IEEE Mass Storage Conference Vendor Reception

Lake Tahoe, NV

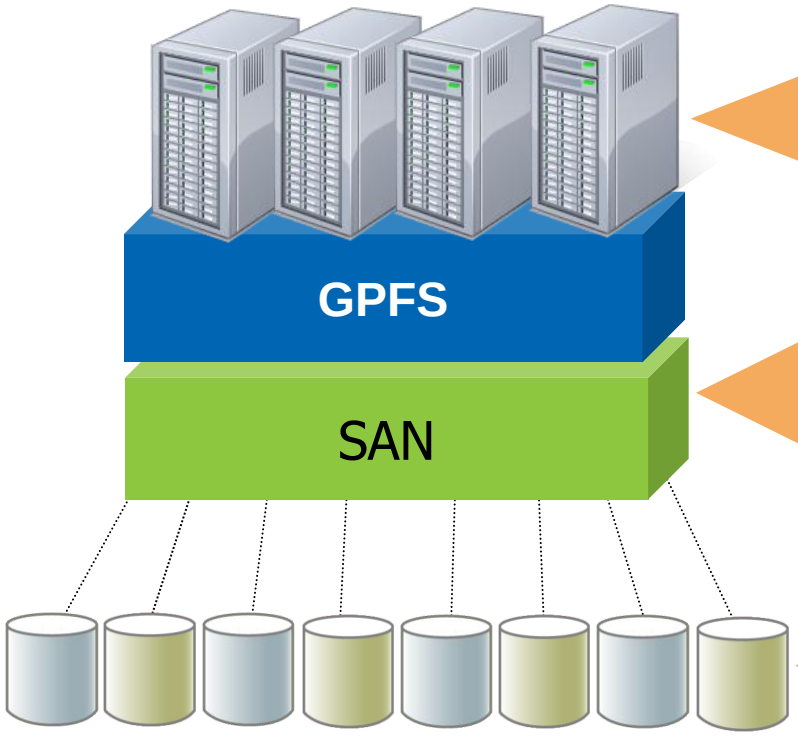
Joe Rotiroti
Client Systems

IBM, Federal
484 433 9756 cell
845 491 5227 fax
rotiroti@us.ibm.com

Manager

May 04, 2010

How is GPFS different?



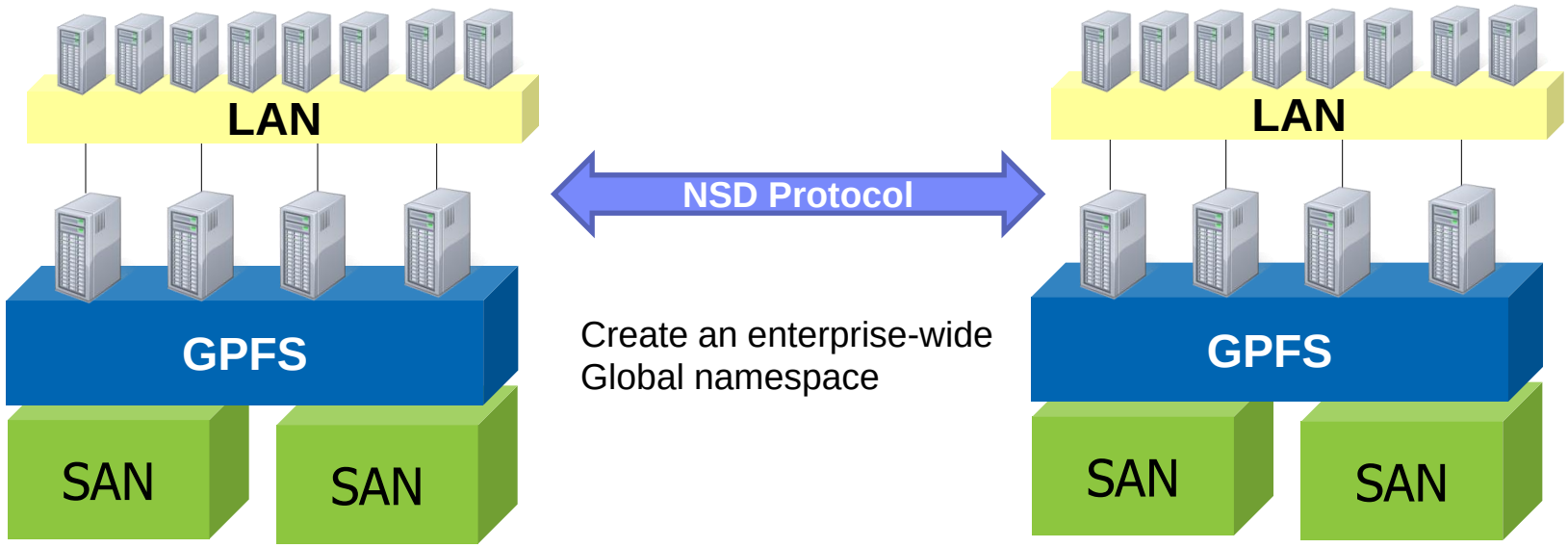
Massive namespace support

Centrally deployed, managed, backed up and grown.

Seamless capacity and performance scaling

All features included - All of the software features: snapshots, replication and multi-site connectivity are included in the GPFS license. There are no license keys besides client and server to add-on, you get all of the features up front.

Enhance enterprise-wide collaboration through multi-cluster



Why?

- Tie together multiple sets of data into a single namespace
- Allow multiple application groups to share portions or all data
- Secure, available and high performance data sharing

File System configuration and performance data

GPFS is already running at data sizes most companies will start supporting five years from now.

Number of files:
- 2 Billion per file system
- 256 file systems
- Max File System Size: 2^99 bytes
- Max File Size = File system size

Disk IO:
- AIX 134 GB/sec
- Linux 66 GB/sec

Number of nodes:
- 1 to 8192

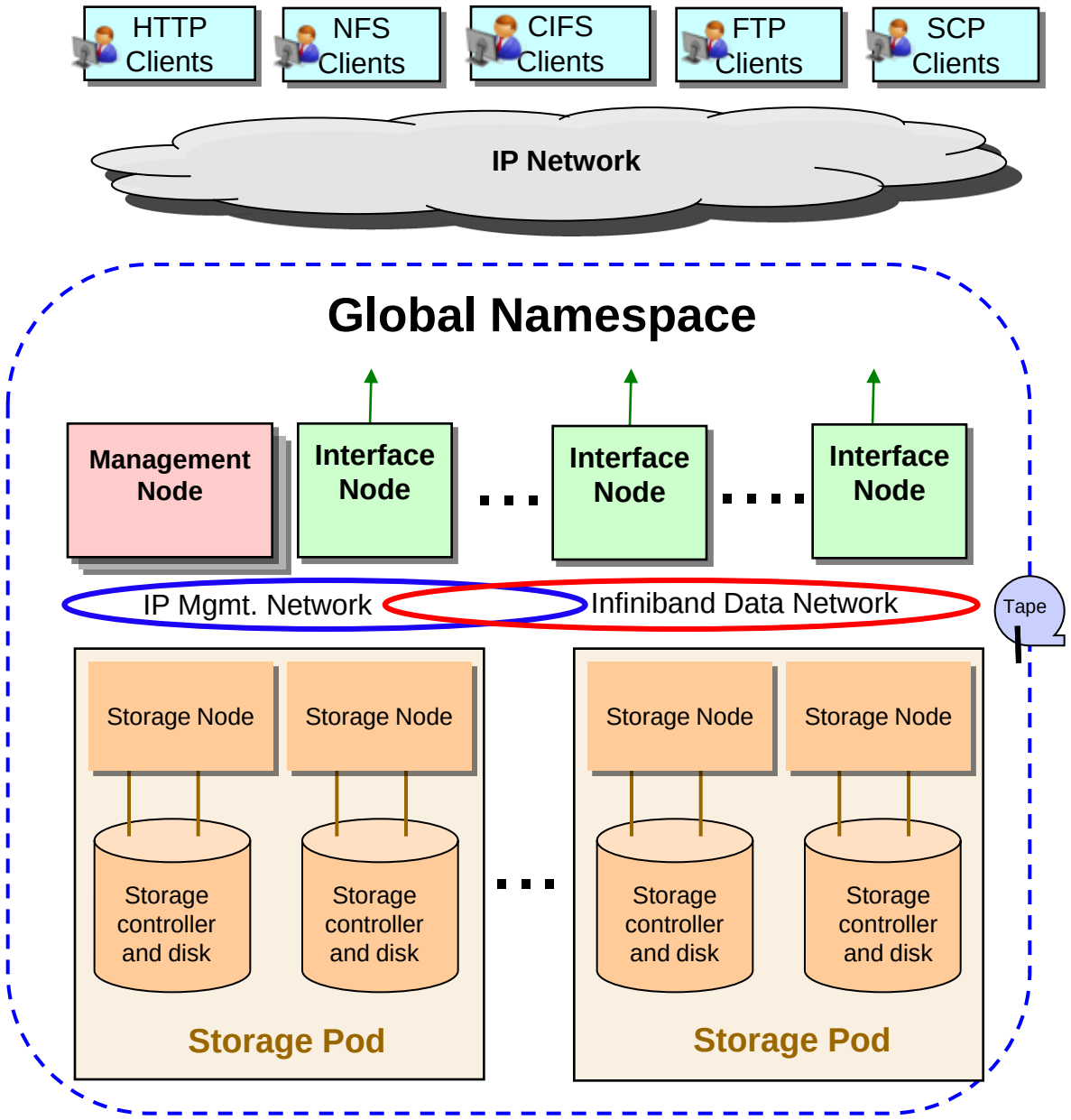
IBM Scale Out NAS Architecture

Parallel Grid architecture provides:

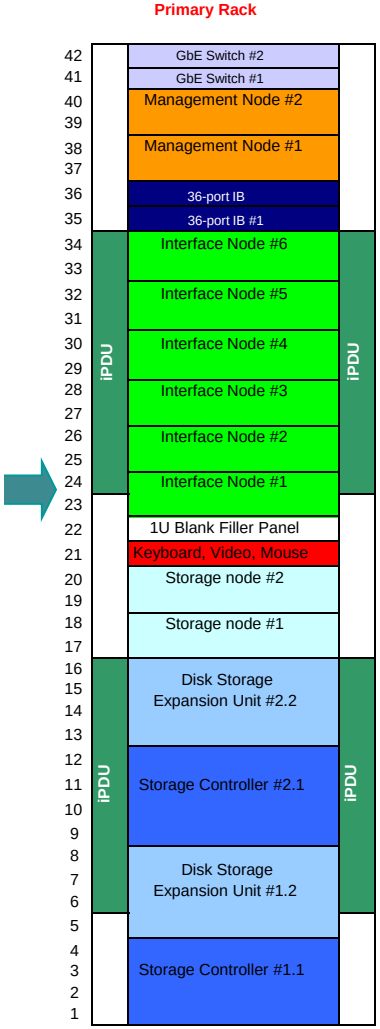
- **Massive scalability**
- **Extreme performance**
- **Automatic Tiered Storage**
- **Global virtual file server**

IBM Scale out NAS is an appliance:

- **Simple:** Only 3 basic orderable parts:
  - ✓ **Management Nodes**
  - ✓ **Interface Nodes**
  - ✓ **Storage Pods**
- All nodes are clustered
- SONAS Software runs on every node
- Uses enterprise commercial off-the-shelf (COTS) components
  - To leverage technology cost and performance advances quickly



IBM Scale Out NAS Appliance



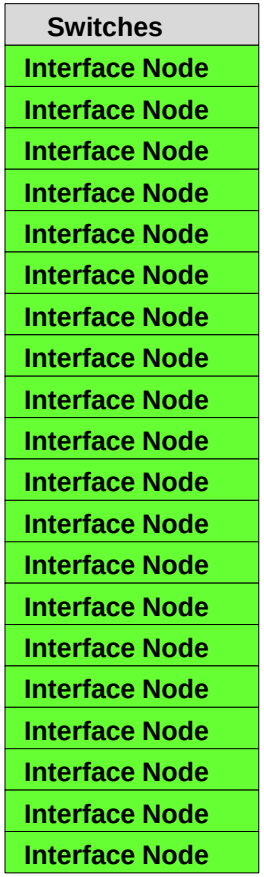
Fully Integrated Appliance

Superscalar Network Attached Storage (NAS) system with a industry-leading software stack

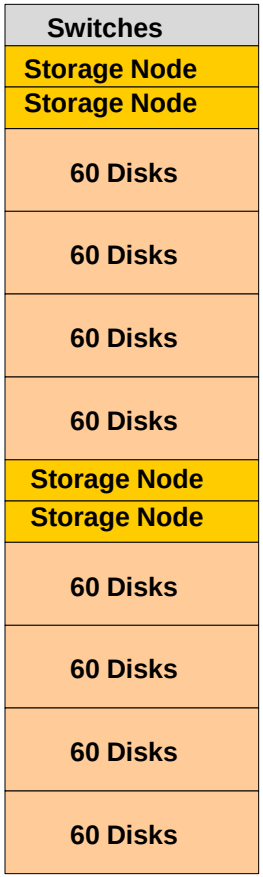
NFS, CIFS, FTP, HTTP and SCP network file protocols

- Near linear performance on a 'per-node' basis to existing market, but with aggregate performance that will be extreme and market leading.
- Global file system namespace
- Synch and Asynch replication of file data
- File system snapshots, quotas,
- Scalable, Integrated Information Lifecycle Management (ILM)
  - Automatic tiered storage
  - HSM support to external Tivoli Storage Manager (TSM) server
- Real-time performance monitoring
- Integrated System HealthCenter for HW monitoring
- LDAP Support, Windows Active Directory support
- 1GbE and 10GbE host interfaces
- High speed extremely low latency (20 Gbit/sec) private Infiniband cluster data network
- High density RAID controller and disk expansion drawer (4U x 60 HDD's)
- Support for high-performance SAS and high-capacity SATA HDD's

IBM Scale Out Network Attached Storage Expansion Options



I/O Dense Configuration



Storage Dense Configuration

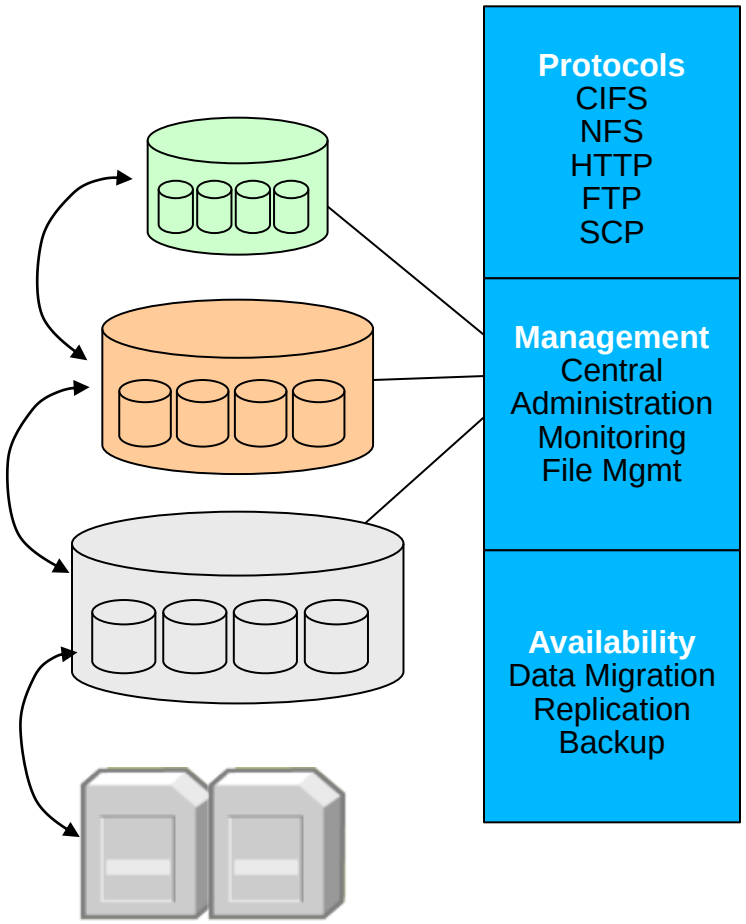
Node Expansion

- Up to 30 interface nodes
- High speed extremely low latency (20 Gbps) private Infiniband cluster data network

Storage Expansion

- Up to 30 storage pods, each with up to 240 HDD's
- Scalable to 14.4 PB of raw storage in single system using 2 TB drives
- Max of 7,200 hard drives

IBM Scale Out NAS: The life of a file from creation to deletion



- **Using Automatic Tiered storage**
  - Storage pool – group of LUNs
  - Fileset - define subtrees of a file system
  - Policies – for rule based management of files inside the storage pools
- **Delivering:**
  - Scalability - billions of files
  - One global file system name space across independent logical storage pools
  - Files in the same directory can be in different pools
  - Files placed in storage pools at create time using policies
  - Files moved between pools via automated policy-driven tiered storage
  - Hierarchical storage based on files
  - Hierarchical to both disk and tape
  - Allows classification of data according to SLAs

SAN Volume Controller

2 nodes

- **IBM's Storage Virtualization product and Single point of control for Storage resources.**
- **Designed to combine Storage capacity from multiple disk systems into a reservoir of capacity**
- **Thin provisioning**
- **Supports the consolidation of Heterogeneous Storage systems from many vendors**
  ftp://ftp.software.ibm.com/common/ssi/pm/sp/n/tsd00254usen/TSD00254USEN.PDF
- **Mirrored virtual disks, which can be mirrored across disk subsystems**

SAN Volume Controller

- **Tiered Storage support**
- **Easy data migration from one Storage to another without Apps down time.**
- **Easy to use Graphical management interface**
- **Supports IBM Tivoli Storage FlashCopy manager application-aware snapshots.**
- **For more information,**

http://www-03.ibm.com/systems/storage/software/virtualization/svc/

- **Storage Performance council Reports**

SPC1 - **http://www.storageperformance.org/benchmark_results_files/a00072_IBM_SVC4.3_SPC1_full-disclosure.pdf**

SPC2 - **http://www.storageperformance.org/results/b00024_IBM-SVC4.2_SPC2_full-disclosure.pdf**

IBM System Storage Disk product family

SAN products

Enterprise storage

DS8000® series

FC/FICON, ESCON

XIV®

FC and iSCSI

Departmental mid-range storage

DS5000™ DS4800

FC

DS4700

DS5020

FC and iSCSI

DS6000™ series

FC and FICON

SAN Volume Controller (SVC) Flexible Disk Virtualization

Entry workgroup storage

DS3200 (SAS)

DS3300 (iSCSI)

DS3400 (FC)

NAS products

N7000 series

N6000 series

N5000 series

N3000 series

N3700

http://www-03.ibm.com/systems/resources/systems_storage_resource_pguide_prodguidedisk.pdf

DS8700 Highlights Disk Road Map

- **New Hardware Platform**
  - P6 model 570 servers
    - 4.7 Ghz processors
      - 2 core and 4 core
    - Enhanced RAS features
  - PCI Express I/O drawer
    - High throughput PCIe 4x Gen2 cable connection.
  - New Model (Model 941/94E)
    - Significant performance increase
  - Enhanced device adapter processor
    - IOPS boost
- **Functional Improvements**
  - High Performance Ficon for System z (multitrack)
  - Encryption enhancements
    - Deadlock Recovery Key Management
    - Dual Platform Key server support
  - Code Load Improvements
    - Service window reductions
- **Inherited Advanced Function from DS8000**

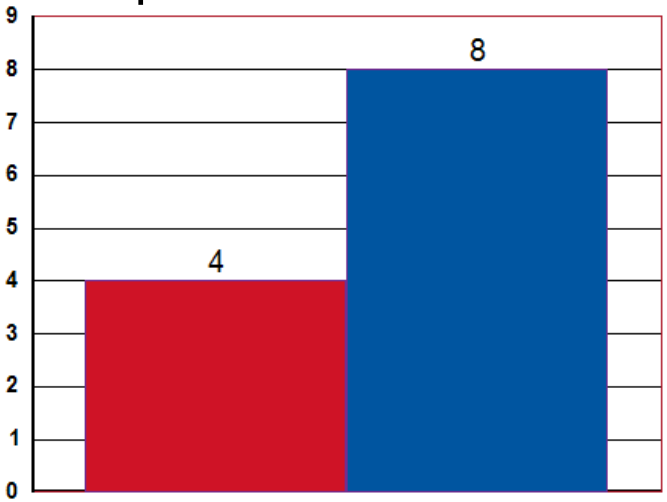
Smart Tools for Optimizing Tiered Storage

Smart data analysis tools

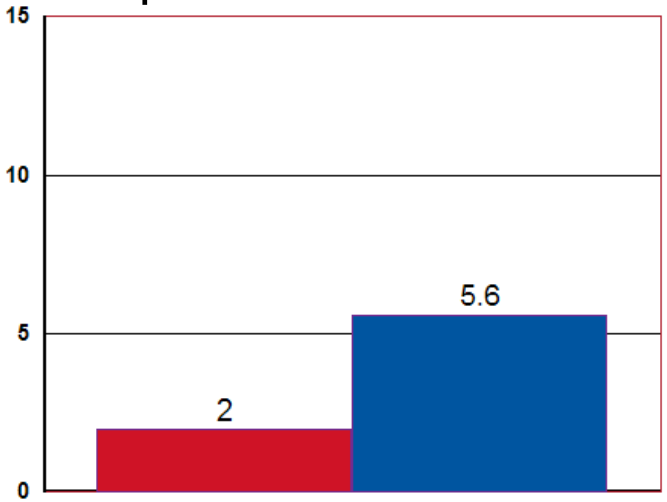
- **IBM FLASHDA** tool for identifying opportunities for SSD exploitation
- **New DS8000 performance analytics** for identifying ***sub-volume*** data to expl tiered storage for both open and zOS
- **System Monitoring Facility (SMF)** for analyzing z/OS trends and peak usage
- **System p AIX filemon** for performance analysis
- **System i workload analysis tools**
- **Tivoli Storage Productivity Center for Disk** for detailed, real-time performar monitoring and historical performance analysis
- **Omegamon XE** for analyzing dataset and volume performance
- **Softek Data Mobility Console for z/OS** offers real time data analysis and use controlled movement of data with concurrent application access

DS8000 POWER6 Estimated Performance

Sequential Reads – GB/s

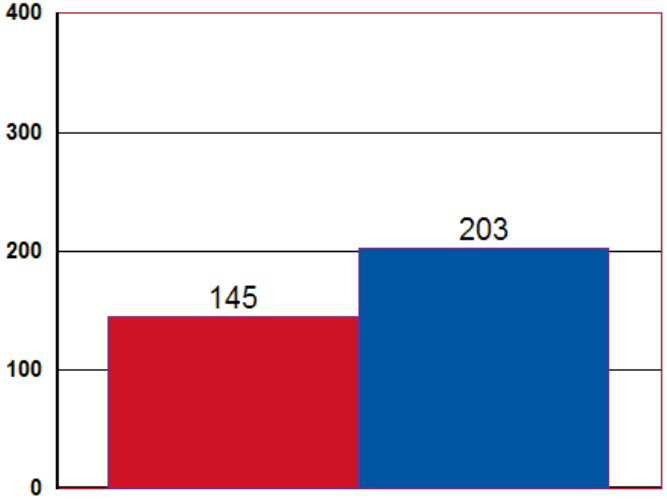


Sequential Writes – GB/s

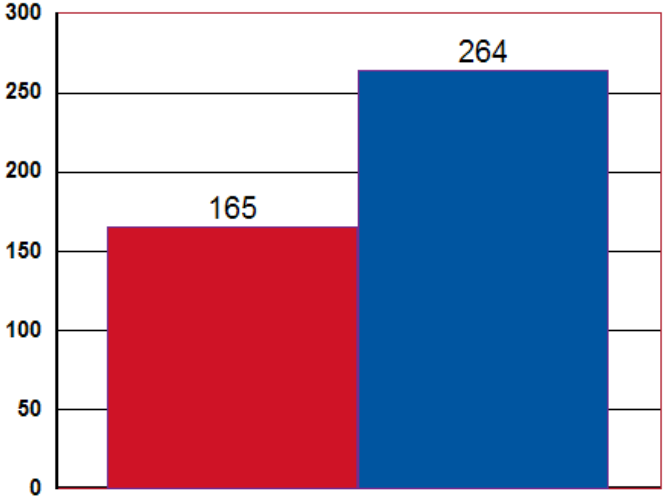


P5+
P6

Mainframe database – 1000 IO/s

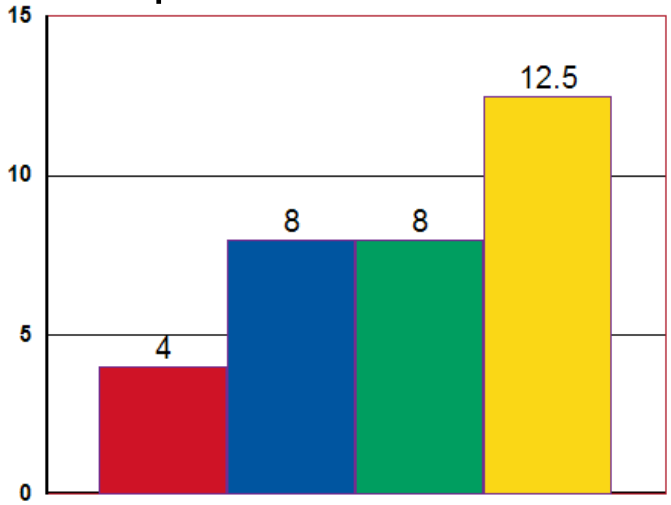


Distributed database – 1000 IO/s

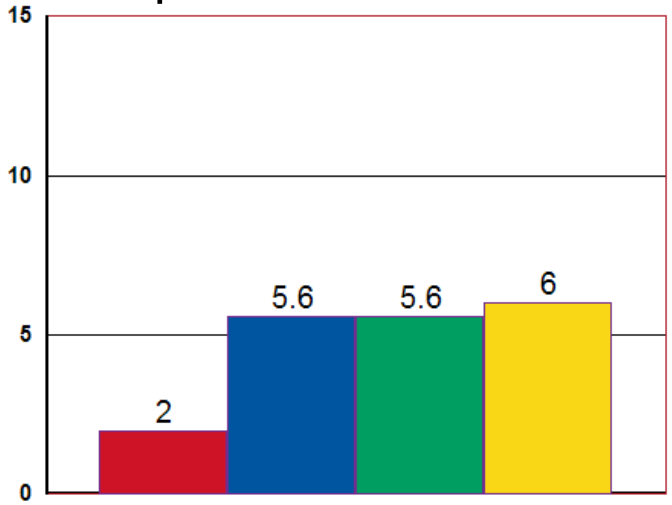


DS8000 Estimated Future Performance

Sequential Reads – GB/s

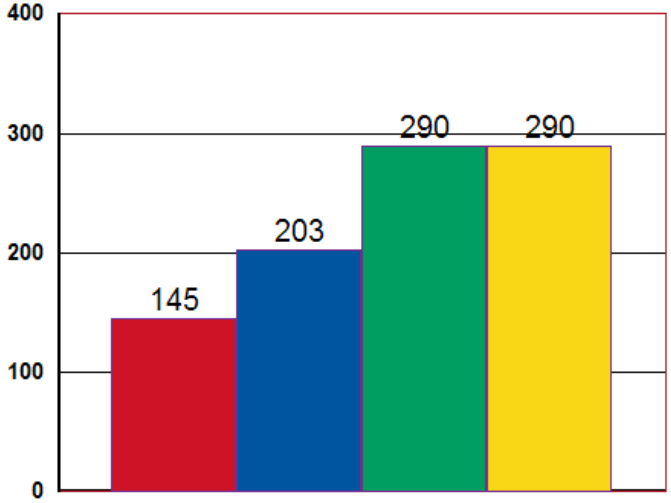


Sequential Writes – GB/s

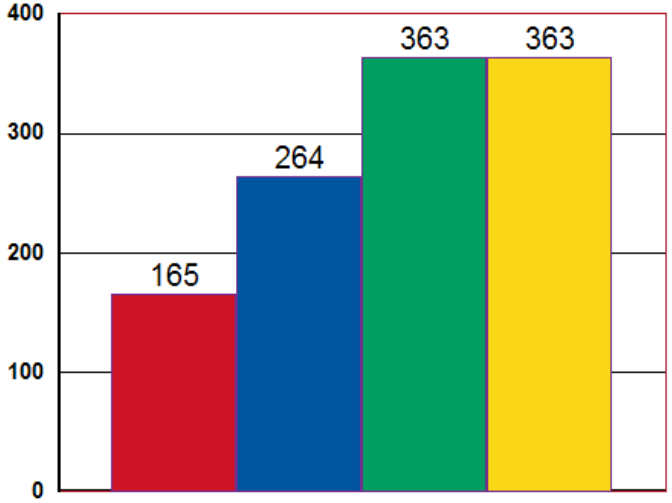


R4.2
R5.0
R5.1
R6.0

Mainframe database – 1000 IO/s

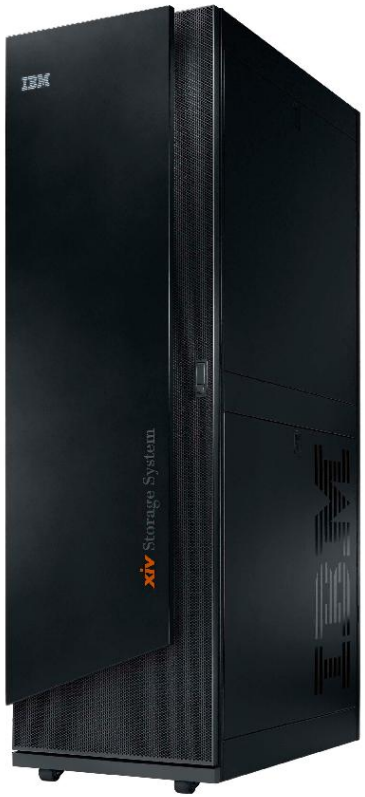


Distributed database – 1000 IO/s



IBM XIV

- **IBM XIV is a Grid-based Storage system of independent modules**
- **Enterprise storage suitable for mixed application use**
- **Ease of storage administration**
- **Scalability options leading up to a full rack**
- **Powerful Copy Services capabilities**
- **Automatic tuning for best performance**
- **Automatic data layout and seemless re-distribution of data when the system changes.**



IBM XIV

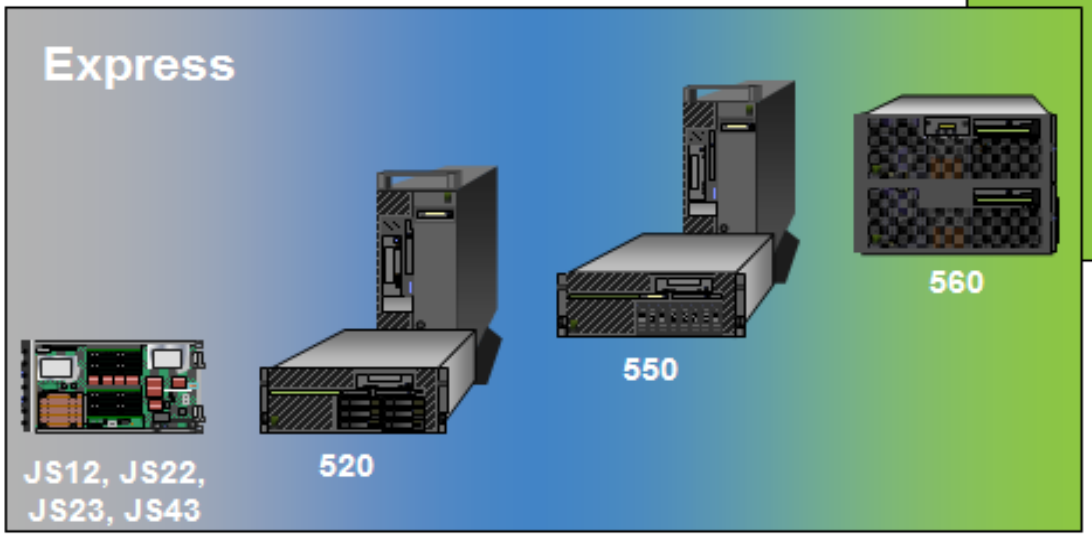
- **Interoperability with IBM AIX, HP-UX, Macintosh, MS-Windows, Novell, Red Hat Linux, Solaris, VM ware ESX and Linux on z.**
- **Supports 1,2,4 Gbps FC rates, iSCSI over Gigabit Ethernet and SCSI3 for both FC and iSCSI**
- **Flexible Interoperability with Virtual layer of storage systems (SAN Volume Controller, Virtual Tape Libraries and Virtual NAS).**
- **Minimum usable 27 TB (6 Modules) and Maximum 79 TB (15 modules).**
- **Storage capacity, spindles can be scaled from base configuration to 7 different configurations and Thin provisioning.**
- **FC and iSCSI host ports support**
- **Min 48GB cache and maximum 120GB.**
- **Various rich data protection features**
  - **Space efficient Snapshots and Consistency groups**
  - **Remote Mirroring (Synchronous and Asynchronous)**
- **High-end architecture with SATA drives**
  - **Provides revolutionary power (cooling) consumption**
- **Net space used by application data and zero lost space**
  - **Thin provisioning**

IBM XIV

- **Reduces the Storage Administration tasks and "Decision-less" storage.**
- **Single Full capacity XIV storage can provide**
  - 105,000 IOPS
  - 2.1GB/s Read
  - 1.3GB/s Write
  - Cache to Disk bandwidth 240 Gbps per module
- **24 x 4 Gb Host ports and 6 x 1 Gb iSCSI ports.**
- **2 x Gigabit Ethernet switches for interconnecting Data modules**
  - **Provides overall internal connectivity of 120 Gbps.**
- **3 UPS – Sufficient for 15 mins of operation without an external power source.**
- **Concurrent and non-disruptive Firmware upgrades**
- **http://www-03.ibm.com/systems/storage/disk/xiv/index.html**

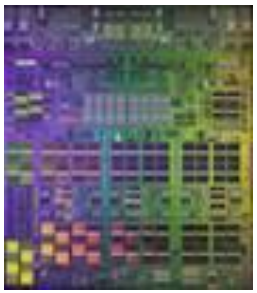
IBM Power Systems

- Consistency
  - Binary compatibility
  - Mainframe-inspired reliability
  - Advanced Virtualization
  - AIX, Linux and IBM i OS
- Complete flexibility for workload deployment



Processor Technology

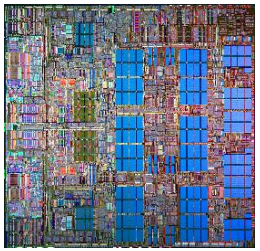
IBM Investment in the Power Franchise



POWER4
180 nm

- Dual Core
- Chip Multi Processing
- Distributed Switch
- Shared L2
- Dynamic LPARs (32)

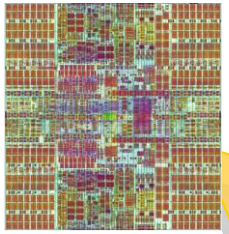
2001



POWER5
130 nm

- Dual Core
- Enhanced Scaling
- SMT
- Distributed Switch +
- Core Parallelism +
- FP Performance +
- Memory bandwidth +
- Virtualization

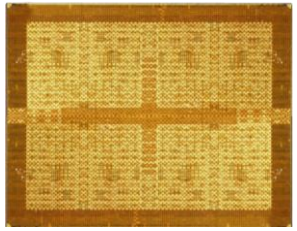
2004



POWER6
65 nm

- Dual Core
- High Frequencies
- Virtualization +
- Memory Subsystem +
- Altivec
- Instruction Retry
- Dyn Energy Mgmt
- SMT +
- Protection Keys

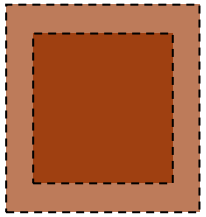
2007



POWER7
45 nm

- Multi Core
- On-Chip eDRAM
- Power Optimized Cores
- Mem Subsystem ++
- SMT++
- Reliability +
- VSM & VSX
- Protection Keys+

2010



POWER8

Future

Dynamic Infrastructure with System x & BladeCenter

X-Architecture™

Enterprise

Cluster 1350 & iDataPlex

Server Consolidation, Large Virtualization and Enterprise Workloads

Massive scale-out HPC, Cloud, Grid, energy efficiency

Scale Up

Infrastructure integration and simplification, energy efficiency

Scale Out

Single, infrastructure Applications

BladeCenter

System x Rack and Tower

IBM® Systems Director

THANK YOU

Joe Rotiroti
Client Systems

IBM, Federal
484 433 9756 cell
845 491 5227 fax
rotiroti@us.ibm.com

Manager

May 04, 2010