Galaxy Community Conference
Penguin – EMC Isilon Life Science Solution Update
Overview

- About Penguin Computing & Isilon
- Penguin – Isilon Life Science Solution
- How To Engage
Penguin Computing Quick Facts

- **1997**
  - Founded and focused on custom-built Linux systems
  - Based in Fremont, CA with offices in San Francisco, Atlanta, Washington DC, Boston
- **2004**
  - Acquired Scyld Software to improve HPC ease of use
- **2010**
  - Launched Penguin-On-Demand (POD) as an HPC as a service cloud offering
- **2011**
  - Offer solutions and services for High Performance Computing, Efficient Data Center, and Cloud Services
  - Provide 24x7 support to customers in over 40 countries
  - 2500+ customers to date
  - Private equity backed
Penguin Experts In Sci. Computing
From Personal Workstations To Super Computers

- 4 processors
- 64 cores
- 512 GB RAM
- 31 TB storage
- NVIDIA Tesla GPGPU or Intel Phi
- Delivered worldwide with Next Business Day support

- 130 TFLOPS
- 13,000 AMD compute cores
- 400TB of High-speed scratch space
- 7500 meters of cable
- Liquid Cooled
- Half a Mega-Watt of Power

Ranked in the Top 50 Largest Supercomputer in the World

© Copyright 2013 EMC Corporation. All rights reserved.
Penguin Experts In Sci. Computing
To a Complete Suite of Targeted Cloud Services

Customer Examples
Cummins Engine
Suzuki Motors(NEC)
University of Memphis
Ion Torrent (Life Technologies)
Cross Commerce Media
Northrop Grumman
Scyld Cluster and Cloud Software Suite

**Insight**
- End-user focused physical or virtual cluster management GUI
- 1:1 match with Scyld CW functionality
- Hardware and workflow monitoring
- Integrated log viewer for maintenance and diagnostics
- Dynamically present cluster and job status

**Cloud Manager**
- HPC focused private and public cloud management
- User and group management
- Integrated resource management and billing
- Ties into existing CRM and financial systems
- Based on Web services
- Built in authentication services

**PODTools**
- Convenient and secure method to submit jobs and data to POD
- Automatically returns results when complete
- Customize workflow through a scriptable interface
- Generate ad hoc reports to see core hour and GB usage

**ClusterWare**
- Cluster management interface
- Minimizes effort required to set-up, update and maintain a cluster
- Guarantees consistency across all servers
- Provides complete, tested and supported cluster software stack
Penguin On Demand (POD)

• Non-virtualized, 'bare-metal' execution for optimal performance
• Low-latency QDR Infiniband and 10GbE fabric on all nodes for great application scalability
• Fast CPUs, GPUs, large memory, local scratch space, fast global storage space
• Over 60 HPC applications installed
• Flexible virtual server environment for job submission and hosted web applications
• Support from experienced HPC system administrators
• Pay-as-you-go only for core hours really used by your jobs - billing at 3-second intervals
EMC Isilon Company Overview
Setting the Standard for Scale-out NAS

- Broad adoption across many markets
  - High Performance Computing (HPC): Life Sciences, Oil & Gas, Electronic Design Automation, Media & Entertainment, Financial Services
  - Enterprise IT: Archive, Home Directories, File Shares, Virtualization, Business Analytics

- Over 4,000+ global customers
  - 2012 60% Y/Y growth rate
  - EMC #1 NAS market*

- Isilon OneFS: Seventh generation, industry-proven, innovative scale-out operating environment

EMC Isilon - Experts in Life Science Storage

- Proven
  - At 250+ life sciences organizations backed by knowledgeable SMEs

- Simple
  - One FTE can administer up to 20 PB using automation, policies and out-of-the-box security, data protection features

- Versatile:
  - Respond quickly to changes in workflow performance and capacity requirements

- Scalable:
  - Meet performance and capacity requirements across of the spectrum of scientific computing workloads
Isilon Scale-Out NAS Architecture

Servers
- Windows
- Linux
- macOS

Client/Application Layer

Ethernet Layer

OneFS Operating Environment

Hadoop

NFS
CIFS
HTTP
FTP
HDFS for Hadoop

Single FS/Volume

Intra-cluster Communication Layer
Isilon Scale-Out NAS Architecture

- Servers
  - Windows
  - Linux
  - Apple

- Client/Application Layer

- Ethernet Layer
  - NFS
  - CIFS
  - HTTP
  - FTP
  - HDFS for Hadoop

- OneFS Operating Environment
  - SmartConnect
  - SmartPools
  - SmartFail
  - AutoBalance
  - SmartQoutas
  - GNA

- Intra-cluster Communication Layer
Storage Tiers in HPC

- **Tier 1**
  - storage for immediate use by active computational processes
- **Tier 2**
  - storage for staging, result collection, raw instrumentation data
- **Tier 3**
  - storage for near line archive, inactive (lukewarm) data with potential
- **Tier 4**
  - permanent archive, cold data
- Isilon fits into Tier 2 and Tier 3 quite well
Isilon Fit Along the Spectrum (Tier 1)

Embarrassingly Parallel
- I/O operations weighted heavily toward sequential reads
- Less data written back and written back sequentially
- Read a number of data files concurrently
- Isilon excels

Fine Grained Parallelism
- Random I/O on a single (or small number) file;
- Block operations within a small number of files.
- Write their output to a single file or transfer data between processes using the I/O subsystem.
- Parallel file system perform much better here
- Consider alternatives to Isilon
Isilon Fit In Tier 2

**Pros**
- Multi-protocol supports a wide variety of users/instruments and high performance networks
- Massive scalability with a single namespace
- High speed (10GbE) networking interfaces

**Cons**
- No IB on the front end
- No RDMA for NFS
- Data movement between HPC storage tiers is cumbersome
Best Practices I/O – Classical HPC

- Tiered access to storage resources
  - Separate HPC workloads from staging and review
    - SmartPools on Isilon can do this within a cluster
    - IRODs and other rule oriented data movers can do this across clusters of storage resources

- NFSv4
  - Greatly enhanced locking functionality
    - No more stale locks
    - Byte-range locking or delegations
  - Compound operations
Best Practices I/O – Traditional HPC

• Load balancing
  – NAS is the easier way to go
    ▪ Better integrated load balancing mechanisms
    ▪ Now easily accessible over IB
  – Isilon SmartConnect balances connections over the cluster and is SmartPool “aware”
  – NAS provides better failover support...no complicated FC mappings, zoning, etc. just IP based redirection
Hybrid Cloud

• Advantages of cloud(s) (private or public)
  – Allows the creation of ad-hoc compute clusters
  – Provide dynamic expansion of existing compute or storage (burst availability)

• Challenges
  – Data locality to compute resources
  – Data transfer to expanded resources
  – Data management and flow
Hybrid Cloud – Best Practices

• Use Common S3/Swift Protocols
  − Common semantics for accessing data when dealing with hybrid resources
  − Isilon now supports Openstack Swift interface with both named access and container/object access
  − Openstack Cinder support coming for virtual block device allocation
  − Some instrumentation vendors (e.g. Illumina) support data output into Swift/S3 containers
Hybrid Cloud – Best Practices

• Provide common application frameworks
  – Gives developers a way to sandbox their app development and deploy to production without changing code
  – Many application frameworks have built in resource schedulers/workflow managers or integrate with them directly
    ▪ Penguin On Demand & Scyld Manager
Grid Best Practices

• Storage challenges are equivalent to classical HPC with the extra twist of widely distributed resources

• Use IRODS
  – Being better instrumented to work directly with Isilon
  – Several meta-schedulers (like MapSeq/UNC) are directly integrated

• Newer meta-schedulers made to make better use of grid resources and provide dynamic reconfig
  – Omero at CRS4 can reconfigure resources for Hadoop and regular resources
  – Apache Zookeeper
  – UNC MapSeq
Spectrum of Resources
Where Do I Fit In?

- Am small academic-private institute that does not have funding to host dedicated research IT
- Do not have staff skilled in setting up bioinformatics solutions
- Want to focus on my science, not administering computers and applications
- Have computing requirements that exceed my local workstation, either because I temporarily need hundreds of computing cores or a large RAM (> 128 Gigs)
- Need to maintain a long-term archive of my experimental data and results for collaboration
Penguin & Isilon At The Center Of Life Sciences Workflows
Penguin – Isilon Life Sciences Solution

- Get all the benefits of a scientific computing environment in a converged infrastructure
- Maximize focus on science while minimizing system administration
- Enabled to support ‘bursty’ workflows through Scyld & POD
- Backed by HPC experts & enterprise global support
- A pre-fab “bespoke” solution built to meet and evolve with your local requirements
Solution Spec’s

Isilon

- 4 X200 (2U) nodes, each with 48GB of RAM, 2x10GbE ports
- 2 QLogic 1U Infiniband switches for Isilon backend
- 2 1U 10GbE switches

Penguin

- 2 1U 10GbE Artica Switches
- 13 1U compute nodes
  - Relion 1800i, dual Intel Xeon E5-2603v2, 64GB of RAM and 1TB DAS
  - 1U compute nodes for as Galaxy host/head node/torque mom
  - 12 U for compute jobs.
How to Engage
Contact

• Penguin Computing
  – David Ingersoll, VP Business Development
  – dingersoll@penguincomputing.com

• EMC Isilon
  – Sasha.paegle@isilon.com
## Arctica 4804x: 48x SFP+ switch with 4x QSFP Uplinks

<table>
<thead>
<tr>
<th>Brand name</th>
<th>Penguin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model name</td>
<td>Arctica 4804x</td>
</tr>
<tr>
<td>Port Configuration</td>
<td>48x 10G(SFP+) 4x 40G (QSFP)</td>
</tr>
<tr>
<td>Switching Capacity</td>
<td>1.28Tbps</td>
</tr>
<tr>
<td>Maximum forwarding rate</td>
<td>960mpps</td>
</tr>
<tr>
<td>Dimension</td>
<td>1RU, 15.5” deep</td>
</tr>
<tr>
<td>Airflow</td>
<td>F2B or B2F</td>
</tr>
<tr>
<td>Redundant power?</td>
<td>Yes</td>
</tr>
<tr>
<td>AC/DC Option</td>
<td>AC or DC</td>
</tr>
<tr>
<td>N+1 hot swap fans</td>
<td>3+1 fixed</td>
</tr>
<tr>
<td>In/Out band mgmt?</td>
<td>Yes</td>
</tr>
<tr>
<td>MAC</td>
<td>128K</td>
</tr>
<tr>
<td>Packet buffer</td>
<td>9MB</td>
</tr>
</tbody>
</table>
IB to IP Switching

• Many HPC customers running dedicated IB networks
• Isilon evaluated and tested IB to IP gateway products from Mellanox
  • These connect the 10GbE frontend and to the Infiniband network.
  • The switch supports proxy-arp between the two networks
  • IPoIB drivers and network setup required on client hosts. No change on Isilon
• Using NFSv3 and 4 with TCP on ipoib we can consistently achieve >95% of the performance of 10GbE
Lustre HSM on Isilon

- Lustre very common as Tier 1 storage
- Isilon working with multiple Lustre HSM solutions
  - Isilon clusters become an archive target for the HSM
  - Intel Lustre allows a degree of mirroring to take place (archive+)
- PLFS based HSM target also available