

The background of the slide features a dark red, textured pattern of overlapping, curved, perforated bands. The DataDirect Networks logo is repeated in a lighter red color across the background. The main logo is in white.

**DataDirect**<sup>™</sup>  
N E T W O R K S

# Interactive Whiteboard Session: An introduction to Storage Design

Robert Mollard  
Systems Engineer, ANZ  
HPC & Life Sciences

# Product Overview

DataDirect  
NETWORKS

File Storage

## Parallel File Storage



**ExaScaler**  
20K Clients  
500GB/s+



**GridScaler**  
1000s of Clients  
500GB/s+, NAS

## SAN File Storage



**xStreamScaler**  
100s SAN/LAN Clients  
HSM Capable

## Enterprise NAS



**NAS Scaler**  
1-16 NAS Servers  
Fully-Featured

## Object Storage



**WOS**  
256 Billion Objects  
GeoReplicated  
**Cloud Foundation**

Storage Systems

## S2A9900



6GB/s in Real-Time  
1,200 Drives in 2 Racks

## S2A6620



2GB/s, Up to 350K IOPS  
120 Drives in 8U

## SFA10K-X, SFA10K-E



15GB/s, 840,000 SSD IOPS  
1,200 Drives in 2 Racks  
Embedded Computing

Disks

## SAS



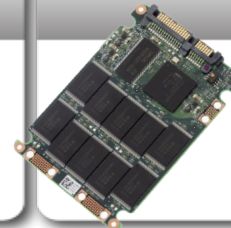
Up to 15K

## SATA



Up to 3TB  
Normal & Low Power

## SSD



200GB, 400GB SSD  
eMLC

# SFA10K-X | Scalability & Density

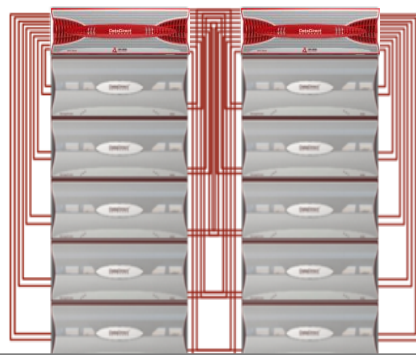
DataDirect  
WORKS

## The World Scalability & Density Leader



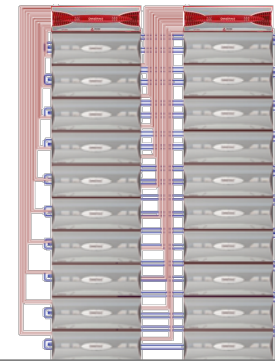
**5 Enclosures**  
**28U: 1/2 Rack**

**Up to 300 Drives**  
**Up to 900TB**



**10 Enclosures**  
**48U: 1 Rack (50U)**

**Up to 600 Drives**  
**Up to 1.8PB**



**20 Enclosures**  
**88U: 2 Racks**

**Up to 1,200 Drives**  
**Up to 3.6PB**

- Simple Cabling: All Enclosures are direct connected (up to 10 enclosures) to the SFA Appliances for easy configuration and maximum reliability.
- Maximum Availability: SFA Storage Systems can lose up to 20% of the available drive enclosures without impacting data availability.

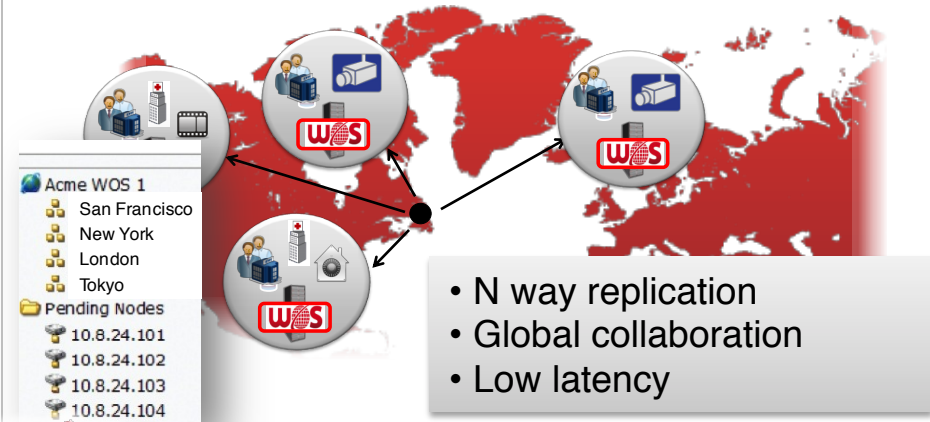
# WOS – Object Storage made simple!

## Hyper-Scale

- Single namespace; 256 billion objects\*
- Start small, grow to 23+ PB\*
- 350MB-700MB/node\*
- 3000-6000 Objects/S @ 50Kb
- 8M Reads/Sec; 2M Writes/Sec\*
- Almost 100% storage efficiency
- Just in time, online growth

\* = Soft Limits

## Global Reach & Data Locality



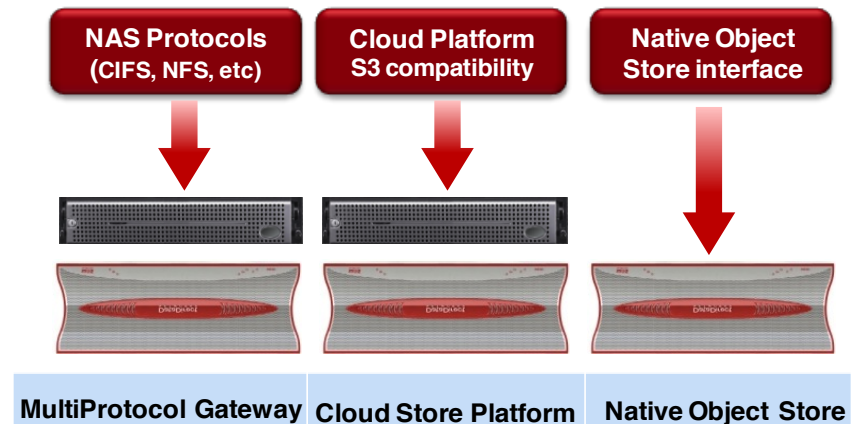
## Efficiency & Resiliency

- Replication or RAID Options
- Declustered Allocation
- Zero-Intervention, Autonomous Healing
- Instant provisioning
- Full drive usage
- Several times faster recovery than trad. RAID

WOS (Replicated or Object Assure)



## Universal Access



### **Medium to High Complexity Storage Environment**

- 5 days Site assessment, Initial planning and design
- 5 days Detailed design
- 5 days Tuning and optimisation
- 5 days Documentation and handover

### **Low Complexity Storage Environment**

- 2 days Site assessment, Initial planning and design
- 3 days Detailed design
- 3 days Tuning and optimisation
- 3 days Documentation and handover

# Where to begin...

1. Application
2. Disk type, array capabilities, block device?
3. File system type?
4. Somewhere in the middle, Operating System?

# Do you understand the data in your environment?

- Data types?
- Usage patterns?
- Understanding historic data usage and patterns will help you plan for the future.
- It will help you tune/optimize your systems over time as the demands change.

# Usage requirements?

- Do you know your Application?
- Is it a general file serving requirement?
  - » Know your data!
- Is it serving a compute cluster?
- Is it serving a Database?
- Is it a HSM requirement?



# Leads and Feeds

- Understand your end points.
  - » Is it a device, like tape
  - » Is it a user at the end of a fast link
- Learn about your workloads
  - » Small Block IO
  - » Large Block IO
- Dedicated paths, is it needed?

# HSM – Understanding Disk Cache Requirements?

- $(\text{Average dataset size} * \text{Number of datasets}) + \text{estimated daily ingest}$
- Can you separate out the inode (metadata) component?

# RAID Stripe Boundaries

## EFI (Extensible Firmware Interface)

EFI labels reserve the first 34 sectors, which misaligns RAID-5/6 "well formed I/O" regardless of the block size

\*You must force an offset to your data slice that lines up the disk sectors on RAID stripe boundaries.

## RAID Stripe Alignment

segment-size (*stripe size or chunk size*)

stripe-width = #Data-Disks \* segment-size

sector-offset = stripe-width / sector-size (typically 512 bytes)

(*ex. 5 disks RAID5 is 4D+1P;  $(4*64*1024)/512 = 640$  sectors*)

\*The Disk Allocation Unit, or LUN Block size should be a multiple of stripe-width.

## SSD erase block alignment

fs block size (*ex. 4096*)

SSD erase block size (*ex. 128k*)

stripe-width: fs-block-size / erase-block-size (*ex.  $128k / 4k = 32$* )