May 1, 2012

DataDirect™ NETWORKS
INFORMATION IN MOTION™

I/O: Toward The Exascale Era

Cray User Group

Keith Miller

Technical Director WW HPC and LS , DDN

DDN | The Scalability Leader

DDN is the world leader in massively scalable storage and processing technology for unstructured & big data applications.

Established: 1998

Revenue: $250M+ Per Year, Profitable & Growing

Headquarters: Chatsworth, California USA

Employees: 500+ Worldwide

Worldwide Presence: 4 Continents, 15 Countries

Installed Base: 1,000+ Customers; 50 Countries

Go to Market: Global Partners, VARs, Resellers

World-Renowned & Award Winning:

DDN | 2012 Update

INFRASTRUCTURE	PEOPLE	R&D
		
Organizational L1/2/3, ProServe **Systems** Oracle	**Organizational** • 140: Data Intensive Computing Team • World's Largest Lustre Front-Line	**Storage Fusion** • 1M+ of State-Machine Code • 100s of Engineers

– The World's Largest Independent Storage Company –

– Heavily Invested In Lustre and The Lustre Ecosystem –

– Driven By HPC –

Current Leadership

DataDirect NETWORKS
INFORMATION IN MOTION

Real-Time, HPC State Machine

- 1M Lines of Zero-Interrupt Storage Engine Code
- Highly-Parallel Storage Processing Architecture
- Adaptive RAM Cache for Mixed Workloads, Journals
- Embedded Virtualization For ExaScaler™ Appliances

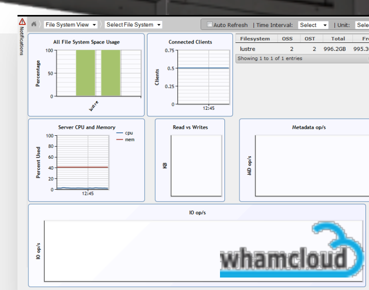
Quality Of Service

- Critical For Strided Writes & Reads
    - SFA Technology Maximizes Cluster Productivity
    - Performance Degredation Less Than 10%
- Real-Time Latency Management

Autonomous, Self-Healing Technology

- Automatic Drive Power Cycling
- Predicts and Prevents Drive Failures
- Minimizes Failure Instances by 80%

SFA12K



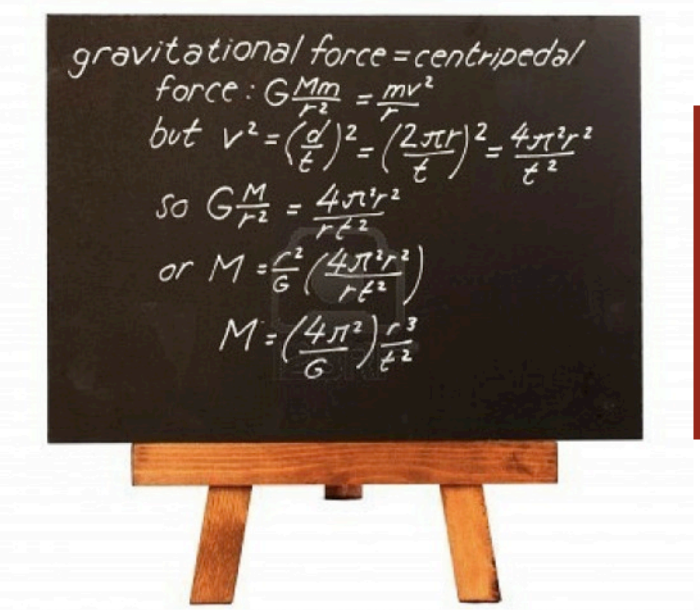
Chroma

DDN & WC | Partners in HPC

DataDirect NETWORKS Information in Motion

DataDirect NETWORKS Information in Motion T3S: Tech. Sales, Support & Service	• The world's largest team of HPC storage solution support and delivery experts • 1 interface; 1 solution
whamcloud Engineering	• The de facto developers of core Lustre
DataDirect NETWORKS Information in Motion Engineering	• 10 years of developing Lustre storage solutions • Real-Time, State Machine Leaders

A Look Toward Exascale



Exascale Systems will be big…

System attributes	2010	"2018"	
System peak	2 Peta	1 Exaflop/sec	
Power	6 MW	20 MW	
System memory	0.3 PB	32-64 PB	
Node performance	125 GF	1 TF	10 TF
Node memory BW	25 GB/s	0.4 TB/sec	4 TB/sec
Node concurrency	12	O(1,000)	O(10,000)
System size (nodes)	18,700	1,000,000	100,000
Total Node Interconnect BW	1.5 GB/s	200 GB/sec	
MTTI	days	O(1 day)	

Billion-Way Parallelism Will Be A Reality

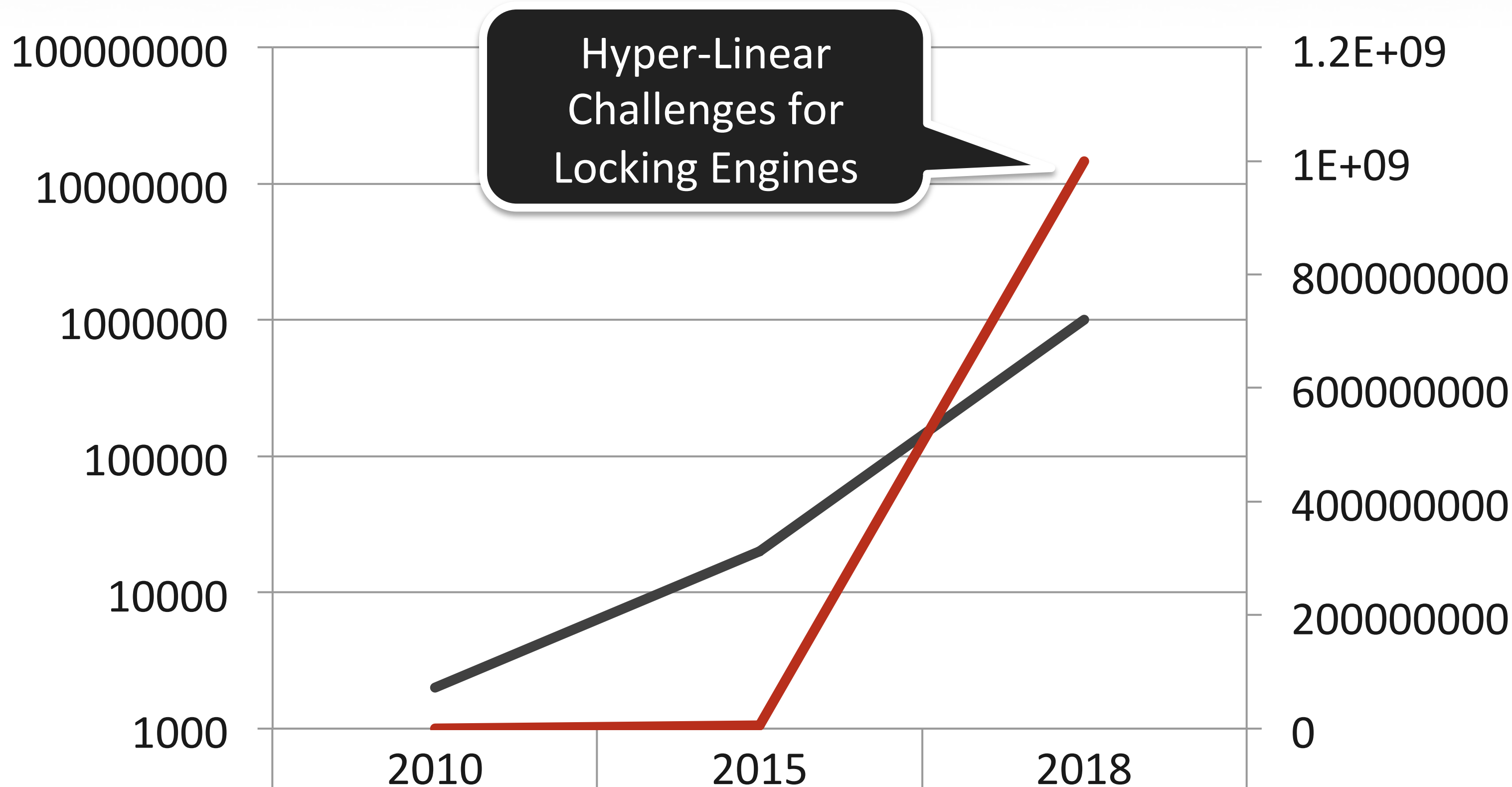
New Configurations Are Outfitting 1000s of Cores Per Data Center Rack

The impact on Storage:

- **B-Way** Parallelism will challenge Exascale file system locking and data integrity mechanisms
- New **memory-class storage** will require new approaches for tiering and data locality
- Exascale consortiums are seeking new methods to **reduce cluster and network workload**

Hyper-Concurrency is a Result of Exascale CPU Evolution

DataDirect NETWORKS INFORMATION IN MOTION



	2010	2015	2018
Performance (TF)	2000	20000	1000000
Concurrency	200000	5000000	1000000000

…Big Data is already in HPC!

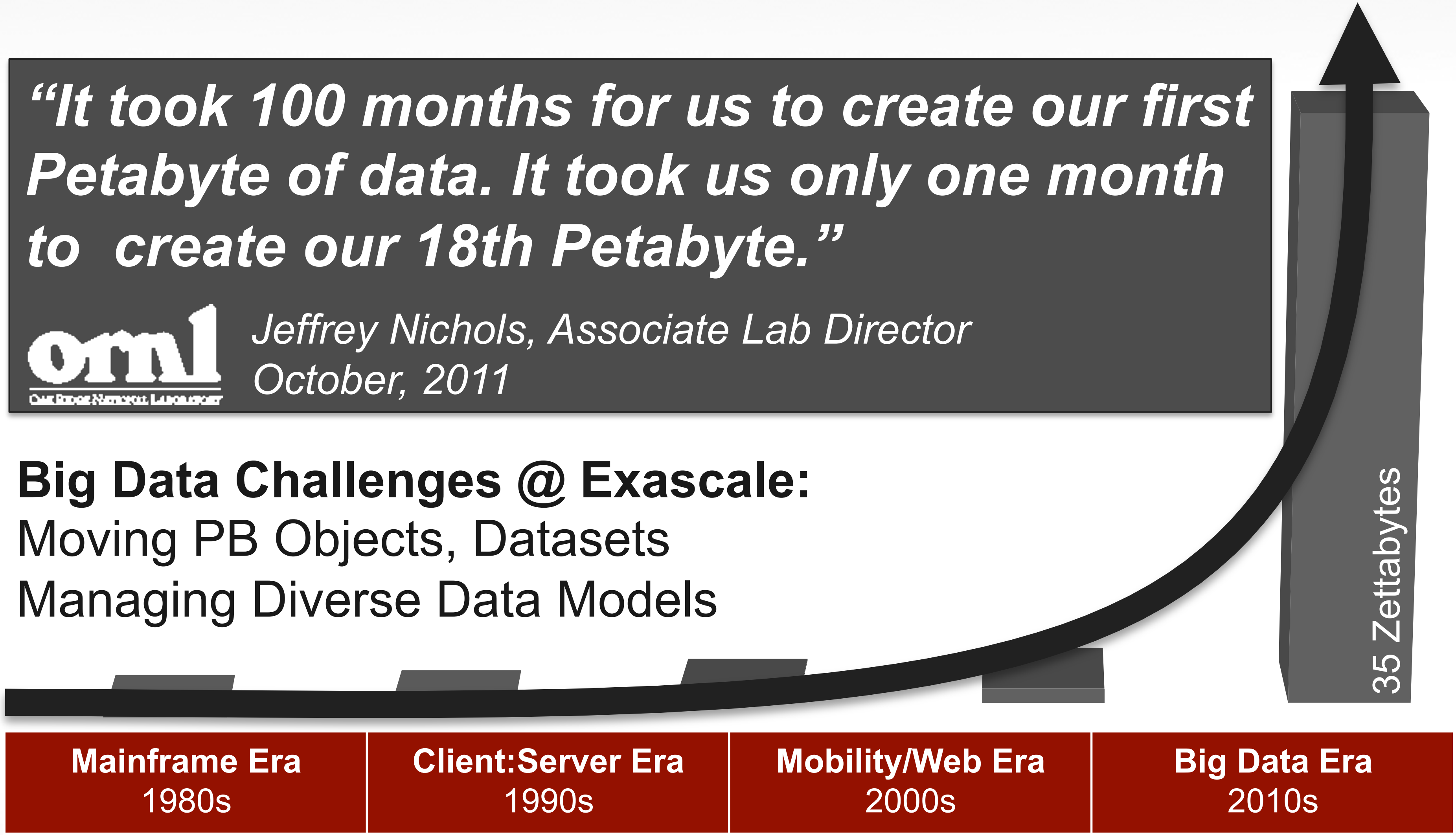
DataDirect NETWORKS
INFORMATION IN MOTION

"It took 100 months for us to create our first Petabyte of data. It took us only one month to create our 18th Petabyte."

ornl
OAK RIDGE NATIONAL LABORATORY

Jeffrey Nichols, Associate Lab Director
October, 2011

Big Data Challenges @ Exascale:
Moving PB Objects, Datasets
Managing Diverse Data Models



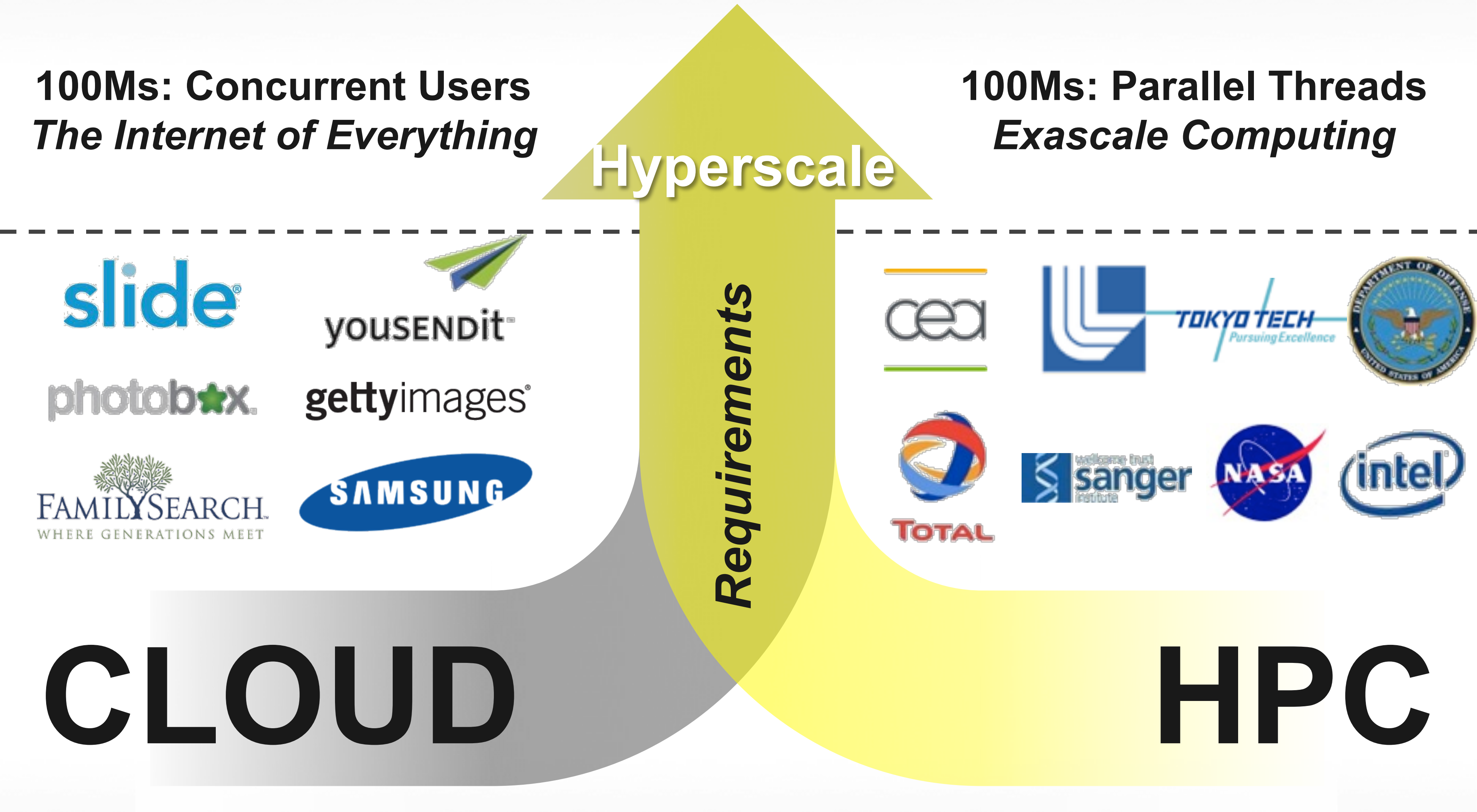
Source: IDC 2011: The Expanding Digital Universe.

Convergence @ HyperScale

DataDirect NETWORKS
INFORMATION IN MOTION

100Ms: Concurrent Users
The Internet of Everything

100Ms: Parallel Threads
Exascale Computing



Today's IO Models: Too Much Clutter

DataDirect NETWORKS INFORMATION IN MOTION

Simplify the layers of Application Stack

- Collapse 5 separate data structures into one

Each layer has its own "Database"

- Whether its named POSIX or SQL

Each layer has failover and redundancy

Each layer has complex recovery

Each layer has garbage collecting and maintenance

Data Transparency will address these issues

Object Storage Is Enabling A Re-Think On Data Layouts

Science – NetCDF, HDF5, SciDB, etc.

- Self-describing data
- Metadata built into the file itself

Web – NoSQL stores

- Flexible data structures (schemas)
- Easy
- Speed, scale

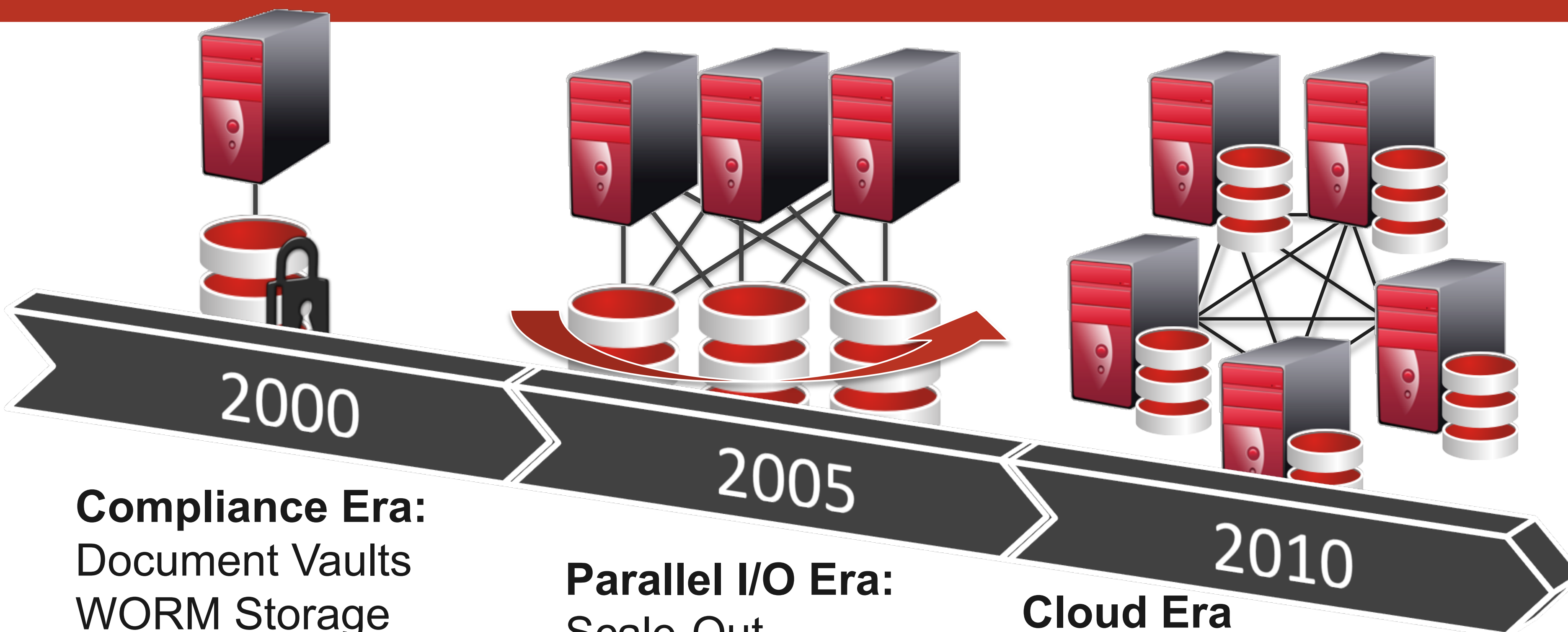
Objects (Cloud)

- Flat, global, arbitrary naming
- Integral sector management
- User-defined metadata

The Evolution of Object Storage

"Hundreds of millions of people use object storage every day – and don't even know it."

StorageMojo



Compliance Era:
Document Vaults
WORM Storage
Eg. EMC Centera

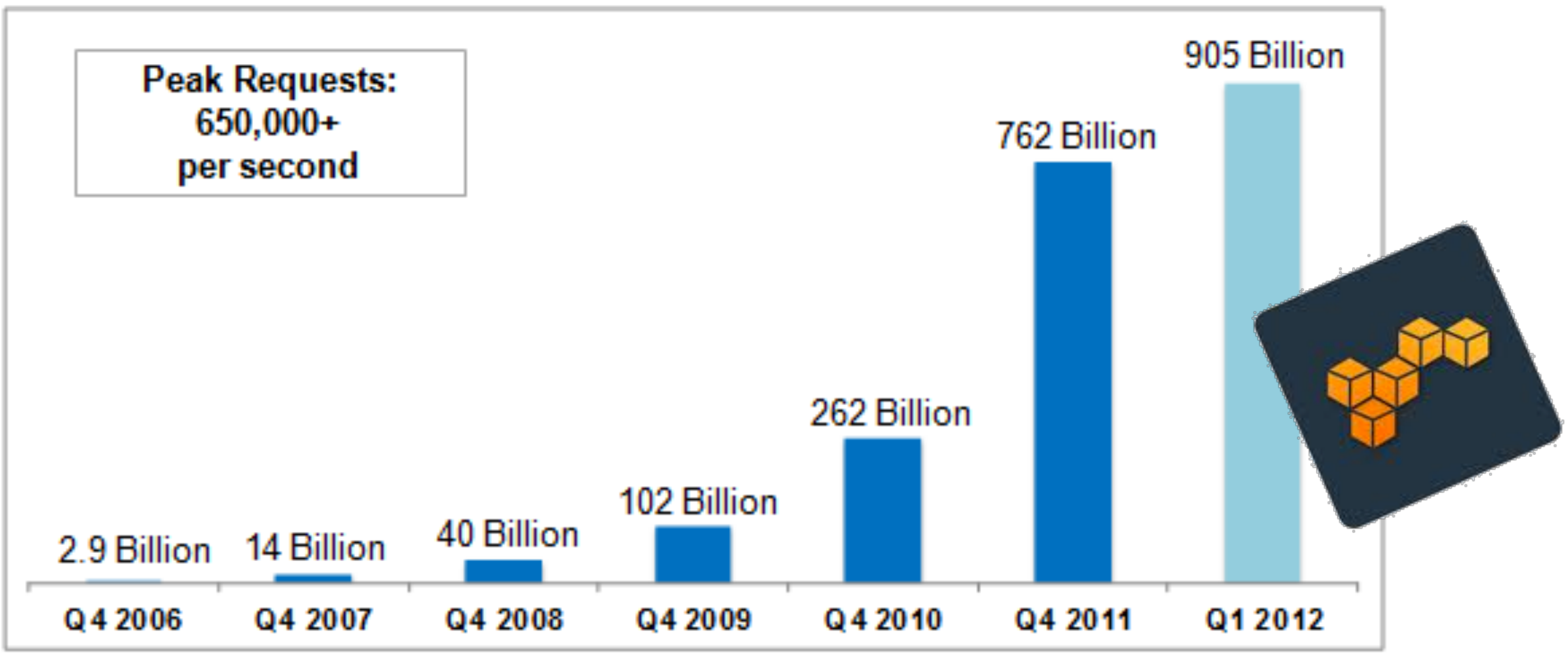
Parallel I/O Era:
Scale-Out
Eg. Lustre, GPFS

Cloud Era
Peer:Peer, Hyperscale
Eg. Amazon S3

A Hyperscale Case Study

DataDirect NETWORKS
INFORMATION IN MOTION

The Cloud Scales: Amazon S3 Growth



Total Number of Objects Stored in Amazon S3

Source: Amazon S3 Blog

Storage: Sources of Latency

Hardware Chain

- Disk drive servo operation
- Multiple SCSI layers
- Multiple bus transitions
- Memory bandwidth limitations
- Network service latencies

Software Chain

- Memory copies
- Kernel operations
- Layers of consecutive operations including the service of V-nodes, I-nodes and FAT
- Serial data transport processes

WOS is designed to reduce latency in all phases of data capture and retrieval

DDN | Hyperscale Initiative

Understand the data usage model in a collaborative environment where immutable data is shared and studied

- A simplified data access system
- Eliminates the concept of FAT, extent lists to maximize efficiency
- Reduce the instruction set to only PUT, GET, & DELETE
- Add the concept of locality based on latency to data and load balance

Abandons storage convention entirely

WOS | Overview

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 Exascale Advantages

WOS is an end to end object placement file system.

- WOS has no concept of fragmentation
- Objects 1MB or less are stored in contiguous space minimizing actuator usage in rotating media and simplifying internal maps in solid state media

WOS is efficient

- Objects are immutable so there is no concept of “File open for xWRITE”
- Locking is completely eliminated
- Like size objects are always stored together in Object Resource Groups (ORG) so that there is no concept of “garbage collection” on a block level basis
- Operations of PUT and GET are accomplished in one concise internal transaction layer

WOS Exascale Advantages (con't)

WOS addresses data corruption in multiple dimensions

- All objects are written with a checksum
- The checksum is evaluated for every GET and every bus transition
- All objects can be written with erasure codes distributed on multiple storage devices
- Nodes automatically recover data in the case of silent data corruption with a rapid, object aware, rebuild operation
- Nodes automatically recover data in the case of media failure

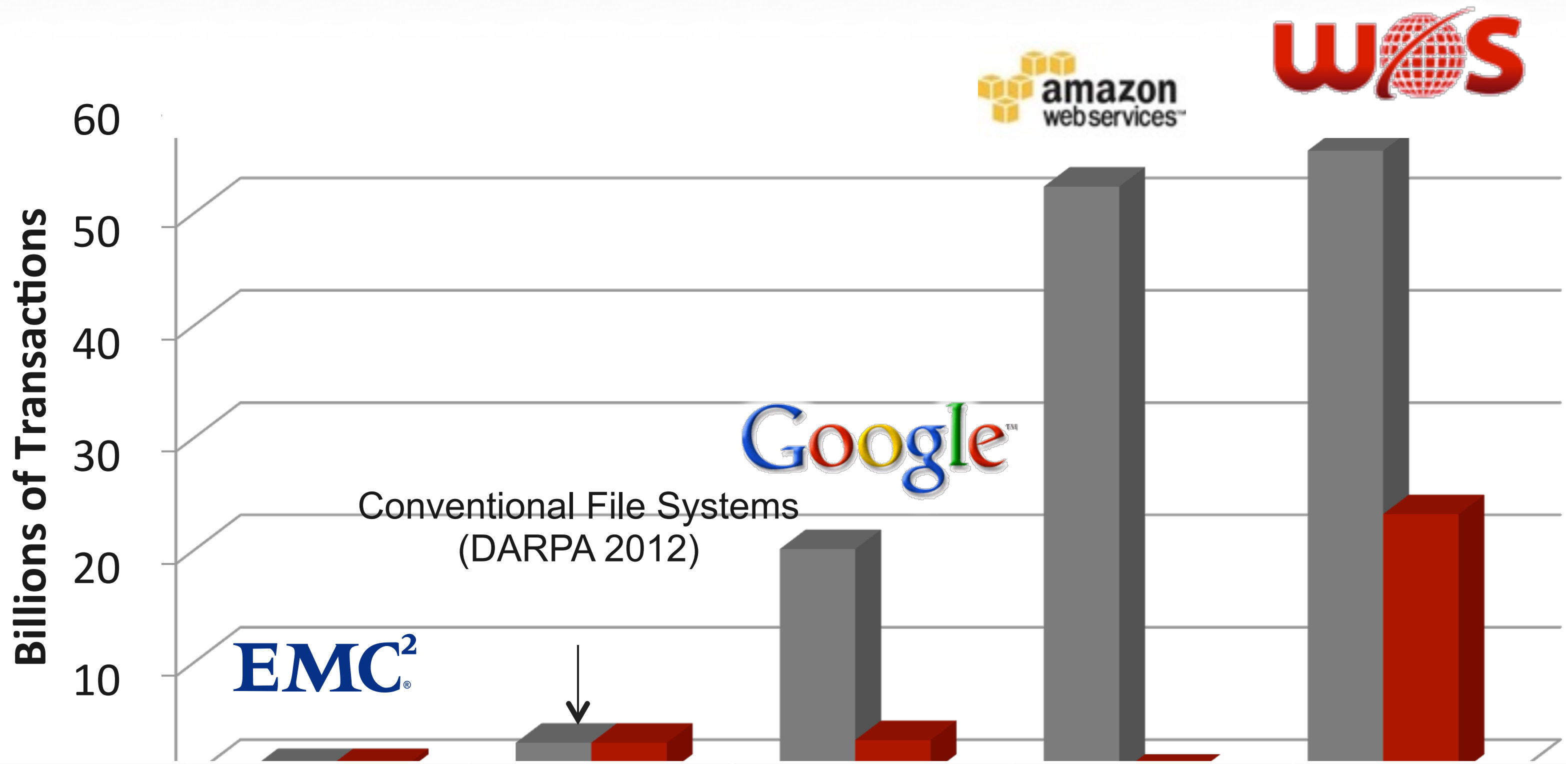
WOS was designed for large scale data operations

- Test limits are 256 billion objects utilizing 256 nodes

WOS is a peer to peer data distribution system which could be utilized to enable collaboration

The Power of Object Storage

DataDirect NETWORKS
INFORMATION IN MOTION



	EMC Atmos	HPCS	Megastore	Amazon S3	WOS
Reads/Day	500,000,000	2,764,800,000	20,000,000,000	32,000,000,000	**55,472,947,200**
Writes/Day	500,000,000	2,764,800,000	3,000,000,000	-	**23,113,728,000**

Amazon: http://aws.typepad.com/aws
EMC: http://reg.cx/1P1E
HPCS: http://www.spscicomp.org/ScicomP13/Presentations/IBM/GPFSGunda.pdf
Megastore: http://highscalability.com/blog/2011/1/11/google-megastore-3-billion-writes-and-20-billion-read-transa.html

Intelligent Object Storage Lays Foundation For Exascale Efficiency

DataDirect NETWORKS INFORMATION IN MOTION

Object Stores

- Versatile data model that works with distributed block management, security and data reliability/recovery
- Tiered, Scalable to B's of reads/writes per second

Converged Storage & Processing

- Pre and post processing functions and achieving integrated ILM services while making applications more data aware

Integrated Analytics

- Integrated Map-Reduce and Analytics services designed to turn PBs of data into knowledge

DDN | Directions

Even More Integrated,
Comprehensive
Appliance Management Tools

Exascale Engineering
Objects, Data Models,
Storage-Class Memory

Stay Tuned….

www.ddn.com

DataDirect™ NETWORKS

INFORMATION IN MOTION™

Thank You