# XT3 – Scaling to New Heights:

# Software Status and Plans

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

**ABSTRACT:** *This paper discusses the current status of XT3 software and development plans for 2006. A review of major milestones and accomplishments over the past year and a half is presented.*

**KEYWORDS:** XT3, software, plans

## Introduction

The general theme of this paper is 'past, present and future'. The first section will provide a perspective of the major software milestones and accomplishments over the past eighteen months. The second section will discuss the current status of the software with respect to development, releases and support. The paper will conclude with a view of future software plans.

## Major Milestones

XT3 software has made significant progress over the last year and a half. And while we still have a ways to go, customers and users are getting excellent results from the machine. Some of the more notable milestones and accomplishments are noted in Figure 1. A key turning point for the system was the release of v1.3 of UNICOS/lc. Coupled with some major enhancements in Lustre and Portals, customers are routinely running applications on large node counts.

Going back to the beginning of 2005, the XT3 system was experiencing a number of significant problems. Fundamental functions such as booting the system were unreliable, the monitoring and management system was barely working, and Portals was extremely unstable. Software developers were working directly with several sites in attempts to stabilize the systems to meet specific milestones. It was not until the summer of 2005 that significant progress was made to identify and resolve a problem that prevented the high-speed network (HSN) from running reliably. With a reliable HSN, significant progress was made in a number of areas such as booting the system and the Lustre file system.

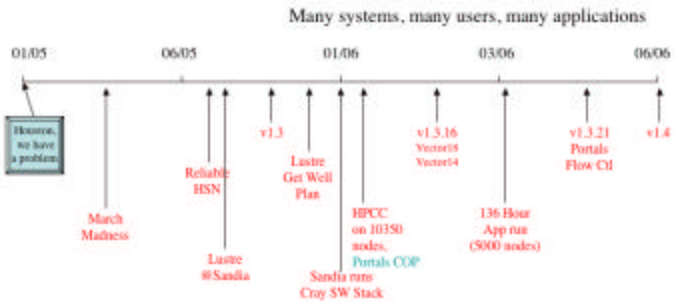


**Figure 1 XT3 major milestones**

The release of UNICOS/lc v1.2 allowed a number of systems to pass acceptance tests but still did not allow most systems to move into production. The release of version 1.3 raised the level of stability of the system and with this an increase in useability. With the increase in load on systems at customer sites, a new set of problems and issues arose, particularly with Lustre and Portals.

Two areas that have undergone major changes are Portals and Lustre. On the Portals side, problems in the communication protocol were being exposed by user applications that were running at increasingly larger node counts. The majority of these issues were resolved by a concerted effort to refactor portions of the Portals code. The focus has been on improving Portals stability and reliability in addition to making incremental improvements in network latency. Portals latency has been reduced from ~20us to less than 7us, with additional improvements in the works. COP and most recently flow control have been added to improve Portals stability and reliability. Customers are now reporting far fewer node dropouts. One of the positive side effects of the COP feature was the ability to reboot compute nodes. This feature is being released in updates to v1.3 and v1.4. This has resulted in a much more stable and resilient implementation of Portals.

In November, we initiated a Lustre get-well plan with our file system partner CFS. At the time, there were numerous reported Lustre problems at customer sites. The file system development group worked closely with CFS to prioritize problems, develop and deliver fixes to critical issues over several updates. The outcome is that Lustre is much more stable. We continue to work with CFS on longer term plans to improve the robustness of Lustre. The outcome of this work was that most customers were able to reliably run large Lustre configurations.

Some but not all XT3 sites experienced problems that resulted in compute nodes failing. Known internally as the Vector 18 and Vector 14 problems (these are generic entry points in the trap handler in the software), the problems proved to be difficult to isolate and fix. These problems were ultimately resolved and the fixes shipped to the field in software updates.

The payoff for the work that went to the various updates to version 1.3 is increases in system utilization and availability. Systems have moved into production mode, the number of users on the systems is increasing, and many applications are being run on XT3 systems with impressive results. A significant milestone was achieved in January 2006 when Sandia National Labs began running the UNICOS/lc 1.3.10 software set (up until this point, Sandia had run a mixture of Cray and Sandia software). This milestone was achieved through vigorous and combined efforts of Sandia and Cray. XT3 systems at Sandia, Pittsburgh Supercomputing Center and Oak Ridge National Laboratory now routinely run applications utilizing thousands of processors (as high as 10350 processors). Of note is one particular application that ran uninterrupted for 136 hours using 5,000 processors.

### XT3 still has a ways to go

Despite the considerable progress that has been made over the last year and a half, there are still significant hurdles that must be overcome. Over the last six months, the XT3 software development group has been focused on addressing software reliability and stability issues. Table 1 shows a list of the software updates for version 1.3. It is easy to see from inspecting the table where the critical areas in XT3 software were. Eleven updates were released since the release of v1.3 in November. This represents an average of two updates a month – a considerable feat for any software development group.

**Table 1 Summary of v1.3 software updates**

| Release Version | Description (focus of effort) |
|---|---|
| 1.3.07 | Base release |
| 1.3.09 | Lustre, Portals, SMW/Diags |
| 1.3.10 | Portals, Lustre, OS, Qk |
| 1.3.11 | SMW/Diags, OS |
| 1.3.12 | Lustre performance, SMW/Diags, Features |
| 1.3.13 | Portals, Lustre, Security |
| 1.3.15 | SMW/Diags, OS, Qk |
| 1.3.16 | Qk, SMW/Diags, IPPO |
| 1.3.17 | Portals, Lustre, Qk (Regression) |
| 1.3.18 | Portals, Lustre, Qk, MPICH |
| 1.3.20 | Portals, Qk, Lustre, SMW/diags |
| 1.3.21 | Portals (flow control), CPA |

The chart in Figure 2 shows the level of activity addressing Software Problem reports (SPR). Over the life of the Red Storm/XT3 product, a total of over 2000 SPRs have been closed. Figure 3 shows that there is still a considerable backlog that must be addressed.

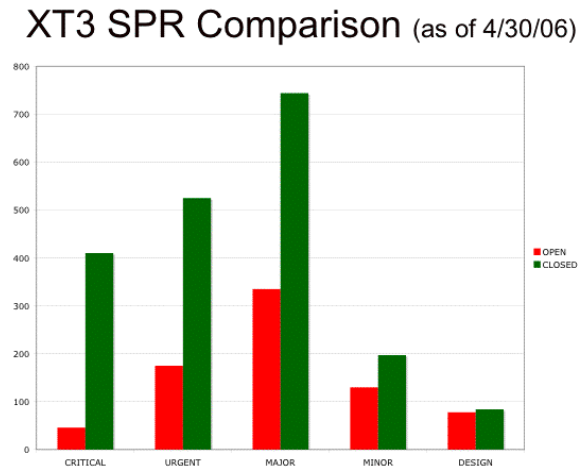


**Figure 2 Open and Closed SPRs**

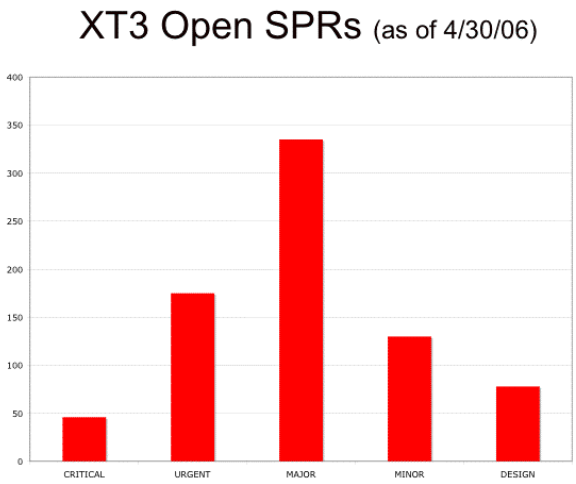


**Figure 3 XT3 SPR backlog**

The SPR backlog is being managed by:

- Planning for Critical SPR interruptions in the development plans

- Scheduling non-critical SPR work as part of the development plans
- Scheduling 'SPR weeks' where emphasis is placed on fixing selected SPRs

Obviously, the best plan for reducing or eliminating the SPR backlog is to improve our software development process such that defects are not released to customers.

## XT3 Software Development

***On-going Process Improvement***

The XT3 software is relatively immature and is therefore going through an extended period of rapid enhancement. Many of the changes are to support new hardware features, customer requested enhancements, and bug fixes. The XT3 development processes reflect the need to support rapid product development and delivery. Over time, it is expected that changes in the product requirements will drive changes to the development processes. These will be discussed later in this paper.

Software releases are planned for the second month of each quarter. Three separate development streams are maintained to allow for simultaneous development, release and support. Figure 4 describes the process. The top line describes the development stream. The twelve-week cycle is broken into two segments:

- The first ten weeks is scheduled for integration of features and fixes. The most impactive features are scheduled for integration first to allow more exposure during testing of the development stream.
- The last two weeks are planned for stabilizing the development stream. During this period, only fixes to address stability issues are integrated.

The entire integration process is managed by a Change Control Board who review all changes being made to the development stream, Unit/Feature testing as well as some regression testing is done concurrently in the development stream. At the end of the twelve-week period, the target release code is split from the development stream, and the next twelve-week cycle begins again.

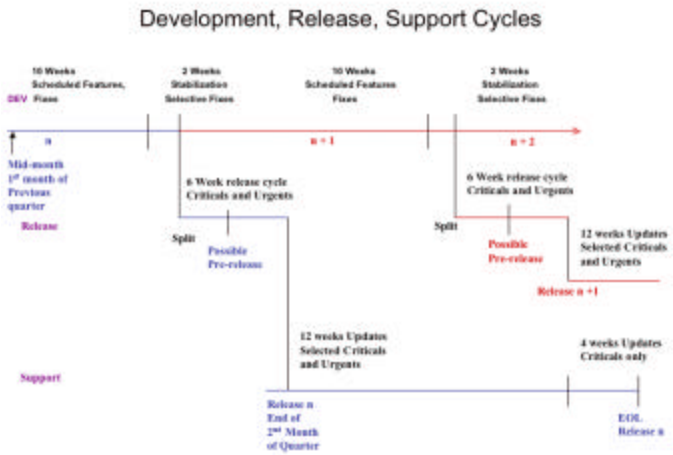


**Figure 4 XT3 development process**

The release cycle consists of six weeks of regression testing and exposure on in-house systems. Te details of the cycle follow:

- Week 1: Build, Regression Test, Expose
- Week 2: Build, Regression Test, Expose
- Week 3: Build, Regression Test, Expose
- Week 4: Build, Installation/Regression Test, Expose – Limited Availability Release
- Week 5: Installation/Regression Test, Expose, Create release errata
- Week 6: Packaging, Final Errata, Media Duplication – General Availability Release

Limited Availability releases are intended to support shipments of new hardware (prior to general availability) as well as a restricted number of pre-release installations. These are intended to provide managed and controlled releases of the software. This change addresses issues supporting sites who received pre-release software which was not under release control.

The bottom line in Figure 4 represents the support stream. The frequency of updates (fix packages) ranges from weekly to 'as required'. Updates are made available for up to twelve weeks following a release. The updates contain fixes to selected critical and urgent problems. Once the next release occurs, updates continue for four more weeks and include fixes to critical problems. This allows customers time to upgrade to the next release.

Over time, the objective of XT3 software development is to implement a process that results in the development source tree to be maintained in a "continuously releasable" state. The goal is the ability to take a snapshot of the development tree, QA and release the system in less than two weeks. Note that this doesn't necessarily mean that there will be releases every two weeks. It means the capability to create a release in two weeks exists. In order to achieve this objective, efforts to minimize destabilization of the source tree must be made:

- All features and fixes are developed in side branches of the source tree. Prior to integration, each and every fix must be tested.
- Features and bugfixes are subjected to peer design and code review. Depending on the scope of the change, a feature may be reviewed by the Software Architecture Committee.
- All changes to the source tree are managed through the Change Control review board.
- Daily build and testing (automated) to identify problems as soon as possible.
- Partner with customers to test on large system configurations.

**Looking Ahead**

There are several forces driving XT3 software development:

- XT cleanup: the backlog of SPRs indicates potential problems with reliability and performance implications. The backlog needs to be reduced to a more manageable state. Some of the bugs are 'just bugs'. Others will require restructuring or refactoring the code to obtain the proper function, performance and maintainability.
- New hardware and software commitments: There is a continuous stream of new hardware, devices and customer requirements. The XT3 will establish the baseline for future Cray scalar systems.
- New and merged components for Common Systems: As mentioned above, the XT3 will be the baseline for future systems. Where it makes sense, efforts will be made to move towards common, shared components across platforms in order to leverage development resources and efforts.

Many of the initiatives started in the last year have not yielded the improvements or resolved all of the problems. XT3 software quality continues to need improvement. The development processes will continue to evolve. Expected changes include:

- Improve the development processes for design and modifications
- Improve bug reporting and fix reporting
- Improving the checkout and review of changes
- Changing the ownership and processes for integrating mods into release streams
- Improving testing at all levels
- Adding developers and testers to key areas

Software releases are currently being delivered quarterly with updates being released on a bi-weekly basis. As machine sizes approach 20K and 30K nodes, checkout of scaling for both features and fixes will become increasingly difficult. It is anticipated that a new development and release model will have to be developed that will incorporate both internal testing as well as customer on-site testing.

### XT3 Software Roadmap

Plans are being developed to map the XT3 software development activities for 2006 through 2008. Figure 5 presents a high-level view of the scalar software roadmap. Key points in the roadmap are:

- Release 1.4: support for single and dual core Opteron processors. A new versions of the compute and service node OS based on Catamount Virtual Node and SUSE SLES9 respectively. A new version of the Lustre file system.
- Release 1.x: support for Hood and customer commitments, support for PathScale compilers.
- Release 2.x: an additional lightweight compute node kernel and job launch utility, support for PathScale compilers.

### Conclusion

The XT3 software system has undergone rapid enhancement and change over the last six months that has resulted in improved reliability and performance. To the end user, this translates to a more robust computing environment that is allowing more applications to run at even greater scales than previously attempted. Despite the progress, there is still much to be done to meet the overall objectives of the XT3 product. Enhancements in the software development processes will result in a more robust computing environment that will serve as the base for future Cray scalar systems.

Scalar Platform High-Level Software Roadmap

| Component | Rel 1.4 (2Q06) | Rel 1.x | Rel 2.x |
|---|---|---|---|
| HW Supported | XT3 Single and Dual Core | XT3 Single and Dual Core, XT4 | XT3 and XT4 |
| OS (SIO) | SLES 9 | SLES 9 | SLES 9 |
| OS (Compute) | Catamount Virtual Node (CVN) | Catamount Virtual Node (CVN) | CVN and Compute Node Linux (CNL) |
| Runtime support | Yod/CPA | Yod/CPA | Yod/CPA or ALPS |
| HSN Protocol | Portals | Portals | Portals |
| File System | Lustre 1.4.6 | Lustre 1.4.6? | Lustre 1.8 |
| System Mgmt | CRMS | CRMS | CRMS |
| Compilers | PGI 6.1, gcc | PGI 6.x, gcc, | +PathScale |
| Tools | Apprentice$^2$, CRAYPAT | Apprentice$^2$, CRAYPAT | Apprentice$^2$, CRAYPAT |

**Figure 5 XT3 software roadmap**

### About the Author

David Wallace is the Software Technical Project Leader for the XT3 project. Previously, he was the Director of System Software during the development of the CRAY T3E and Director of Software Engineering at SGI. David can be reached at dbw@cray.com.