第15回PCクラスタシンポジウム
ストレージ製品とデータ・ソリューションのご紹介
クレイ・ジャパン・インク
Safe Harbor Statement

This presentation may contain forward-looking statements that are based on our current expectations. Forward looking statements may include statements about our financial guidance and expected operating results, our opportunities and future potential, our product development and new product introduction plans, our ability to expand and penetrate our addressable markets and other statements that are not historical facts. These statements are only predictions and actual results may materially vary from those projected. Please refer to Cray's documents filed with the SEC from time to time concerning factors that could affect the Company and these forward-looking statements.
クラスタシンポジウム

第15回PCクラスタシンポジウム

Supercomputing

Computation

XC40システム
CS400システム

Analytics

Urika-XA
Urika-GD

Storage & Data Management

Big Data

クレイ・アダプティブ・スーパーコンピューティング
The Cray DataWarp technology in the Trinity system will provide the first multi-petabyte multi-terabyte/sec IO burst handling capability, ever.

--Gary Grider, HPC Division Leader, LANL
Storage Trends
Convergence of Supercomputing and Big Data
Computing Memory and Storage Trends

Current Model

- CPU
- Memory (DRAM)
- Parallel Storage (HDD)
- Archive Storage (HDD & Tape)

On Node

Off Node (External)

Future Model

- CPU
- Near Memory (HBM/HMC)
- Far Memory (DRAM/NVDIMM)
- Near Storage (Flash)
- Parallel Storage (HDD)
- Archive Storage (HDD & Tape)

On Node

Off Node (Internal/HSN)

Off Node (External)
Cray Storage Solutions - Span Data Lifecycle

Sonexion Tier:
- Fast throughput
- Fully parallel access
- True scalability

TAS Tier:
- Best cost efficiency
- Data fully accessible
- Extensive scalability

Cray TAS Connector
Cray Storage Offerings

- **Cray DataWarp**
  - I/O Acceleration
- **Cray SONEXION**
  - High Performance Storage
- **Cray Tiered Adaptive Storage**
  - Efficient Long-term Storage

**Features**:
- Pure performance
- Breakthrough efficiencies
- Balanced and cohesive design
- Protect and store
- Access data forever
- Easily sustain long-term storage
- Efficiently scale
- Innovate faster
- Be confident
Summary – Optimized End-to-End Storage Solutions

- End-to-End Optimization
- Cray System Architectures
  - Aries, InfiniBand, 40GbE
- DataWarp, Sonexion, TAS
- Application I/O Optimization
- Parallel File Systems Leadership
- Scalable Networking Experts
- Best in Class Storage Systems

Performance Optimization
High Performance Flash Storage & I/O Acceleration
Cray DataWarp™

Flash Storage IO Acceleration System for Cray XC40

Performance
- Scales from 70 thousand to 40 million IOPS
- Accommodate wide range of workloads

Efficiencies
- 5x the performance of disk - same cost
- Offloads I/O intensive workloads

Cohesive
- Flexible usage models
- Automated workflows

Bridges the performance gap between compute and storage
DataWarp Benefits

- **DataWarp absorbs spikes**
  - Reduces size of underlying PFS
  - Provision PFS for sustained performance instead of peak
  - Accommodates range of applications

- **Improves efficiency**
  - Flash tier: 3-5x the performance of disk at the same cost
  - Machine efficiency improved
  - Better use of compute resources

Spikes drive up cost of storage

DataWarp absorbs spikes & reduces size of PFS
Use Case: Local Storage on Demand

Per Node Scratch

• Each compute node in a job is assigned a private part of the allocated SSD space
• Much faster than “faking it” with a parallel file system

Per Node Swap Space

• Dynamic compute node swap space
Use Case: Shared Fast /ssd

- High Bandwidth access to shared files
- Files can be striped across multiple DataWarp Nodes
- Space can be temporary for the job, or be marked as persistent to work between jobs
Use Case: Checkpoint / Restart

Fast Checkpoint / Restart

- User asks for enough SSD to cover the number of concurrently resident checkpoints
- High Bandwidth checkpoints are written to SSDs
- Followed by an asynchronous explicit or transparent copy out to rotating storage
Automated Coordination Across Storage Tiers

Scheduler Directives

Cray HPC DataWarp

Cray Sonexion Storage Lustre Parallel File System

Persisten Namespace

Automated Policies

Cray TAS Archive Storage

HDD Pool

Tape Pool

Data Loosely Coupled

Data Tightly Coupled
High Performance Primary Storage
Cray Sonexion™

Efficient, proven, scale-out Lustre system for any Linux HPC environment

Innovate Faster
- Deploy faster
- Achieve results faster

Scale Efficiently
- Sustained performance over 1 TB/s
- 25% fewer components
- Over 3 PBs usable in one 42U rack

Be Confident
- Proven Cray architectures
- Single point of support

Optimize performance and capacity at maximum density
## Efficiency Comparison – 1TB/s Example

<table>
<thead>
<tr>
<th>Deployment Comparison</th>
<th>Modular Storage with External Servers</th>
<th>Cray Sonexion (NCSA Blue Waters)</th>
<th>Monolithic Storage with External Servers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacity</td>
<td>22 PBytes</td>
<td>22 PBytes</td>
<td>32 PBytes</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>1 TByte/sec.</td>
<td>1 TByte/sec.</td>
<td>1 TByte/sec.</td>
</tr>
<tr>
<td>LNET routers</td>
<td>942</td>
<td>482</td>
<td>440</td>
</tr>
<tr>
<td>Storage units</td>
<td>472</td>
<td>180</td>
<td>360</td>
</tr>
<tr>
<td>Hard drives</td>
<td>28,320</td>
<td>15,120</td>
<td>20,160</td>
</tr>
<tr>
<td>External servers</td>
<td>942</td>
<td>0</td>
<td>294</td>
</tr>
<tr>
<td>Director IB switches</td>
<td>6</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>IB cables</td>
<td>5,468</td>
<td>482</td>
<td>1512</td>
</tr>
<tr>
<td>Racks</td>
<td>94</td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td>Cost</td>
<td>$$$</td>
<td>$</td>
<td>$</td>
</tr>
</tbody>
</table>
Simplify Management
Tiered Data Management
Cray Tiered Adaptive Storage (TAS)

Protect and access your high-value data when you need it, for as long as you need it

Protect

• Preserve assets efficiently
• Transparently migrate data

Access

• Continuously accessible
• Flexible access models

Sustain

• Break vendor lock-in
• Multigenerational data preservation

Efficient storage management with tiers
Traditional HSM for HPC – Complex

Data Ingest

IB Fabric

fs1

fs2

fs3

DM

DM

DM

Ethernet

Lustre Movers

HSM

HSM

HSM

HSM Movers

HSM

HSM

HSM

Archive Media

Archive Media

Archive Media

Disk Cache

Archive Media

Archive Media

Archive Media

QDR

FDR

FC

Ethernet
Simplified Data Management for Big Data and HPC

Data Movement and Transparent User Access

Common Access Protocols:
Lustre, NFS, SMB, HTTP, FTP

Shared Virtualized Storage Pool

IB or FC Fabric

fs1
fs2
fs3

QDR
FDR
FC
Ethernet
Protect Data over its Lifespan

- **Transparent tiering for users**
  - Data always accessible regardless of tier
  - File system appears infinitely large
  - Files always visible from the file system

- **Automated data management**
  - Policy-based data management
  - 24x7 data management
  - Multiple copies and disaster recovery

- **Works with any Lustre 2.5**
  - Cray TAS Connector for Lustre HSM
Policy-based Data Movement

- Familiar Actions & Policies
- Transparently *Archive* from disk cache to archive media
- Manage disk space or *Release* archived files from disk
- Automatically *Stage* released files back when accessed
Sustain Long-term Repositories

- **Open data format**
  - Based on POSIX TAR
  - Data is accessible without TAS
  - No vendor lock-in

- **Data protected at scale**
  - Support for 100's of PB of managed data
  - Integration with Lustre HSM

- **Future-ready technology migration**
  - Support for multigenerational data management
  - Migrate with technology
Parallel Archiving with Cray TAS

TAS Gateway

Archiver

Customer Network

Storage Networks

Customer Network

Archiver

Archiver

Archiver

Archiver
Cray TAS – Tiered Data Management Summary

Protect and access your high-value data when you need it, for as long as you need it

● **Store and Protect**
  ● Up to 5 copies—across any media
  ● Transparently migrate data across tiers, from ingest to archive

● **Access Data Forever**
  ● All data always accessible to apps and users
  ● Your choice of file protocol

● **Stay Open – Sustainable Infrastructure**
  ● Open formats break vendor lock-in
  ● Data transparently migrates across generations of storage infrastructure, disk and tape
● お問い合わせ先:
クレイ・ジャパン・インク
プリセールスチーム
礒野智之 isono@cray.com
The future is seldom the same as the past

Seymour Cray
June 4, 1995