Early Experience with Red Storm

April 19, 2005
Sue Kelly and Bob Ballance
Sandia National Laboratories
smkelly@sandia.gov, 505-845-9770
Outline of Talk

• Overview of Red Storm
• Where we are
• How we got here
• Where we’re going
• What we learned
Red Storm is a Massively Parallel Processor

Service  Compute Partition  Parallel I/O

Users  /home

Net I/O
27×16×24 3D Mesh/Torus + I/O

Torus Interconnect in Z

10,368 Compute Node Mesh

X=27
Y=16
Z=24

256 Service & I/O Nodes

256 Service & I/O Nodes
Usage Model

Unix (Linux)
Login (Service)
Node
with Unix environment

Compute Resource
I/O

Net I/O
Space Sharing of Jobs

• Jobs occupy disjoint regions simultaneously
• Example – red, green, and blue jobs:

10,368 Compute Nodes

X=27

Y=16

Z=24
Red Storm Layout

(27 X 16 X 24 mesh)

Normally Classified
Switchable Nodes
Normally Unclassified

I/O and Service Nodes
Disconnected Cabinets

Disk storage system not shown

Sandia National Laboratories
Red Storm Design Goals

Balanced System Performance  -  CPU, Memory, Interconnect, and I/O.

Usability  -  Functionality of hardware and software meets needs of users for Massively Parallel Computing.

Scalability  -  System Hardware and Software scale, single cabinet system to ~30,000 processor system.

Reliability  -  Machine stays up long enough between interrupts to make real progress on completing application run (at least 50 hours MTBI), requires full system RAS capability.

Upgradability  -  System can be upgraded with a processor swap and additional cabinets to 100T or greater.

Red/Black Switching  -  Capability to switch major portions of the machine between classified and unclassified computing environments.

Space, Power, Cooling  -  High density, low power system.

Price/Performance  -  Excellent performance per dollar, use high volume commodity parts where feasible.
Red Storm Design Parameters

True MPP, designed to be a single system.

Fully connected high performance 3-D mesh interconnect.

Topology  -  27 X 16 X 24 compute nodes and 2 X 8 X 16 service and I/O nodes

108 compute node cabinets and 10,368 compute node processors. (AMD Sledgehammer @ 2.0 GHz)

~10 TB of DDR memory @ 333 MHz (1.0 GB per processor)

Red/Black switching   -   ~1/4, ~1/2, ~1/4.

8 Service and I/O cabinets on each end (256 processors for each color)

240 TB of disk storage (120 TB per color).
Red Storm Design Parameters

Functional hardware partitioning   -   service and I/O nodes, compute nodes, and RAS nodes.

Functional system software partitioning   -   LINUX on service and I/O nodes, LWK (Catamount) on compute nodes, stripped down LINUX on RAS nodes.

Separate RAS and system management network (Ethernet).

Router table based routing in the interconnect.

Less than 2 MW total power and cooling.

Less than 3,000 square feet of floor space.
Red Storm RAS System

RAS Workstations

Separate and redundant RAS workstations for Red and Black ends of machine.
System administration and monitoring interface.
Error logging and monitoring for major system components including processors, memory, NIC/Router, power supplies, fans, disk controllers, and disks.

RAS Network
- Dedicated Ethernet network for connecting RAS nodes to RAS workstations.

RAS Nodes
- One for each compute board
- One for each cabinet
Red Storm Performance

Peak of \( \sim 40 \) TF

Expected MP-Linpack performance \( >20 \) TF

Aggregate system memory bandwidth \( \sim 55 \) TB/s

Interconnect

- Aggregate sustained interconnect bandwidth \( > 100 \) TB/s
- MPI Latency \( 2 \) \( \mu \)s neighbor, \( 5 \) \( \mu \)s across machine
- Bi-Section bandwidth \( \sim 2.3 \) TB/s
- Link bandwidth \( \sim 3.0 \) GB/s in each direction

Disk and External Network I/O

- Sustained 50 GB/s each color parallel disk I/O
- Sustained 25 GB/s each color external network I/O
## Comparison of ASCI Red and Red Storm

<table>
<thead>
<tr>
<th></th>
<th>ASCI Red</th>
<th>Red Storm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full System Operational Time Frame</td>
<td>June 1997 (Processor and Memory Upgrade in 1999)</td>
<td>August 2004</td>
</tr>
<tr>
<td>Theoretical Peak (TF)</td>
<td>3.15</td>
<td>41.47</td>
</tr>
<tr>
<td>MP-1.inpack Performance (TF)</td>
<td>2.379</td>
<td>&gt;20 (est)</td>
</tr>
<tr>
<td>Architecture</td>
<td>Distributed Memory MIMD</td>
<td>Distributed Memory MIMD</td>
</tr>
<tr>
<td>Number of Compute Node Processors</td>
<td>9,460</td>
<td>10,368</td>
</tr>
<tr>
<td>Processor</td>
<td>Intel P II @ 333 MHz</td>
<td>AMD Opteron @ 2.0 GHz</td>
</tr>
<tr>
<td>Total Memory</td>
<td>1.2 TB</td>
<td>10.4 TB (up to 80 TB)</td>
</tr>
<tr>
<td>System Memory B/W</td>
<td>2.5 TB/s</td>
<td>55 TB/s</td>
</tr>
<tr>
<td>Disk Storage</td>
<td>12.5 TB</td>
<td>240 TB</td>
</tr>
<tr>
<td>Parallel File System B/W</td>
<td>1.0 GB/s each color</td>
<td>50.0 GB/s each color</td>
</tr>
<tr>
<td>External Network B/W</td>
<td>0.2 GB/s each color</td>
<td>25 GB/s each color</td>
</tr>
<tr>
<td>Interconnect Topology</td>
<td>3-D Mesh (x, y, z) 38 X 32 X 2</td>
<td>3-D Mesh (x, y, z) 27 X 16 X 24</td>
</tr>
</tbody>
</table>
# Comparison of ASCI Red and Red Storm

<table>
<thead>
<tr>
<th></th>
<th>ASCI Red</th>
<th>Red Storm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interconnect Performance</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MPI Latency</td>
<td>15 μs 1 hop, 20 μs max</td>
<td>2.0 μs 1 hop, 5 μs max</td>
</tr>
<tr>
<td>Bi-Directional Link B/W</td>
<td>800 MB/s</td>
<td>6.0 GB/s</td>
</tr>
<tr>
<td>Minimum Bi-section B/W</td>
<td>51.2 GB/s</td>
<td>2.3 TB/s</td>
</tr>
<tr>
<td><strong>Full System RAS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAS Network</td>
<td>10 Mbit Ethernet</td>
<td>100 Mbit Ethernet</td>
</tr>
<tr>
<td>RAS Processors</td>
<td>1 for each 32 CPUs</td>
<td>1 for each 4 CPUs</td>
</tr>
<tr>
<td><strong>Operating System</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compute Nodes</td>
<td>Cougar</td>
<td>Catamount (Cougar)</td>
</tr>
<tr>
<td>Service and I/O Nodes</td>
<td>TOS (OSF1)</td>
<td>LINUX</td>
</tr>
<tr>
<td>RAS Nodes</td>
<td>VX-Works</td>
<td>LINUX</td>
</tr>
<tr>
<td><strong>Red Black Switching</strong></td>
<td>2260 - 4940 - 2260</td>
<td>2688 - 4992 - 2688</td>
</tr>
<tr>
<td><strong>System Foot Print</strong></td>
<td>~2500 sq ft</td>
<td>~ 3000 sq ft</td>
</tr>
<tr>
<td><strong>Power Requirement</strong></td>
<td>850 KW</td>
<td>1.7 MW</td>
</tr>
</tbody>
</table>
Red Storm System Software

**Operating Systems**
- Linux on service and I/O nodes
- Catamount LWK on compute nodes
- Linux on RAS monitors

**Run-Time System**
- Logarithmic Job launch (yod)
- Node allocator (CPA)
- Batch system – PBS Pro

**File Systems**
- Lustre

**User Environment**
- PGI Compilers - Fortran, C, C++
- Libraries - MPI, I/O, Math, MPI-2
- Showmesh
- Debugger - TotalView
- Performance Monitoring

**Network**
- 50 x 10GigE to each RoSE cluster
- 10 x 1GigE to login nodes
- 1 GigE to Mgmt Workstations

**System Mgmt and Admin**
- Accounting
- Red Storm Management System
Outline of Talk

• Overview of Red Storm
• Where we are
• How we got here
• Where we’re going
• What we learned
3/31 Milestone Achievements

1. All Red Storm hardware is on site, powered up, and integrated into the system.
2. We have booted 60 (3x20) compute cabinets as a single system and the remainder as two smaller systems. (Now booting as two systems total: 3x27 and 1x27.)
3. We have run all 7x applications.
4. We ran multiple applications simultaneously and demonstrated space sharing of the machine.
5. We have demonstrated sufficient stability of the platform and the software to verify normal operation of each application.
6. We used HPL (High Performance Linpack) to obtain a first look at the performance of the system.

Milestone #30: Bring Red Storm to initial operation and run the 7x applications
We have run all 7x applications using a 7x dataset and normal I/O

<table>
<thead>
<tr>
<th>Application</th>
<th>Problem Set</th>
<th>Processor Count</th>
<th>Run Time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTH</td>
<td>shape charge</td>
<td>2500 (5120)</td>
<td>60 (11)</td>
</tr>
<tr>
<td>Partisn</td>
<td>sntiming</td>
<td>2575</td>
<td>51</td>
</tr>
<tr>
<td>Partisn</td>
<td>sntiming 5x5</td>
<td>4096</td>
<td>Not reported</td>
</tr>
<tr>
<td>sPPM</td>
<td>inputdeck</td>
<td>2550</td>
<td>81</td>
</tr>
<tr>
<td>sPPM</td>
<td>inputdeck</td>
<td>3600</td>
<td>7</td>
</tr>
<tr>
<td>ITS</td>
<td>starsat</td>
<td>2370 (3500)</td>
<td>51 (2)</td>
</tr>
<tr>
<td>Sage</td>
<td>timing_c</td>
<td>2575</td>
<td>6</td>
</tr>
<tr>
<td>UMT2k</td>
<td>3d_slit</td>
<td>2575 (2799)</td>
<td>19 (1)</td>
</tr>
<tr>
<td>Salinas</td>
<td>Cube</td>
<td>2197</td>
<td>54</td>
</tr>
<tr>
<td>Alegra</td>
<td>ic_scaling</td>
<td>2048</td>
<td>205</td>
</tr>
<tr>
<td>Presto</td>
<td>walls</td>
<td>1024 (1536)</td>
<td>60 (30)</td>
</tr>
<tr>
<td>Calore</td>
<td>teratest</td>
<td>512 (1536)</td>
<td>119 (130)</td>
</tr>
</tbody>
</table>

Numbers in parentheses are runs with higher processor count, but with limited I/O or non 7x dataset
Stability Metrics

Runs Versus Uptime
(Data from 3/1/05 - 3/26/05 on Row 0 System)

Points near origin are early attempts at larger configurations
Outline of Talk

• Overview of Red Storm
• Where we are
• How we got here
• Where we’re going
• What we learned
Condensed History of Red Storm Development

- Cray/Sandia partnership formed in July, 2002 with a letter contract
- Hardware arrived between September 2004 and February 2005
- Even though hardware had arrived, software was immature and had little exposure to real hardware
  - In particular RAS and networking software were deficient
- System was divided up into ~15 smaller systems for various development and risk mitigation efforts
- Many, many long days, nights, and weekends got us to 3/31
Sample Application Scaling Progress
From Row 0 Runs

<table>
<thead>
<tr>
<th>Date</th>
<th>Boot</th>
<th>HPL</th>
<th>sPPM</th>
<th>CTH</th>
<th>ITS</th>
<th>Partisn</th>
<th>Sage</th>
<th>UMT2k</th>
<th>Salinas</th>
<th>Alegra</th>
<th>Calore</th>
<th>Presto</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/21/2005 Sandia</td>
<td>44</td>
<td>16</td>
<td>44</td>
<td>16</td>
<td>44</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/28/2005 Sandia</td>
<td>96</td>
<td>64</td>
<td>24</td>
<td>64</td>
<td>16</td>
<td>44</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/04/2005 Sandia</td>
<td>96</td>
<td>70</td>
<td>24</td>
<td>70</td>
<td>7</td>
<td>48</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/11/2005 Sandia</td>
<td>448</td>
<td>188</td>
<td>24</td>
<td>188</td>
<td>188</td>
<td>48</td>
<td>5</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/18/2005 Sandia</td>
<td>960</td>
<td>188</td>
<td>92</td>
<td>256</td>
<td>200</td>
<td>252</td>
<td>80</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/25/2005 Sandia</td>
<td>1344</td>
<td>284</td>
<td>560</td>
<td>564</td>
<td>564</td>
<td>564</td>
<td>188</td>
<td>188</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/04/2005 Sandia</td>
<td>1344</td>
<td>284</td>
<td>560</td>
<td>1024</td>
<td>564</td>
<td>564</td>
<td>188</td>
<td>284</td>
<td>0</td>
<td>8</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>3/11/2005 Sandia</td>
<td>1915</td>
<td>1024</td>
<td>1300</td>
<td>1024</td>
<td>1024</td>
<td>564</td>
<td>1300</td>
<td>1331</td>
<td>64</td>
<td>8</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>3/18/2005 Sandia</td>
<td>2562</td>
<td>1900</td>
<td>1872</td>
<td>1900</td>
<td>1872</td>
<td>1900</td>
<td>1900</td>
<td>1900</td>
<td>1000</td>
<td>512</td>
<td>1536</td>
<td>1536</td>
</tr>
<tr>
<td>3/24/2005 Sandia</td>
<td>2582</td>
<td>2562</td>
<td>2375</td>
<td>2562</td>
<td>2560</td>
<td>2575</td>
<td>2575</td>
<td>2575</td>
<td>2197</td>
<td>1024</td>
<td>1536</td>
<td>1536</td>
</tr>
<tr>
<td>3/31/2005 Sandia</td>
<td>2582</td>
<td>2562</td>
<td>2560</td>
<td>2562</td>
<td>2560</td>
<td>2575</td>
<td>2575</td>
<td>2575</td>
<td>2197</td>
<td>2048</td>
<td>1536</td>
<td>1536</td>
</tr>
</tbody>
</table>
Latency performance over time

Mpilatency Test Results - Zero Byte Message

- OS based
- Firmware based

Axes:
- X-axis: Dates (11/25/04, 12/15/04, 1/4/05, 1/24/05, 2/13/05, 3/5/05, 3/25/05)
- Y-axis: Microseconds (0, 5, 10, 15, 20, 25, 30)

Graph showing the performance over time with different markers representing OS based and Firmware based tests.
Outline of Talk

• Overview of Red Storm
• Where we are
• How we got here
• Where we’re going
• What we learned
Upcoming Milestones

• Begin classified operation in June for LANL runs only
• Complete 3 ASC milestones by September 05, which will result in a production-ready system for a limited set of users
Short Term Plans

• Service partitions need to move beyond:
  – Minimal number of nodes
  – Minimal configurations
  – Not operating as intended
  – PBS/Job batch policies not installed
  – No normal login nodes and login procedures
  – Very little Sandia network integration
• Lustre needs more testing
• Need more reliability testing
Outline of Talk

• Overview of Red Storm
• Where we are
• How we got here
• Where we’re going
• What we learned
Lessons Learned

- Seek first to emulate
  - Learn from the past
  - Simulate the future
- Successful technology transitions require people transformations
  - Need technology philosophers
    - Tilt Meters
    - Historians
- Even Tiger Woods has a coach
- The big bang only worked once
  - Deploy test platforms early and often
- Build de-scalable, scalable systems
  - Don’t forget that you have to get it running first!
  - Leave the support structures (even non-scalable development tools) in working condition, you’ll need to debug some day
- Only dead systems never change
  - Nobody ever built just one system even when successfully deploying just one system
  - Nothing is ever done just once
- End to end arguments apply within large systems AND teams
- Build scaffolding that meets the structure
  - Is build and test infrastructure in place FIRST?
  - Will it effectively support both the team and the project?
Conclusion

• Red Storm is on target to be the 7X Son of Red

• Early experience shows good scalability

• We have much to do in the next 5 months