
    - – Automated storage management
    - – Integrated, multi-workload storage consolidation
    - – Integrated save solutions

- **IBM ESS (Enterprise Storage Server)**
  - ▶ Fibre channel attachment
    - – 2 Gbps Fibre channel speed (V5R2)
  - ▶ Solutions:
    - – External, multi-server storage consolidation
    - – More multi server storage flexability
    - – Copy Services
      - • PPRC
      - • FlashCopy

Integrated iSeries Storage

HSL

Fibre channel

ESS - Enterprise Storage Server

ESS = Enterprise Storage Server (a.k.a. "Shark")

iSeries storage architecture

- **Autonomic**
  - ► OS/400 automatically spreads data across all disks - parallel I/O
  - ► Designed for Ease-of-Use Self-Management
    - − Added disk capacity is utilized automatically
    - − No continuous disk monitoring needed
  - ► Integrated Linux and Windows environments benefit too!
  - ► Integrated Disk Caching - Expert Cache
  - ► Applies to ESS storage too

- **Easy operations - single GUI management and wizards**

- **High Performance for all environments**

- **iSeries integrated storage availability**
  - ► RAID 5 or RAID 10 (RAID 51 !)
  - ► OS/400 save-while-active (SWA)

| Windows 2000 | UNIX | DB2 | Linux |

OS/400

CPU, L1, L2, L3 Cache, Main Storage and Expert Cache

OS/400 Automated Storage Management

Integrated I/O, SCSI, Fiber Channel

Integrated iSeries Storage or ESS
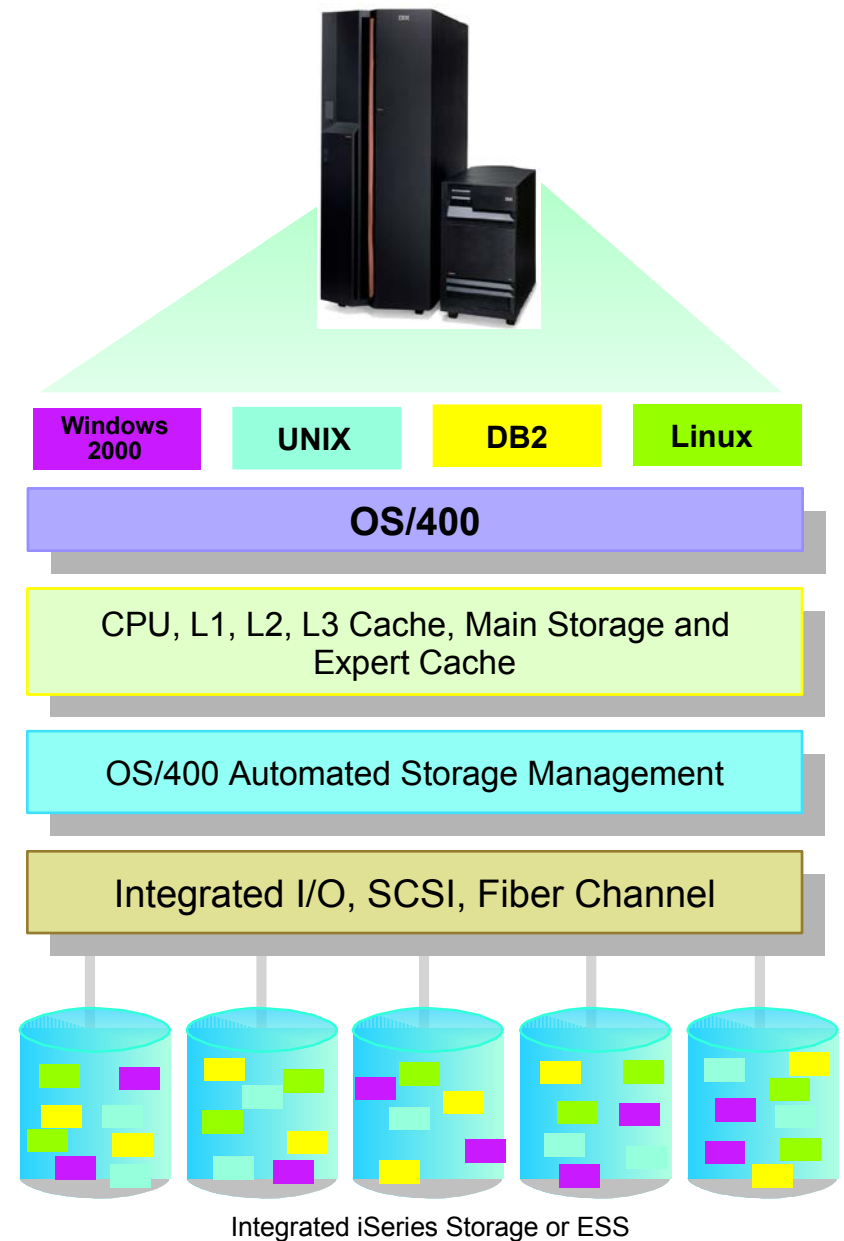
Notes: Logical view of iSeries storage

Autonomic
- OS/400 automatically spreads data across all disk units Parallel I/O results in optimum performance and disk utilization
- No user planning or management required - no manual data placement
- Automation drastically reduces operator involvement
- The operator never assigns disk storage, it's all transparent
  - ▸ No individual "disk full" conditions
  - ▸ No continuous disk performance monitoring
- Added disk capacity is automatically utilized
- Consolidated backups and tapes
- Integrated Linux and Windows environments benefit too!

This all applies to both integrated (disks) and external storage (LUNs)

iSeries storage
- Always treated as virtual storage
- Always multi-user, multi-application
- Machine Interface (MI) hides storage management complexities
- Non-disruptive growth through automated data striping and load balancing
  - ▸ Add and balance concurrent with production workload *and* ASP balancing commands
- Disk drives are essentially viewed as virtual memory store
- Integrated SAN like features built into OS/400
- Single view of storage to applications
- Hardware and software are independent
- Expert cache, parallel I/O
- Workload optimization
- Integrated save utilities perform object level saves to enable easy recovery

Auxiliary Storage Pools *(ASP)*

System ASP (ASP1)

- All systems ship with a system ASP

- Increased granularity and availability through user ASPs
  - ▶ Up to 31 user ASPs supported
  - ▶ Full range of object types supported

User ASP - Journal Receivers and Save Files

- New in Version 5: Independent ASPs
  - ▶ Enhanced granularity and availability
  - ▶ Up to 223 independent ASPs supported
  - ▶ Provides:
    - – Multiple DB2 databases
    - – Potentially reduced reclaim storage time
    - – Increased availability
      - • iASP failure isolation
      - • Switch iASP Clusters

Independent ASP - Company 1

Independent ASP - Company 2

Independent ASP - Payroll
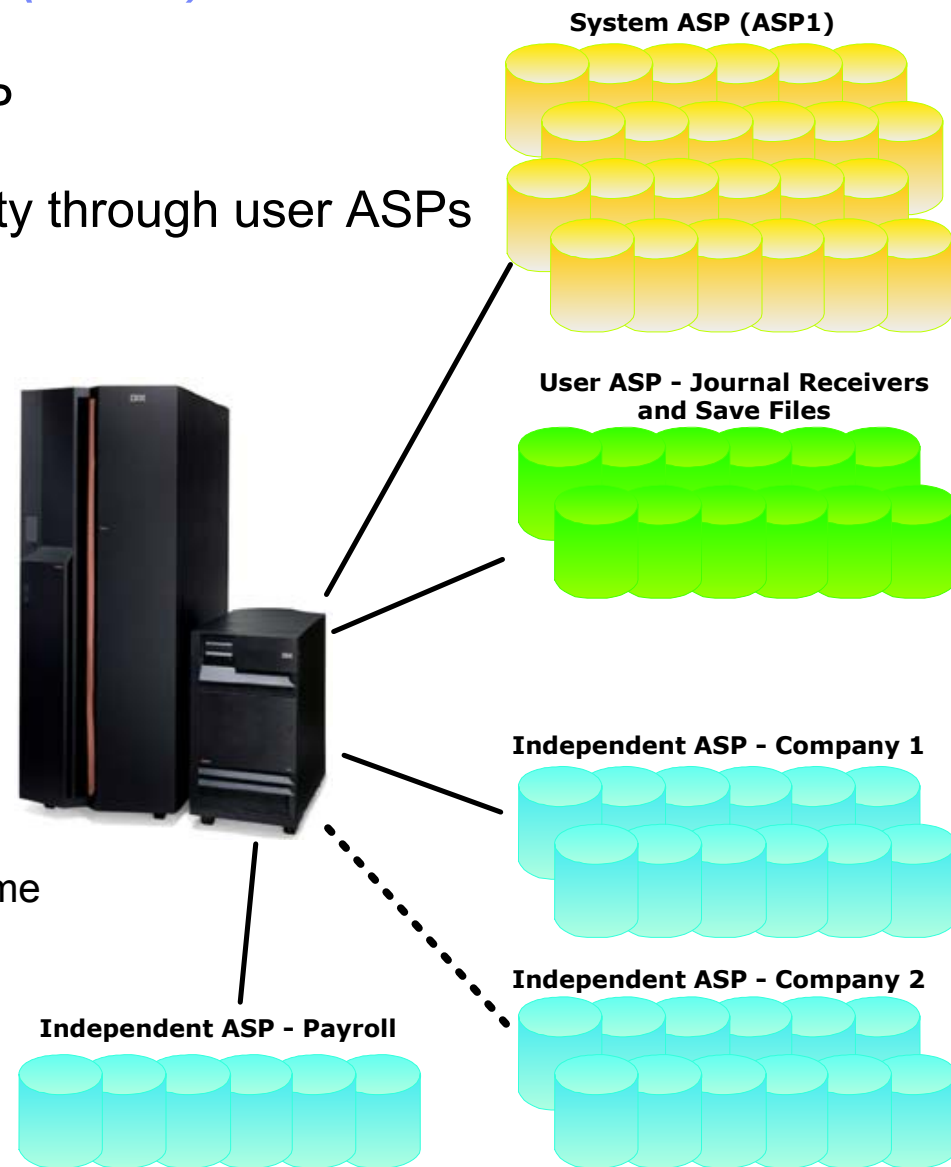
IBM eServer iSeries Independent ASPs:
A Guide to Moving Applications to iASPs
SG24-6802

Storage protection methods for iSeries

- **ESS SAN solutions for iSeries:**
  - ► Dual RISC cluster controllers with cache and Nonvolatile Storage
  - ► RAID-5 or RAID-10
  - ► RAID-5 + OS/400 disk mirroring (RAID-51 !)
  - ► Limited fiber path redundancy - no SDD available yet

- **iSeries integrated storage solutions:**
  - ► RAID-5
    - – Capacity, balance, or performance
  - ► Disk mirroring (e.g. RAID-10)
    - – Device, IOA adapter, or bus level
    - – IOA level recommended at a minimum

FC

SAN
Switches, Hubs, Routers, etc.

| Host Bay | **HBA** | | **HBA** | Host Bay |

Processor	**SSA Adp.**	Processor
Cache	**SSA Loop**	Cache
NVS		NVS

Enterprise Storage Server

OS/400 Storage Management

HSL bus   **HSL** bus

Processor   **Processor**

| Cache (SCSI) | **I/O Adapter** | | Cache (SCSI) | **I/O Adapter** |

Copy A   Copy B
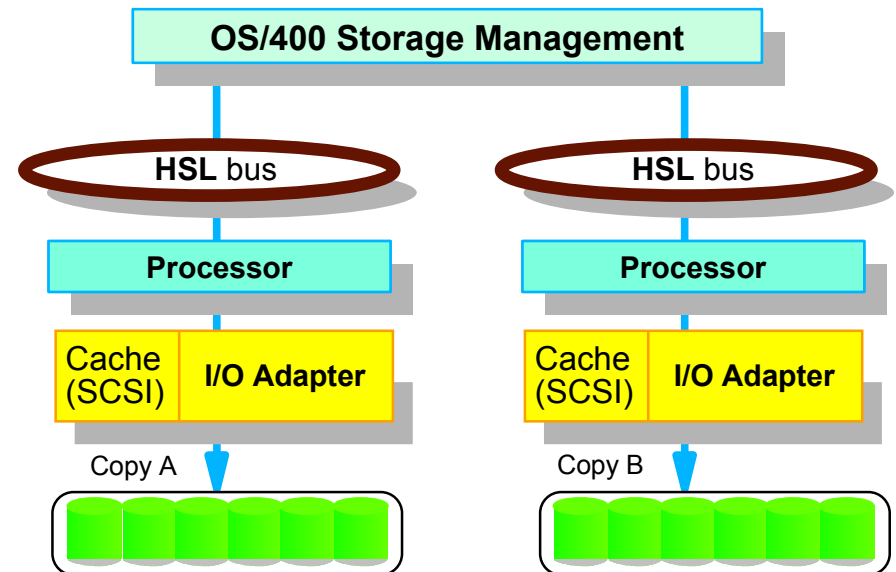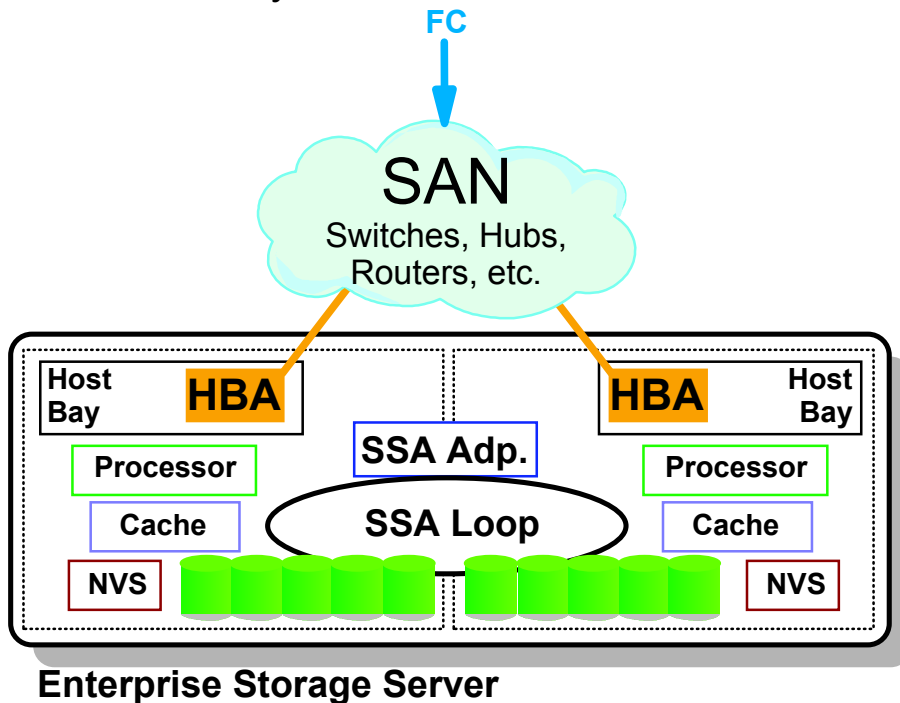
iSeries Disk Mirroring full bus redundancy example
(ESS or iSeries disks)

Remote OS/400 disk mirroring

- iSeries internal or ESS storage mirrored to second ESS

- Benefits
  - ► Full redundancy - I/O path and disk protection
  - ► Simple campus/local disaster recovery
  - ► No additional software required

- Considerations
  - ► Data loss - does not protect transactions or data
  - ► Unplanned fail-over can be many hours
  - ► Performance considerations
  - ► Unplanned fail-over can be many hours
    - – IPL required, plus any DB recovery
  - ► Not a choice for concurrent software and hardware upgrades

LSU

LSU

LSU

Remote ESS - iSeries
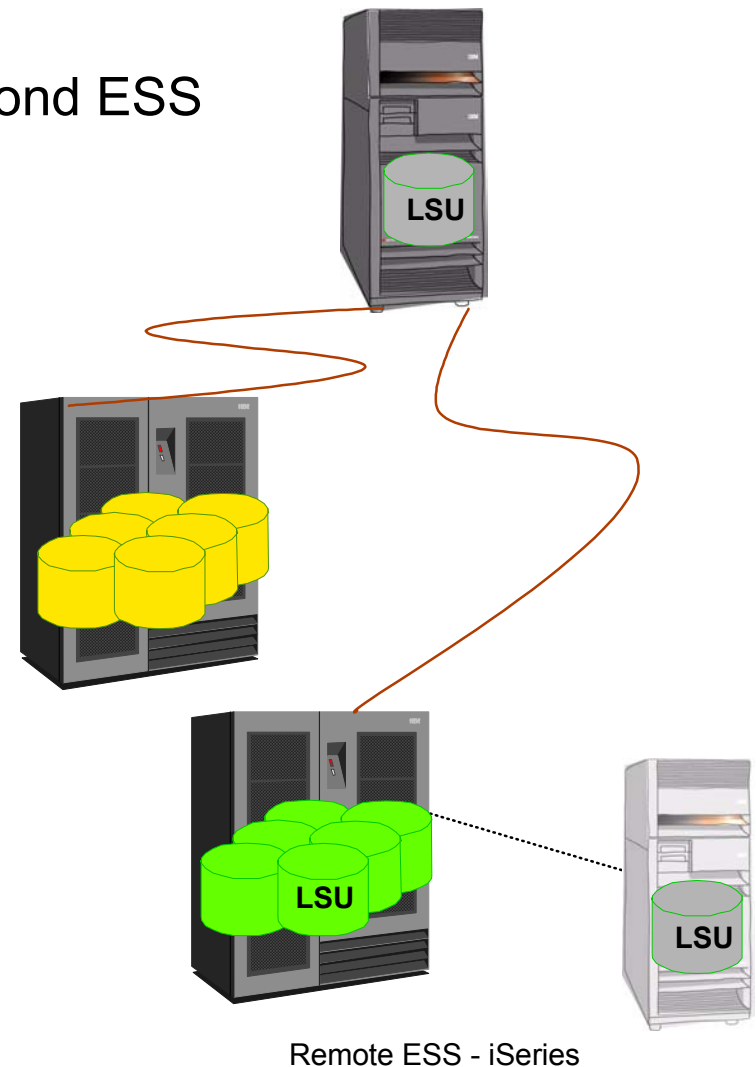
Before you begin sizing

Understand the iSeries workload

- iSeries will typically have a mix of workloads active at different times
  - ► Traditional database
  - ► File serving via Netserver or integrated xSeries
  - ► Webserver and byte stream file I/O
  - ► Domino
  - ► etc.

- Installed DASD configuration
  - ► All disk in the system ASP?
  - ► Disk allocated to user ASPs?
  - ► For journaling?
  - ► For save files?
  - ► Other reasons?
  - ► Disk allocated to independent ASPs?

- Capacity (GBs) requirements
  - ► Used GBs
  - ► Desired freespace
    - – Less than 75-80% is good
    - – 80-85%, are you investigating a purchase?
  - ► 85-90%, is the new DASD ready to be installed?

Considerations for sizing and performance planning

- Never size disk based on cost or capacity alone
  - ► Focus areas and considerations:
  - ► Workload related:
    - – I/O rates
    - – Transfer sizes
  - ► All hardware related
  - ► Number of disk units
  - ► Disk speed and size
  - ► Mixed sizes
  - ► ESS unique
    - – Number of LUNs and size
    - – Number of iSeries fiber channel adapters (#2766)
    - – ESS sharing
    - – Raid rank sharing

2766 FC Adapter

LSU

ESS Physical RAID Array

iSeries LUNs
(32 per FC)

LUN Sizes
8.589
17.548
35.165
36.003
70.564
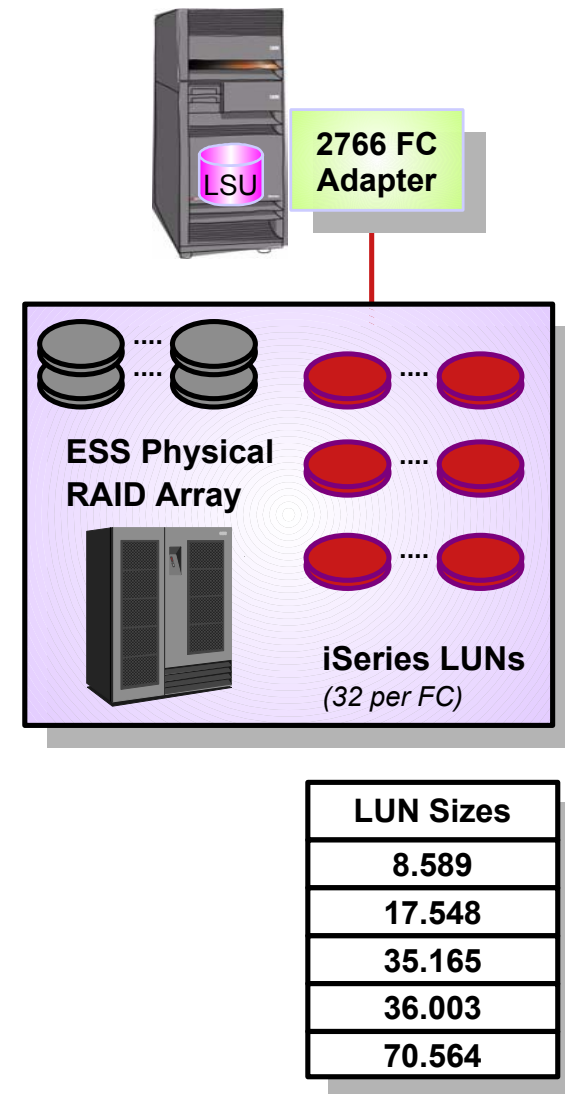
Storage summary

- Analyze storage requirements and server availability requirements separately
  - ► You have choices in iSeries storage
    - Use the storage solution that meets your business needs:
      - Capacity
      - Performance
      - Flexibility
    - When to use integrated iSeries storage
      - Server requires best performance, lowest cost, best integration
    - When to use SAN storage
      - Requirement to move disk between systems or platforms on a frequent basis
  - ► iSeries offers many choices in availability solutions
    - Understand your availability requirements
      - Which solution meets your business RTO and RPO?
      - What outage types that must be covered?
        - ◆ Disaster recovery?
        - ◆ Planned outages?
        - ◆ Unplanned outages?
    - Deploy the storage and iSeries availability technologies to meet your availability goals

- Always use disk protection (RAID-5 or OS/400 Disk Mirroring)

Systems management and consolidation

Systems management and consolidation challenges

- Adequate staffing and skills

- Managing multiple vendor/platform environments

- Performance and capacity management

- Automated operations

Productivity *	iSeries Solution	Standard Intel Architecture Solution	UNIX-based Solution
Users per server	**375.1**	112.8	200.7
Servers per IT Staffer	**3.5**	1.3	2.2
Server operating life (yrs)	**8.0**	4.5	6.6

iSeries supports more users per server than other solutions. iSeries requires only two thirds of the support staff of Unix solutions. iSeries requires only 40% of the support staff of Standard Intel Architecture Solutions

* Source "Server Cost of Ownership in ERM Customer Sites: A Total Cost of Ownership (TCO) Study" IDC September 2001   ** IBM Server Group Market Intelligence 2Q 00

iSeries operational environment and management

- Electronic support services

- iSeries Navigator

- Operational issues
  - IPLs
  - PTFs
  - Avoiding limits
  - DASD management

- Software upgrades

- Integrated xSeries management

- Linux partition management
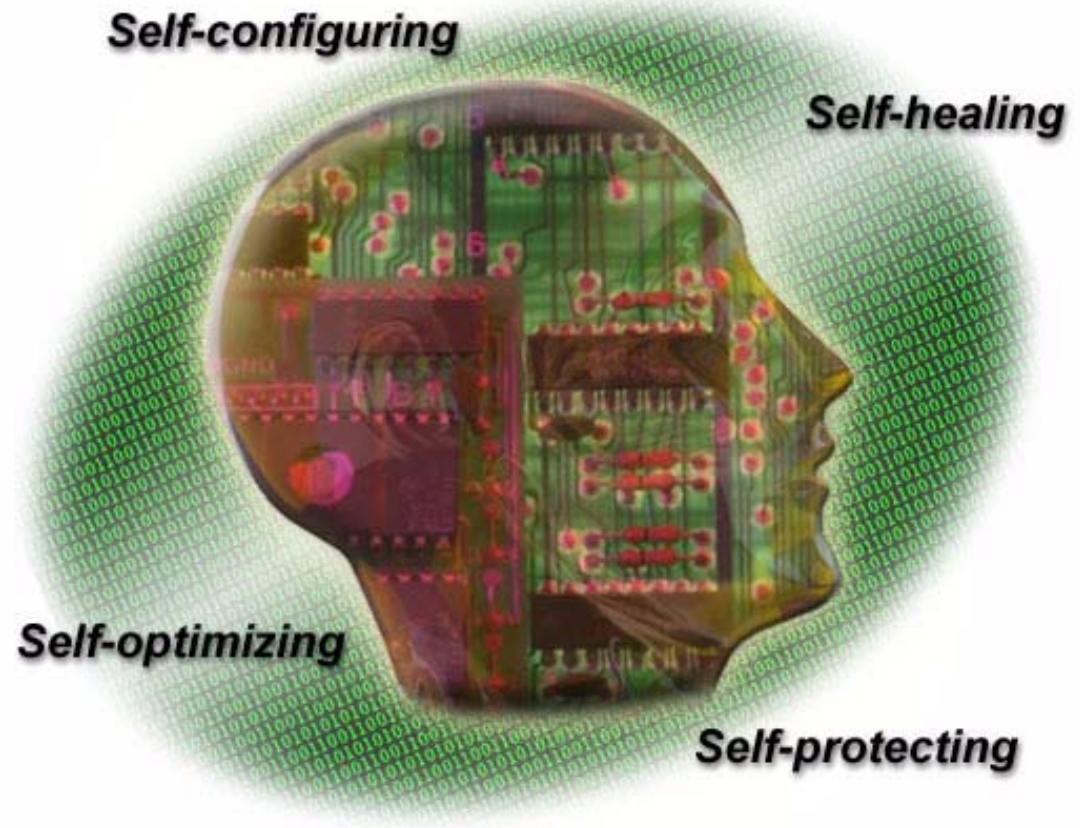
Dilbert "problem management"
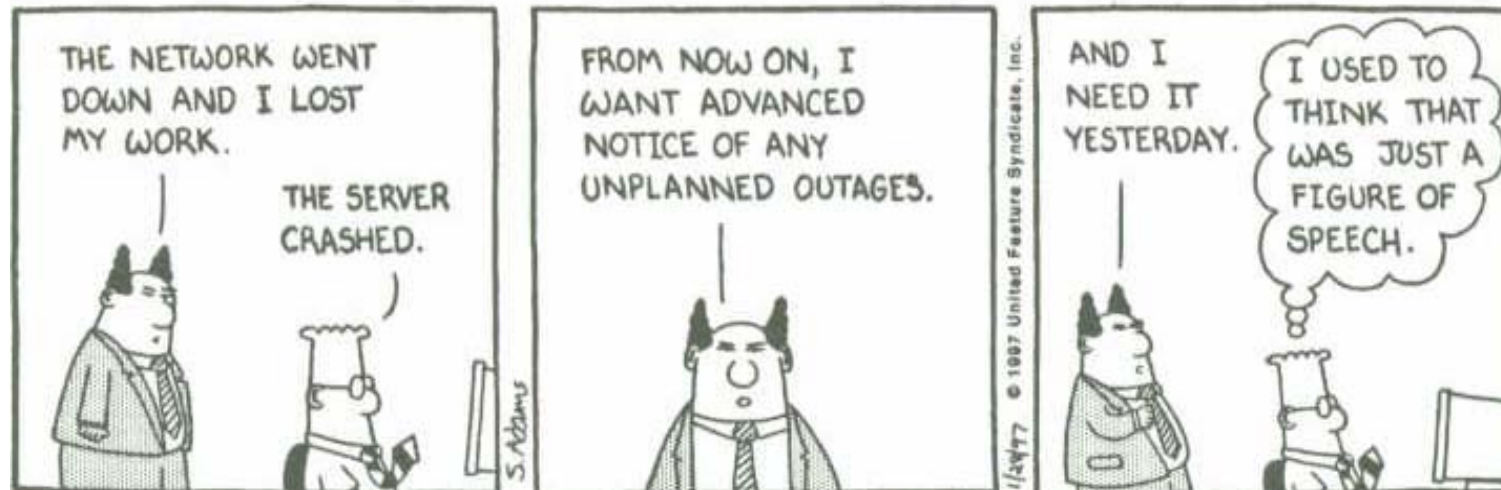
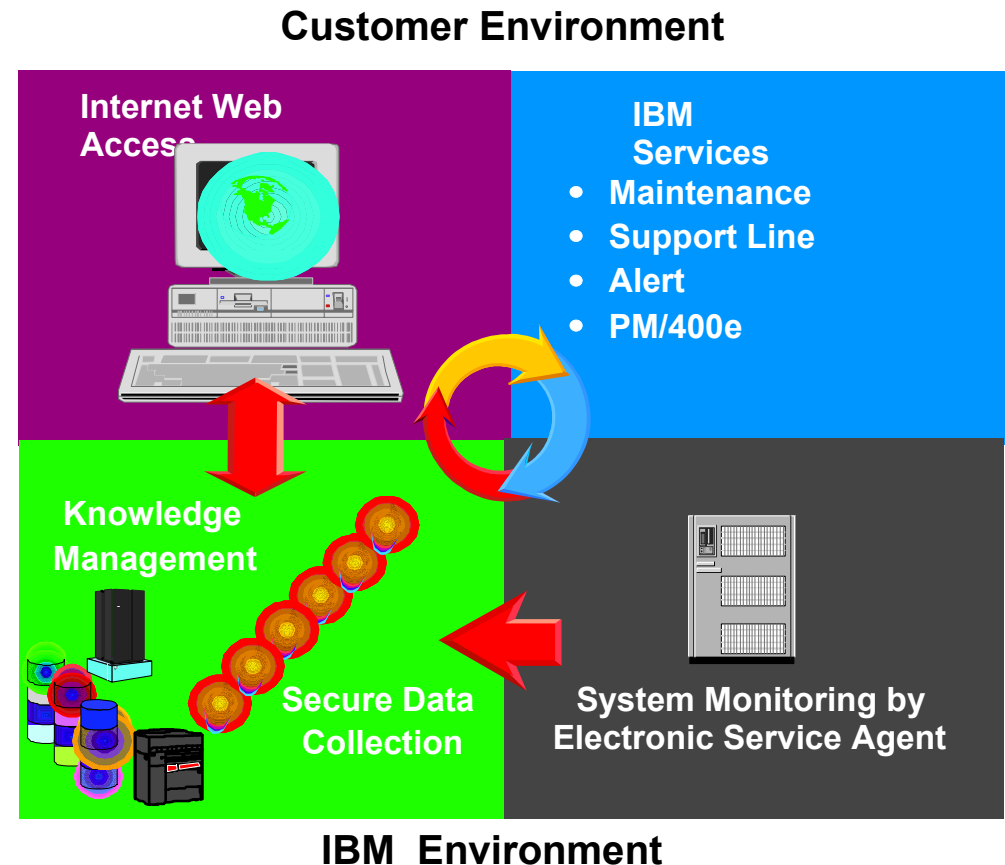# Systems management architecture and process components

Central box: **Availability Management**

Surrounding boxes:
- Change & Problem Management
- Performance Management
- Configuration Management
- Security Management
- Service Level Management
- Recovery Management
- Operations Management
- Event & Incident Management
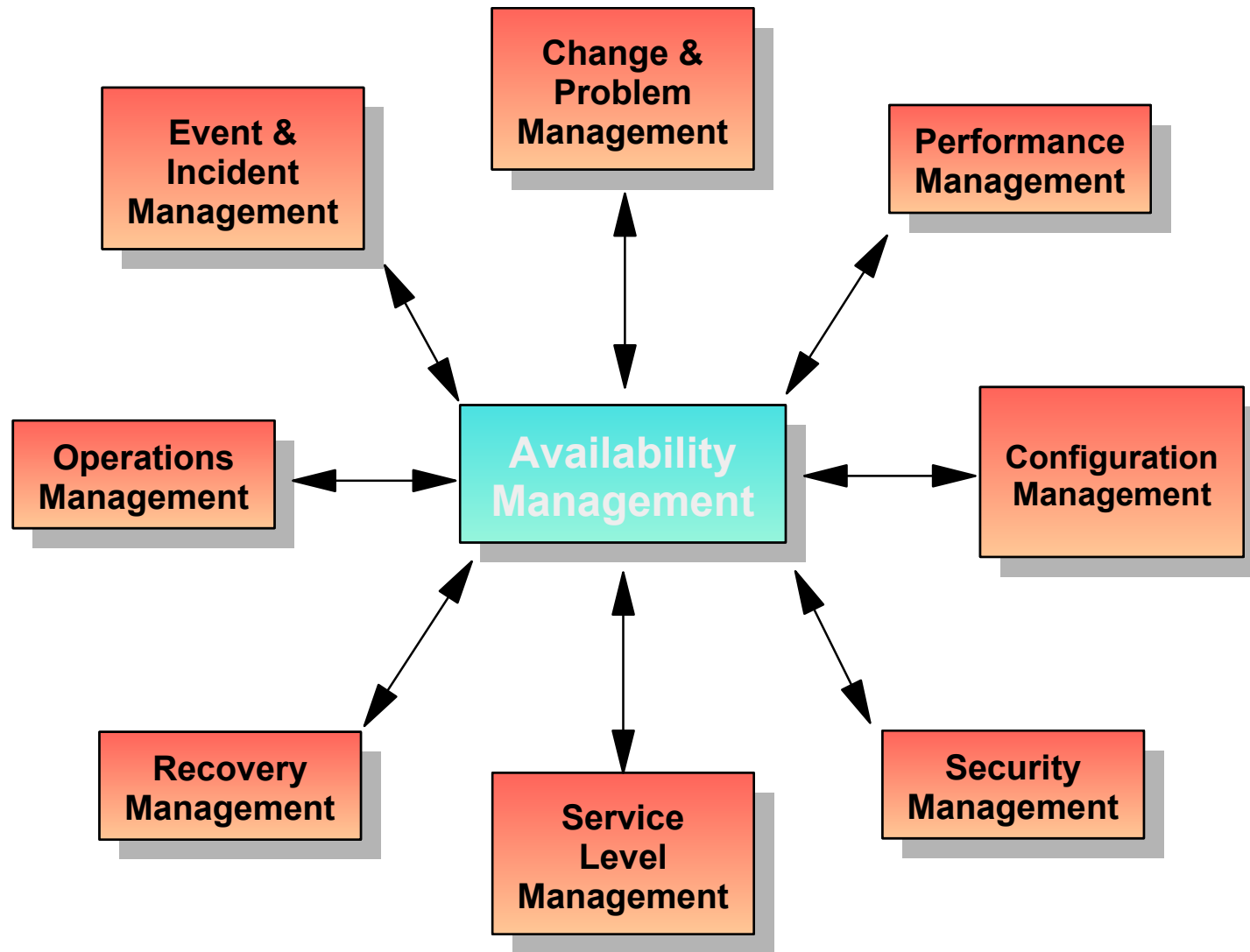
Critical skill sets for administrators

- Database
  - ► Performance analysis
  - ► Security administration

- Operations
  - ► Network
  - ► Performance analysis

- Security
  - ► Database
  - ► Performance analysis

- Network
  - ► Security
  - ► Integration between servers
  - ► Protocols

Electronic technical support and services

- **Technical support portal**
  - ▶ Sales and support tools and information
  - ▶ Personalized web access to technical information
  - ▶ Order software enhancements

- **Remote service and support capability**
  - ▶ Capacity monitoring and reporting
  - ▶ Fixes before failure
  - ▶ Remote connection for problem determination

- **Learning services**
  - ▶ World-class education and training
  - ▶ Flexible delivery formats
  - ▶ Skills development series

Customer Environment

Internet Web Access

IBM Services
- **Maintenance**
- **Support Line**
- **Alert**
- **PM/400e**

Knowledge Management

Secure Data Collection

System Monitoring by Electronic Service Agent

IBM Environment

IBM Electronic Service Agent

a cornerstone for Technical Support and Services

Fully Automatic

Sends error symptoms to IBM

Service agent detects a hardware problem

Data Catcher

Problem Database

Sends Symptoms for Diagnosis

Electronic Response

Dispatch CE

Voice Support

Analysis of Problem Record
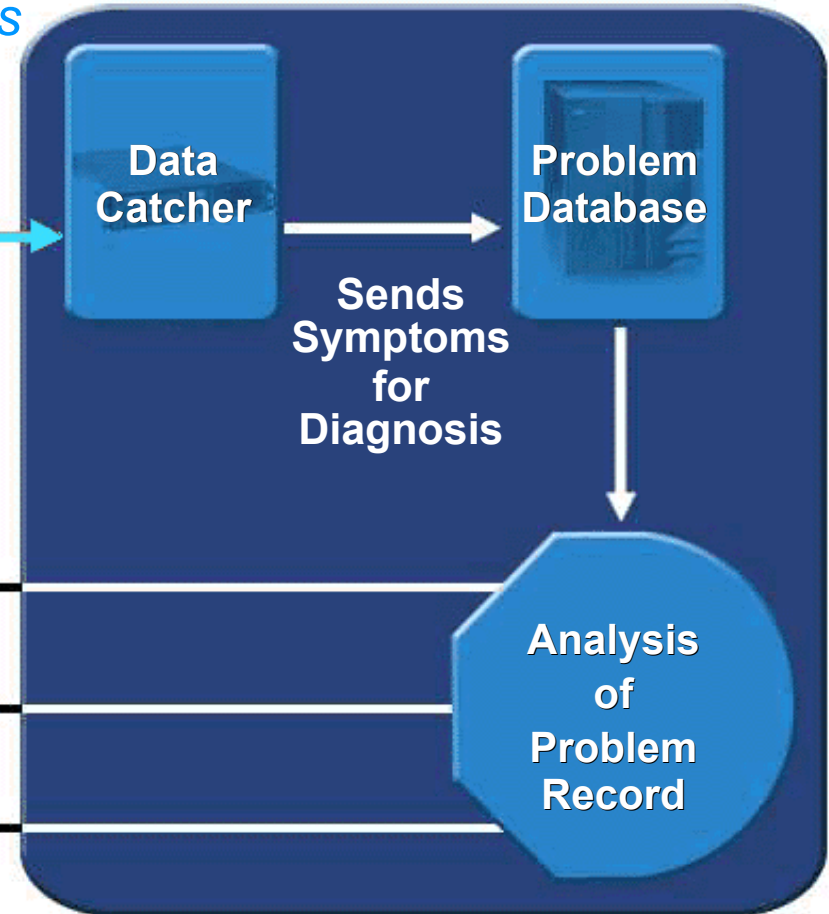
- *Faster problem resolution*
- *Higher availability/resiliency*
- *Lower maintenance cost*

iSeries Navigator

Our goal: Automated systems management

Multiple navigation paths replaced with a single Interface

Operational issues

- When to IPL?
  - ► IPLs that result from normal, scheduled maintenance are usually adequate
  - ► Running out of addresses
    - Address space on RISC machines (V3R6+) is extremely large
    - 18.4 quintillion bytes!
    - This should not occur during your career
  - ► Reduce job and system shutdown with ENDSBSOPT(*NOJOBLOG) parameter
    - Eliminates production of joblogs
    - Up to 30% less time to end subsystems and jobs

- Not all hardware installations require an IPL

- Full hardware testing requires an IPL

- Abnormal system termination
  - ► Resulting IPL is faster with V4 and above
  - ► Use access path protection

Operational issues (continued)

- **Applying PTFs**
  - ► IPL not needed for all types of PTFs
  - ► Other than HIPERs, applying quarterly cumulative and group packages may be sufficient for IPL requirements

- **Manage on single or multiple servers via iSeries Navigator**
  - ► Collect PTF
  - ► Send and install
  - ► Copy from media
  - ► Search inventory
  - ► Compare and update
  - ► Cleanup
  - ► Uninstall
  - ► Cancel actions

Self-configuring disk management

- **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**

Avoid *nasty* DASD full conditions

Disk migration while active

- **Concurrent data migration**
  - ► Enabled through start ASP balance (STRASPBAL) command
  - ► Majority of 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
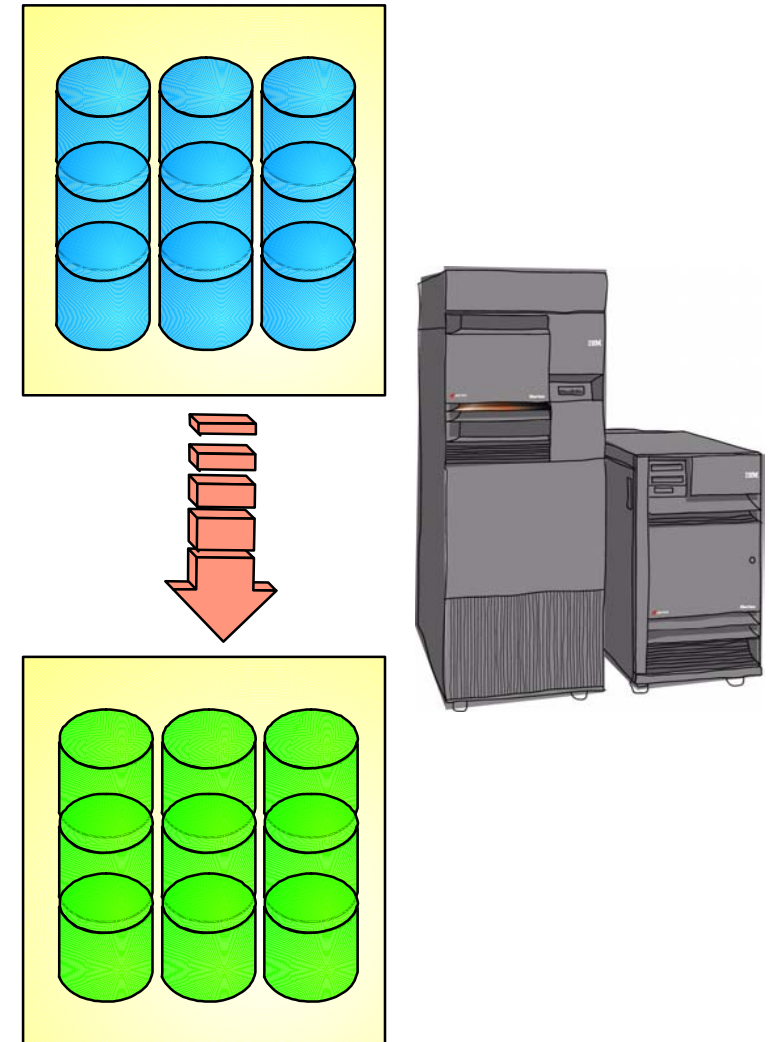
Virtual media for software upgrades and maintenance

- **Virtual media install**
  - ► Automate OS/400 and LP release load through virtual media
    - – Isolate media errors
    - – Reduce user intervention
    - – Improve system availability
  - ► Enabled with OS/400 V5R2
    - – Extended to support V5R1 through PTFs
  - ► Easy configuration steps
    - – Create a virtual image catalog device
    - – Build image catalog to transfer IBM software from distribution media to a user defined IFS directory
    - – Restart system to load from virtual device
    - – Point licensed program installation or PWRDWNSYS command to IFS path
  - ► PTF groups and cumulative packages, too!

New with V5R2

QIBM.... 1
QIBM.....2
QIBM.....3
..
QIBM....10

Before
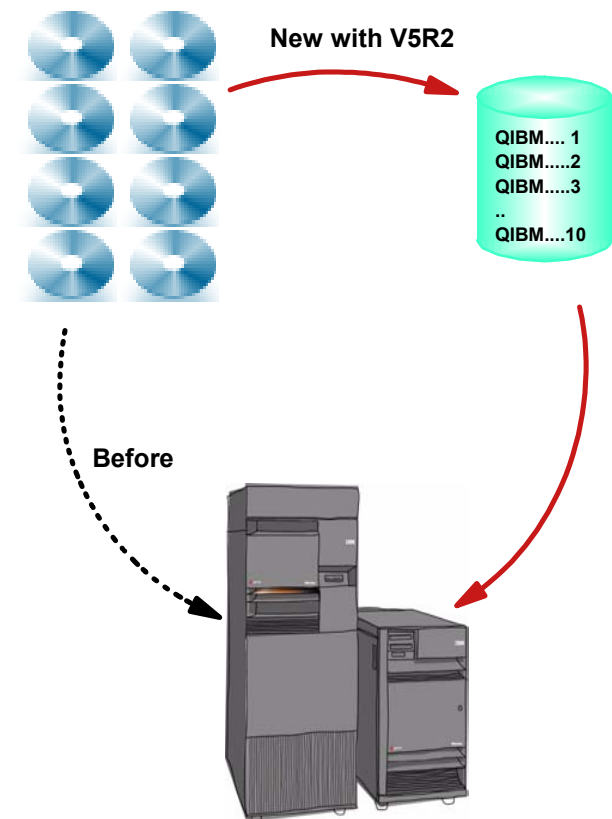
Know your limits

	V5R2	V5R1
Table size	1 TB	1 TB
Index (access path size)	1 TB	1 TB
Rows in a table	4.2 billion	4.2 billion
Number of journal entries	9.9 billion	9.9 billion
Journal receiver size	1 TB	1 TB
Number of jobs on the system	485,000	485,000
Number of jobs in a subsystem	32,767	32,767
Temporary storage per job - *NOMAX or	2 TB	2 TB
Entries per user profile	10 million	5 million

The System Administrator's Companion to AS/400 Availability and Recovery - SG24-2161

Notes: Know your limits

V5R2 maximum capabilities are documented at:
   http://www.redbooks.ibm.com/redpapers/pdfs/redp0204.pdf

V5R1 and earlier releases maximum capabilities documentation can be found in the iSeries Support Software Knowledge Base by searching for the phrase "maximum capabilities" or these specific document numbers:
   V5R1 - 23931042
   V4R5 - 19690531
   V4R4 - 15704052
   V4R3 - no updates
   V4R2 - 12761736

The System Administrator's Companion to AS/400 Availability and Recovery - SG24-2161

Enhanced Windows server operations

► **Greater flexibility for server and disk utilization**

► **Improved backup & recovery**

► **Manage multiple applications and services from a single iSeries**

► **Lower the cost of user administration**

► **Provide flexible Storage Area Network like management for Windows servers**

► **Improve Windows server up time and stability**

Systems management and consolidation summary

- **Investigate single server availability options**
  - ► Use the latest techniques to minimize planned downtime
  - ► Keep current with high impact PTFs
  - ► Take advantage of Electronic Support Services

- **Server consolidation for increased availability**

- **Use iSeries clusters when availability requirements exceed a single server capabilities**

Live Ethernet adapter

Primary IXS

system ASP

user ASP

IXS hot spare

Standby Ethernet adapter
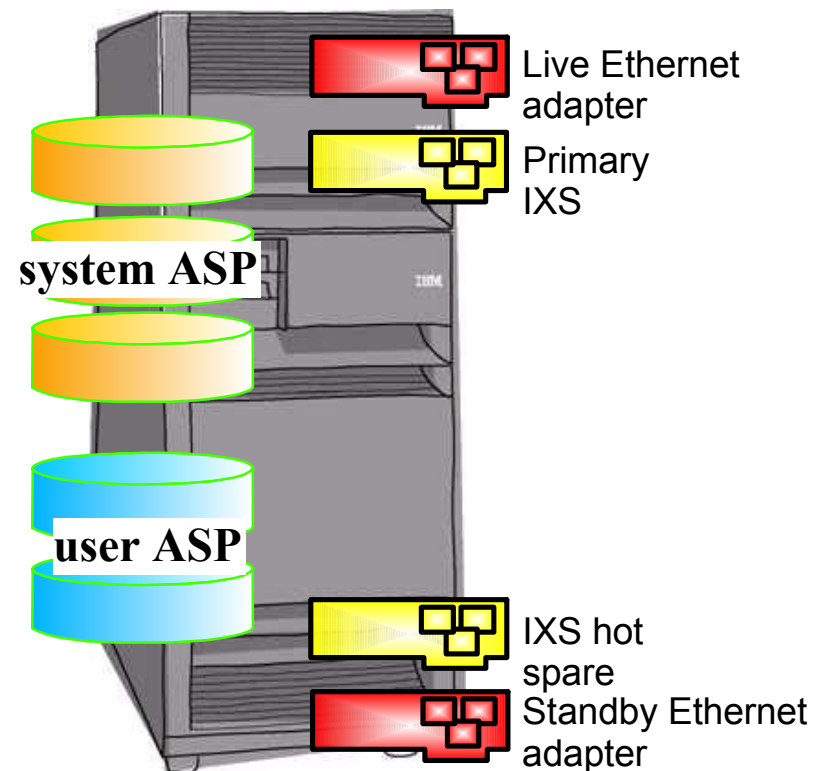
Backup and recovery

Backup and recovery challenges

- The story of the incredibly shrinking window
  - ► Backup
  - ► Recovery Time Objective (RTO)
  - ► Recovery Point Objective (RPO)

- Balancing
  - ► Complexity
  - ► Frequency
  - ► Duration

Backup and recovery strategies and tips

- Understanding save/restore performance

- How to make it go faster

- Minimizing save window

- Achieving recovery point objectives

Overview of factors influencing save performance
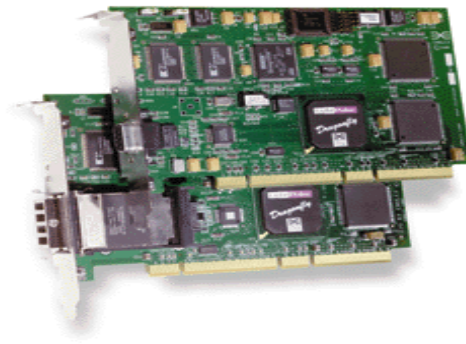
Tape Type + Interface Type + Connection Speed +
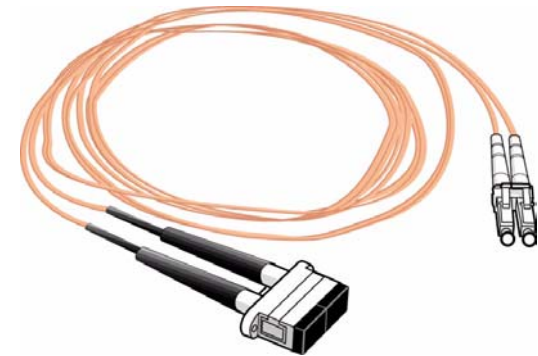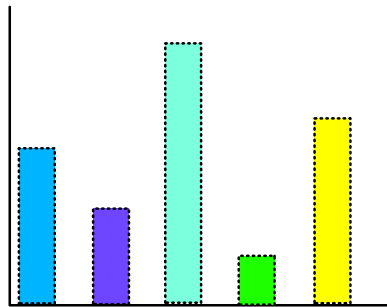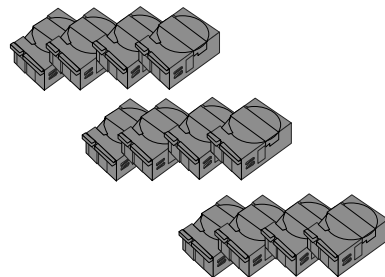
Save Workload Type + Storage Configuration and Performance + CPU Type and Number = GB per hour saved or restored
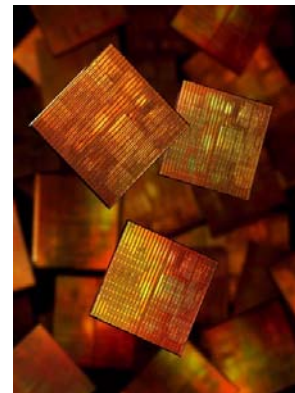
Your Save Performance

Understanding tape drive and IOP capabilities

Tape drive capabilities

VXA-2

► **6 MB/s**

358x

► **15 MB/s**

358x

► **Fibre - 35 MB/s**

3590

► **B-Ultra - 9 MB/s**

► **E/H - 14 MB/s**

IOP capabilities

SPD technology

► **FC 6501 - 17 MB/s**

► **FC 6534 - 17 MB/s**

PCI technology

► **FC 2729 - 13 MB/s**

HSL PCI technology

► **FC 2749 - 38 MB/s**

► **FC 2765 - 88 MB/s**

► **FC 5702 - 73 MB/s**

Notes: Understanding save performance

Tape drive capabilities

	VXA-2	LTO 1	LTO 2	3590
Data rates	− 6 MB/s native, 18 MB/s compressed	− 15 MB/s native, 45 MB/s compressed	− 35 MB/s native, 105 MB/s compressed	− B Ultra-SCSI - 9 MB/s native, 27 MB/s compressed − E/H - 14 MB/s native, 42 MB/s compressed
Announced	− May, 2003	− August, 2000	− January, 2003	− B - April, 1995 − B Ultra - January, 1999 − E - April, 1999 − H - June, 2002
Capacities	− 21 GB (60 meter) − 59 GB (170 meter) − 80 GB (230 meter)	− 100 GB − Capable of 3:1 compression	− 200 GB (purple) − Capable of 3:1 compression	− High Performance ['J' (Blue)] ● B - 10 GB ● E - 20 GB ● H - 30 GB − Extended High Performance ['K' (Green)] ● B - 20 GB ● E - 40 GB ● H - 60 GB − Capable of 3:1 compression

IOP/IOA capabilities

► SPD and PCI IOPs are available on 7xx and earlier technologies and in migration towers for 8xx technology.

► HSL PCI IOPs are only available on 8xx and later technology.

► There are two types of SCSI interfaces
  − Feature code 5702 is LVD SCSI (ultra SCSI 4) with sustained data rates of up to 73 MB/sec (future performance enhancements possible).
  − Feature code 2749 is HVD SCSI (ultra SCSI 2) with rates of up to 40 MB/sec possible.

Influence of workload

Source files
3 GB library
12 GB library
4 GB file
8 GB file
16 GB file
1 Dir / Many Obj
Many Dir / Many Obj
Domino
Network storage
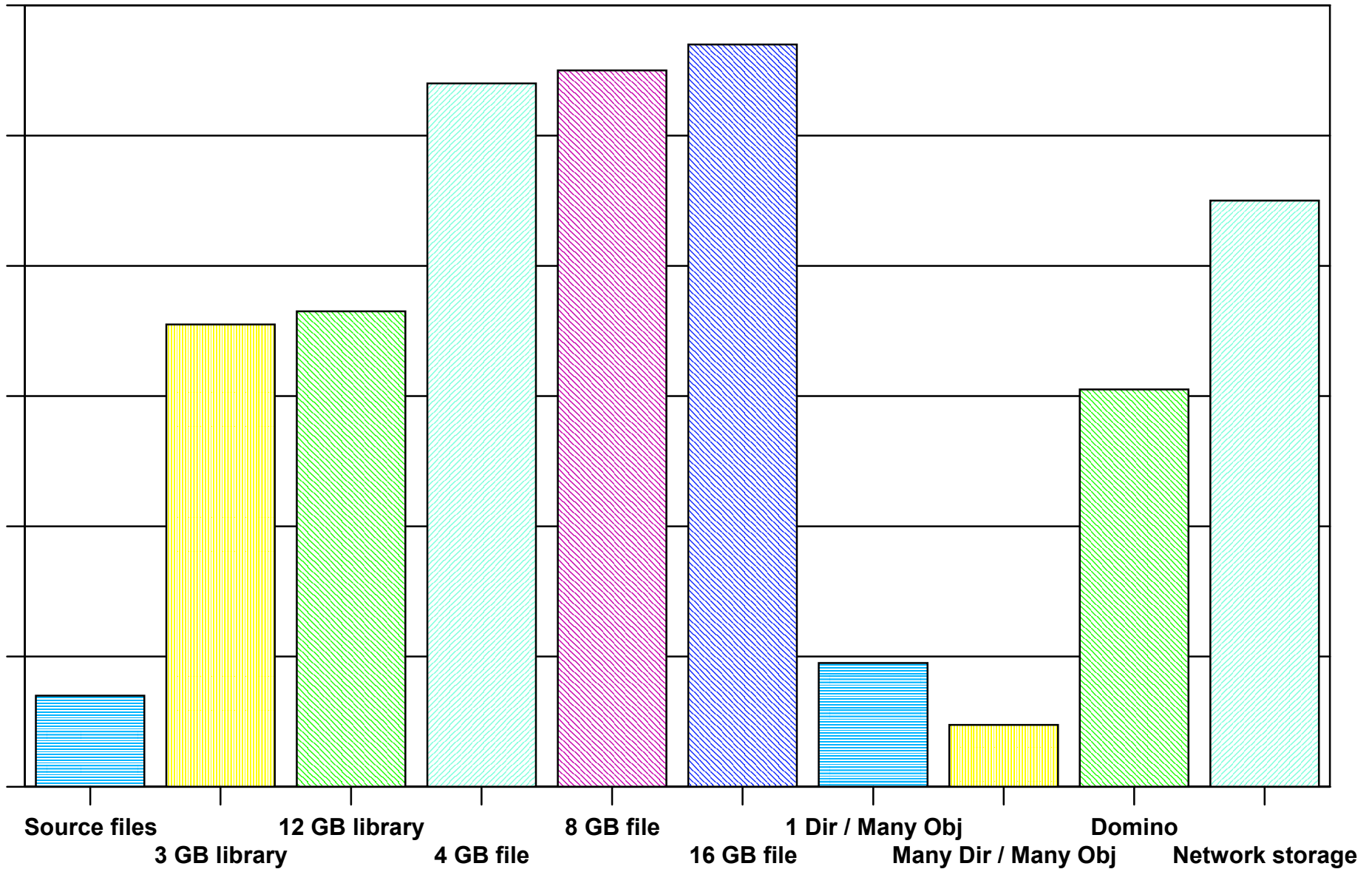
Quantity of DASD arms impact

| 15 arms | 30 arms | 45 arms | 60 arms | 75 arms |

■ 16 GB file ■ 6 GB library
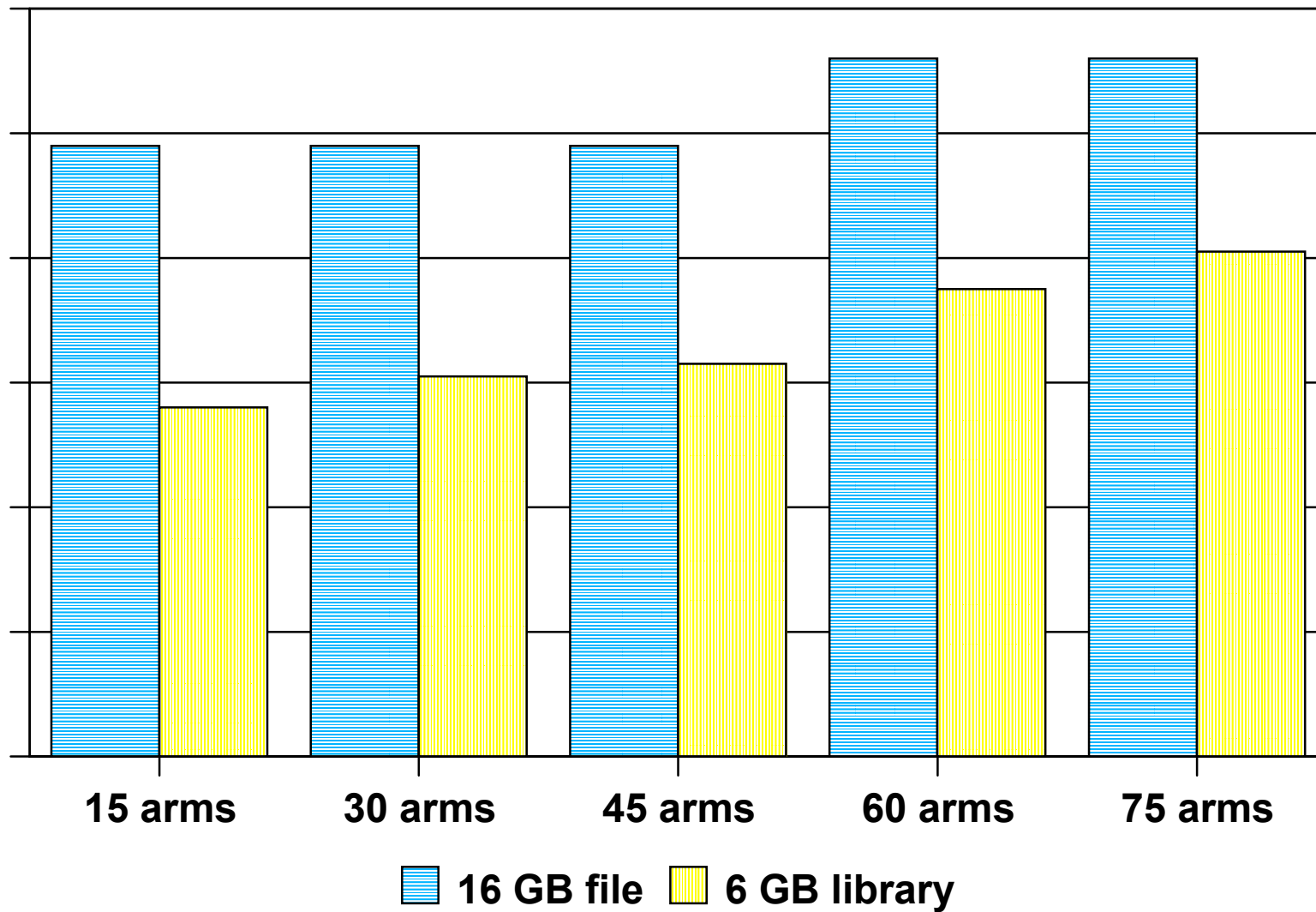
Bus and CPU influence
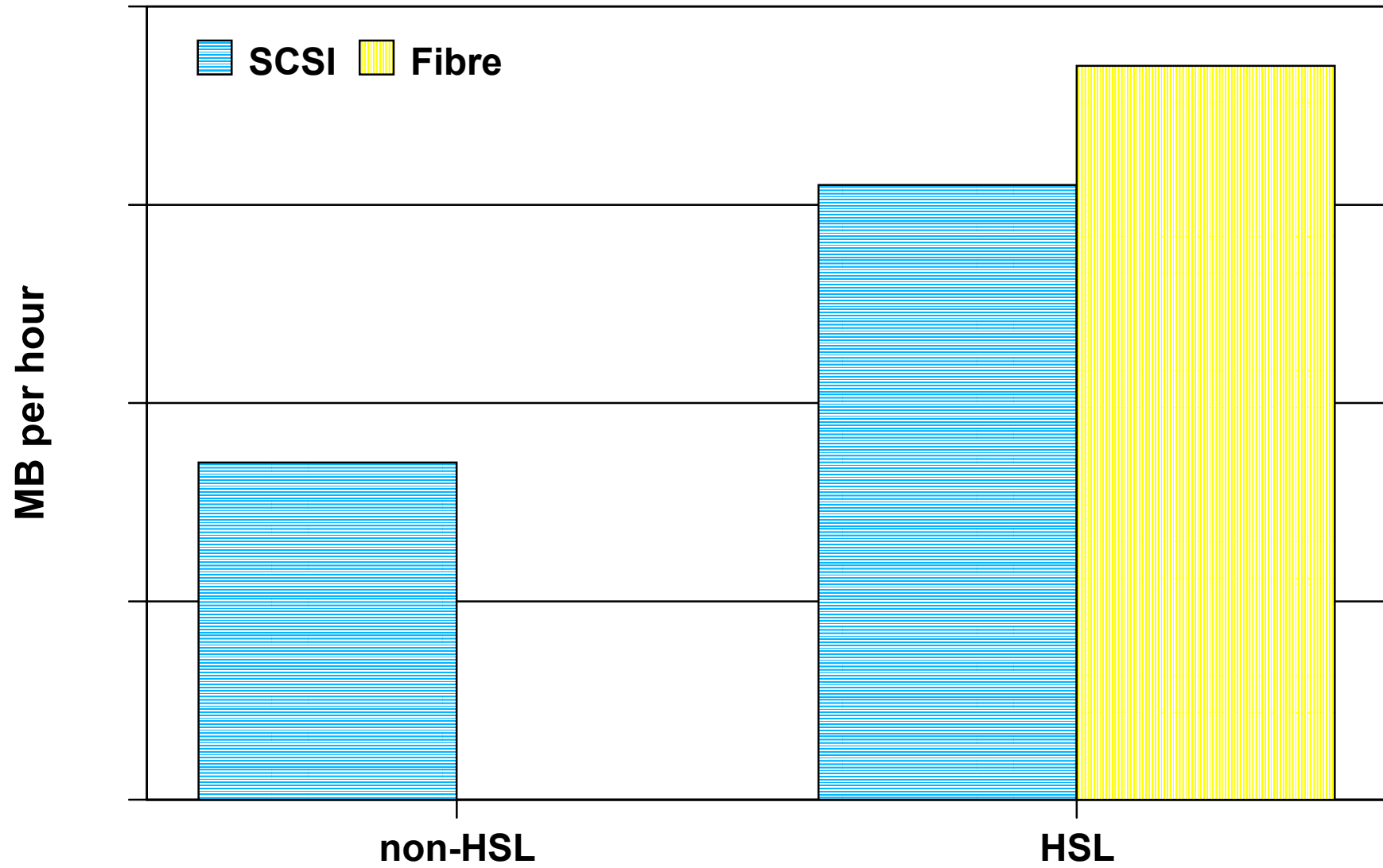
Save performance benchmark - traditional workloads

Save performance comparison, single tape drive, single save command

305 GB / hr

Legend:
- Source file
- 3 GB library
- 12 GB library
- 4 GB file
- 8 GB file
- 16 GB file

GB per hour (y-axis: 0, 50, 100, 150, 200, 250, 300, 350)

Categories: DVD-RAM *YES, DVD RAM *NO, SLR60, SLR100, VAX-2, 3580 gen1 - SCSI, 3580 gen1 - fibre, 3580 gen2, 3590E SCSI, 3590E FIBRE
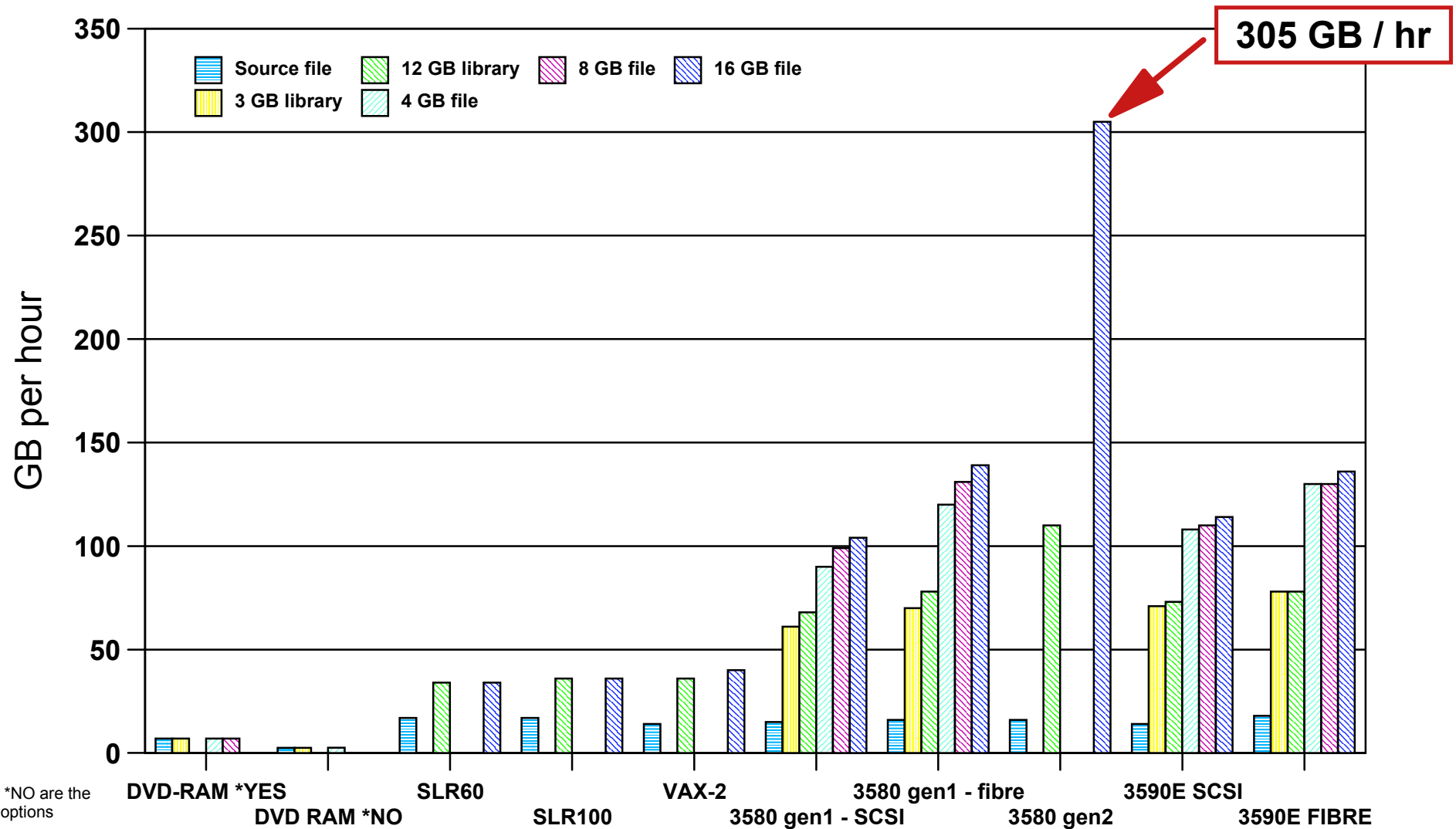
*YES and *NO are the DTACPR options used

Notes: Save performance benchmark - traditional workloads

All save/restore benchmarks were performed with USEOPTBLK(*YES) when the device supported this parameter.

For detailed information about the save/restore performance benchmarks, consult the *iSeries Performance Capabilities Reference*. This document is available on the Internet at URL http://www-1.ibm.com/servers/eserver/iseries/perfmgmt/resource.htm

The Performance Capabilities Reference refers to the following tape drives by indicated terms:

Tape drive	Reference term
16GB 1/4" cartridge	MLR1-S
25GB 1/4" cartridge	MLR3
50GB 1/4" cartridge	SLR100

The Performance Capabilities Reference workloads are:

Guide Workload Name	Presentation Workload Name	Description
NSRC1GB	Source files	96 source files with approximately 30,000 members.
NUMX3GB, NUMX12GB	3 GB library, 12 GB library	The User mix data is contained in a single library and made up of a combination of source files, database files, program and command objects, data areas, menus, query definitions, as well as other common iSeries objects found in libraries. NUMX12GB contains 52900 objects.
SR4GB, SR8GB, SR16GB	4 GB, 8 GB, 16 GB file	The large file workload is a single database file with members 4GB in size to create the file size being tested.

Save performance benchmark - IFS workloads

Save performance comparison, single tape drive, single save command
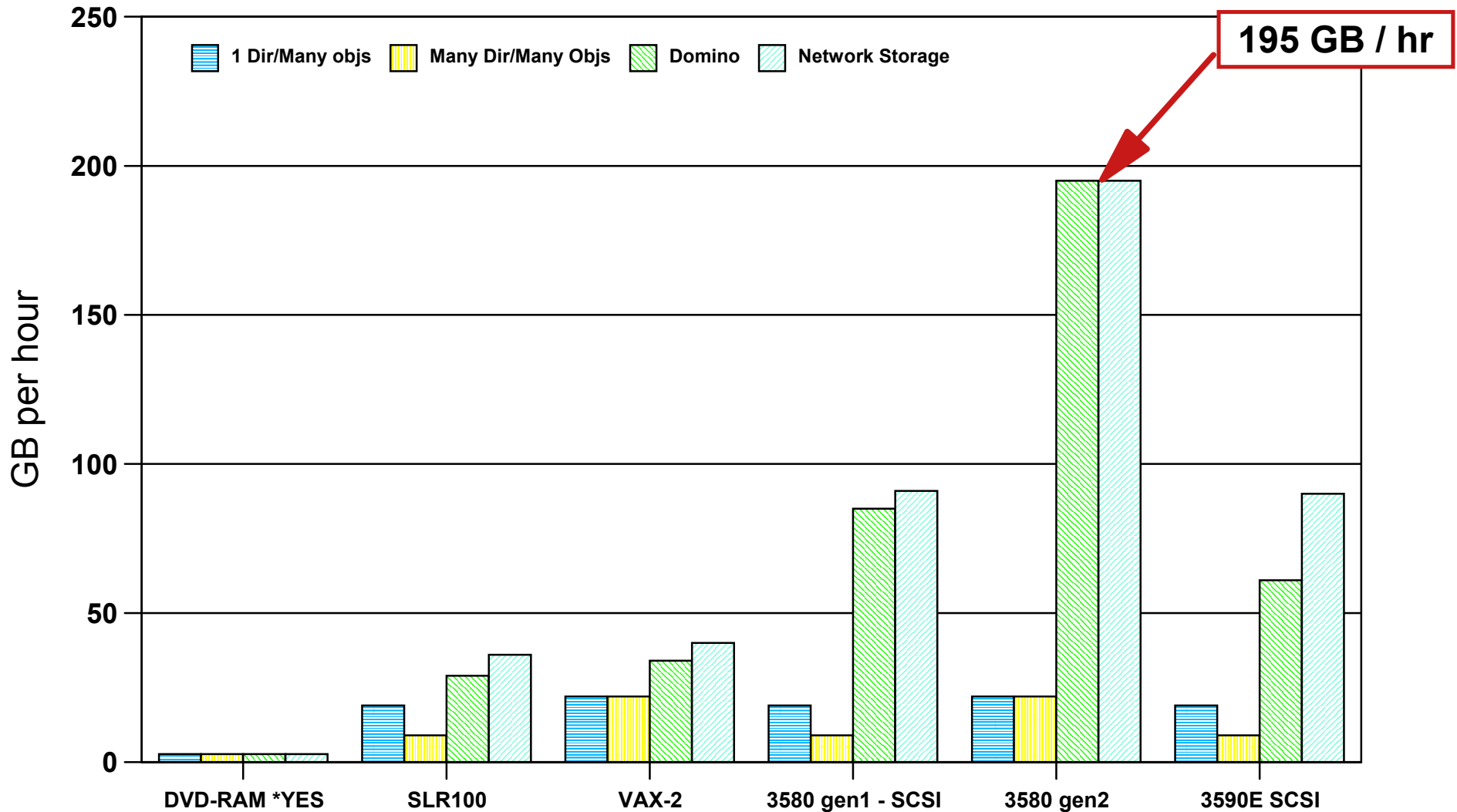
Notes: Save performance benchmarks - IFS workloads

For detailed information about the save/restore performance benchmarks, consult the *iSeries Performance Capabilities Reference*, available on the web at http://www-1.ibm.com/servers/eserver/iseries/perfmgmt/resource.htm.

Consult section "15.3 Workloads" for additional information about Integrated Files System related workloads.

1 Dir / Many Obj
    111,111 stream files in a single directory, approximately 4 GB total size
    Type2 directory structure

Many Dir / Many Obj
    111,111 directories with 111,111 stream files, approximately 5 GB total size
    Type2 directory structure

Domino
    Single directory with 90 mail files, each mail file 152 MB, approximately 13 GB total size

Network storage
    Linux storage space of approximately 6 GB.

Tips for IFS save/restore performance

Many directories, relatively few files? Take advantage of performance enhancing PTFs
        V5R1 - SI05856
        V5R2 - SI05599
    consider upgrading to Type2 directories
        See Information Center for migration information
        Consult information APAR II13161 for V5R1 PTF information.

Save performance benchmark - target SAVF

Save performance comparison, single tape drive, single save command

See details and notes on following page

Planning for Continuous Availability - 60
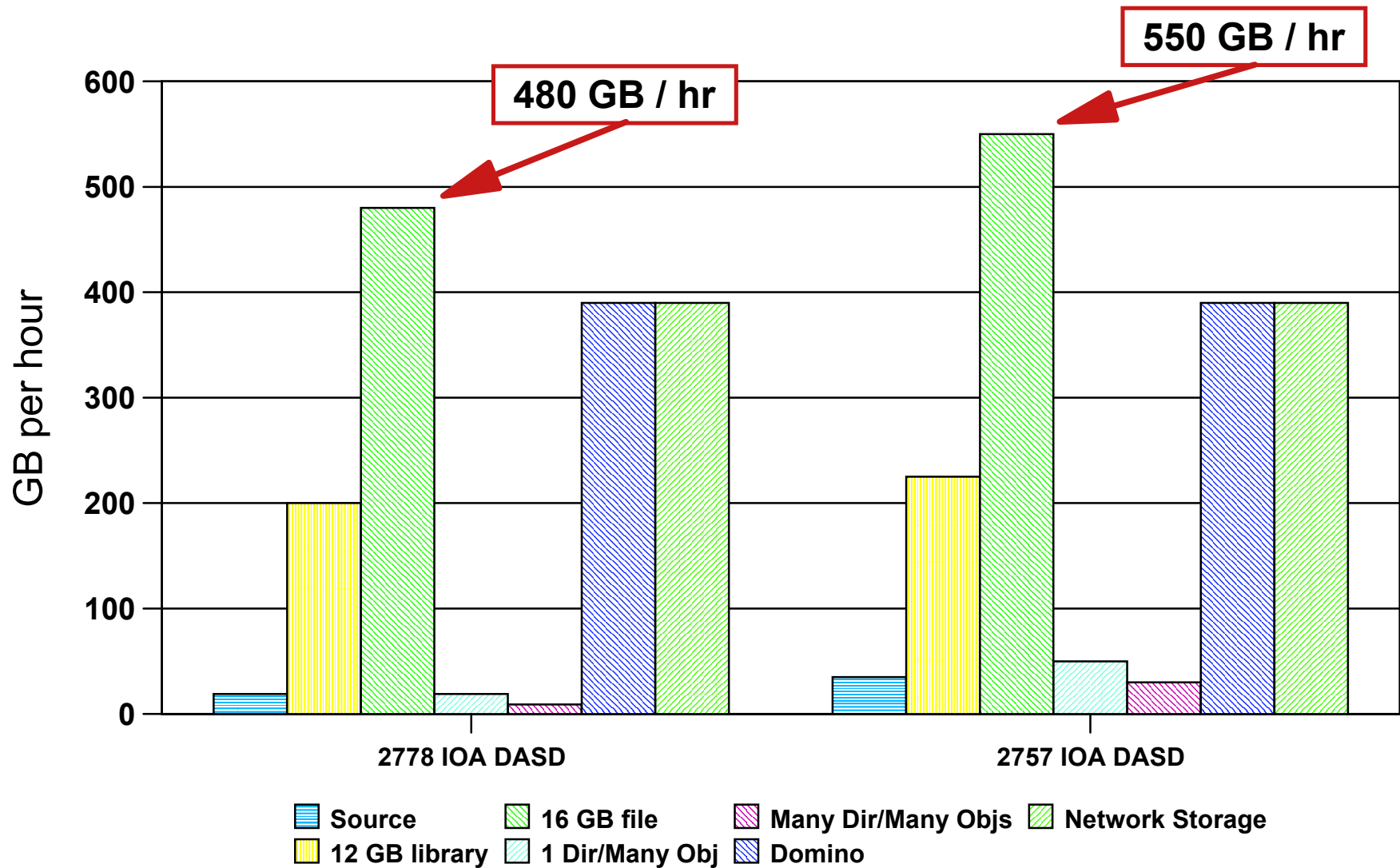
Notes: Saving to save files

Performance achieved depend upon available resources.

Save file benchmark performed in the most favorable environment to achieve maximum save rates

    24-way 840

    128 GB main storage

See section 15.17 of the **Performance Capabilities Reference Manua**l for additional information on maximum rates achieved.

New compression options available on QSRSAVO and QsrSave APIs in V5R2

    Reduces space needed on disk

    Degrades save performance significantly

        *LOW - least compression (1.5: 1), least CPU impact

        *MED - intermediate level of compression (2.7:1), intermediate CPU impact

        *HIGH - best compression (3:1), consumes most of CPU

    BRMS only supports *LOW compression

    Section 15.17 contains benchmark results for these different compression options.

Using SAN for sharing tape devices

- **In V5R1 - one (1) drive per adapter per system or partition**

Zone 1

Zone 2
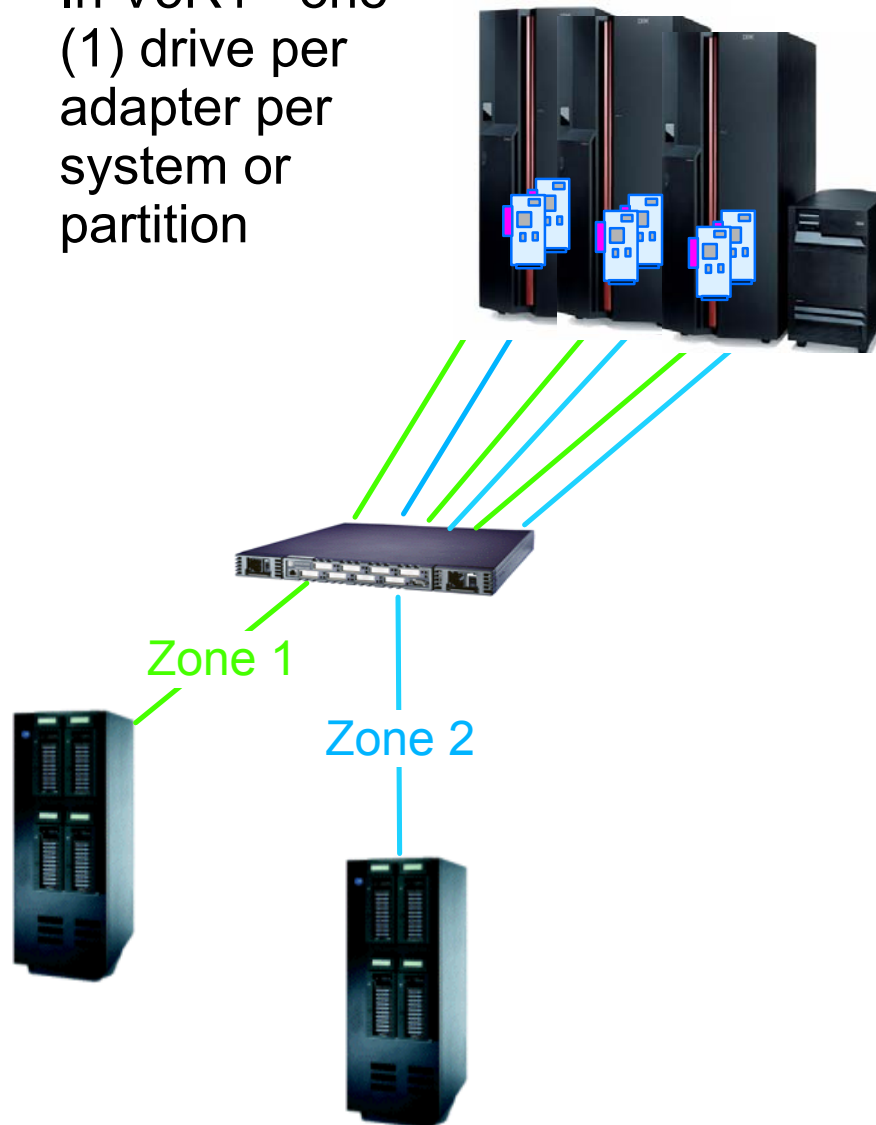
- **In V5R2**
  - ► Up to 16 devices per adapter per system or partition
  - ► Be aware of bandwidth requirements
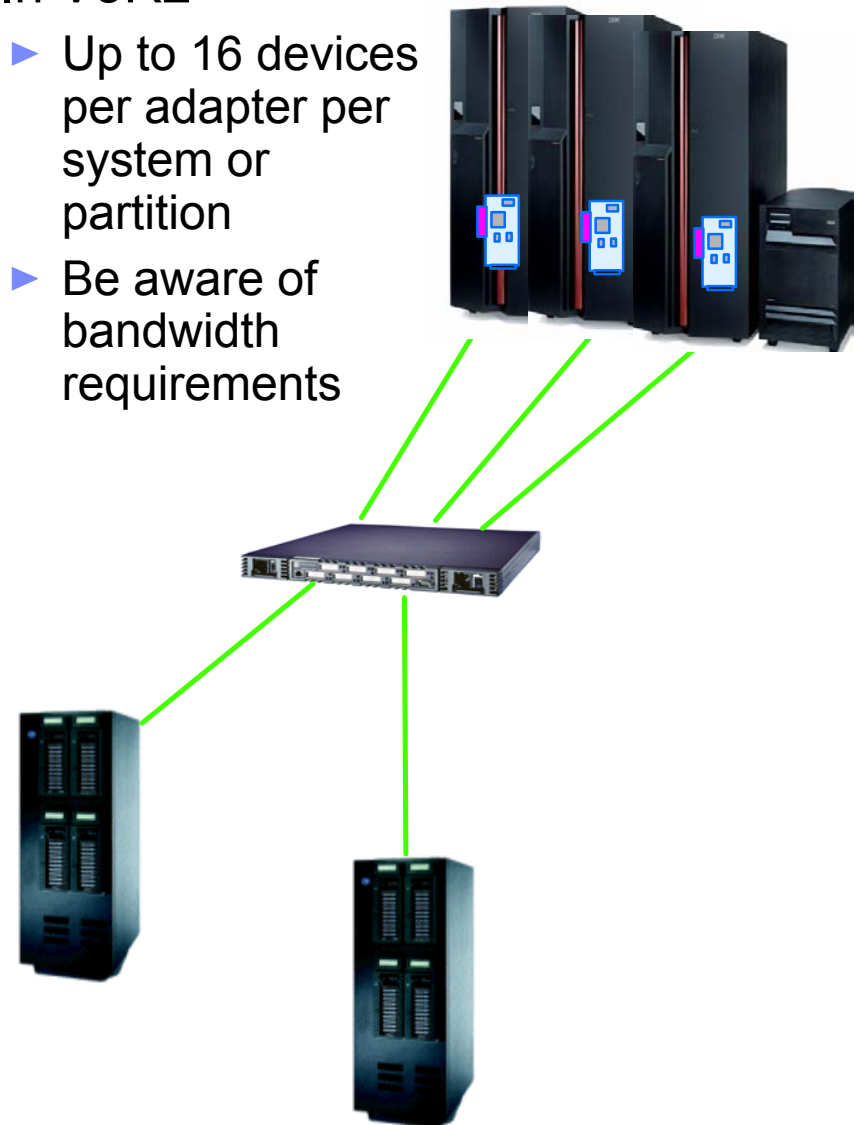
Notes: Alternate installation device versus alternate IPL device

Alternate IPL device must connect to IOP on bus 1

Alternate installation device can be on any bus.

    Defined through dedicated service tools (DST).

    Required for fibre channel attached tape drives

Alternate installation device **requires** IPL from licensed internal code CD for prior to recovering from tape.

Options for multiple devices

Concurrent Save

Save Job

Save Job

Save Job

Save Job

iSeries

Parallel Save

Save Job

iSeries
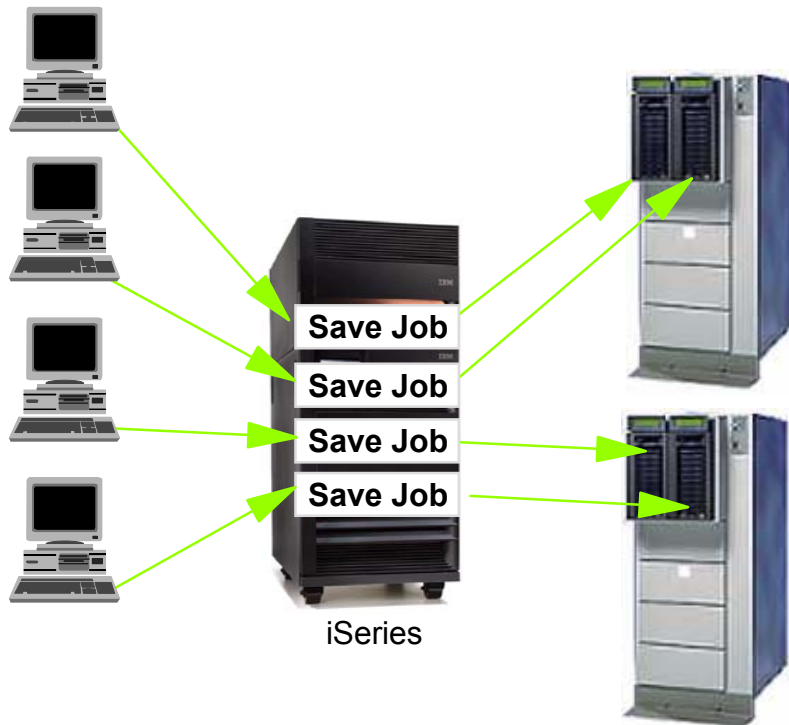
Notes: Options for multiple devices

Concurrent saves

Enhanced in V4R1 to allow multiple concurrent saves against single library or set of libraries. Enabled on SAVLIB, SAVOBJ, SAVCHGOBJ, and QSRSAVO API.

Generic values on LIB and OMITLIB parameters enhance concurrent support on SAVLIB and SAVCHGOBJ commands. For example:
   Omit groups of libraries via OMITLIB(TEMP*).
   SAVLIB LIB(*ALLUSR) using 2 tape drives concurrently:
      −SAVLIB LIB(*ALLUSR) OMITLIB(#LIBRARY A* B* ...L*)
      −SAVLIB LIB(*ALLUSR) OMITLIB(M* N* ...Z*)

Run multiple BRMS backup control groups concurrently to reduce save and restore times. User responsible for balancing the control group data so jobs end at same time.

Parallel saves

Starting in V4R4, parallel save and restore support enables a a large library, a large object, or a list of large objects to be saved to multiple tape drives in parallel, spreading the data across tape cartridges. It improves save performance for very large objects. And does not require the use of the same number of drives during the restore operation.

Expanded parallel support (V5R1 and higher) allows special values *ALLPROD, *ALLTEST, *ALLUSR, *IBM, *ASP01−*ASP32, and generics to be used. Type of spreading determined by a combination of special value, and an algorithm based on number of objects and their size in addition to the speed of the tape device. Restores for objects saved in parallel using special values are in serial restore mode (single tape device).

OS/400 Save-While-Active (SWA)

Enhanced in V5 !

- Integrated OS/400 utility for save window reduction
- No additional hardware or software required
- Virtual point in time copy on individual objects and libraries
- Allows granular restores from tape as needed

#1. Start of save request.

#2. Checkpoint complete - object changes now allowed.

#3. Unchanged pages and copied pages are saved to tape.

#4. Checkpoint image completely saved to media. End of save request.

#5. Results!

Objects in memory to be saved - **Get a checkpoint (snapshot) of each object.**

Original pages copied to temp storage before a change.

Copied pages in temp storage are released.

Objects on system after save.

Original objects as they were at checkpoint are written to tape.

Eric Hess/Sue Baker
IBM iSeries ATS 2001

Notes: Save-while-active

Significant enhancements since V4R1:

    Once checkpoint reached, locks for most objects dropped.

    More applications can restart after checkpoint message CPI3712 sent.

    Operations that can run after checkpoint message:

        Remove or rename members (RMVM, RNMM).

        Delete, move, or rename most objects including files (DLTF, MOVOBJ, RNMOBJ).

        Start a subsystem (STRSBS).

        S/36E operations such as LIBRLIBR.

Faster save-while-active checkpoint starting with V5R1

    Database changes to achieve faster checkpoint

V5R2 allows save-while-active to proceed for files with unrelated commit cycles

        • Limited to saving a single object

If you pursue a save-while-active philosophy to minimize outages, you may also want to consider utilizing OMIT options on the SAVSYS command to reduce save times.

    Configuration data - OMIT (*CFG)

    Security data - OMIT (*SECDTA)

    Both - OMIT (*CFG *SECDTA)

Utilize SAVCFG to concurrently save configuration data

Utilize SAVSECDTA to save user profiles and security attributes while the system is active

FlashCopy

- 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
  - Creation of test environment
  - Multiple platforms supported

- Considerations:
  - Must quiesce (maybe power off) primary server
  - Up to 2-hour outage or more on primary server
  - New complexities
    - Manual processes to switch
  - May impact disk performance
  - For nightly backups, may not be faster than OS/400 Save-While-Active

Primary iSeries

Backup iSeries or LPAR (inactive)

LSU

LSU

Point in time copy

Flash Copy

LSU

LSU

Single ESS
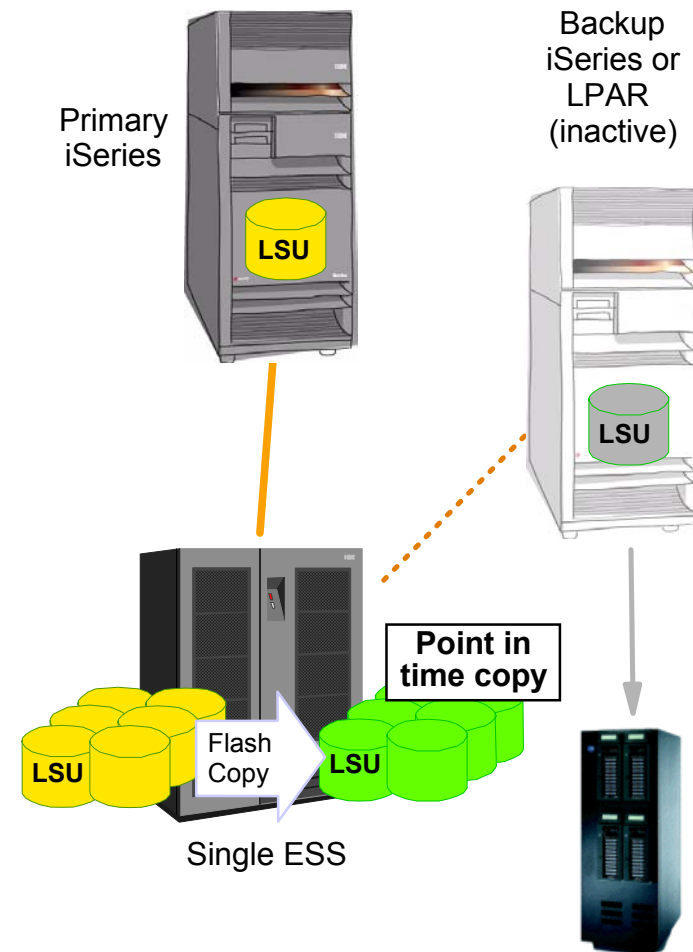
Notes: FlashCopy

FlashCopy Overview

FlashCopy can make an instant point-in-time (T0) copy of a volume or LUN. FlashCopy completes in two to five seconds. In two to five seconds you have a guaranteed point-in-time copy of the data, and you can access the source and target volumes.

When FlashCopy is invoked, the ESS builds control tables, then responds back almost instantly that the copy is complete. Next, the physical copy is made in a background task. During the copy process, if the requested data has not been copied to the target volume, a copy-on-demand command for the requested data will be issued, and the data will be copied immediately. In the unlikely event of a failure during the copy process, the ESS can recover back to T0, or the starting point.

For Open Systems, FlashCopy is invoked from the StorWatch ESS Specialist. FlashCopy is selected by clicking on the Copy Services button on the main ESS Specialist window.

After you select Copy Services, identify the source and the target volumes by a left click on the source and a right click on the target.

The ESS Specialist shows the volumes, or LUNs, by their ESS internal serial numbers. You will first have to find out from the host, the serial number of the source and target volume you intend to copy, in order to identify the volume on the ESS Specialist.

Next, you select the task that you want to perform on the volume pair. In this case it is FlashCopy. You have the option to suppress the background copy task. This can be useful if you need the copy only for a short time.

Another option lets you remove, or withdraw, the relationship between source and target. You only need this if you had specified the NOCOPY option; otherwise, the relationship will automatically end when the physical copy has completed.

Online Backup of Lotus Servers

Lotus® Domino™ 6

Lotus® iNotes™

Lotus® QuickPlace™

Lotus® Sametime®

- Save while active with no checkpoints
- Enabled through BRMS

Notes: Online Backup of Lotus Servers

Full save:

Save Domino server while active

V4R4 BRMS PTF SF65062 - integrated into V4R5 and higher

Domino 5.02c

No need to reach checkpoint - true online backup

No need to end server or replicate to another server for backup

BRMS/400 handles recovery process

Allows 24X7 up time of Domino server

Save Quickplace server while active

Quickplace 1.03

Incremental save

Incremental backups - Domino servers

V4R4/V4R5 BRMS PTF SF65708 - integrated into V5R1 or higher

Domino 5.08.01

V5R1 BRMS PTF SI01742

BRMS/400 handles recovery process

Allows point in time restores

Incremental backups - Quickplace servers

Quickplace (TBD)

For more information go to: http://www-1.ibm.com/servers/eserver/iseries/service/brms/domIncremental.htm

Integrated servers save/restore

- **IXS/IXA:**
  - ► Save storage space via IFS SAV tools *(fast)*
  - ► Save with 2000/NT backup utilities *(granular)*
    - – From IXS/IXA using iSeries attached tape or save files
    - – From a network attached 2000/NT workstation using workstation tape
  - ► Save with Tivoli Storage Manager *(Enterprise)*
    - • TSM running on IXS/IXA
    - • TSM running on network attached 2000/NT workstation

- **Linux - Virtual disks:**
  - ► Save storage space via IFS SAV tools *(fast)*
  - ► Save with Linux backup utilities *(granular)*
    - – From Linux partition using iSeries attached tape or save files

- **Linux - Dedicated disk** (Linux controls iSeries disks)
  - ► Save with Linux backup utilities *(granular)*
    - – From Linux partition using iSeries attached tape or save files

IXS/IXA save restore

iSeries tape drive

Storage Space (IFS save)

TSM server

2000/NT tape drive

LAN
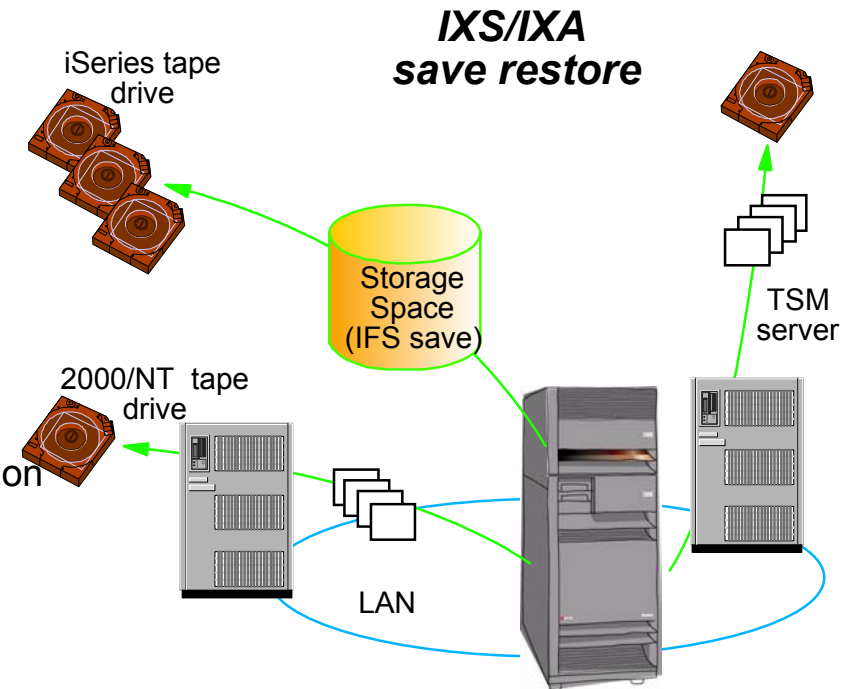
Protect your data between *saves*

Journaling

- **Benefits beyond data recovery:**
  - ► Completely secure
    - – Entries can only be added not removed
  - ► Enhances security by providing an audit trail
    - – Including user defined entries for any activity, any object
  - ► Basis for data replication

- **New in V5R2:**
  - ► Feature 42 of OS/400 - HA Journal Performance
    - – Standby mode
    - – Journal Caching

- **New in V5R1:**
  - ► Additional object types:
    - – Byte stream files (IFS files)
    - – Data queues
    - – Data areas
  - ► Journal only changed data (DB2 objects only)

FILEA   FILEB   FILEC

CHANGES   A D D S   DELETES

Journal

Detached receivers, ready for save to tape

Receiver #2   iver #1
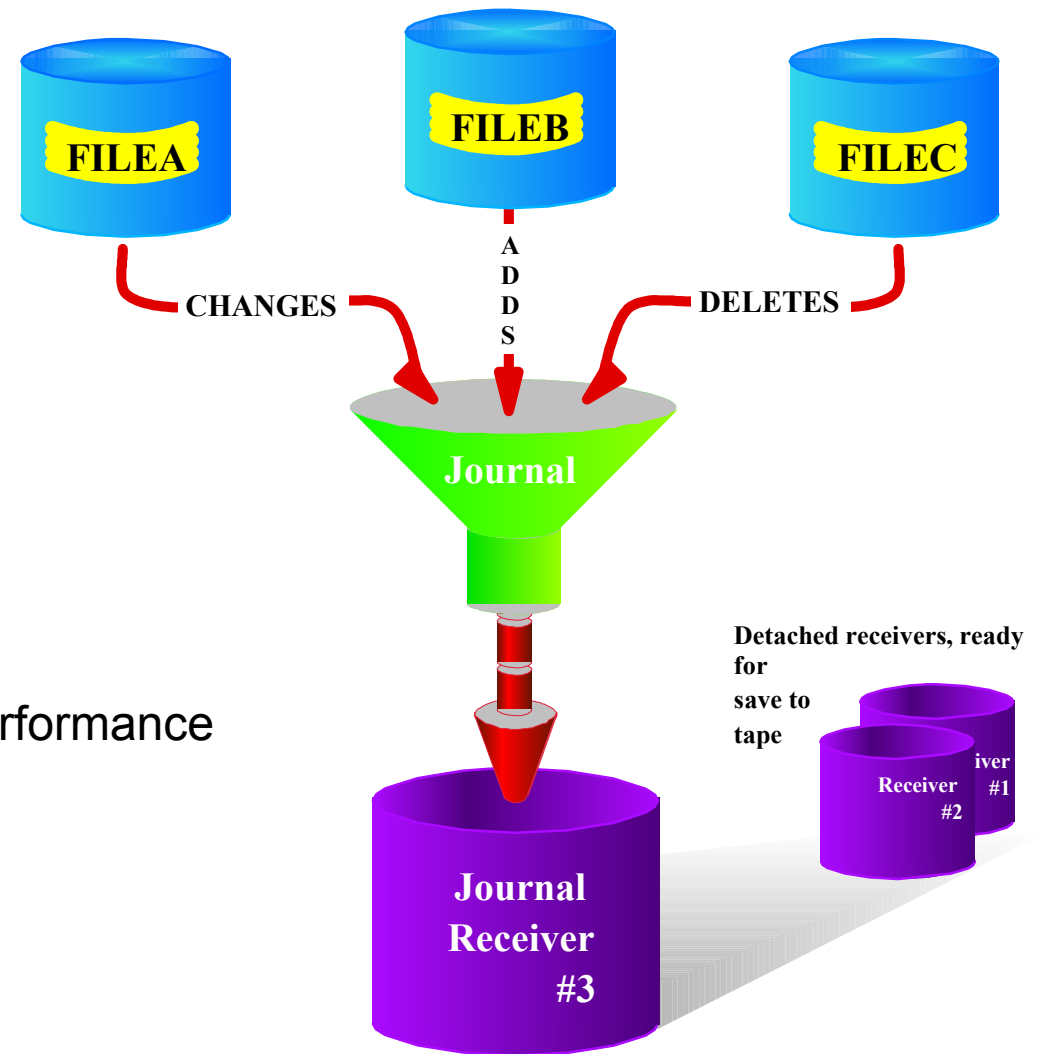
Journal Receiver #3

Using journal to recover

- **Restore most recent save**

- **Restore, if necessary, journal and receivers**

- **Use APYJRNCHG command to replay the entries**
  - ► Make this go fast by using multiple APYJRNCHG commands ...
    - – By library
    - – By file
  - ► One command per processor

- **Utilize Apply Journal Changes Extended PRPQ (5799-AJC) to reduce manual interventions**
  - ► www.ibm.com/eserver/iseries/ha/hajournal.htm

FILEA

FILEB

FILEC

CHANGES

A D D S

DELETES

Journal

Applied receivers

Journal Receiver #3

Receiver #2

iver #1
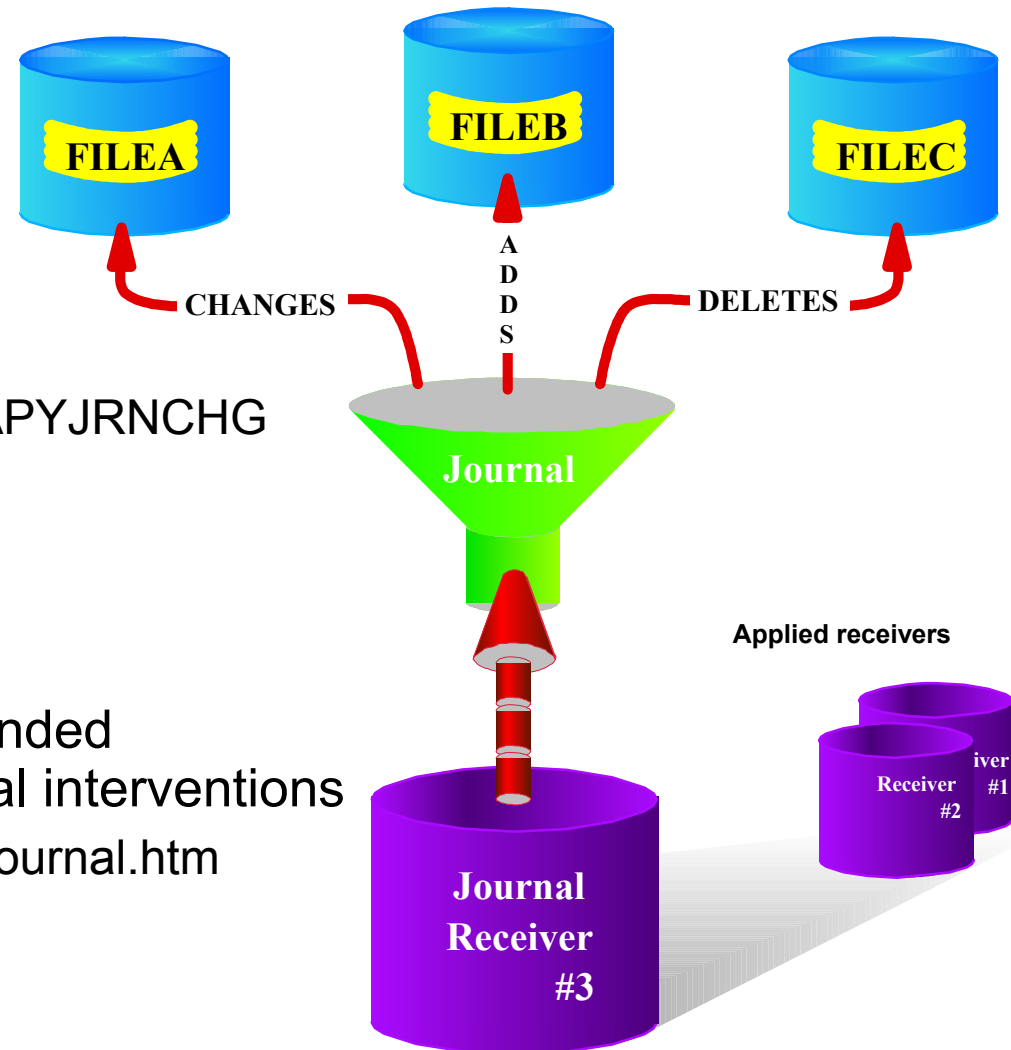
Journal performance myths

- Consumes up to 25% CPU resource

    ✓ Recent ITSO residency measured CPU overhead at 2-3%

- Difficult to manage

    ✓ System management eliminates the headaches

- Requires a user ASP

    ✓ User ASP is recommended for highest availability, but not required

- Runtimes will double or worse

    ✓ It depends!

Batch Job 5 Million DB operations (10% Adds, 90% Updates) 9 Million journal entries	Elapsed Time (minutes)
Ordinary journaling enabled	163
Journal cache enabled	24

Target system apply process	Database replay rate (trans/hr)
Without caching	600,000
With caching	2,400,000

Optimized journal environment

- Use fastest disk arms and largest IOP write cache

- Use system managed receivers

- Use Batch Journal Cache PRPQ or OS/400 feature 42 HA Journal Performance
  - ► Significant performance benefit for some environments
  - ► Changes behavior of journaling for add, update, delete database operations
  - ► Benefit for both primary and target systems within a cluster

- Consider the use standby mode on target data replication system

- Recommendations:
  - ► 1 CPU / 1000 entries per second
  - ► Minimum of 2 arms per "active" receiver
  - ► Isolate activity to user ASP

FILEA **FILEB** **FILEC**

CHANGES ADDS DELETES

Journal

Detached receivers, ready for save to tape

Receiver #1
Receiver #2

Journal Receiver #3
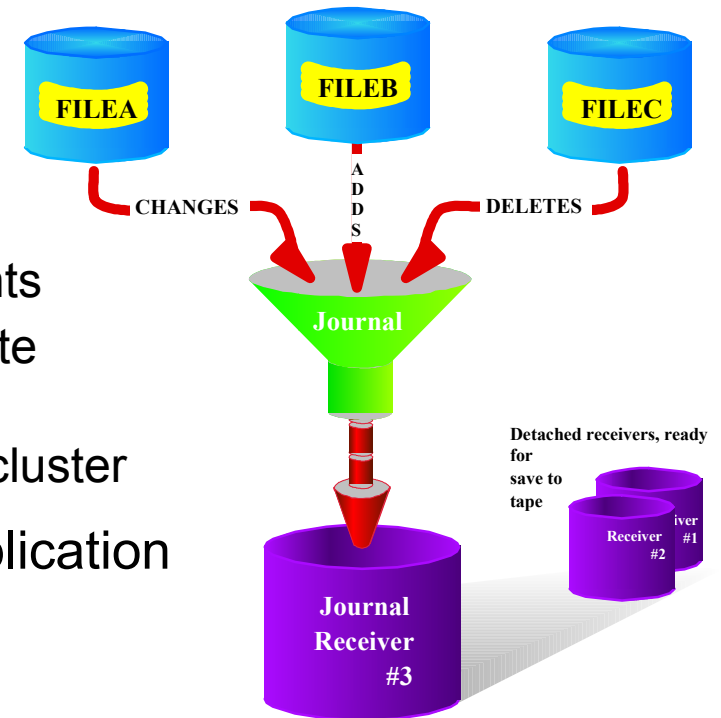
Striving for Optimal Journal Performance on DB2 Universal Database for iSeries
IBM Form Number: SG24-6286

Remote journaling

Basic Configuration

Source ○———▶ Target

LJ   RJ1
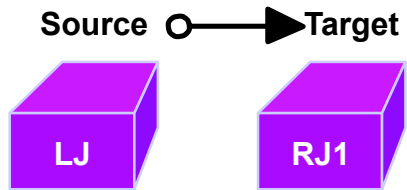
- **Synchronous or asynchronous modes**
- **Basic, cascade or broadcast configurations**

Cascade Configuration

Source ○——▶ Target   Source ○——▶ Target

LJ   RJ1   RJ2   RJ3

Source ○——▶ Target

Targets:

Broadcast Configuration

RJ1

LJ ●———▶ RJ2

Source

RJ3
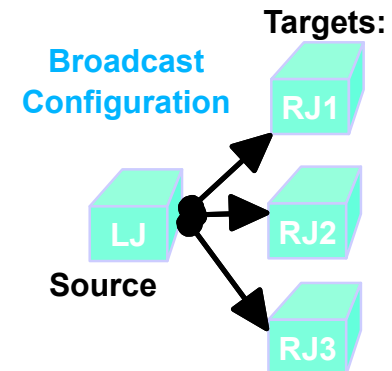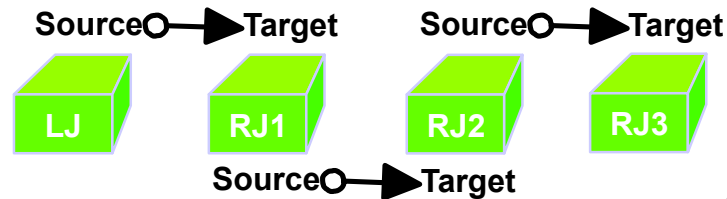
AS/400 Remote Journal Function for High Availability and Data Replication
IBM publication number SG24-5189

Redbooks

Access path protection

- **Access path journaling**
  - ► Provides ability to protect selected mission critical access paths
  - ► Can be used in conjunction with SMAPP to insure protection of critical access paths

- **System managed access path protection (SMAPP)**
  - ► Controlled (on/off/statistics) via access path commands
    - – EDTRCYAP to change
    - – DSPRCYAP to view current exposures

Understand access path exposure

Rebuild at file ope   Rebuild after IP   Rebuild at IPL
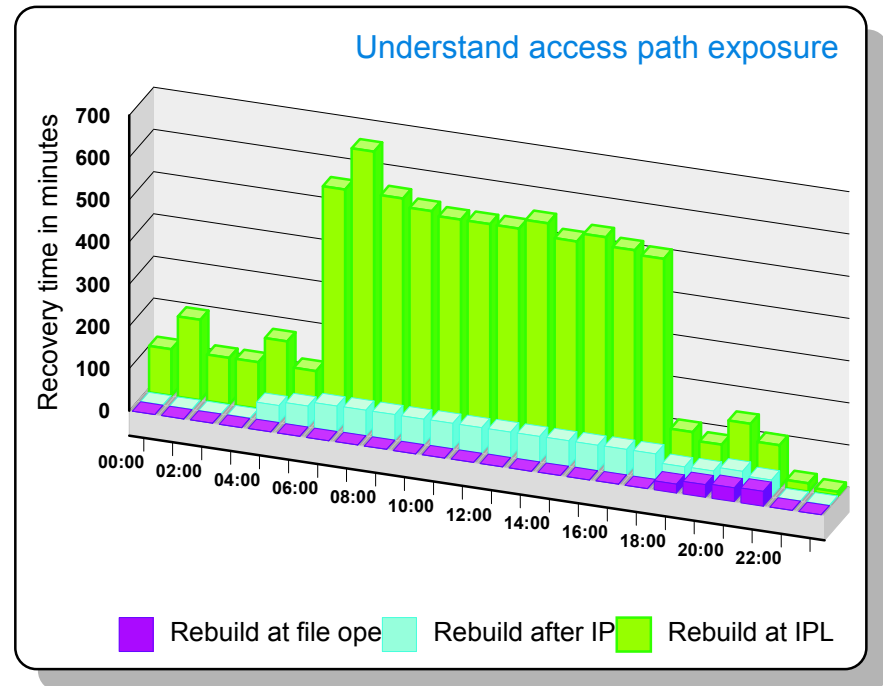
```
                Display Recovery for Access Paths

Estimated system access path recovery time  . . . . :          19
Minutes
Total disk storage used . . . . . . . . . . . . . . :         .138    MB
% of disk storage used  . . . . . . . . . . . . . . :         .000
System access path recovery time  . . . . . . . . . :         150


     ----------Access Path Recovery Time----------  --Disk Storage
Used---
ASP      Target (Minutes)        Estimated (Minutes)   Megabytes
ASP %
 1             *NONE                     219                  .098      .000
 2             *NONE                       0                  .040      .000
```

V5R2 System Managed Access Path Protection changes

- **Default changed to 70 minutes**

- **Access paths for physical files attached to standby journals are exempt from SMAPP protection**
  - ► These access paths are considered "ineligible" for protection

```
                Display Not Eligible Access Paths


                           Estimated
                           Recovery
File       Library    ASP       Time      Reason Not Eligible
INDEX14F   DATA1      1         00:22:14  FRCACCPTH(*YES)
INDEX2     IASPLIB1   IASP39P   00:08:11  FRCACCPTH(*YES)
INDEXA     IASPLIB1   IASP39P   00:06:04  Multiple Journals
INDEX22S   DATA1      1         00:04:02  JRNSTATE(*STANDBY)
```

- **New displays**
  - ► Top 500 ineligible access paths per ASP
  - ► Largest "at risk, yet protected"

- **New option to select SMAPP behavior**

```
               Display Protected Access Paths


                             Estimated
                             Recovery
File       Library    ASP       Time
INDEX2A    DATA1      1         00:52:42
INDEX3A    IASPLIB1   IASP39P   00:32:42
INDEX3B    IASPLIB1   IASP39P   00:10:42
INDEX27A   DATA2      1         00:08:42
INDEX33A   DATA1      1         00:08:42
INDEX2A    DATA2      1         00:00:22
INDEX27A   DATA2      1         00:00:21
INDEX3A    DATA1      1         00:00:21
```

mal

Save/restore summary

- **Use the latest techniques to minimize planned downtime**
  - ► Multiple high speed tape devices
  - ► Concurrent and parallel saves
  - ► Save-while-active and on-line Lotus server saves
  - ► ESS FlashCopy
  - ► Protect orphan transactions with journaling

- **Test, test, and test again**
  - ► Recovery time objective met?
  - ► Recovery point objective met?

- **Use iSeries clusters when availability requirements exceed a single server's capability**

Storage Disaster Recovery Solutions

(a.k.a. Copy Services)

ESS Copy Services

- **FlashCopy**
  - ► Local copy
    - – Tape backups, test copies

- **Peer-to-Peer Remote Copy**
  - ► Local, synchronous copy
    - – Campus or local disaster recovery

- **Peer-to-Peer Remote Copy - XD**
  - ► Extended distance
  - ► Asynchronous remote copy
    - – Remote data migration

FlashCopy

Local or campus secondary ESS

Backup iSeries (inactive)

PPRC

Local or campus secondary ESS

Backup iSeries (inactive)
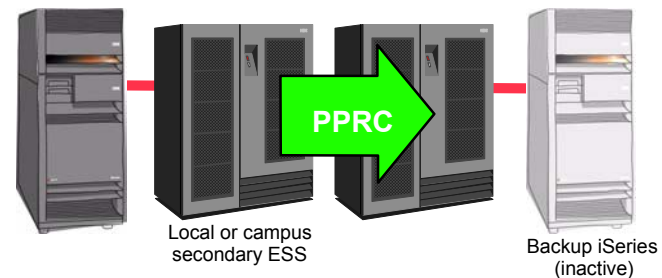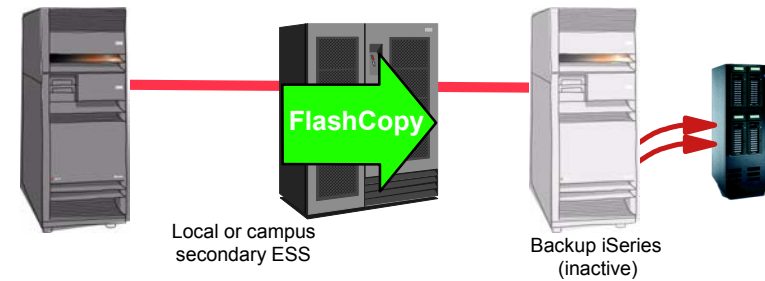
Peer-to-Peer Remote Copy (PPRC)

- Continuous, synchronous copy to a second ESS

- Benefits:
  - ► Campus/city disaster recovery (up to 103km)
  - ► Multiple platforms supported

- Considerations:
  - ► Data loss - does not protect transactions or data in server memory
  - ► Unplanned failover can be many hours
    - – IPL required, plus any DB recovery
  - ► New complexities
    - – Manual processes to switch or recover
  - ► Will impact disk performance
  - ► High communication costs
  - ► Second ESS disk copy and second iSeries are unavailable for other uses
  - ► Not viable for masking outages from software and hardware upgrades and maint.

Primary iSeries

Primary ESS

LSU

LSU

PPRC continuous copy (sync up to 103km)

ESCON

LSU

LSU

Local or campus secondary ESS

Backup iSeries (inactive)
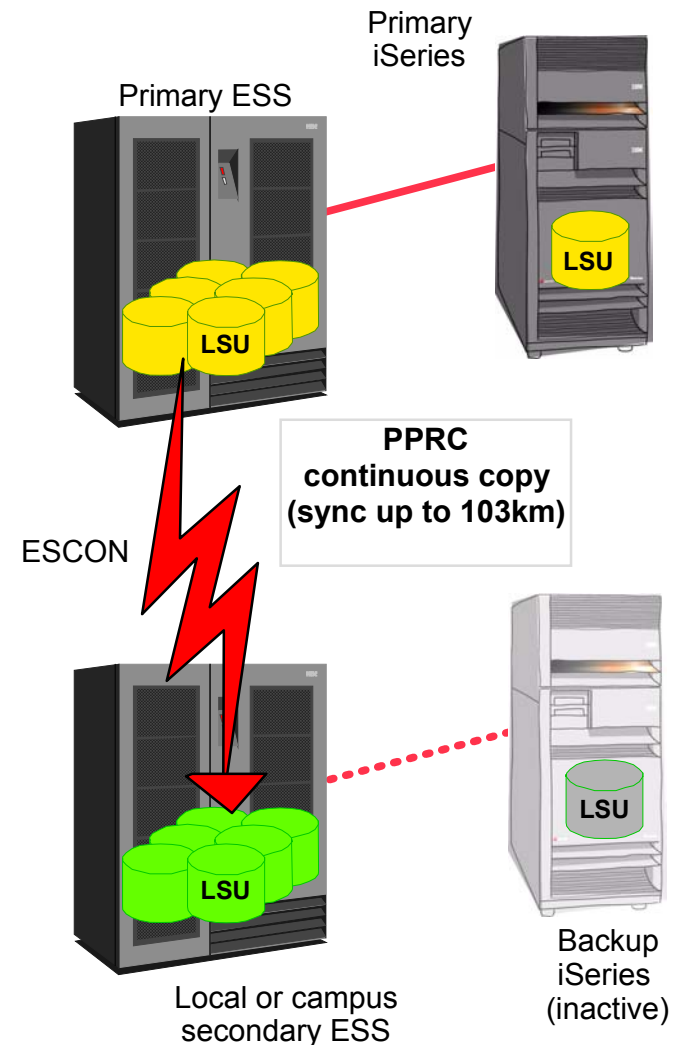
PPRC - XD (Extended Distance)

- **Asynchronous, remote copy**
  - ► Remote ESS copy is 'fuzzy'
  - ► Fuzzy copy is unusable until it is synchronized with primary ESS

- **Benefits:**
  - ► Data migration over long distances
  - ► Multiple platforms supported

- **Considerations:**
  - ► Not a good DR solution
  - ► Daily outage (up to two hours) on primary server to create 24-hour checkpoint
  - ► Not a good Disaster Recovery solution for the money (i.e., 24-hr. checkpoint)
  - ► One day's worth of transactions lost
  - ► May impact disk performance
  - ► High communication costs
  - ► New complexities
    - – Manual processes to switch or recover

Primary iSeries

Primary ESS

LSU

LSU

network

PPRC - XD

Fuzzy copy

LSU

Flash Copy

Remote ESS

FlashCopy

LSU

LSU

Backup iSeries

Point in time copy
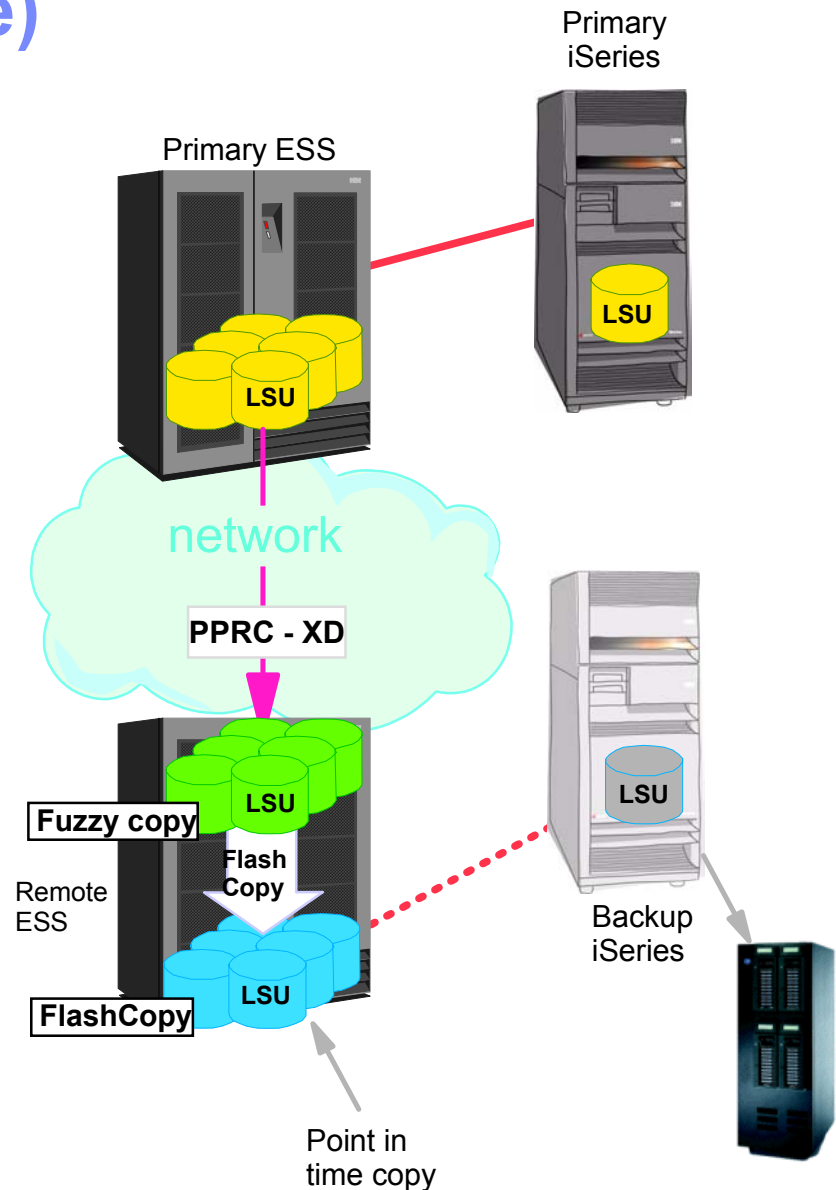
Clustering

Why clusters?

- ■ **Clusters are demanded because...**
  - ► RAID/SAN is not enough
    - − Data and applications must be <u>resilient</u>
    - − Networks, power, Internet connections must be <u>redundant</u>
  - ► Fail-overs and switch-overs are <u>required</u>

single system image

Any network

Appl Srvrs   **Data Srvrs**

Routers
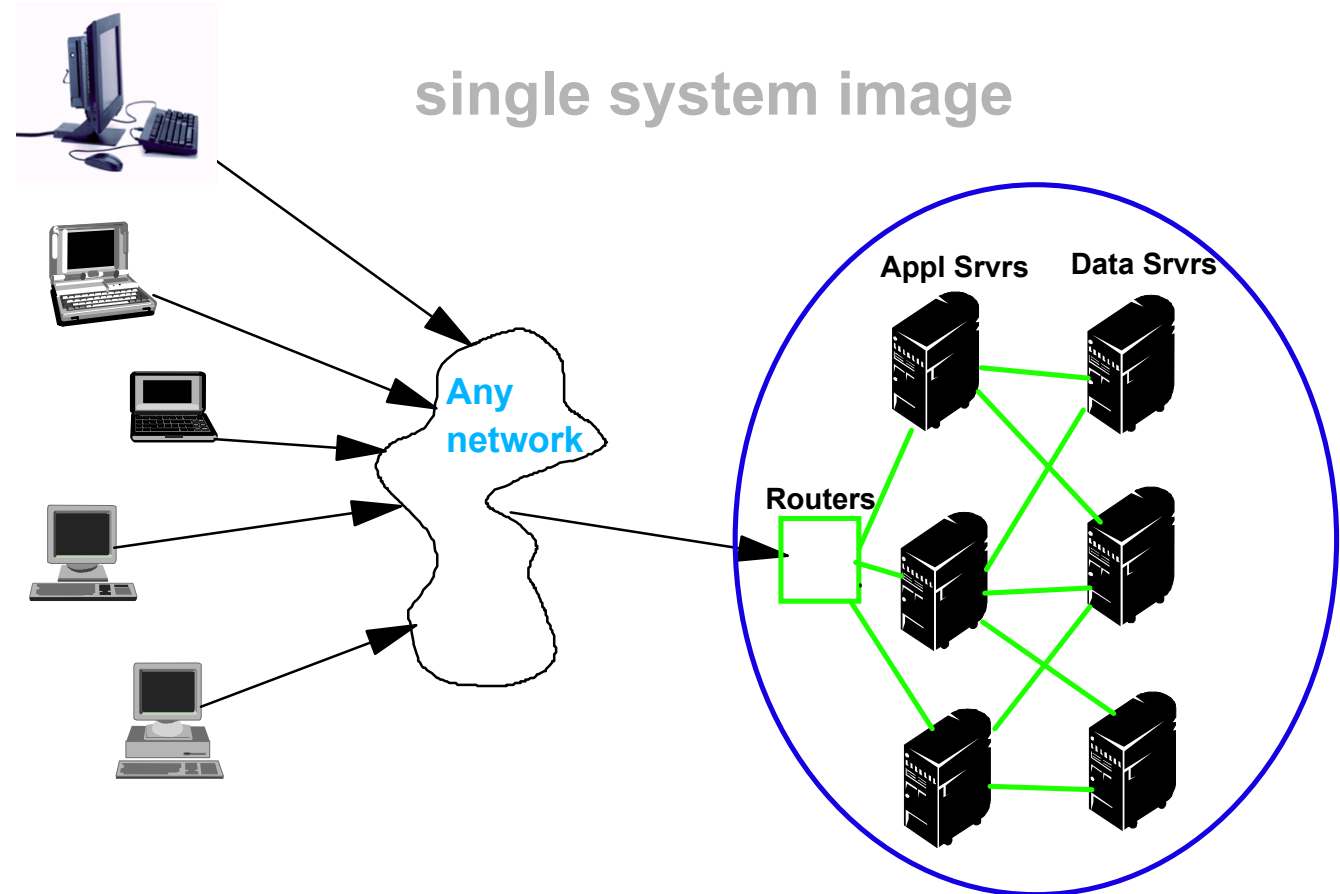
iSeries clustering solutions

**iSeries
data replication
HA cluster**

**iSeries
switched disk
cluster**

HSL

**Microsoft Cluster
Services**
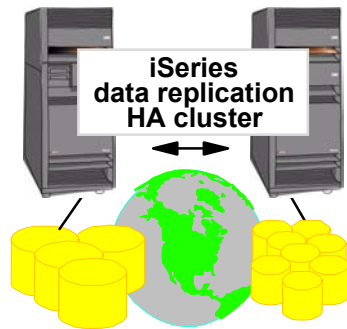
- ■ Data and application resiliency through replication
  - *Local or remote coverage; unlimited distance*
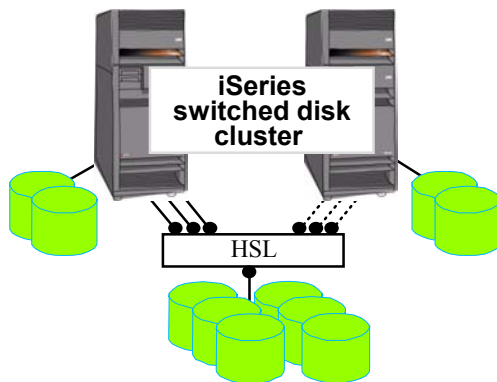
- ■ Device and application resiliency via Switch Disk
  - ► iSeries Switched Disk Clusters
  - *Campus coverage up to 250 meters*
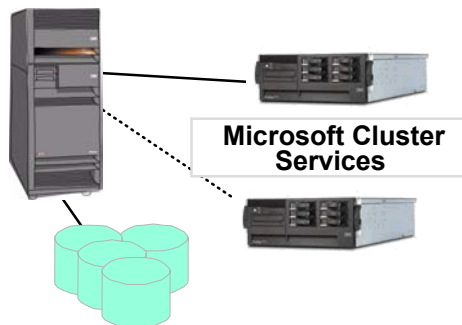
- ■ Integrated xSeries resiliency *(IXS/IXA)*
  - ► Hot spare, switched disk, Microsoft Cluster Services, replication cluster
  - *Campus coverage up to 250 meters*

- ■ Linux partitions
  - ► Local or remote coverage

iSeries data resiliency clusters

Data & CPU concurrently useable on backup server

Primary iSeries

Local or remote backup iSeries (active)

Cluster Middleware:

Object Replication

Cluster Management & Synchronization

LAN-WAN-HSL-OC400

Cluster Middleware:

Object Replication

Cluster Management & Synchronization

Tape

Nightly backups using SWA, and parallel and concurrent saves
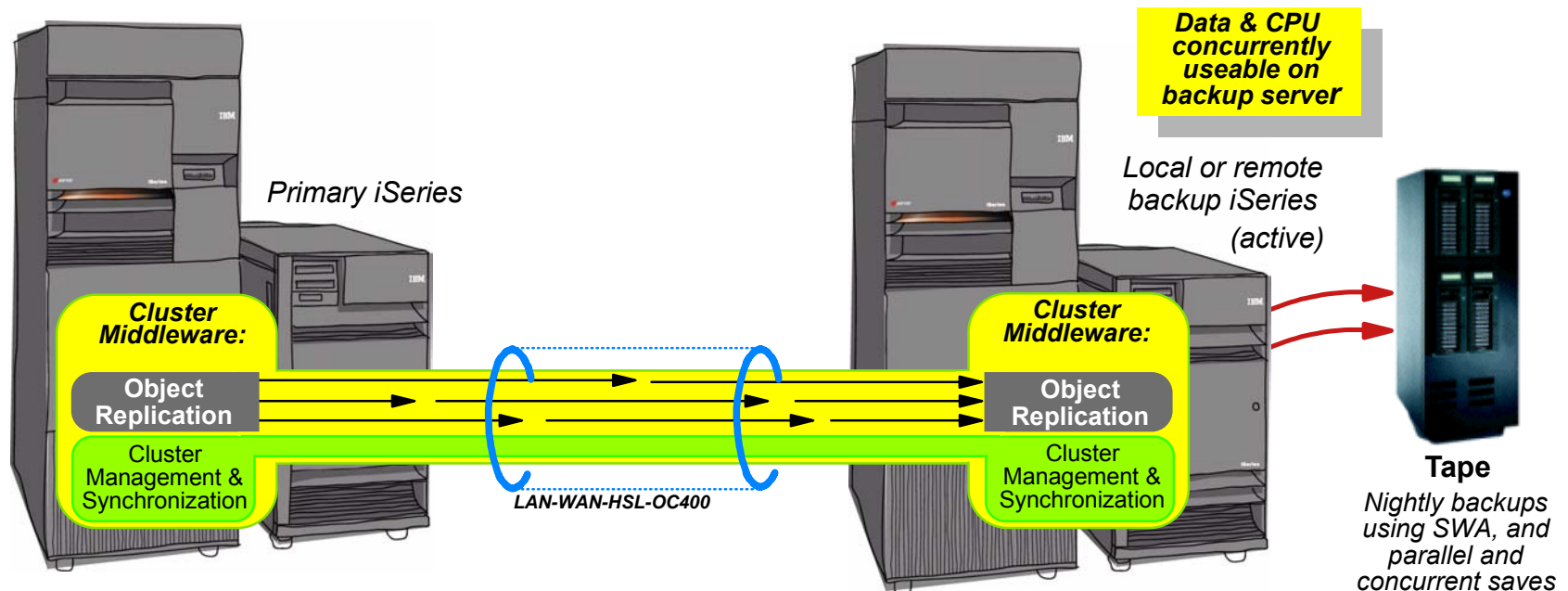
- ■ Data resiliency
  - ► System level approach to maximizing availability
    - − Addresses planned and unplanned outages
    - − Can optionally address disasters
    - − Can offload tape backups and read only workload to backup (secondary) server
  - ► Provides maximum availability levels and versatility for iSeries

Replication using local journal

Primary server

Applications

OS/400

Audit journal

JRN RCV

Replication software

audit JRN entry retrieve process

JRN entry retrieve process

Send Process

Cluster Management & Synchronization

LAN-WAN-HSL-OC400

Secondary server

DB file

DB file

Apply process

Staging area

Receive Process

Cluster Management & Synchronization
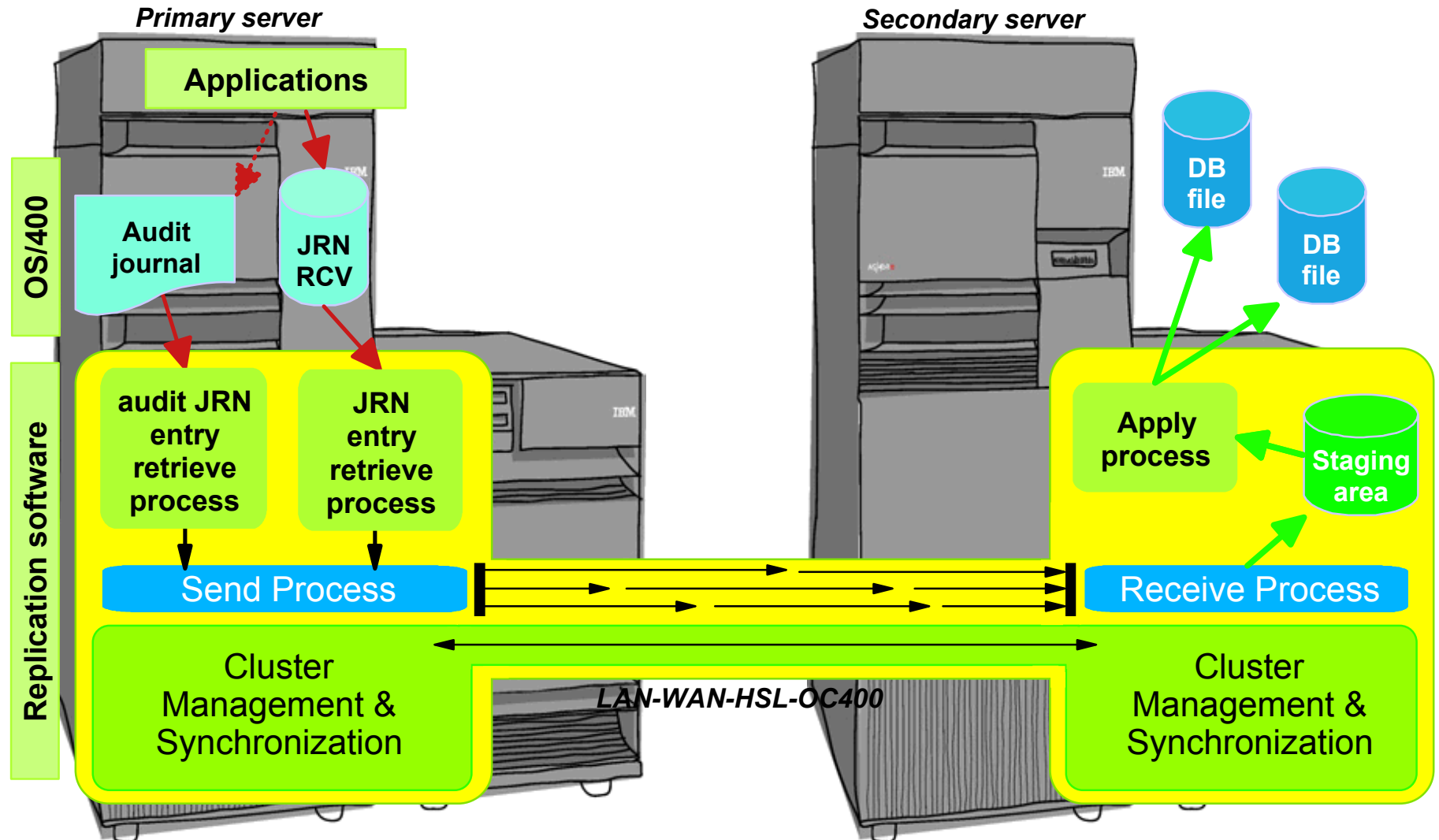
Replication using remote journal

Primary server

Applications

OS/400

Audit journal

JRN RCV

LAN-WAN-HSL-OC400

Replication software

audit JRN entry retrieve process

Send Process

Cluster Management & Synchronization

LAN-WAN-HSL-OC400

Secondary server

RMT JRN RCV

DB file

DB file

Apply process

Staging area

Receive Process

Cluster Management & Synchronization
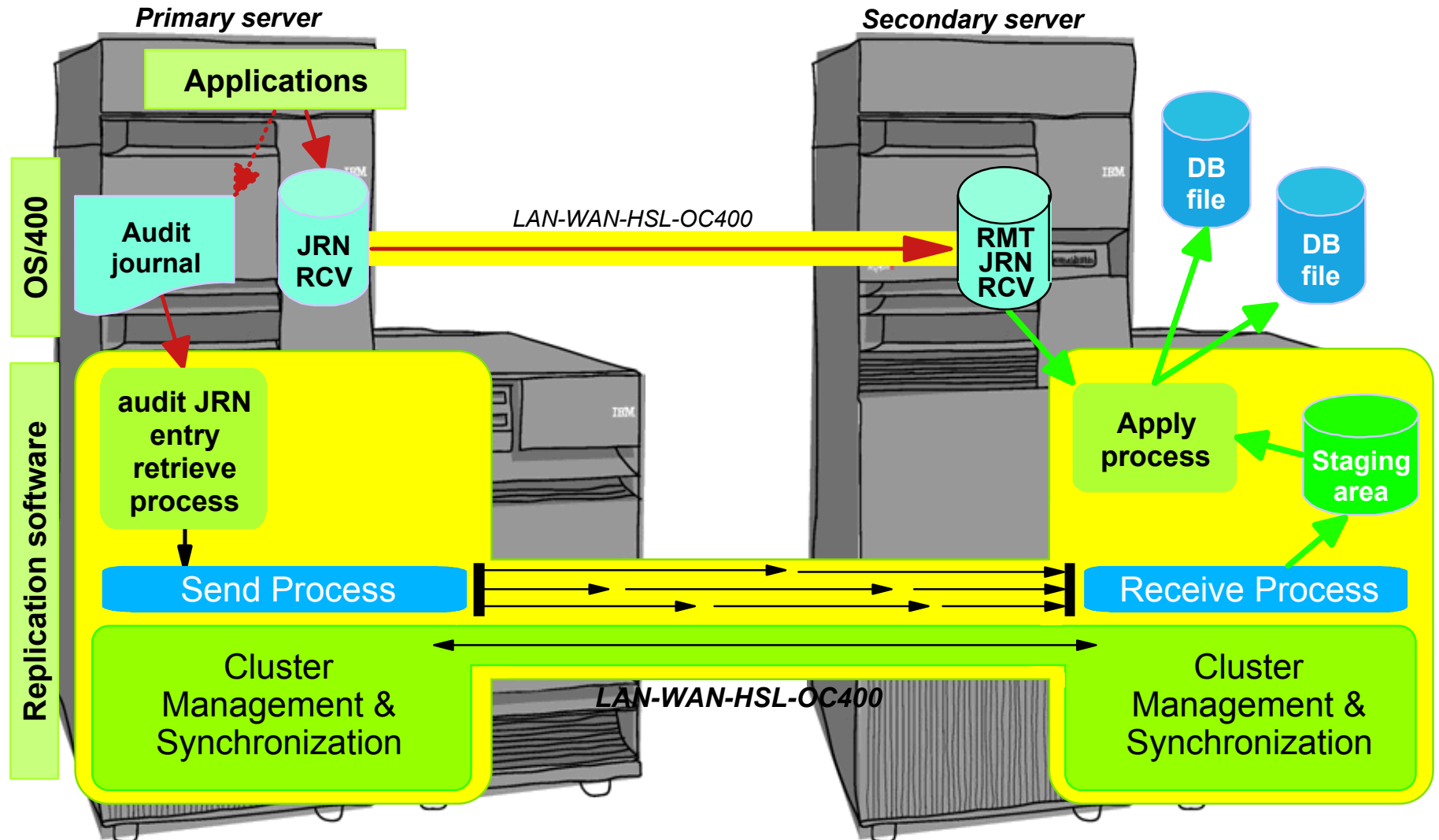
Reduce target latency with journal standby mode

Standard operational mode

FILEB
FILEC
FILEA

Journal

Local Receiver

Local iSeries - active

LAN-WAN-OC400

FILEB
FILEC
FILEA

HABP Cluster Middleware Apply Process

Remote Receiver

Journal

Standby Mode

Standby Receiver

Backup iSeries - standby mode
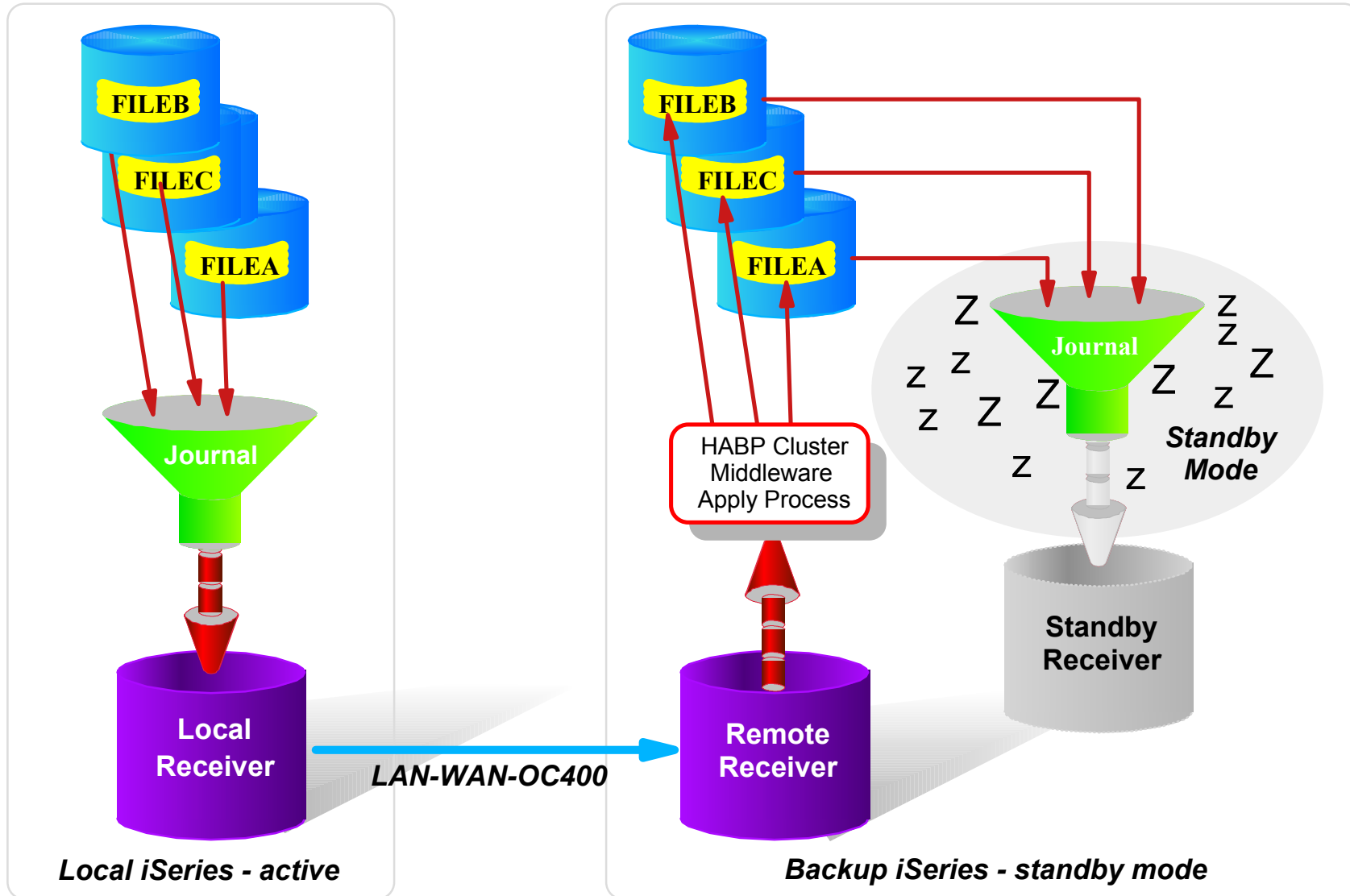
Switching out of standby mode

FILEB
FILEC
FILEA
Journal
Local Receiver
Local iSeries - off-line

LAN-WAN-OC400

FILEB
FILEC
FILEA
HABP Cluster Middleware Apply Process
Journal
Standby Mode Off
Remote Receiver
Local Receiver
Becomes active
Backup iSeries - now active

iSeries switched disk clusters

Storage Tower is switched between servers

Standard or ClusterProven Applications

OS/400 Opt. 41 HA - Switchable Resources

OS/400 Cluster Services

Cluster Management

HSL

Primary iSeries

Independent Auxiliary Storage Pools (IASP)

HSL Switchable Storage Tower

Standard or ClusterProven Applications

OS/400 Cluster Services

OS/400 Opt. 41 HA - Switchable Resources

Cluster Management

HSL

Backup iSeries

Option 41 of OS/400

- Device resiliency
  - IASP is switched between primary and backup server
    - Addresses some planned and unplanned outages
    - Not for save window reduction or disaster recovery
    - Limited distances

Clustering and IASPs for Higher Availability on the IBM eServer iSeries Server - SG24-5194

iSeries Copy Services for ESS

New!

Combining switchable iASP and copy services

iSeries Cluster & Device Domain

Primary iSeries (active)

Backup iSeries (active)

LSU

SLIC, OS/400, temporary objects, paging space, User profiles, spool files

System ASP

LSU

System ASP

iSeries Copy Services for ESS (service & toolkit)

Single ESS

Point in time copy

Independent Auxiliary Storage Pools (IASP)

Data

Flash Copy

Data
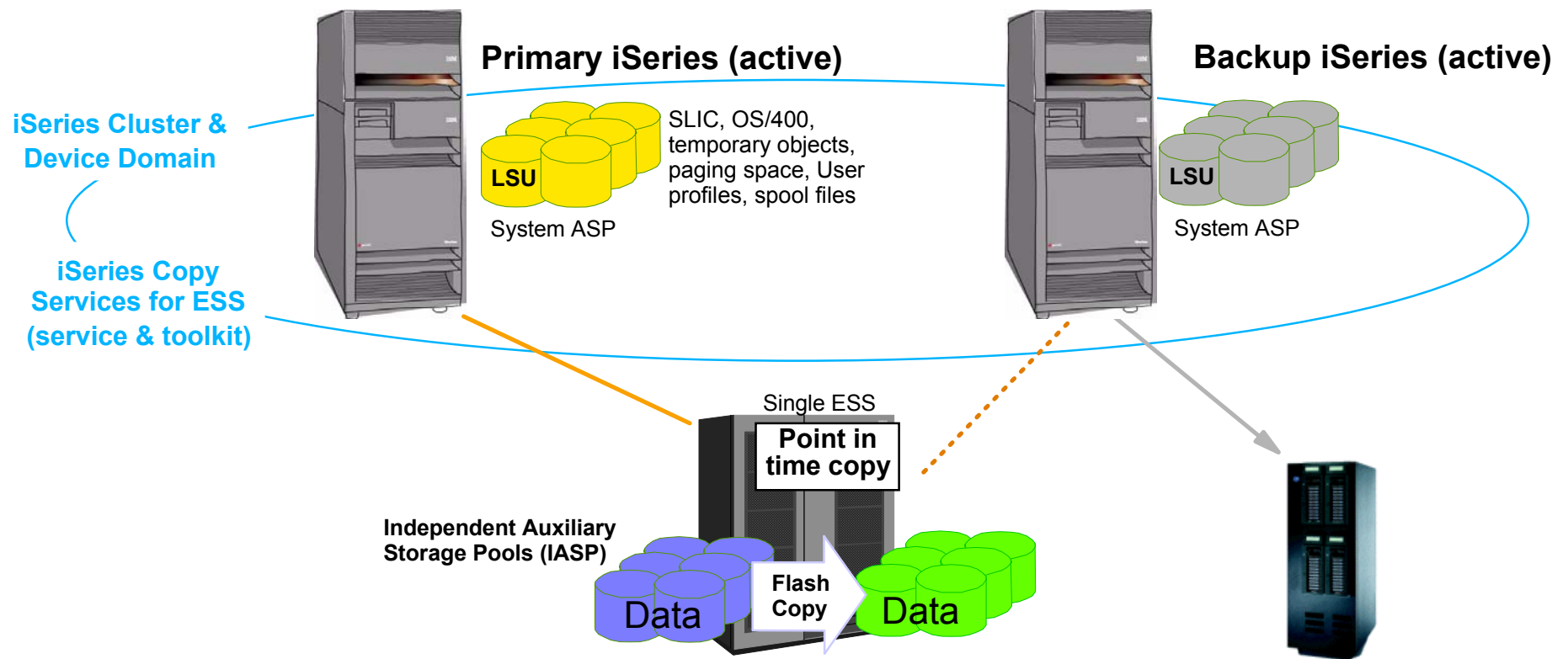
- Provides support for ESS FlashCopy in a Switched Disk Cluster
  - ► Eliminates need for IPL with FlashCopy - save to tape by IASP
  - ► 'iSeries Copy Services for ESS' is a for-fee service offering and toolkit offered by the iTC
  - ► FlashCopy only at this time, PPRC in the future

iSeries data and application resiliency clusters

Primary iSeries

Data & CPU concurrently useable on backup server

Local or remote backup iSeries (active)

Standard or ClusterProven Applications

Cluster Middleware:

OS/400 Cluster Services

Object Replication

Cluster Management & Synchronization

LAN-WAN-HSL-OC400

OS/400 Cluster Services

Standard or ClusterProven Applications

Cluster Middleware:

Object Replication

Cluster Management & Synchronization

Tape

Nightly backups using SWA, and parallel and concurrent saves

iSeries Cluster Middleware Providers:

DataMirror
VISION SOLUTIONS
LAKEVIEW TECHNOLOGY

Understanding application relisiency

Somthing happened?

Checkpoint restart

Start over

Name ___
Addr: ____
SSN: ____
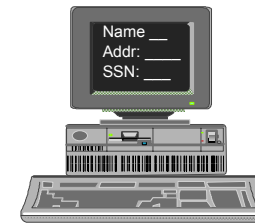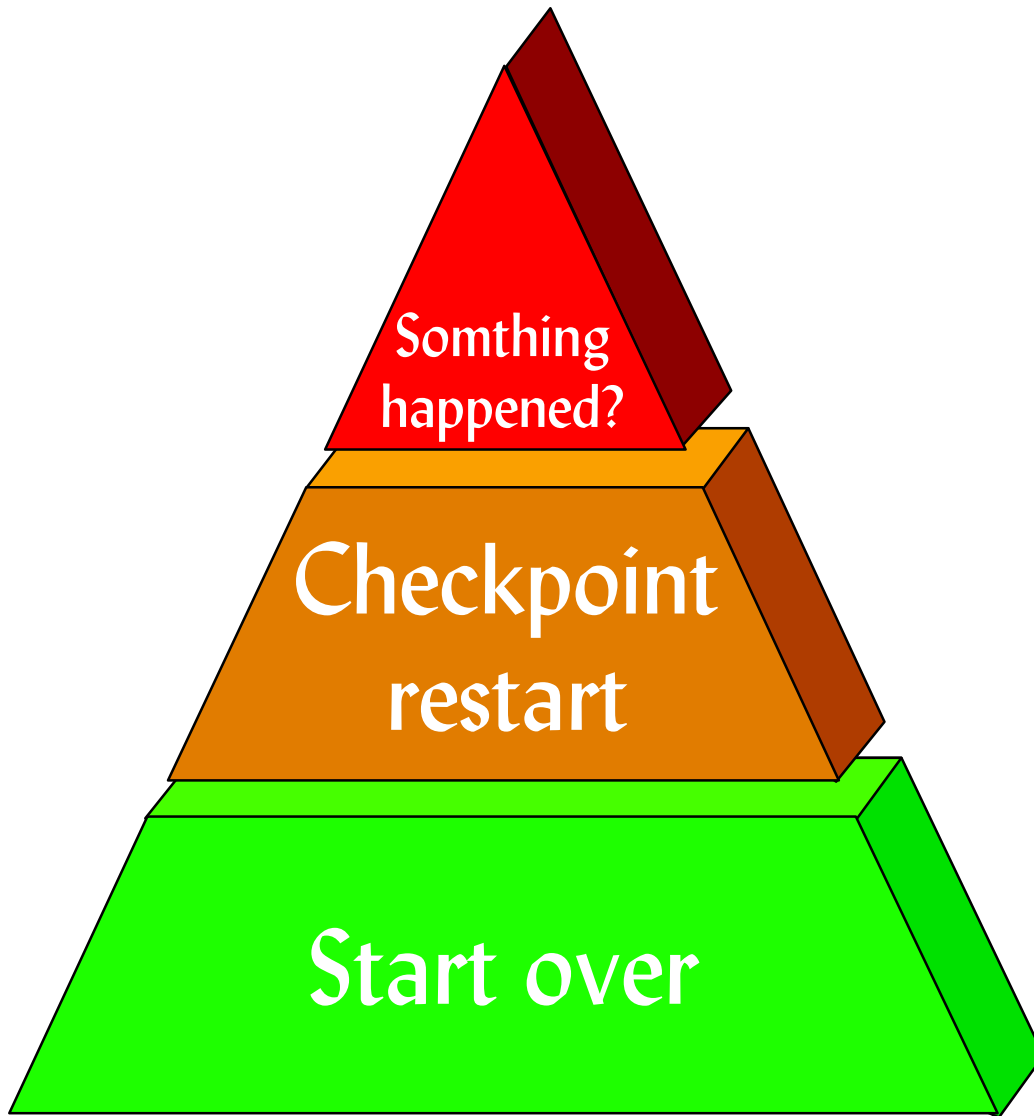
Advanced
Cluster**Proven**

Data Entry

Options
1- Insert
2- delete

Application Options

Application
- Create
- Modify
- Status

Cluster**Proven**
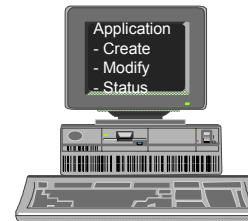
Application Menu
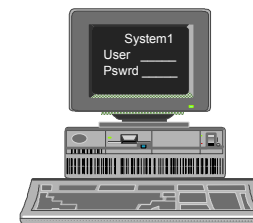
System1
User _____
Pswrd _____

System Sign-on

Cluster resource services to aid application resiliency

- Recovery domains via cluster resource groups (CRGs)

- Exit programs

- IP address takeover

- Cluster heartbeat every 3 seconds (tunable) with redundant path support

- Broadcast distress message
  - ► Panel actions
    - – White button IPL power off
  - ► Utility power loss
    - – Battery backup (BBU)
    - – UPS
  - ► System commands
    - – PWRDWNSYS *immed
    - – ENDSYS, ENDSBS *ALL
    - – ENDTCP
    - – ENDSBS QSYSWRK

message →
acknowledge

message →
acknowledge

message →
acknowledge

message →
acknowledge

message →
acknowledge
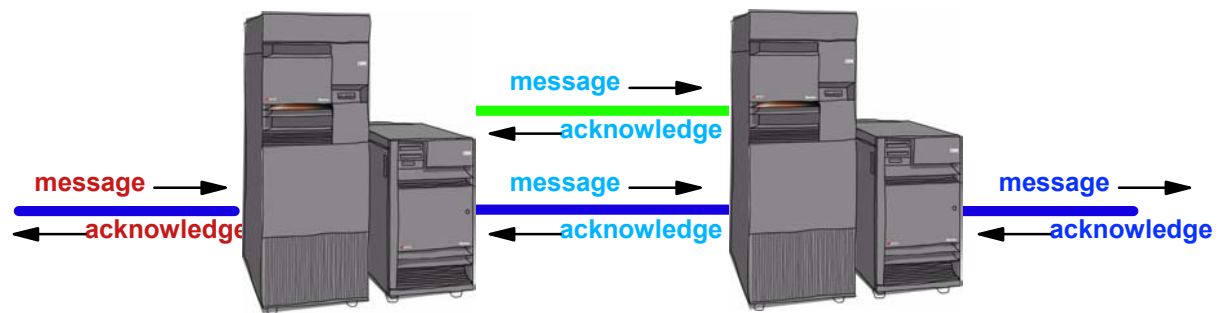
Protect critical transactions

- Commitment control
  - ► Ensures all transaction changes are complete for all files affected
    - – Eliminates partial database updates in the event of a failure

  - ► Incomplete transactions can be rolled back
    - – Manually or by the application program or manually
    - – Or during the IPL following abnormal termination

  - ► Applicable to a single system and high availability dual systems
  - ► Most effective means of recovering to a known point

program

transaction

file write/update

file A   file B   file C
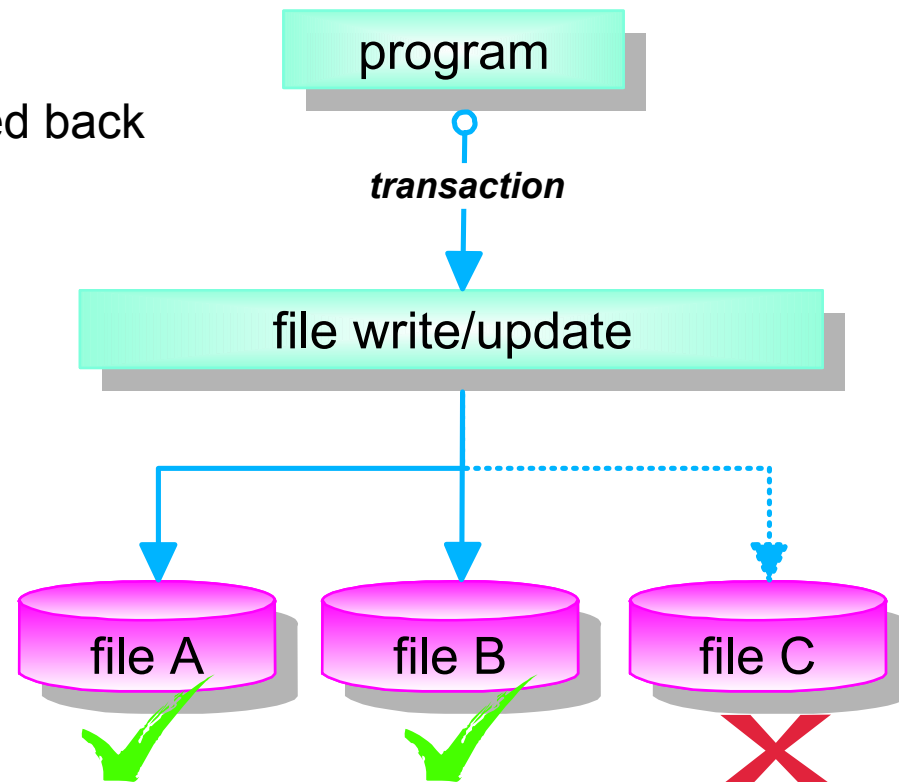
Clustering summary

- **You have choices!**
  - ► Data resiliency through replication or switched disk
  - ► Application resiliency via ClusterProven or cluster enabled

- **Solutions are neither 'shrink-wrap' nor 'plug-and-play'**
  - ► Apply resources and project management processes as you would to a new business critical application

- **The best clustering implementations are a team effort, including all IT infrastructure**

- **Some applications perform better than others in the clustering environment**
  - ► Learn which ones you have
  - ► Invest in application and operations tips and techniques, which can make applications work better with clustering

- **Invest early in education and consulting**
  - ► Both Business Partners and IBM Global Services supply these services

Meeting various needs

(examples)

Tradional example

- Planned and unplanned outages

- Nightly backups

- Read only workload

- DR

HABP Cluster Replication
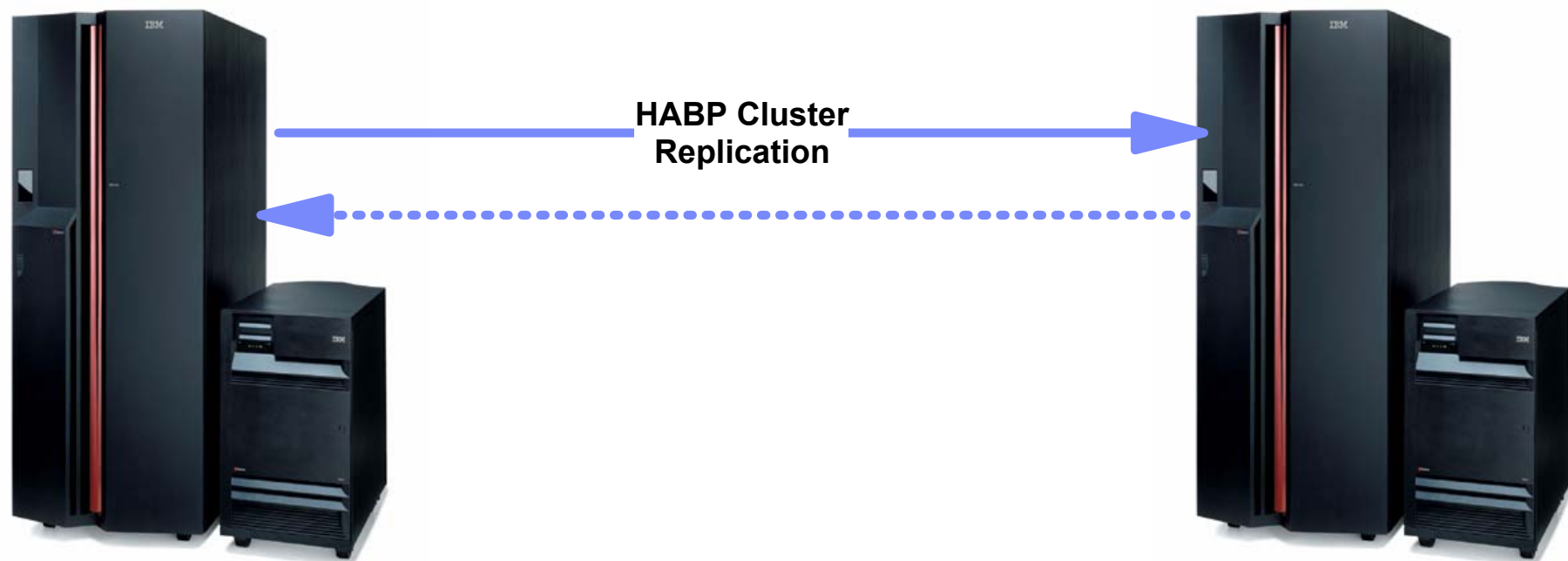
Addressing backups, development and remote DR

- Separate production from development and testing
  - ▶ Release, PTF, application testing and development

- Non-disruptive tape backups

- Remote disaster recovery with CuOD

Remote 6-way iSeries

SYS3

Remote hot backup for disaster recovery

OS/400
PPPPPP

Local 8-way iSeries

SYS1	SYS2
Test/Backup environment	**Production environment**

HABP Cluster

HABP Cluster

OS/400	**OS/400**
P **PP**	**PPPPP**

Primary partition

Enterprise class, flexible and scalable
- ▶ *Shared processor support*
- ▶ *Dynamic resource movement*
- ▶ *Up to 32 partitions*
- ▶ *Virtual Ethernet LAN*
- ▶ *Graphical management*
- ▶ *New APIs provide custom automated management*

Notes: Do it in-house, or use a business continuity provider?

Highly successful companies recognize the value of a technology solutions provider that can help plan, implement and manage an ongoing business continuity program. As client/server computing became more widespread in the early 1990s, and the price/performance ratio for computing and storage continued to improve, many companies implemented in-house data mirroring, redundant storage arrays and other high-availability techniques to create duplicate online or near-online copies of data.

Although these strategies can provide virtually continuous availability of data at very attractive price points, enterprises that test and validate their ability to recover from an outage have learned that business continuity encompasses many more challenges. To provide true continuity for critical business processes—not just critical data—companies using the in-house approach must also:

- Ensure that sufficient latent capacity will be immediately available to assure rapid failover and recovery

- Test capacity availability without disrupting ongoing operations

- Install redundant network capacity dedicated to business continuity
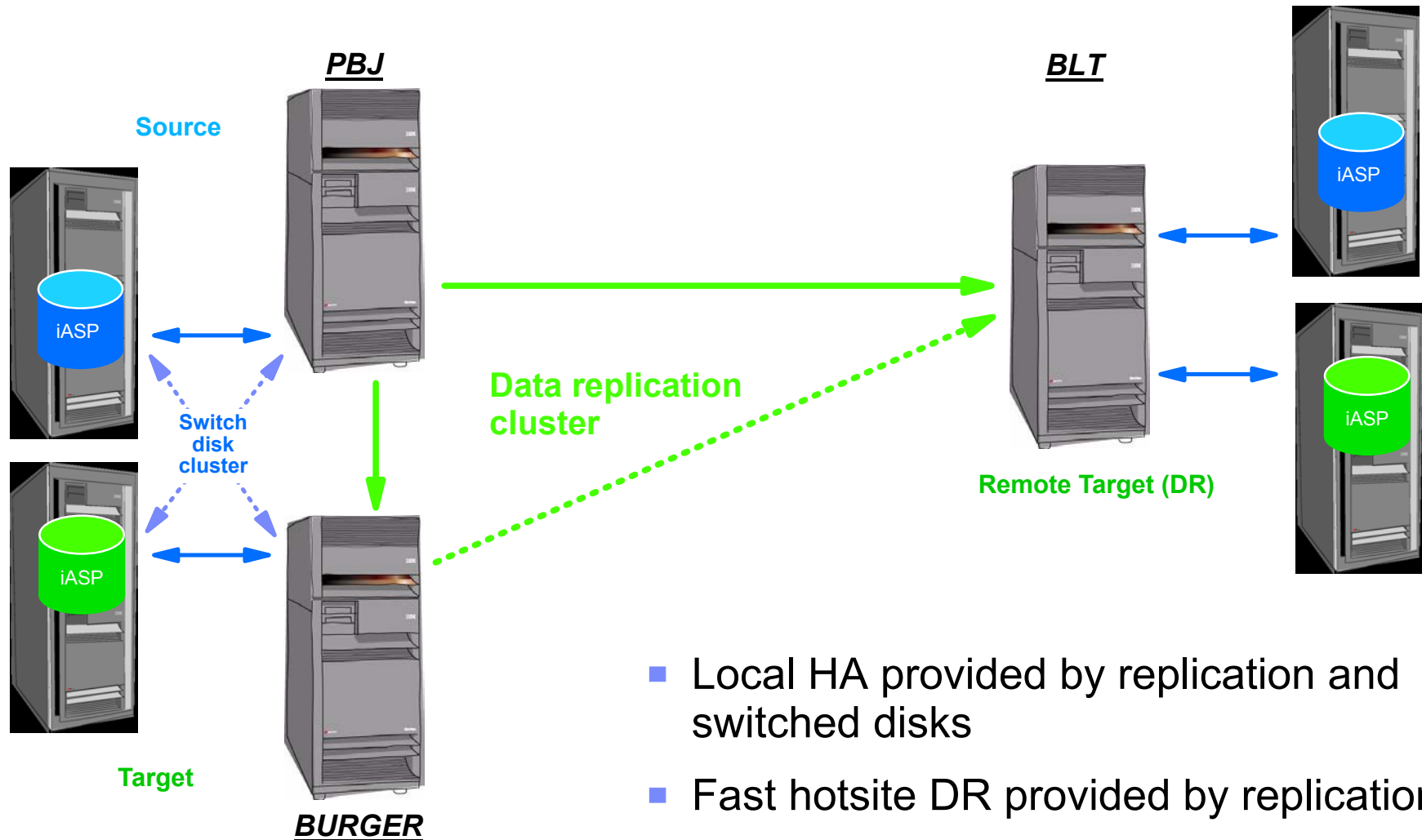
- House failover equipment in a separate location from the main production equipment and provide further redundancies, such as sourcing electrical supplies from different power grids

- Establish and maintain relationships with vendors to assure quick delivery of replacement PCs, network hardware, desks, chairs, telephones, etc., in the event of a facilitywide disaster

- Secure adequate funding from end-user departments to implement and maintain adequate critical business continuity protection Acquire, train and retain skilled personnel who can manage the complex interdependencies and specialized elements of business continuity Make adequate provisions for adding recovery support staff in the event of a regional or natural disaster.

Cluster using both replication and switched disks

PBJ

Source

iASP

Switch disk cluster

iASP

Target

BURGER

Data replication cluster

BLT

iASP

Remote Target (DR)

iASP

- Local HA provided by replication and switched disks
- Fast hotsite DR provided by replication

Cluster using replication, switched disks, FlashCopy

Source   **_PBJ_**

DR Site

HABP Cluster Replication
(Remote Application DR)

HABP Cluster Replication (Sysbase)

BLT   **Target**

Sysbase

Sysbase

Sysbase

Switch disk cluster

Sysbase

I/O Tower
(with fibre
channel)

Backup
(HA)

Mirrored
SYSBASE
(ASP1)

Active
(Production)

Backup
(saves)

Point in
time copy

**Independent
Auxiliary
Storage Pools
(IASP)**

1. Local HA provided by IASPs,
   Switched Disks and HABP

   ► User profile and spool file
      replication via HABP solution

2. Backups provided by 'iSeries
   Copy Services for ESS' and
   FlashCopy

3. DR provided by HABP

Flash
Copy

Data   Data

Single ESS

Local or campus
secondary ESS

Backup iSeries
(inactive)

Solution summary

- When to use iSeries HA functions
  - ► Requirement for quick failover or switchover
  - ► Transaction level integrity is important
  - ► Coverage for all outage types is required (planned, unplanned and local or remote disaster recovery)

- When to consider external SAN copy functions
  - ► When used in conjunction with iSeries Clusters
  - ► For basic, local DR when longer down times are acceptable
  - ► For basic tape backups, when longer downtime is acceptable
  - ► External disk copy functions are place for other platforms and you don't require the iSeries HA functions listed above

Event Summary

iSeries availability summary

- **The iSeries server provides:**
  - ► Resilient hardware and software
  - ► Integrated systems management
  - ► Integrated storage management for consolidated, multi-platform environments

- **iSeries availability solutions are:**
  - ► Scalable - able to grow with your business
  - ► Flexible - granular
    - – Address planned, unplanned, and disaster/recovery
    - – Provide additional workload capacity
    - – Supports different OS release levels and hardware configurations

- **iSeries storage solutions provide:**
  - ► Multi-server storage consolidation
  - ► ESS Copy Services for local DR and tape backups

Capacity on Demand

ESS and SAN

Multi-server solutions

Integrated storage Mgt.

iSeries Navigator

BRMS/400

iSeries Clusters

Switch Disks

High speed tape

ESS Copy Services

Independent ASPs

xSeries Clusters

IBM BCRS

iSeries LPAR

iSeries availability summary

- Choices for:
  - ► High availability choices:
    - Replication clusters
    - Switched disk clusters
    - Native features and functions
  - ► Disaster recovery choices:
    - Outsourcing - turnkey, faster implementation, requires less business resource commitment
      - ◆ Recovery time longer than an in-house solution
    - In-house - fastest recovery time from a disaster
      - ◆ Custom, longer implementation - requires more resource commitment
    - DR Technologies include:
      - Replication clusters
      - ESS copy services (e.g. PPRC)
      - Business Continuity and Recovery Services (BCRS)

Capacity on Demand

ESS and SAN

Multi-server solutions

Integrated storage Mgt.

iSeries Navigator

BRMS/400

iSeries Clusters

Switch Disks

High speed tape

ESS Copy Services

Independent ASPs

xSeries Clusters

IBM BCRS

iSeries LPAR

What to do next

- Hold a "Simulated Disaster" to help determine hidden costs and recovery times and issues
  - ► You may find "hidden" affects on different departments
  - ► An untested or undocumented plan rarely works as intended in a real emergency

- Determine your business availability requirements
  - ► Cost of downtime by application
  - ► Cost of downtime by organization

- Match your findings to one of the iSeries solutions available to you

- Recommendations
  - ► Use experienced consulting services for all solutions
  - ► Obtain and use solution references
  - ► Staff appropriately
  - ► Educate and maintain your staff's skills
  - ► Test, test, and test the solution !

The nd, Thank You!

Appendix

Trademarks and Disclaimers

© IBM Corporation 1994-2003. All rights reserved.
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:

AS/400	IBM
AS/400e	IBM (logo)
eServer	iSeries
*e*server	OS/400

Lotus and SmartSuite are trademarks of Lotus Development Corporation and/or IBM Corporation in the United States, other countries, or both.

MMX, Pentium, and ProShare are trademarks or registered trademarks of Intel Corporation in the United States, other countries, or both.
Microsoft and Windows NT are registered trademarks of Microsoft Corporation in the United States, other countries, or both.
Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
SET and the SET Logo are trademarks owned by SET Secure Electronic Transaction LLC.
C-bus is a trademark of Corollary, Inc. in the United States, other countries, or both.
UNIX is a registered trademark of The Open Group in the United States and other countries.
Other company, product or 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.

About this IBM presentation

Some of the content of this presentation represents a good faith effort to communicate the current investment and development direction for iSeries products and functions. As IBM may change or cancel its plans at any time, this content is not to be used as an implied or stated commitment to performance or function.

IBM may have patents or pending patent application covering subject matter in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to the IBM Director of Licensing, IBM Corporation, 208 Harbor Drive, Stanford, CT 06904, USA.

The information contained in this document has not been submitted to any formal IBM test and is distributed AS IS. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

Any performance data contained in this document was determined in a controlled environment, and therefore, the results that may be obtained in other operating environments may vary significantly. Users of this document should verify the applicable data for their specific environment.

Notes: About the speakers

Sue Baker

e-mail: smbaker@us.ibm.com

Sue Baker has spent over 20 years working with IBM midrange systems. Over the years, she has worked with customers in the industries of manufacturing, transportation, distribution, education, and telecommunications. While working with these customers, Sue has assisted them in developing solutions for capacity and performance issues, high availability requirements, systems management/operations, and programmer productivity environments. Sue currently focuses on assisting customers with developing and implementing the appropriate performance, capacity planning, and operations management techniques that are needed in the more complex multiple system and/or high availability environments.

Eric Hess

e-mail: texas@us.ibm.com

Eric has been with IBM for over 20 years working with the System/38, AS/400 and iSeries, specializing in availability and recovery, and systems management. He is a speaker at the iSeries Technical Conference, various user groups meetings and iSeries availability planning seminars around the country . Eric currently focuses on helping customers to enhance their availability, improve their save/restore strategies and improve their systems management practices.

Cluster Middleware Partners

**3100 Steeles Ave E, suite 700
Markham, Ontario, Canada L3R 8T3
800-362-5955 or 905-415-0310
web page: http://www.datamirror.com
e-mail: info@datamirror.com**

**2301 W 22 St, suite 206
Oakbrook, IL 60523
630-537-0440
web page: http://www.lakeviewtech.com
e-mail: info@lakeviewtech.com**

**17911 Von Karman Ave, 5th floor
Irvine, CA 92614
800-683-4667 or 949-253-6500
web page: http://www.visionsolutions.com
e-mail: info@visionsolutions.com**