# T09

# IBM® TotalStorage® Enterprise Storage Server (ESS)
## Storage Optimized for an on Demand World

**Jim Tuckwell**

**Marketing Program Director**

**Systems Technology Group**

IBM **@server** xSeries
Technical Conference

Aug. 9 - 13, 2004

**Chicago, IL**

*With information on demand, businesses can respond with flexibility and speed to any customer requirement, market opportunity, or external threat*

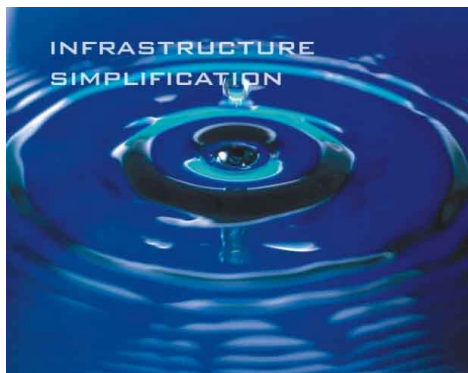**Getting there involves:**

**1.** Simplification **of the underlying IT infrastructure and its management are required to support the changes in the business and lower cost and complexity**

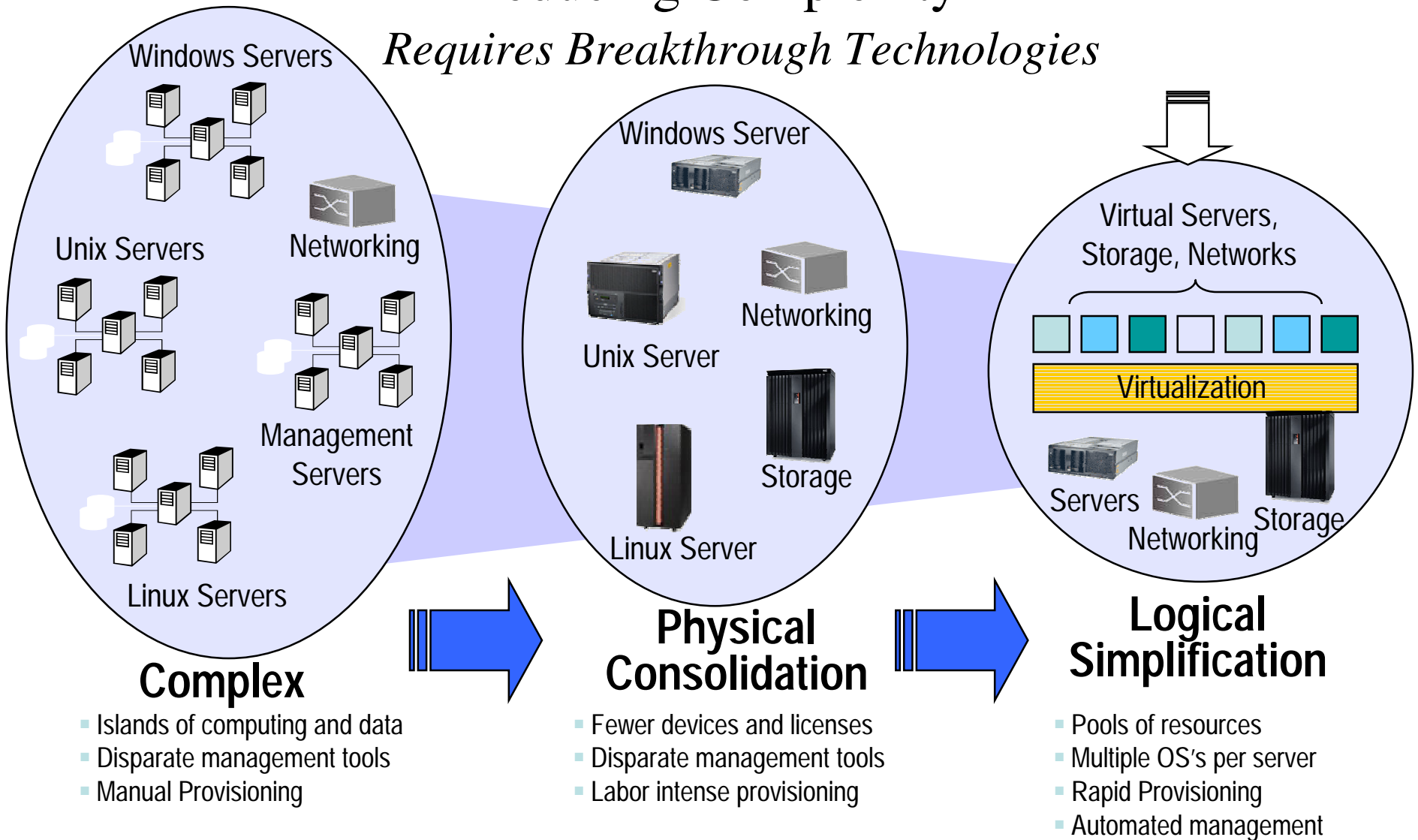**2.** **Assuring** business continuity**, security and data durability**
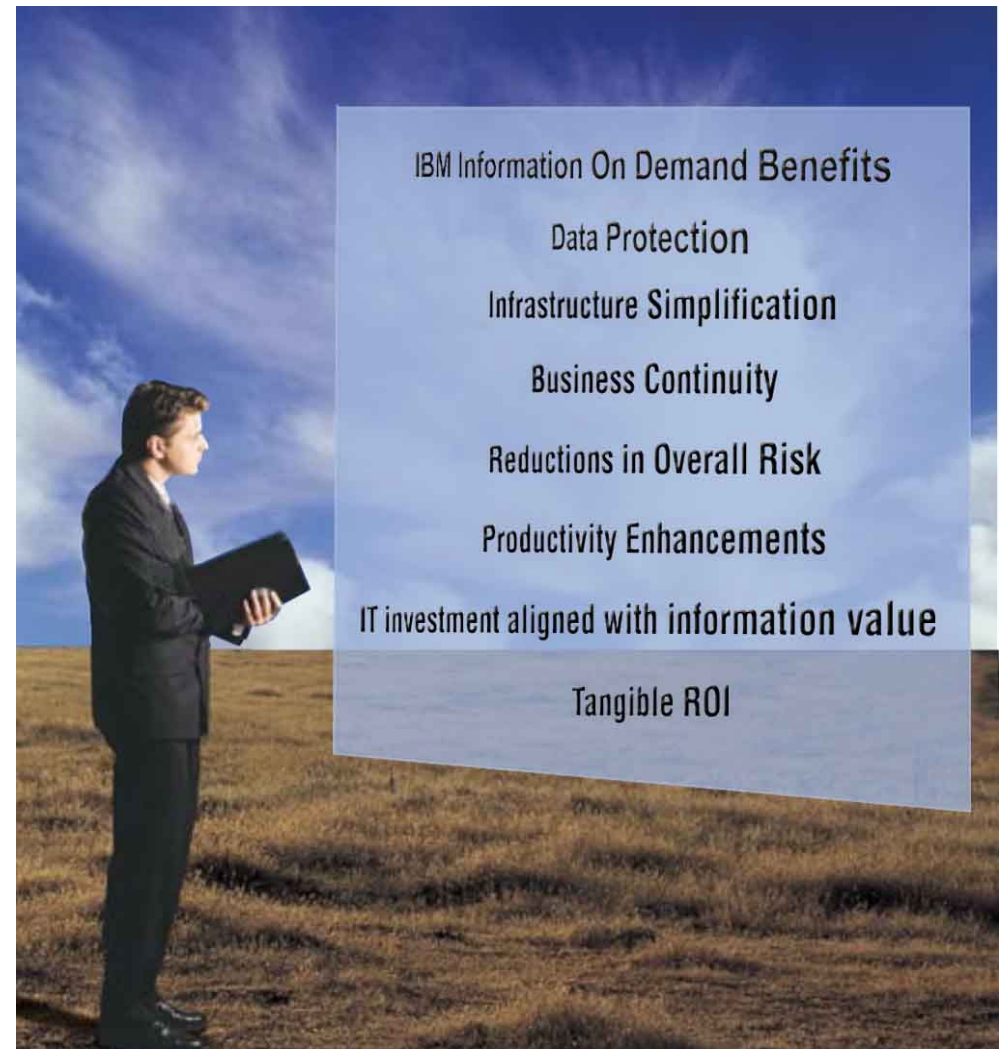
**3.** **Efficiently managing information** over its lifecycle

# Reducing Complexity
## *Requires Breakthrough Technologies*

### Complex

- Islands of computing and data
- Disparate management tools
- Manual Provisioning

**Windows Servers**

**Unix Servers**

**Networking**

**Management Servers**

**Linux Servers**

### Physical Consolidation

- Fewer devices and licenses
- Disparate management tools
- Labor intense provisioning

**Windows Server**

**Unix Server**

**Networking**

**Storage**

**Linux Server**

### Logical Simplification

- Pools of resources
- Multiple OS's per server
- Rapid Provisioning
- Automated management

**Virtual Servers, Storage, Networks**

**Virtualization**

**Servers**
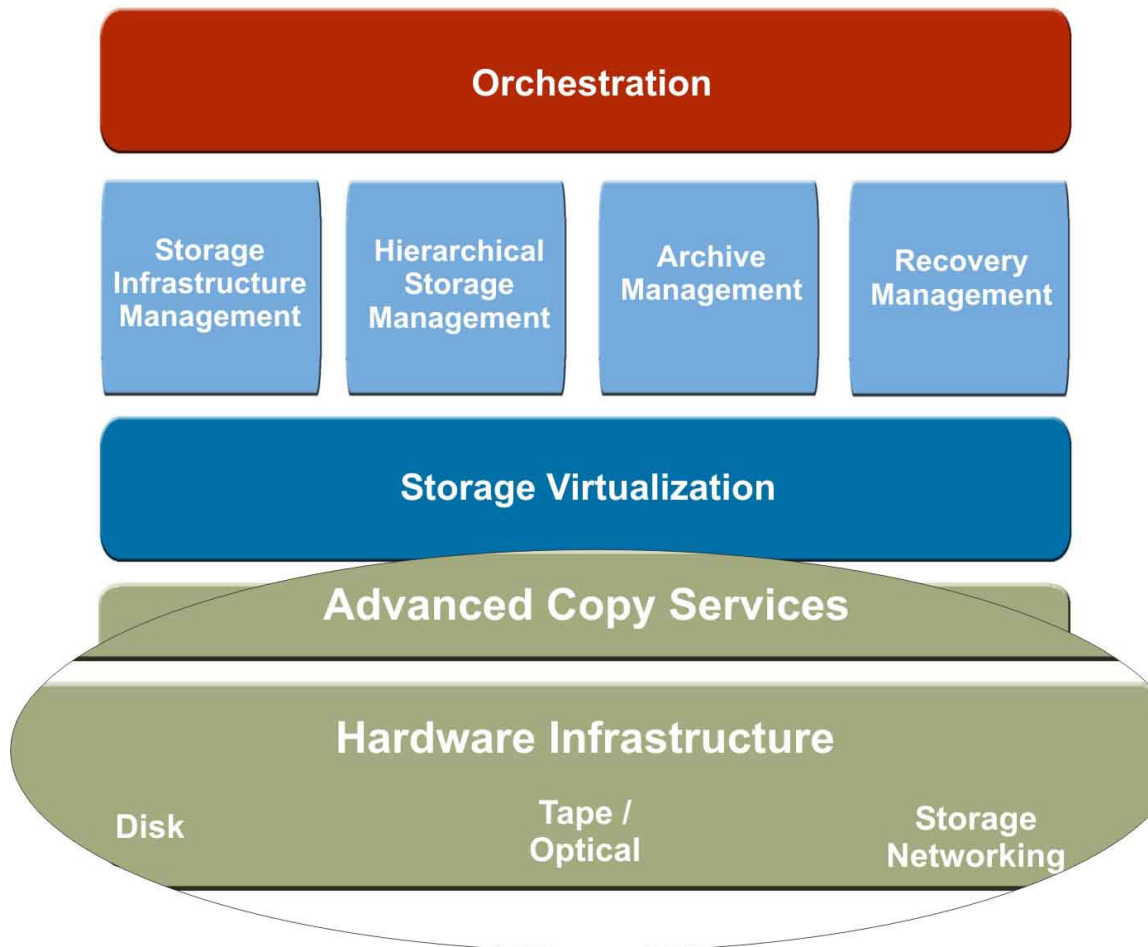
**Networking**

**Storage**

# IBM provides the industry's most complete set of capabilities for Information On Demand

- Based on industry standards

- A broad set of capabilities that is modular in design

- Supports heterogeneous environments

- Built on innovative technology

- Enables end to end solutions

*At the heart of IBM's strategy is a <u>simplified</u>, and <u>resilient</u> TotalStorage environment that helps you <u>efficiently manage information over its Lifecycle</u>.*

**Orchestration**

| Storage Infrastructure Management | Hierarchical Storage Management | Archive Management | Recovery Management |

**Storage Virtualization**

**Advanced Copy Services**

**Hardware Infrastructure**

Disk    Tape / Optical    Storage Networking

**Advanced Copy Services**
- *Flashcopy*
- *Global and Metro Mirror for ESS, FAStT, and VTS*

**Hardware Infrastructure**
- **Enterprise disk**
  - **ESS 750, 800 , 800turbo**
- *Mid-range disk with FC or S-ATA*
  - *FAStT100, FAStT600, FAStT700, FAStT900*
- *Enterprise tape*
  - *3592, 3590, 3494, 3494-VTS*
- *Midmarket tape*
  - *3580, 3581, 3582, 3583, 3584*
- *Storage Networking*
  - *SAN Partnerships*
    - *Cisco, McDATA, CNT, Brocade*
  - *IP Attached*
    - *NAS Gateway 500*

# ESS Model 750 & 800 – Dual Server Based Architecture

- "5 – 9s" Availability
- 2 Gb Fibre Channel / FICON®
- RAI D-5 and/or RAID-10
- Mirrored Write Cache
- SPC-1 Published Performance
- Standard zSeries® Optimization features
- Advanced copy services

- **Competitive Advantages**
  - Designed to support 24 X 7 Operations
    - Designed to avoid single points of Failure/Repair
  - 64 non-arbitrated, pipelined paths to disks
  - 20 + Autonomic features designed to offer excellent efficiency
  - Imbedded flexibility offers optimal solution design
  - Asset Protection - Non disruptive upgrade design
    - Improved application availability
    - Improved TCO/ROI
  - Designed to deliver low Total Cost of Ownership
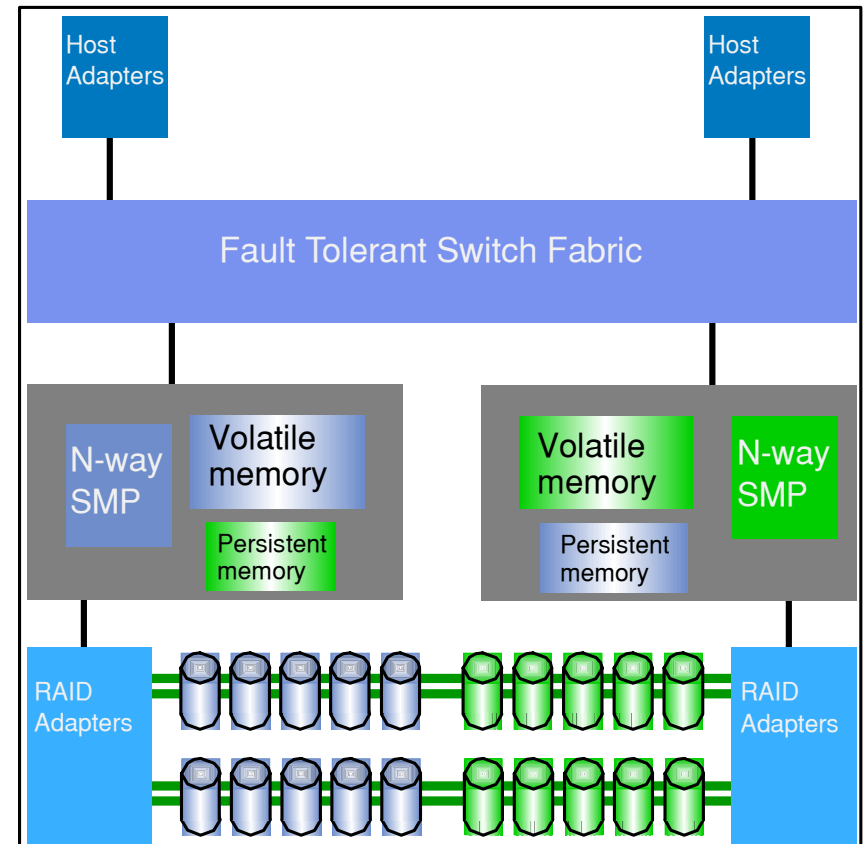
# ESS…

## Outstanding Price/Performance, TCO and continually enhanced value

### Key Features:

- Superior price performance - up to 3.0Xthroughput of previous IBM models
- Designed to support continuous operations
- Investment protection with advanced functions, designed to support non-disruptive upgrades
- Advanced long distance copy services for business continuance
- Expandable from 582GB up to 55.9TB physical capacity in base and turbo options
- Integrated storage solutions
- Low TCO

- Server Based Architecture
- Third Generation in ESS evolution
- Designed to help prevent single point of failure/repair
- Redundant, "failoverable" hardware
- 64 non arbitrated paths to disk
- Base (4-Way) and turbo (6-Way) Processor
- 18.2, 36.4, 72.8 GB (10K/15K rpm), 145.6 GB (10Krpm) physical HDDs
- Intermixable disk capacities/speeds
- RAID-5, RAID-10 (intermixable)
- Support for z/OS®, S/390®, iSeries®, Linux, Unix, AIX®, Windows® NT, Windows 2000®
- Browser, command line and open (ESS API) management options
- Attachment via ESCON, FICON ™, Fibre Channel, 2 Gigabit FC/FICON and Ultra SCSI
- FlashCopy® V1 and V2 / NOCOPY point in time copy efficiencies
- PPRC V1, V2 and XRC disaster recovery solutions for open and z/OS
- Standby Capacity on demand features

### February 2003 Enhancements:
**Enhanced open solutions and performance**

- ESS API ( SMI-S) open interface
- zSeries® Linux Enhancements
  - FICON
  - Copy Services
  - Additional Linux Distributions
- SGI Origin Services (IRIX)
- 15K rpm, 72.8 GB physical capacity HDD

### May 2003 Enhancements:
**Enhanced Business Continuance Solutions and performance**

- FlashCopy V2 - Data Set FlashCopy, Multiple Relationships, Significant Performance Boost
- PPRC V2 - Asynchronous Cascading PPRC

### October 2003 Enhancements:
**Enhanced Business Continuance, Open API and Performance**

- PPRC over Fibre Channel
- ESS API - Copy Services Management
- Arrays Across Loops
- Turbo II

### April 2004 Enhancements:
**Enhanced Flexibility, Resiliency**

- ESS Model 750
  - ESS functionality and mid range pricing
  - Non-disruptive growth to ESS 800
- Global Mirro
  - Dramatically improved global mirroring functionality, flexibility and TCO

# IBM ESS Technology Leadership

## ESS exhibits up to 4+ year shipment lead time on these items

| Item | IBM ESS | EMC Symmetrix | HDS Lightning |
|---|---|---|---|
| 15K RPM disks | 2q2002 (18 and 36GB) 3q2003 (73GB) | 2004 ( only 73GB disks) | 4q2003 (73GB disks) |
| true RAID-5 (striped data and parity) | 1999 (1994 in RAMAC) | 2004 | supported |
| copy-on-first-write volume replication optimization | 2000 | 2003 (EMC Snap) | not announced |
| concurrent logical config changes under user control | 1999 | still being augmented as of 2004 | available as of? |
| SMI-S support | 1h2003 | 2004 | available as of ? |
| nondisruptive model upgrades within family | 2q2004 | not announced | not announced |
| continuous asynch remote copy with m to n cross system self-managed congroups | 2q2004 | not announced | not announced |
| iSeries - Fibre channel support | 2q2001 | 3q2003 in DMX (after period of no iSeries support) | iSeries not supported |
| Mainframe items: (IBM licenses *specifications* after ESS delivers the *function*) | | | |
| PAV (volume I/O parallelism) | 1999 (static and dynamic; 256 concurrent I/Os implemented) | dynamic=2q2001 (8 concurrent I/Os implemented) | licensed to Hitachi in 2001 (16 concurrent I/Os implemented) |
| Priority I/O Queuing | 1999 | not announced | not announced |
| XRC | 1q2000 (initially available 199?) | EXRC announced 2003; delivery=? | supported (HXRC) |
| HyperSwap | 3q2002 | not announced | available as of ? |
| FICON 1Gb/s | 3q2001 | 1q2002 | 1q2002 |
| FICON 2Gb/s | 3q2002 | 2h2003 | available as of ? |
| CUIR (availability item) | 4q2001 | not announced | not announced |

# ESS Family of Enterprise Storage Servers

- **ESS Model 750**
  - For moderate capacity needs - supports up to 4.6 TB
  - Designed to meet lower tier customer workload requirements
  - Designed to offer attractive pricing for modest capacity requirements
  - Designed to offer non-disruptive path to ESS Model 800

- **ESS Model 800**
  - Excellent capacity scalability supporting up to 55.9 TB
  - High performance scalability designed to meet/exceed the vast majority of customer workload requirements
  - Wide array of disk capacities and speeds

- **ESS Model 800 with Turbo option**
  - Up to 30% greater throughput than ESS base model 800
  - Designed to meet the needs of very high performance environments
  - Designed to offer non-disruptive Field upgradeability from ESS Model 800 Base

## Included in all ESS Models

Designed to avoid single points of Failure/Repair
Designed to support 24X7 operations
Offers flexibility in efficient management options
Wide array of heterogeneous host attachments

20+ autonomic functions
Rich set of optional advanced functions
Designed to offer low total cost of ownership

# What Customers say about ESS

*"The Enterprise Storage Server is based on what I consider to be the most powerful and stable processor on the market, the four-way RISC symmetric multiprocessor (SMP).  You cannot find a better, faster, more reliable engine for a storage server...We now have a flexible, rock-solid central storage solution that can be expanded as needed."*

Wally Knapp, Director of IT, Community College of Baltimore County

*"We chose the ESS because it is a very solid architecture:  built to perform, built to last."*
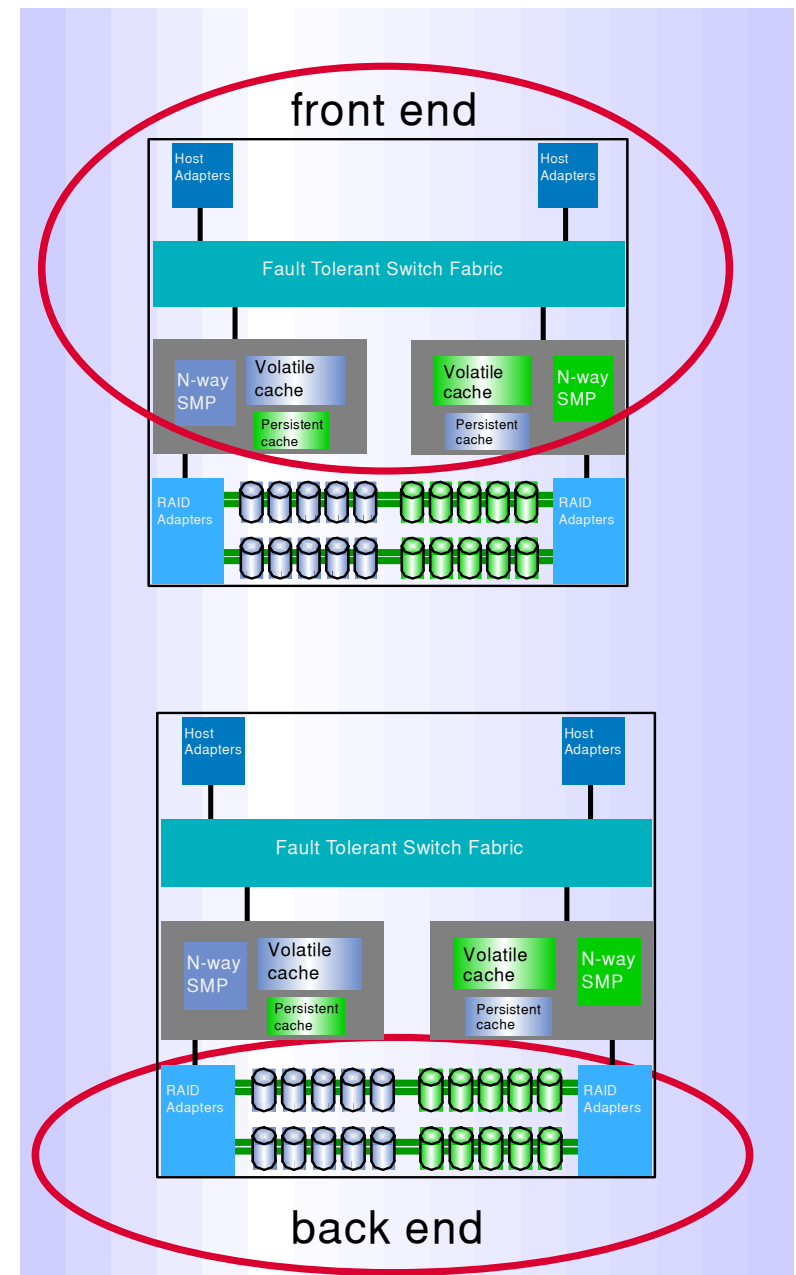
Chad Carter, technical services manager, Grand River Hospital

*"We chose the IBM ESS storage system to be the cornerstone of our data consolidation project because it offers the performance, scalability and advanced functionality we need to manage and protect the availability of our customers' assets."*

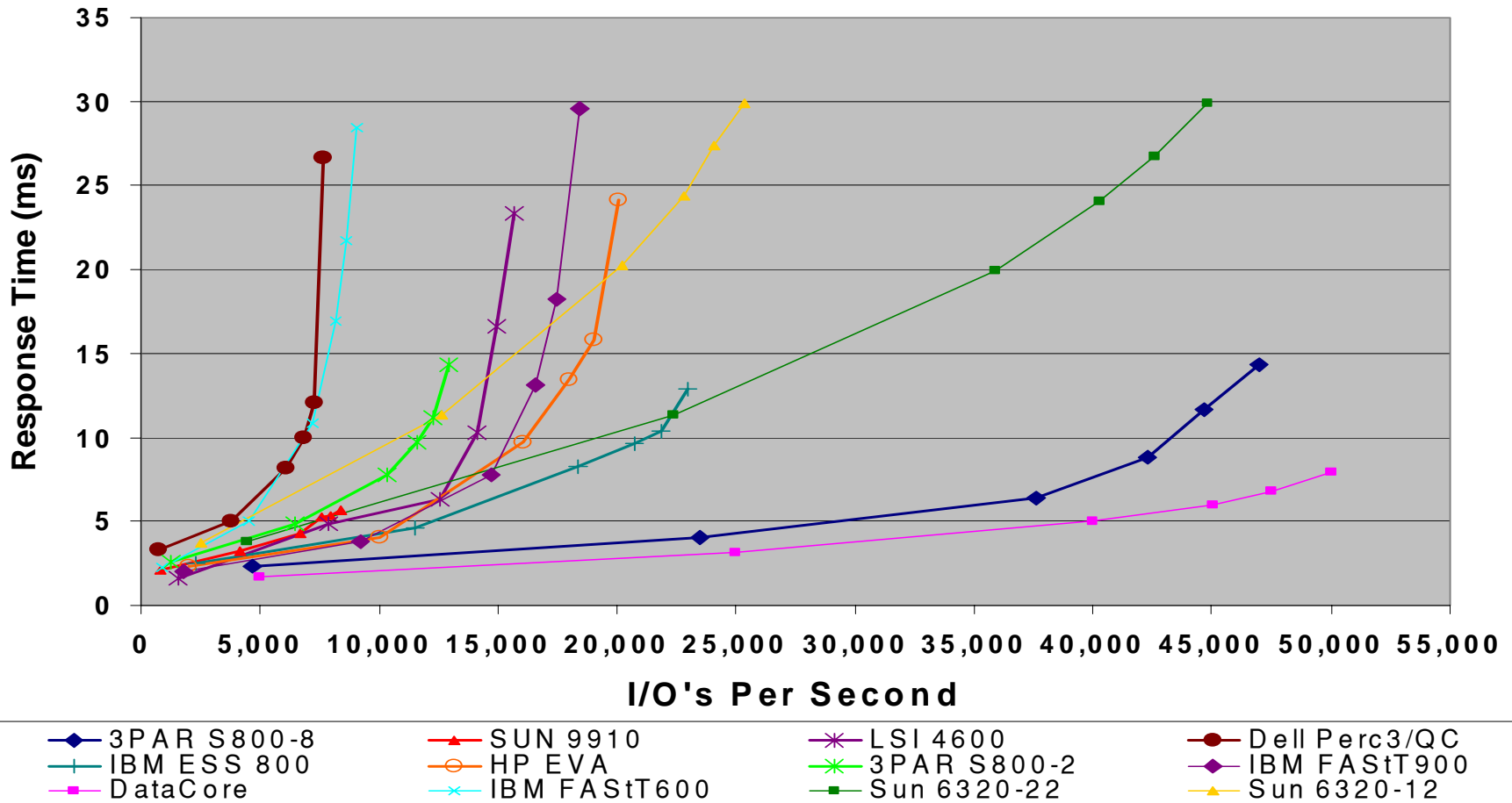Peter Kraemer, Global head of IT Production, Commerzbank

# Self-Optimizing Performance

- No restrictions on concurrent path xfer

- Optimized cache (self-tuning, efficient space allocation)

- Dual active clustered servers
  - Dual n-way SMPs, Turbo II accelerator option

- Server multi-path I/O load balancing

- High-speed disks (10/15K RPMs)

- Efficient data paths to disks
  - 64 paths, pipelined and unarbitrated, 2.56GB/s bandwidth

- Stripe-all design (automated and granular)

- High-performance disk adapters
  - RAID and sparing I/O traffic offloaded from rest of the system

- zSeries accelerators (PAV/MA, Priority I/O Queuing, channel commands)



front end
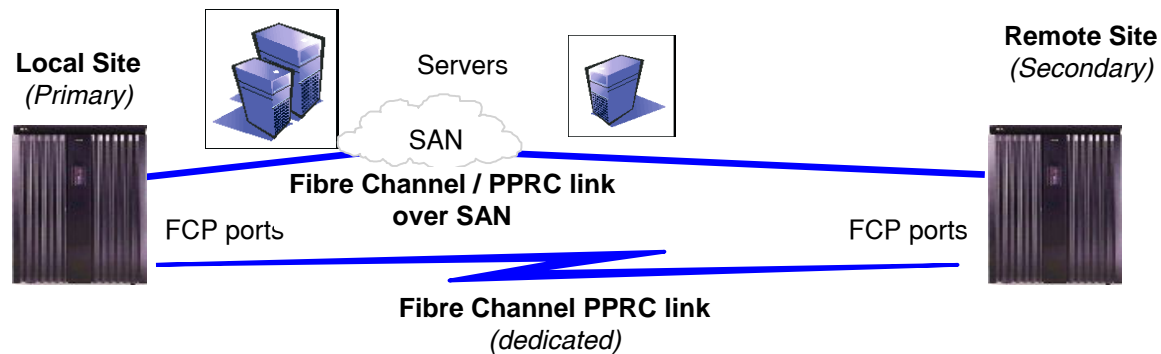
Host Adapters    Host Adapters
Fault Tolerant Switch Fabric
N-way SMP    Volatile cache    Volatile cache    N-way SMP
Persistent cache    Persistent cache
RAID Adapters    RAID Adapters

Host Adapters    Host Adapters
Fault Tolerant Switch Fabric
N-way SMP    Volatile cache    Volatile cache    N-way SMP
Persistent cache    Persistent cache
RAID Adapters    RAID Adapters

back end

# SPC-1 Results



**SPC Posted Benchmark Results 09/16/03**
(mirrored cache results)

Y-axis: Response Time (ms) — 0, 5, 10, 15, 20, 25, 30, 35

X-axis: I/O's Per Second — 0, 5,000, 10,000, 15,000, 20,000, 25,000, 30,000, 35,000, 40,000, 45,000, 50,000, 55,000

Legend:
- 3PAR S800-8
- SUN 9910
- LSI 4600
- Dell Perc3/QC
- IBM ESS 800
- HP EVA
- 3PAR S800-2
- IBM FAStT900
- DataCore
- IBM FAStT600
- Sun 6320-22
- Sun 6320-12

Source: Randy Kerns, Evaluator Group

# ESS Fibre Channel Mirroring

- ## ESS mirroring solutions can use 2 Gbit Fibre Channel (FCP) links for disk mirroring between ESS systems
    - Support included in existing PPRC V2 licenses at no additional charge
    - Supported on ESS Models 750, 800, and 800turbo



**Local Site**
*(Primary)*

Servers

SAN

**Remote Site**
*(Secondary)*

**Fibre Channel / PPRC link over SAN**

FCP ports

FCP ports

**Fibre Channel PPRC link**
*(dedicated)*

- ## Very High Performance, High Capacity Disk Mirroring
    - Utilizes existing ESS M800 Fibre Channel adapters
    - Up to 8x speed and throughput improvement *per link* as compared to ESCON
        - Only two Fibre Channel links per ESS are needed for most customer workloads
        - Competing alternatives may require four or more Fibre Channels per equivalent subsystem TBs
        - Minimum 4 to 1 collapse for migration of ESS PPRC ESCON to PPRC Fibre Channel disk mirroring links
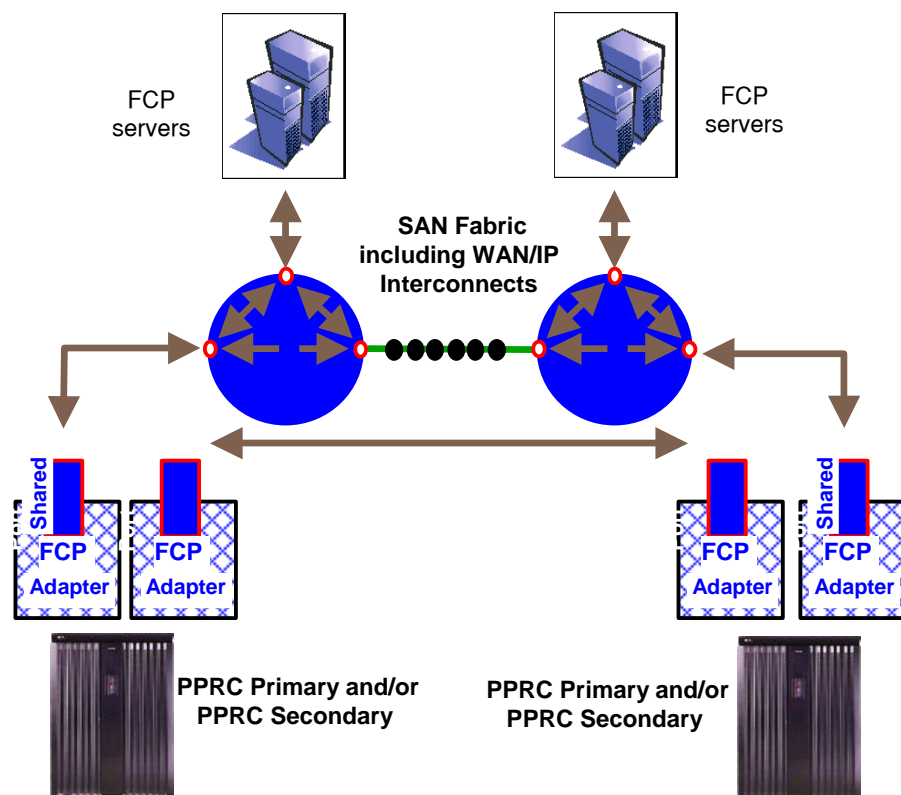
Price / Performance

Value

**Sets _New Standard_ for High Performance, High Capacity Disk Mirroring**

# ESS…

Leveraging IBM Fibre Channel Mirroring Technology

FCP servers

FCP servers

**SAN Fabric including WAN/IP Interconnects**

Shared

Shared

**FCP Adapter**

**FCP Adapter**

**FCP Adapter**

**FCP Adapter**

**PPRC Primary and/or PPRC Secondary**

**PPRC Primary and/or PPRC Secondary**

- **Flexible, Dynamic Connectivity for Disk Mirroring**
  - ➜ ESS Fibre Channel mirroring links are designed to be:
    - • *Bidirectional, full duplex, simultaneous send/receive*
  - ➜ Designed to provide simultaneous sharing of ESS PPRC links between storage and server workloads, in both directions
    - • *Competitive disk mirroring Fibre Channel alternatives may require more ports*
    - • *Previous ESCON implementations required more ports*
  - ➜ PPRC paths designed to be routed through a SAN fabric and exploit an existing Fibre Channel infrastructure and capacity
    - • *Including Wide Area Network and IP interconnects*

- **Benefits:**
  - ➜ **Designed to help significantly reduce infrastructure requirements and cost**
  - ➜ **Designed to provide enhanced performance with lower TCO**
  - ➜ **Designed to provide highly flexible, dynamically changeable configuration possibilities**
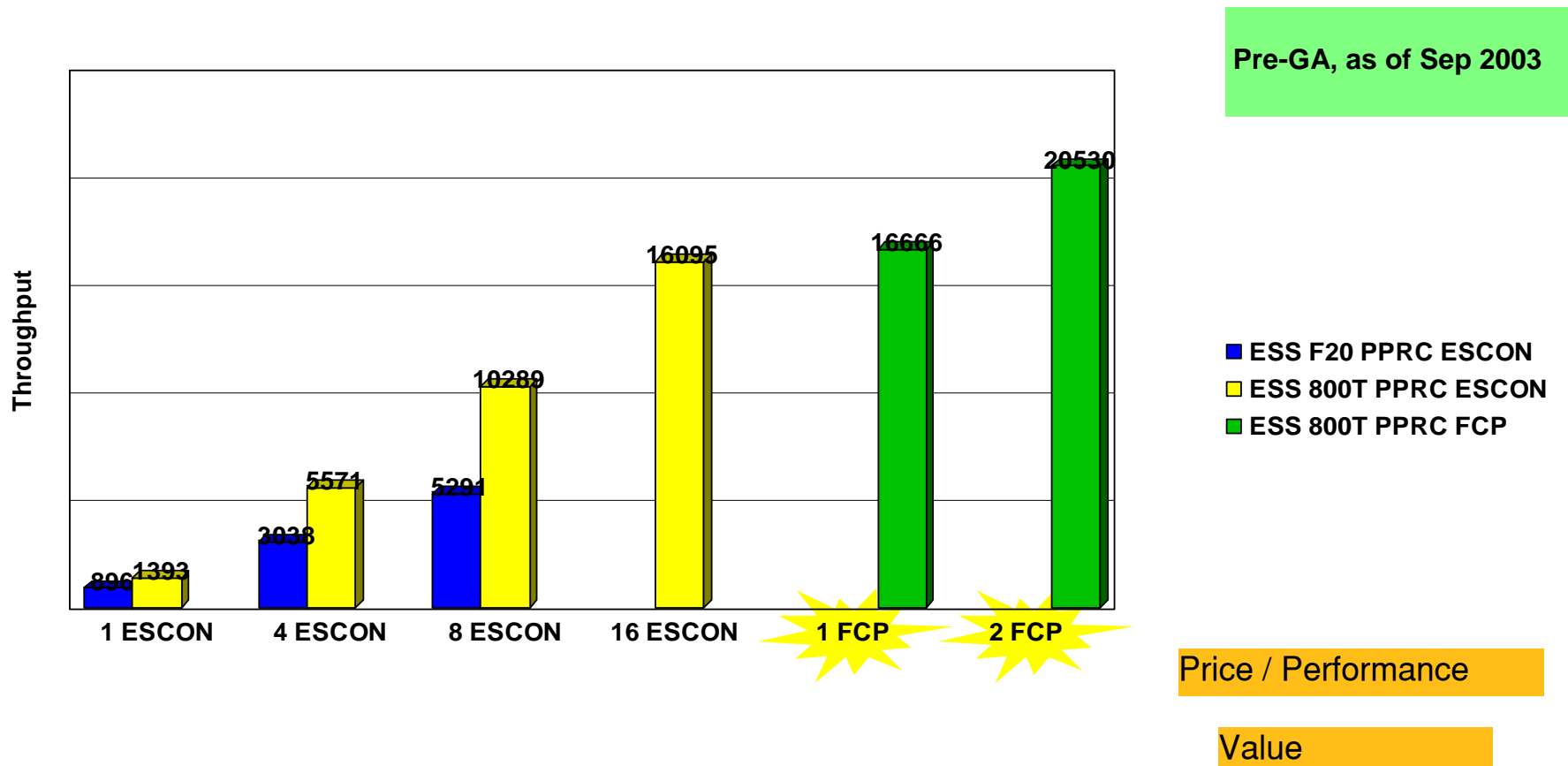
Price / Performance

Value

**Sets _New Standard_ for Flexible Connectivity in Disk Mirroring**

# A <u>New Standard</u> in Disk Mirroring Performance

## Metro Mirror over FCP at 0 km - 100% 4 KB Write Hit Workload (Open)

**Pre-GA, as of Sep 2003**

Throughput

| | | | | | |
|---|---|---|---|---|---|
| 896 1393 | 3038 5571 | 5291 10289 | 16095 | 16666 | 20530 |
| 1 ESCON | 4 ESCON | 8 ESCON | 16 ESCON | 1 FCP | 2 FCP |

■ ESS F20 PPRC ESCON
□ ESS 800T PPRC ESCON
■ ESS 800T PPRC FCP

Price / Performance

Value

# Why Asynchronous Disk Mirroring?

▪ **Metro Mirror (Synchronous PPRC) over Fiber Channel is now supported for distances up to 300 KM**

  ▪ This can protect from a local disaster, but not a regional area disaster such as hurricane or earthquake

▪ **Synchronization overhead** associated with 'zero-data-loss' Metro Mirror storage mirroring may be unacceptable for some important customer environments
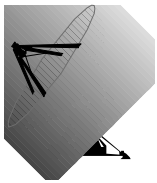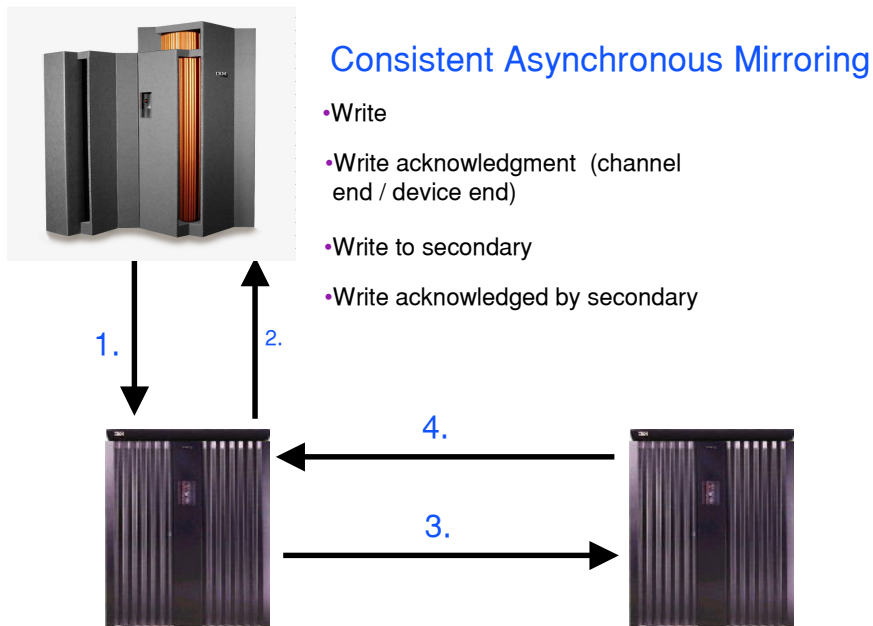
▪ Asynchronous disk mirroring can **reduce configured bandwidth** requirements between source and target, thus reducing infrastructure costs

▪ Although asynchronous copy allows the remote site to lag the local site in data currency, this is considered **acceptable** as a tradeoff for long distance, as long as:
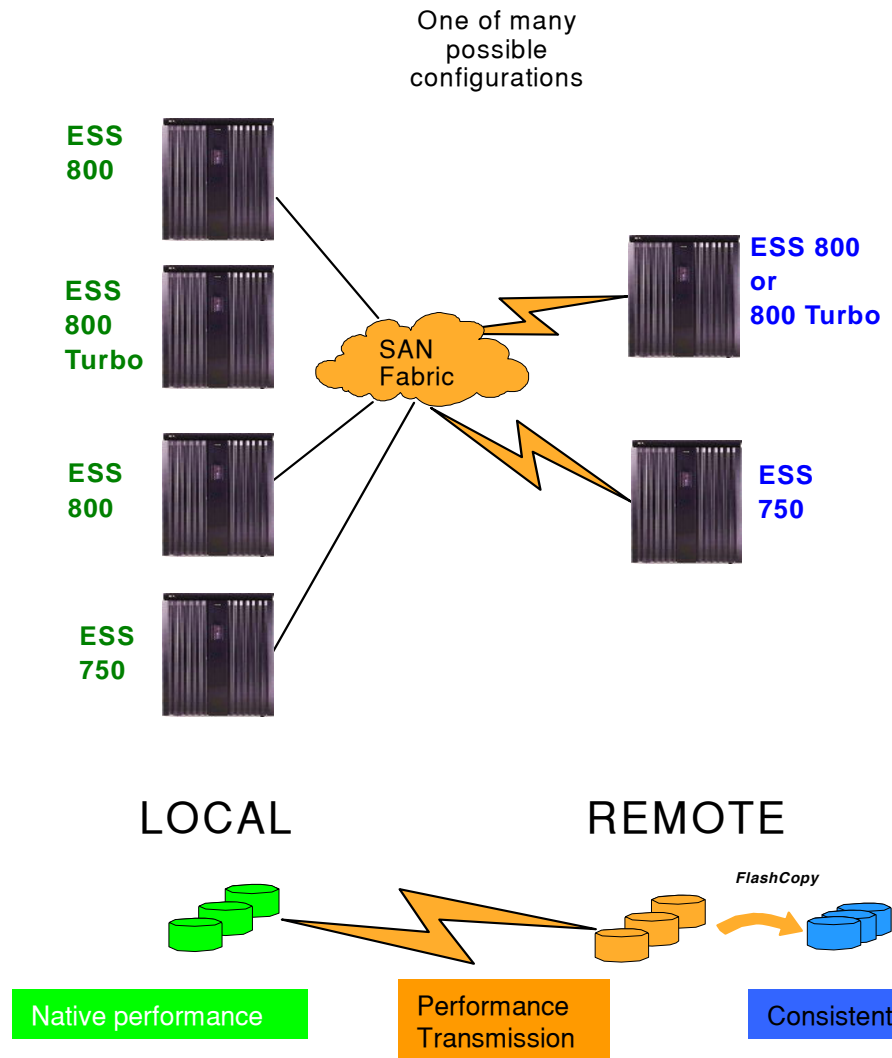
RPO

Production =>    Data Lag

**Failure**
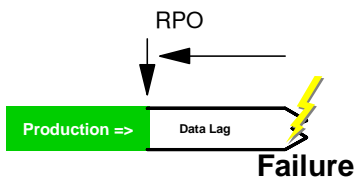
  ▪ The remote copy is data consistent and databases restartable
  ▪ The data loss is on the order of seconds, rather than minutes or hours
  ▪ The less data that can be lost, the better the solution

# Characteristics of a Data Consistent Asynchronous Solution

### Consistent Asynchronous Mirroring

• Write

• Write acknowledgment  (channel end / device end)

• Write to secondary

• Write acknowledged by secondary

1.

2.

4.

3.

▪ No application response time degradation

▪ Mirror copy of data created independent from primary application write

▪ Consistency always maintained at mirrored site

▪ Mirror lags in currency to the smallest degree possible

▪ Link Bandwidth utilization improved over synchronous solutions

▪ A tertiary copy is always required to preserve consistency

▪ Data Loss limited to data in queue or in    transit

# Announcing: Storage Consolidation and Disaster Recovery with *Global Mirror for ESS*

One of many possible configurations

ESS 800

ESS 800 Turbo

ESS 800

ESS 750

SAN Fabric

ESS 800 or 800 Turbo

ESS 750

LOCAL          REMOTE

*FlashCopy*

Native performance

Performance Transmission

Consistent Data

- Designed to Provide:

  - Global Distance: **Two-site, unlimited distance, data consistent asynchronous disk mirroring**

  - Scalability: **Consistency Group supported across up to 8 total ESSs in Global Mirror session**

  - Flexibility: **Many possible configurations**

  - Heterogeneous: **Data can span zSeries® and open systems data, and can contain a mix of zSeries and open systems data**

  - Application Performance: **Native**

  - Mirroring Performance: **Two ESS Fibre Channel disk mirroring links per ESS sufficient for almost all workloads**

- Benefits

  - Autonomic: **No active external controlling software required to form consistency groups**

  - Saves cost: **No server cycles required to manage consistency groups, fewer**

  - Lowers TCO: **Improved performance, global distances; and lower costs**

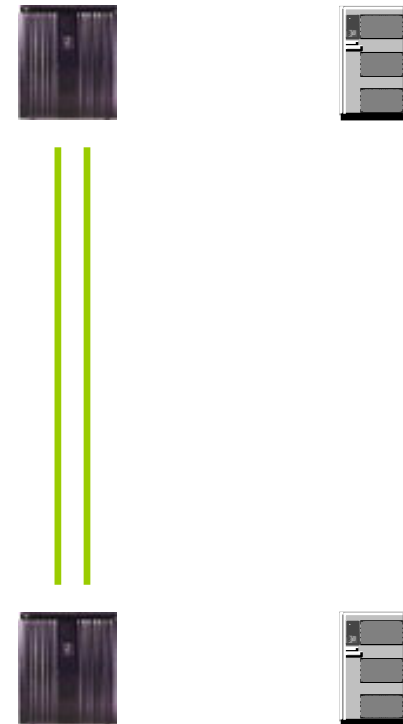# *Global Mirror for ESS* Characteristics and Customer Value
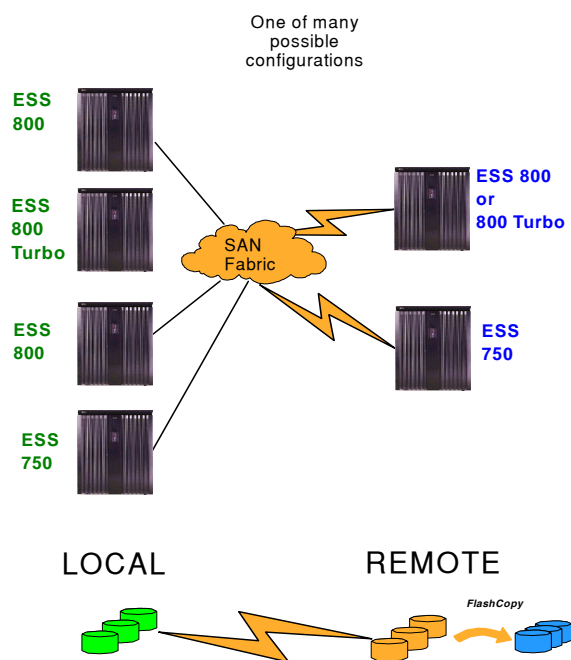
**Currency of Consistent Data**

**Link Requirements**



**Consistency Group**

**Data is more current =>**
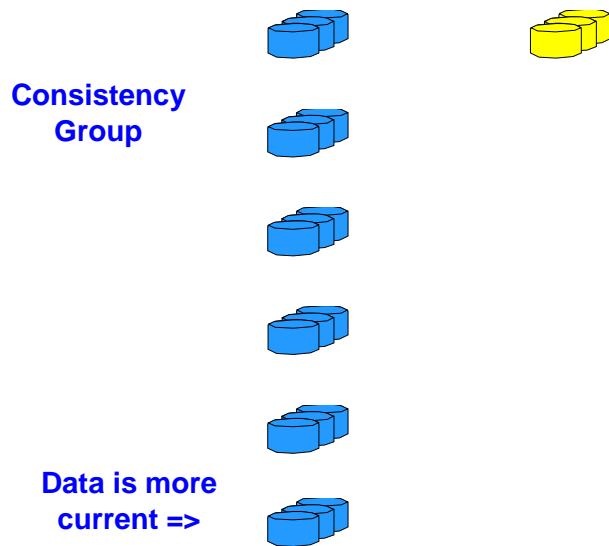
**Outage**

**10 TB:  Improved currency of 25 seconds could save *8000* OLTP transactions**

**ESS reduced number of mirroring links can save significant telecom costs**

# Summary: Global Mirror for ESS

One of many
possible
configurations

ESS
800

ESS
800
Turbo

SAN
Fabric

ESS 800
or
800 Turbo

ESS
800

ESS
750

ESS
750

LOCAL        REMOTE

FlashCopy

- Maintains data consistency at the mirrored site without host/server automation SW
  - Consistent Mirror is maintained completely by the ESS microcode
  - Currency can be configured to be as little as 3-5 seconds, bandwidth and transit time to remote permitting
  - No perceptible response time degradation due to Global Mirror for ESS on applications
  - Unlimited distances, exploits SAN distance extension

- Autonomic, Self-Tuning Characteristics:
  - ►Autonomic self tuning
  - ►Does not require server cycles for data movement
  - ►Ride out peak write periods
  - ►Reduced Configured Bandwidths
  - ►Can handle peaks in write activity without impacting host performance

- Scalability: up to 8 ESSs total in Global Mirror session

- Heterogeneous solution for Open, zSeries, and iSeries
  - ►Open, zSeries, and iSeries may all be in the same Consistency Group
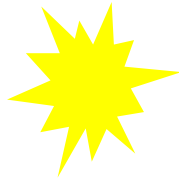
# eServer and ESS Business Continuity Solution Integration Packages

| GDPS Storage Manager | or | AIX HACMP 5.1 | or | eRCMF (Open Systems) |
|---|---|---|---|---|

**Synchronous PPRC**

**Primary Site**
*(PPRC primary)*

**Secondary Site**
*(PPRC secondary)*

**Scalable Data Integrity**

- **GDPS®\* Storage Manager** for ESS mirroring for *zSeries*

- **AIX HACMP 5.1** provides new exploitation of ESS mirroring for *pSeries® AIX*

- **Windows® GeoDistance** provides integration of ESS mirroring for *Microsoft Windows Clustering*

- **D/R solutions for *Heterogeneous Platforms***
  - **eRCMF\*\* Storage Manager** for ESS mirroring for *Open Systems*
  - **GDPS Open LUN Management** for ESS mirroring for *zSeries and Open*

- **Benefits:**
  - **Tuned and integrated for specific environments**
  - **Simplified** implementation and powerful management of ESS PPRC, optimized for specific environments
  - **Automated Data Integrity** for applications and databases
  - **Highly scalable** coordination and control
  - **Integrated solutions** with **attractively priced software/services offerings**
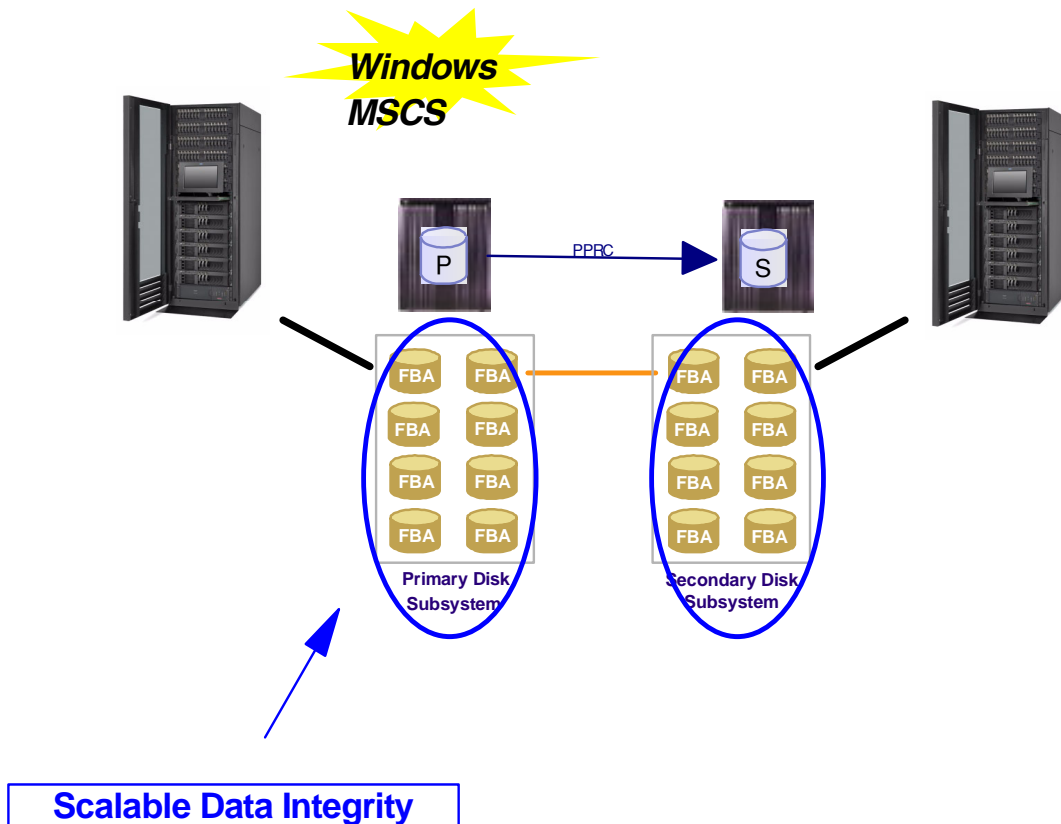  - **Worldwide 24X7 support** by IBM Global Services
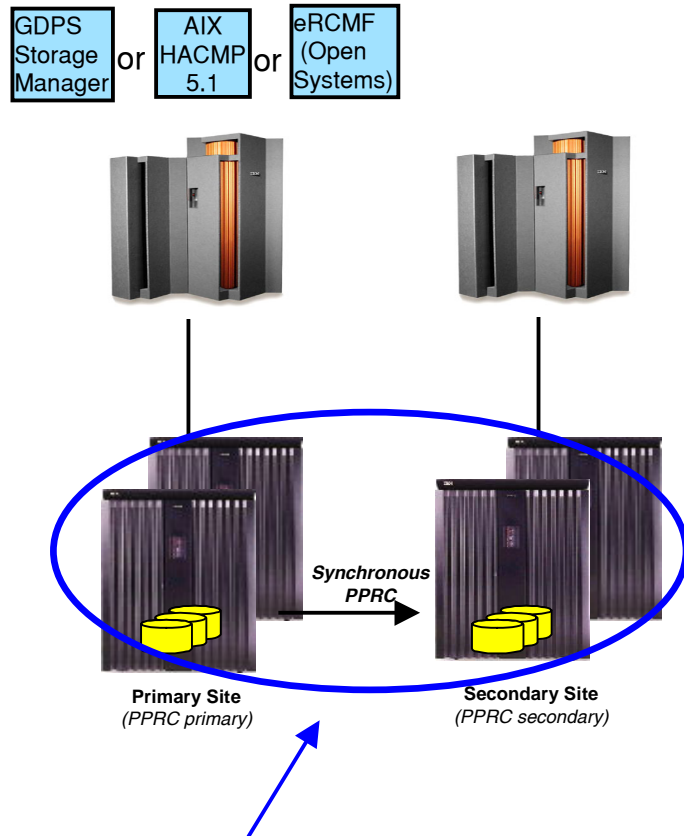
Solutions

# Windows GeoDistance Solution
## ESS Mirroring Integration with Windows Microsoft Clustering Services (MSCS)



**Windows MSCS**

P → PPRC → S

Primary Disk Subsystem

Secondary Disk Subsystem

**Scalable Data Integrity**

- Benefits:
  - Designed to enhance geographically dispersed Windows MSCS clustering with integrated exploitation of ESS mirroring
  - Designed to provide 'power outage data consistency' at remote Windows MSCS site for unplanned outages

- Windows GeoDistance is:
  - **Designed to provide Data Integrity:** Windows MSCS-based code integrated with ESS mirroring Consistency Group data integrity function

  - **Designed to be Scalable:** Can *span* multiple ESS storage footprints

- Solution applies to:
  - Any Windows MSCS cluster using Metro Mirror
  - Supports one cluster node at each site

Solutions

# Geographically Dispersed Sites (GDS) Overview

- Extend Microsoft Cluster Server (MSCS) to metropolitan distances and beyond
- Exploit mirroring capability of the Enterprise Storage Server
- Provide automated disaster recovery
- Support application-level failover/failback
- Allow manual failovers for planned maintenance

# GDS Benefits

- Adds enhanced availability and disaster tolerance to MSCS
- Works with any cluster-aware application
    - Databases (MS SQL Server, IBM DB2)
    - Messaging (MS Exchange, Lotus Domino)
    - File Sharing, Print, IIS
- Uses Microsoft Cluster Administrator interface – no new interface to learn
- Supports distances of up to 103km (or more)

# Typical Hardware Configuration
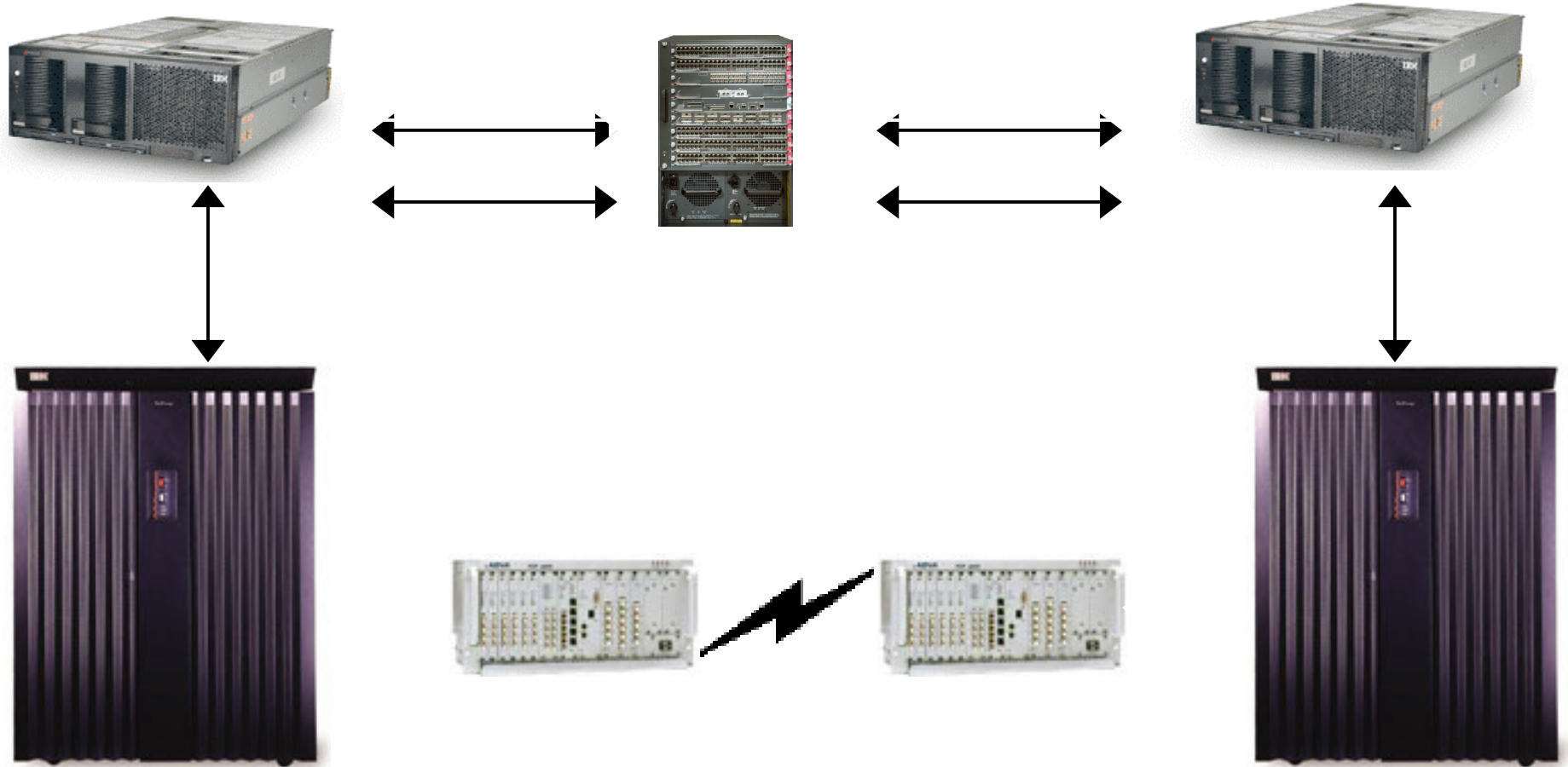
# GDS Components

- Failover Service
    - The quorum arbiter
- SDD Service
    - Interfaces with ESS
- Extension DLL
    - Interfaces with MSCS Cluster Administrator
- Resource DLL
    - Manages disk resources
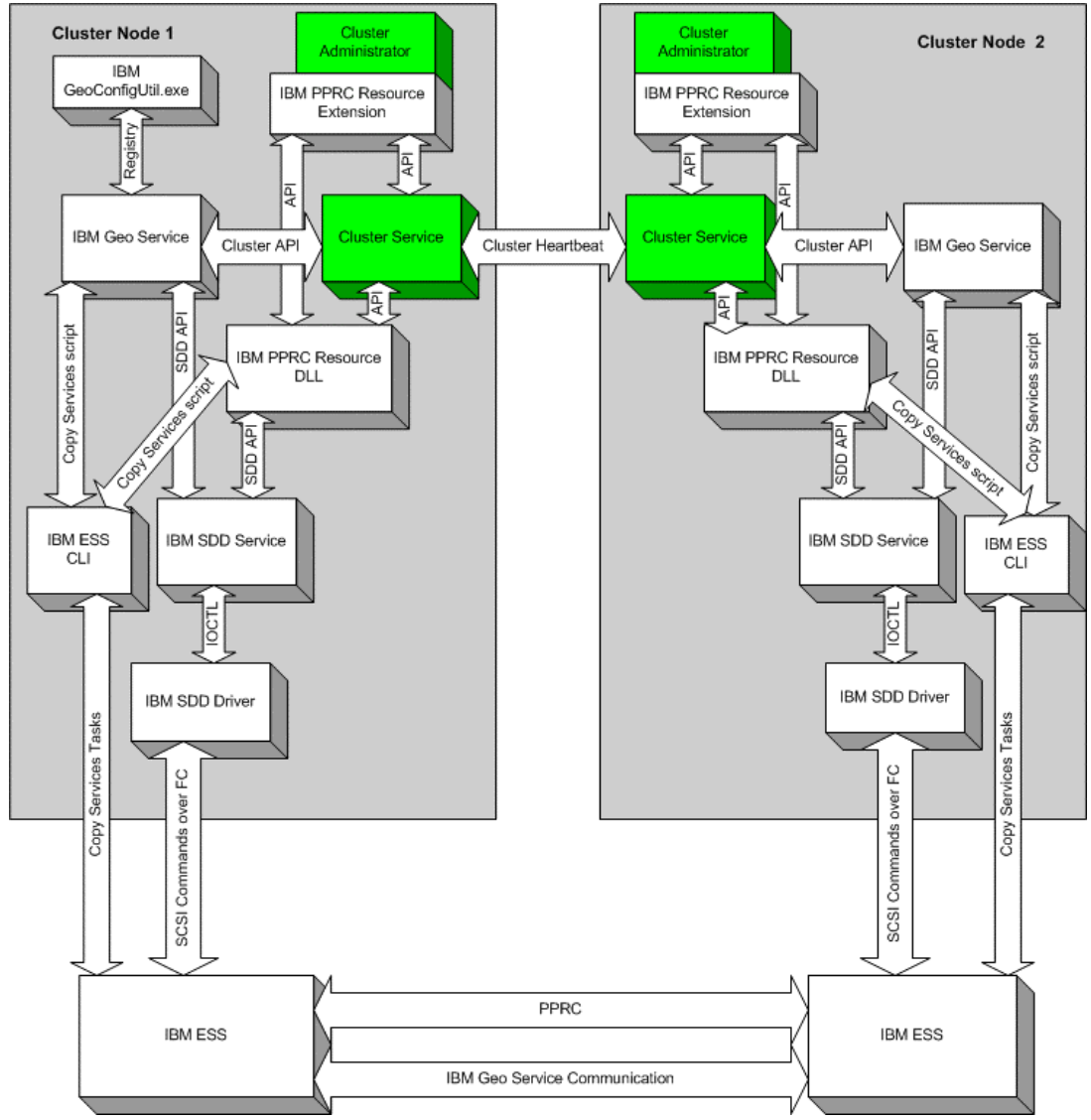
# IBM GDS Summary

- Improves data availability with continued service during hardware, software and network failures or upgrades

- Adds disaster tolerance to MSCS

- Seamless integration of PPRC with MSCS

- No new external interfaces to learn

# ESS xSeries Interoperability Support

| Linux | NetWare | VMWare | Windows |
|---|---|---|---|
| Red Hat Linux | Novell NetWare 4.11, 4.2 Standby Server, 5.0, 5.1, 6.0. 6.5 | VMWare ESX 2.0.1 | Microsoft Windows NT - Server 4.0 - Server 4.0 with MSCS |
| Red Hat Enterprise Linux | | VMWare Guest Support - Win2K - United Linux - Red Hat Enterprise Linux Advanced Server 2.1 | Microsoft Windows 2000 - Server - Advanced Server including MSCS |
| SuSE Linux | | | Microsoft Windows Server 2003 - Standard Edition - Enterprise Edition inc. cluster service |
| SuSE Linux Enterprise Server | | | |
| United Linux | | | |

With information on demand, clients can respond with flexibility and speed to any customer requirement, market opportunity, or external threat



*Information On Demand Solutions featuring IBM TotalStorage*

# Appendix

# St. Anthony's Medical Center

## Business Need

–High availability and failover storage architecture

–24x7 availability and protection for its PACS data

## Solution

–10TB ESS with PPRC, 3494, SAN switches

–3TB backed up/archived per day

–Automated, unattended backup via TSM

## Benefits

–Provides online, real-time images to physicians

–10 fold reduction in Recovery Time Objectives

–Automated , unattended backup/recovery

–25% reduction in administrative costs overall

# St. Michael's Hospital

- **Need**

  – **Establish an online PACS system to store and retrieve images/data electronically**

  – **Enterprise wide disaster recovery plan**

  – **Better leverage IT staff and resources**

  – **Automated, centralized backup/recovery**

  – **Scalable, reliable storage solution**

- **Solution**

  – **IBM Tivoli Storage Manager for automated, centralized backup.**

  – **IBM TotalStorage SAN Volume Controller with FlashCopy and PPRC**

  – **IBM eServer pSeries and xSeries servers**

  – **IBM TotalStorage 3584 Ultrium UltraScalable Tape Library for centralized tape**

  – **IBM TotalStorage 3583 Ultrium Scalable Tape Library at secondary data center**

  – **IBM TotalStorage Enterprise Storage Server for main data center**

**ST. MICHAEL'S HOSPITAL**

*A teaching hospital affiliated with the University of Toronto*

**Benefit**

- **Continuous access to data**

- **Increased staff productivity due to automated, centralized backup/recovery**

- **Protection of critical PACS repository**

- **Highly available, resilient IT infrastructure (99.999% availability)**