Early Experiences on Janus

Dr. Richard Loft
Director of Technology Development
CISL/NCAR
Janus Supercomputer

- Joint project between CU Boulder, NCAR and CU Denver
- Number 78 on the Nov 2011 Top 500 List with 152.2 teraflops on Linpack
- Funded in part through a NSF MRI project, PI/CO-I team includes
  - Henry Tufo – CU, Boulder/NCAR
  - Jan Mandel – CU, Boulder, Denver
  - James Syvitski - CU
  - Richard Loft -NCAR
  - Keith Julien – CU, Boulder
Janus Supercomputer Facts

• 1368 compute nodes (Dell C6100)
  – two, 2.8 GHz, 6 core Intel Westmere processors per node
  – 2 GB/core; 24 GB per node
  – 16,428 total cores
  – One QDR NIC per node
• Fully non-blocking QDR Infiniband network
• 960 TB of usable Lustre-based scratch storage
  – 16-20 GB/s max throughput
• No local storage on the compute nodes
• No battery backup of the compute nodes

Janus is a good analog to Yellowstone
Janus Lessons Learned - Management

• Bi-institutional Allocation Process
  – http://www2.cisl.ucar.edu/docs/allocations/janus-request-form
  – Correctly communicating accounting records
• Weekly concalls between CU and NCAR important to resolve issues and build trust and cooperation.
• Front Range Consortium for Research Computing (FRCRC)
  – Forum for reaching a broader consensus and cooperation across institutions.
  – https://www.frcrc.org/
Resource Management Lessons

- Hoards of small/short jobs submitted by (e.g. by biologists) wreaks havoc
- Long-running single processor jobs (e.g. by physicists) clogs system.
- Led to queuing system and policy changes (February 2012)
  - Shared vs exclusive nodes
  - Jumbo (>5761 core) queue
  - Long (7 day) queue
- Moved compilers on native Xeon architecture
# Janus Queue Structure

## Janus Queues

<table>
<thead>
<tr>
<th>Queue</th>
<th>Cores</th>
<th>Max time</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>janus-debug</td>
<td>1-5760</td>
<td>1 hour</td>
<td></td>
</tr>
<tr>
<td>janus-short</td>
<td>1-5760</td>
<td>4 hours</td>
<td></td>
</tr>
<tr>
<td>janus-long</td>
<td>1-960</td>
<td>7 days</td>
<td></td>
</tr>
<tr>
<td>janus-small</td>
<td>12-240</td>
<td>24 hours</td>
<td></td>
</tr>
<tr>
<td>janus-normal</td>
<td>241-960</td>
<td>24 hours</td>
<td></td>
</tr>
<tr>
<td>janus-wide</td>
<td>961-5760</td>
<td>24 hours</td>
<td></td>
</tr>
<tr>
<td>janus-jumbo</td>
<td>5761+</td>
<td>6 hours</td>
<td>Runs 2 times per month for 24 hours</td>
</tr>
</tbody>
</table>
Janus: User Environment Challenges

• Configuring NCAR-specific libraries: NetCDF, etc.
• Establishing 1.5 TB /contrib type storage for initial data sets
• Working parallel environments (MPI/Open MP)
• Compiler correctness issues (both Intel & PGI)
Validating CESM was a challenge

- Running under PGI on Opteron/Cray ≠ PGI on Xeon/Dell
- Jan 4, 2012 – CESM successfully compiled under PGI
- March 6, 2012 – CESM validation runs completed in 250 year run.
- Good to retire many of these issues now, rather than on Yellowstone!
Janus User Community

• 54 projects allocated
  – 38 small (<50 KCPU-hrs)
  – 16 large (>50 KCPU-hrs)
• 20 projects active
  – 7 University
  – 12 NCAR
  – 1 other (NOAA)
• 34 projects inactive
Janus usage in CPU-Hours/month

CESM 250 year Validation Run
Breakdown of Janus Usage by Institution Type (as of May 1st)

- NCAR: 421 KCPU-HRS
- WYO: 156 KCPU-HRS
- UNIV: 113 KCPU-HRS
- OTHER: 0 KCPU-HRS
Possible Causes of Low Utilization Levels

• Adoption curve effect
  – System only *really* usable since February

• Frost Fratricide
  – IBM BG/L still heavily used
  – Frost will be shut down on May 31, 2012

• Waiting for Yellowstone
  – People porting/testing but not doing science runs

• Compiler and RT changes are larger barriers than anticipated *(most worriesome for Yellowstone)*
  – Born out by CESM experience
  – Some evidence prior Linux cluster users fair better
Planned Response

- Survey user experiences
- Emphasize training to familiarize users with new environment – also good idea for Yellowstone
- Continue to work Janus technical issues with CU Research Computing Team
- Allocate additional resources e.g. ~1.5 M CPU-hrs this round.
THANKS