RS/6000 SP: Planning Vol. 1, Hardware and Physical Environment

Introducing the RS/6000 SP

This chapter is an overview of the IBM RS/6000 SP. It introduces the processor nodes, model and expansion frames, the standard and optional hardware, and installation site planning. Figure 1 is a conceptual illustration of an RS/6000 SP system.

Figure 1. Conceptual RS/6000 SP system

View figure.

RS/6000 SP system overview

The IBM RS/6000 SP is IBM's family of scalable, parallel computing solutions. It provides a state-of-the-art parallel computing system and industry-leading application enablers and applications. The RS/6000 SP runs the AIX operating system along with the Parallel System Support Programs (PSSP) system software on the control workstation and all processor nodes. For complete information on SP system software issues, see *Planning Volume 2, Control Workstation and Software Environment*.

The scalable architecture of the SP system, its high-performance communication, PowerPC and POWER3 processors give you the power to handle data-intensive, compute-intensive and I/O-intensive jobs with ease. You can execute both serial and parallel applications simultaneously, while managing your system from a single workstation. For scientific and technical applications, the SP system delivers the power and connectivity for rapid turnaround; from structural analysis and seismic modeling to circuit simulation and molecular modeling. Multiple users can run complex queries against very large amounts of data and obtain results interactively. This makes the SP system an ideal solution for database query, online transaction processing, business management, and batch business applications.

The IBM software offerings for the SP system provide an interlocking set of licensed programs designed to address a wide range of system and application needs. The open architecture, based on the AIX operating system (IBM's implementation of UNIX), enables you to easily integrate the SP system into your existing environments. The software architecture is closely tuned to the SP system hardware design for maximum usability and performance.

The SP system family continues the AIX and RS/6000 policy of adherence to open systems standards. Connection to I/O devices, networks of workstations, and mainframe networks is a key element of the SP system. Ethernet, HIPPI, SCSI, FDDI, Token-Ring, ATM, SSA, ESCON, BMCA and Fibre Channel 266 and 1 GB interfaces are supported by the SP system.

In addition to the standard network interface cards, the IBM RS/6000 SP system offers versatile, high-speed network connections using extension nodes. Also, the SP system offers an optional e-business interface using the IBM RS/6000 7017 Enterprise Server as an SP-attached server.

Hardware planning

The basic components of the RS/6000 SP system are:

- Processor nodes (includes SP-attached servers)
- Frames with integral power subsystems
- Switches
- Extension nodes
- Control workstations (a high availability option is available)
- Network connectivity adapters
- External disk drives

These components connect to your existing computer network through a local area network (LAN), making the RS/6000 SP system accessible from any network-attached workstation.

Processor nodes

The IBM RS/6000 SP System is scalable from one to 128 processor nodes. Up to sixteen thin, eight wide, or four high processor nodes can be mounted in a tall frame while a short frame can hold up to eight thin or four wide nodes.

SP systems with more than 128 processor nodes are available on a special order basis; for details, consult your IBM account representative.

There are four types of RS/6000 SP processor nodes:

- High nodes
- Wide nodes
- Thin nodes
- SP-attached servers

Nodes have either a Symmetric MultiProcessor (SMP) configuration or a uniprocessor configuration. SMP-type nodes use Peripheral Component Interconnect (PCI) architecture, while the withdrawn uniprocessor nodes use Micro Channel Architecture (MCA).

See the following chapters for complete details on the available processor nodes:

High nodes

- 375 MHz POWER3 SMP High Node (F/C 2058).
- POWER3 SMP High Node (F/C 2054).

Wide nodes

- 375 MHz POWER3 SMP Wide Node (F/C 2057).
- Withdrawn POWER3 SMP Wide Node (F/C 2053) (withdrawn from production).
- 332 MHz SMP Wide Node (F/C 2051).

Thin nodes

- 375 MHz POWER3 SMP Thin Node (F/C 2056).
- Withdrawn POWER3 SMP Thin Node (F/C 2052) (withdrawn from production).
- 332 MHz SMP Thin Node (F/C 2050).

SP-attached servers

An SP-attached server is an IBM RS/6000 7017 Enterprise Server configured to operate with the RS/6000 SP System. Enterprise Servers are available as models S70, S7A, and S80. Each is a high-end, PCI based, 64-bit SMP unit that supports concurrent 32-bit and 64-bit applications.

Like a standard SP system processor node, the SP-attached server can perform most SP system processing and administration functions. However, unlike a standard SP system processor node, the SP-attached server is housed in its own frame. Thus, the SP-attached server has both node-like and frame-like characteristics.

For SP system attachment, the Enterprise Server requires the installation of several SP system-specific cables and communications adapters.

For details, see SP-attached servers (RS/6000 M/T 7017 and 7026).

Frames

IBM RS/6000 SP system frames contain and provide power for processor nodes, switches, Hard disk drives and other hardware. A frame feature code provides an empty frame with its integral power subsystem and ac power cable. You order nodes and other components using their respective feature codes.

Frames are offered in a list of five options:

- Tall (1.93 m) ***model*** frames
- Tall ***expansion*** frames
- Short (1.25 m) model frames
- Short expansion frames
- SP Switch frames

Frames have locations known as ***drawers*** into which the processor nodes are mounted. Tall frames have eight drawers and short frames have four. Each drawer location is further divided into two ***slots***. A slot has the capacity of one thin node. Wide nodes occupy one full drawer while high nodes fill two drawers.

Complete details on current production frames can be found in the following chapters:

Tall model and expansion frames

Tall frames - 1.93 m (Model 550, 555, F/C 1550, 2031, 2032).

Short model and expansion frames

Short frames - 1.25 m (Model 500 and F/C 1500).

SP Switch frames

SP Switch frames (F/C 2031, 2032).

Control workstations

When planning your control workstation, you can view it as a server to the SP system applications. The subsystems running on the control workstation are the SP system server applications for the SP system nodes. The nodes are clients of the control workstation server applications. The control workstation server applications provide configuration data, security, hardware monitoring, diagnostics, a single point of control service, and optionally, job scheduling data and a time source.

Since most control workstation planning decisions are based on the ability of the control workstation to handle your software requirements, planning information for this item is covered in *Planning Volume 2, Control Workstation and Software Environment*. Included in that book is the latest list of supported control workstations. Also included are the requirements for a high-availability option (**F/C 1245**), which eliminates the control workstation as a single point of failure.

Control workstation planning information in this book is limited to the following:

- I/O adapter features for both PCI and MCA-type control workstations; see Control workstation interface adapters.
- Service Director; a system monitoring and fault reporting application. For details, see Service Director.

Switches

Switches provide a message-passing network that connects all processor nodes with a minimum of four paths between every pair of nodes. The SP Switch series can also be used to connect the SP system with optional external devices.

For information on planning inter-frame switch cabling, see Planning for switch cabling.

SP Switch

SP Switches are available in both 8-port (**F/C 4008**) and 16-port (**F/C 4011**) configurations. For detailed information, see SP Switch (F/C 4011 and F/C 4008).

SP Switch2

The SP Switch2 is available only in a 16-port (**F/C 4012**) configuration. This switch can be used only in SP systems populated exclusively with POWER3 High Nodes. For detailed information, see SP Switch2 (F/C 4012).

SP Switch Router

The IBM RS/6000 SP Switch Router is a licensed version of the Ascend GRF switched IP router that is enhanced for direct connection to the SP Switch. Network connections through SP Switch Routers are typically faster and have better availability than network connections through SP system nodes.

Connections between the SP system and the SP Switch Router require an SP Switch mounted in the SP system and an SP Switch Router adapter mounted in the router. A switch cable is required to complete the connection between the SP switch and the SP Switch Router adapter. For detailed information on the router, its components, and installation requirements, see SP Switch Router (M/T 9077 04S and 16S).

Network communications

SP systems have several communication requirements, including the following:

- All SP systems require an SP Ethernet LAN for system administration.
- Switch-configured systems require a frame-to-frame switch cable network.
- SP systems connected to external networks (or with networks between SP system partitions) require additional communication adapters.

The required SP Ethernet LAN that connects all nodes to the control workstation is needed for system administration and is to be used for that purpose exclusively. If you attempt to route non-administrative traffic over the SP Ethernet, and it interferes with administrative traffic, you will have to reroute the non-administrative traffic. Further network connectivity is supplied by various adapters, some optional, that can provide connection to I/O devices, networks of workstations, and mainframe networks. Ethernet, FDDI, Token-Ring, HIPPI, SCSI, FCS, and ATM are examples of adapter types that can be used as part of an RS/6000 SP system.

On boot/install server nodes, some adapters are needed to support systems that contain nodes running on different PSSP software release levels. For details, see Communication cabling and *Planning Volume 2, Control Workstation and Software Environment*.

Communication cabling

Cables are among the most important things you plan for when you set up a communication network.

Cable planning for your SP system communication network includes:

- Placing hardware items on your floor plan (control workstation, processor frames, routers, and servers) so that cables are the correct length to reach all necessary connection points.
- Ensuring that all wiring standards are followed for the cable types you install.
- For switch-configured systems; laying out the frame-to-frame switch cable network and determining the quantity, length, and type of switch cable you need for each switch-equipped frame.

For details on completing these tasks, see Communication cabling.

Communication and I/O adapters

If you plan to connect your entire SP system (or partitions within your SP system) to external networks, you must install communication adapters. If you have an SP system with multiple system partitions, you can enhance system performance by using optional network adapters instead of the SP Switch network for communication between partitions.

Two different communication bus architectures are used in SP systems; Peripheral Component Interconnect (PCI) and Micro Channel Architecture (MCA).

Use of PCI adapters is limited to the following nodes:

- POWER3 SMP Nodes
- 332 MHz SMP Nodes
- SP-attached servers

Use of MCA communication adapters is limited to the following nodes:

- 160 MHz Uniprocessor Thin Nodes (withdrawn from production)

- 200 MHz SMP High Nodes (withdrawn from production)
- 135 MHz Uniprocessor Wide Nodes (withdrawn from production)

Planning for communication I/O adapters

Planning for I/O adapters involves two high-level tasks:

1. Determining ***which*** adapters to use
2. Determining ***how many*** adapters to use

Determining which adapters to use

Factors for deciding which adapters to use include:

- Performance capability
- What is included with the adapter
- What you must supply for the adapter, including cables if needed
- Software requirements

These details for each PCI-type adapter can be found in PCI communication adapters.

For MCA-type adapters, see Appendix C, MCA communication adapters.

Determining how many adapters to use

Each communication adapter has limitations placed on it including:

- The maximum number allowed per node type
- Performance requirements
- Placement restrictions

This information is shown in tables and lists for each PCI-type adapter in PCI bus I/O adapter requirements for SMP nodes.

For MCA-type adapters, see MCA bus adapter requirements.

Hard disk drives

Hard disk drives for the SP system can be either of two types as follows:

- Internal (contained within the node)
- External (mounted separately, outside of the node)

Internal hard disk drives

Internal hard disk drives are installed in bays within the node. If a node does not use an external boot device, it must be configured with internal hard disk drives. In this configuration, nodes have a minimum and maximum limit for hard disk drives. For details, see Internal hard disk drives.

External hard disk drives

Some later nodes can be configured with no internal hard disk drives. Nodes having only external drives use them as the source of boot information. In this application, external hard disk drives are called ***external boot***

disks. For details, see External boot disks.

For available features to extend node storage, see External hard disk drives.

Installation site planning

Electrical power planning

Power and electrical requirements describes power planning issues, such as the electrical subsystem used in SP system frames and the power requirements of individual SP system components. These issues are covered in the following sections:

- Scalable Electric Power Base Unit (SEPBU) describes the power subsystem.
- SEPBU input branch circuit requirements provides branch circuit information for each frame configuration.
- Power cord specifications for SEPBU-equipped frames describes the power cables provided with each SP system and lists the plugs and connectors required.
- Planning for power requirements of SP frames and features discusses SP system component and auxiliary equipment power requirements.
- Obtaining power system backup devices provides information on obtaining uninterruptible power supplies that are properly matched with your SP system.
- Upgrading power systems in early SP frames describes the equipment supplied in the features for upgrading early SP system power subsystems.

Environmental planning

Environmental factors covers the following topics:

- Environmental specifications of the RS/6000 SP lists temperature and humidity specifications for the SP system.
- Acoustical environment of the RS/6000 SP system has a table listing acoustic emissions of nodes and frames.
- SP thermal output and cooling requirements describes the procedures you use to calculate air conditioning requirements for your installation.
- Evaluating electromagnetic interference.
- Evaluating electrostatic discharge.
- SP frame tie-down considerations provides drawings showing the tie-down locations on SP system frames.

Floor planning

In addition to providing weights and dimensions for SP system frames and their shipping containers, Floor planning provides information needed to develop the work space for your SP system. Some of the topics related to this are:

- Raised floor installations.
- Non-raised floor installations.
- Service clearance specifications for frames has tables and illustrations showing the floor space required for SP system frames and the locations of cable openings and leveling casters for each frame type.
- Multi-frame system floor planning and illustrations contains tables and illustrations showing floor layouts for a variety of multi-frame SP system installations.

- RS/6000 SP system floor load analysis contains the formula you use to calculate the actual load placed on your installation site floor.

RS/6000 SP system upgrades

Many upgrades, conversions, and feature additions are available to enhance SP system performance and capability. Examples of typical upgrades include adding frames to your SP system, installing new higher-performance nodes, and converting MCA-type nodes to PCI nodes. For details, see Appendix A, SP system upgrades, conversions, and feature additions.

[Top of Page | Previous Page | Next Page | Table of Contents | Index]

RS/6000 SP: Planning Vol. 1, Hardware and Physical Environment

375 MHz POWER3 SMP High Node (F/C 2058)

Description

375 MHz POWER3 SMP High Nodes (F/C 2058) use PCI bus architecture and have four, eight, twelve, or sixteen 375 MHz 630FP 64-bit processors per node. Your IBM RS/6000 SP system must be operating at PSSP 3.2 (or later) to use these nodes.

The 375 MHz POWER3 SMP High Node occupies two full drawer locations, thus four nodes can be housed in a tall frame. These nodes require a 1.93 m tall, deep frame (**Model 550**) or expansion frame (**F/C 1550**); they are not supported in the withdrawn 2.01 m frame or in the 1.25 m frame. These nodes can be placed in the first node slot of a frame without requiring additional nodes.

The 375 MHz POWER3 SMP High Node provides additional hard disk drive and PCI adapter capacity by connecting to SP Expansion I/O Units; for details, see SP Expansion I/O Unit (F/C 2055).

Note:

375 MHz POWER3 SMP High Nodes are not compatible with High Performance switches (**F/C 4010** and **F/C 4007**).

PCI bus description

The 375 MHz POWER3 SMP High Node PCI bus contains one 32-bit and four 64-bit PCI slots for I/O adapters.

Additional PCI adapters can be attached to the bus by using up to six optional SP Expansion I/O Units. Each expansion unit has eight 64-bit PCI adapter slots.

Requirements and options

F/C 2058 requirements

375 MHz POWER3 SMP High Nodes occupy two full node drawers. Up to four 375 MHz POWER3 SMP High Nodes can be installed in one tall/deep frame. Mandatory prerequisites are:

- PSSP 3.2 (or later) on the processor node, control workstation, and backup nodes
- Four processors (on one card, mounted in one slot)
- 1 GB of memory
- 9.1 GB of mirrored storage (with internal booting)

F/C 2058 options

Available options include the following:

- Four processor slots allowing a maximum of sixteen processors per node
- Four memory slots supporting up to 64 GB of memory
- Five PCI slots (four 64-bit and one 32-bit) for communication adapters
- A dedicated Mezzanine Bus (MX) slot for an optional SP Switch adapter
- A dedicated port for an optional SP Switch2 adapter
- Integrated Ethernet with BNC and RJ45 ports (only one port can be used at a time):
  - 10BASE2 Ethernet (BNC)
  - 10BASE-T or 100BASE-TX Ethernet (RJ45)
- Support for up to six SP Expansion I/O Units (**F/C 2055**)
- Two internal hard disk drive bays supporting up to 72.8 GB of storage (36.4 GB mirrored)
- Integrated Ultra SCSI network
- Two external, nine-pin, RS-232 connectors on the planar S2 and S3 ports. The S3 port is supported only for HACMP serial heartbeat; the S2 port is not supported for this use. A 9-pin to 25-pin converter cable is included with the node for this connector.
  - Node-to-node HACMP cable (**F/C 3124**)
  - Frame-to-frame HACMP cable (**F/C 3125**)

Processor requirements and options

375 MHz POWER3 SMP High Nodes require a minimum of four 375 MHz, 630FP processors mounted on one card. You can order up to three additional four-processor cards (**F/C 4350**) to configure the node with a total of sixteen CPUs.

Table 1. Processor options for 375 MHz POWER3 SMP High Nodes (F/C 2058)

Feature Code	Description	Minimum per node	Maximum per node
4350	One processor card with four CPUs	1	4

Memory requirements and options

375 MHz POWER3 SMP High Nodes have one to four memory cards, require a minimum of one GB of memory, and support a maximum of 64 GB. Memory is supplied by 128, 256, and 512 MB DIMMs which must be mounted in banks of eight DIMMs. Different capacity DIMMs (in banks of eight) can be mixed on the memory cards.

For the best memory-access bandwidth, memory DIMMs should be distributed evenly across four memory cards. As an example, you can realize better bandwidth by using four banks of 128 MB DIMMs (4 GB total) distributed evenly over four memory cards rather than by using one bank of 512 MB DIMMs (4 GB total) on one memory card. The following list illustrates this:

- 1 to 16 GB memory mounted on one card yields 16.8 - 24.3% of peak bandwidth
- 2 to 32 GB mounted on two cards yields 33.5 - 48.5% of peak
- 4 to 64 GB mounted on four cards yields 67 - 97% of peak

The configurator rules used for memory placement in these nodes are designed to yield the best memory performance. Any plans to increase the amount of memory in the future should always be taken into consideration when deciding what size DIMMs to use and the quantity of memory cards to order.

Table 2. Memory features for 375 MHz POWER3 SMP High Nodes (F/C 2058)

Feature Code	Description	Minimum Node Requirement	Maximum Features Per Node
4880	Base Memory Card	1	4
4133	1 GB - (8 x 128 MB DIMMs)	1	16
4402	2 GB - (16 x 128 MB DIMMs)	0	8
4403	3 GB - (24 x 128 MB DIMMs)	0	4
4404	4 GB - (32 x 128 MB DIMMs)	0	4
4134	2 GB - (8 x 256 MB DIMMs)	0	16
4412	4 GB - (16 x 256 MB DIMMs)	0	8
4413	6 GB - (24 x 256 MB DIMMs)	0	4
4414	8 GB - (32 x 256 MB DIMMs)	0	4
4421	4 GB - (8 x 512 MB DIMMs)	0	16
4422	8 GB - (16 x 512 MB DIMMs)	0	8
4423	12 GB - (24 x 512 MB DIMMs)	0	4
4424	16 GB - (32 x 512 MB DIMMs)	0	4

Hard disk drive requirements and options

375 MHz POWER3 SMP High Nodes can have one pair of internal hard disk drives attached through an integrated Ultra SCSI network. The node can have either no internal hard disk drives (with external booting) or from 9.1 GB to a maximum of 36.4 GB of mirrored, internal disk storage.

Additional hard disk drives can be attached to the 375 MHz POWER3 SMP High Node by connecting up to six SP Expansion I/O Units. Each expansion unit has four hard disk drive bays. For details, see SP Expansion I/O Unit (F/C 2055).

Optional internal hard disk drives are available as follows:

- 9.1 Gigabyte Ultra SCSI disk pair (**F/C 2909**)
- 18.2 Gigabyte Ultra SCSI disk pair (**F/C 2918**)
- 9.1 Gigabyte Ultra SCSI 10K RPM disk pair (**F/C 3804**)
- 18.2 Gigabyte Ultra SCSI 10K RPM disk pair (**F/C 3810**)
- 36.4 Gigabyte Ultra SCSI 10K RPM disk pair (**F/C 3820**)

External storage devices can be accessed through optional Ultra SCSI adapter (**F/C 6207**) and SSA adapter (**F/C 6230**).

Switch and communication adapter requirements and options

Switch Restrictions

- 375 MHz POWER3 SMP High Nodes are not supported with the SP Switch-8. You must use either the SP Switch, 16-port (**F/C 4011**) or the SP Switch2, 16-port (**F/C 4012**).
- 375 MHz POWER3 SMP High Nodes are not compatible with the older High Performance series of switches. They can only be used with the SP series switches.

Switch adapters

For systems using the SP Switch, these nodes require the SP Switch MX2 Adapter (**F/C 4023**); for details, see Switch adapters (F/C 4020, 4022, and 4023).

For systems using the SP Switch2, these nodes require the SP Switch2 Adapter (**F/C 4025**); for details, see SP Switch2 Adapter (F/C 4025).

I/O adapters

The 375 MHz POWER3 SMP High Node has five PCI (Peripheral Component Interconnect) adapter slots. A full line of PCI adapters is offered for these nodes including:

- SCSI-2
- Ethernet
- Token Ring
- FDDI
- ATM
- Async
- Wide Area Network (WAN)
- SSA RAID5
- S/390 ESCON
- Serial HIPPI

For more information about these adapters, see PCI bus I/O adapter requirements for SMP nodes and PCI communication adapters.

Note:

A 10BASE2 or 10BASE-T/100BASE-TX Ethernet adapter for the SP Ethernet is integrated into the POWER3 High Node and does not occupy a PCI slot.

[Top of Page | Previous Page | Next Page | Table of Contents | Index]