8-30-12

DataDirect™ NETWORKS
INFORMATION IN MOTION™

File and Object Storage Solutions for Big Data

DDN's Object-based Cloud Storage System

Mario Vosschmidt

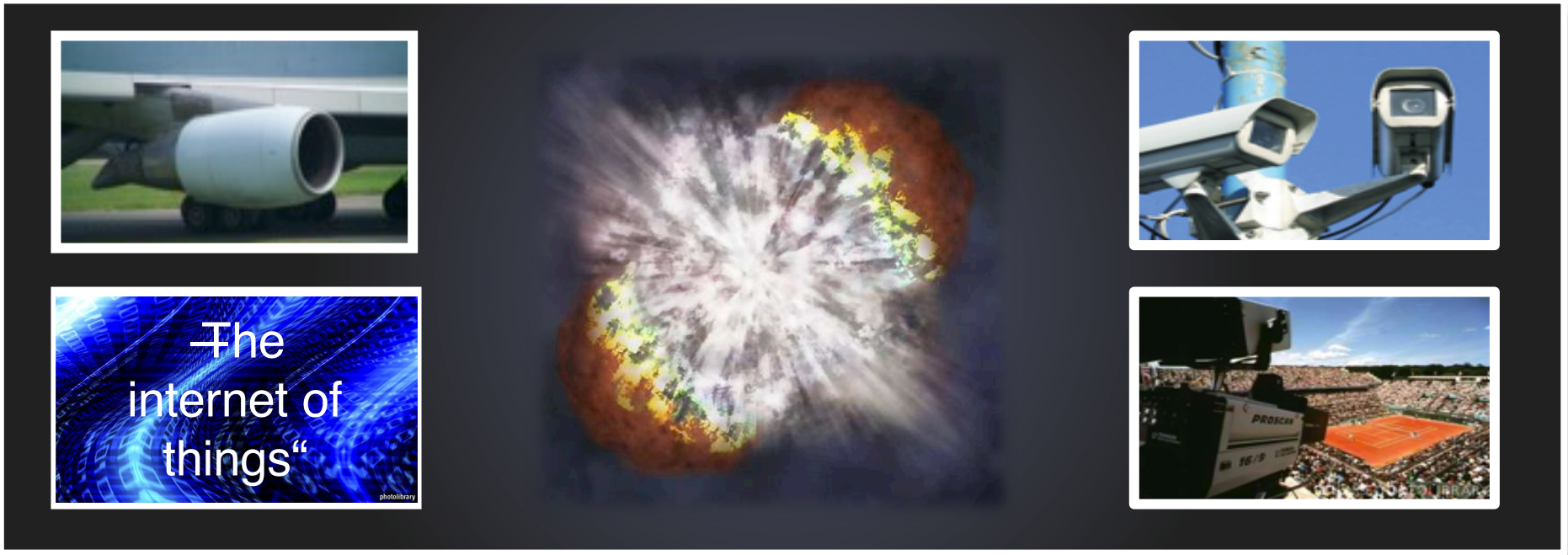
DDN Systems Engineering

Gridka Summer School 2012

Information Explosion 2.0: The New World

DataDirect NETWORKS
INFORMATION IN MOTION

BIG DATA



Hyper-scalable • Distributed • Collaborative

DDN: The Leader In Massively Scalable Storage Solutions & Services

DataDirect NETWORKS
INFORMATION IN MOTION

Key Statistics

- Delivers highly scalable and highly efficient storage solutions that enable customers to accelerate time to results, scale simply as data sets continue to grow, and gain competitive advantage through resolving performance and capability scaling challenges
- **Established:** 1998
- **Financials:** Over $200M Annually, Very Profitable
- **Headquarters:** Chatsworth, California USA
- **Employees**: Approximately 500 Worldwide
- **Customers:** Over 1,000 Worldwide
- **Footprint:** 17 Industries, 4 Continents, 49 Countries
- **Go to Market:** Global Partners, VARs, Resellers
- **Key Market Segments:**
  - **High Performance Computing & Life Science**
  - **Cloud & Web Content**
  - Rich Media
  - Intelligence/Federal
  - Surveillance

Industry Validation

#1 **DDN is the Largest Private Storage Company in the World**
Branded External Storage – By Revenue

#4 **DDN is the 4th Largest Storage Company Worldwide**
Branded External Storage – By Revenue

1000+ World-Leading Customers

SONY AIRBUS Microsoft
OAK RIDGE National Laboratory
yousendit
GENERAL DYNAMICS
IBM
TOTAL
photobox
BOEING
MEDIASET

DDN TOP 500 Presence

DataDirect NETWORKS
INFORMATION IN MOTION

- 70% 7 Of the Top10
- 65% 13 Of the Top20
- 64% 32 Of the Top50
- 61% 61 Of the Top100
- 29% 143 Of the Top500
- over 50% Of the Top500 Bandwidth
- over 75% Of the Top500 Lustre Sites
- over 60% Of the Top500 GPFS Sites

SFA12K™ | Data Center Configuration

DataDirect NETWORKS
INFORMATION IN MOTION™

World Leading Enterprise Storage Data Center Efficiency

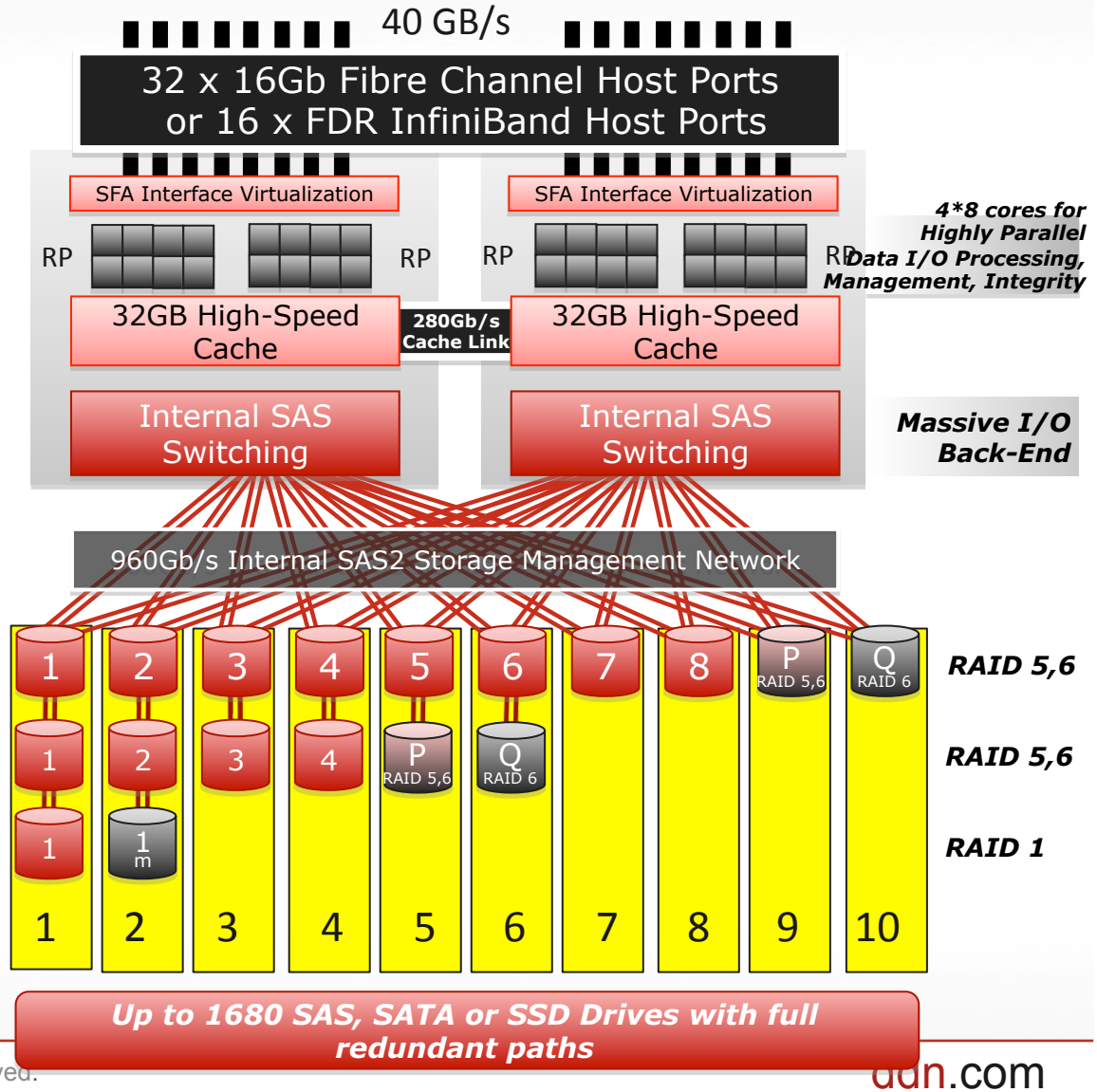
5 Enclosures	10 Enclosures	20 Enclosures
1/2 Rack	1 Rack (50U)	2 Racks
Up to 420 Drives	Up to 840 Drives	Up to 1,680 Drives
Up to 1.68PB	Up to 3.66PB	Up to 6.72PB
12 GB/s	20 GB/s	40GB/s

SFA12K-40

DataDirect NETWORKS
INFORMATION IN MOTION

Highly Parallelized SFA Storage Processing Engine

- Active/Active Design
- 1.7 Million Burst IOPS from 48GB Mirrored, Non-Volatile Cache
- Up to 1.4M Sustained Random Read IOPS from SSDs
- 40GB/s Raw Sequential Read & Write Speed
- Internal PCIe Gen 3 bus
- 280Gb/s ICL
- RAID Levels 1, 5 and 6
- Intelligent Write-Through Striping
- SATAssure Data Protection
- GUI, SNMP, CLI
- 32 x FC-16 ports or 16 x FDR-IB ports

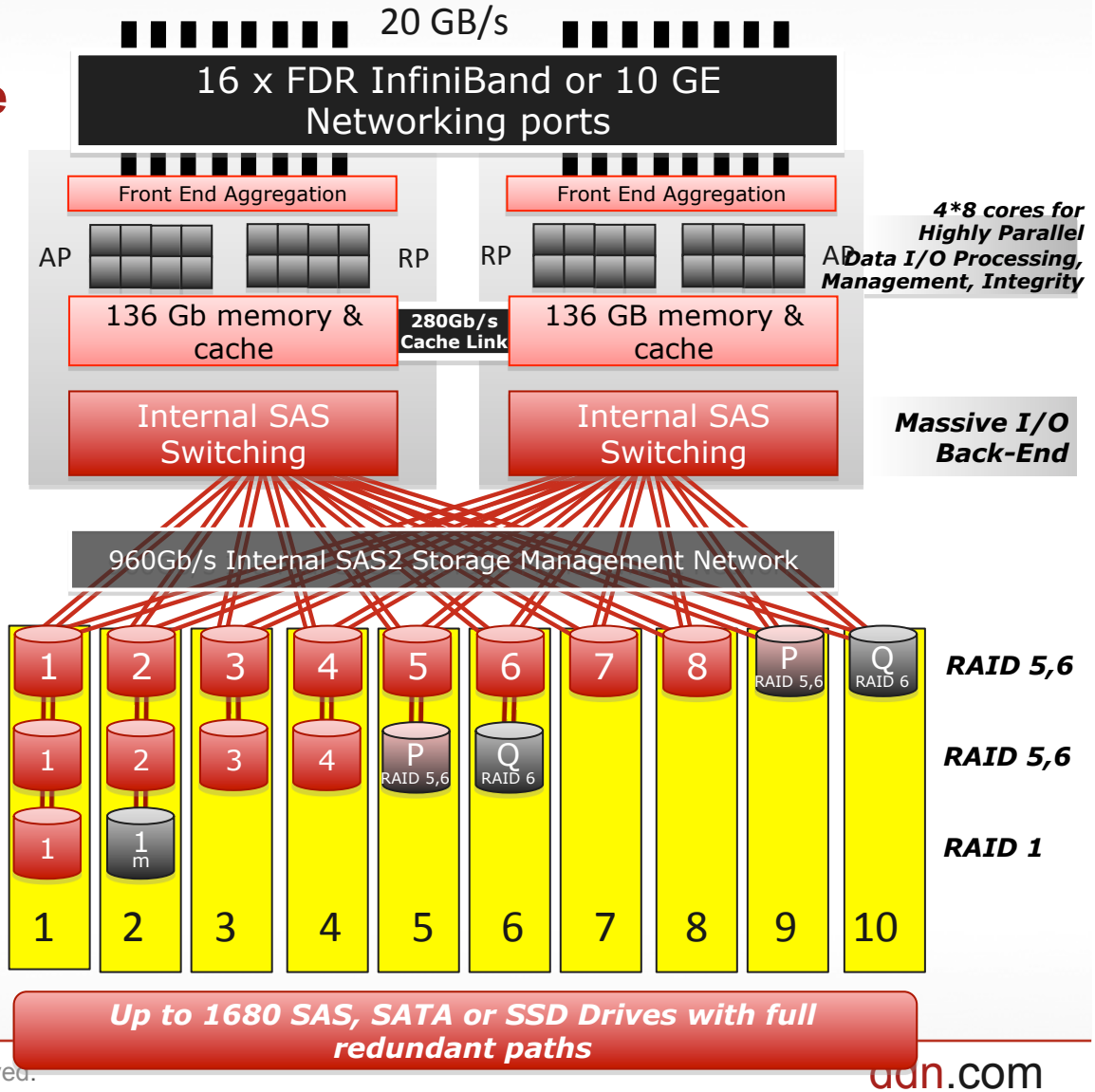


SFA12KE-20

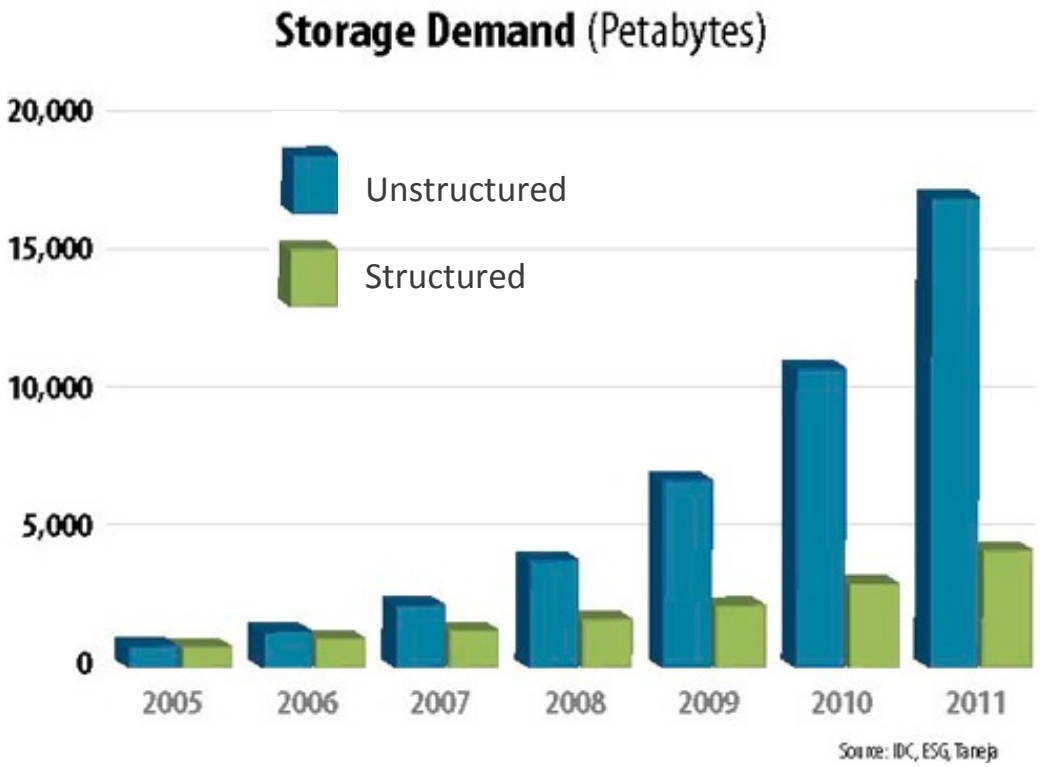
DataDirect NETWORKS
INFORMATION IN MOTION

Highly Parallelized SFA-E Embedded Filesystem Engine

- Active/Active Design
- Appiication and RAID within one System (AP & RP)
- Low Latency in-memory communication
- Up to 20GB/s filesystem throughput
- GPFS or Lustre embedded engine
- RAID Levels 1, 5 and 6
- Intelligent Write-Through Striping
- SATAssure Data Protection
- GUI, SNMP, CLI
- Directmon consolidated monitoring
- 16 x FDR-IB or 10 GE ports



Unstructured Data Explosion



Unstructured Data Examples

- Photos
- Music
- Images
- Video
- Documents
- Email

Unstructured Data Characteristics

- **Big**
- **Immutable: less subject to change**
- **Accessed by many**
- **From anywhere**

"Unstructured, file-based data will continue to grow at a blistering pace and IT will continue to struggle to manage it." -Enterprise Strategy Group

Conventional Storage

DataDirect NETWORKS
INFORMATION IN MOTION

Well understood, but stretched to its limits for unstructured data cloud storage

- Introduced in the 90's
  - With 16TB file system sizes that many still have
- Management intensive
  - LUNs, Volumes, Aggregates,…
  - Heroically management intensive at scale
- Antiquated resiliency techniques that don't scale
  - RAID
  - Cluster failover, "standby" replication, backup
  - File Allocation Tables, Extent Lists
- Inefficient usable space utilization
- Focused on structured data transactions (IOPS)
  - File locking overhead adds cost and complexity
- Prohibitively expensive at cloud scale
  - To purchase, run, manage, protect and expand

Hyperscale Initiative at DDN

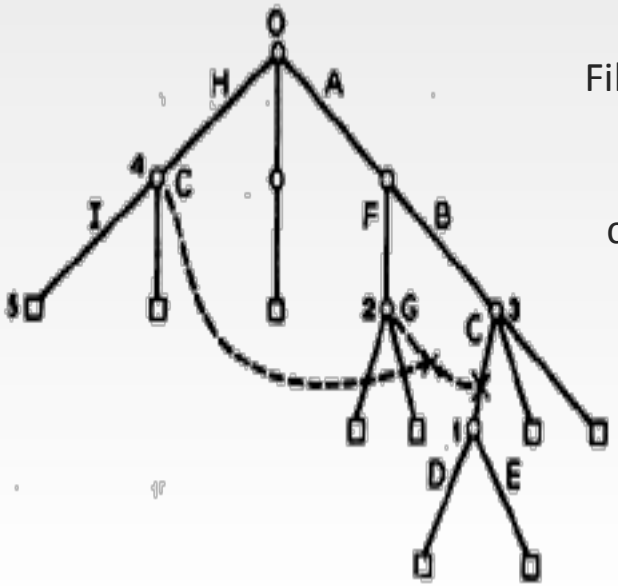
DataDirect NETWORKS
INFORMATION IN MOTION

- Worldwide scientific collaboration
  - Understand the data usage model where immutable data is shared and studied
- Eliminate the complexity of data storage and distribution
  - Simplify the data access system
  - Remove the concept of filesystem allocation and to maximize efficiency
  - Reduce the instruction set to only PUT, GET, & DELETE
- Data must appear local - everywhere
  - Add the concept of locality based on latency to data and load balance

Object Storage Explained

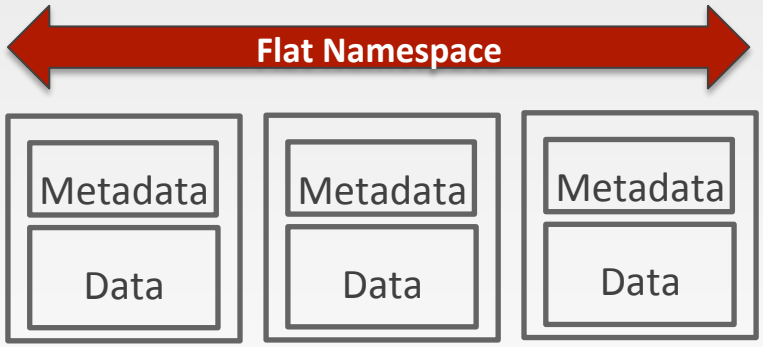
- Object storage stores data into containers, called objects
- Each object has both data and user defined and system defined metadata (a set of attributes describing the object)

File Systems



File Systems were designed to run individual computers, then limited shared concurrent access, not to store billions of files globally

Objects



Objects are stored in an infinitely large flat address space that can contain billions of files without file system complexity

1965, Fall joint computing conference, Multics filesystem

Hyperscale Storage | Web Object Scaler

DataDirect NETWORKS
INFORMATION IN MOTION

WOS

DataDirect NETWORKS

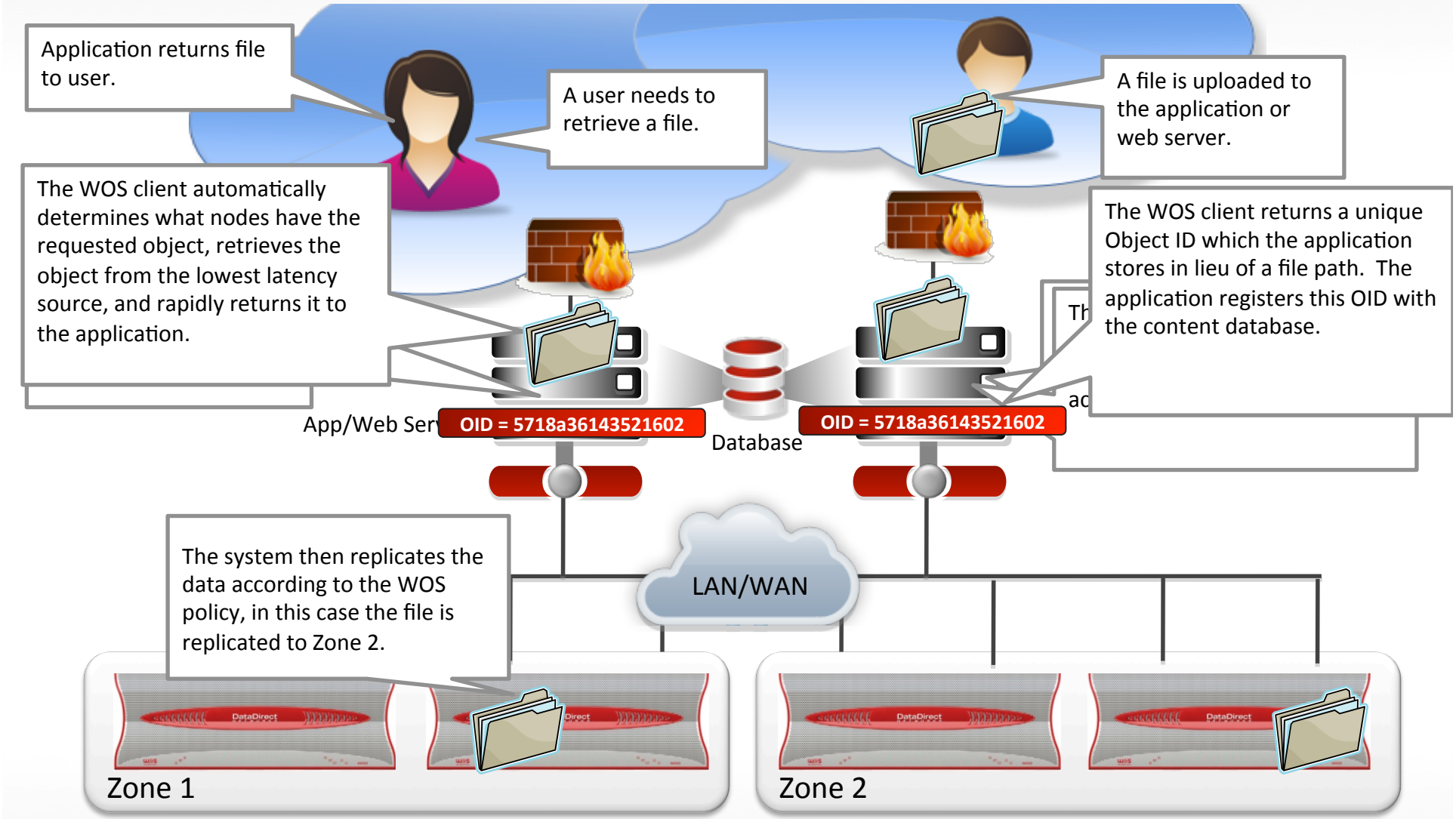
NoFS

Hyperscale
Distributed
Collaboration

This slide needs to be viewed in PowerPoint presentation mode.

Static display such as editing mode or printed slides will not convey anything meaningful due to the interactive nature of this slide.

WOS Concepts



Best viewed in presentation mode

WOS (Web Object Scaler)

DataDirect NETWORKS
INFORMATION IN MOTION

GeoDistributed, Scale-out Object Storage System

- Hyper-Scalable Cloud Storage Foundation

TRUE End:End Object Storage

- Maximum Performance From Every Media

Easy to Manage at Hyperscale

- Single namespace, Single Global Cluster Interface

Autonomous, Self-Healing Big Data Infrastructure

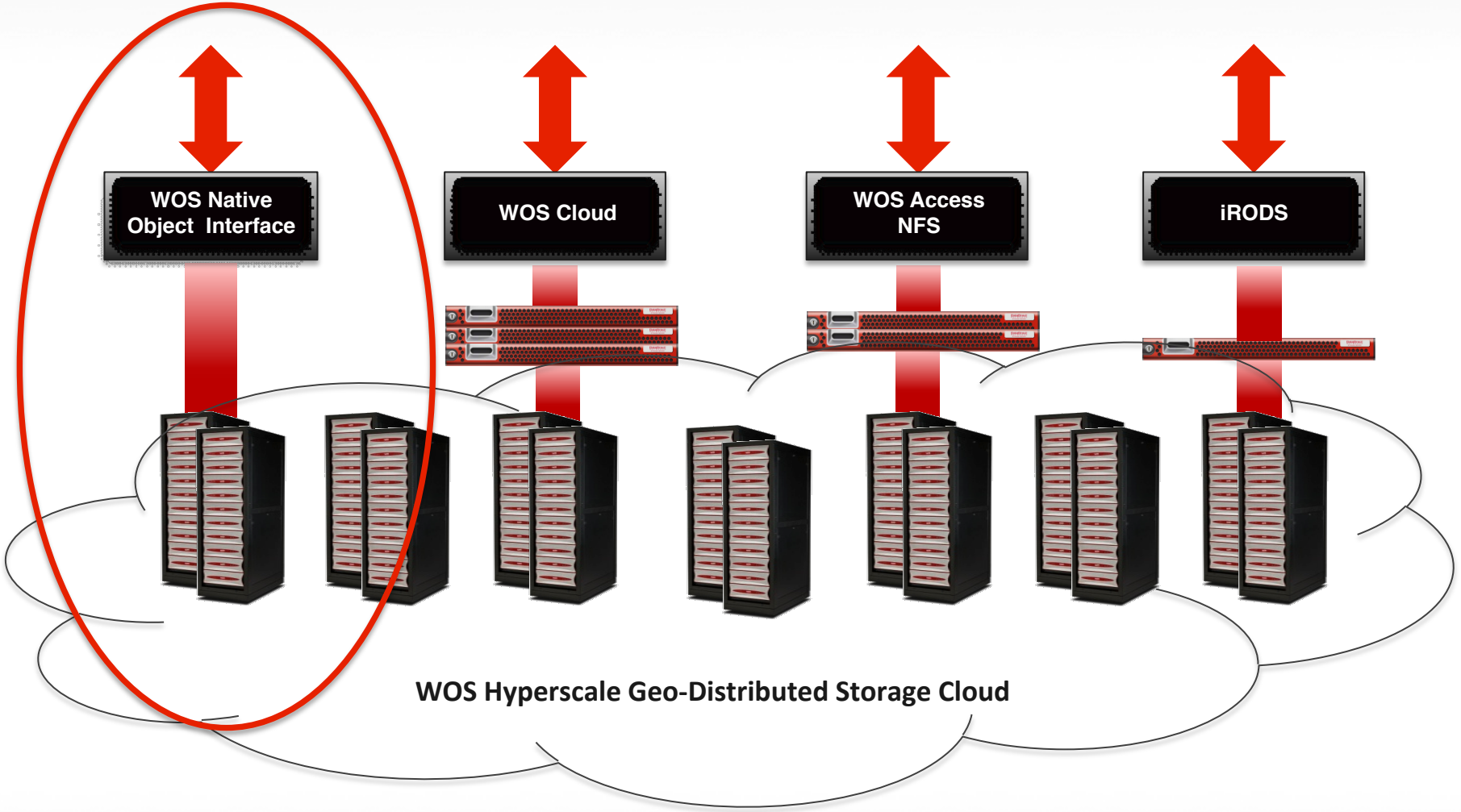
- Intelligent, Fail-In-Place Architecture

Flexible Cloud Storage Service Platform

- Multimodal Access Featuring Billing & Multi-Tenancy

WOS Interfaces Drive HPC and Life Sciences Solutions

DataDirect NETWORKS
INFORMATION IN MOTION



WOS Asynchronous Replication

Data replication to remote WOS nodes occurs asynchronously to increase performance

- Two copies of data are always written to "local" (lowest latency) node before "PUT" call returns OID to minimize/eliminate risk of data loss
- Replication occurs in background, once replication successfully completes, extra local copy is removed

Previously, data needed to be replicated to remote sites to insure data protection before acknowledging a successful "write" to the application

- Taking extra time to accomplish

Especially useful for big files and big data sets

ObjectAssure Single Copy Data Protection

ObjectAssure **erasure-code** based **declustered** data protection

- An erasure code provides redundancy by breaking objects into smaller fragments and storing the fragments across different disks
- Data can be recovered from any smaller combination of fragments

ObjectAssure is the first erasure code protection mechanism for hyper-scale, high-performance cloud storage

With ObjectAssure, each WOS node can withstand up to two concurrent drive failures without loss of data or data availability

- Data protection without the cost of replication

Enables single copy environments

- Expands usable storage space and further reduces overall storage costs

Why is WOS so *Fast*?

DataDirect NETWORKS
INFORMATION IN MOTION

Traditional storage does a lot of work to write/read data

- Expends excess disk operations
  - 5-12 Disk - Operations per File Read
- Multiple levels of translation and communication
  - Metadata lookups and directory travelling
  - Extent list fetches
  - RAID & block operations

WOS delivers performance through simplicity

- None of the constructs of traditional systems
- Single-Disk-Operation Reads, Dual-Operation Writes
- Reduced latency from SATA Disks since seeks are minimized
- Millions of file/ops per second with ¼ of the disks

WOS Building Blocks

DataDirect NETWORKS
INFORMATION IN MOTION

High-Density Scale-Out Cloud Storage Appliances

Linear performance scaling with low latency and high efficiency

WOS 6000

4U, 60-Drive System
8 x GbE per Chassis (2 nodes)
2PB/Rack, 23PB/Cluster
25B Objects/Rack

WOS 1600

3U, 16-Drive System
4 x GbE per Node
544TB per Rack

WOS – Architected for Big Data

DataDirect NETWORKS
INFORMATION IN MOTION

Hyper-Scale

3U, 16-drive WOS Node (SAS/SATA)

4U, 60-drive WOS Node (SAS/SATA)

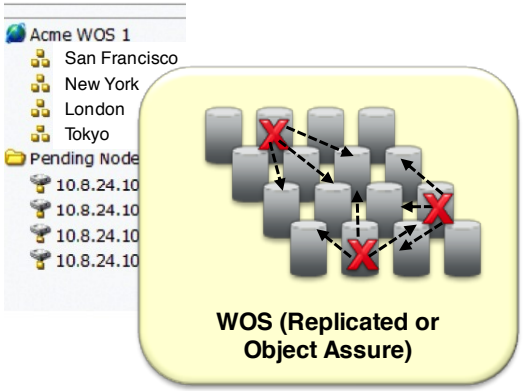
2PB / 11 Units per Rack

- 256 billion objects per cluster
- Scales to 23PB
- Start small, grow to tens of Petabytes
- Network & storage efficient

Global Reach & Data Locality

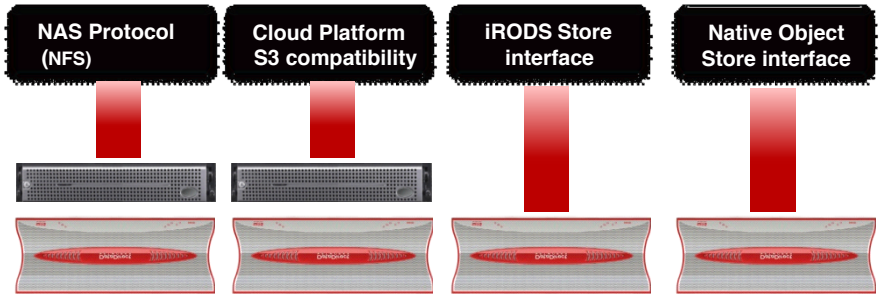
- Up to 4 way replication
- Global collaboration
- Access closest data
- No risk of data loss

Resiliency with Near Zero Administration



- Self healing
- All drives fully utilized
- 50% faster recovery than traditional RAID
- Reduce or eliminate service calls

Universal Access



WOS NAS Access
- NFS protocol
- Scalable
- HA & DR Protected
- Federation across existing NAS

Cloud Store Platform
- S3-Compatible & WebDAV APIs
- Multi-tenancy
- Reporting & Billing
- Personal storage

iRODS
- Custom rules – a rule for every need
- 28 access methods

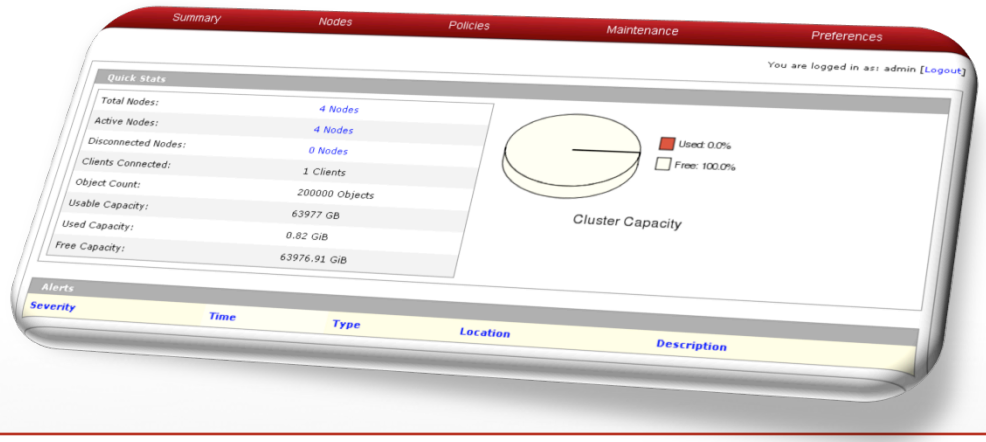
Native Object Store
- C++, Python, Java, PHP, HTTP REST interfaces
- Simple interface

Near Zero Administration

DataDirect NETWORKS
INFORMATION IN MOTION

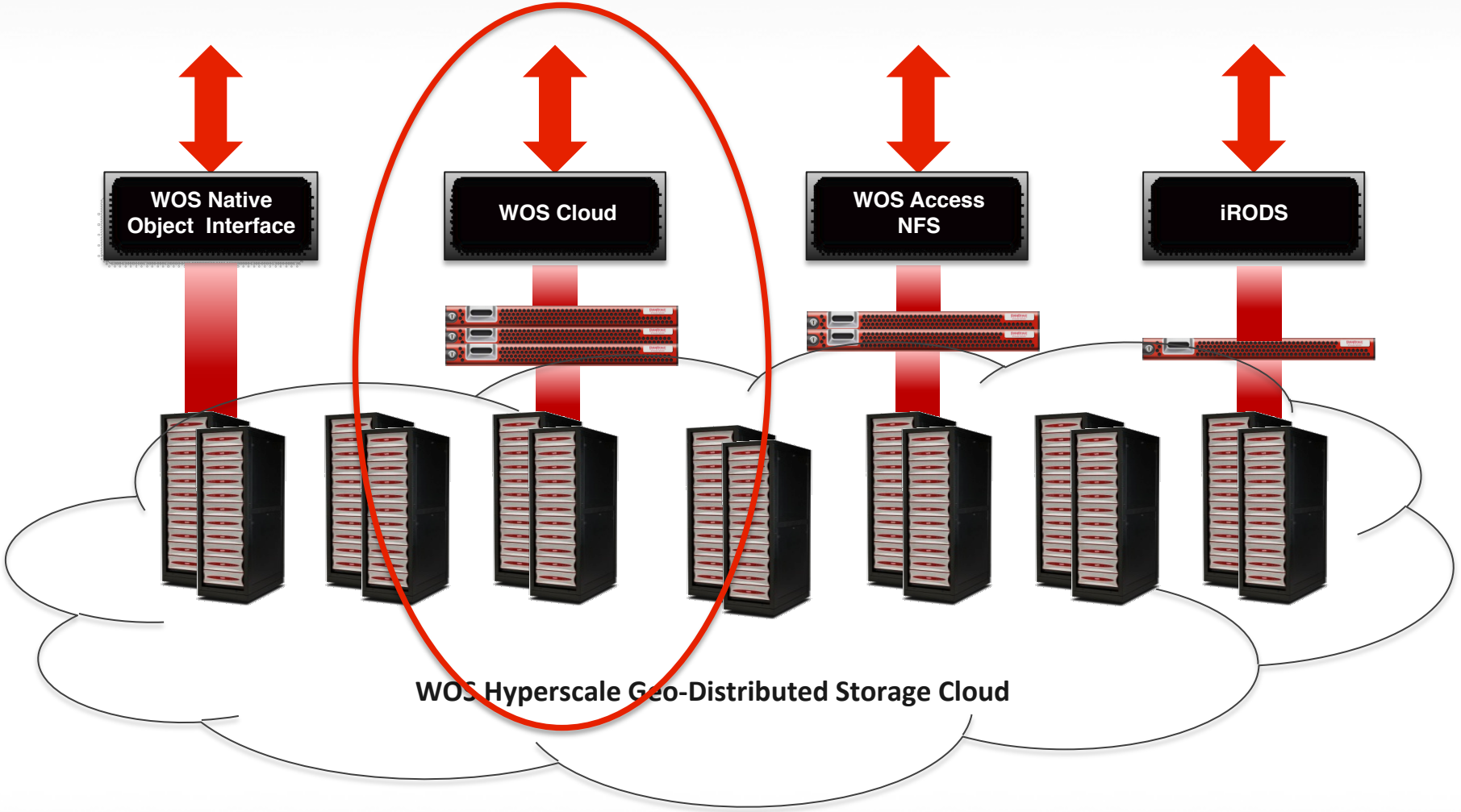
Multi-petabyte WOS deployments require less that 1 FTE to manage

- Deploy additional WOS nodes in 10 minutes or less
  - Install Node, Assign IP Address, Point and Click!
- Single, global namespace easy to configure and extend
- Policy-driven replication is automatic and transparent
- Data and disaster protection is built-in
- WOS System self-heals from disk failure
- Zero-Intervention via (planned) fail in place graceful degradation



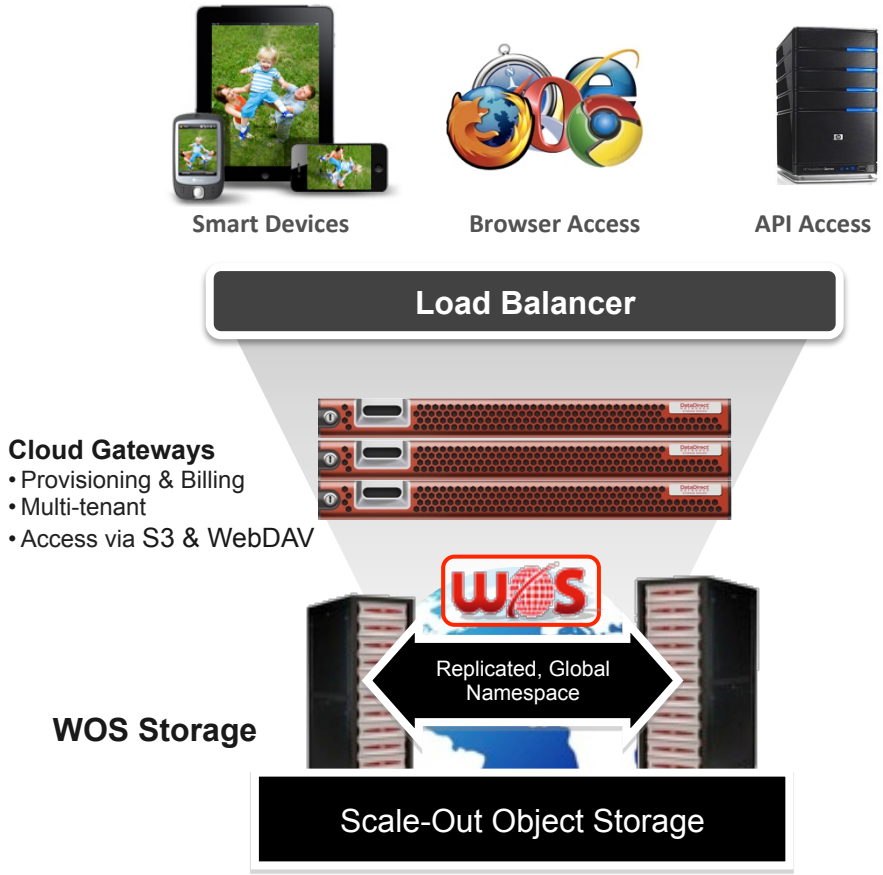
WOS Interfaces Drive HPC and Life Sciences Solutions

DataDirect NETWORKS
INFORMATION IN MOTION



Complete Enterprise Cloud Storage
SOLUTION

- Drop-in file storage and sharing
- S3 compatible & WebDAV APIs
- Multi-tenancy support
- Full CDMI compliance
- Integrates with provisioning & billing systems
- Security
  - 128-bit SSL encrypted communications
  - 256-bit AES encryption for stored files
  - Integration with LDAP, Active Directory



Amazon S3 Compatible Cloud Storage

DataDirect NETWORKS
INFORMATION IN MOTION

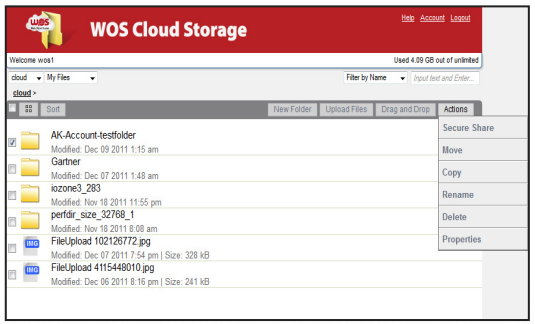
- S3 has become the defacto standard cloud storage API
- With all the benefits of WOS
  - Data protection, geo-distributed low latency access, easy management
- Plus
  - Multitenant provisioning, admin, usage reporting
  - Interfaces to billing systems
  - Packaged and supported for quick deployment

amazon web services™
S3 Simple Storage Service

WOS

WOS Cloud Drop-in File Access

- Mobile Access
  - Native Clients for iPhone, iPad
  - Easy-to-use navigation with native menu functions
  - Upload and download files
  - View files using native viewer applications
- Security
  - SSL encrypted communications
  - 256-bit AES encryption for stored files
  - Integration with LDAP, Active Directory
- Browser Access
  - Internet Explorer, FireFox, Chrome and Safari
  - Centralized view of all stored content



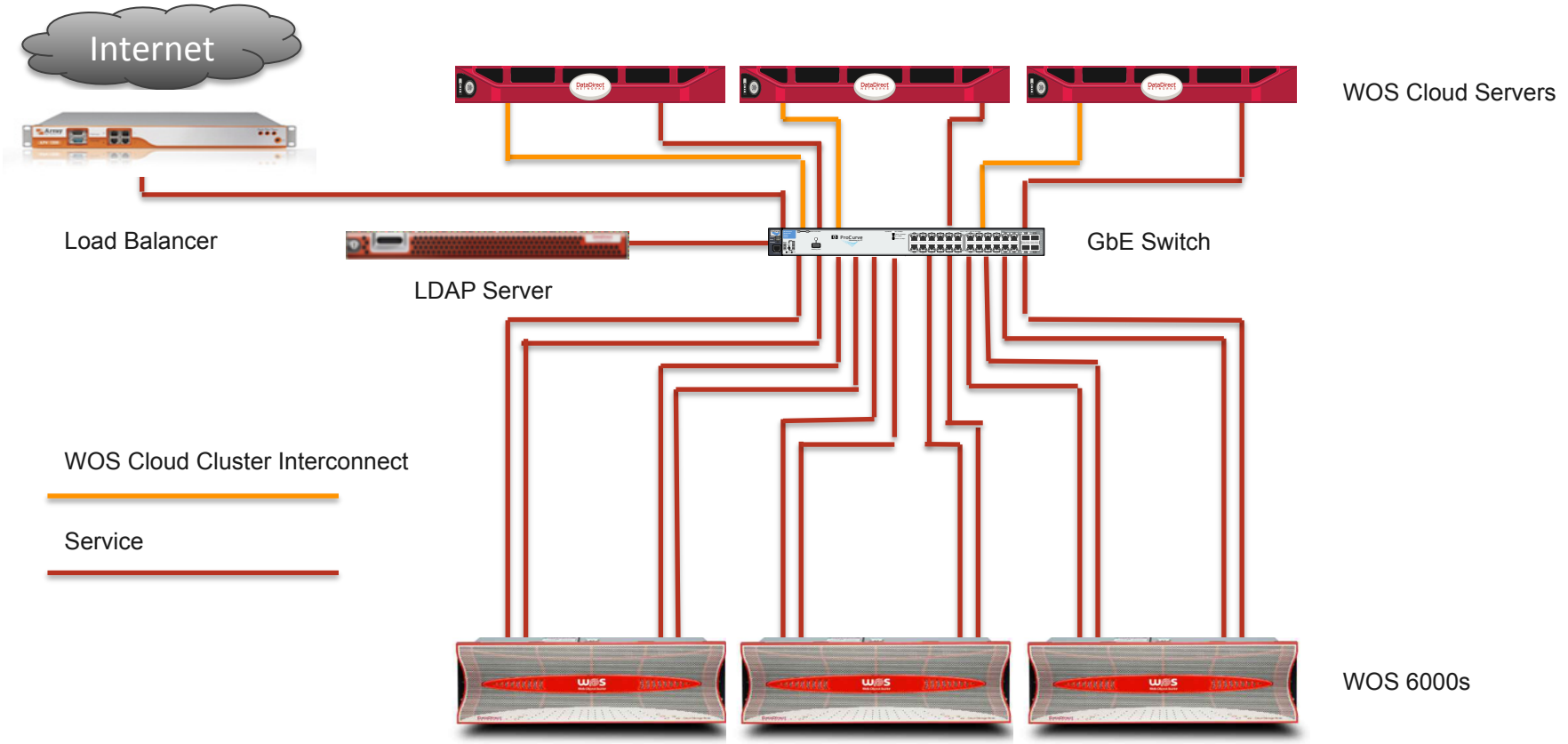
Example HPC & Life Sciences Solution TokyoTech

DataDirect NETWORKS
INFORMATION IN MOTION

Solution

- The DDN solution was the 360TB WOS Cloud system, consisting of 3 x WOS Cloud servers and 3 x WOS6000 (120TB each)
- Files are uploaded via their PC through the WOS Cloud Client
- Data from a variety of equipment throughout the Academy, including images and other biologic sensor data, voice data, movies and text data are uploaded by S3 applications developed by TokyoTech
- TokyoTech selected WOS Cloud because it offered behind the firewall security, high density, modular scalability and S3 compatibility
- The ACLS has about 600 users that access WOS Cloud
- The ACLS especially liked WOS's object storage replication mechanism as it met their plans to provide future storage and access at an additional location

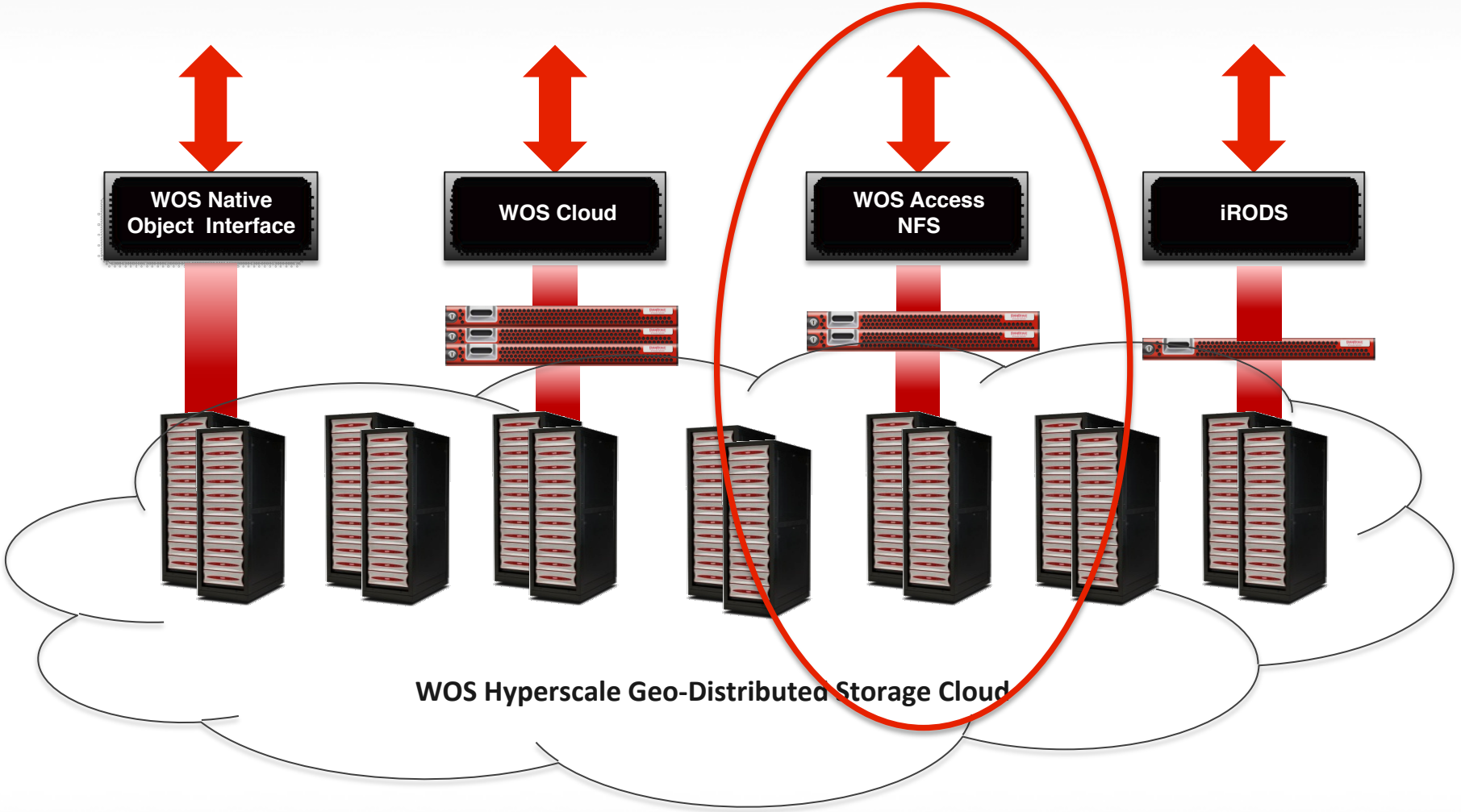
TokyoTech WOS Cloud Network



TokyoTech WOS Cloud Network

WOS Interfaces Drive HPC and Life Sciences Solutions

DataDirect NETWORKS
INFORMATION IN MOTION



WOS Access NFS

WOS Access NFS is a software/server solution that supports NFS V3 and V4

NFS Access Control Lists (ACLs), group/user levels, and identity authorization

- Authorization is per file/directory

Synchronizes NFS Gateways across multiple sites

- Single Federated NFS Namespace to 23PB

HA protected

- Active/passive failover

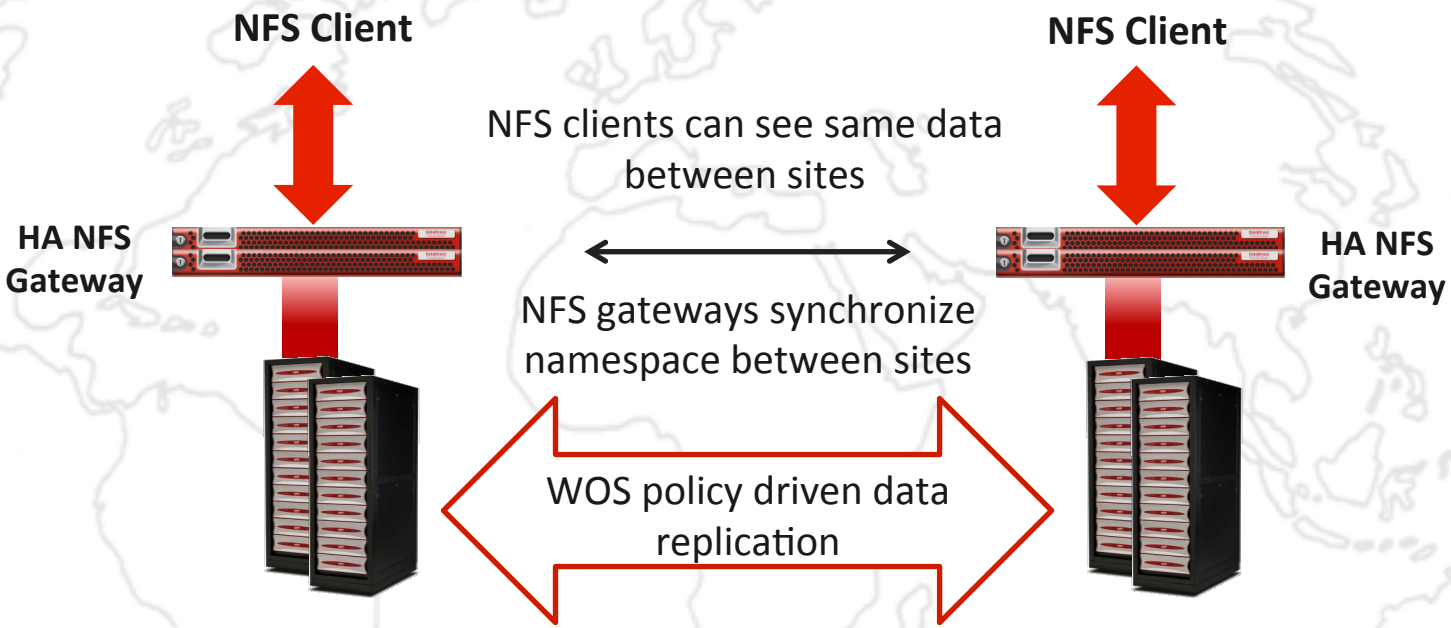
Intuitive Graphical User Interface (GUI)

Available as software only

WOS Access NFS GUI

WOS Access NFS

Geographically Distributed Single NFS Name Space



WOS NFS Access Key Differentiators

World's Largest, Easiest, Lowest Cost, Global NFS File System

Data Access is Location-aware for Lowest Latency

Industry Leading Data Protection and Data Integrity

- Self healing, instantaneous recovery from disk failure, not days
- Built in data integrity, no silent data corruption
- Replication built in, not added on

Industry Leading TCO

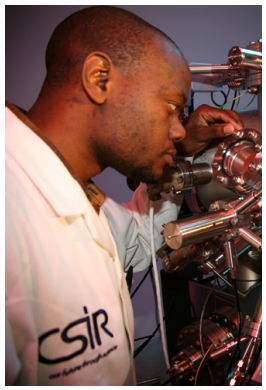
- Efficiency, manageability and reduced service response requirements
- WOS utilizes 99% of the disk platter v. 65-70% for FS/NAS implementations
- High density reduces power and cooling costs

WOS Access NFS Example User

DataDirect NETWORKS
INFORMATION IN MOTION

The Council for Scientific and Industrial Research (CSIR) in South Africa

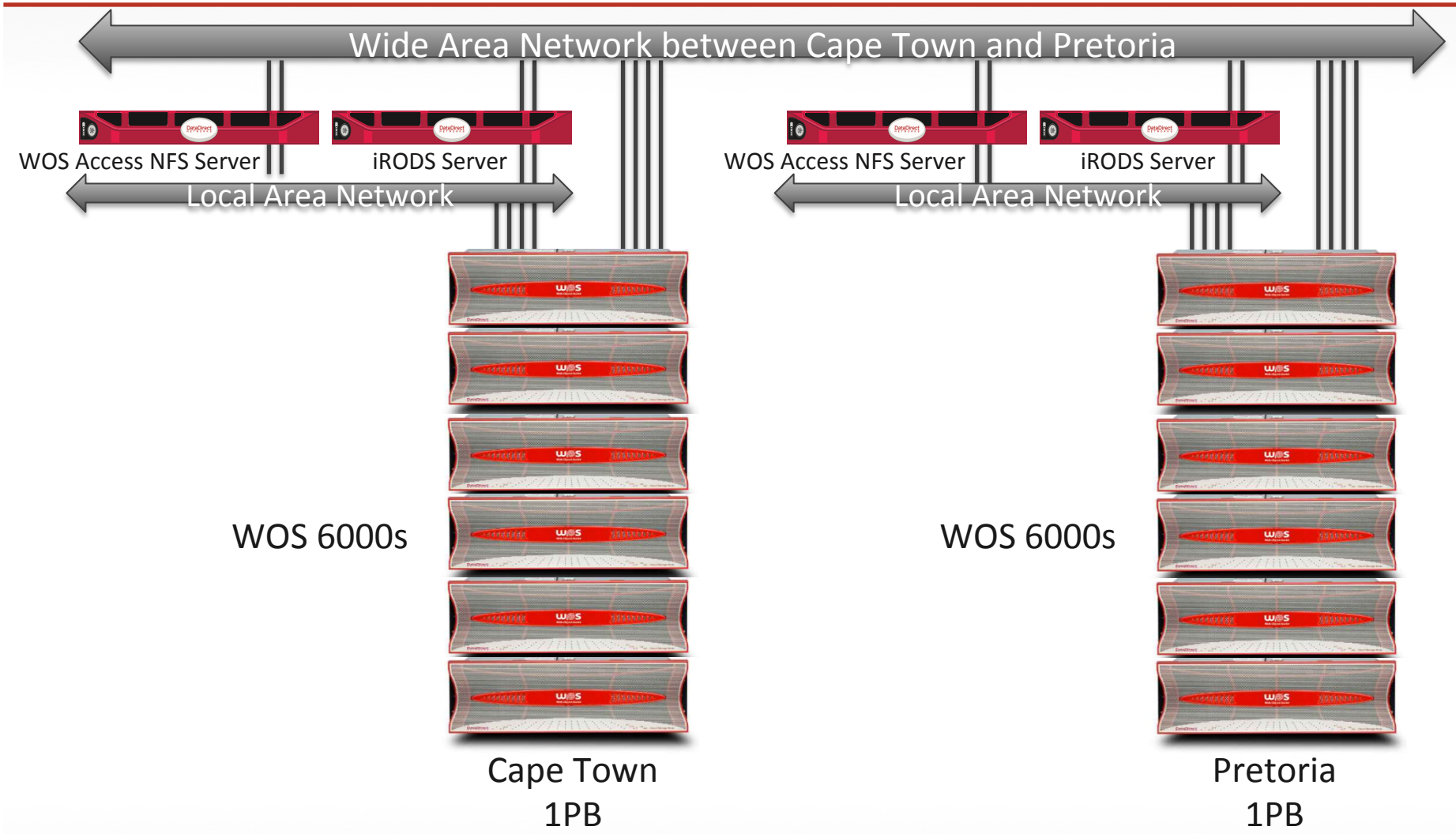
- The CSIR is one of the leading scientific research, development and technology implementation organizations in Africa
- WOS Stores Data for the Southern African Large Telescope, the largest optical telescope in the southern hemisphere, and the Karoo Array Telescope
- WOS also stores data for space science, bio-informatics, climate change and the paleao-sciences





CSIR
our future through science

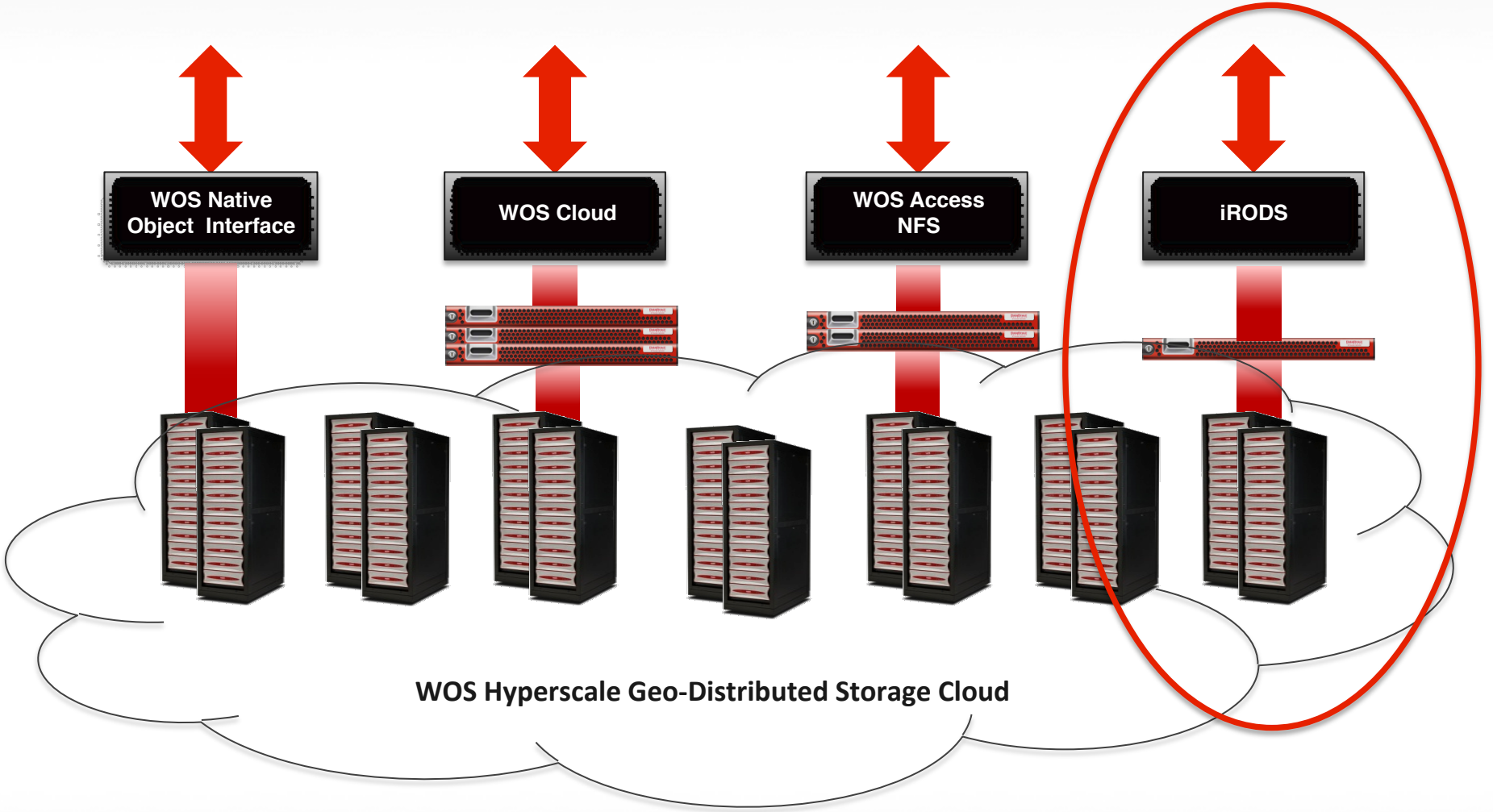
CSIR Configuration



Total Capacity for the CSIR WOS Cluster : 2 PB

WOS Interfaces Drive HPC and Life Sciences Solutions

DataDirect NETWORKS
INFORMATION IN MOTION



Automated, High Throughput Genome Sequence Processing Workflow

DataDirect NETWORKS
INFORMATION IN MOTION

3) iRODS initiates processing on compute cluster for image processing, base calling and further processing to produce final DNA data

Next Generation Genome Sequencer

Compute Cluster

iRODS Server / iCAT Database in 10KE VMs

DDN 10KE

i·R·O·D·S
Integrated Rule-Oriented Data System

WAN

iRODS Server

iRODS Server

DDN WOS 6000

DDN WOS 6000

DDN WOS 6000

1) 10KE running GPFS receiving data from a sequencer

2) iRODS registers sequencer output completed

4) iRODS moves the processed sequence data off to WOS

5) WOS replicates data across the WAN for protection and remote access and analysis

High Performance Storage
Data acquisition and processing

Distributed Data Repository
Data preservation and sharing

Integrated Rules Oriented Data System (IRODS)

DataDirect NETWORKS
INFORMATION IN MOTION

IRODS brings two things to the table

- A rules engine that automates data management
  - IRODS can execute a number of rules on how data is managed over time
  - Example:
    1. Data captured on very fast, expensive tier 1 storage
    2. At a certain time, migrate data to tier 2 storage
    3. Automatic replication to disaster recovery sites
    4. At a certain time, archive to tape or delete
  - Very powerful for Big Data HPC Workflows

- Annotates data with rich, domain specific metadata
  - Searching/retrieving data of interest is much harder than simply storing the data
  - Vital to the end users

WOS and IRODS

DataDirect NETWORKS
INFORMATION IN MOTION

What does WOS bring to the table?

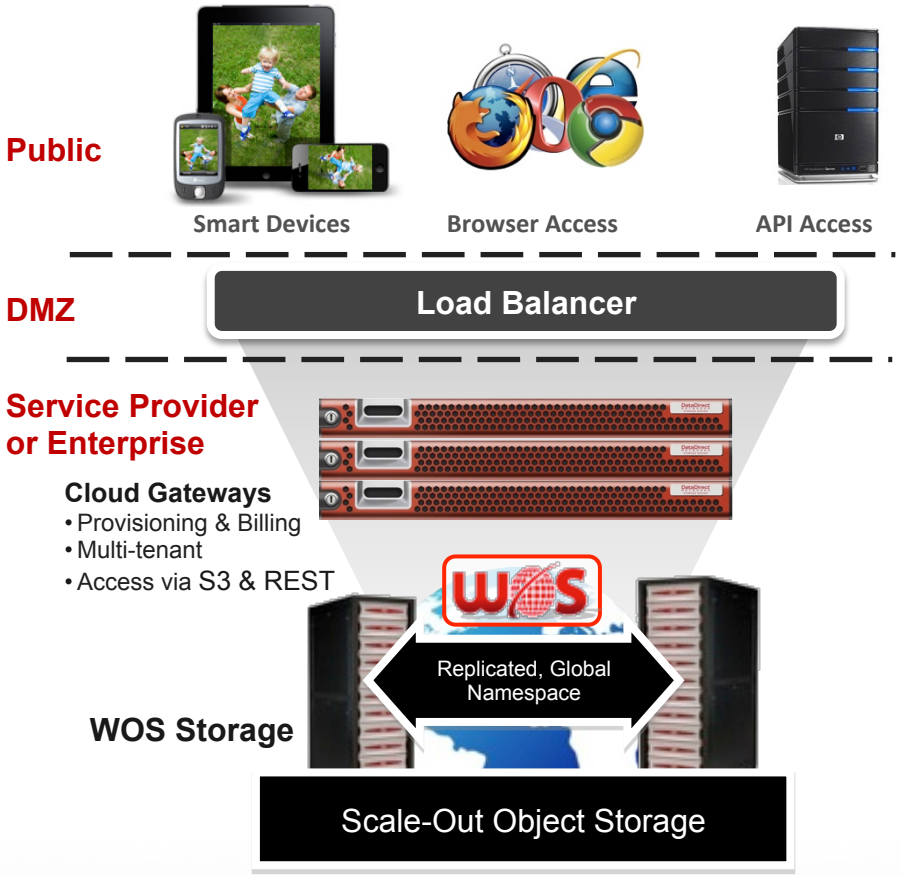
- Ease of Management
  - The same no matter how big or how distant
- Efficiency
  - Up to 2X more usable storage
- Data Integrity
  - Detect and repair bit-rot
- Data Availability
  - No RAID
- Failure Tolerance
  - Very important at petabyte scale
- Performance

WOS Cloud

DataDirect NETWORKS
INFORMATION IN MOTION

Efficiently build hyperscale storage for Public and Private Clouds

- **Offer industry leading, differentiated service**
  - Better service delivery, support, performance, cost, robustness, and SLAs
  - Flexible, pay as you grow scale
  - Remote manageability with no physical access
- **Cloud Platform Software**
  - Multi-tenancy support
  - S3 compatible & WebDAV APIs
  - Full CDMI Compliance
  - Integrates w/ existing provisioning & billing systems
  - Geographic location controls
  - Native smart client access



IRODS and WOS in HPC and Life Sciences

DataDirect NETWORKS
INFORMATION IN MOTION

"We have demonstrated the power of DDN WOS to facilitate collaboration on a level not previously achievable."
- Jos van Wezel, Data Access and Data Management group lead, Steinbuch Centre for Computing at the Karlsruhe Institute of Technology

KIT
Karlsruhe Institute of Technology

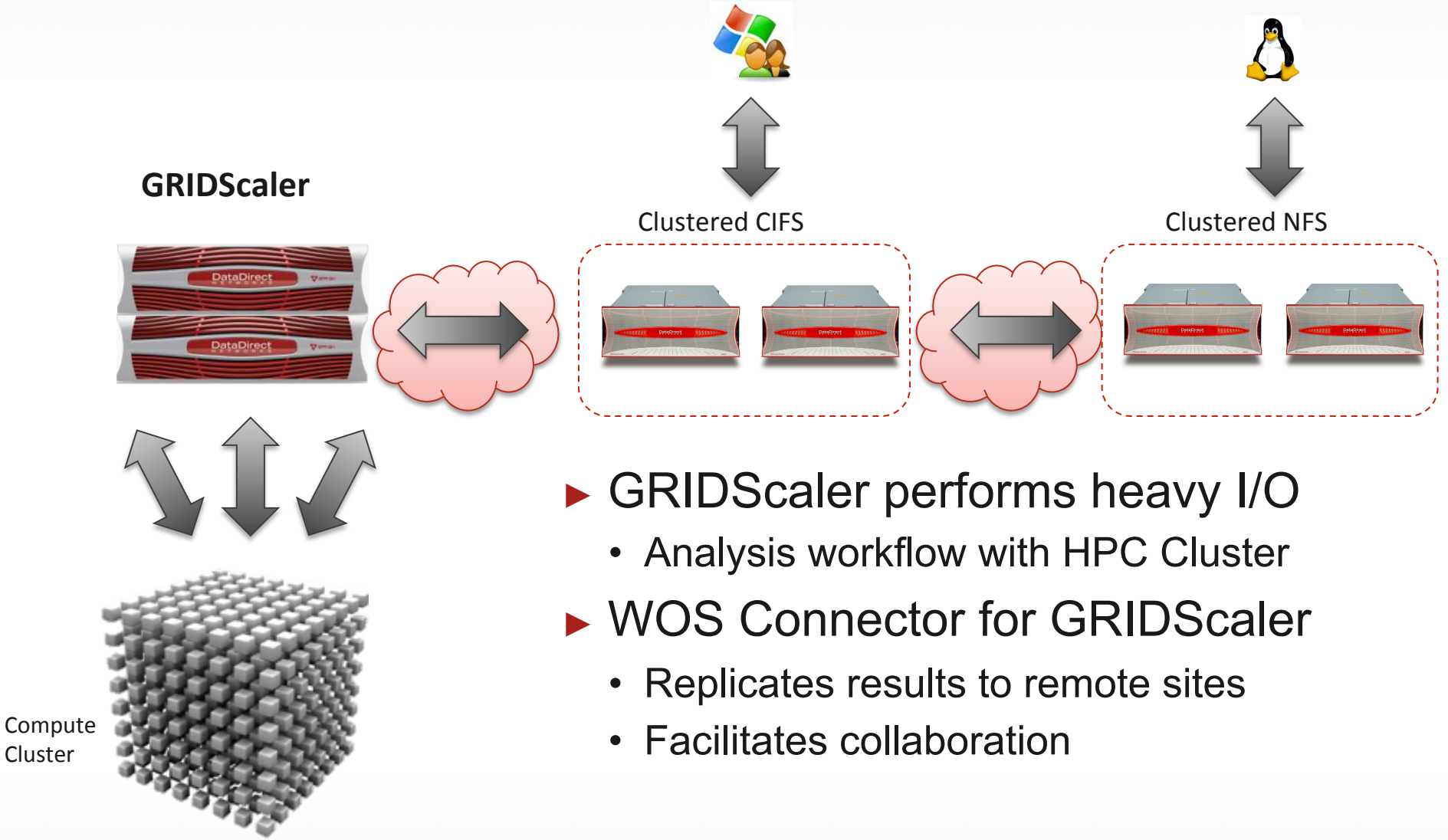
"Highly reliable and flexible storage at the multi-petabyte scale is critical for the research community. We are able to use WOS to create a globally accessible storage cloud."
- Chris Jordan, Data Management and Collections Group Lead, Texas Advanced Computing Center, the University of Texas at Austin

TACC

"DDN's WOS puts CSIR on a much lower cost curve compared to conventional storage, while also allowing our scientists at both Cape Town and Pretoria instant access to the immense amounts of scientific data we are generating."
– Dr. Happy Sithole, Director of the Center for High Performance Computing at CSIR in South Africa

CSIR
our future through science

GRIDScaler Bridge to WOS (Future directions)



- GRIDScaler performs heavy I/O
  - Analysis workflow with HPC Cluster
- WOS Connector for GRIDScaler
  - Replicates results to remote sites
  - Facilitates collaboration

Integration with Web Object Scaler

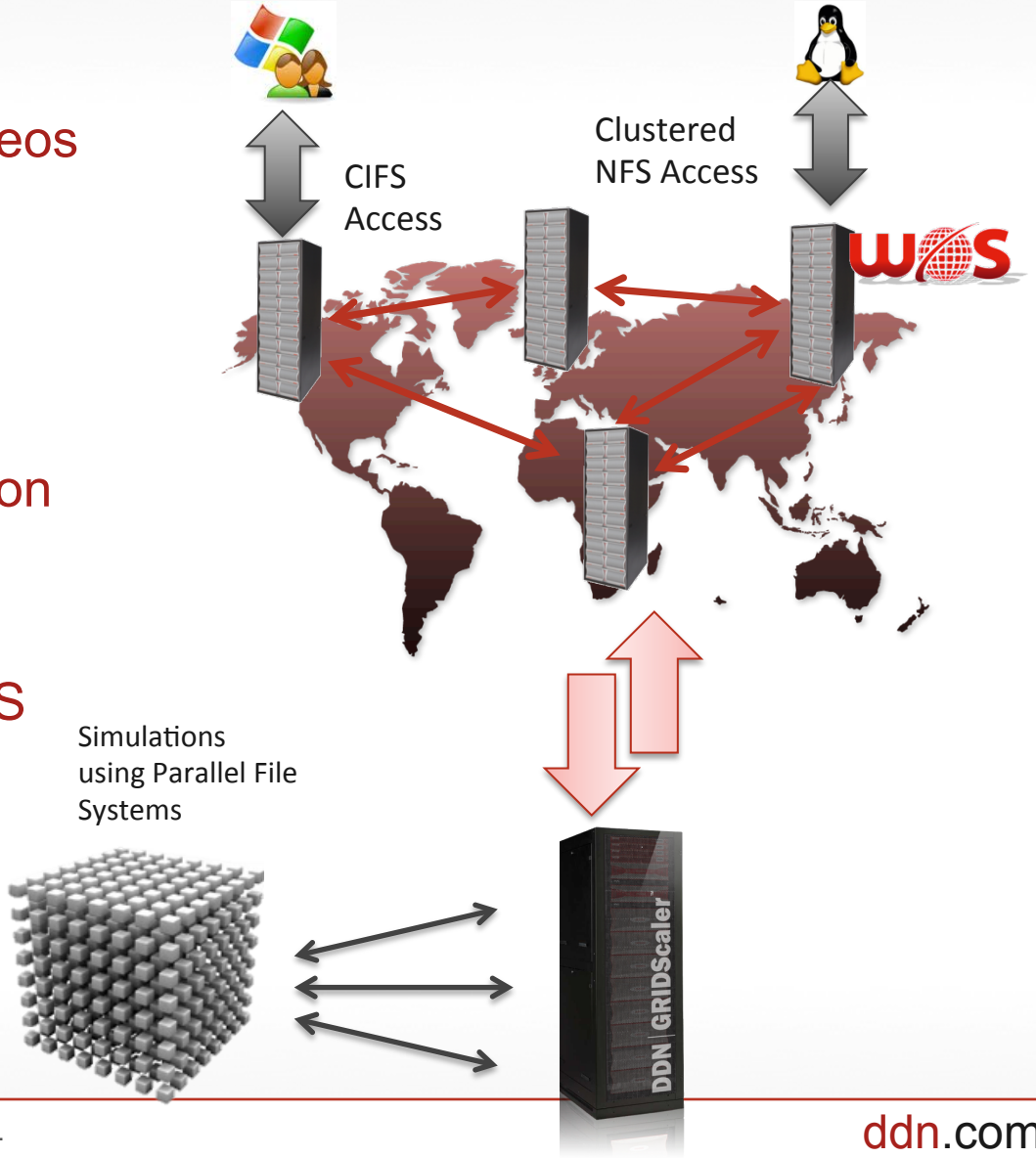
DataDirect NETWORKS
INFORMATION IN MOTION

Built for collaboration across geos

Simulate on GRIDscaler and distribute using WOS

Ingest using WOS access (NFS and CIFS) and simulate on GRIDscaler

Back up files safely to the WOS cloud for disaster recovery



WOS (Web Object Scaler) Summary

DataDirect NETWORKS
INFORMATION IN MOTION

GeoDistributed, Scale-out Object Storage System

- Hyper-Scalable Cloud Storage Foundation

TRUE End:End Object Storage

- Maximum Performance From Every Media

Easy to Manage at Hyperscale

- Single namespace, Single Global Cluster Interface

Autonomous, Self-Healing Big Data Infrastructure

- Intelligent, Fail-In-Place Architecture

Flexible Cloud Storage Service Platform

- Multimodal Access Featuring Billing & Multi-Tenancy

DataDirect™
NETWORKS
INFORMATION IN MOTION™

DDN | Web Object Scaler®

Thank You.