

A photograph of solar panels under a blue sky with clouds, viewed from a low angle looking up. The panels are blue with a grid pattern.

# Lustre User Group 2009

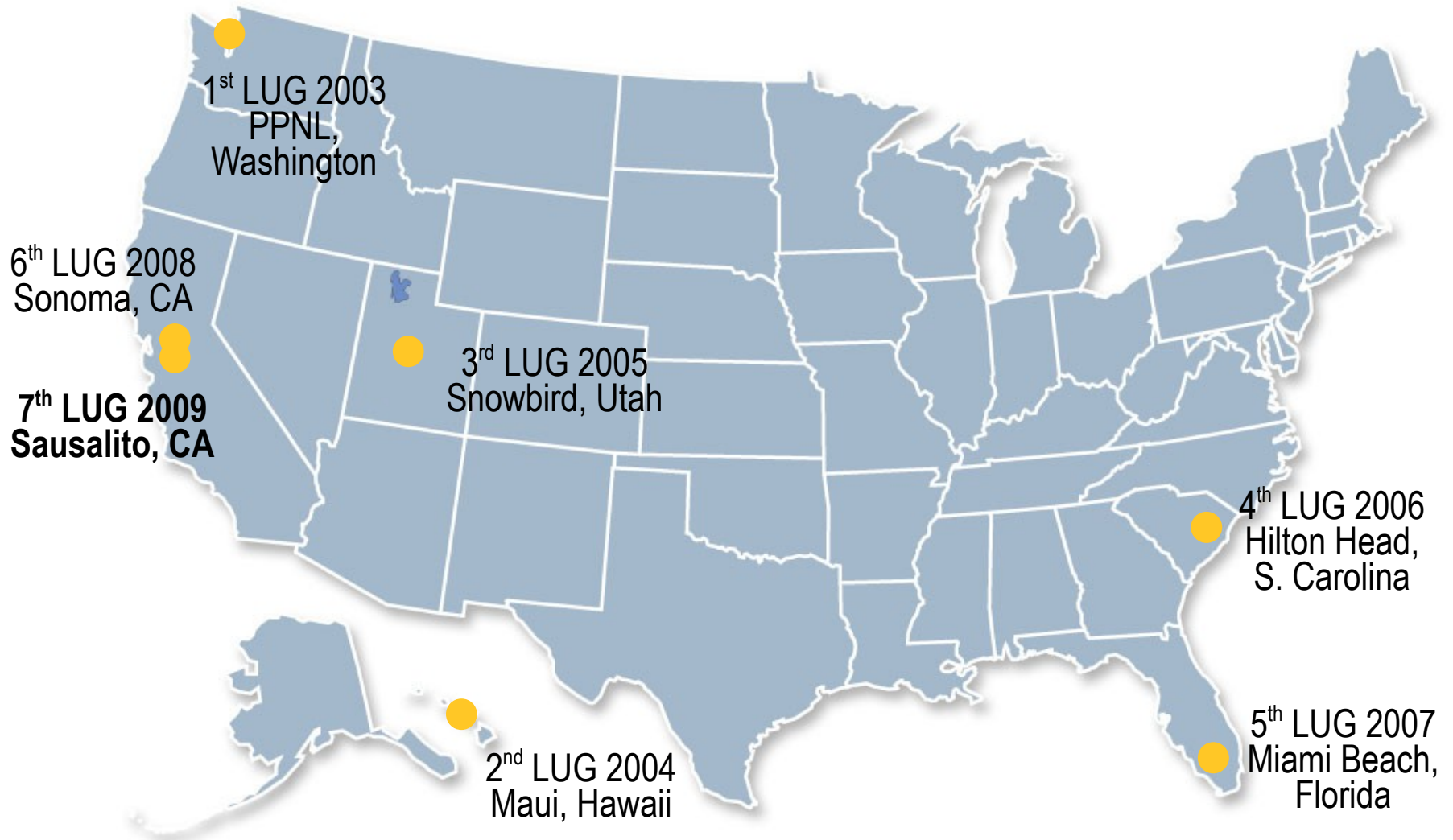
Keynote Presentation

**Peter Bojanic**  
**Sun Microsystems**

# Welcome to Cavallo Point Lodge, Sausalito California







# LUG Participation

Year	Location	Total Registrants	Organizations
2007	Miami Beach, FL	68	40
2008	Sonoma, CA	122	43
<b>2009</b>	<b>Sausalito, CA</b>	<b>125</b>	<b>43</b>

# Thank You

- Sponsors



**DataDirect**  
NETWORKS

- John Fowler – Fireside chat this evening
- Co-coordinators – Dan Ferber & Rich Brueckner
- Support team – Lindsey Stack & Rosie Krupski
- All of our presenters & facilitators

# New Lustre.org Wiki

The screenshot shows a browser window titled "Main Page" with the URL [http://wiki.lustre.org/index.php/Main\\_Page](http://wiki.lustre.org/index.php/Main_Page). The browser's address bar includes a search engine (Google) and a search input field containing "Search Lustre.org". The page header features the Lustre logo on the left and navigation links for "Log In / Create Account" and "Help" on the right. Below the header is a blue navigation bar with the Sun Microsystems logo and a search bar.

**High Performance and Scalability**

For the world's largest and most complex computing environments, the Lustre™ file system redefines high performance, scaling to tens of thousands of nodes and petabytes of storage with groundbreaking I/O and metadata throughput.

[More on Lustre performance, service, and support at \[sun.com/lustre\]\(http://sun.com/lustre\)](#)

**What's New**

**Upcoming Release of Lustre 1.8**

Lustre 1.8 is in the final cycles of release testing and expected to GA in April 2008. Lustre 1.8 will introduce several robust, new features including:

- Adaptive Timeouts
- OSS Read Cache
- OST Pools
- Version-based Recovery

Read more about 1.8 features and why you should upgrade.

**Lustre User Group - April 16-17, 2009**

LUG is the premier event to learn about Lustre technology, acquire best practices, and share knowledge with other users. Don't miss this once-a-

**Download**  
Download this open source software for free.  
» [Get Lustre](#)

**Learn**  
Find out about Lustre interoperability, features and publications.  
» [About Lustre](#) » [Publications](#)

**Use**  
User resources to make Lustre perform at its best.  
» [Install and Configure](#) » [Administer](#)

**Contribute**  
Developer resources and tools to contribute to Lustre.  
» [Code](#) » [Test](#) » [Pre-Release Versions](#)

**Get Involved**  
Find groups and community developer projects.  
» [Community](#) » [Lustre Centres of Excellence](#)

# Lustre Support Matrix

The screenshot shows the 'Lustre Releases' page on the wiki. The page title is 'Lustre Releases' and the URL is 'http://wiki.lustre.org/index.php/Learn:Lustre\_Releases'. The page features a navigation menu with links like 'Main Page', 'Download', 'Learn', 'Use', 'Contribute', and 'Get Involved'. A 'Wiki Toolbox' is also present with links for 'What Links Here', 'Related Changes', 'Special Pages', 'Printable Version', and 'Permanent Link'. The main content area is titled 'Lustre Releases' and contains the 'Lustre Support Matrix' section. This section includes a paragraph explaining that the matrix lists supported networks, kernels, and e2fsprogs versions, and a note about detailed information for Lustre 1.6.x releases. The support matrix table is as follows:

Supported Kernels	Lustre Version				
	1.8.0	1.6.7.1	1.6.6	1.6.5.1	1.6.4.3
<b>RHEL 5</b>					
2.6.18-92.1.17.el5		X			
2.6.18-92.1.10.el5			X		
2.6.18-53.1.14.el5				X	
2.6.18-53.1.13.el5					X
<b>SLES 10</b>					
2.6.16.60-0.31		X			
2.6.16.60-0.27			X		
2.6.16.54-0.2.5				X	X

# Vibrant OEM Ecosystem

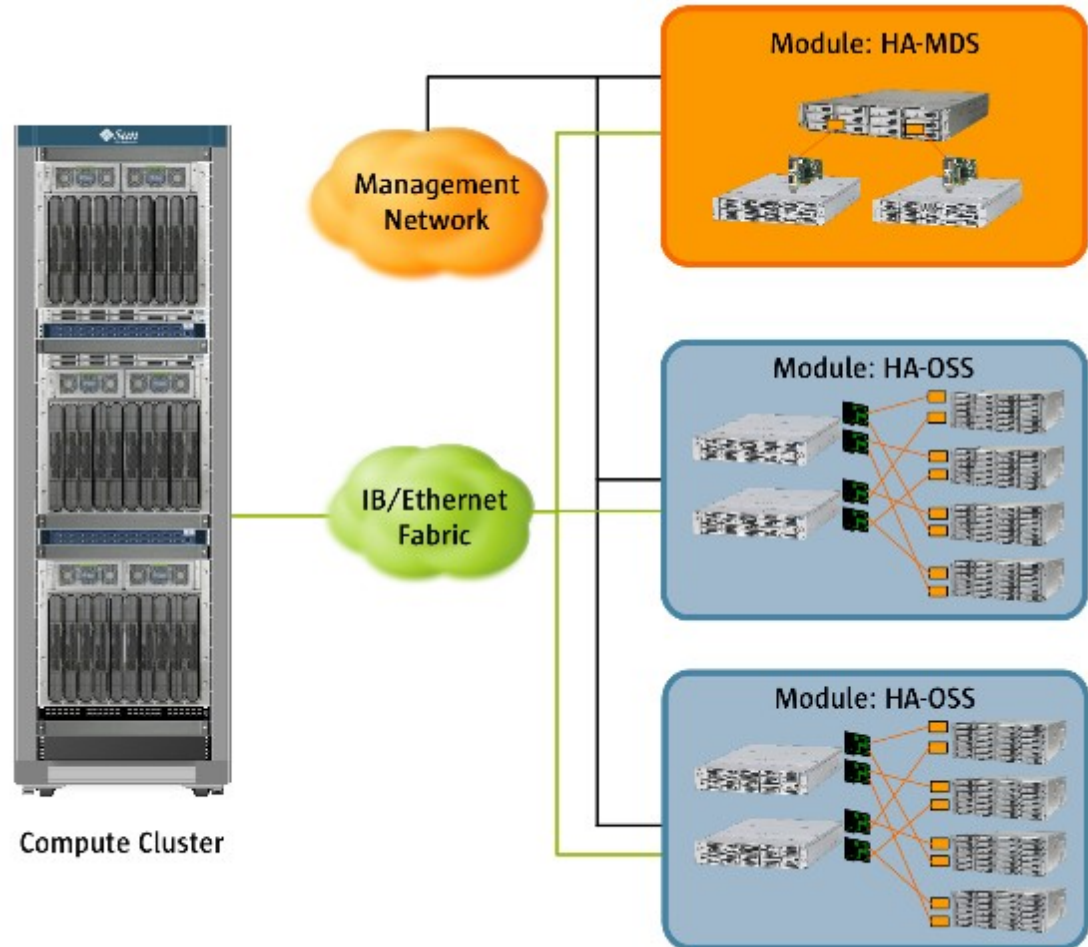


More to come in 2009...



# Sun Lustre Storage System

- Fully integrated Lustre storage solution based on proven reference architecture
- Offers compelling price/performance value through Sun Open Storage products
- Optimized software stack with automated install and configuration tools
- Sun Professional Services and Support to ensure success



# Professional Services & Support

Sun  
expertise  
for your

**SUCCESS**

## Enterprise Installation Service

- HPC and storage installation services

## Lustre Implementation Services

- Highly flexible, modular design
- Design, implementation and information transfer

## Sun Spectrum Hardware Services Plans

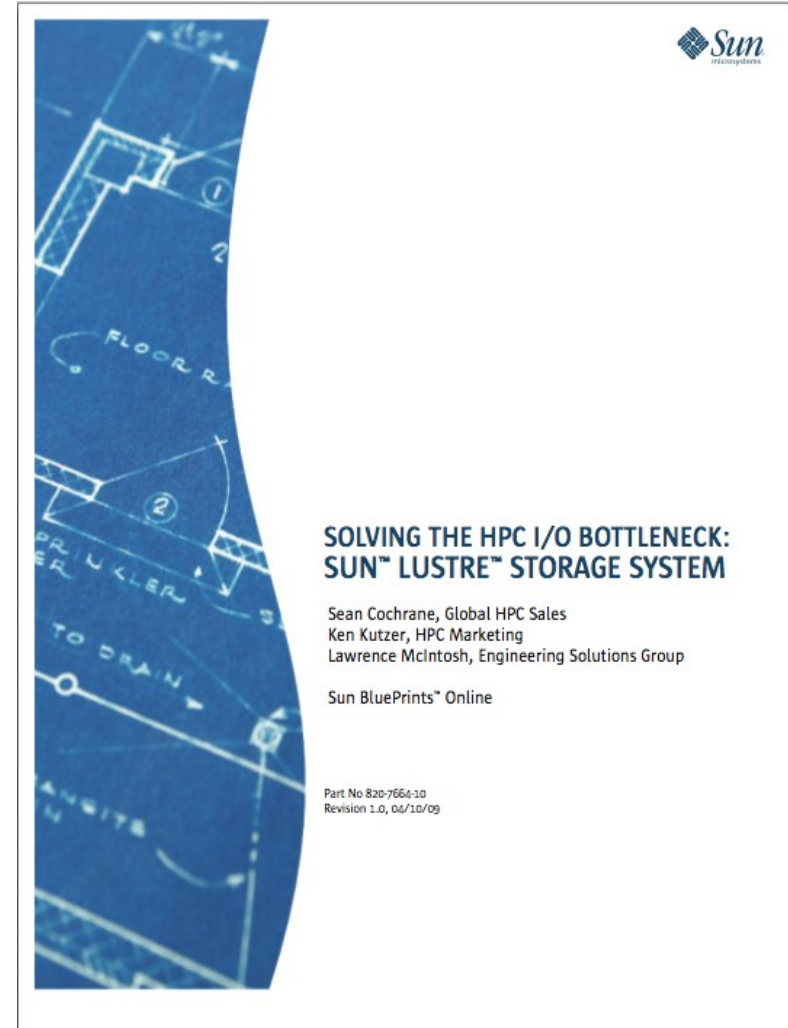
- Dependable maintenance coverage and technical assistance


## Sun Spectrum Software Services Plans

- Online technical support, updates (patches), and knowledgeable

# More About Lustre Storage System

- Lustre on Sun.com  
<http://sun.com/lustre>
- Sun Lustre Storage System on Sun.com  
<http://sun.com/scalablestorage>
- White paper – Solving the I/O Bottleneck: Sun Lustre Storage Solution  
<http://www.sun.com/offers/details/820-7664.html>



Sun  
microsystems

**SOLVING THE HPC I/O BOTTLENECK:  
SUN™ LUSTRE™ STORAGE SYSTEM**

Sean Cochrane, Global HPC Sales  
Ken Kutzer, HPC Marketing  
Lawrence McIntosh, Engineering Solutions Group

Sun BluePrints™ Online

Part No 820-7664-10  
Revision 1.0, 04/10/09

# Lustre Roadmap Update



# Accomplishments in 2008

- 7 of top 10 Supercomputers in the World
  - > Almost all customers are now running Lustre 1.6
- New IO performance records
  - > 190 GB/s on ORNL Jaguar
- Raising the quality bar
  - > Substantially greater scale testing (Hyperion, etc.)
  - > Community outreach on test plans
- Advancement on development goals
  - > Hundreds of bugs fixes and enhancements to HEAD
  - > Lustre 2.0 alpha release
  - > Lustre DMU server pre-alpha release

# Challenges with Lustre 2.0 & CMD

## HEAD to Good

- The Hendrix project development branch diverged from the Lustre production release branch for 18 months
- In the past 18 months we've ported or rewrote several hundred bug fixes and minor improvements back to HEAD

## Recovery

- CMD implemented only the minimal set of recovery use cases required to satisfy the Hendrix acceptance tests
- There remained considerable architecture and implementation work to complete full recovery

## Raising the Quality Bar

- Continually raising standards for product quality assurance testing
- Large scale testing and testing on Cray equipment are now part of the standard quality practice
- We have farther to reach with HEAD than we originally measured and anticipated

# Lustre Releases Status

- 2.0 Alpha
  - > Available for download now  
[http://downloads.lustre.org/public/lustre/Lustre\\_2.0\\_Alpha/](http://downloads.lustre.org/public/lustre/Lustre_2.0_Alpha/)
  - > Acceptance small has run; IOR and Simul runs up to 74 clients
  - > CMD and Quotas are unqualified for the alpha
- 1.8.0
  - > Failover issues found by HP in RC5 testing
  - > We'll release as soon as these are fixed
- 1.6.7.1
  - > Directory corruption fix for 18695 now available for download

# Call to Action: Lustre 2.0 Testing

- Intermediate Release Milestones Until GA
  - > New builds every 4-6 weeks; next milestone is May 7
  - > Builds will be available for download via FTP
  - > Release milestone criteria and test results will be published on [lustre.org](http://lustre.org)
- We need ample community testing and feedback
  - > Please install, test and provide feedback
  - > Post defects in Bugzilla and comments/suggestions to [lustre-discuss](http://lustre-discuss)
  - > We are recruiting beta test sites for the fall (talk to Jessica Popp for details)



# Lustre Release Plan

Lustre 1.6 Supported  
until end April, 2010

April 2009	Q4 2009	2010	2011+
<p><b>Lustre 1.8.0</b></p> <ul style="list-style-type: none"> <li>• OST Pools</li> <li>• OSS Read Cache</li> <li>• Adaptive Timeouts</li> <li>• Version based recovery (VBR)</li> </ul> <p><b>Lustre 1.8.x</b></p> <ul style="list-style-type: none"> <li>• Simplified Interoperation with 2.x</li> </ul>	<p><b>Lustre 2.0.0</b></p> <ul style="list-style-type: none"> <li>• Server and Client Restructure for CMD and ZFS</li> <li>• Clustered MetaData Early Evaluation (No Recovery)</li> <li>• Security GSS Early Evaluation</li> <li>• Server Change Logs</li> <li>• Commit on Share</li> <li>• MDS Performance Enhancements</li> </ul>	<p><b>Lustre 2.x Release(s)</b></p> <ul style="list-style-type: none"> <li>• ZFS Lustre GA</li> <li>• Improved SMP Scaling</li> <li>• Clustered MetaData Early Eval w/Recovery</li> <li>• Size on MDS</li> <li>• Imperative Recovery</li> </ul>	<p><b>Lustre 3.0.0</b></p> <ul style="list-style-type: none"> <li>• Clustered MetaData GA</li> <li>• Beginning of Other HPCS Enhancements</li> </ul> <p><b>Lustre 3.x Release(s)</b></p> <ul style="list-style-type: none"> <li>• Online Data Migration</li> <li>• Write Back Cache</li> <li>• Proxies</li> </ul>
<p>RHEL 5, SLES 10 RHEL 6 in 1.8.x after RHEL 6 GA SLES 11 in 1.8.x</p>	<p>RHEL 5 &amp; 6, SLES 10 &amp; 11</p>	<p><b>Release in 2.x or 3.x Depending on Readiness</b></p> <ul style="list-style-type: none"> <li>• HSM/HPSS</li> <li>• HSM/SAM-QFS</li> <li>• Windows Native Client</li> <li>• Security GA</li> <li>• Network Request Scheduler</li> <li>• pNFS Exports</li> <li>• Scalable health monitoring</li> </ul>	
<p>1.4 EOL is June 2009 1.6 EOL is 12 months after 1.8 GA</p>			

# EOL for Current Lustre Releases

- Lustre 1.4
  - > Support for 1.4.x will officially cease on June 30, 2009; no further 1.4.x maintenance releases
- Lustre 1.6
  - > Regular quarterly 1.6.x maintenance releases will cease with Lustre 1.8.0; ad-hoc 1.6.7.x releases for critical issues only (e.g. 1.6.7.1)
  - > Support for 1.6.x will officially cease one year after 1.8.0 (ends April, 2010)
- Lustre 1.8
  - > This becomes our active production code line until Lustre 2.0.0
  - > We'll deliver 1.8.x maintenance releases on a quarterly basis (bug fixes, minor performance improvements)

# Metadata Performance Project

- Joint Sun/Cray/ORNL project team launched to improve single metadata server performance
- Mission
  - > Create a suite of metadata performance tests
  - > Benchmark baseline performance metrics
  - > Identify and implement improvements
  - > Measure and quantify incremental improvements
- Release vehicles
  - > Lustre 1.8 – modest performance improvements with minimal risk (RPC reduction; async operations; LNET SMP scaling)
  - > Lustre 2.0 – More ambitious improvements specific to the new metadata code base on the HEAD branch (Size on MDS)
  - > Lustre 3.0 – CMD, Metadata Writeback Cache

# Top Lustre Group Goals for 2009

- Maintain Lustre lead in performance and installed base for "extreme" HPC
- Get the Lustre HEAD code base into ship shape
- Integrate Lustre with ZFS/DMU (kernel space)
- Expand support, service, and OEM partnerships around the world



# Lustre and the Future

- Continued focus on extreme HPC
- Capacity
  - > Exabytes of storage
  - > Trillions of files
  - > Many client clusters each with 100,000's of clients
- Performance
  - > TB's/sec of aggregate I/O
  - > 100,000's of aggregate metadata ops/sec
- Community Driven Tools and Interfaces
  - > Management and Performance Analysis



# Questions?

A close-up photograph of water splashing, with white foam and blue water, occupying the top half of the slide.

# THANK YOU

[pbojanic@sun.com](mailto:pbojanic@sun.com)