Multi-OS Platform “i-PX9000”

TANIOKA Takahiro, KAWANABE Masazumi, MURAKAMI Shoichi

Abstract

Multi-OS platform i-PX9000 and various software products based on it make possible server integration in which servers running multiple open OSs including HP-UX, Linux and Windows can be accommodated in a single cabinet while inheriting the existing ACOS assets. This can lead to an open linkage system that can deal flexibly with changes in the job environment with reduced operation/management costs.

In addition to the above, this paper also describes the outline of the new system architecture including the ACOS Platform Technology, which is one of the key technologies of the i-PX9000.

Keywords

i-PX9000, ACOS platform technology , ACOS Access Toolkit, refam VX, OLF-DB

1. Introduction

Since its introduction in 1975, the ACOS-4, which is one of NEC’s mainframe computers, has made great progress in meeting the needs of the times including those for functional enhancement, specification expansion, as well as linkage and coexistence with open servers.

The environments of corporate information systems are progressing as we speak, and the enterprise servers assuming the key operations of private corporations are presently required to offer the flexibility to deal with changes in the information environments that are becoming increasingly advanced and complex, as well as the solidness that can ensure stable system operations.

The ACOS series i-PX9000 is the latest server platform that can meet the above changes in the information environments by means of the high reliability, high availability and compatibility with existing systems of mainframes that NEC has been cultivating for the past 30 years, the ACOS-4VX that is the latest ACOS OS compatible with i-PX9000, and the “multi-OS” environment that can integrate servers by natively running the open OSs such as the HP-UX 11iv2, Red Hat Enterprise Linux v.4 and Windows Server 2003.

2. Features of i-PX9000

The ACOS series i-PX9000 has been developed as the successor model to the previous parallel ACOS series i-PX7600 and i-PX7800.

The hardware platform uses the Intel Itanium2 featuring the most advanced 64-bit architecture and the chipsets developed originally by NEC. A powerful, rigid system infrastructure that the corporate information systems of the next generation deserve is built by introducing high reliability and high-speed interfaces at the same time as the latest open technology.

In the development of the i-PX9000, we put the emphasis on absolutely guaranteeing compatibility with existing systems. We newly developed an ACOS architecture virtualization technology called “ACOS Platform Technology (APT)” to ensure the compatibility (binary compatibility, operation compatibility) with past ACOS series at the same time as maintaining the high reliability and high availability unique to the ACOS that NEC has cultivated over many years.

Photo External view of i-PX9000 (Model SX).

The line of the base hardware models of the i-PX9000 include Model S that is the standard model and Model A that features high extendibility. When it is required to construct a system in which the open environment is integrated with the ACOS in a multi-OS environment, the open OS can be run by adding the "Open OS Loading Function" to a base model and expanding the hardware resources for the open OS.

In December 2005, we also started marketing of Model SX, which enhances the open OS extension function further in the performance of the Model S class in order to support more flexible system configuration (**Photo**)

3. System Architecture of i-PX9000

The system architecture of the i-PX9000 features interconnection of the CPU boards called "cells" and I/O boxes by means of crossbar (**Fig. 1**).

A cell consists of a CPU board that incorporates NEC's originally developed chipset in addition to four Itanium2 processors and memory. An I/O box can accommodate up to 14 I/O interface cards incorporating open technologies such as a high-speed optical loop channel based on the Fibre Channel Arbitrated Loop (FC-AL) with a transfer rate of 200MB/sec. and 1Gbps LAN (1000BASE-T/SX), and features the possibility of smooth linkage with open systems.

This system allows existing ACOS peripherals such as CGMT devices, WAN and printer devices to be used continuously because they can be connected to the IOPs (Input/Output Processors) of the I/O boxes through interface cards. Models A and SX can accommodate up to 8 sets of cells and I/O boxes respectively, and Model S can accommodate up to 2 sets of cells and I/O boxes respectively.

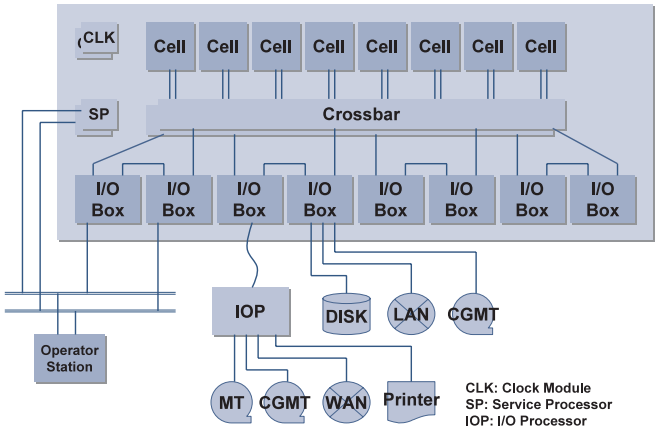

Fig. 1 System architecture of i-PX9000 / Model A.

The SPs (Service Processors) are the management boards, which are connected to the cells, I/O boxes and crossbar to attain operation management functions including initial hardware diagnostics, hardware configuration control and power supply control as well as fault processing for the main system hardware including fault log collection, log analysis and fault surveillance. With i-PX9000, they cooperate with the Operator Station that is connected to them through the network and the DGPFW (DiaGnostic Processor FirmWare) that is part of the ACOS Platform Technology to be described later in order to provide similar maintenance/management functions and user interface functions to the traditional ACOS systems.

From the viewpoint of high availability, all of the i-PX9000 models including Model S have built-in redundancy for the power supplies and fans. In addition, Model A provides redundant paths for data communications between the SPs, CLK modules, crossbar, cells and I/O boxes so that faulty locations can be isolated and operations can be resumed and continued in case of emergency. In addition, the communication paths for the Operator Station are also provided in dual redundancy.

For the RAS (Reliability, Availability, Serviceability) function, all of the data paths are capable of detecting or correcting 1-bit errors and the controller inside the chipsets has protection circuitry for parity, etc., to ensure high reliability. The chipsets also incorporate functions for higher availability, such as a mechanism for recovering the hardware uncorrectable errors of memory data using firmware or software, and the I/O recovery function that ensures continuous use of I/O boxes containing faults without having the system go down. Before running the ACOS system, these functions can be enabled in the chipset setup to support the advanced error recovery provided by the ACOS Platform Technology described below.

4. ACOS Platform Technology (APT)

The APT used in the i-PX9000 is the generic name given to the virtualization technology that implements an ACOS architecture on Itanium2-based hardware (**Fig. 2**).

(1) EPU (Execution Processing Unit) Virtualization Technology

In place of the NOAH (NEC One-chips ACOS Hardware engine) series of custom processors used in traditional ACOS CPUs, the APT employs the newly developed EPUFW (EPU FirmWare) that operates natively on Itanium2. The EPUFW runs on the individual processors that are assigned as the EPUs among the Itanium2 processors in the hardware, and executes at high speeds the ACOS-4 and applications' instruction codes stored in the ACOS MU (Memory Unit) that

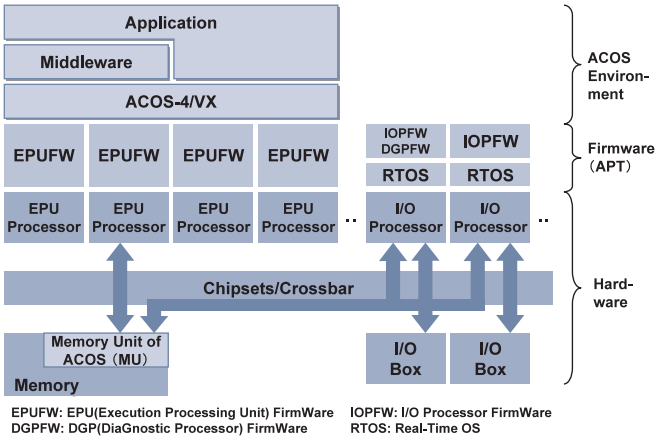

Fig. 2 Concept of ACOS Platform Technology.

is assigned in part of the hardware memory space, by converting the instruction codes into the native codes of Itanium2. This allows the ACOS-4 software environment implemented on APT to recognize individual Itanium2 processors as the EPUs of the ACOS architecture.

To execute ACOS instruction codes having the complicated CISC architecture at high speeds, traditional NOAH processors adopted a structure in which the ACOS instruction obtained from the memory is temporarily converted into a simple microcode sequence called a core instruction inside the processor before being executed. The relationship between each Itanium2 processor and EPUFW in APT can be regarded as similar to the relationship between the hardware and core instruction in the NOAH processors.

This construction has made EPUFW possible to implement the advanced RAS functions that have been available with the NOAH processors. The processor relief function is activated when an EPU fails to hand the register information, etc. inside it to another EPU for continuing the execution of the program being run on the EPU, and the instruction retry function makes it possible to retry execution of an ACOS instruction in the same EPU. Even in case of a fault with an EPU, it is switched to a reserve EPU to maintain the operation.

(2) I/O Processor (IOP) Virtualization Technology

Traditional IOPs incorporated special processors and firmware in them and have been configured as an independent system from the CPU system. When an input or output operation is activated by the ACOS-4 OS, the IOP performs the input or output from or to a peripheral device such as a disk by executing the input/output instruction sequence called the channel program in the MU.

With the i-PX9000, we accommodated the IOP function in the main system by newly developing the IOPFW (IOP FirmWare) that performs native operations on Itanium2 as part of the APT functions. Since the IOPFW needs to execute concurrently the multiple inputs and outputs to and from the channels (I/O cards) in the I/O boxes located under the IOPFW, it is packaged as a process on the RTOS (Real-Time Operating System) running on Itanium2.

In the development of the IOPFW for the i-PX9000, it was indispensable to provide the advanced RAS functions available with traditional IOPs as well as advanced I/O performance comparable to those of traditional systems. With the IOPFW, we succeeded in implementing a logic structure with extremely high resistance against I/O faults and achieved a high availability proper to ACOS by providing the recovery function in case of I/O operation fault and the continued operation function based on dynamic partial degeneration of the faulty position.

(3) Diagnostic Processor Firmware (DGPFW)

To provide the ACOS architecture implemented with the APT with operation management and fault processing functions, the DGPFW is provided as a process running on the real-time OS in the same way as the IOPFW. The DGPHW is connected to the Operator Station through the network in the same way as the SPs in the hardware, and assumes the configuration control of the EPUs and IOPs that are virtualized on the APT, the operation control such as startup of the ACOS-4 OS and the fault processing functions such as the fault log collection for firmware-detected faults and fault surveillance, in cooperation with the SPs.

5. Multi-OS Functions

(1) Partitioning Mechanism

The new feature of the i-PX9000 that is not found with traditional ACOS systems is the open OS execution function based on the hardware partitioning mechanism. Partitioning refers to separation of the entire system into a maximum of 8 partitions with Models A/SX and up to 2 partitions with Model S, by setting the cells and I/O boxes as the units of separation. The partitions can be separated so that interference of software will not occur between them by the hardware setup, so that each partition can launch and run an OS as if it is an independent computer. This multi-OS function makes it possible to integrate open servers including HP-UX, Linux and Windows servers in the same cabinet as shown in the example of server integration using Model SX in **Fig. 3**.

In the multi-OS integrated environment made possible by the multi-OS function, the allocation of hardware resources to

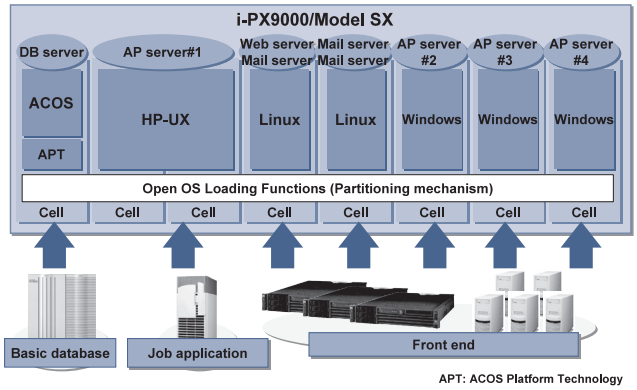

Fig. 3 Example of server integration of i-PX9000/Model SX.

each partition can be moved across the separation units such as cells. As a result, the system configuration can be changed flexibly according to fluctuations in job loads, for example by moving cells in the ACOS partition to the open OS partition or exchanging/leasing cells between open partitions.

(2) Software Product for Server Integration

At NEC, we have been releasing a series of open linkage products as the base products for use in the construction of systems that can make use of the advantages of both the ACOS machine featuring high reliability and abundant assets and the open server featuring high productivity in new job development, as the product group supporting the ACOS-open server linkage systems.

Traditionally, these products assumed that NX servers and Express servers are installed around an ACOS machine, and were targeted at the PA-RISC and IA32 processor-based servers.

As the multi-OS function of the new i-PX9000 incorporating Itanium2 has made it possible to run both the ACOS and open servers on a single i-PX9000, we decided to provide the open linkage products with compatibility with Itanium2-based hardware and make possible the convergence and integration of ACOS and open server mixed systems.

We first provided Itanium2 compatibility for the open linkage base products including the ACOS Access Toolkit (for transaction linkage), refam VX (for storage disk sharing), OLF-DB/EX and RX (for database linkage) and commercialized them at the same timing as the release of the i-PX9000.

For the future, we are also planning to provide Itanium2 compatibility to the Universal Backup (for backup integration) that is a higher-level product to the base products above and commercialize the shared file servers (for file sharing between multiple OSs of i-PX9000) as a new product.

In addition, aiming at the simplification of operation management, we are also planning to provide the i-PX9000 compatibility for the HA/Global Master, which is the function for unified surveillance of multi-OS environment for reallocation of resources amongst the OSs.

These pieces of software will make it possible to build systems that have both the advantages of ACOS's high reliability and open OS's high productivity on the i-PX9000.

6. Conclusion

In the above, we outlined the system architecture available with the ACOS series i-PX9000 including the ACOS Platform Technology, and introduced the multi-OS environment that can be implemented using the i-PX9000. In the future, we are determined to continue the development of the i-PX9000 as a "multi-OS platform" enabling integration of distributed open servers while also inheriting the existing ACOS assets.

---

*Intel and Itanium are trademarks or registered trademarks of Intel Corporation and its subsidiaries in the USA and other countries.
*Windows is a registered trademark of Microsoft Corporation in the USA and other countries.
*HP-UX is a trademark of Hewlett-Packard Company.
*Linux is a trademark or registered trademark of Mr. Linus Torvalds in the USA and other countries.
*Red Hat is a registered trademark or trademark of Red Hat, Inc. in the USA and other countries.
*The official name of Windows Server 2003 is Microsoft® Windows Server™ 2003.
*Other brand names and product names are registered trademarks or trademarks of their respective owners.

Authors' Profiles

TANIOKA Takahiro
Principal Engineer, 1st Engineering Department,
Computers Division,
1st Computers Operations Unit,
NEC Corporation

KAWANABE Masazumi
Manager, 1st Computers Software Division,
Computers Software Operations Unit,
NEC Corporation

MURAKAMI Shoichi
Assistant Manager,
ACOS System Marketing-Promotion Department,
Platform Marketing-Promotion Division,
1st Computers Operations Unit,
NEC Corporation

●The details about this paper can be seen at the following.
Related URL : ACOS Club: http://www.sw.nec.co.jp/acosclub/