DataDirect™
NETWORKS
INFORMATION IN MOTION™

Exploring Efficiencies in Data Reduction, Analysis, and Distribution

IEEE Mass Storage Symposium

April 18, 2012

Dave Fellinger

Chief Scientist, Office of Strategy & Technology

DDN Summary

Highly-scalable & high-efficiency storage solutions that enable customers to accelerate time to results & scale simply as data sets grow, gaining competitive advantage through resolving performance & capacity challenges.

- **Established:** 1998
- **Financials:** Over $250M Annually, Profitable, Growing
- **Headquarters:** Chatsworth, California USA
- **Employees**: Approximately 400 Worldwide
- **Customers:** Over 1,000 Worldwide
- **Footprint:** 17 Industries, 4 Continents, 49 Countries
- **Go to Market:** Global Partners, VARs, Resellers
- **Key Market Segments:**

  HPC & Life Science, Cloud & Web, Rich Media, Security

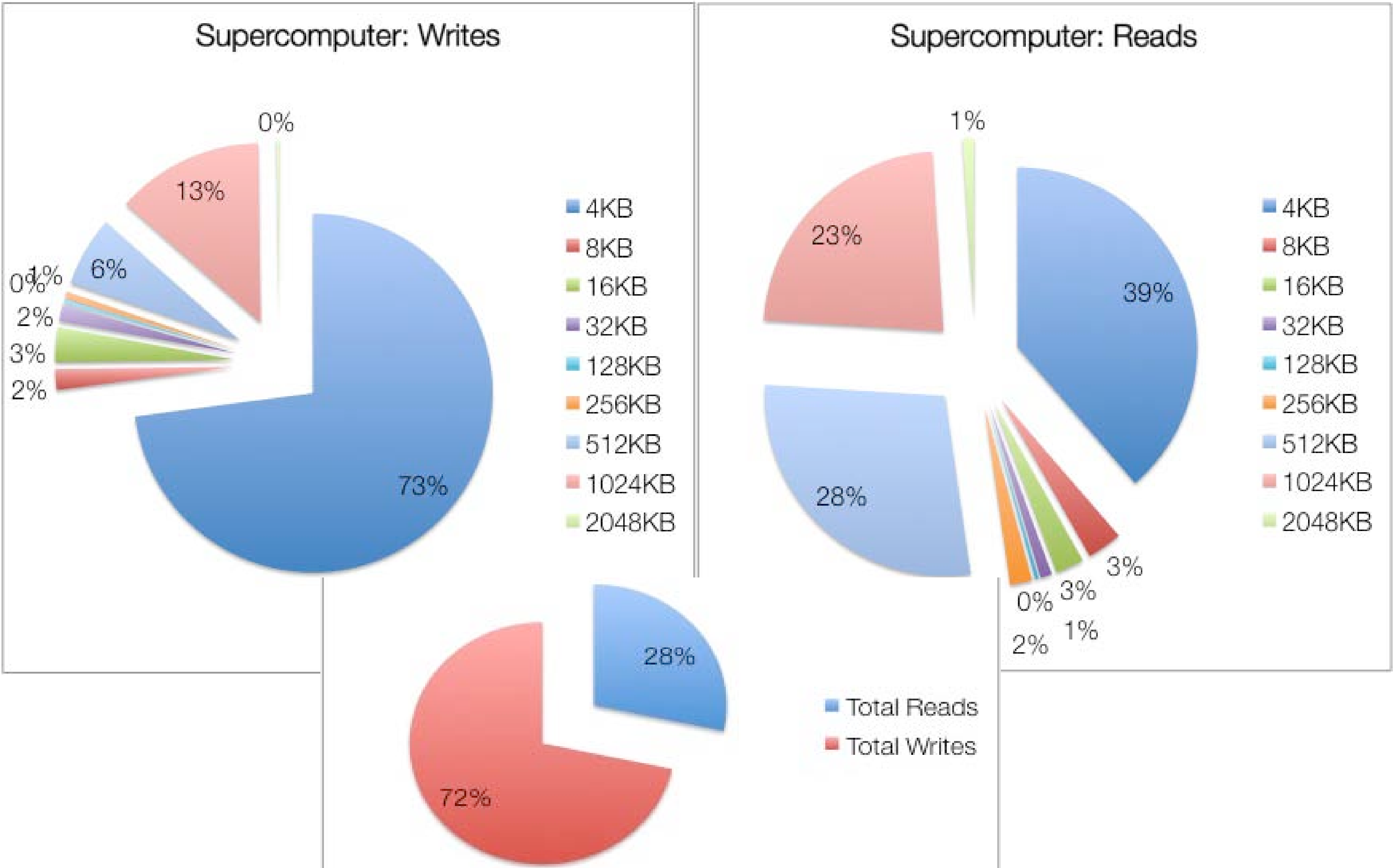
Challenges of Extracting Knowledge From Data

- Data reduction and distribution is a critical process in the feedback loop of iterative science.
  - Huge amounts of data must be captured and stored
  - Processes used to execute data conversions or reductions
  - Reduced data must be distributed and will be analyzed locally by globally distributed researchers
  - Collaborations are generally established to visualize the results
  - The entire process can then feedback required changes to the process
- Every process has inherent latencies that must be reduced or eliminated if possible.

Sources of Latency

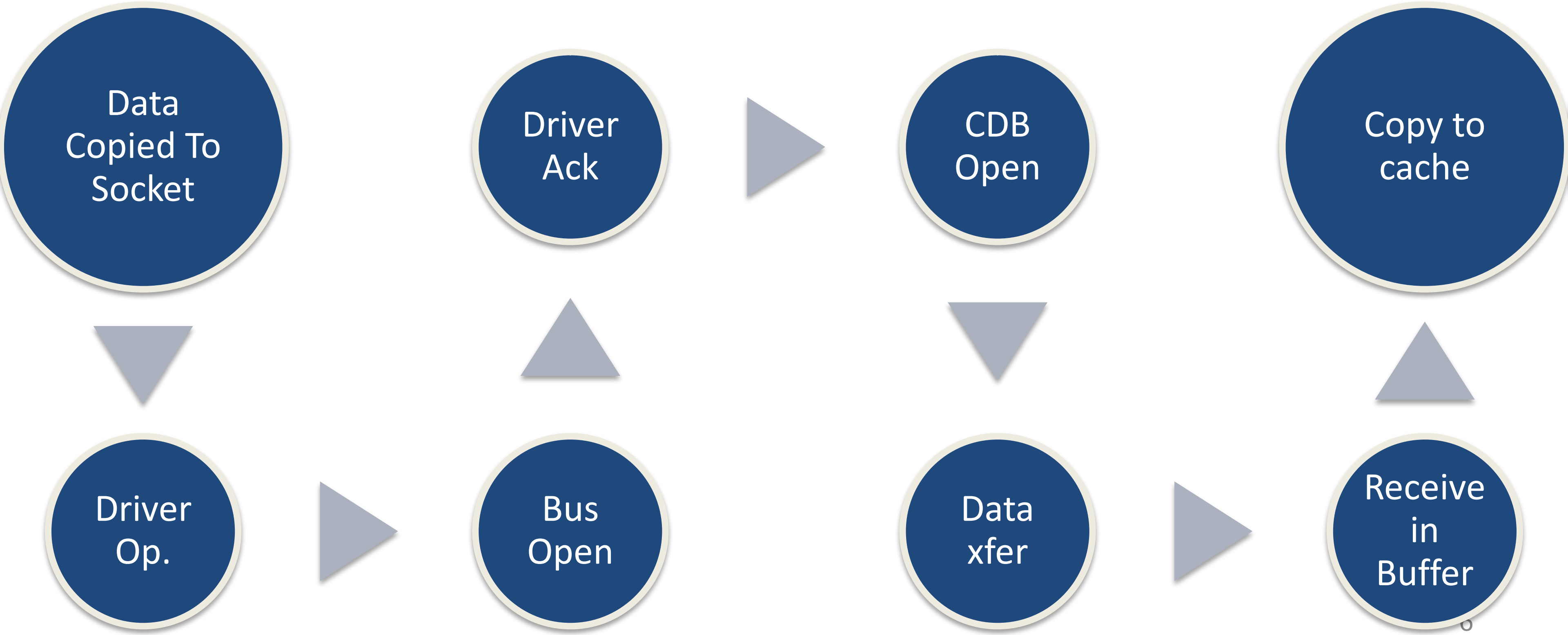
- Hardware Chain
  - Disk drive servo operation
  - Multiple SCSI layers
  - Multiple bus transitions
  - Memory bandwidth limitations
  - Network service latencies
- Software Chain
  - Memory copies
  - Kernel operations
  - Layers of consecutive operations including the service of V-nodes, I-nodes and FAT
  - Serial data transport processes

HPC Data Patterns



“Traditional” File Access

- File servers are connected to storage devices by serializing devices such as HBAs or HCAs
- Multiple steps executed to move data to/ from a server;



Data Copied To Socket

Driver Op.

Bus Open

Driver Ack

CDB Open

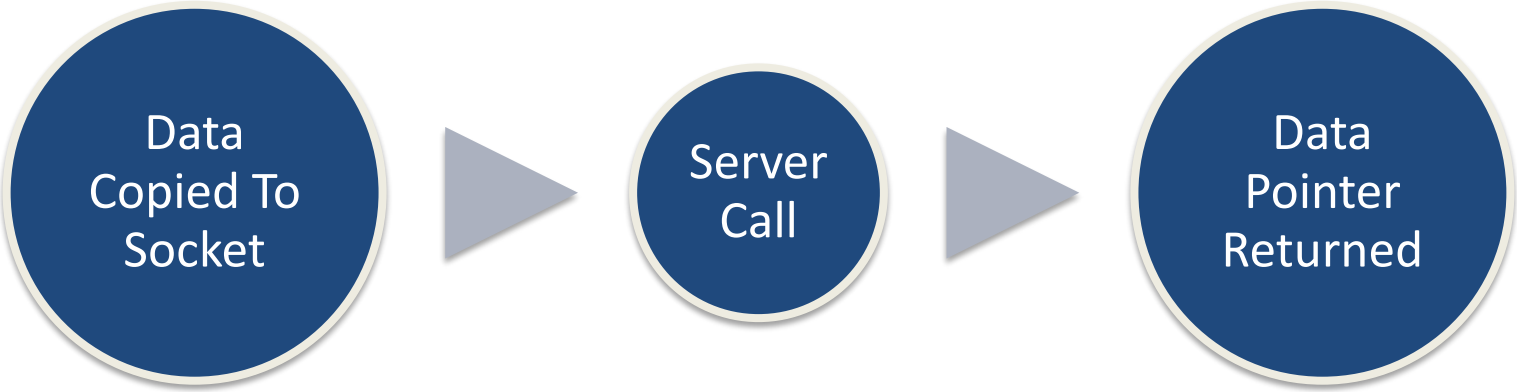
Data xfer

Receive in Buffer

Copy to cache

Efficient Alternative File Access

- File servers are run as virtual machines within the storage system in a shared memory environment with the storage cache
- The steps to move data to or from the storage;



Data Copied To Socket

Server Call

Data Pointer Returned

“Traditional” Data Reduction

- Data reduction, manipulation, filtering, resolution shifting, etc. is done in an external network connected processor.
- The steps required to execute a process must first include moving data to the processor;
  1. The file server builds a front end socket
  2. A routine is called that appends headers and footers to data packets so that a TCP transport layer can move the data from the server to the processor
  3. The processor receives the frames from a switch, strips the headers and footers, reorders the packets if required, and places the data in user space for manipulation

Efficient Alternative Data Reduction

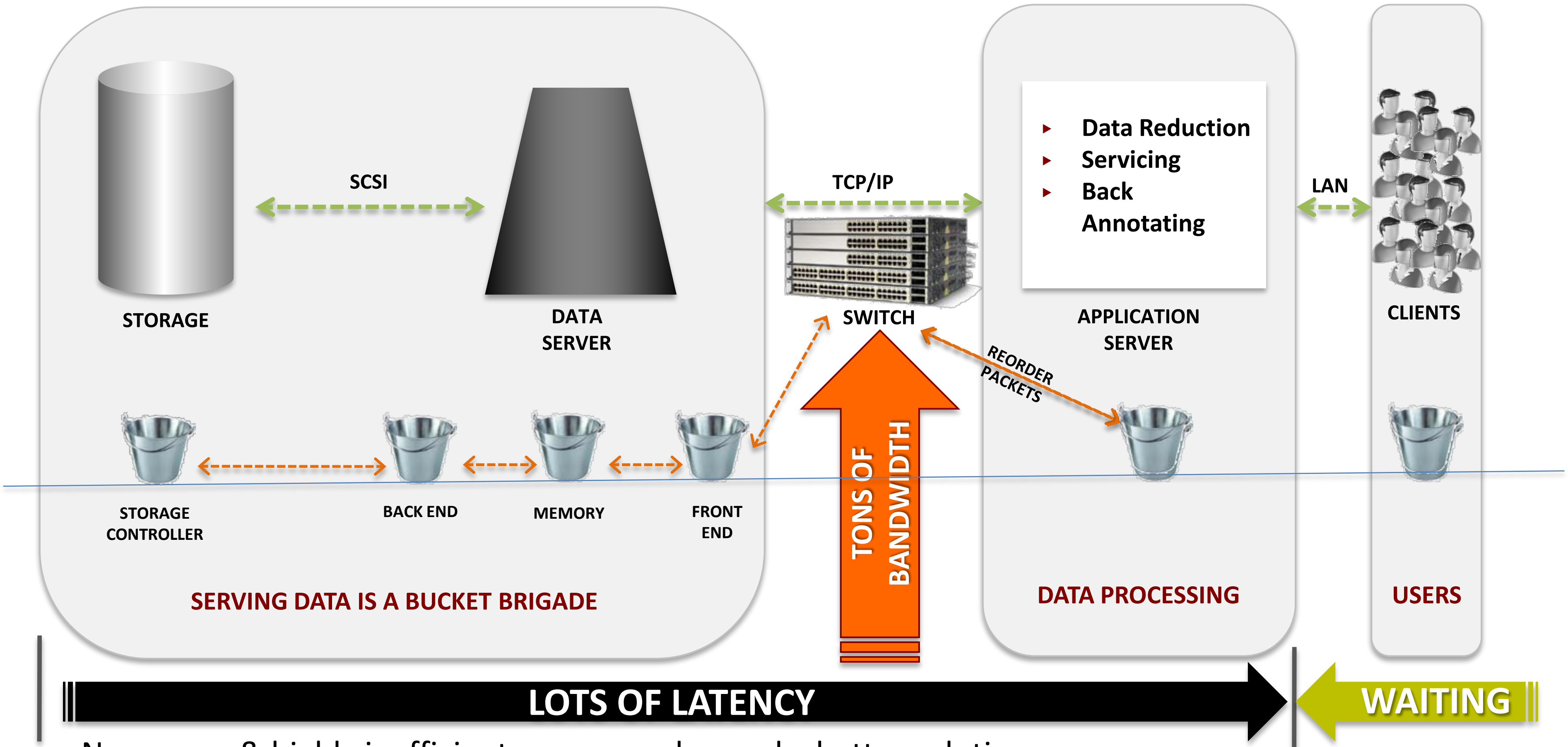
- An image including the file system and the data reduction process is run as a virtual machine on the storage system in multiple cores with dedicated cache.
- Steps to obtaining reduced data are executed removing both the SCSI bus transaction and the TCP/IP layer;
  - The user requests data through the process
  - Reduced data is transported through the TCP layer
- An alternative is to start a scheduled process on the storage that processes data from a “raw” data LUN to a LUN containing processed data.

i·R·O·D·S

Integrated Rule-Oriented Data System

DDN has partnered with RENCI & Dice to simplify & accelerate data reduction with DDN in-storage processing with iRODS.

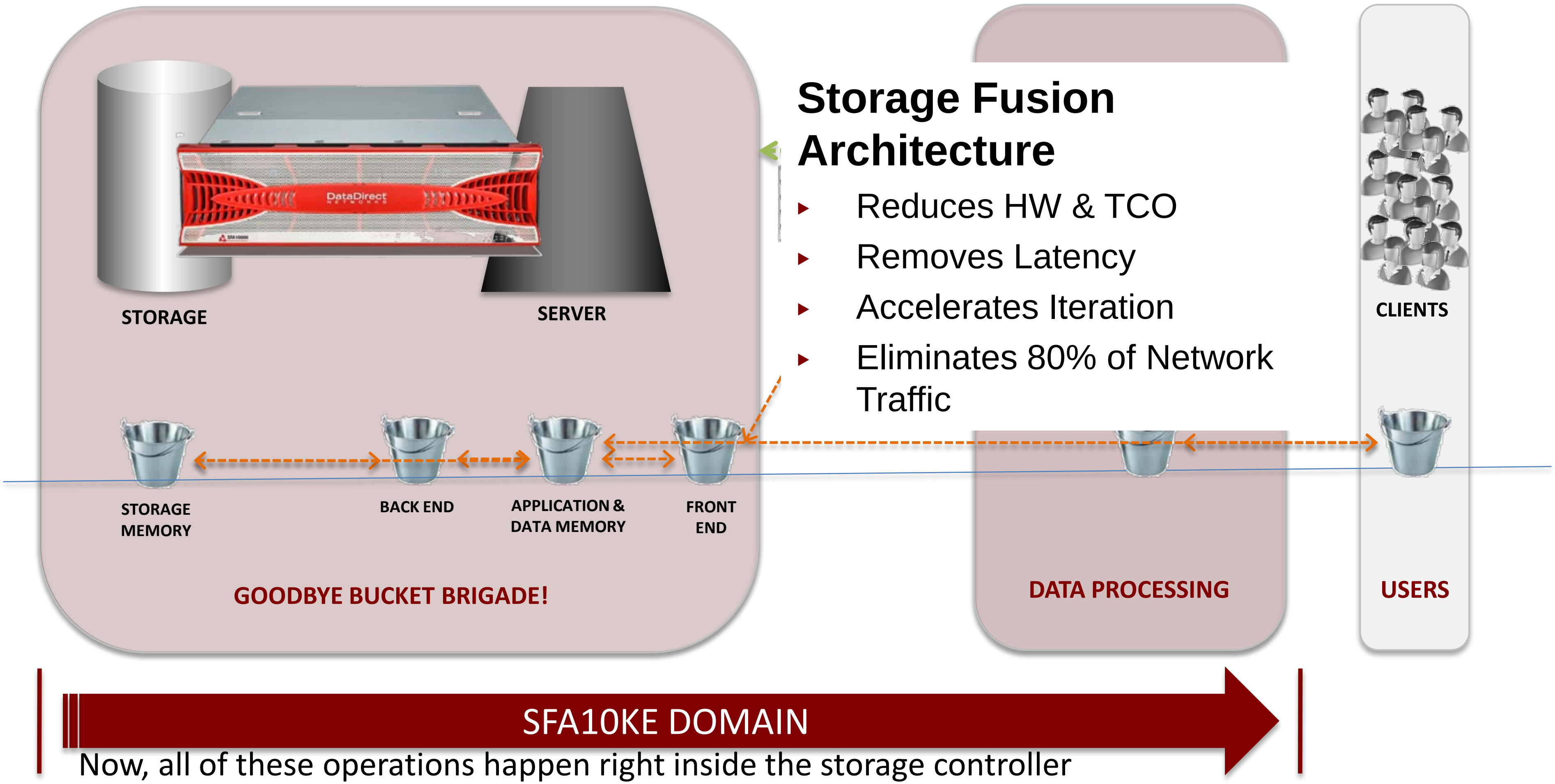
Inefficient Bucket Brigade of Protocols



- Data Reduction
- Servicing
- Back
Annotating

Numerous & highly inefficient processes demand a better solution

Reduction of Latency and Network Traffic



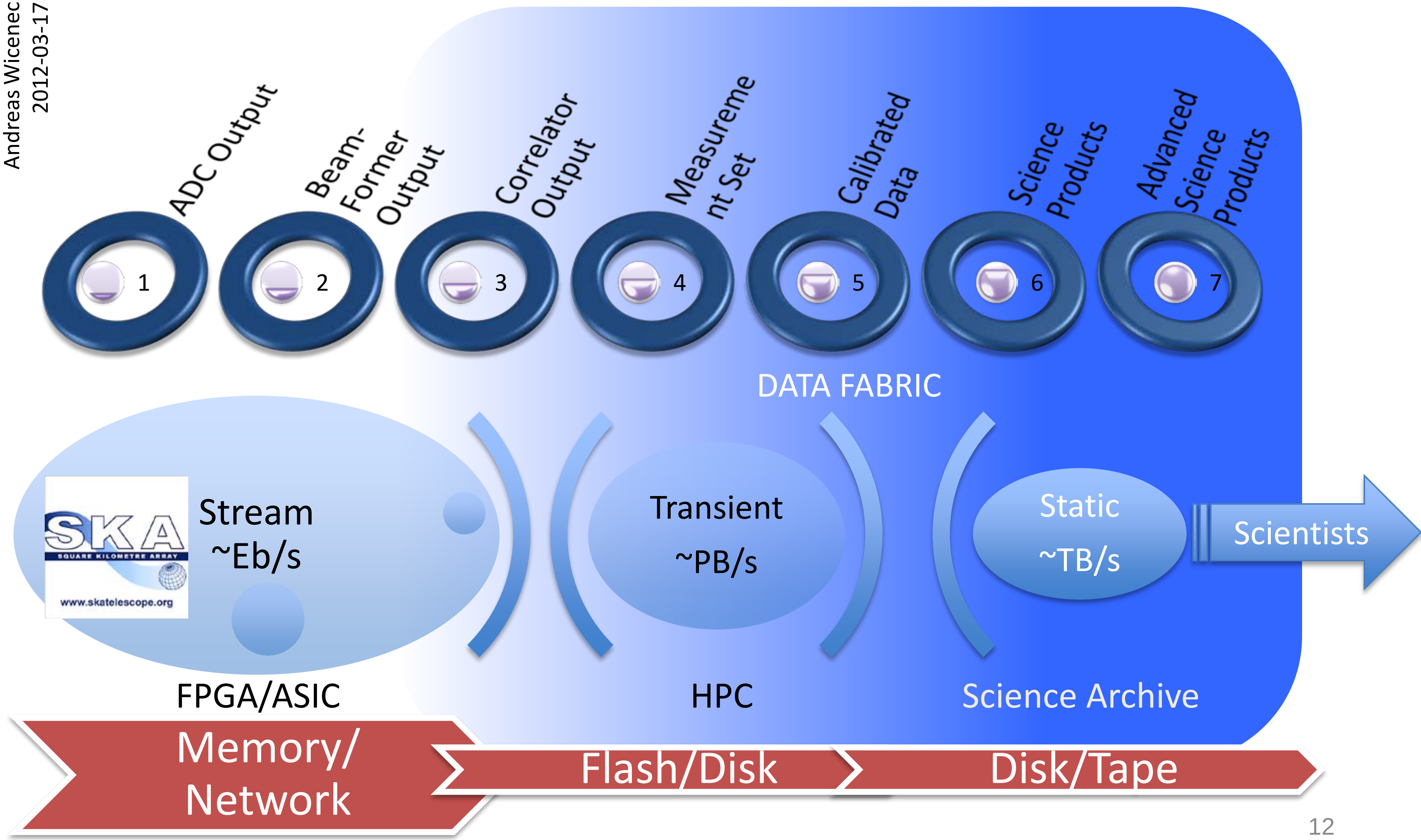
Storage Fusion Architecture

- Reduces HW & TCO
- Removes Latency
- Accelerates Iteration
- Eliminates 80% of Network Traffic

Now, all of these operations happen right inside the storage controller

Practical Example : ICRAR/SKA

Andreas Wicenec
2012-03-17



ICRAR / SKA Research (con't)

Andreas Wicenec
2012-03-17

- Data capturing covering a whole continent
- Globally distributed data reduction/analysis
- Requirement to be able to trace the whole process and optimize the data life-cycle to minimize operational costs while maximizing the scientific return.
- Solution: Assign unique object IDs to data products as early as possible and trace them throughout the data fabric (memory, disks, tapes).
- Move processes to data, co-locate processing & data
- Research in multi-resolution lossy and lossless compression distributed over HPC storage hierarchy. Storage performs compression/de-compression.

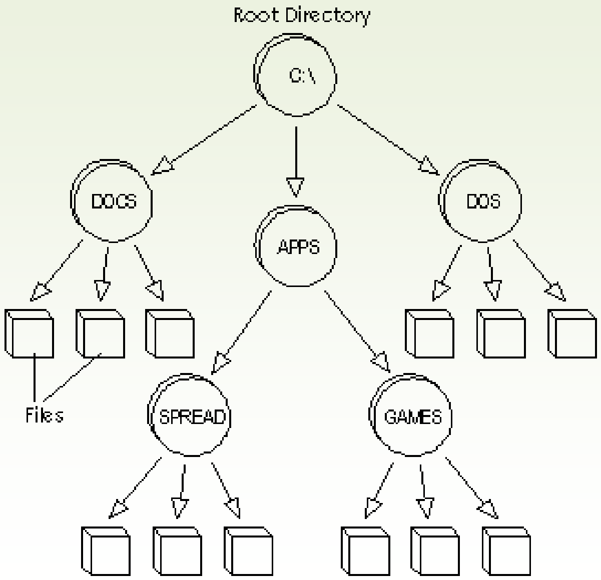
WOS, Data Distribution @ Hyperscale

- Understand the data usage model in a collaborative environment where data is shared and studied but never modified.
- Build a simplified data access system minimizing parsing and block level layers.
- Completely eliminate the concept of I-node references including FAT and extent lists.
- Incorporate a reduced instruction set including only PUT, GET, and DELETE.
- Incorporate a native concept of policy based network locality to enable collaboration and data continuance.

WOS Efficiency

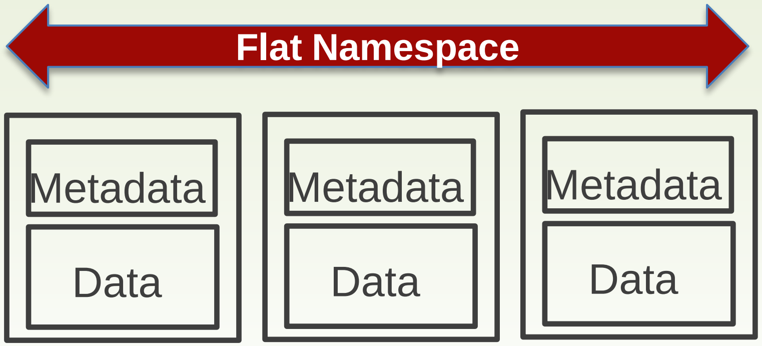
- Object storage stores data into containers, called objects
- Each object has both data and user defined and system defined metadata (a set of attributes describing the object)

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

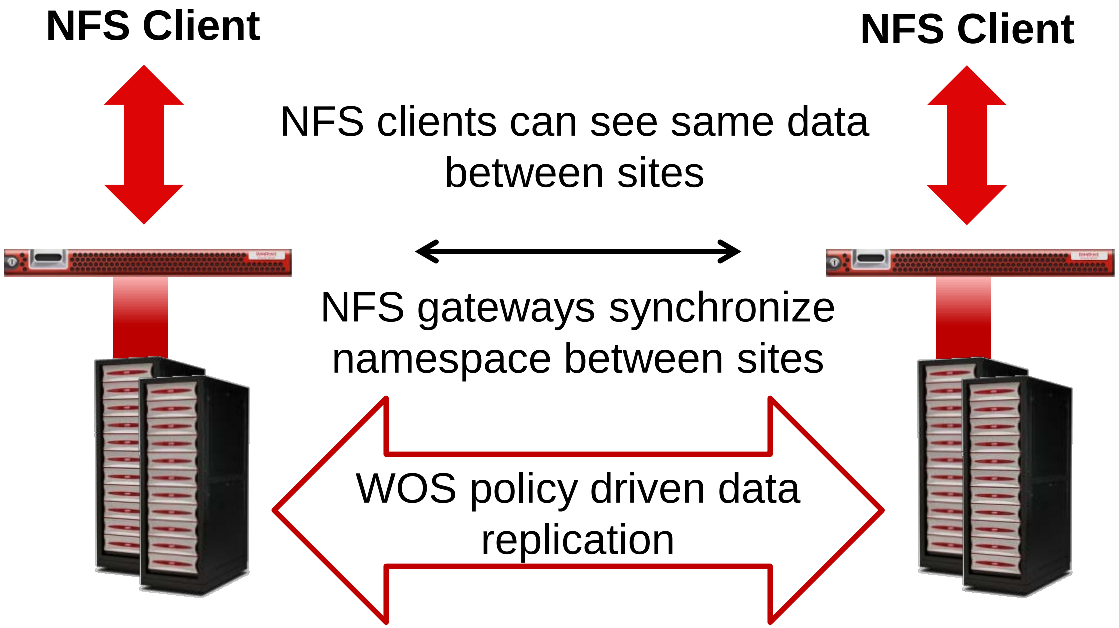
Enabling Collaboration

Access and Update Data Simultaneously Across Multiple Sites

- Multiple copies of data can be replicated globally based on policy for both disaster recovery and low latency access
- Data PUSH ensures immediate access to the latest project data
- Globally distributed users collaborate as part of a powerful workflow
- Accelerates discovery and time to market

Mirrored Data Distribution Centers

Geographically Distributed Single NFS Name Space



NFS Client

NFS Client

NFS clients can see same data between sites

NFS gateways synchronize namespace between sites

WOS policy driven data replication

- WOS Access NFS will have significant manageability, robustness and inherent load balancing advantages v. other solutions when there is
  - Immutable, unstructured data at scales > 2PB
  - Multisite access and/or disaster recovery requirements

Conclusion

- As data sets and transaction densities grow, data systems must become more efficient at every level.
- Latencies are expensive in multiple dimensions often requiring additional hardware as a “work around” to enable usable performance.
- The fundamental concept of “just enough” is not a luxury but rather a necessity for data capture, reduction, and distribution.
- Processing must be adjacent to the data.
- Service times to end users should become a portion of the overall figure of merit for any research system.

Thank You

Questions?