Adoption Trends for Solid State Technologies in Big Data Sites

Bret Weber
Executive Vice President Global Engineering Operations and CTO

May 3, 2016
IEEE MSST
History

Founded in ’98
World’s Largest Private Storage Company
Double Digit Growth, Profitable, Self Funded

Headquarters
Santa Clara and Chatsworth, CA

World-Renowned & Award Winning
Financial Services
Powering 40% of Leading Global Investment Banks

Manufacturing
Powering 30% of the World's Top Aero & Auto Manufacturers

Web, Cloud & Telco
Powering the Largest Web Scale Global Service Providers

Media
Powering over 600 HD Workflow Customers

Supercomputing
Powering Over 2/3 of the Top100®

Life Sciences
Powering Over 1/3 of the Top Sequencing Centers
Multi Dimensional Data Growth is infrastructure Landscape

It's not just about this.

MORE...

DATA

SOURCES

COLLABORATION

ANALYSIS

Burst Buffer

Active Archive

Tape

© 2016 DataDirect Networks, Inc. * Other names and brands may be claimed as the property of others. Any statements or representations around future events are subject to change.
RFI/RFP Trends & Flash
Where Flash is Growing Fastest

- NVMe Burst Buffer Layer
- SSD Accelerated Metadata
- SSD Accelerated Application / Data Set
- Lower Latency for select cloud apps
Recent Customer RFI/RFP Trends

- Over the past several years*, the % of primary storage bids with SSD has increased from 25% to 75%.

- SSD, as a percent of drives in deals, has increased over 500% in the last several years*.

- Commercial accounts are 3x more likely to include SSD in a purchase than traditional HPC accounts***.

- Commercial accounts SSD / HDD purchasing is ~3x higher than Traditional HPC***.

- Biggest recent technology inquiry** growth areas:
  - SSD
  - Object Storage
  - Active Archive
  - Cloud
    - Private
    - Hybrid

* Based on RFI Participation
** Based on DDN RFI / RFP Participation CY 2015
*** Based on DDN Internal Purchase Data 2012 - present
"We needed a storage platform that could support our needs in two areas: first, ingesting data quickly into the research environment and secondly, simplifying our ability to analyze the data..." Dr. Nidhan K. Biswas, Computational Biologist and Young Biotechnologist Awardee at NIBMG
Traditional Storage Caches Are Congested By Competing Streaming and Random Workloads, Resulting in Poor Performance In Mixed Workloads

Real-Time Adaptive Cache Technology
ReACT™

DDN’s SFA ReACT Cache Analyzes Data Composition In Real Time To:
• Write-Through Sequential Data: Avoid write mirroring penalty, parallel
• Mirror and Cache Random Data: More headroom for random, unaligned I/Os
“Simple data queries that used to take two minutes now take two seconds”  Mike Shuey
Research Infrastructure Architect, Purdue University
DDN Success Story: Accelerating Academic Research

“DDN’s SFX delivered a 900% improvement in read capability at a low cost while enabling us to access millions of small files on dedicated solid-state modules while continuing to stream very large data files simultaneously. Simple data queries that used to take two minutes now take two seconds.”

Mike Shuey, Research Infrastructure Architecture, Purdue University
What is IME?
The *Infinite Memory Engine*

- A S/W Application Accelerator which leverages NVM and SSD to remove system level performance bottlenecks
  - High bandwidth
  - Low latency (Read & Write, Large & Small, Aligned & Random)
  - Data integrity & protection
  - Massive scalability
  - No application changes required

1. POSIX compatibility for Commercial Big Data Applications
2. Solid-state cache provides line-speed performance under almost any I/O profile
3. Re-aligns I/O greatly increasing file system performance
4. API for job scheduler & application integration
DDN IME Deployment in Technical HPC Facility

IME and PFS clients deployed here (MPI-IO and POSIX)

Compute nodes can bypass IME

Persistent storage and authoritative namespace.

IME tier consists of multiple servers organized into redundancy groups (pools)

Network capacity for burst bandwidth

Network capacity for Stage-Out / Replay (IME-to-PFS) or Stage-In / Pre-Stage (PFS-to-IME) bandwidth

IME BURST BUFFER

PARALLEL FILESYSTEM

ACTIVE & DEEP ARCHIVES

COMPUTE CLUSTER
IME Examples
Test results reflect elapsed times for computing a single “shot” which includes all the computation, MPI communications and the I/O time. One-time initialization time excluded.

Speed-ups are normalized relative to Lustre-only performance.

IME approaches the performance of smaller scale in-memory runs, and benefit of IME increases with scale.

IME yields 3x full-app speed-up over Lustre alone.
End-To-End DDN Architecture

Log In Nodes
- N/W
- Burst Buffer on IME
- Home, Secure & Projects on GRIDScaler

Compute Cluster
- EDR Infiniband N/w
- Scratch on EXAScaler
- PFS Stats Collection & Monitoring
- 10 Gbps Ethernet
- Archive on WOS
- NAS Gateways

ddn.com
IME Motivation – Application IO Acceleration

**POSIX IO AND PFS BOTTLENECKS**

PFSs were not designed for today’s mixed I/O and massively parallel I/O access patterns.

**STORAGE LATENCY**

HDD seek times & network queuing effects add latency.

**FRAGMENTED I/O PATTERNS**

Mal-aligned apps slow down the PFS for all applications sharing the PFS.

**OUT OF CORE DATA**

Many datasets are too big to fit in DRAM.

---

**BURST BUFFER**

Most cost & space-efficient way to provision peak performance.

**PFS ACCELERATOR**

Mal-aligned apps slow down the PFS & entire cluster.

**APP OPTIMIZER**

Dynamically aligns mal-formed I/O into striped writes without code mods.

**CORE EXTENDER**

Unlike DRAM, no dataset is too big for IME with TBs or PBs of fast, cost effective NVM.
Thank You!

Keep in touch with us

sales@ddn.com

@ddn_limitless

company/datadirect-networks

9351 Deering Avenue
Chatsworth, CA 91311

1.800.837.2298
1.818.700.4000

© 2016 DataDirect Networks, Inc. * Other names and brands may be claimed as the property of others. Any statements or representations around future events are subject to change.