DataDirect™
N E T W O R K S

Lustre User's Group

- **Performance**

  More Performance Per Pipe and Per OSS

  S2A directRaid does real time hardware Raid

- **Scalability**

  Linear scalability with drive expansion

  S2A enables Storage Network Scaling

- **QOS**

  Unique True Active/Active Dual Controllers

  Real Time Parity Check

  The S2A can raid the JBODs too

  S2A building blocks use up to 10 times less Disk Daisy Chaining

- **Virtualization**

  The S2A can export all Luns to All Ports and enable parallel access

  True Lustre OST fail over

  True Lustre Load Balancing via ALL ports

- **Lustre Network RAID1**

  Log Based Lustre Network Raid1 can take advantage if S2A virtualization

- **OST Addition and Deletion**

  S2A makes online OST addition and deletion agnostic from the storage side

- **Performance Enabling**

  S2A directRaid can saturate the Host Pipes (FC or IB) & the disks

- **Enabler of Failover and Load Balancing**

- **Best Internal and External bandwidth**

  Less controllers and Servers to manage with higher bandwidth per server

- **Best Cost ($/MB/s)**

- **Ease of management and Integration**

Cheetah 1 FC

- Dual ported at 100MB/s
- 1GB capacity
- Sustained reads at 5MB/s
- 6.5mS full stroke seek
- Block reassign in ~1.5s

Cheetah 7 FC

- Dual ported at 200MB/s
- 300GB capacity
- Sustained reads at 50+MB/s
- 6.5mS full stroke seek
- Block reassign in ~2.5s

The challenge is to achieve dramatic performance increases with no change in disk random performance

High Performance Silicon Based Storage Controller

- Parallel access for hosts
- Parallel access to a large number of disk drives
- True performance aggregation
- Reliability from a parallel pool
- Quality of Service
- Scalability
- Drive error recovery in real time
- True State Machine Control

DataDirect NETWORKS

DDN S^{2}A9500 Content Access:
Host Parallelism and PowerLUNs
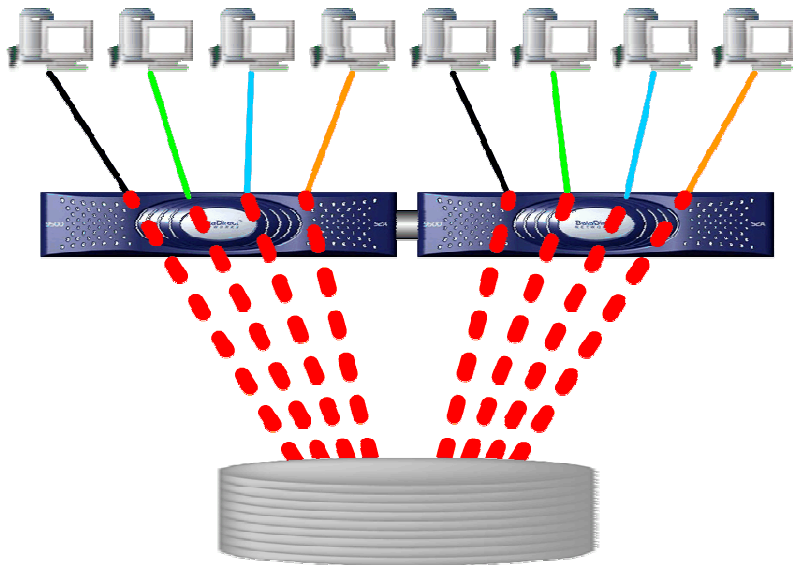
Generic RAID
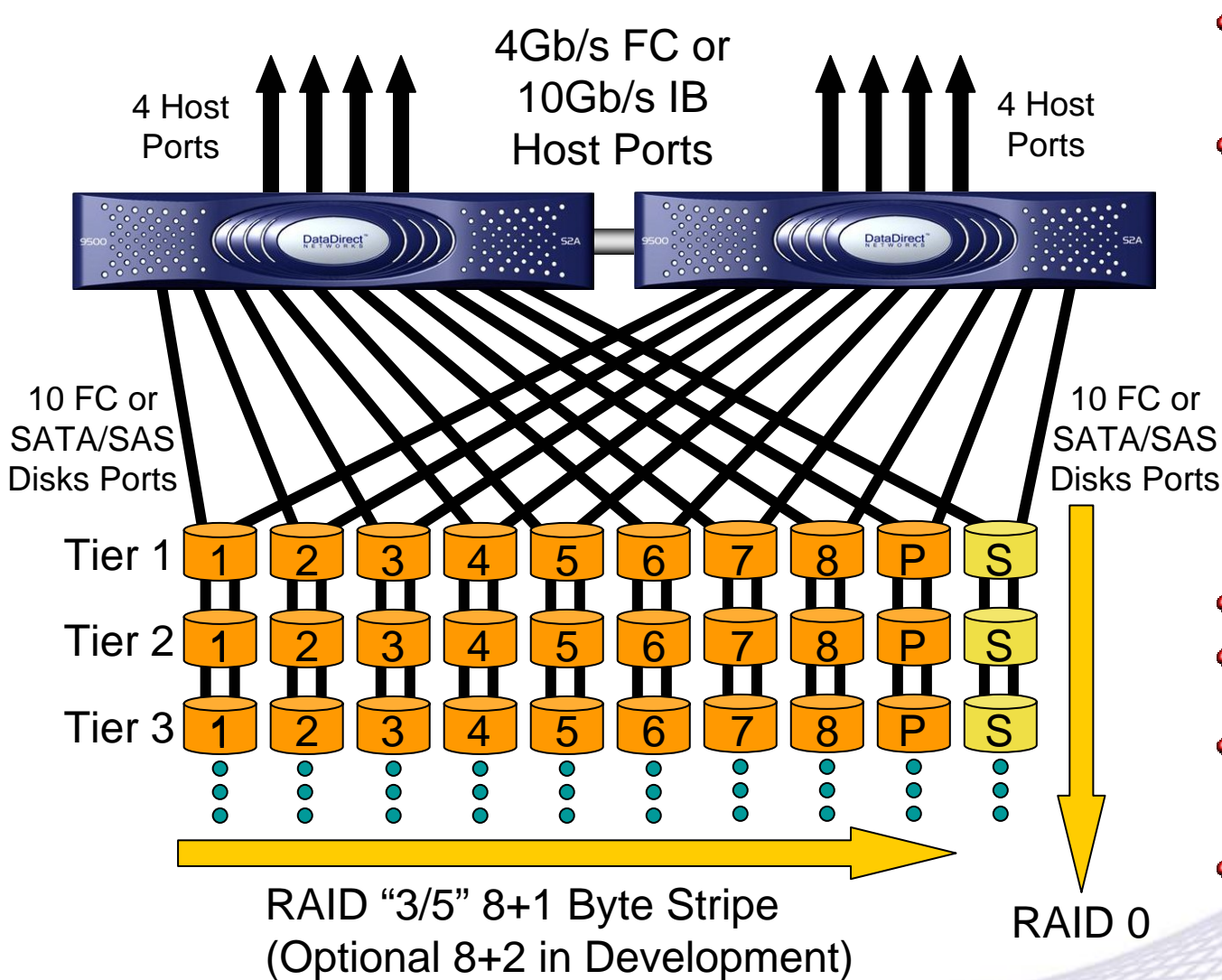## SAN Architecture

Fibre-Channel Switches

12-Way FS Stripe
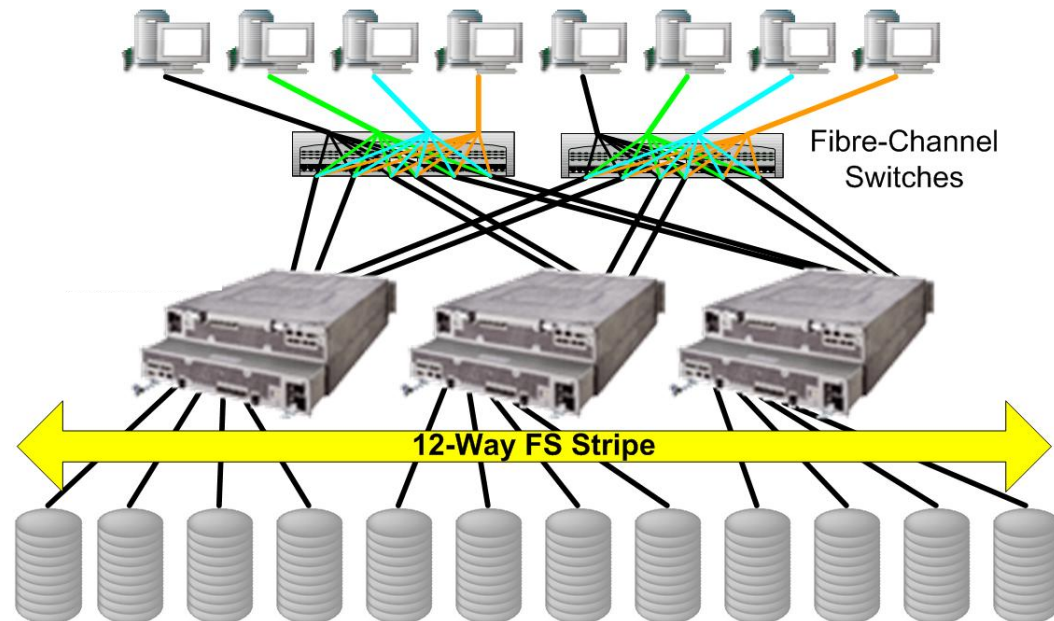
Like straws in a glass of water
- No Switching Latencies
- Greatly reduced Port contention
- No Striping Overhead
- Tested up to 53% improvements just due to host parallelism and PowerLUNs with only 8 hosts
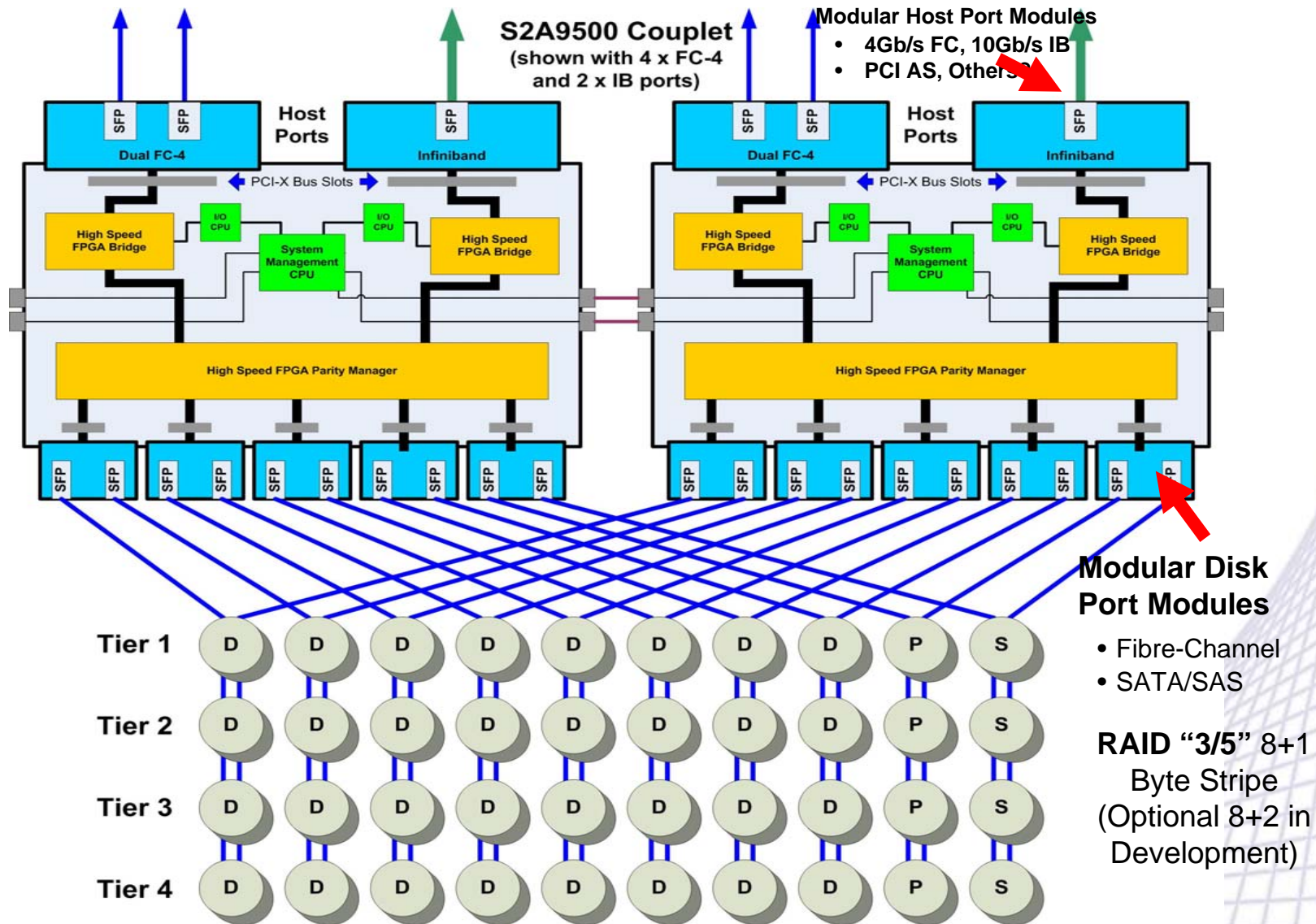
- Congested, Complicated Fabrics
- Lots of Switching Latencies
- Lots of Port Contention
- Host Striping robs CPU Performance
- Small I/O size per Storage Device
- Many more components (higher complexity)
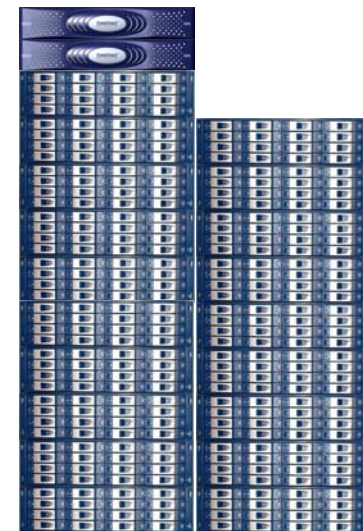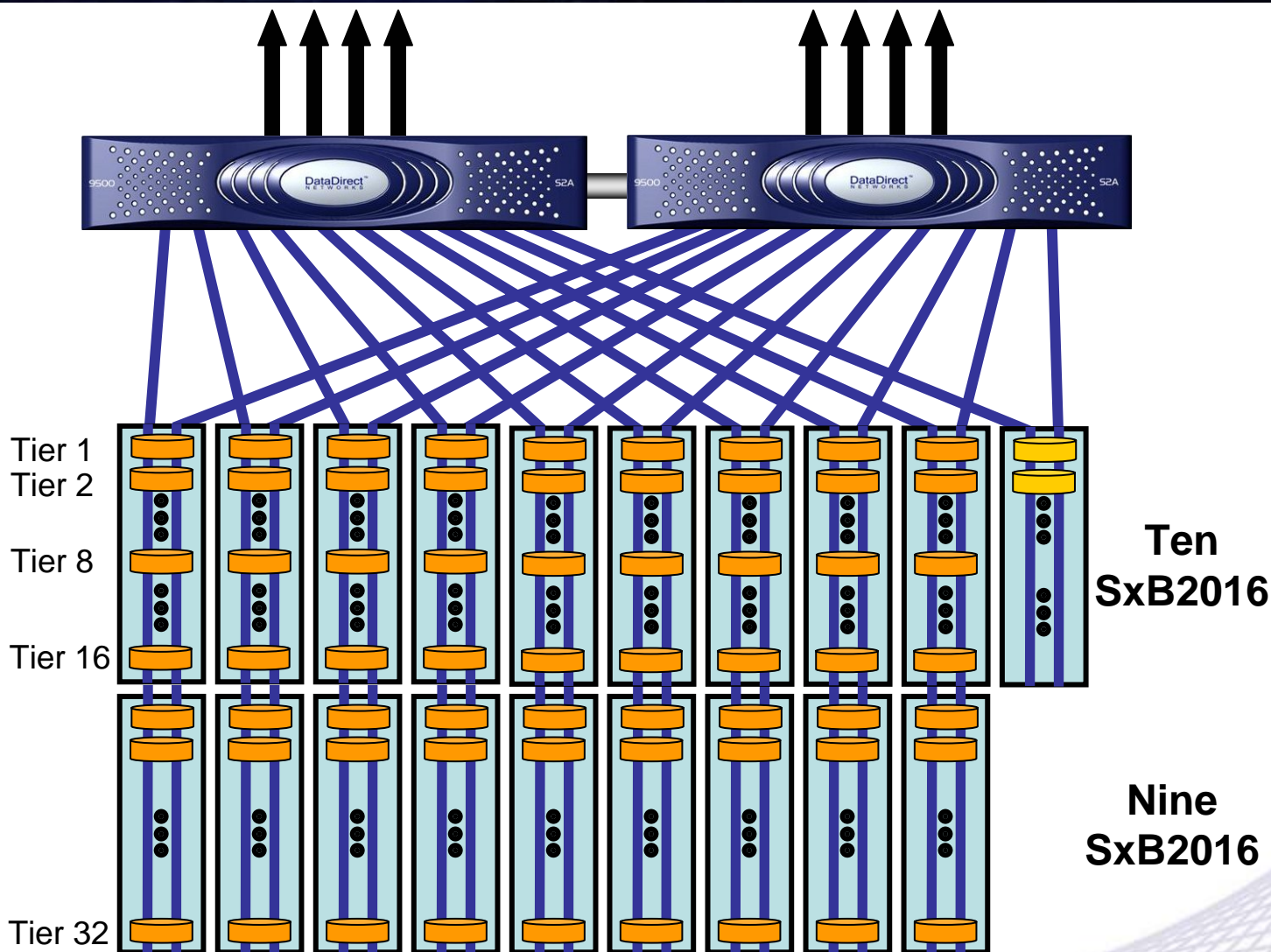
S2A9500 Basic Configuration

4Gb/s FC or 10Gb/s IB Host Ports

4 Host Ports

4 Host Ports

10 FC or SATA/SAS Disks Ports

10 FC or SATA/SAS Disks Ports

Tier 1 | 1 2 3 4 5 6 7 8 P S

Tier 2 | 1 2 3 4 5 6 7 8 P S

Tier 3 | 1 2 3 4 5 6 7 8 P S

RAID "3/5" 8+1 Byte Stripe
(Optional 8+2 in Development)

RAID 0

- **PowerLUNs can span arbitrary number of Tiers**

- **directRAID**
  - **Equivalent READ & WRITE performance**
  - **No performance degradation in crippled mode**
  - **Tremendous back-end performance for very low-impact rebuild, disk scrubbing, etc.**

- **RAIDed Cache**

- **Parity Computed on Writes AND Reads**

- **Multi-Tier Storage Support, Fibre Channel, SATA and SAS Disks**

- **Up to 1250 disks total**
  - **1000 formattable disks**

DataDirect NETWORKS

S2A9500 Couplet
(shown with 4 x FC-4
and 2 x IB ports)

Modular Host Port Modules
- **4Gb/s FC, 10Gb/s IB**
- **PCI AS, Others?**

Host Ports

Dual FC-4 | Infiniband | Dual FC-4 | Infiniband

SFP

PCI-X Bus Slots

High Speed FPGA Bridge | I/O CPU | System Management CPU | I/O CPU | High Speed FPGA Bridge

High Speed FPGA Parity Manager

Modular Disk Port Modules
- Fibre-Channel
- SATA/SAS

RAID "3/5" 8+1
Byte Stripe
(Optional 8+2 in
Development)

Tier 1 D D D D D D D D P S
Tier 2 D D D D D D D D P S
Tier 3 D D D D D D D D P S
Tier 4 D D D D D D D D P S

Tier 1
Tier 2

Tier 8

Tier 16

Ten SxB2016

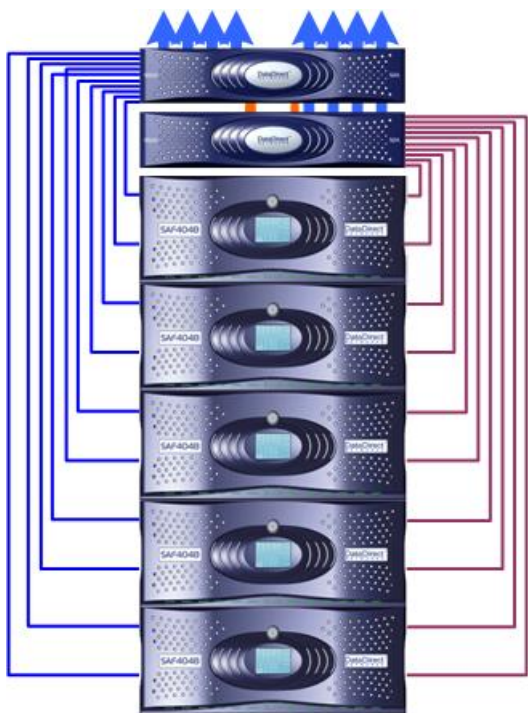**Nine SxB2016**

Tier 32

Full JBOD Redundancy

SAFX248 SATA Chassis

- 48 Slots, 4U
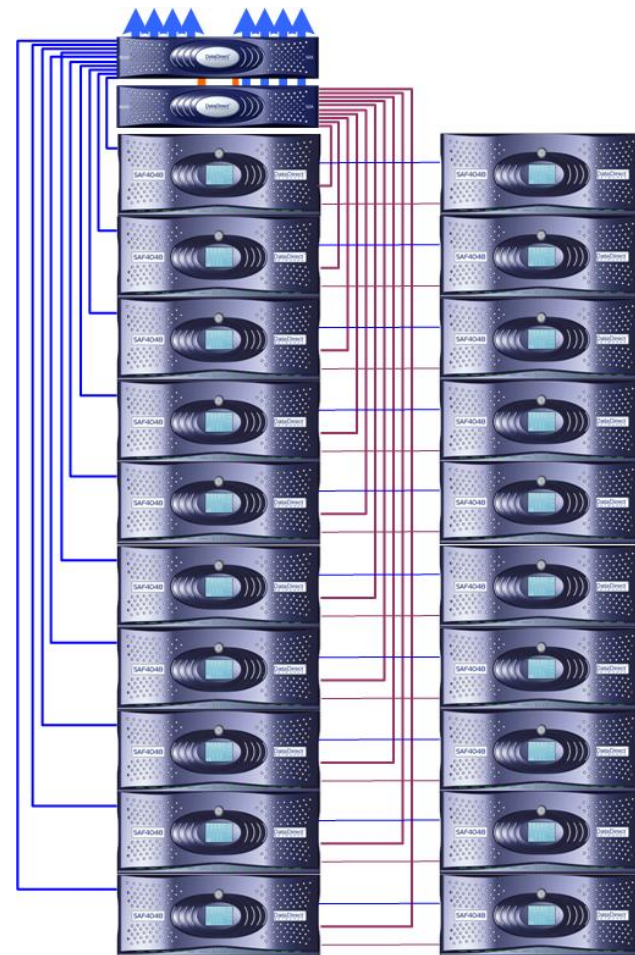- 480 Disks per Rack
- 240TB per Rack
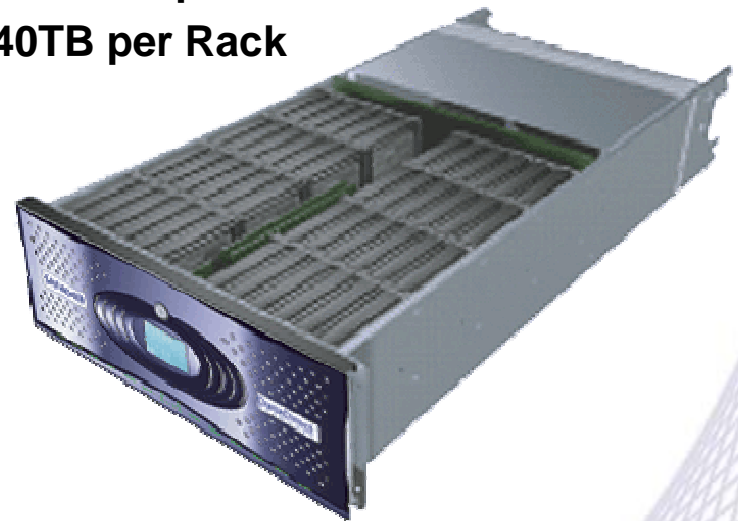
S2A9500 with

- **Five 48-Slot JBODs**
- **Two Dual Loop per JBOD 240 Disks**
- **120TB SATA using 500GB Drives**

or

- **Twenty 48-Slot JBODs**
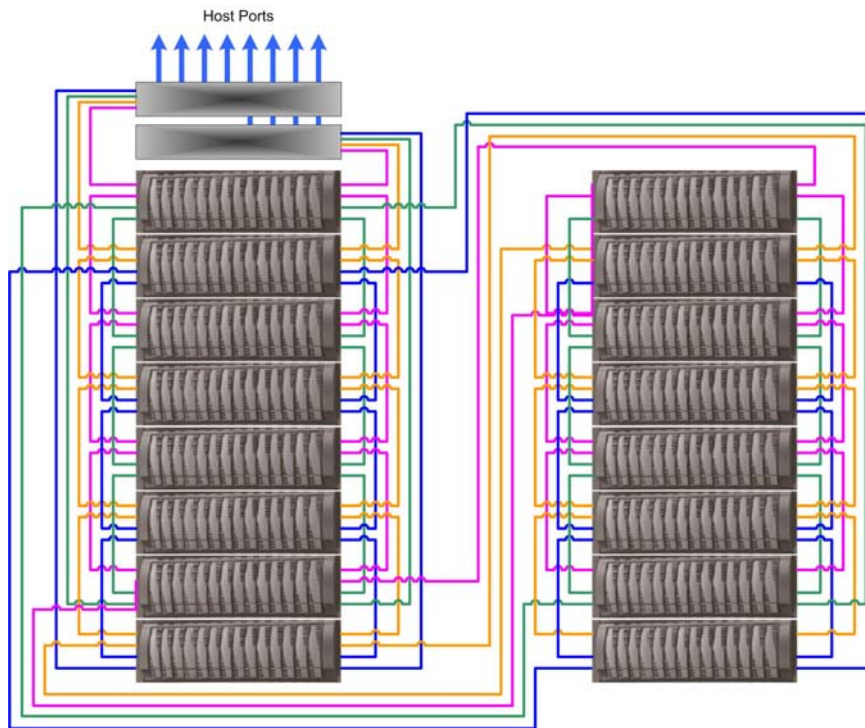- **Two Dual Loop per JBOD 960 Disks**
- **480TB SATA using 500GB Drives**

S2A

- 20 48-Slot JBODs
- Single Daisy-Chain JBOD
- 960 Disks
- 480TB SATA

4x Capacity

4x Simpler

2-Rack Footprint

Standard RAID

- 16 14-Slot JBODs
- 224 Disks
- 112TB SATA

Note:
- 500GB SATA Disks

Technology Roadmap

DataDirect
N E T W O R K S

- **S2A9500 FC-4**
  - 3 GB/sec aggregate bandwidth
 - Production Q4/05

- **S2A9500 IB 4x (SRP)**
  - 3 GB/sec aggregate bandwidth
  - Production Q4/05

- **You can mix and match FC-4 and IB 4x host side pipes on an S2A9500**

- **Stage One: SAS/SATA Enabled Enclosure**
  - Production in 12/05

- **Stage Two: SAS/SATA Back End S2A**
  - Production in Q2/06

- **Stage Three: Integrated Virtual Drives**
  - Production in Q2/07

- **Stage One: SAS/SATA Enabled Enclosure**
  - Production in 12/05

- **Stage Two: SAS/SATA Back End S2A**
  - Production in Q2/06

- **Stage Three: Integrated Virtual Drives**
  - Production in Q2/07
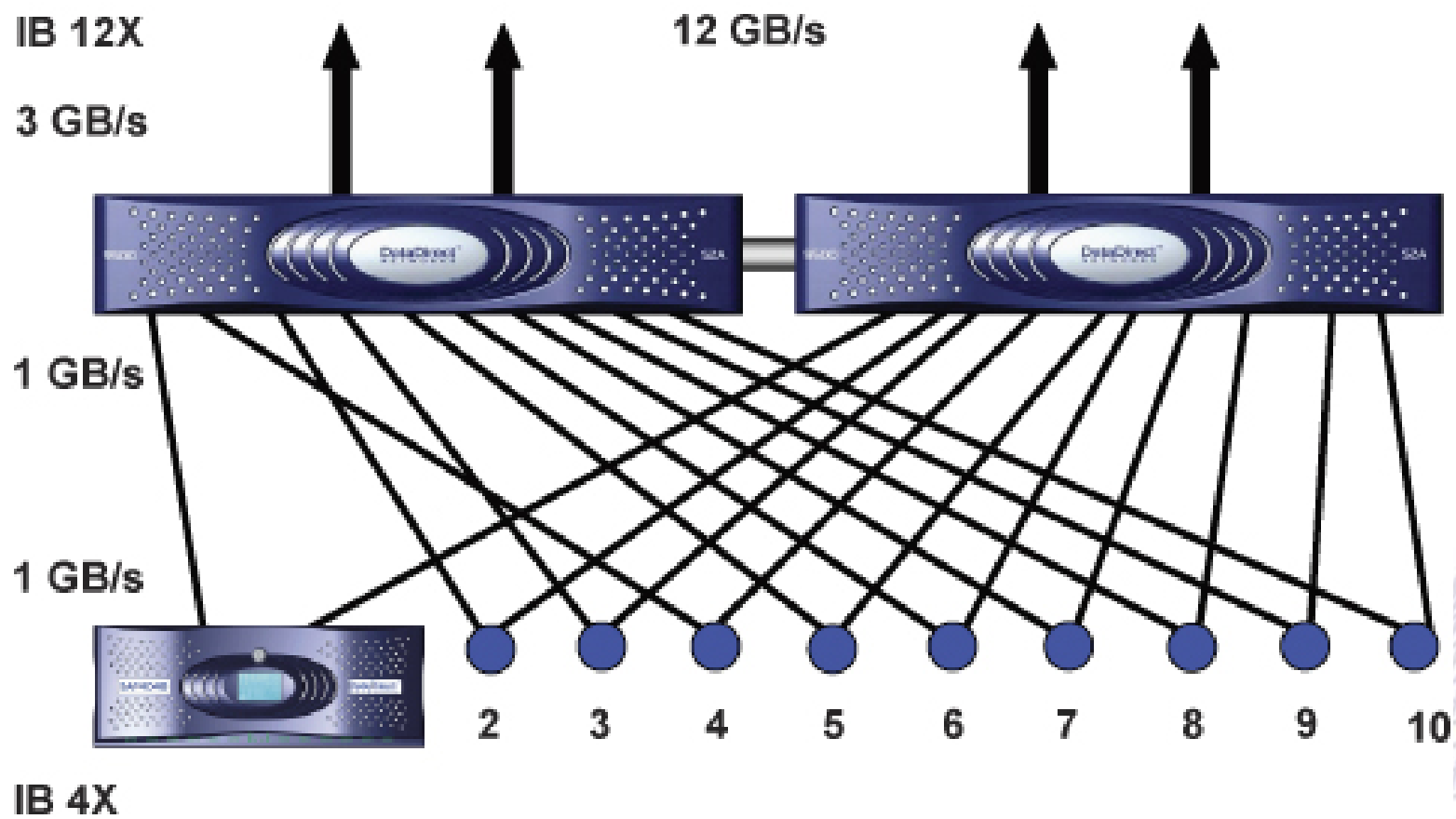
- 12GB/s sustained from each Couplet

- Virtual drives with distributed cache created from each 48 drive enclosure

- Virtual drives can be daisy chained for huge system capacity

- Internal bandwidth >20GB/s/Singlet

Lustre User's Group