G18

IBM DS6000 and DS8000 Implementation for System z

Curtis Neal

IBM
SYSTEM z9 AND zSERIES EXPO
October 9 - 13, 2006

Orlando, FL

Abstract

- **G18 IBM DS6000/DS8000 Implementation for System z**

  Curtis Neal, IBM

- The DS8000 family of disk systems brings the best of breed attributes including massive scalability, industry leading performance and capability for the z environment. The DS6800 brings all the features of the earlier IBM Enterprise Storage Server (ESS) to a new, smaller rack mounted package. This session will discuss these system's features, how to successfully apply the products and some newly announced capabilities.

- **Level: Standard**

AGENDA

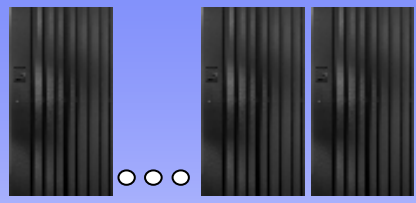
- Understand the DS6800 and DS8000 hardware performance characteristics.
- Understand logical configuration considerations related to performance
- Performance considerations related to the DS8000 hardware components and logical resources will be discussed.
- Understand principles of DS performance optimization
  - Isolation
  - Resource sharing
  - Spreading

IBM DS8000 – Enterprise Disk Functionality

- Setting a "*New Standard*" in Cost Effectiveness
  - ***Balanced Performance*** – Up to 6X ESS Model 800
  - ***Linear Scalability*** – Up to 192TB (designed for > 1PB)
  - ***Integrated Solution Capability*** – Storage System LPARs
  - ***Flexibility*** – Dramatic addressing enhancements
  - ***Extendibility*** – Designed to add/adapt new technologies
  - ***Storage Management*** – All New Management Tools
  - ***Availability*** – Designed for 24X7 environments
  - ***Resiliency*** – Industry Leading Copy and Mirroring Capability

- **IBM DS8300 hits 101,101 IOPS2 - New SPC-1 benchmark record - October 2005.**
  - Top performer for scale up disk system technology

New Standard in Functionality, Performance, TCO



DS8000

- **Outstanding copy services**
- **Supports lower administrative costs with common management tools and interfaces**
- **Designed to provide enterprise class reliability to help support continuous operations**

DS8000 – R2 Announcement Summary

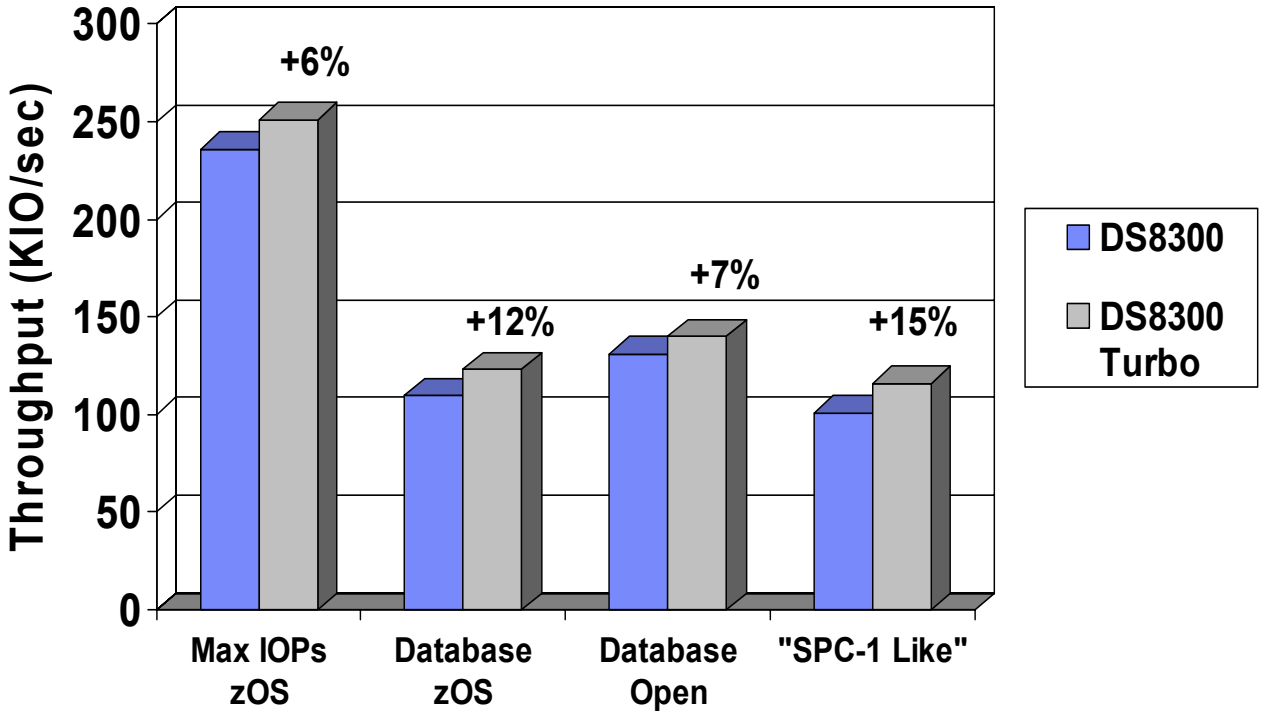
Announcing new models - Turbo Models 931, 932, and 9B2

- Supports previously announced features / functions, as well as all new R2 features / functions

Announcing new features for ALL models:

- ***IBM POWER5+ processor – standard on Turbo models***
  - Compared to the current IBM POWER5 processor, the POWER5+ processor may enable up to 15% performance improvement in I/O operations per second in transaction processing workload environments
- ***4Gb FCP / FICON adapter***
  - Designed to offer up to 50% improvement in a single port MB/second throughput performance, helping to enable cost savings with potential reduction in the number of host ports needed.
- ***500GB 7,200 rpm FATA drives***
  - Drives can be added to DS8000 series models to support various fixed content, data archival, reference data, and near-line applications that require large amounts of storage capacity at lower cost per MB
- ***3-site Metro / Global Mirror***
  - Provides fast failover / failback to any site, fast re-establishment of 3-site recovery (without production outages), and quick resynchronization to any site with incremental changes only (incremental resyncing)
- ***Earthquake resistance kit***
  - Seismic kit for stabilizing the storage unit rack, Feature needed on each frame
- ***Ethernet adapter pair (for TPC RM support)***
  - TPC for Replication provides management of DS8000 series business continuance solutions, including FlashCopy and Remote Mirror and Copy functions.
- ***Performance Accelerator (Models 932, and 92E only)***

DS8300 vs. DS8300 Turbo



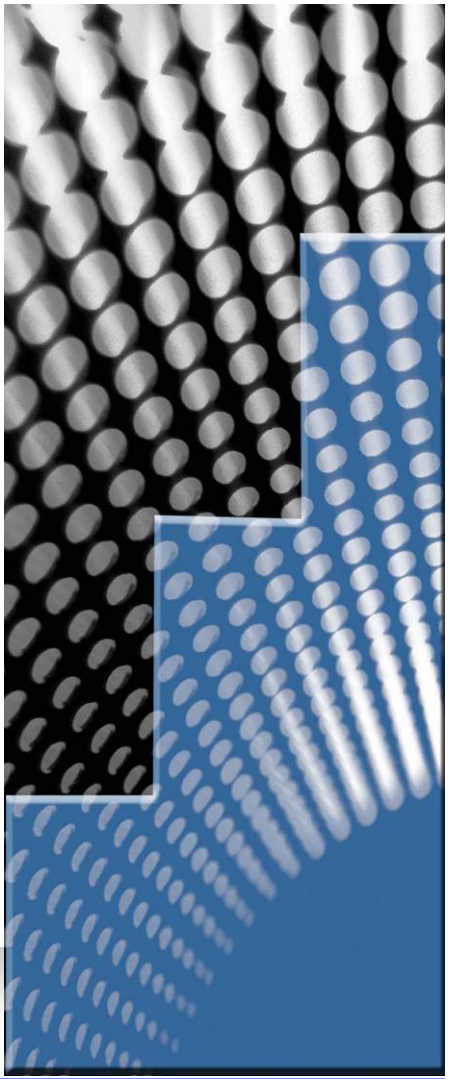
DS8300 RAID 10 – (32) 2Gbit FC ports and 256GB cache

DS8300 Turbo RAID 10 – (32) 2Gbit FC ports and 64GB Cache

** The "SPC-1 Like" tests reported used the same mix of I/O workload components as defined in the SPC-1 specification, but were not performed under the conditions required for SPC-1 audit certification.*

IBM System Storage Disk Delivers New Functions to Support a Cost-Optimized, Tiered Infrastructure

- New lower cost disk drive alternative for IBM DS6000 and DS8000
  - 500 GB 7200 rpm Fiber Channel ATA (FATA) disk drive for nearline usage supports low cost and high capacity scalability
  - Enables 2nd tier of disk storage within the disk subsystem
    - Integrates into ILM philosophy to manage data placement based on access need/value throughout its lifecycle
  - Helps to reduce costs for customers needing large amounts of less frequently used data. Good storage for backup data, archiving, document imaging and retention, reference data
  - IBM FlashCopy® and Remote Mirroring functions designed to support copying data from Fiber Channel drives to FATA drives and vice versa



Good complement to the IBM System z9 Business Class (z9 BC) for cost effectiveness

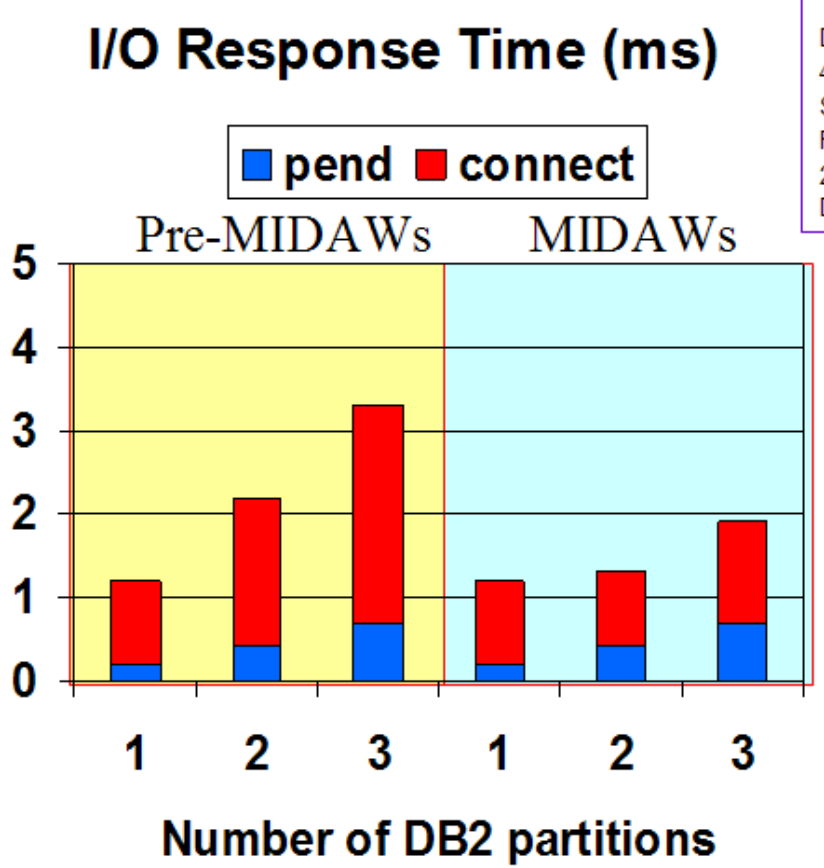
What is the MIDAW Facility?

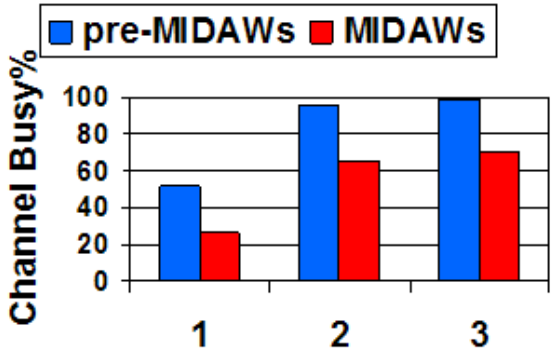
- **MIDAW (Modified Indirect Data Address Word) facility – new system architecture and software exploitation designed to improve the performance of many applications by reducing FICON channel, director, and control unit overhead.**
  - Requires the System z9 processor and z/OS® 1.7 or 1.6 + APARs
    - APARs OA10984, OA13324, OA13384
  - Implemented by Media Manager
  - Can significantly improve FICON performance for certain applications such as DB2 sequential workloads that use Media Manager to process small records especially with extended format datasets
    - For the same amount of FICON channel processor utilization, achieve higher levels of MB/sec
    - Allow more channel programs to benefit from faster link speeds
    - Increase the efficiency of the links by transferring more data frames and less command frames
    - Increase the efficiency of the cu port host adapters

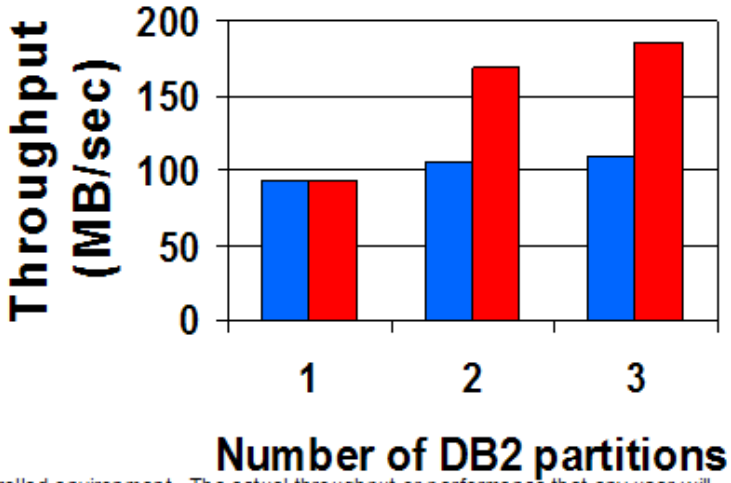
Parallel DB2 Table Scan, Non-EF 4K (single channel)

Configuration:
MIDAW : z/OS 1.7
Pre-MIDAW: z/OS 1.4

DB2 for z/OS Version 8
4000 byte row size
System z9 109
FICON® Express2
2 Gbit/sec link
DS8000 control unit







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 numbers stated here.

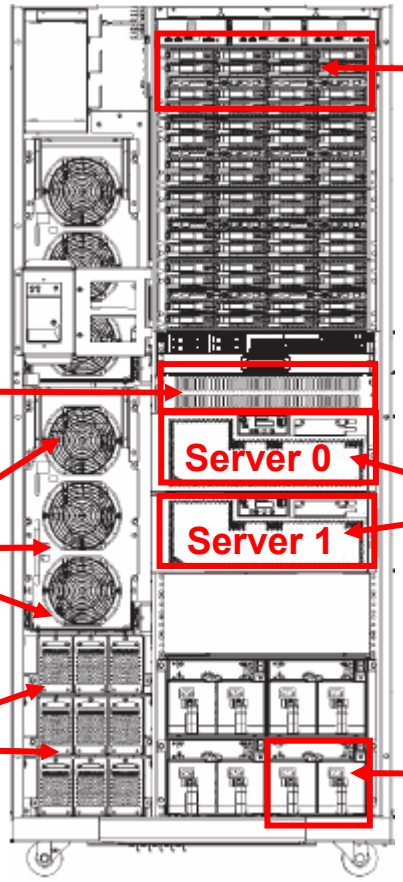
DS8000 Hardware Performance Considerations

- **Front-end performance considerations**
  - I/O ports
  - Host adapters (HAs)
  - I/O enclosures
  - Servers (processors and memory)

- **Backend performance considerations**
  - Ranks (and array type)
  - Device Adapter (DA) pairs
  - Servers (processors and memory)

DS8000 Hardware Components

DS8000 Front

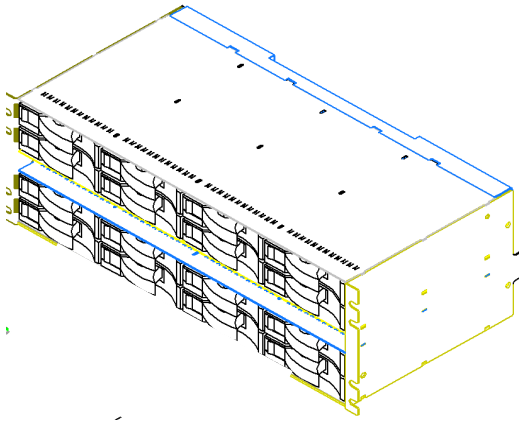


Storage Hardware Maintenance Console
Configuration
Guided Maintenance/Call Home
Supports Multiple Systems

Redundant Power

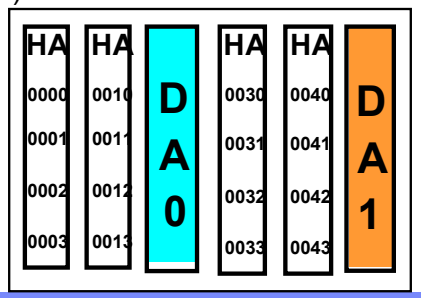
Battery

Dense HDD Packaging
16 drives per pack
128 drives in base rack



IBM eServer p5 570
Dual 2-way or Dual 4-way
Processor complexes
- Logical volume management
- Processors
- System memory (Cache)
- Persistent memory (NVS)

4 I/O Bays
- Each bay supports
- 4 Host Adapters and
- 2 Device Adapters



I/O Enclosure

DS8000 Memory Considerations

- Processor memory
  - *System memory ("Cache")*
    - Primarily affects read performance
  - *Persistent memory ('NVS')*
    - Primarily affects write performance
    - DS8000 persistent memory scales with processor memory size

Processor Memory	Persistent Memory	Support
16GB	1GB	2107 Model 921 only
32GB	1GB	
64GB	2GB	
128GB	4GB	
256GB	8GB	2107 Models 922/9A2 only

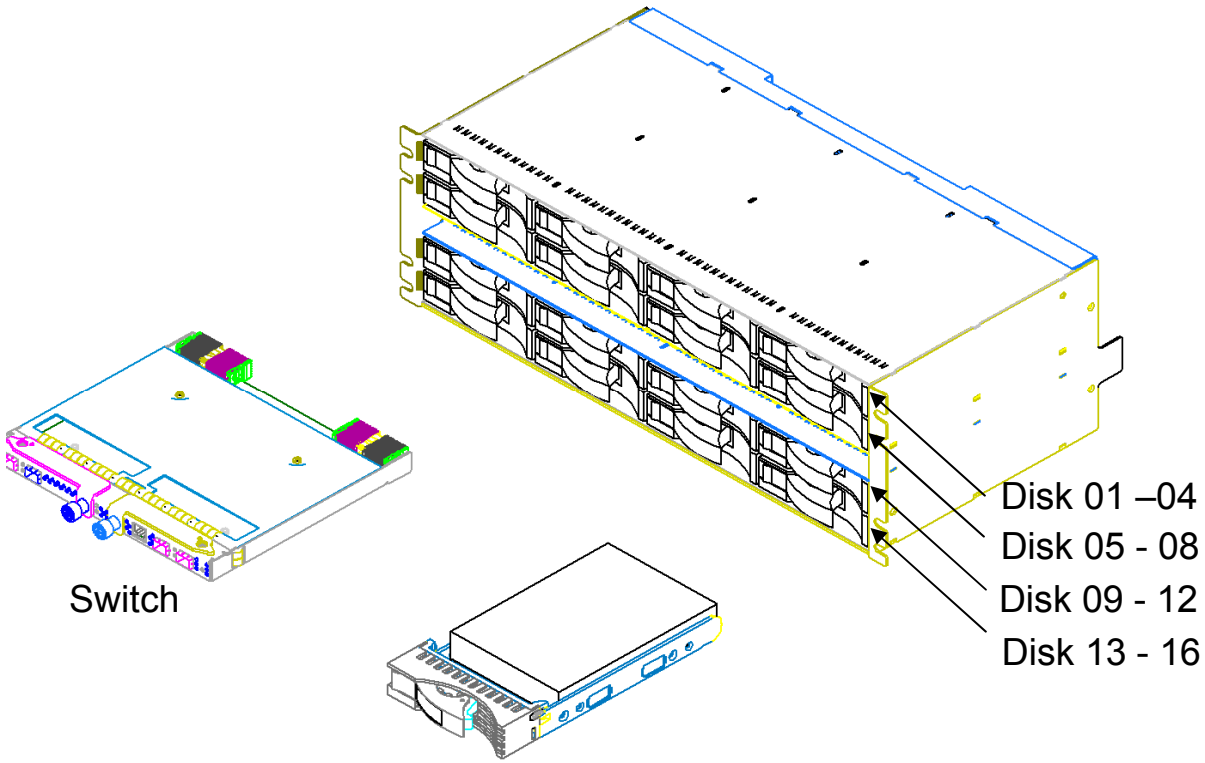
DS8000 Disk Enclosure Pair

Disk enclosure

- 16 disks
  - Cabled to Device Adapter

Disk enclosure pair

- 2 adjacent disk enclosures
  - Front (16 disks)
  - Rear (16 disks)
- Maximum of 32 disks



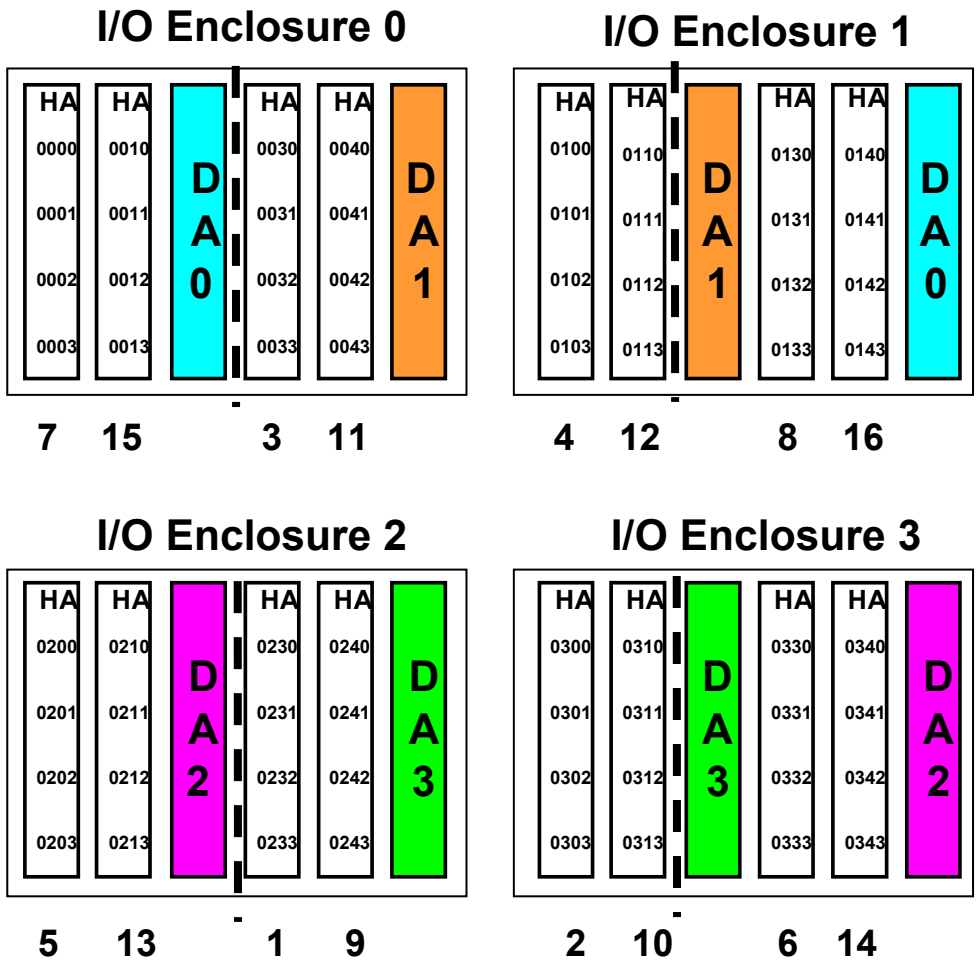
DS8000 Host and Device Adapter Ports

Host Adapter

- *Connects host systems to DS8000 servers*
- *I/O Port*
  - Cabled to z/OS or open systems server(s) directly or via Storage Area Network (SAN)
  - Host adapters are typically ordered in pairs for availability
- *4-port FCP/FICON Host Adapter*
  - Each I/O port may be configured to support either FCP or FICON protocol
  - For high performance connections, 2 I/O ports per FCP/FICON Host Adapter recommended
    - z/OS FICON Express2 to DS ratio of 1:1recommended
    - z/OS FICON Express to DS ratio of 2:1 may be acceptable

Device Adapter

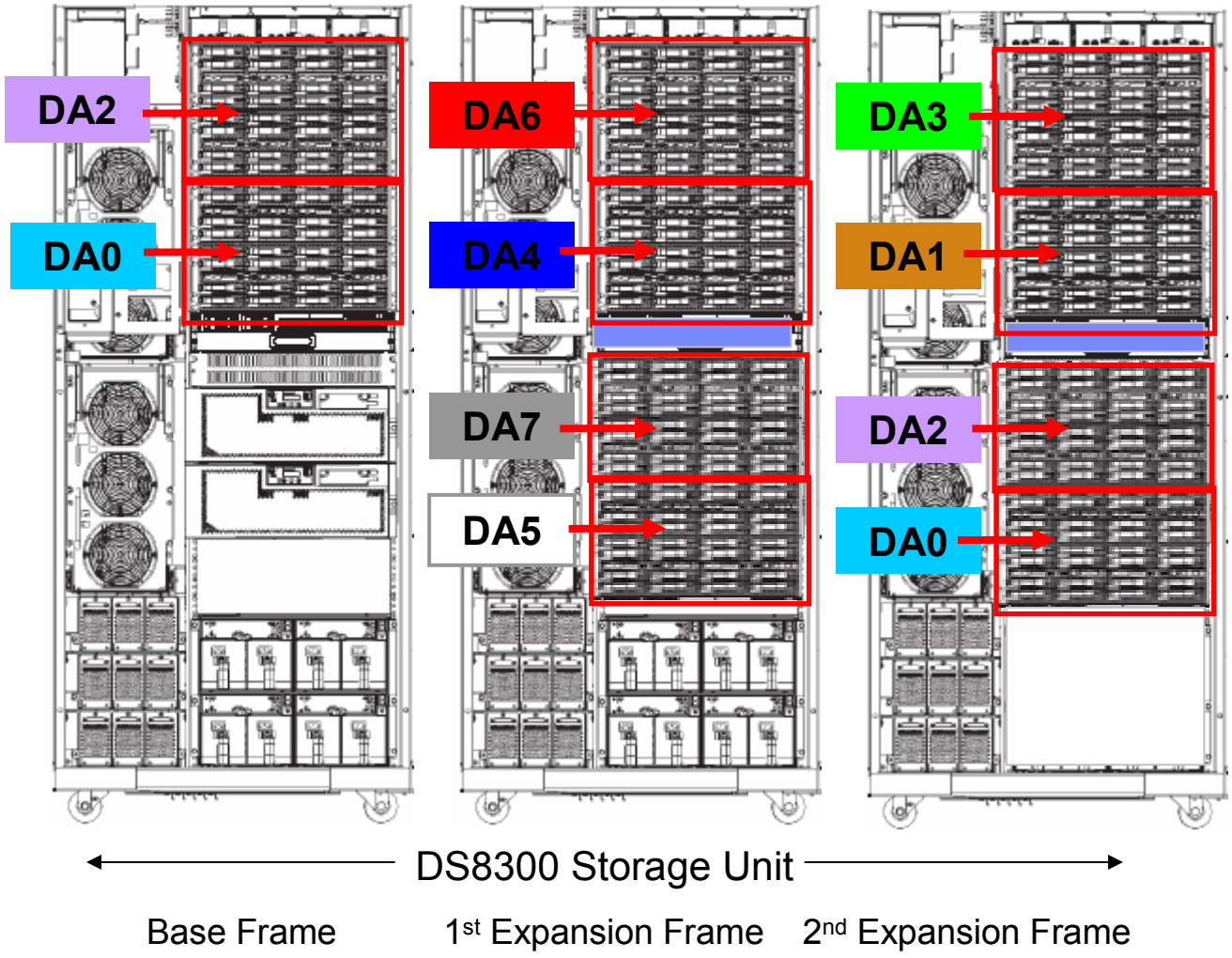
- Connects disk enclosures to DS8000 servers
- Cabled to Fibre Channel switches in disk enclosures
- Device adapter pair
  - 2 Device adapters in adjacent (left and right) I/O enclosures supporting the same disk enclosures



DS8300 Storage Unit

2107 Model 922

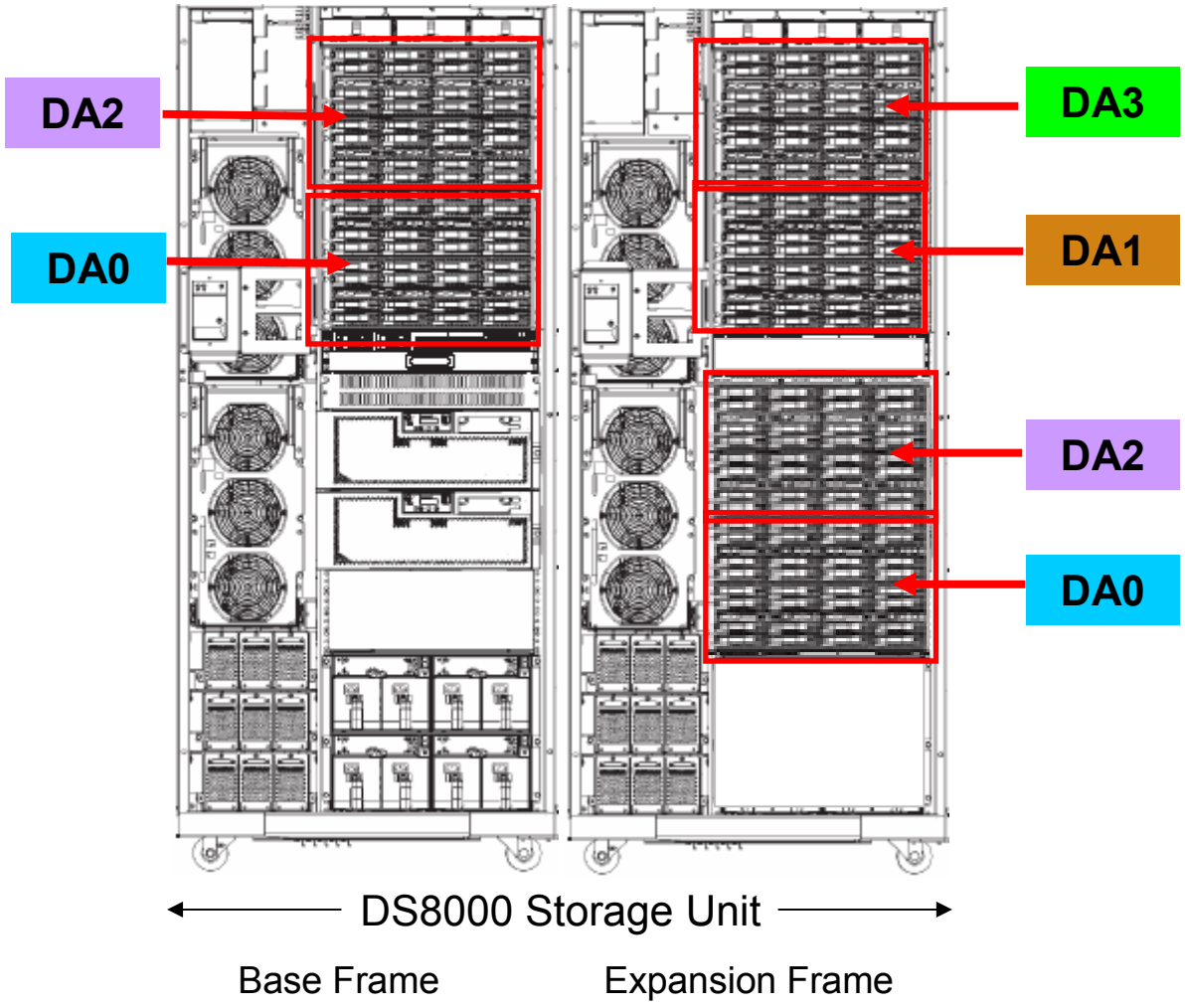
- 2 server I/O loops
- Maximum of 8 DA pairs
- Order of installation of disks on DA pairs
  - 2,0,6,4,7,5,3,1,2,0
  - 64 disk increments
  - Disks ordered per frame
  - Larger disks installed first
- Fully-populated 922
  - 640 disks
  - DA2 and DA0 - 128 disks each
  - Other DA pairs – 64 disks each
- 922 with 512 disks
  - Balanced configuration
  - Uses all 8 DA pairs equally



DS8100 Storage Unit

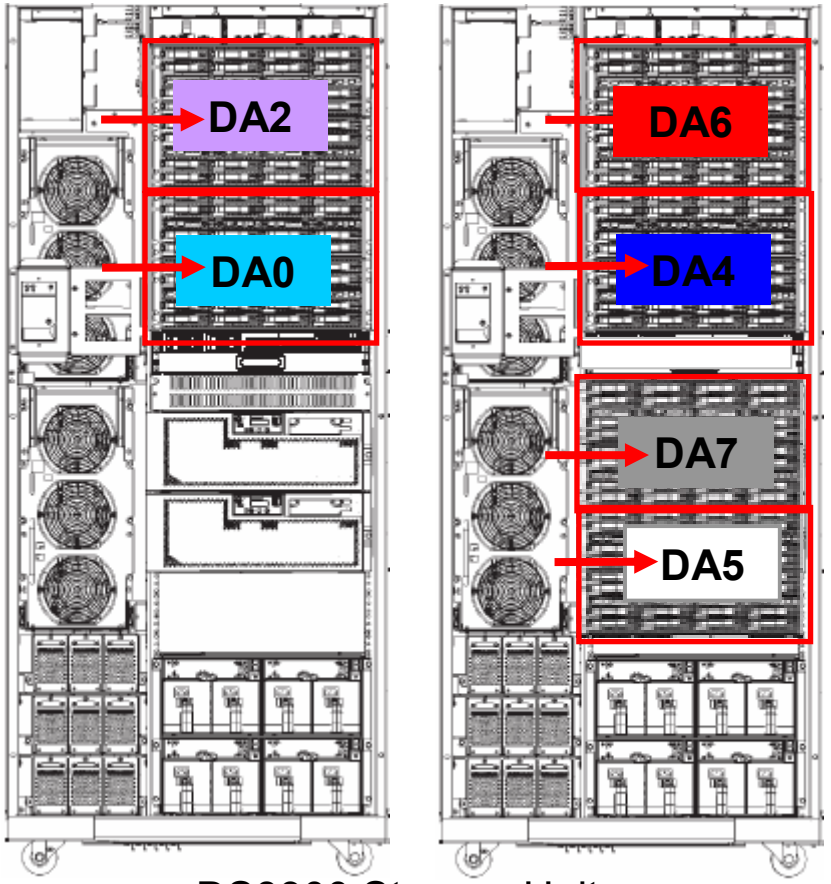
2107 Model 921 Disk Enclosures and Device Adapter Pairs

- One server I/O loop
- Maximum of 4 DA pairs
- Installation order of disks on DA pairs
  - 2,0,3,1,2,0
  - 64-disk increments
  - Disks ordered per frame
  - Larger disks installed first
- Fully-populated 921
  - 384 disks
  - DA2 and DA0 - 128 disks each
  - DA1 and DA3 - 64 disks each
- 921 with 256 disks
  - 'Balanced configuration'
  - Uses all 4 DA pairs equally



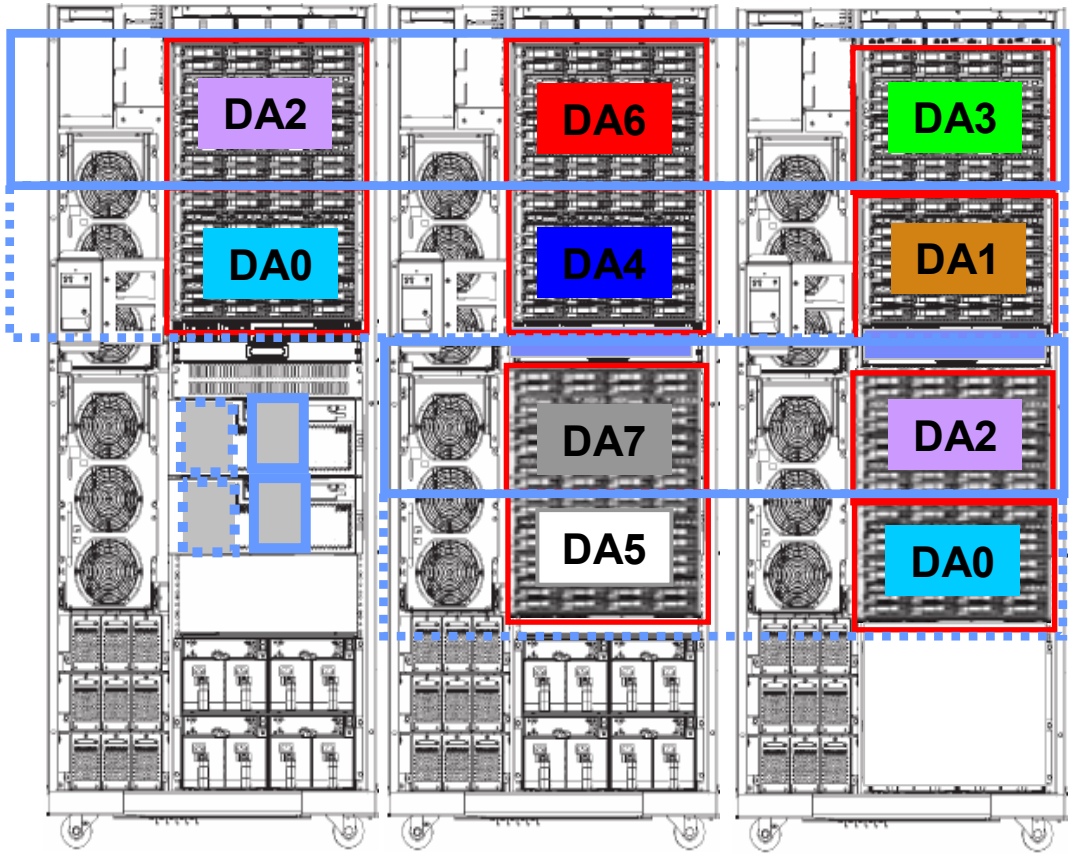
DS8300 Application Performance Accelerator RPQ

- **For workloads requiring high bandwidth but not a large number of disk drives**
  - Normally, 64 disks required per DA pair before adding next DA pair
  - RPQ requires only 32 disks per DA pair
  - Currently DS8300 model 922 A+B frames only
- **Allows use of more device adapters to achieve higher bandwidth for a given amount of disk capacity**



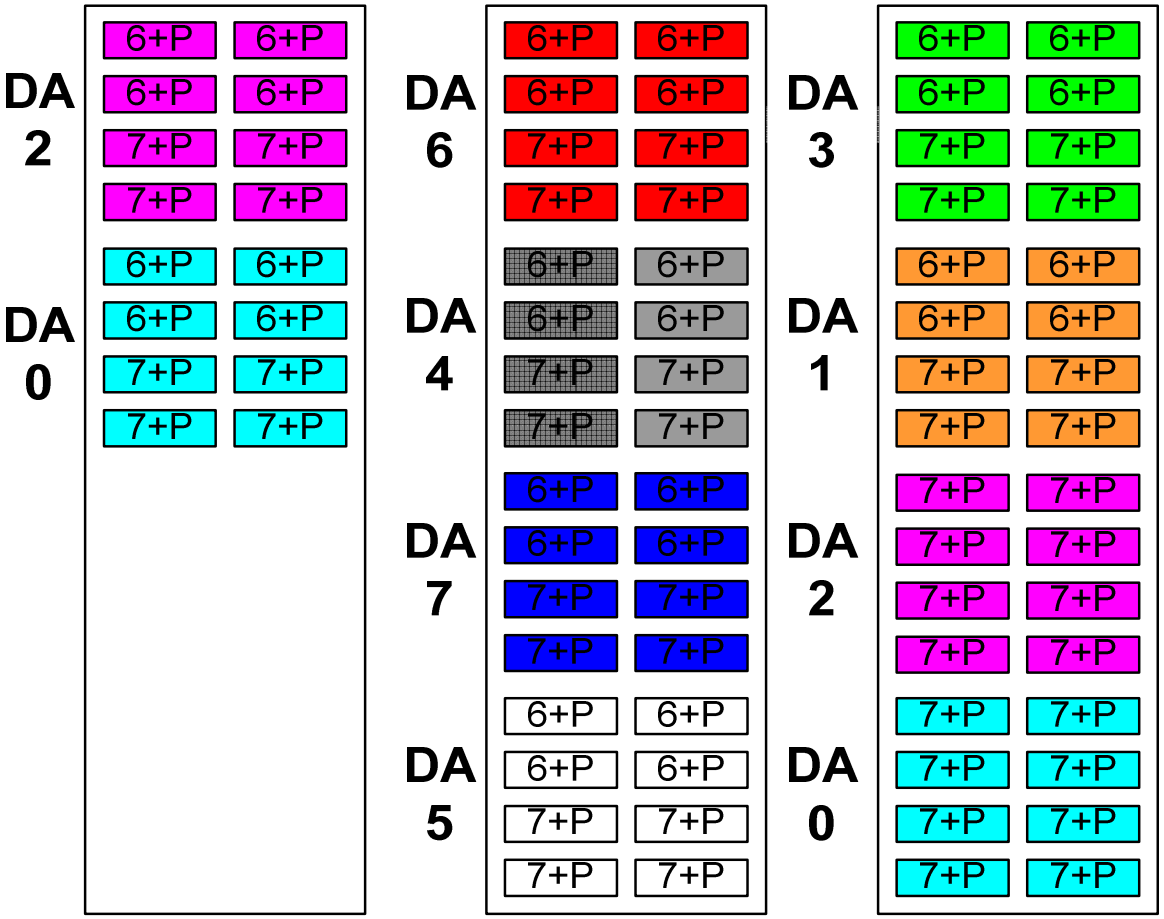
DS8000 Storage Image
2107 9A2

- Dual Storage Image DS8000 (2107 Model 9A2)
  - — Dedicated 50/50 resource split
    - Processors
    - Memory
    - Disk enclosures
    - I/O enclosures
    - Different numbers of disks and host adapters may be installed in the enclosures dedicated to each Storage Image
- Identified by serial number (s/n)
  - –DS8000 -- Storage Image ID ends in 1 or 2
    - *Dedicated to Storage Image 1*
    - — *DA0, DA4, DA5, DA1, DA0*
  - –*Dedicated to Storage Image 2*
    - —DA2, DA6, DA7, DA3, DA2



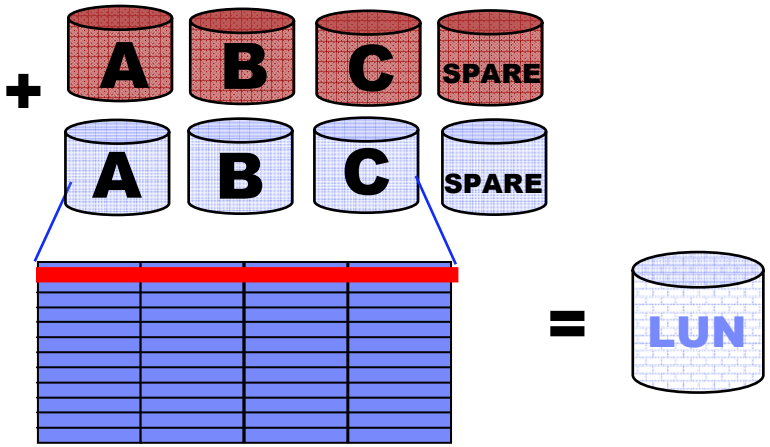
2107 Model 922 RAID5 Example
(Single Disk Type)

- **Each block represents one 8-DDM RAID5 array**
- **Minimum of 4 spares per DA**
  - –First 4 arrays created on DA contain spares
- **Single disk type example**
  - –Mixed disk types or RAID types may result in creation of more 6+P arrays
- **Ranks with spares are balanced across server0 and server1 extent pools**
- **Ranks on each DA are balanced across server0 and server1extent pools**



DS8000 Hot Spares

- **DS/8000**
  - DDMs purchased in DDM Install Groups of 16 DDMs
- **A minimum of one spare is required for each Array Site defined until the following conditions are met**
  - Minimum of 4 spares per DA pair
    - *the spares are balanced between the two device interfaces*
  - Minimum 4 spares of the largest capacity Array Site on the DA pair
    - *the spares are balanced between the two device interfaces*
  - Minimum 2 spares of capacity and rpm greater than or equal to the fastest Array Site of any given capacity on the DA pair
    - *spares are balanced between the two device interfaces*
- **All spares are available to all Array Sites on that DA pair**



Disk Type	RAID 10	
	Idle	Moderate workload
73 GB 15K	0:20	0:50
146 GB 10K	0:45	1:50
146 GB 15K	0:45	1:40
300 GB 10K	1:15	3:40
500 GB 7.2K	2:40	6:45

DS8000 Specifications

	DS8100	DS8300
DDMs	16-384	16-640
DDM Interface	FC-AL	FC-AL
DDM Types	73,146,300,500 GB	73,146,300,500 GB
RAID Types	RAID 5,10	RAID 5,10
Max Capacity w/146 GB DDM	56 TB	93.4 TB
Max Capacity w/500 GB DDM	192 TB	320 TB
Max Sequential Bandwidth	2 GB/s	4 GB/s
LUNs/CKDs	64K Total	64K Total
Max N-Port Logins/Port	510	510
Max Process Logins	2K	2K
Max Logical Paths / CU	512	512
Max LUN Size	2 TB	2 TB
Dynamic Provisioning	Add/Del	Add/Del
Cache // NVS	16-128 GB // 1-4 GB	32-256 GB // 1-8GB
Processor	DS8000 ML (SMT) 2 Way	DS8000 ML (SMT) 4 Way
Host Adapters	ESCON x2 FC(4 Gb/s)x4	ESCON x2 FC(4 Gb/s)x4
Host Adapter Slots	16	32
Max Host Adapter Ports	64	128
Interface Protocols	SCSI-4GB or 2 Gb FCP/FICON	SCSI- 4Gb or 2 Gb FCP/FICON
PPRC Fabric	FCP	FCP
DA Slots	8	16
DA Throughput	720 MB	720 MB

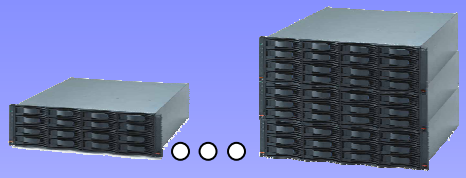
IBM DS6800 – Enterprise Disk Functionality

- **DS6000 Models - 522/EX2**
  - 3U package for controllers and 7 disk expansion units
  - High storage density footprint -16 drives per 3U package
  - Maximum of 1024 host and 8192 LUNs (2TB max)
  - **One year, same-day, 24x7 IBM onsite repair (IOR) warranty**
  - **500 GB 7,200 rpm FATA disk drives**
    - Provide additional price, and capacity flexibility to help address specific application and business requirements
    - Intermix is at the drawer level
  - **TPC Replication Manager Support**
  - **CALL HOME SUPPORT VIA TELEPHONE LINE**
    - Provides the ability to E-Mail a customer at Call Home

Key Differentiators

- zSeries and iSeries native attachment
  - Including Parallel Access Volumes support
- Same advanced software features as DS8000, ESS 800
- High availability features not generally found in midrange storage products

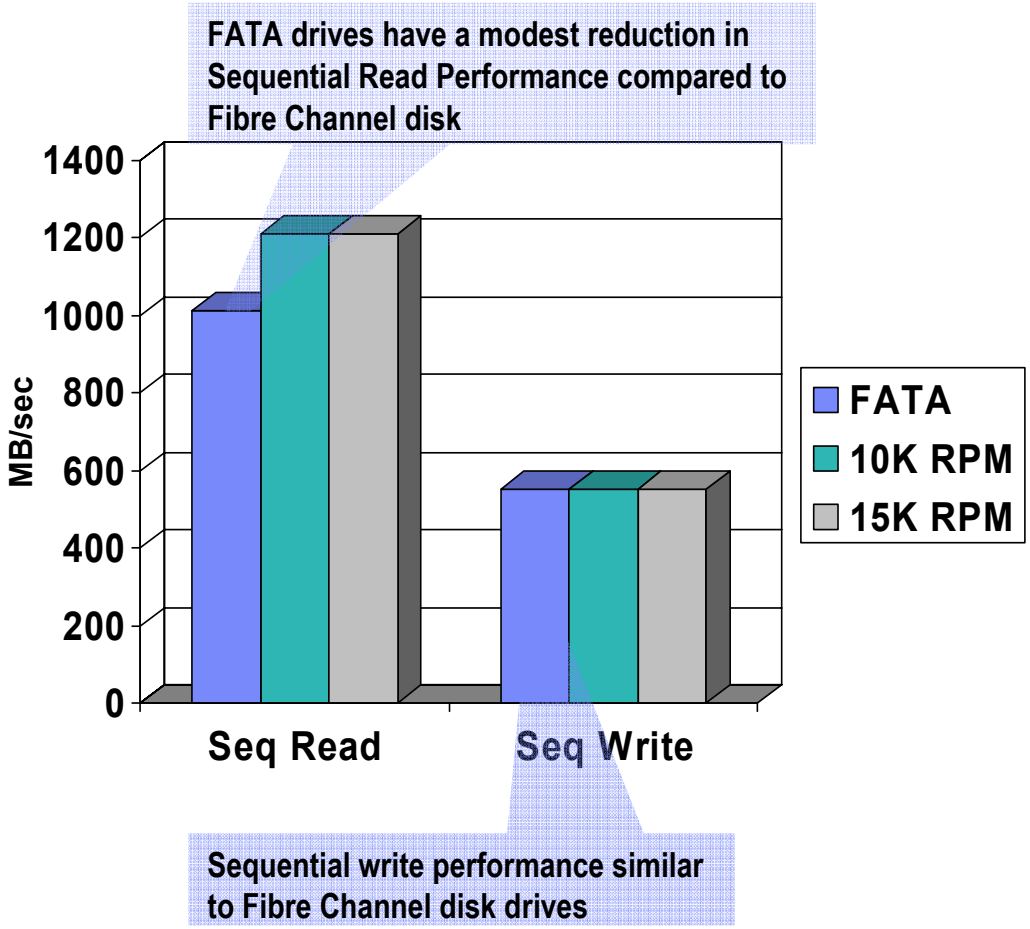
New Standard in Pricing and Packaging

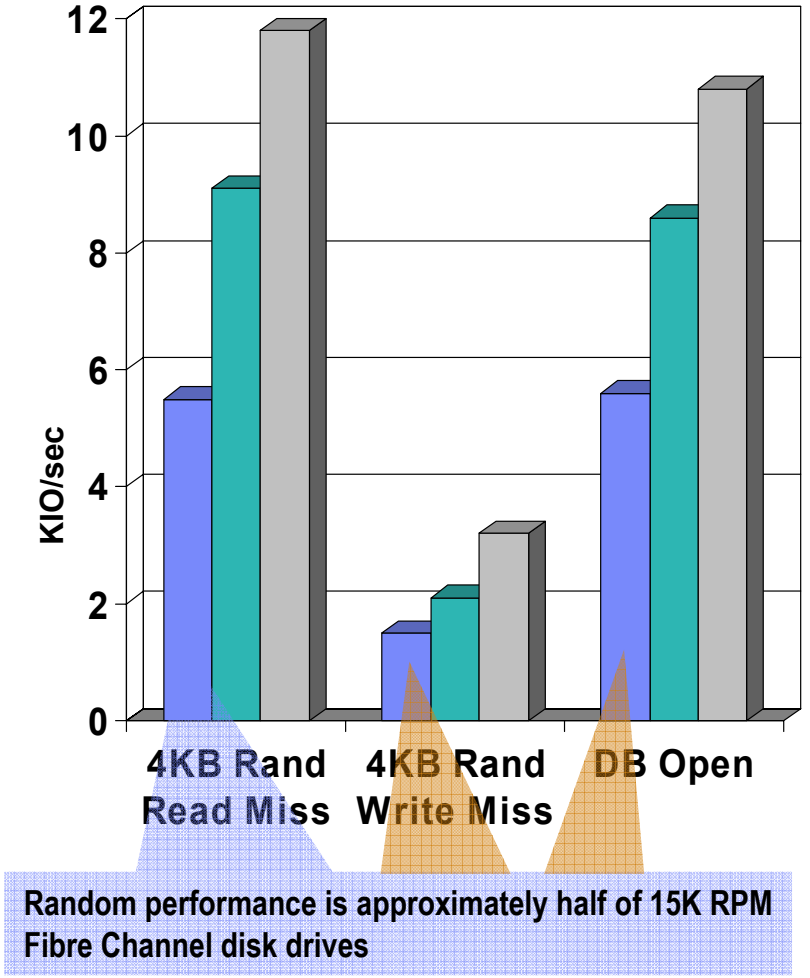


DS6000

- **Attractive price with all the capability of traditional enterprise products**
- **Designed for easy install, easy service, modular packaging**
- **Great performance in a small package**
- **Can start small and grow large in physical capacity – 38 TB**
- **Up to 8 FC/FICON® host ports, 4 GB Cache**

DS6800 Performance with FATA Disk Drives
with 64 disks





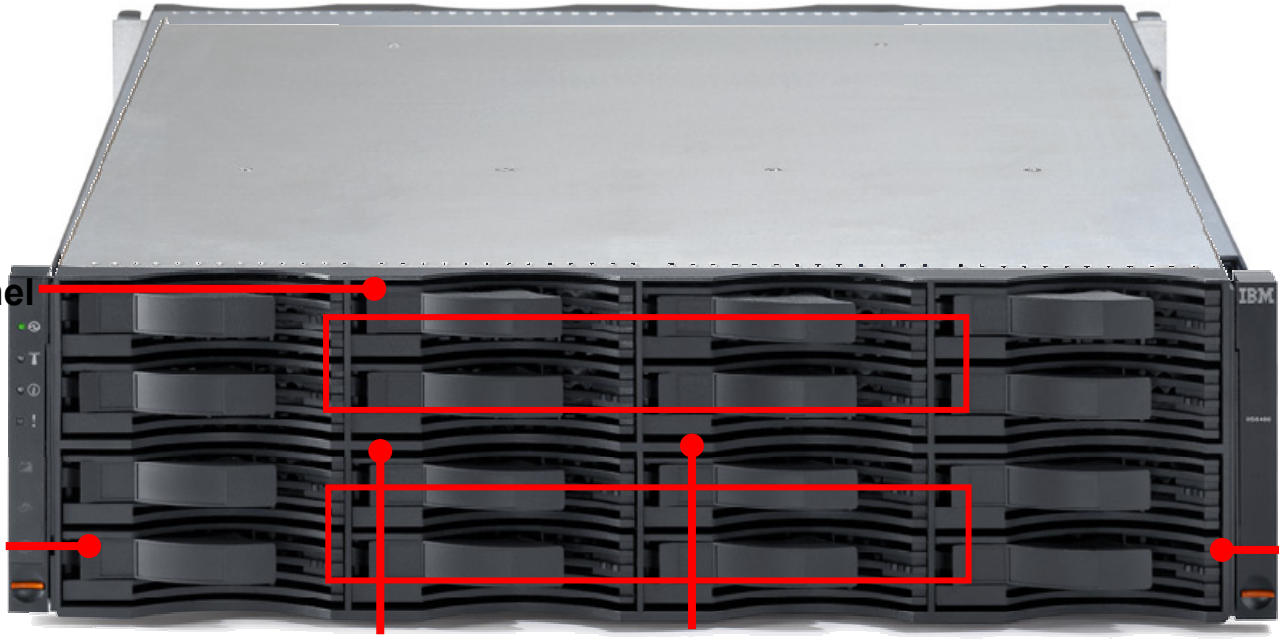
Note: Preliminary performance results; actual performance may change at GA

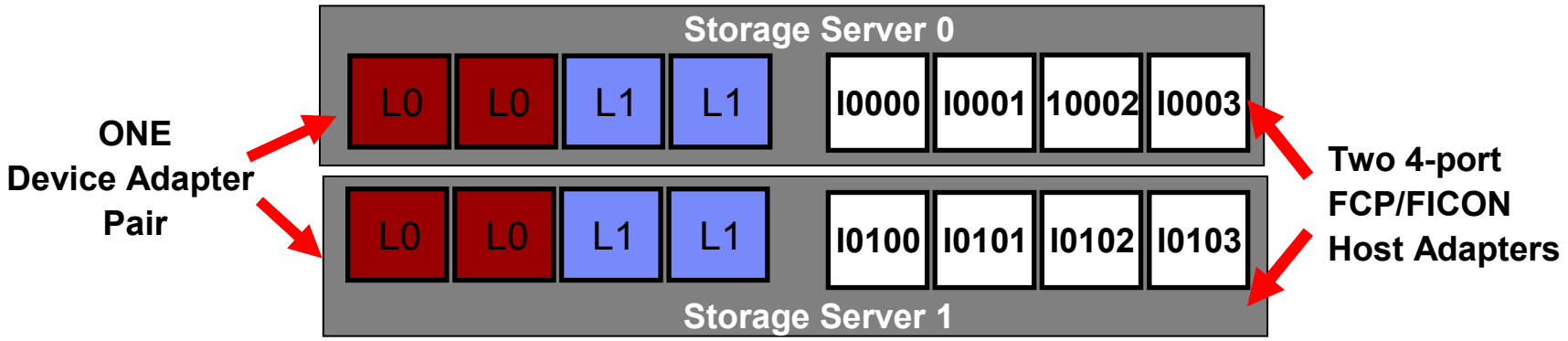
DS6000 Hardware Performance Considerations

- **Front-end performance considerations**
  - I/O ports
  - Host adapters
  - Servers (processors and memory)

- **Back-end performance considerations**
  - Ranks (and array type)
  - Loops
  - Servers (processors and memory)

DS6800 High Availability Design

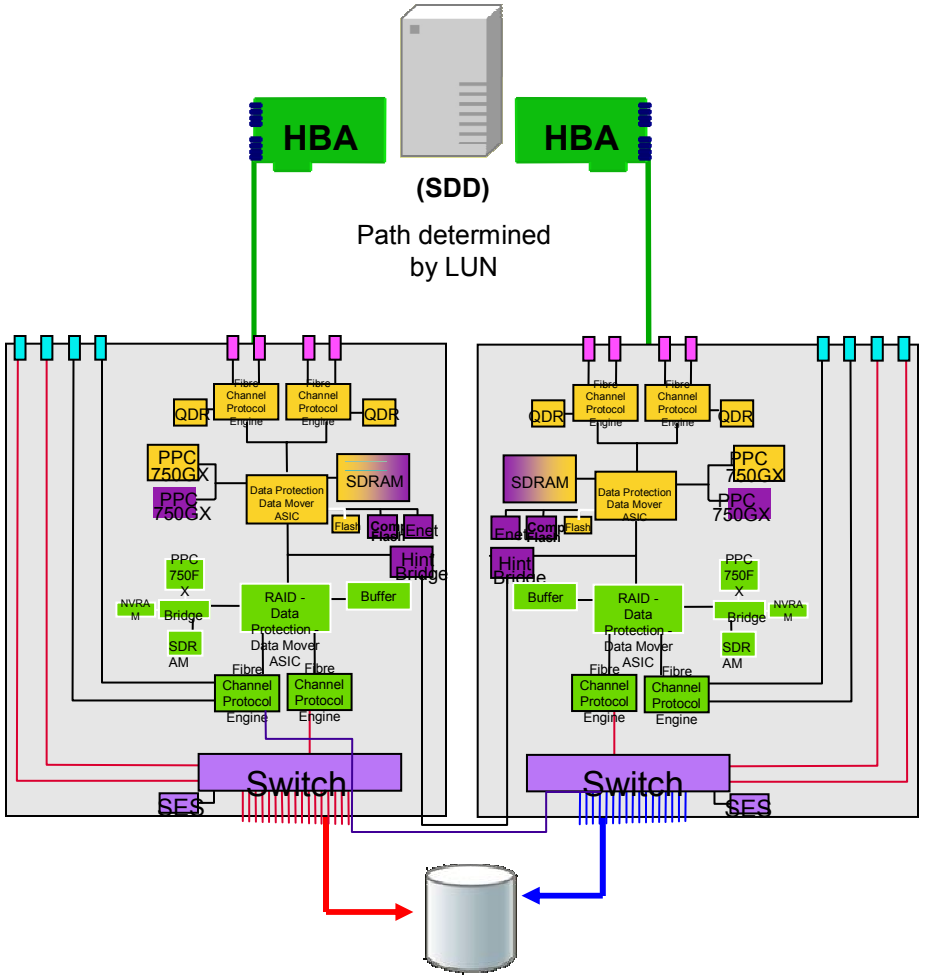




Preferred Path Data Operations

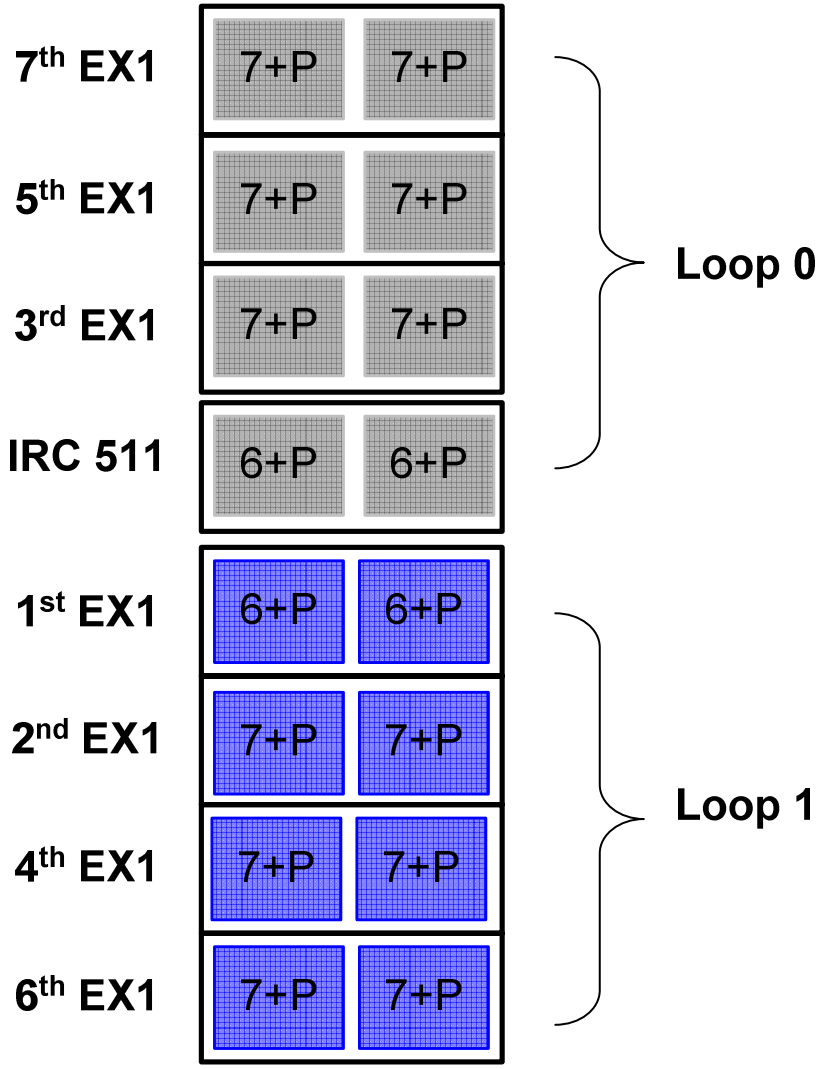
Path determined dynamically

- Chooses the most efficient and optimum path to store and retrieve data from the storage systems
  - Beyond load balancing
  - Our system selects the best path at each stage of the request
- *Competitors – can do preferred path, but path is selected at time of initial request*
- Preferred Path
  - Dynamically selects the most efficient and optimum path to use at each data interchange during read and write operations
  - A level above load balancing used by competition



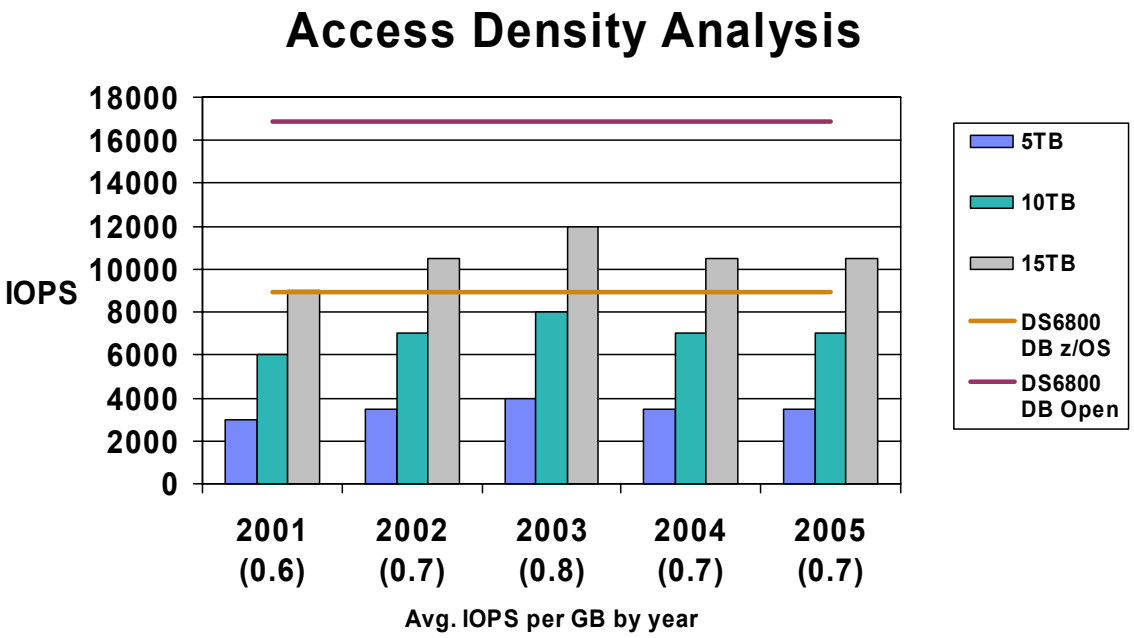
DS6000 RAID5 Example (Single Disk Type)

- **Each block represents one 8-DDM RAID5 array**
- **Minimum of 2 spares per 64 DDMs**
- **Spares are assigned to array sites and remain with those array sites, whether an array is created or not**
  - –Not necessarily on IRC and 1st EX1 as shown
  - –Array capacity should be confirmed after creation
- **Single disk type example**
  - –Mixed disk types may result in more 6+P arrays



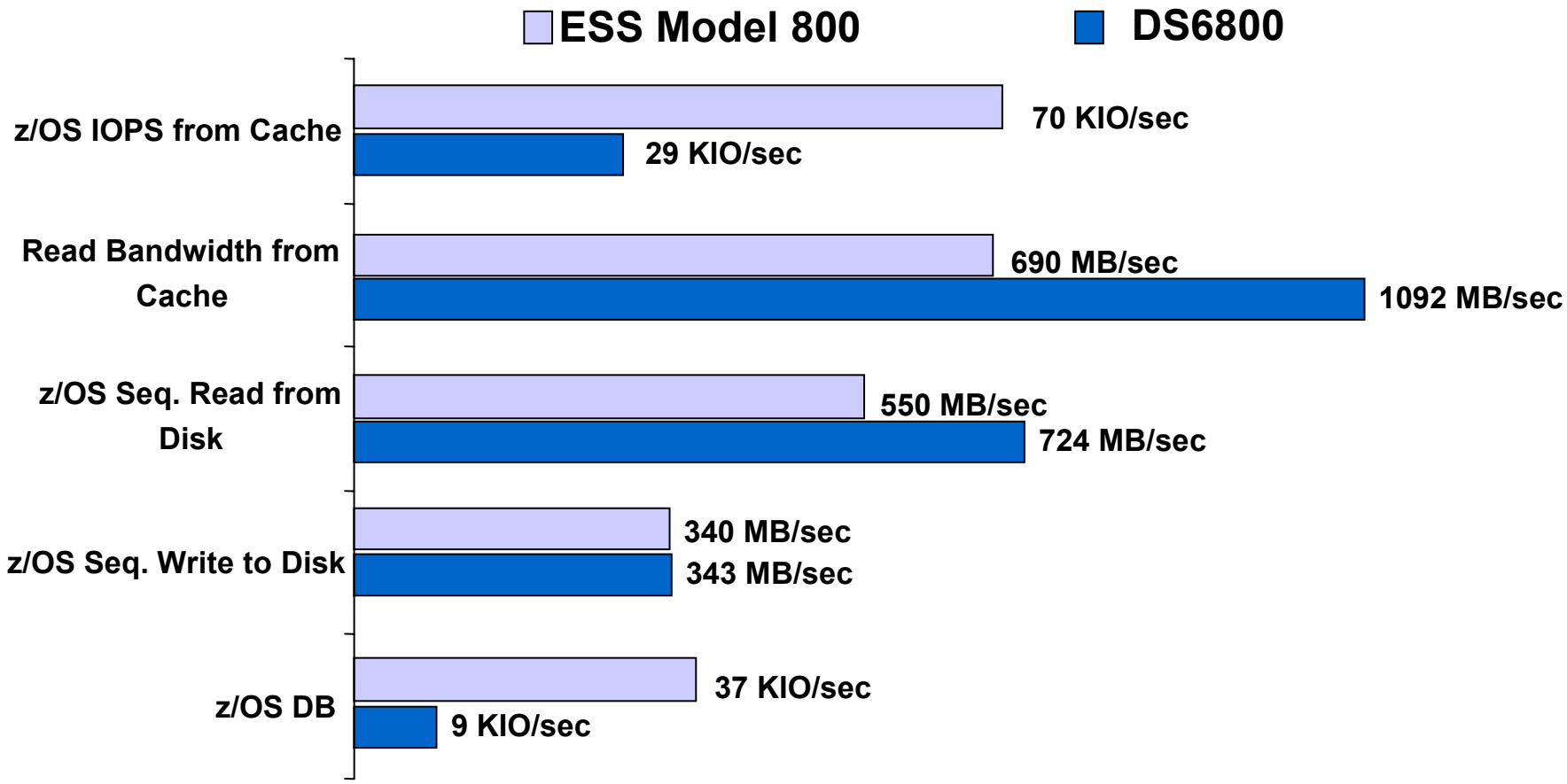
Performance in System z Environments

- **Access density – average IOPS per gigabyte of storage**
- **DS6800 exceeds the performance needs of smaller System z servers**
  - Bronze line indicates DS6800 IOPS in System z database environments
- **DS6800 has proven performance to handle open system database needs**
  - Dark red line



Performance Comparison – DS6800 to the ESS Model 800

IBM System z™ - Maximum Configuration



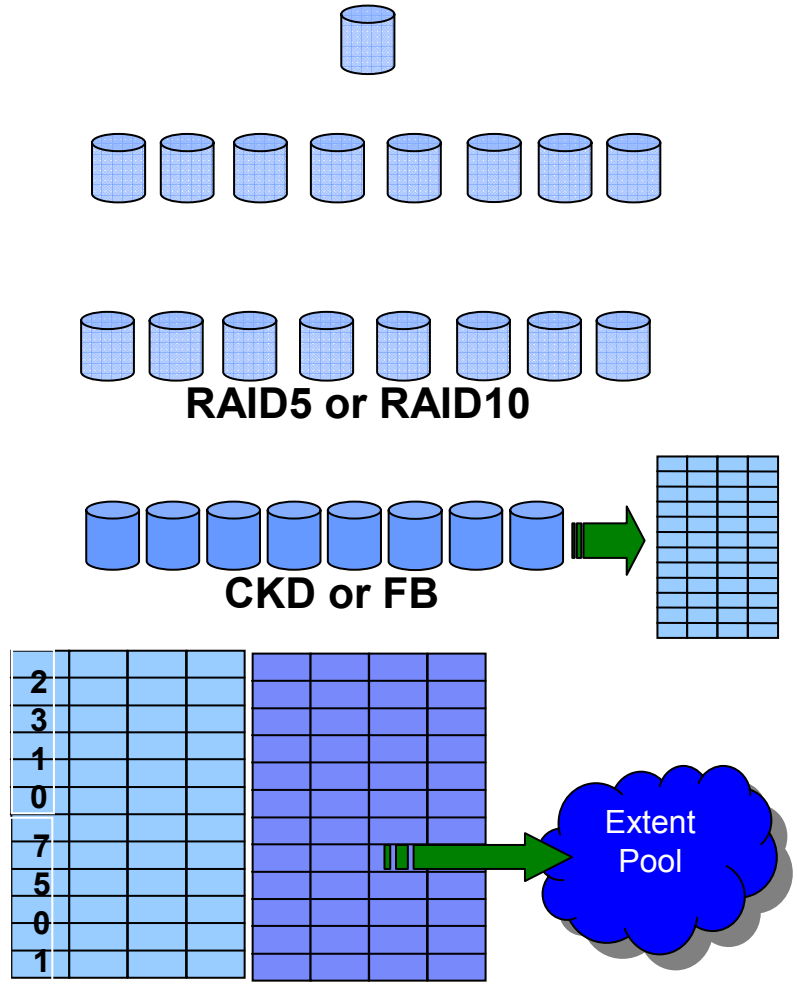
Full box numbers use maximum number of host ports per system (8 on DS6800, 16 on ESS 800). IOPS benchmarks use 4K Byte block transfers. MB/s benchmarks use 64KB transfers.

DS6800 Specifications

	DS6800
Controllers	**Dual Active**
Max Cache	**4 GB**
Max Host Ports	**8-Ports; 2Gb FC/FICON**
Max Hosts	**1024**
Max Storage / Disks	**224**
Disk Types	**FC 10K: 146GB, 300 GB** **FC 15K: 73, FATA 7.2: 500 GB**
Max Expansion Mod	**13**
Max Disk Loops	**4 (2 dual redundant)**
Max LUNs	**8192** **(up to 2 TB LUN size)**
RAID Levels	**5, 10**
RAID Array Sizes	**4 or 8 drives**
Operating Systems	**z/OS, i5/OS, OS/400, AIX, SUN Solaris, HP UX, VMWare, Microsoft Windows, Linux**
Packaging	**3U – Controller & Expansion Drawers**
Power consumption	**Controller: 0.69 kVA** **Expansion drawer: 0.48 kVA**

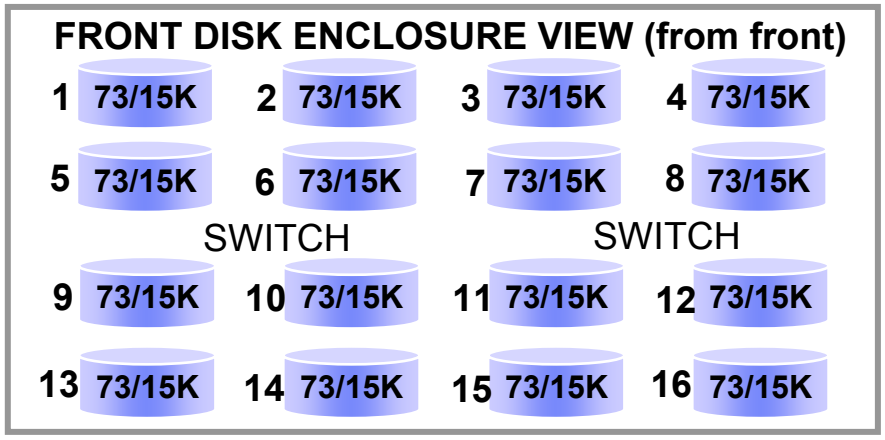
Storage Hierarchy

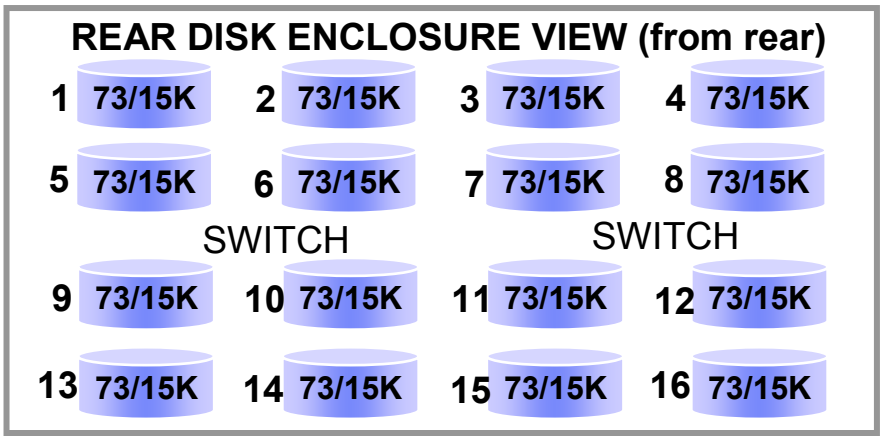
- Disk
  - Individual DDMs
- Array Sites
  - Pre-determined grouping of DDMs of same speed and capacity (8 DDMs for DS8000)
- Arrays
  - One 8-DDM Array Site used to construct one RAID array (DS8000)
- Ranks
  - One Array forms one CKD or FB Rank (8 DDMs)
  - ***No fixed, pre-determined relation to LSS***
  - Rank can be added to pool at any time
  - Rank can be removed from pool if no extents on rank are currently assigned to LUNs/volumes
  - Thresholds - Warning of space utilization
  - Reserve space option
- Extent Pools
  - 1-N Ranks form an Extent Pool (N=1/2 total ranks)
  - All Extents in a Pool are same storage type (CKD/FB); same RAID recommended
  - Associated with server0 or server1 (DS8000)



DS8000 Array Site

- Logical grouping of disks on which an array and rank will be built
  - 8 disks
  - Same capacity
  - Same speed
- Grouping determined by software
  - No fixed or pre-determined relationship to physical disk locations
  - During installation each DS8000 array site is initially created from disks in both the front and rear disk enclosures of a single disk enclosure pair (2 disk enclosures/32 disks)
  - After sparing, the DS8000 array site may include a disk from another disk enclosure pair on the same Device adapter (2-4 disk enclosure pairs/64-128 disks)
- Each DS8000 disk enclosure supports a maximum of 4 array sites

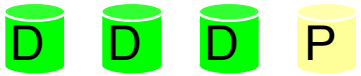




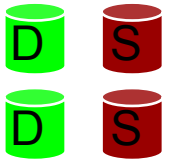
DS6800 Array Site

- **4-DDM arrays (DS6000 only)**
  - DS6000 – 1 array site
  - RAID5
    - 2+P (+ 1 spare)
    - 3+P (no spares)
    - Parity is striped across all disks in array but consumes capacity equivalent to one disk
  - RAID10
    - 1+1 (+ 2 spares)
    - 2+2 (no spares)

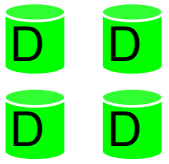
RAID5 2+P



RAID5 3+P



RAID10 1+1



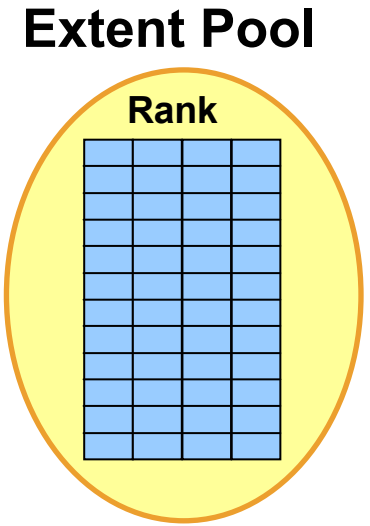
RAID10 2+2

Rank Considerations

- A rank (and the underlying RAID array and array site) has no relationship to server0 or server1 until after it has been assigned to an extent pool
  - Rank ID (Rx) does not indicate a server association unless specifically configured to do so
  - Adopts server0/1 affinity when assigned to an extent pool
  - Acquires LSS when volumes are created
- Ranks built on arrays containing spares should be balanced across server0 and server1 extent pools
- RAID array with storage type defined
  - CKD or FB
  - Logically divided into *extents*, the units of space allocation from which volumes will be created
    - FB - 1GB
    - CKD - 1113 cylinders (3390M1; .94GB)
- One-to-one relationship between array & rank
  - One RAID array becomes one rank
  - DS8000 – 8 DDMs

Define an Extent Pool

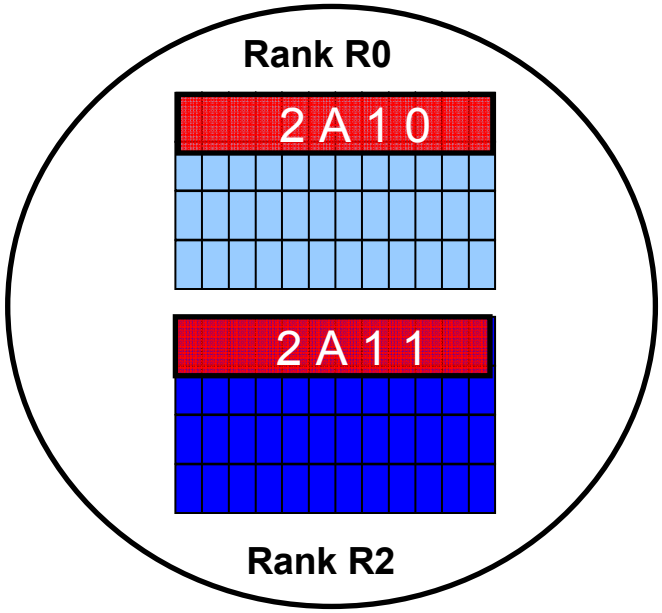
- Logical grouping of extents from one or more ranks
  - One or more ranks are user-assigned to an extent pool
    - Extents of assigned ranks make up the extent pool
  - Logical volumes are created from the extents in the pool
  - *Considerations:*
    - Adopts server0/1 affinity when assigned to an extent pool
    - Acquires LSS when volumes are created
    - Single extent pool is recommended unless larger LUNs are required
- Extent pool is one storage type
  - 1 extent type
  - Either CKD (z/OS) or FB (open systems)
- Extent pool is user-assigned to either server0 or server1
  - Volumes created from extents in pool will have an affinity to either server0 or server1
    - Volumes are accessible to either server for availability
  - Pools assigned to server0 will have 'even' IDs (e.g. P0, P2)
  - Pools assigned to server1 will have 'odd' IDs (e.g. P1, P3)
  - Extent pools and ranks assigned to pools should be balanced across server0 and server1



Extent pools should be balanced across server0 and server1

Multi-Rank Extent Pool Considerations

- All extents for a volume are allocated on one rank if possible
  - –May change in the future
  - –No striping of extents for a single volume across multiple ranks
  - –Volume may ‘spill’ across ranks if not enough extents are available on one rank
  - –Volume may be larger than a single rank
- Each new volume is allocated on the rank with the most free extents
  - –May change in the future
  - –Standard volume size will result in round robin allocation of volumes across ranks
  - –Same LSS will be used on multiple ranks
  - –DSCLI *showfbvol, showckdvol & showrank* commands may be used to map volumes to ranks



Logical Subsystem/Logical Control Unit Considerations

- For open systems, LSSs do not directly affect application performance
- For z/OS, more LCUs will provide additional addresses for PAVs which can improve performance
- One or more unique LSS/LCU IDs consistently associated with a specific rank and DA or application may simplify performance analysis
- Using one LSS/LCU across multiple ranks
  - Allows cross-rank copy services consistency grouping
  - May make performance analysis more complex

Logical Subsystem (LSS)

- **An LSS is a logical grouping of 256 volume addresses**

- **DS6000 and DS8000 use 4-digit hexadecimal volume IDs (volume x'210A')**
  - 1st two digits are the LSS ID
    - x'21' (LSS x'21')
    - The DS6000 has a maximum of 32 LSSs (x'00' – x'1F')
    - The DS8000 has a maximum of 255 LSSs (x'00' – x'FE')
    - Determine association with server0 (even LSSs) or server1 (odd LSSs)
  - 2nd two digits indicate volume within LSS
    - x'0A' (volume x'0A', the 11th volume in the LSS)
    - Each LSS has a maximum of 256 volumes (x'00' through x'FF')

- **16 LSSs in sequence (e.g. x'20'-x'2F') must be the same storage type**
  - Either CKD *or* FB
  - Called an Address Group
  - 1st digit of volume ID (and LSS ID) is the Address Group ID
    - X'21' (Address Group x'2')

Logical Subsystem (2)

- **Used by z/OS**

- **An Address Group (16 LSSs in numerical sequence) must be a single storage type (CKD or FB)**
  - CKD (z/OS)
    - CKD LSSs are commonly referred to as Logical Control Units (LCUs)
    - CKD LSSs (LCUs) are *explicitly* created through DS6000 or DS8000 LCU definition
      - User specified LSS/LCU ID

- **ESCON protocol (z/OS) is only able to access LSSs x'00' – x'0F' (4096 addresses)**
  - LSS x'00' – x'0F' (Address Group 0) must be CKD storage type for ESCON access
  - ***Do not create FB LUNs in LSSs x'00'-x'0F' if ESCON access will be required***

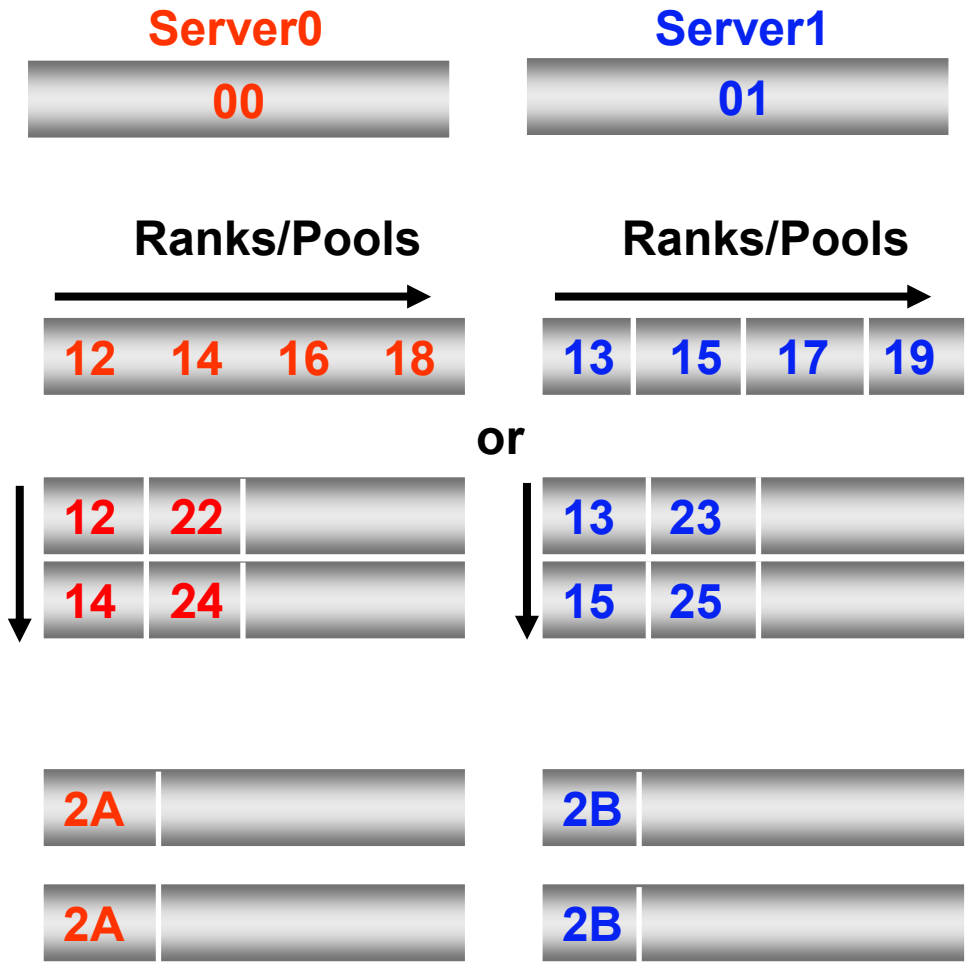
- **Parallel Access Volumes (PAVs)**
  - Additional 'alias' addresses for a volume for parallel access
  - May be shared *within* an LSS
  - Consume LSS addresses

Logical Subsystem (3)

- **An Address Group (16 LSSs in numerical sequence) must be a single storage type (CKD or FB)**
  - CKD (z/OS)
    - CKD LSSs are commonly referred to as Logical Control Units (LCUs)
    - CKD LSSs (LCUs) are *explicitly* created through DS6000 or DS8000 LCU definition
      - User specified LSS/LCU ID
  - FB (open systems)
    - FB LSSs are *automatically* generated during FB volume creation
    - ***Selection of FB volume ID determines LSS (1st 2 digits of volume ID)***
- **ESCON protocol (z/OS) is only able to access LSSs x'00' – x'0F' (4096 addresses)**
  - LSS x'00' – x'0F' (Address Group 0) must be CKD storage type for ESCON access
  - ***Do not create FB LUNs in LSSs x'00'-x'0F' if ESCON access will be required***

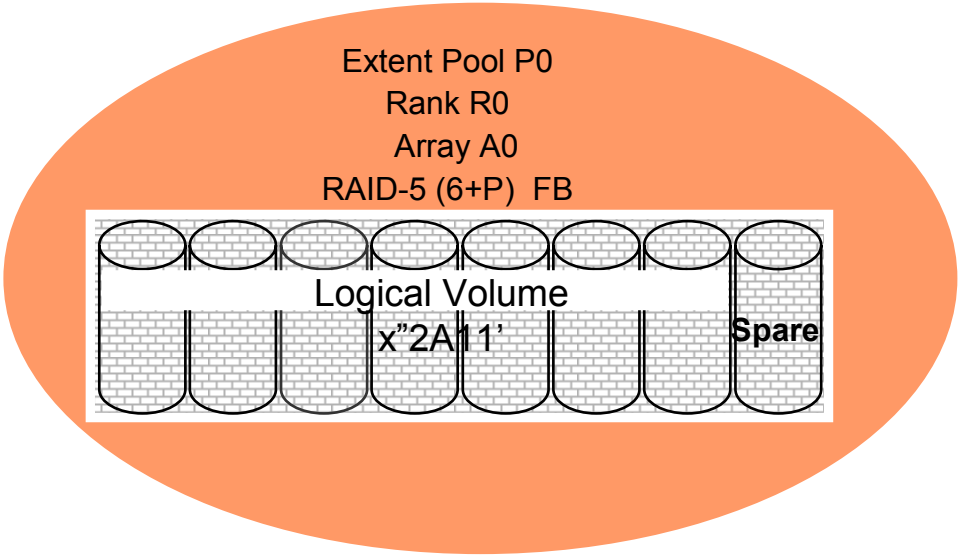
Logical Subsystems and Ranks – 3 Approaches

1. One unique LSS/LCU per rank
   - Simplifies management
2. Multiple LSSs/LCUs on one rank
   - Provides more addresses (& PAVs) for rank
   - Allows utilization of large drives with small volumes
3. One LSS/LCU on multiple ranks
   - Enables cross-rank Copy Services consistency group
   - This will be typical for DS8000 multiple-rank extent pools
   - May make management more complex



DS8000 / DS6800 Create Logical Volume

- Logical volume is striped across all disks in the RAID array/rank
- Volume capacity may be specified in
  - Binary GB
  - Decimal GB
  - Blocks (512 bytes)
- Logical volume is created from extents in one extent pool
- In a multi-rank extent pool, a logical volume is created from extents on one rank in extent pool if possible
  - Volume may be larger than a single rank
  - Volume may 'spill' across ranks in multi-rank pool
- In a multi--rank extent pool, each new volume is created on the rank with the most free extents
- Volumes may be dynamically deleted and their extents reused



Logical Volume Considerations

- **Volume size**
  - Volume size does not *necessarily* affect performance
  - For open systems, for a given amount of capacity, choose a volume size small enough to allow volumes to be spread appropriately across all ranks available to an application workload
  - For z/OS, larger volumes may require more aliases (PAVs)

- **Volume placement**
  - Logical volume placement on ranks, DAs (DS8000), loops (DS6000) and servers (server0 and server1) has a significant effect on performance
  - Logical volumes for each application workload should be allocated according to isolation, resource sharing and spreading principles

Principles of DS Performance Optimization

- **Allocation of logical volumes and host connections for an application workload**
  - Isolation
  - Resource sharing
  - Spreading
- **These principles are described in detail in Chapter 4 (4.1-4.3) of**
  - ***IBM TotalStorage DS8000 Series: Performance Monitoring and Tuning SG24-7146.***

Workload Isolation

- Dedicating a subset of hardware resources to one workload
  - **Ranks**
  - **I/O ports**
  - **…**
- Logical volumes and host connections for the workload are isolated to the dedicated resources
- Provides increased probability of consistent response time for an important workload, but…
  - Maximum potential performance limited to the set of dedicated resources
  - Contention still possible for any resources which are not dedicated (e.g. processor, cache, persistent memory)
- Can prevent less important workloads with high I/O demands from impacting more important workloads
- Rank level isolation may be appropriate for heavy random workloads
- DA level isolation (DS8000) may be appropriate for large blocksize, heavy sequential workloads

Workload Resource Sharing

- Multiple workloads use a common set of resources
  - Ranks
  - I/O ports
  - …
- Logical volumes and host connections for multiple workloads are assigned to the shared set of resources
- Provides higher potential performance by making a larger set of resources available to a workload, but…
  - Possible contention for all shared hardware resources (e.g. ranks, I/O ports, Device Adapters as well as processors, cache and persistent memory)
- Good approach when not enough workload information is available to identify isolation requirements
- Also good for workloads that:
  - Will not try to consume all of the hardware available
  - Peak at different times

Workload Spreading

- **Workload is balanced and distributed evenly across all allowed hardware resources**
  - Applies to both isolated and resource-sharing workloads

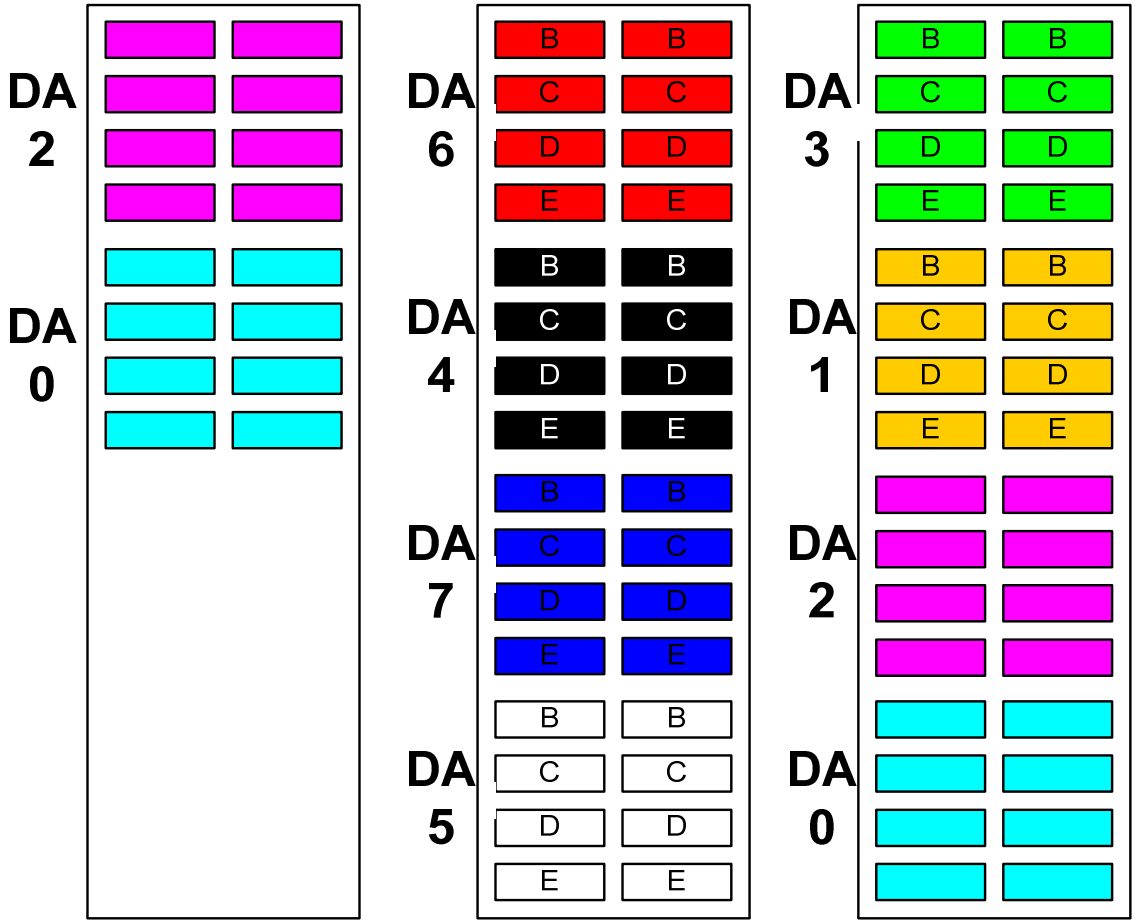
- **Logical volumes**
  - —Logical volumes for a workload are spread across:
    - Ranks
    - Device adapters (DS8000)
    - Loops (DS6000)
    - Server0 and server1
    - New logical volumes are allocated on least-used shared resources
  - —Additionally, host logical volume striping may be used
  - —Logical volume spreading exception
    - Files or datasets which will never be accessed simultaneously
    - Multiple log files for a single application may be placed on the same rank

Workload Spreading (2)

- **Host connections**
  - —Host connections for a workload are spread across:
    - I/O ports
    - Host adapters
    - I/O enclosures (DS8000)
    - Server0 and Server1
      - Both host adapters (DS6000)
      - Left side I/O enclosures and right side I/O enclosures (DS8000)
    - New host connections are allocated on least-used shared resources
  - —Additionally, multipathing software may be used

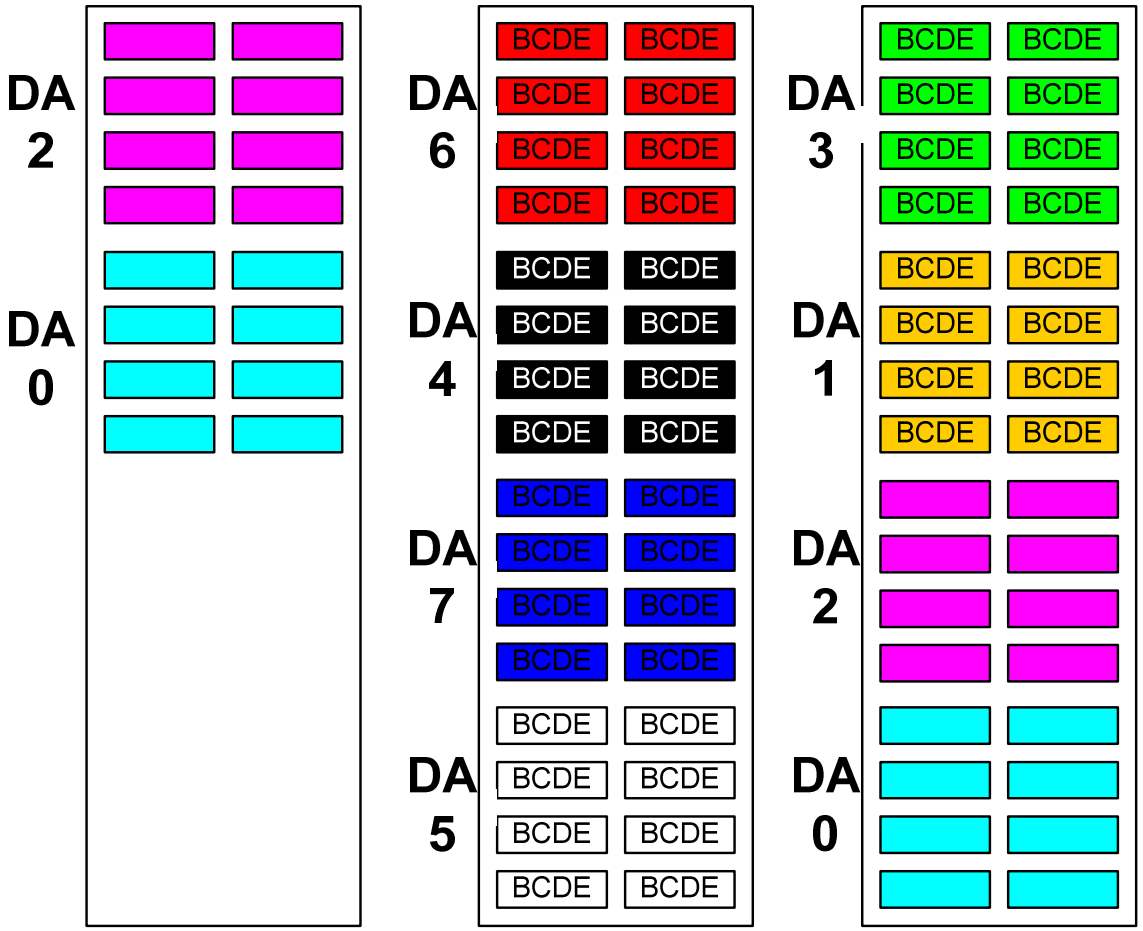
DS8000 Rank Level Isolation Example

- **4 workloads (B,C,D,E) share 6 DA pairs**
- **Each workload has 12 dedicated ranks**
  - –2 on each of the shared DA pairs
- **Logical volumes for each workload are allocated on its dedicated ranks on the shared DA pairs**
- **Device adapters are a shared resource**
- **Additionally, host level striping may me used across multiple volumes**



DS8000 Rank Level Resource Sharing Example

- 4 workloads (B,C,D,E) share 48 ranks on 6 DA pairs
- Ranks for the workloads are spread across DAs as well as server0 extent pools and server1 extent pools
- Additionally, host level striping should be considered



The secret to good performance is balance

- **There is only one basic principle:**
  - Spread the load uniformly across the resources
  - Spread across clusters
  - Spread across Device Adapters
  - Spread across **RANKS**.
- **Spread across ranks is the only real tool!**
  - Use ranks uniformly
  - Use ranks spread across the Device Adapters
  - Use ranks across clusters, even across **Boxes!**
- **You need tools to tell whether you have succeeded**
  - It won't always work the way you think it will!
  - TotalStorage Productivity Center, PDCU, Perfstats, ESS Expert
  - Skew across ranks and rank loading are the most common issue

DS8000 References

IBM TotalStorage DS8000 Series: Concepts and Architecture	SG24-6452
IBM TotalStorage DS8000 Series: Implementation	SG24-6786-01
IBM TotalStorage DS8000 Series: Copy Services with IBM eServer zSeries	SG24-6787-01
IBM TotalStorage DS8000 Series: Copy Services in Open Environments	SG24-6788-01
IBM TotalStorage DS8000 Series: Performance Monitoring and Tuning	SG24-7146
IBM TotalStorage DS8000 Introduction and Planning Guide	GC35-0495
IBM TotalStorage DS8000 Installation Guide	SY27-7641
IBM TotalStorage DS8000 User's Guide	SC26-7623
IBM TotalStorage DS8000 Host Systems Attachment Guide	SC26-7628
Multipath Subsystem Device Driver User's Guide	SC30-4096

DS6000 References

IBM TotalStorage DS6000 Series: Concepts and Architecture	SG24-6471
IBM TotalStorage DS6000 Series: Implementation	SG24-6781-01
IBM TotalStorage DS6000 Series: Copy Services with IBM eServer zSeries	SG24-6782-01
IBM TotalStorage DS6000 Series: Copy Services in Open Environments	SG24-6783-01
IBM TotalStorage DS6000 Series: Performance Monitoring and Tuning	SG24-7145
IBM TotalStorage DS6000 Introduction and Planning Guide	GC26-7679
IBM TotalStorage DS6000 Installation, Troubleshooting & Recovery Guide	SC26-7678
IBM TotalStorage DS6000 Host Systems Attachment Guide	SC26-7680
Multipath Subsystem Device Driver User's Guide	SC30-4096


Thank
You

NOTICES AND DISCLAIMERS

Product data has been reviewed for accuracy as of the date of initial publication. Product data is subject to change without notice. This document could include technical inaccuracies or typographical errors. IBM may make improvements and/or changes in the product(s) and/or program(s) described herein at any time without notice. Any statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business. Any reference to an IBM Program Product in this document is not intended to state or imply that only that program product may be used. Any functionally equivalent program, that does not infringe IBM's intellectually property rights, may be used instead.

THE INFORMATION PROVIDED IN THIS DOCUMENT IS DISTRIBUTED "AS IS" WITHOUT ANY WARRANTY, EITHER EXPRESS OR IMPLIED. IBM EXPRESSLY DISCLAIMS ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NONINFRINGEMENT. IBM shall have no responsibility to update this information. IBM products are warranted, if at all, according to the terms and conditions of the agreements (e.g., IBM Customer Agreement, Statement of Limited Warranty, International Program License Agreement, etc.) under which they are provided. Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products in connection with this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. IBM makes no representations or warranties, expressed or implied, regarding non-IBM products and services.

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents or copyrights. Inquiries regarding patent or copyright licenses should be made, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

Legal Information and Trademarks

The following are trademarks of International Business Machines Corporation in the United States and/or other countries.

AIX*	IBM logo*	S/390*	iSeries*
DB2 Universal Database*	FAStT*	Tivoli*	pSeries*
e-business logo*	IBM eServer*	TotalStorage*	xSeries*
ESCON*	Magstar*	WebSphere*	xSeries*
FICON*	OS/390*		
IBM*	OS/400*		

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

Intel is a trademark of the Intel Corporation in the United States and other countries.
Java and all Java-related trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc., in the United States and other countries.
Lotus, Notes, and Domino are trademarks or registered trademarks of Lotus Development Corporation.
Linux is a registered trademark of Linus Torvalds.
Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.
SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.
UNIX is a registered trademark of The Open Group in the United States and other countries.

* All other products may be trademarks or registered trademarks of their respective companies.

The following are trademarks or registered trademarks of Network Appliance, Inc.

- Data ONTAP, FlexVol, FilerView, SecureAdmin, RAID-DP, SecureAdmin, FlexClone, FlexCache, MultiStore, SnapLock, SnapShot, SnapMirror, SnapMover, SnapRestore, SnapVault, SyncMirror