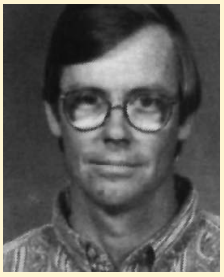
The Architecture Behind AIX 4.1

More than a release, AIX 4.1 is a new version of AIX. It provides additional functions and capabilities in many significant areas—from kernel and application programming interfaces to the "look and feel" that end users will experience—while retaining the functionality and application binary compatibility with AIX Version 3 on which existing customers depend.

We asked Michael Day, a lead AIX system architect in Austin, Texas, to tell us about new features in AIX 4.1 and their implications for developers and customers.

What was your overall objective in designing AIX 4.1?

Day: The overall theme of AIX 4.1 development was to make the AIX experience more appealing to a greater number and more diverse set of customers and markets. We wanted AIX to continue to be the operating system of choice for the existing POWER-based systems and to be a player in the wave of new PowerPC™-based products being introduced by IBM and others. AIX 4.1 is well positioned technically to take advantage of IBM's new Power Personal Systems line of products as well as the upcoming high-end Symmetric Multiprocessor (SMP) systems. While AIX was already well optimized to take advantage of the POWER Architecture™, additional optimization has been added to leverage the family of processors based on the PowerPC Architecture™.



Michael Day

It is quite an undertaking to build one operating system that covers what IBM calls the "laptops to teraflops" range of machines. How have you provided a good operating system for each machine, while avoiding additional overhead costs of supporting a range of platforms?

Day: We started with a good base. AIX Version 3 was already designed with scalability and system flexibility in mind with its use of large virtual address spaces, pageable kernel, dynamically loadable kernel extensions, and self-configuring device and subsystem support. This capability was extended to deal with more diverse system designs required to handle the expanding family of POWER and PowerPC processors, and differing system buses and I/O hardware used by systems being introduced by IBM and others.

What are the key enhancements in AIX 4.1?

Day: I have separated the enhancements and new functions into six categories. From the bottom to the top of the operating system, they are as follows:

- Enhanced kernel functions
- Industry standards
- Application binary compatibility
- Enhanced system management capabilities
- New graphical end-user environment
- New packaging

Let's start with the enhanced kernel functions.

Day: Since AIX 3.2.5, AIX has had platform-specific modules for each processor chip architecture. In Version 3.2.5, we added support for the POWER2 processor and the PowerPC 601™. At boot time, AIX detects the type of system, and then overlays and fills in branch tables for the platform-specific modules required by the system. No memory is wasted by holding unneeded code. In AIX 4.1, we added modules for the

PowerPC 603™ and 604 chips, the upcoming PC-based Power Personal Systems line of products, and the upcoming SMP systems.

Kernel changes were also made to support 2- to 8-way symmetric multiprocessor systems. This is classic SMP support—fully symmetric I/O with shared memory. We used the IEEE POSIX™ 1003.14 MP-specific commands for system administration and management.

At the same time, we implemented threads within the kernel, creating a one-to-one relationship between user threads and kernel threads with the pthread library based on POSIX 1003.4A (Draft7). This should be beneficial to object-oriented, multimedia, real-time, and database programming.

In standard UNIX®, if you had multiple threads of execution, you had to fork and exec processes, using pipes or shared memory to communicate between them. With thread support in the kernel, these applications no longer need that overhead. Developers can create multiple threads of execution within one process, and the threads have access to all the same data and resources. In addition, this change makes it easier for multithreaded applications to exploit multiprocessing. Each thread can automatically be dispatched to the next available processor, since the kernel now handles the execution priorities and dispatching at the thread level.

We have done a lot to parallelize the critical kernel subsystems in anticipation of an SMP machine—including Logical Volume Manager (LVM), Virtual Memory Manager, filesystems, subsystems such as Streams and tty, and many device drivers. Since AIX 4.1 has implemented locking and serialization strategies, many existing applications can run efficiently on an MP system.

We have also replaced our home-grown extensions of the Berkeley Software Distribution (BSD) tty subsystem. AIX 4.1 uses a new tty subsystem based on UNIX System V.4-compatible Streams that makes it easier to port Streams-based applications to AIX. AIX 4.1 still supports the AIX 3.2 tty functionality from an applications perspective.

We have added two new features in the Journaled File System (JFS). One is called *fragmentation support*. Previously, disk space was allocated in blocks of 4 KB. This method often wasted space on small files or files with final blocks containing only a few bytes. AIX 4.1 allows the system administrator to set up fragmented filesystems so that the last block of a file can be allocated in 512-byte blocks instead of 4 KB, resulting in much better disk utilization. In fact, the `/home` and `/var` filesystems are installed as fragmented by default.

The other new feature is an on-the-fly software data compression/decompression option that achieves about a 50% reduction in disk space requirements. It uses a fast LZ1-type algorithm written at IBM that is similar to compression techniques available on personal computers. Even with software compression enabled, performance is very respectable.

Other kernel enhancements include the following:

- Increased filesystem size up to 64 GB (individual files must still be less than 2 GB)
- Support of LVM disk striping for increased performance on very large files
- Improved CD-ROM support including Rockridge filesystem format, photo CD, and multisession capabilities
- Support of NFS 4.2, providing NFS updates and SMP support
- Improved application link times by 2X to 10X
- The newly adopted Open Software Foundation® (OSF®) Common Data Link Interface (CDLI) for LAN-based device drivers enables one driver to support protocols over both sockets and Streams (Previous device drivers could only deal with one or the other; the new interface provides much better path length, and drivers are much smaller.)

AIX 4.1 is well positioned technically to take advantage of IBM's new Power Personal Systems line of products and the high-end SMP systems.

Let's talk about new support for standards.

Day: There are several significant features of AIX 4.1:

- It supports POSIX 1003.1 systems interfaces.
- It supports POSIX 1003.2 shells and commands.
- Thread support is based on Draft 7 of the POSIX 1003.4A threads specification. The upcoming AIX 4.1-based Distributed Computing Environment (DCE) support will provide a mapping layer to convert from Draft 4 interfaces to Draft 7 interfaces.
- AIX is designed to be XPG4 base branded. XPG4 is a superset of the POSIX 1003.1 and 1003.2 standards. It provides more commonality in commands and interfaces, and better internationalization support than XPG3.

- AIX 4.1 is designed to conform to the emerging X/Open™ SPEC™ 1170, which is intended to provide a common UNIX interface specification.

Whenever IBM talks about change, users and developers worry about backward compatibility. What's the situation there?

Day: Since AIX 4.1 is a new version of the operating system with extensive enhancements to core technology, binary compatibility with existing applications was a major development and test theme for the product. For example, we had shifted to System V.3-compatible libcurses and X11R5 support, but we found that many existing application binaries depended on the previous libcurses, X11R3 and X11R4 interfaces. To handle this problem, we added a new Load-Only object type to shared libraries, so that we could transparently support the old interfaces for application binaries while ensuring that any applications rebinding would access only the new interfaces.

We have also identified old commands, links, and libraries that are non-standard or obsolete and have packaged them into optional compatibility packages. If AIX 4.1 is installed as a migration from AIX 3.2, these compatibility packages will be automatically installed to ensure a smooth migration for current customers. However, we do request that application providers refrain from relying on these facilities when releasing new versions of their software.

Although we have an excellent compatibility story for applications executing in the user space, the compatibility picture is not as good for applications with complex kernel extensions. Because of the restructuring of key components of the kernel for threads and SMP support, these complex kernel extensions (such as filesystems, complex device drivers, and communications subsystems) will require a moderate level of porting. "Well-behaved" character and block device drivers—those that do not directly access areas like the U area or `proc` structures, should run without trouble. On SMP systems, we will even run them *funneled,* that is, bound to a specific processor.

We have tested more than 90 of the key AIX 3.2 applications on AIX Version 4 with very good success. We believe this achievement will help quicken software vendor certification of applications on AIX 4.1.

The system management facilities and customization capabilities of AIX 4.1 have been greatly expanded.

System management has always been a strength of AIX. What's new with AIX 4.1?

Day: The system management facilities and customization capabilities of AIX 4.1 have been greatly expanded to support a more diverse set of users with the upcoming introduction of the Power Personal systems. Some system management and GUI enhancements are as follows:

- Repackaging of AIX for custom installation
- Simplified and faster operating system installation
- Installation assistant (graphical and ASCII versions)
- Graphical visual system management
- AIXwindows® Desktop, based on Common Desktop Environment (CDE) technology

These enhancements, along with many others, are a direct response to our customers.

One of the larger efforts in the development of AIX Version 4 was to completely restructure the packaging of the operating system components to support custom installation and improved modularity. Our objective was to automatically install the minimum operating environment for the system being installed, and then provide the facilities for the customer to customize the system by installing additional operating system components when needed.

In AIX Version 3, all the RISC System/6000 device support was installed on all systems, even if the hardware was not present or supported on the system. As the number of systems supported by AIX expanded and diversified, this procedure negatively affected software installation time and disk space usage. To alleviate this problem, AIX Version 4 now automatically installs a customized minimum operating system environment. The device configuration subsystem has been extended to automatically install only the device packages required to support the installed hardware. Custom install will determine if a minimum graphical environment or an ASCII-only based environment should be installed. It will also automatically install the country-specific message packages, locales, and helps for the primary language of the system to support the base operating system and all previously installed applications.

Installation of the base operating system is much simpler and faster than it was in AIX Ver-

sion 3. Now with fewer prompts and screens, it supports an optional no-prompt install using a data file on a diskette. AIX Version 4 supports four installation modes:

- **Overwrite:** Completely re-installs the system
- **Preservation:** Re-creates the root and user filesystems and leaves everything else alone
- **Migration:** Saves all the user's files, configuration information, and applications, then migrates the installed operating system components to the latest level
- **Installation:** Restores the system from a previously created system backup tape

A new Network Installation Manager (NIM) supports centrally administered remote system installation of networked systems. This method offers flexibility for centralized operating system and application installation through push installs to one or more network clients, pull requests from clients, and support of diskless and dataless clients from network servers. NetView® Distribution Manager (NetView DM) also offers remote application installation; however, NIM expands this capability by supporting remote operating system installation as well. These installation capabilities are ideal for customers who have a large number of similar or identical installations. For these customers, a centralized shop can use NIM, system backup, or data-driven Base Operating System (BOS) install to support their clients.

Once the minimum operating environment has been installed on the system, the new Install Assistant is started. This facility helps users through the customization process by taking them through either graphical- or text-based menus to complete the installation and system setup. The graphical version provides drag-and-drop interfaces for tasks such as adding users, installing additional software, and so on. The Install Assistant is essentially a customization help system that launches the appropriate System Management Interface Tool (SMIT) or Visual System Management subtask to perform the required customization.

I previously mentioned the repackaging of AIX and AIXwindows and its custom install. To help customers choose what to install, we have provided a software install facility called *bundles*. Bundles are collections of installable operating system software components and Licensed Program Product (LPP) components that are grouped together and can be installed with one selection. AIX 4.1 supports both system-defined and user-defined bundles. We have created the following prepackaged bundles of the operating system and LPP components:

- **Runtime client:** Primarily runs end-user applications; requires 120 MB of disk space for the ASCII user interface and up to 160 MB for the graphical user interface
- **Server:** Contains components for server functions such as network name serving, routing, more than two concurrent users, directory services, service aids, and performance analysis tools
- **Application Development:** Provides the tools and header files necessary to develop AIX-based applications
- **Personal Productivity:** Provides components for using the AIXwindows Desktop, Windows Application Binary Interface (WABI), Ultimedia® multimedia services, and DOS diskette support

This bundle support depends on the `installp` subsystem provided in AIX Version 4. This subsystem supports the operating system and program product software packaging as a hierarchy consisting of program products that contain installable packages of updatable options or filesets. These filesets are named so that the customer can understand the function provided from a high level. For example, you will no longer find filesets called `bosext1` and `bosext2`, as in AIX Version 3. Instead, you will find filesets named `bos.acct` (accounting services) and `bos.net.tcp.client` (the client support for TCP/IP network services). Textual descriptions are also provided in addition to the fileset name. Enhancements have also been made to the install facility to provide preview and deinstall capabilities.

Installation of the base operating system is much simpler and faster than it was in AIX Version 3.

Does this new packaging and install help others—such as OEMs—incorporate AIX technology?

Day: Certainly. For example, a repackaged AIX is currently being used as part of the IBM 6611 router. I also believe that Apple® and other OEMs are planning to use AIX 4.1 technology in product offerings on PowerPC-based systems. This granular packaging enables OEMs and other IBM groups to utilize the components of the system without costly tailoring and repackaging to

remove unneeded components and add needed ones.

Even the AIX kernel is in its own separate package and offers two flavors: uniprocessor and multiprocessor. The two packages contain different kernel binaries that are compiled from the same source. The operating system auto install determines which package to install on the system. We worked hard to keep the system-specific functions (buses, device support, and diagnostics) out of the runtime environment package containing the core UNIX commands and interfaces. This enables the runtime to be more easily adapted to diverse system definitions that contain a POWER or PowerPC processor. Most system-specific functions are in individual device packages that are custom installed, depending on the target system configuration.

The repackaging of AIX into more granular functions and better understood dependencies provides a solid base for building an excellent update strategy.

What about operating system updates?

Day: The repackaging of AIX into more granular functions and better understood dependencies provides a solid base for building an excellent update strategy. Our early AIX Version 3 update process was less than pleasing to customers. Because the operating system packaging was very coarse, customers often applied updates to software components they did not use. The tools we used for determining co-requisites between packages were also very coarse and conservative. This meant that when updates were applied to fix a particular problem, much of the system itself was updated as well. This side effect, along with our procedure of tracking updates with Program Temporary Fix (PTF) numbers, resulted in lengthy update times and some risk of changing unrelated areas. Using PTF numbers to track the level of software updates was also difficult, because there was no easy way to determine which PTF numbers had superseded previous PTF numbers.

In AIX 4.1, we wanted an easy, effective way to track the update process. In this new AIX release, each installable entity has its own version, release, modification level (for maintenance), and fix level (customer-required change between modifications). We call it a "VRMF" scheme. A higher VRMF number supersedes lower ones. To identify the requirements for a certain fix, you simply need to know the APAR number, available from the product support center. The system uses the APAR number to determine the filesets and the minimum VRMF levels that must be installed to provide the fix. Once the APAR is selected, the system will then install the fixes. In addition, the `oslevel` command knows which filesets and the minimum VRMF level necessary to report that the system is operating at some known maintenance level or release level.

When will we see the next refresh of AIX?

Day: A refresh of AIX 4.1, planned for October, will include features such as additional SMP performance tuning, Streams-based tty performance improvements, additional system hardware support, and worldwide translated message packages. Although the full internationalization facilities are in AIX 4.1, only the English message packages were available for shipment in August. No changes to application interfaces will be made, so there should be no effect on ISV application certification assessments. For existing AIX 4.1 customers, handling of the refresh will be similar to applying a new maintenance level to the system, not a reinstall.

What advice would you give developers as they move to AIX 4.1?

Day: First, ISVs with existing applications on AIX Version 3 should determine if their existing releases can be certified on AIX 4.1 without any changes. We have tested over 90 key AIX applications to ensure operating system compatibility, but we do not certify these applications. We use them with our extensive test cases to flush out incompatibilities.

Here are some suggestions for ISVs.

- If you are planning a new release of your application on AIX 4.1, obtain a copy of the AIX 4.1 publication, *All About AIX 4.1,* from IBM.
- If your applications use complex kernel extensions, you may need to make straightforward modifications to port them to the new kernel environment. You should also make the kernel extension thread and MP safe.
- We recommend that you compile your applications in the default `COM` mode of the new AIX 4-based compiler so that your applications will use the common instruction set. This will enable your application to run efficiently on the existing POWER and PowerPC processors as well as future PowerPC processors.

- Be sure to look at the contents of the `bos.compat` and `X11.compat` packages and determine if you are relying on commands, links, or libraries that we have determined to be obsolete. If so, you should remove your dependency on them.
- Determine if your applications are dependent on compatibility options for commands and convert them to the new POSIX 1003.2 or XPG4 specifications.
- For help or advice concerning porting your application, we have an excellent staff at our AIX porting centers backed by the AIX development team.

What are some points of incompatibility between Versions 3.1 and 4.1?

Day: Complex kernel extensions must be ported to Version 4 because of changes in support for both the MP environment and threads. Extensions that directly access the U area and `proc` structures must also be modified. There are now services for accessing those areas safely.

Moving to the POSIX 1003.2 command and utilities standard changed a few flags and the functions of some commands. We were often able to add the new functions and keep the old ones, but there may be some deletions that particularly affect shell scripts.

Hardware vendors will find that the new CDLI capability makes developing and supporting LAN device drivers quicker, easier, and less costly than the previous Common I/O (CIO) model found in AIX Version 3. However, to take advantage of the performance and common support for Streams- and sockets-based protocols, they must be modified to utilize the new interfaces.

Was anything taken out in AIX 4.1?

Day: We have removed the High Function Terminal (HFT) support that provided multiple sessions on graphics terminals without AIXwindows. It was causing us problems in upgrading to new versions of Xwindows from the X Consortium. In addition, it would have required us to port it to the new Streams-based tty subsystem. We chose to implement a small tty emulator called LFT in place of HFT and concentrate on making AIXwindows more robust and reliable.

A set of packages called *compat options* (such as `X11.compat` and `bos.compat`) contains compatibility support code enabling some existing applications to run unmodified on AIX 4.1. Moving functions to these packages provides an early warning to developers that we plan to eliminate them in future releases of the operating system.

What sets AIX 4.1 apart from other versions of UNIX?

Day: Here are a few examples of the many features that make AIX 4.1 such a great operating system:

- Operating system and compilers highly optimized for performance on POWER and PowerPC processors
- Scalability—the pageable and preemptable kernel enables AIX to scale well
- Dynamically loadable kernel extensions
- Built-in Journaled File System (JFS)
- Built-in Logical Volume Manager (LVM)
- Extensive graphical and ASCII system management facilities
- Custom installation for loading only what you need onto your system
- Automatic device installation and configuration
- Comprehensive 2-D and 3-D graphic support
- Broad range of connectivity types and options

In this new AIX release, each installable entity has its own version, release, modification level, and fix level.

New PowerPC Magazine

PowerPC News, a complimentary publication from IBM, brings you the latest milestones and major successes of the PowerPC microprocessor architecture. The *PowerPC News* is published monthly. To add your name to the mailing list or for additional information, call 1-800-PowerPC. ■