UNICOS/mk Status and Update

Jim Harrell, System Software Division, Cray Research, A Silicon Graphics Company, 655F Lone Oak Drive, Eagan, Minnesota 55121

ABSTRACT: *The UNICOS/mk operating system has been introduced recently on the CRAY T3E mainframe. This operating system is a modularized version of the UNICOS operating system and is composed mainly of code from UNICOS. The following status and update describes the current project directions, the release plans, feature plans, and the status of the system in the field.*

1 Introduction

Since the last Cray User Group (CUG) Conference in Barcelona, there have been changes in the plans and status of the UNICOS/mk project. The merger with Silicon Graphics Inc. (SGI) and project delays have caused some of the changes. The following review of the changes and current plans is divided into four sections. The first section reviews the changes in project focus since the last CUG conference; the second section reviews the current release plans; the third section describes the feature contents of the releases; and the fourth section describes the status of the systems in the field

2 Changes in Project Focus

The original UNICOS/mk project direction had been first, to provide a system for the CRAY T3E mainframe; second, use UNICOS/mk on future systems; third, migrate UNICOS/mk to the parallel vector (PV) systems over several releases. The value of having a single operating system on all platforms makes it easy for users to move from one platform and Cray Research also benefited by being able to concentrate on a single software system.

After the merger, a review of future plans took place. The plans for future Cray systems (referred to as scalable node (SN)), were similar to the plans for future systems at SGI. Thus, the SGI SN and Cray SN plans have been merged, which includes the decision to use MIPs-based processor and SGI's IRIX operating system.

To use IRIX on the SN systems, SGI has plans to modify it to produce a system called Cellular IRIX. This system is built with a client/server capability. It is already under development, and supports MIPs processors and the MIPs ABI. UNICOS features and functionality required by Cray customers will be ported to the Cellular IRIX system.

This decision led to a review of the need to port UNICOS/mk to PV systems. The current UNICOS used on PV systems is a full-featured system that is extremely stable. In light of the previously described merger plans, replacing this system does not provide obvious value to customers, thus the plans to provide UNICOS/mk on PV systems has been dropped. Cray Research will continue to support UNICOS on the PV systems.

The CRAY T3E system has just begun to ship and has a series of hardware and software enhancements scheduled for the next several years.UNICOS/mk will remain the system for the current CRAY T3E system and for enhanced versions of the CRAY T3E systems.

3 Release Plans

The current UNICOS/mk release is 1.2.4. It is actually a pre-release that is being made available to all CRAY T3E systems. The 1.2.4 system life will be short. Over the next six months Cray expects to have frequent releases and updates.

The UNICOS/mk project will do this by using a modified UNICOS IDS process, which allows us to quickly deliver features and fixes to CRAY T3E customers. Because of the frequency of releases, the support of a release or update is limited to the current release until the UNICOS/mk 1.5 release.

The release plans are based on initially providing functionality and reliability. Later releases will focus on UNICOS compatibility (later meaning within the next 6 to 9 months).

The current release schedule is for a 1.3 release at the end of December 1996. The 1.4 release is scheduled for mid-February 1997. The current plans call for the 1.5 release to be shipped in

May 1997. The next release is tentatively called 1.6 and is planned for a 4th quarter 1997 release.

The 1.2 release, which is the feature version of the 1.2.4 release, is a stand-alone UNICOS system with support for the CRAY T3E system. The system is capable of running interactive, and batch in a serial mode like UNICOS where individual processes run on separate PEs. Simultaneously the system can run multi-PE applications in partitions of PEs allocated for the purpose of running multi-PE applications.

The main services of the operating system has been distributed across several PEs called OS PEs. The serial and application PEs also have operating system servers which provide many common services and manage the local PE hardware. The distribution of services has been aimed at providing scalability by doing as much as possible at the local PE.

The feature list for the UNICOS/mk 1.2 release is as follows:

- Configurations: scalability up to 256 PEs
- Process management
- Basic scheduling
- Common commands
- Signals
- UNICOS filesystem (NC1)
- Kernel statistics
- System administration tools
- Terminal/console
- Boot/dump/monitor
- UNICOS accounting
- Kernel logging of system messages
- Resource management
- User database
- Basic limit support
- Initial tape support
- User-level striping
- Distributed I/O system call (distio)
- File system assistant (I/O Scalability)
- Parallel disk and packet servers
- TCP/IP
- NFS server and client
- NIS/ONC
- CVT 3.1

The 1.3 release, scheduled for the end of 1996, contains fewer features and is devoted mostly to reliability improvements. However, there are several significant features that should be mentioned.

Support for process migration appears in the 1.3 release. This feature can be used to help move processes, including multi-PE applications to different PEs. Initially, this will be a manual process, but will be the basis for automatic load levelling in the 1.4 release.

Another important feature is pcache. This is the replacement for ldcache in UNICOS and provides caching for physical disk devices that substantially improves performance.

The complete planned feature content for the UNICOS/mk 1.3 is as follows:

- PM process migration
- Shared-text support
- TCP/IP commands & daemons (phase 2)
- Socket server performance enhancements (phase 1)
- OS caching (pcache)
- SNMP
- Kerberos V4 Support
- ToolTalk
- Tape subsystem serverization (phase 2)
- The UNICOS/mk configuration tool (CT)
- Configuration server: inform registered servers of upgrades
- Configuration server: correctness verification
- Configuration server: consistency checking
- Configuration server: multiple server instances
- New online tape diagnostic - VTT

The UNICOS/mk 1.4 release, scheduled for mid February of 1997, provides substantial progress towards UNICOS compatibility. In this release disk quotas, DMF, and initial MLS functionality are provided. This release also provides additional scalability features such as remote mount, which will allow multiple file servers on a CRAY T3E system.

The current feature list for the UNICOS/mk 1.4 release is as follows:

- Political scheduling
- Remote mount
- Disk quotas
- Shared File System (SFS) support
- Serverized DMF support
- Socket server performance enhancements (phase 2)
- TCP/IP over GigaRing (host-to-host)
- Configuration server: dynamic configuration
- Enable MLS functionality: ACLs, privilege management, credential management
- Disk server minor device number support
- F-Packet support in hdd driver
- Enhanced hardware error reporting
- Initial hardware error recovery
- T3E diagnostic and client

The UNICOS/mk 1.5 release completes the port of UNICOS functionality into UNICOS/mk. In this release, the check-

point/restart feature is provided, resource limits support is completed, as is MLS support, and support for DCE/DFS is provided.

The current feature list for the UNICOS/mk 1.5 release is as follows:

- Checkpoint/restart
- Complete resource limits
- MLS: mandatory access control (MAC) and networking authorization list (NAL) support
- SysV IPC message queues and semaphores
- Support for hardware enhancements
- DCE/Core services
- DCF/DFS client and server
- Cray ReelLibrarian (CRL)

The next UNICOS/mk release after 1.5 is tentatively scheduled for the 4th quarter of 1996. The feature content for this release has not yet been completely defined.

Currently the expectation is that features in the areas of system resilience and system performance will be provided as our knowledge and experience with the CRAY T3E system continues to grow.

The ability to provide multiple system software partitions has been discussed for this release. System partitioning will provide a functionality like UNICOS Guest for the CRAY T3E system. The difference is that instead of running multiple versions of the OS on the same CPUs, the CRAY T3E system partitioning will allow PEs be allocated to different partitions and different versions of UNICOS/mk to be run in each partition.

Another group of features being studied for possible inclusion into this release are IRIX features that are going to be ported to UNICOS and UNICOS/mk. These features are being ported in order to help customers transition to the future operating systems that will be provided by Cray Research and Silicon Graphics.

4 Field Status

At this time (the middle of October 1996), Cray Research has shipped 18 CRAY T3E systems. The smallest is 8 PEs and the largest is 256 PEs. Customers are successfully running distributed applications, a native programming environment, and simultaneous interactive and batch environments. There are some problem areas, specifically boot/dump, install/configuration, and some CRAY T3E hardware-dependencies.

UNICOS/mk reused a large percentage of UNICOS code. Cray believes that being able to reuse so much of UNICOS will give UNICOS/mk a substantial boost towards reliability because UNICOS is so reliable.

In many areas of the operating system, the UNICOS and UNICOS/mk sources are identical. The parts of the operating system that tend to be the most alike are the “upper layers”. These layers represent a substantial part of the operating system. The upper layers are also not hardware dependent, so they tend to remain constant across hardware platforms. Examples are the NC1 filesystem and socket support.

For the UNICOS/mk project, Cray took the approach of “pipelining” the software development. The aim here was to verify as much of the system as possible on other machines before trying to run it on the CRAY T3E. The operating system was first merged with the Chorus microkernel on a CRAY YMP. The programmers were familiar with both the CRAY YMP and with UNICOS, so that the new portion of the system was limited to one variable.

The UNICOS/mk system was then ported to a CRAY T3D to develop and verify the MPP features needed for the CRAY T3E system. The variables in this port were the CRAY T3D and the parallel features.

UNICOS/mk has been running on CRAY T3Ds in “batch mode” since June of 1995. Extensive testing was performed to verify that the system interface was still the same as the UNICOS interface and that the new MPP features functioned correctly.

The last phase of the pipeline was the port the code to the CRAY T3E. This port concentrated on the CRAY T3E hardware dependencies, the new system workstation (SWS), and GigaRing support.

Cray’s diagnosis of the current problems is that they seem to be located in areas associated with this last port to the CRAY T3E. The prognosis for the future is that the system will stabilize quickly because the hardware dependent layers are critical, but a small proportion of the total system. The upper layers of the operating system have had extensive checkout and should be much more stable. Cray is working hard to ensure that these systems meet expectations.