Cray System Software Features for Cray X1 System

Don Mason

Cray, Inc.

1340 Mendota Heights Road

Mendota Heights, MN 55120

dmm@cray.com

ABSTRACT

This paper presents an overview of the basic functionalities the Cray X1 system software. This includes operating system features, programming development tools, and support for programming model

Introduction

This paper is in four sections: the first section outlines the Cray software roadmap for the Cray X1 system and its follow-on products; the second section presents a building block diagram of the Cray X1 system software components, organized by programming models, programming environments and operating systems, then describes and details these categories; the third section lists functionality planned for upcoming system software releases; the final section lists currently available documentation, highlighting manuals of interest to current Cray T90, Cray SV1, or Cray T3E customers.

Cray Software Roadmap

Cray's roadmap for platforms and operating systems is shown here:
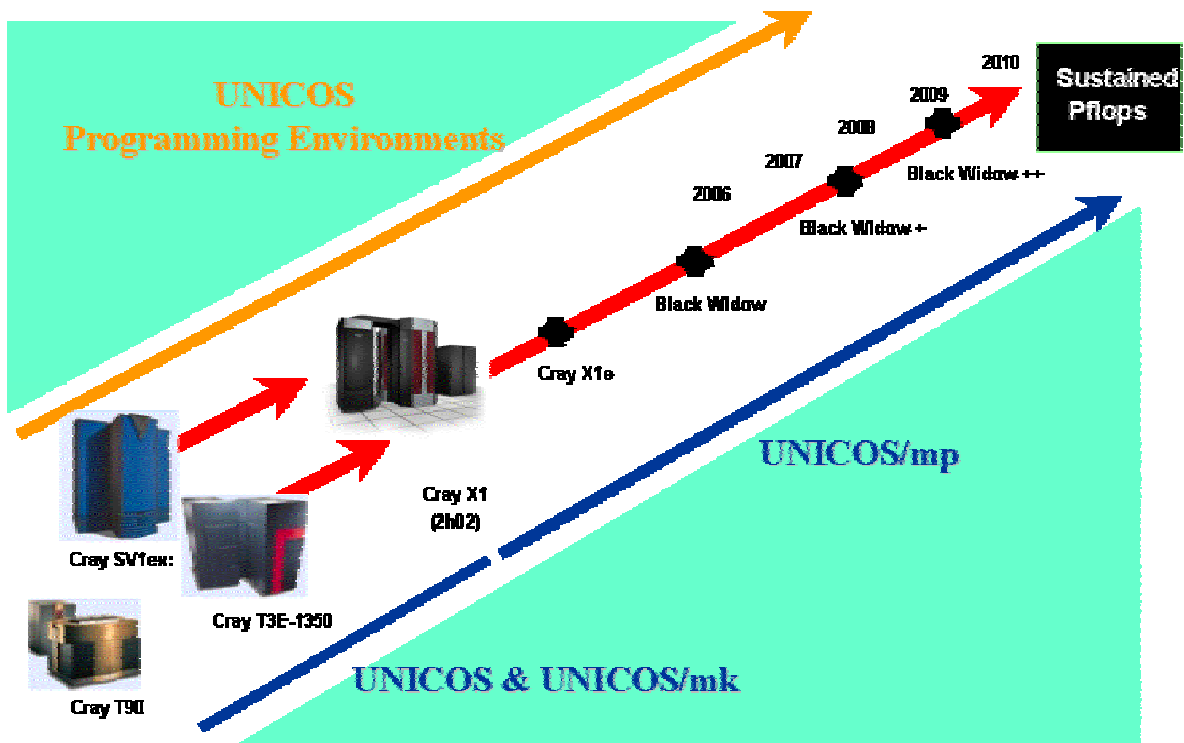
Figure 1: Cray Software Roadmap

The Cray X1 system is the first scalable vector processor system that combines the characteristics of a high bandwidth vector machine like the Cray SV1 system with the scalability of a true MPP system like the Cray T3E system. The Cray X1e system follows the Cray X1 system, and the Cray X1e system is followed by the code-named Black Widow series systems. This new family of systems share a common instruction set architecture.

Paralleling the evolution of the Cray hardware, the Cray Operating System software for the Cray X1 system builds upon technology developed in UNICOS and UNICOS/mk. The Cray X1 system operating system, UNICOS/mp, draws in particular on the architecture of UNICOS/mk by distributing functionality between nodes of the Cray X1 system in a manner analogous to the distribution across processing elements on a Cray T3E system.

UNICOS/mp will be utilized on the Cray X1 system and follow-on products as shown in the Figure 1.

The UNICOS Programming Environment compilers utilize the vector technology developed over a period of more than 25 years by Cray. The common code generator on the Cray X1 system is the same as those used on the Cray T90 system and Cray SV1 system.

Cray System Software Components

The components of the Cray X1 system software are shown here:
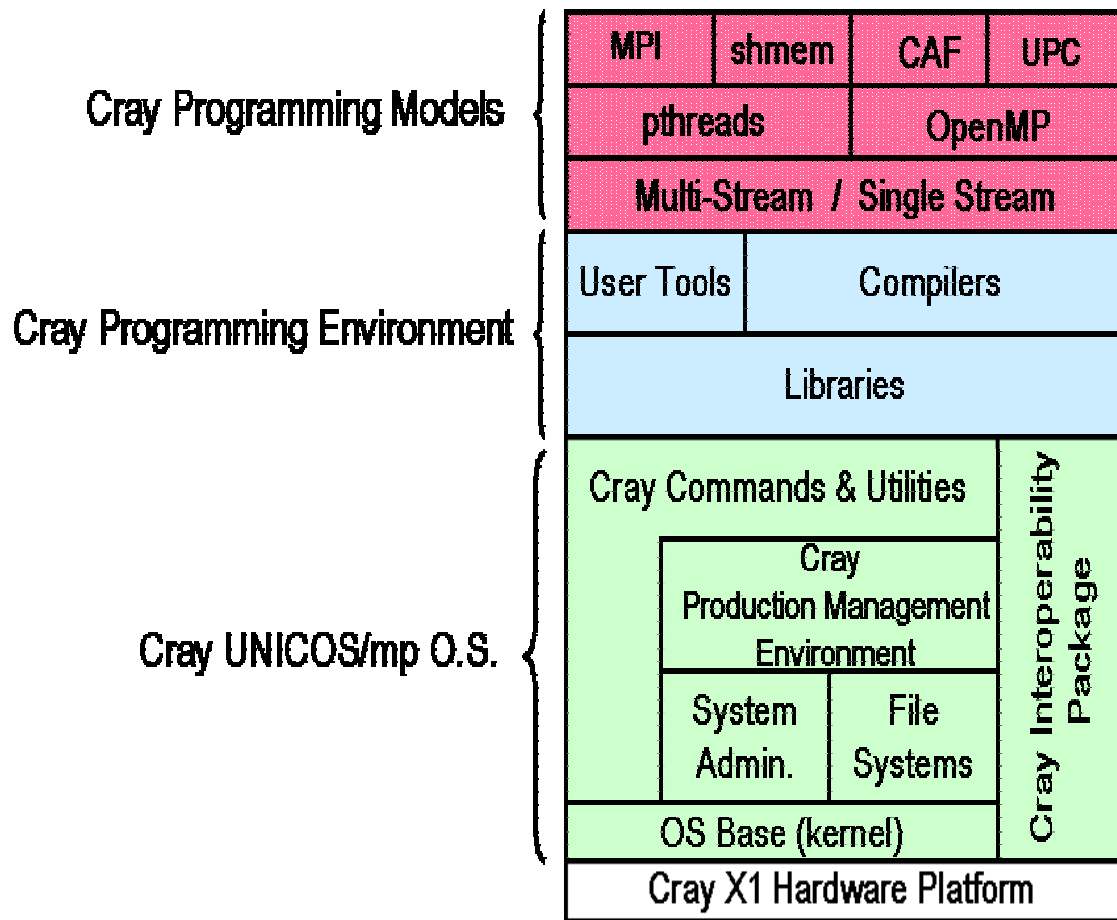
Figure 2: Cray X1 System Software Components

In this figure, the base of the component set is the Cray X1 system hardware. Layered on top of the base are the components of the Cray X1 system UNICOS/mp Operating System, then the Programming Environment components and finally the Programming Models used in Cray X1 system applications. These areas are discussed in the following sections.

Cray Programming Models

Message Passing Interface (MPI), our most important programming model, takes advantage of the architectural features of the Cray X1 system: high bandwidth and global addressing. MPI applications, like all distributed programming models, can span the entire set of application nodes on a Cray X1 system. Our MPI implementation supports the important features of the MPI-2 standard including MPI I/O, one-sided messaging, and enhanced collectives.

Distributed Memory Models: UPC and CAF

Cray also supports two newer distributed programming models that offer many forward-looking advantages: Unified Parallel C (UPC) and Fortran Co-arrays (CAF). Both models are 'in language' so that their expression of parallelism is much simpler. For example, type information does not need to be re-expressed in subroutine calls; and in fact calls themselves are not used to improve performance. The Cray compilers also use in language expression to overlap system communication and computation to further enhance performance.

Cray supports the interoperability of UPC and CAF within MPI applications, meaning that developers can optimize sections of MPI code with either UPC or CAF. Furthermore, as the U.S. Department of Energy has identified the importance of supporting effective parallel programming models for tera-scale computing (and beyond), they are sponsoring the development of portable implementations of both UPC and CAF. (See www.pmodels.org for more information.)

Shared Memory Models: Pthreads and OpenMP

Shared memory models execute on single Cray X1 system nodes. For shared memory programming, Cray supports pthreads and OpenMP. OpenMP support will be available mid-2003 for both C and Fortran. Some applications may benefit from a layering of MPI (or other distributed models) on top of shared memory models such as OpenMP. This mix of programming models is supported on the Cray X1 system.

MSP and SSP Mode Support

Automatic compiler support for the Cray X1 system multi-streaming processor (MSP) underlies all programming models. Unique synchronization hardware in the Cray X1 system is utilized to tightly couple four cache sharing streams of execution for peak performance.

Certain applications have been shown to perform more effectively by using single streams independently, also known as Single-Stream Processor (SSP) execution. Applications may be compiled (and linked with Cray libraries) to run in either SSP or MSP mode and all programming models are supported for each mode.

Programming Environment Components

Cray Programming Environments provide all tools needed for application development. These include compilers and libraries, as well as performance analysis and debugging tools.

Compilers

Fortran 90, C, and C++ compilers and the CAL assembler are available on the Cray X1 system. Because of full Cray X1 system hardware support for 32 bit operations, default compiler data types now follow industry standards. For example, C 'int's are 32 bits. Porting applications from other platforms has been demonstrated to be easier than on previous architectures. For backward compatibility, a 'default 64' mode is available for Fortran.

Fortran adheres to the Fortran 90 standard and supports select Fortran 2000 features such as Fortran/C interoperability.

C conforms to the C98 standard. With C++ Cray provides the Dinkumware class libraries.

C, C++ and Fortran utilize the Cray PDGCS back-end technology under development for several Cray product generations. PDGCS support for vectorization has been enhanced to take advantage of the large Cray X1 system cache and register set. Scratch memory use for temporary storage is greatly reduced with the compiler enhancements.

Libraries

Math and scientific libraries that deliver a high percentage of Cray X1 system peak performance are available for ease of performance optimization. Scientific libraries include FFTs, BLAS, LAPACK and SCALAPACK.

Visualization libraries include X11R6 and Motif.

IO libraries include Cray's ffio for high performance local data cache management.

Debugging and Performance Analysis

Cray offers Etnus TotalView as the default debugger for the Cray X1 system. Cray has enhanced TotalView to support Cray X1 system features, and provides first line support for the product. Cray tracks ongoing TotalView releases as they are made available.

For performance analysis, Cray offers the Cray Performance Analysis Tool (CrayPat). CrayPat combines the functionality of previous Cray performance analysis tools such as flowtrace and perftrace into one package. CrayPat is also the main tool for accessing the many hardware performance counters on the Cray X1 system.

Operating System Components

UNICOS/mp is an evolved operating system that has descended from UNICOS (Cray SV1 and Cray T90) and UNICOS/mk (Cray T3E). Its architecture closely parallels that of UNICOS/mk in that select operating system functions run on application nodes and most system services run on system nodes.

UNICOS/mp is a full-featured, POSIX-compliant operating system that supports system scalability to hundreds of mult-stream processors or thousands of single-stream processors. Also available are POSIX commands and utilities, several shells (csh, ksh, sh, tcsh), many UNICOS commands and the Cray Open Source (COS) package. COS includes commonly available tools such as perl, bash, tcl, cvs and many others.

UNICOS/mp includes support for high performance file IO with features such as direct IO and asynchronous IO. The XFS file journaling file system manages terabytes of data. Volume management supports RAID disks and automatic failover.

Like UNICOS and UNICOS/mk, UNICOS/mp includes support for many production management features shown here:

Cray Production Management Environment

Cray Application Placement (psched)	Cray Checkpoint/Restart	Cray Accounting
• Application Migration • Gang Scheduling • Pre-emption • Allocation Control	• Application, system or user initiated • Restart in different placement • Distributed applications	• Standard SVR4 • Project accounting (acids) • Process limits
Cray Resiliency Management • Module addition • Down CPUs • Impact only apps with resources assigned	**Cray System Partitioning** • Run multiple OS versions • Operate from CWS	**Cray Security** • National Information Assurance Partnership (NIAP) EAL2 security assessment.

Figure 3: UNICOS/mp Production Management Features

Of special note, the Cray Application Placement scheduler (psched) is an evolution of the UNICOS/mk psched. Psched interoperates with Altair PBSPro to manage optimal placement of applications in Cray X1 system nodes. It supports application migration to ensure high utilization as well.

Partitioning

New with the Cray X1 system is support for system partitioning. Cray X1 system nodes may be grouped into arbitrary partitions that can then managed as independent system images, each with its own copy of UNICOS/mp. Customers wishing to evaluate new operating system releases may run different versions of UNICOS/mp on different partitions for testing and performance comparison.

UNICOS/mp provides a high degree of interoperability for Cray X1 systems in a computing environment. The Cray Network Subsystem is a special routing box that delivers high performance Gigabit Ethernet connectivity to and from Cray X1 systems. For larger block sizes, transfer rates are a high percentage of peak media numbers.

Storage Area Network

New on the Cray X1 system under UNICOS/mp is the ADIC StorNext   SAN file system. StorNext allows different vendor systems to share files directly attached to the Cray X1 system at high data transfer rates. A suite of StorNext data management products such as backup and migration services is also accessible from the Cray X1 system via the StorNext SAN.

Interoperability features are summarized in the figure below.
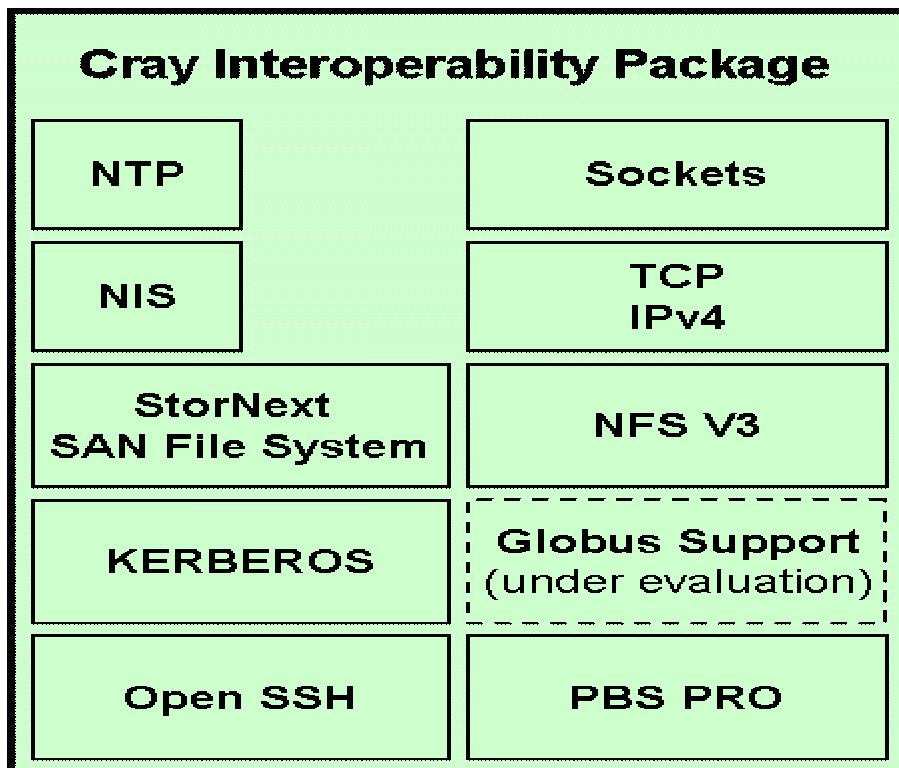
Figure 4: Interoperability Features

Planned 2003 Software Releases

The following overviews the releases planned for 2003.

Programming Environment

PE 5.0 (est. June 2003)

- ∑ SSP mode support

- ∑ SUN hosted cross-compilation package

- ∑ Libsci enhancements adding SCALAPACK and BLACS and performance improvements for BLAS, FFTs and LAPACK

- ∑ Improved formatted IO performance

- ∑ Loopmark for C

PE 5.1 (est. September 2003)

- ∑ OpenMP

TotalView

- ∑ GUI interface (May)

Distributed programming model support (September)

Motif

- ∑ Library support (June)

MPT 2.2

- ∑ SSP mode support (June)

- ∑ MPI-2 extensions (June)

- ∑ RMA derived data handling

- ∑ Send/receive "in place"

UNICOS/mp

UNICOS/mp 2.2 (July 2003, estimated)

- ∑ Initial 128 MSP support (general 128 MSP support will be available in an August 2.2.x update)

- ∑ Psched support for PBSPro application limits

UNICOS/mp 2.3 (est. October 2003)

- ∑ Initial 256 MSP support

Σ   ADIC SAN support on Cray X1 systems (only) with Fabric support

Cray X1 System Software Documentation

The following manuals are available for Cray X1 system software. Four of the manuals: the Cray X1 System Overview, the Cray X1 User Environment Differences manual, the Migrating Applications to Cray X1 Systems manual, and the Cray X1 System Administration Differences manual will be of special interest to Cray customers with Cray T90, Cray SV1 or Cray T3E systems. These manuals provide comparisons between Cray X1 system software with that available on the previous systems.

There are over 40 manuals currently available for Cray X1 software. The Cray X1 System Overview provides an overview of Cray X1 hardware and software capabilities, and points readers to other documentation for complete information.

Cray customers with Cray SV1 or Cray T3E systems may find the following manuals interesting since they document differences between Cray X1 software and software available on the older systems:

The Cray X1 User Environment Differences

Cray X1 System Administration Differences manuals

Migrating Applications to Cray X1 Systems provides information on migrating applications that currently run on Cray T3E or Cray SV1 systems to applications that run on Cray X1 systems.

Optimizing Applications on the Cray X1 System documents techniques you can use to analyze and optimize the performance of Fortran, C++, or C code on Cray X1 systems.

Manuals

Cray X1 System Overview

Optimizing Applications on Cray X1 Systems

Cray X1 User Environment Differences

Cray Programming Environments Release Overview and Installation Guide

Cray C and C++ Reference Manual

Cray Fortran Compiler Commands and Directives Ref. Manual

Fortran Language Reference Manual, Volume 1

Fortran Language Reference Manual, Volume 2

Fortran Language Reference Manual, Volume 3

Cray Fortran Co-array Programming Manual

Cray X1 Application Programmer's I/O Reference Manual

Cray Message Passing Toolkit Release Overview

Migrating Applications to Cray X1 Systems

Dinkum C++ Library Documentation

Etnus TotalView Getting Started

Etnus TotalView New Features

Etnus TotalView Reference Guide

Etnus TotalView Users Guide

TotalView Release Overview, Installation Guide, and User's Guide Addendum for Cray X1 Systems

FLEXlm End Users Guide

Cray Assembly Language (CAL) for Cray X1 Systems Reference Manual

Cray Open Software Release Overview/Installation Guide

Cray X1 System Administration Differences

UNICOS/mp General Administration Guide

UNICOS/mp Networking Facilities Administration

Cray X1 System Configuration and Operations Guide

UNICOS/mp Disks and File Systems Administration

UNICOS/mp Resource Administration

UNICOS/mp Release Overview

UNICOS/mp Installation Guide

Cray Network Subsystem (CNS) Overview/Installation Guide

Cray Workstation (CWS) Installation Guide

Cray Workstation (CWS) Release Overview

Cray Programming Environment Server (CPES) Release Overview/Installation Guide

Cray Programming Environment Server (CPES) Administration Guide

PBS Pro 5.3 Administrator Guide

PBS Pro 5.3 User Guide

PBS Pro Release Overview, Installation Guide, and Administration Addendum for Cray Systems

PBS Pro 5.2 External Reference Specification, PBS-3BE01

Common Installation Tool (CIT) Reference Card

CrayDoc Installation and Administration Guide