# State of XT Software: The Year in Review

**David Wallace**
**Cray Inc.**

## Abstract

The Cray XT Software Annual Report will highlight the accomplishments over the last twelve months. The presentation will provide detailed status on software development projects including XT4 Quad Core and Cray XT5, statistics on Software Problem Reports, and software releases. The presentation will provide a brief view of software roadmap and plans for the next twelve months.

Keywords: CNL, Catamount

## Introduction

This presentation reviews the accomplishments over the last twelve months by the Cray Software group with respect to the Cray XT system. A year ago in my XT update presentation, I talked about the UNICOS/lc 1.4 and 1.5 releases, dual core and a preview of the UNICOS/lc 2.0 release and CNL. All of that is behind us now and today I'd like to touch on the accomplishments over the past 12 months.

## Accomplishments

The last twelve months since the previous CUG have been extremely busy for the Software group. Representative of the work that's been done is the number of software updates that have been shipped to the field. Table 1 shows a list of UNICOS/lc 1.5 updates and their respective release dates. You can see that the trend for 1.5 has been fewer releases over time. This is indicative of the increasing stability of UNICOS/lc 1.5. Software continues to support the 1.5 release however releases are now focusing on primarily critical fixes only.

**Table 1 UNICOS/lc 1.5 updates**

| Revision | Release Date |
|---|---|
| 1.5.45 | 03-MAY-2007 |
| 1.5.47 | 11-MAY-2007 |
| 1.5.52 | 29-JUN-2007 |
| 1.5.55 | 20-JUL-2007 |
| 1.5.57 | 10-AUG-2007 |
| 1.5.59a | 08-OCT-2007 |
| 1.5.60 | 02-NOV-2007 |

The release of UNICOS/lc 2.0 and Compute Node Linux (CNL) marked a major achievement for the Software Development group. The release of CNL is the culmination of over two years of development to create an implementation of Linux that scales and performs well on very large systems. Our partnerships with ORNL and NERSC played a significant part in this achievement. I'd like to thank all of the people who contributed to the CNL project, particularly all of the developers and site

support personnel at ORNL who spent many long weekends and nights testing CNL.

**Table 2 UNICOS/lc 2.0 releases and updates**

| Revision | Release Date |
|---|---|
| 2.0.LA | 02-JUL-2007 |
| 2.0.14 | 30-JUL-2007 |
| 2.0.17 | 13-AUG-2007 |
| 2.0.20 | 10-SEP-2007 |
| 2.0.GA | 10-OCT-2007 |
| 2.0.33 | 06-DEC-2007 |
| 2.0.35 | 20-DEC-2007 |
| 2.0.36 | 08-JAN-2008 |
| 2.0.39 | 24-JAN-2008 |
| 2.0.40 | 01-FEB-2008 |
| 2.0.41 | 21-FEB-2008 |
| 2.0.44 | 10-MAR-2008 |
| 2.0.49 | 18-APR-2008 |

The adoption rate of CNL has been significant as well. Over 50% of the customer cabinets are now running CNL and all new systems are being installed with CNL. CNL is experiencing growing pains, but users have reported significant results using CNL.

CNL is experiencing some growing pains and the frequency of updates reflects this (see Table 2). This is similar to early releases of Catamount although the focus of fixes is much narrower. Many of the current critical or urgent problems are related to how Linux stresses the hardware and network differently from how Catamount stresses the system. We are finding and fixing a number of latent Portals problems which result in a significant reduction in the number of system interrupts and a rapid increase in overall system stability.

Software development has also been very active in working with the Support organization on a variety of customer projects and work related to improving support tools. Some of the support work is highlighted in the list below:

- Completed joint CNL assessment with NERSC.
- Helped support the Army HPCRC migration to Compute Node Linux (CNL).
- Engineering and managerial support of CNL bring-up on NERSC and ORNL machines.
- Completed a STREAMS performance analysis for NERSC.
- Spent a significant amount of time on analyzing OS jitter. The outcome of this investigation was an analysis paper as well as implemented improvements.
- Assisted Service organization on most system acceptances.
- Worked with the XTreme group to identify future software requirements with special attention given to requirements related to system administration.

Software Development is involved in most aspects of system acceptances including benchmarking and system support.

In addition to the many software updates for UNICOS/lc 2.0, Software has also been active in planning and implementing new

software enhancements. Several of these enhancements are currently being field tested at selected customer sites. The UNICOS/lc 2.0 release has been retrofitted with a number of new features or enhancements. Some of these features/enhancements are listed below:

- Unified Boot
- Ldump
- Linux Kernel
  - support for Quad Core and SMP
  - Family 0x10 processor support
  - Critical bugfixes and patches
- AMD Quad Core processor support
- Compute Node Health Daemon (Phase 1)

Software is currently focused on the upcoming major software release for the Cray XT systems. The list below highlights some of the major changes and enhancements included in the Cray Linux Environment 2.1 release.

- SLES10 SP1 on SIO nodes
- Great improvements to XTInstall tool!
- Perfmon 2.3 2.6.5X
- Comprehensive System Accounting
- Cray Data Virtualization Service
- Service node failover and warmboot (Phase 1)
- Affinity/pinning with SDB support (segment tables)
- Restructuring of the software build/RPMs
- Portals performance optimizations on CNL
- Kernel Huge page support
- Improvements to Out-of-Memory (OOM) killer on the XT Compute Nodes

Details of the new features and changes in the CLE 2.1 release can be found in the CLE 2.1 Release Preview document so I will not attempt to describe them here. I do want to highlight the work that has been done to improve the XT software install tool. As some of you will recall, the upgrade from 1.3 to 1.4 which included an upgrade from SLES 8 to SLES 9 required a complete re-install of the software. This typically took about a week to accomplish including the site customizations that were required. Internally, we've done numerous upgrades to our systems including upgrades from 1.5 to 2.0, and 1.5 to 2.1. Feedback from sites who upgraded from 1.5 to 2.0 was very positive. And further enhancements have improved the process further. Much of the prep work can now be done off-line and specific attention has been given to minimizing the amount of dedicated time required to complete the upgrade.

One of the objectives of Software has been to implement a common software strategy for all of the Cray systems. To this end, Software has made great strides in integrating the common software strategy. Both XMT and XT have shipped systems using XT SIO nodes. Included in this effort, XT and X2 are using common Linux source code for the OS.

While not strictly related to Cray XT software, the release of the first Cray X2 system marked the introduction of Cray's hybrid computing product. Figure 1 illustrates a hybrid XT and X2 system. The file system, network, login and system nodes are XT. And you see both X2 and XT blades all connected using the SeaStar network. As previously mentioned, X2 and XT use the same kernel source base and in addition, the

Programming Environment shares a number of common elements.

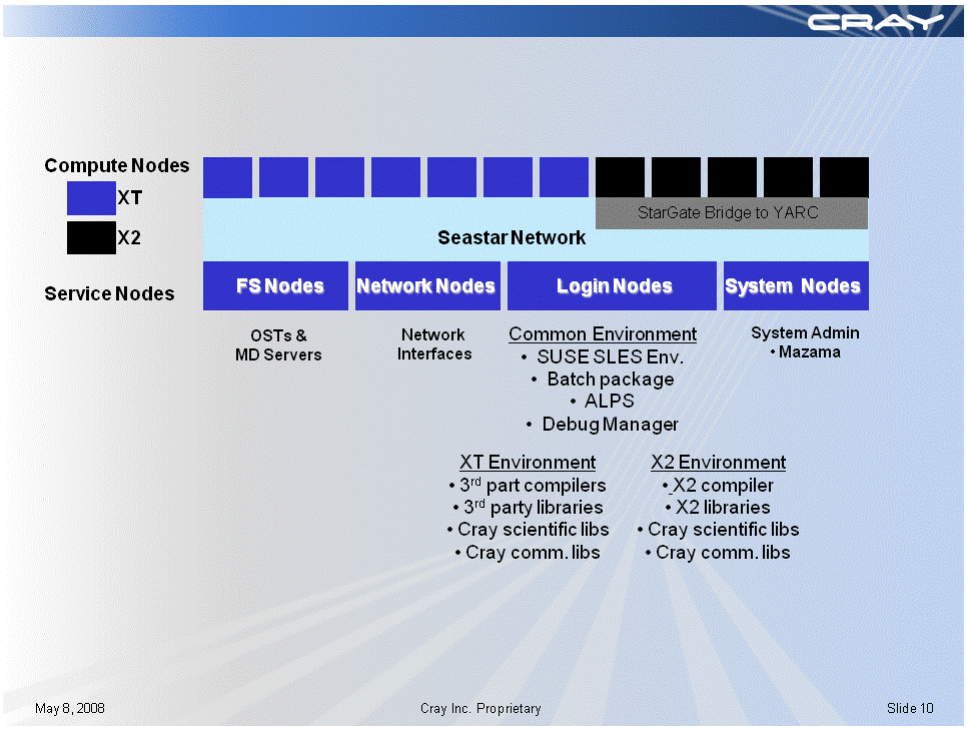


**Figure 1 Hybrid User Environment**

A list of accomplishments wouldn't be complete without covering Software's efforts to support new Cray hardware. Software Development has been very active in supporting new hardware including support for XT4 Quad Core, XT5, PCI-Express, as well as several hardware issues including one particularly nasty problem represented by SPR 740520. The Quad core work deserves special recognition because Software was able to qualify a new Linux kernel and introduce the change into the mainstream UNICOS/lc release without any regressions. The new software to support QC was available to customers in advance of the actual quad core hardware so customers could decouple the software upgrade from the actual hardware upgrade. In fact, ORNL upgraded and was running the new software in production a full month before beginning their hardware upgrade.

Another accomplishment was Software's effort to reduce the SPR backlog. The Software group established quarterly goals for reducing the number of unresolved or open SPRs and we succeeded in reducing the backlog by 40% over the year. Scoring criteria was established that took a number of factors into account including severity and age of the SPR. This was largely to make sure that managers didn't attempt to game the system in order to achieve their SPR reduction targets.

### A glimpse of the future…

Behind the scenes, the Software group is working diligently to define the software roadmap for the operating system and programming environment that covers the next three to five years. Not only is this roadmap intended to guide the development of new features and enhancements, it is

largely intended to provide a peek at the long-term plans which will allow customers (and Sales and Marketing) to plan for future software upgrades. Details of the current Software Roadmap can be found in Peter Young's CUG presentation "Cray Software Update".

A few things to note:

- The Linux operating system and user environment will track closer to the SUSE releases. SLES 10 SP1 will be supported in the CLE 2.1 release. SLES 10 SP2 is planned to be supported in the CLE 2.2 release.
- CLE will track closer to the SUN/CFS Lustre releases. This is a goal. The OSIO group will test and qualify each Lustre update to maintain and enhance file system resiliency, reliability and performance. It is our goal to only release Lustre updates that meet the established release criteria.
- SIO node failover and reboot will be supported in updates beginning with CLE 2.1.
- The first phase of the CNL node health project will be introduced in CLE 2.1. This project, phased over several releases, is aimed at reducing job failures and system outages by detecting, reporting and taking appropriate action on nodes that are not functioning properly.

**Summary**

The crowning achievement since the last CUG in May 2007 was the release of UNICOS/lc 2.0 and Compute Node Linux. In addition, the Cray Software group has been very busy supporting all aspects of the company including Customer Service, System Test and Checkout, Manufacturing, Sales and Marketing. A number of software releases and updates were distributes (UNICOS/lc 1.5 and 2.0) and plans for the next release, CLE 2.1, are in the final stages of preparation. Work on CLE 2.2 has already begun. Longer term plans in the form of a Software Roadmap are being developed to assist customers in planning for system upgrades.