Summary of OSF RI System Technology Accomplishments for 1992 and Directions for 1993

IRA GOLDSTEIN

December, 1992

1. System Technology Objectives and Accomplishments

In 1989, the RI initiated a program to realize the full potential of Mach technology as a foundation for open systems. The goal was to employ Mach as a modular, extensible architecture capable of providing an economic and common base for multiple operating environments. The vision was to unify open and proprietary environments with a common foundation that delivered new benefits to the end user including real-time responsiveness, higher levels of trust, and scalability to clusters and massively parallel systems. To understand the algorithms and data structures needed to achieve this new system technology foundation, the RI has been developing two advanced prototypes: OSF/1 MK and OSF/1 AD.

1992 was an exceptional year. At industry conferences during the second week in November, RI technology was being demonstrated as the basis for new operating systems on platforms ranging from PS-2's to massively parallel supercomputers.

OSF/1 MK rearchitects OSF/1 into a pure Mach kernel, or microkernel, and a user space server. The goal has been to obtain a modular and scalable framework with maximum code re-use. This latter objective was established to maximize quality and minimize time-to-market. Ultimately, the OSF/1 server may well be decomposed into smaller, functional services but we see this as an incremental task for the future. The MK project began in Fall 1990.

OSF/1 AD is a set of extensions to MK that support scalability to massively parallel computers. These are systems that can support 1000's of processors on a high speed communication mesh. Both multicomputers and clusters represent examples of this kind of hardware system. To insure scalability, there is no requirement for shared memory (as there is for symmetric multiprocessors). New algorithms were developed in cooperation with CMU to enable the system to scale successfully and present the illusion of shared memory and a single process space to users. The AD project began in Summer, 1991.

The RI's Mach research has been exceptionally successful. 40 sites have received the MK technology. One of the most significant transfers has been to IBM. The PC group at

Boca Raton is exploring the use of Mach as a software architecture for future low-end PS-2 systems. At the fall Comdex, IBM demonstrated a multi-personality implementation that employed the OSF/1 server for the Unix personality and an IBM implementation for the OS-2 and Windows personalities. This design was proposed to IBM by the OSF RI in 1990. IBM has received very positive reviews by the industry press. The OSF/1 server has worked sufficiently well that it is being employed as the programming environment by IBM and IBM partners to develop new Mach device drivers and additional personalities. Under legal procedures established by IBM to insure unencumbrance, the RI is presently adding its two years of bug fixes and functional enhancements to Boca's version of Mach in order to accelerate their advanced development efforts. This includes the RI's real-time extensions to Mach.

In the same week that IBM was demonstrating its low-end Mach-based environment, Intel Supercomputer Systems Division was demonstrating the first generation of OSF's scalable AD technology at the Minneapolis Supercomputer Conference. Intel will be shipping AD release 1.0 as the operating system for its Paragon multicomputers in 1993. Convex has made a public commitment to this technology and Hitachi has assigned sabbaticals to the RI to assist in its development. Many companies are presently assessing it on a confidential basis.

The MK and AD prototypes from the RI do not achieve product quality. But they are sufficiently high quality that early adopters can move quickly to market with a program that completes their industrialization. This is a reasonable paradigm since there is an inevitable amount of customization that is needed to meet the specialized needs of different vendors,. Technical development and actual market experience provided by early adopters also make any subsequent OSF Engineering program less expensive and less risky.

To insure portability, the RI prototypes have had their interfaces carefully documented. This documentation is widely used by the community. CMU distributes it and IBM is using some of this documentation as a starting point for its own system.

2. Funding

The RI obtained substantial supplementary funding to augment the base level of funding provided by the OSF sponsors. Hardware systems, sabbaticals and/or cash were contributed by Compaq, Convex, Cray, Digital, Hewlett Packard, Hitachi, IBM and Intel. In addition, a major research proposal was accepted by DARPA for the development of advanced operating system technology for high performance computing.

One purpose of the National Cooperative Research Act (under which OSF is organized) is to encourage economy through common efforts. The RI has succeeded well in this dimension. It has provided a highly leveraged site at which vendors and government agencies can invest their research dollars.

3. Communication

The RI hosted a symposium, two workshops and two tutorials in 1992. Over 200 engineers attended the Feb. Symposium on Mach and over 80 attended the July workshop on High Performance Computing. A January workshop took place in Europe at the RI's Grenoble site. At these meetings, the membership had an opportunity to articulate their requirements and their critique. The RI listens closely and adjusts the program as needed. Real-time has become an important goal as a result of feedback from the OSF membership.

4. Recognition

An RI staff member was asked to chair the 1993 Mach Usenix Conference and another RI staff member is the local arrangements chairman. The RI has also been invited to participate in Unix International's working group on system architecture and in DARPA's HiPer-D program to develop high performance distributed data acquisition systems. We were also asked to organize and chair panels on system technology at DARPA's annual technical meetings, at Uniforum and UnixExpo.

5. 1993 Goals

System technology is critical to harnessing the power of new, parallel architectures and faster interconnects. Insuring that our architecture is extensible and can satisfy critical end user requirements will be our primary focus. Here are three areas that are already on our agenda for 1993:

- *I/O processing speed commensurate with system architecture.* I/O performance for supercomputers is very demanding. For 1993, we plan to deliver a new faster IPC implementation, a revised file system architecture, and a framework which preserves the modular structure of AD without the requirement for (and performance penalty associated with) separate address space servers.

- *Real-time response.* We plan to execute Phase 2 of our real-time program by making the kernel preemptive, supporting local real-time IPC, and synthesizing outside real-time developments from some of our collaborators.

- *Fault detection and recovery.* High availability is a multi-year effort which we will start in 1993 with an emphasis on fault detection, notification and general robustness.

Our major efforts in 1993 will be directed at insuring that early adopters of the RI technology are successful. We will be analyzing and improving our scalability architecture based on actual experience from systems running >100 processors. We will also be studying the performance of multi-personality versions of Mach-based systems and proposing architectural enhancements to achieve improved performance and reliability.

In addition we will initiate a few activities which are targeted for mid to late 1994 delivery:

- We will put in place a rigorous software engineering program to define, clarify, and begin the re-engineering the Mach internal availability. We expect to work collaboratively with NSA, security vendors and equipment vendors.

- We will begin a research program aimed at extending OSF's architecturally neutral distribution format (ANDF) for massively parallel architectures.

- If funding permits, we will conduct preliminary investigations into the relationship of open systems to a new generation of user interfaces known as *virtual reality*. We expect that this may be a new area for the RI in 1994. It meshes well with the RI's expertise in scalable system technology. Virtual reality environments require a great deal of computation. We will investigate how to apply virtual reality to improve the user's ability to "mine" information and manipulate simulations. Such investigations may also contribute to the development of standards for interoperability among distributed systems that support a virtual reality interfaces.

6. Conclusions

The RI has contributed to a redefinition of the entire industry's approach to operating systems and delivered prototypes that are being used as the basis for real products by key open systems companies. It has accomplished these goals by means of broad support from industry and government. By any measure, the RI had an exceptional year. We expect 1993 to be equally productive.