Cray XC System Software Update

The Beat Goes On

kmark@cray.com
https://www.linkedin.com/in/kelly-mark
AGENDA

• Introducing CLE 7.0!
  • Deep dive into Hugepage memory accounting
• XC Patching Strategy Change
• XC Software Roadmap
• XC Software Support Matrix
• Summary
• Q&A
WHY THE EXCITEMENT?
XC – IN THE NEWS

Deep Learning
Fighting Cancer

Improving Drug Success Rates

Understanding the Human Brain

Clean Energy
Research

Performance and Quality

Cray XC
Power to Push Boundaries

Advanced Weather Forecasting

Capture and Convert Ocean Wave Energy
BRIDGE TO SHASTA
INTRODUCING CLE 7.0

Consistent upgrade process – same time requirements

- Same as CLE 6.0 UP06 -> CLE 6.0 UP07

SLES15, Lustre 2.11, NVIDIA 10.0

Performance improvements for ARM

- Turbo
- Native hugepages

Other Improvements

- SSA Client
- Logging & Networking
- eLogin/CLE common behavior (Ansible plays for standardized networking, ntp)
- Direct Attached Lustre (DAL) moved from CentOS to SLES
- Hugepage memory accounting
CLE 7.0 UP00 - Hugepages – Memory cgroup limits

• Memory cgroups now count and include hugepage usage in the overall memory usage

• Now the memory limits set by a WLM and enforced by the kernel will be more comprehensive

• Some jobs may fail because their use of hugepages is correctly tracked and counted against their memory cgroup limits

• This tracking is enabled by default but can be disabled with a boot time parameter

CLE 6.x process address space

CLE 7.x process address space

basex page region

huge pages

not tracked by mem cgroup

base page region

tracked by mem cgroup
Reserved hugepages are now counted in addition to in use hugepages by the out-of-memory (OOM) killer when computing the OOM score for each process. This means that the OOM killer will be better able to target the correct process. This behavior can be disabled at runtime.
The resident set size (RSS) calculations performed by the kernel now include all hugepages that are in use or reserved by the process.

As an example, the RSS data shown in `/proc/<pid>/status`, the summary line displayed by aprun after a job completes, and the `max_rss` value gathered by the RUR taskstats plugin will all now include hugepage usage.

This behavior can be disabled at runtime.
XC PATCH STRATEGY GOALS

• Patches are Released as Generally Available
  • Enables patches to be fully tested like a standard release
• Standardize Installation Across all Patchsets
  • Leverage tools used for Major/Minor releases
  • Reduces install errors
• Eliminate Patch Dependency Chains
  • Patches are cumulative – and contain any dependencies
• Patches are Released on a Regular Cadence
  • Monthly (or as needed)

Patches are generally available (GA) to all customers at the same time
WHAT IS IN A CUMULATIVE PATCHSET?

• README
  • One README.txt instead of three
  • Install instructions removed (unless exception)

• Standard installer (and removed INSTALL and LOAD scripts)

• Patchset manifest – rpm list of what changed in SMW/CLE ISOs

• CLE, SMW, SLE Update ISO’s

• Revised Publication: S-2559
  • Standard instructions regardless of Patchset content
XC PATCH UPDATE - CURRENT

- **CLE**
  - Download: README, INSTALL, LOAD, ISO (rpms)
  - Copy to Directory /var/adm/cray/release/patchsets
  - Run LOAD (creates backup snapshot of current system)
  - Run INSTALL (updates current snapshot)
  - Build Images
  - Run Manual steps depending on patch (zypper; zap; reboot cabinet controller)
  - Boot CLE

- **SMW**
  - Download: README, INSTALL, LOAD, ISO (rpms)
  - Copy to Directory /var/adm/cray/release/patchsets
  - Run LOAD (creates backup snapshot of current system)
  - Run INSTALL (updates current snapshot)
  - Build Images
  - Run Manual steps as directed in README
  - Boot SMW

- **SLE**
  - Download: README, Loadfile, License, Changes.txt, Md5Sum, Record.fn, ISO (rpms)
  - Copy to Directory /root/isos
  - Run Manual steps as directed in README
  - Boot SLE

3 patch processes depending on patch type

Each patch required unique instructions

Inconsistent snapshot instructions
XC CUMULATIVE PATCH PROCESS

1 patch process
Same process as software updates
Fully documented, few exceptions

Boot SMW to Patch Set snapshot
Update:
- Configuration
- Cabinet Controller
- Firmware

Boot CLE
Update eLogin
## CUMULATIVE PATCHSET - EXAMPLE

<table>
<thead>
<tr>
<th></th>
<th>PS01 March</th>
<th>PS02 April</th>
<th>PS03 May</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLE 7.0.UP00</td>
<td>March changes</td>
<td>March changes</td>
<td>March Changes</td>
</tr>
<tr>
<td></td>
<td>April Changes</td>
<td>April Changes</td>
<td>May Changes</td>
</tr>
<tr>
<td>SMW 9.0.UP00</td>
<td>March Changes</td>
<td>March Changes</td>
<td>March Changes</td>
</tr>
<tr>
<td></td>
<td>April Changes</td>
<td>April Changes</td>
<td></td>
</tr>
<tr>
<td>SLE Update</td>
<td>SLE Update-April</td>
<td>SLE Update April</td>
<td></td>
</tr>
</tbody>
</table>
# XC PATCH UPDATE – SCHEDULE

<table>
<thead>
<tr>
<th>Release</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>July</th>
<th>August</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLE 7.0/SMW 9.0 UP00</td>
<td>PS02 4/3</td>
<td>PS03 5/1</td>
<td>Week1</td>
<td>Week1</td>
<td>Week1</td>
</tr>
<tr>
<td>CLE 6.0/SMW 8.0 UP07</td>
<td>PS31 4/10</td>
<td>PS32 5/8</td>
<td>Week2</td>
<td>Week2</td>
<td>Week2</td>
</tr>
<tr>
<td>CLE 6.0/SMW 8.0 UP06</td>
<td>n/a</td>
<td>PS36 5/15</td>
<td>Week3</td>
<td>Week3</td>
<td>Week3</td>
</tr>
</tbody>
</table>
XC PATCH UPDATE – CUSTOMER FEEDBACK

• NO COMPLAINTS (yet)!!
  • 10 customers have applied 6.0.UP07 PS30
  • 6 customers have applied 7.0.UP00 PS02

• Some praised new approach

• A few questions
  • One customer was concerned about patch size, we were able to remove multiple CUDA rpms and significantly reduced ISO size
  • One customer noticed a documentation error in PS30
### 2019 XC ROADMAP – SOFTWARE

<table>
<thead>
<tr>
<th>2019</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CLE 7/SMW9 UP00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Security Pack</td>
<td>April Patchset</td>
<td>May Patchset</td>
<td>June Patchset</td>
</tr>
<tr>
<td></td>
<td>CLE 7/SMW9 UP01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>July Patchset</td>
<td>Aug Patchset</td>
<td>Sept Patchset</td>
<td>Oct Patchset</td>
</tr>
<tr>
<td></td>
<td>Nov Patchset</td>
<td>Dec Patchset</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**CLE 7.0/SMW 9.0 UP00**
- SLES15
- Hugepage Memory reporting
- ARM B2 support
- CascadeLake
- NVIDIA Cuda 10.0
- Workload Managers w/SLES15 support
- Security Pack
- Lustre 2.11

**CLE 7.0/SMW 9.0 UP01**
- Bug Fixes
  - Logging
  - PE Image Push/sqpush
- Security Packs
  - 6.0.UP06
  - 6.0.UP07
  - 6.0.UP00
- Workload Manager qualifications
- NVIDIA Cuda 10.1
- Lustre 2.12

**Key:**
- Patchset includes quarterly security pack
## 2020 XC ROADMAP – SOFTWARE

<table>
<thead>
<tr>
<th></th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CLE 7.1 UP00</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan</td>
<td>Patchset</td>
<td>Feb Patchset</td>
<td>Mar Patchset</td>
<td>April Patchset</td>
</tr>
<tr>
<td><strong>CLE 7.1/SMW 9.1 UP00</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLES15 SP1</td>
<td>Security Packs</td>
<td>Security Packs</td>
<td>Security Packs</td>
<td>Security Packs</td>
</tr>
<tr>
<td><strong>CLE 7.1/SMW 9.1 UP01</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.0UP07</td>
<td>7.0UP00</td>
<td>7.0UP01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Key:**
- Patchset includes quarterly security pack
XC Software Support Matrix
## CLE/SMW Support Matrix

<table>
<thead>
<tr>
<th>CLE/SMW Major or Minor Releases</th>
<th>CLE/SMW Update Releases</th>
<th>Release Date</th>
<th>*Full Support Ends</th>
<th>*Extended Support Ends</th>
<th>*Legacy Support Ends</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLE 7.0/SMW 9.0 Final Major Release in support of XC systems</td>
<td>UP03</td>
<td>3Q30</td>
<td>TBD</td>
<td>TBD</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>UP02</td>
<td>1Q20</td>
<td>CLE 7.0.UP02 GA</td>
<td>TBD</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>UP01</td>
<td>3Q19</td>
<td>CLE 7.0.UP01 GA</td>
<td>8/27/2020</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>UP00</td>
<td>2/28/2019</td>
<td>CLE 7.0.UP01 GA</td>
<td>2/27/2020</td>
<td>N/A</td>
</tr>
<tr>
<td>CLE 6.0/SMW 8.0</td>
<td>UP07</td>
<td>7/12/2018</td>
<td>2/28/2019</td>
<td>2/28/2020</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>UP06</td>
<td>3/1/2018</td>
<td>7/12/2018</td>
<td>8/28/2019</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>UP05</td>
<td>10/5/2017</td>
<td>3/1/2018</td>
<td>7/12/2018</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>UP04</td>
<td>6/27/2017</td>
<td>10/5/2017</td>
<td>3/1/2018</td>
<td>N/A</td>
</tr>
<tr>
<td>CLE 5.2/SMW 7.2 Final Major Release in support of XE/XK systems</td>
<td>UP04</td>
<td>09/25/2015</td>
<td>12/31/2017</td>
<td>1/31/2019</td>
<td>Ends 1/31/2022</td>
</tr>
</tbody>
</table>
Summary

• CLE 7.0 – performance and stability improvements
• Cumulative Patches – improved visibility for customers, improved quality
• Predictable roadmap cadence - 2 releases per year with monthly patches
• Look forward to more of your achievements on XC
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.
THANK YOU

QUESTIONS?

kmark@cray.com
https://www.linkedin.com/in/kelly-mark

cray.com
@cray_inc
linkedin.com/company/cray-inc-