

Chapter 3

Chapter Three

Juniper Networks Router Architecture

When the routers produced by Juniper Networks first hit the market in 1998, they brought simplicity of design, a logical UNIX-style CLI, and robust troubleshooting tools. The engineers who designed the routers wanted to build a device that made sense. In doing so, they filled a void in the market and appealed to other engineers who wanted a router that moved packets through as quickly as possible.

Juniper Networks' product offerings come in the form of the M-Series routers. The M40 was the first router of its kind capable of scaling to meet current Internet needs. The initial M40 release was based on the Internet Processor I. The Internet Processor I was the fundamental core of the packet forwarding engine (PFE). The PFE consisted of a shared memory, a single forwarding table, and a one-write, one-read architecture. The entire PFE was capable of forwarding 40Mbps, more than 100 times the capacity of other available router architectures at the time.

Although the M40 was quite progressive, Juniper Networks was able to improve on its available functionality by upgrading the processor, raising the available memory, adding redundancy, and including the ability to filter traffic through ACLs in later iterations of the product.

This chapter introduces you to the router models and architectural differences of each product. In addition, we will describe each hardware and software piece of the router and explain how that piece contributes to the overall logic of the device. This will include the routing engine, the PFE, the switching fabrics and control boards, the interfaces available, and the differences between the available router models. It will also include an explanation of the router's boot process and how to upgrade the JUNOS software.

Since the architecture and operating-system commands of these routers differs from those of other vendors, such as Cisco, with which you may already be quite familiar, the material covered here should help you to understand the rest of the material in this book. It will also undoubtedly help in your pursuit of Juniper Networks career certifications.

3.1 Juniper Networks Router Models

This section describes the different models of Juniper Networks routers. The physical dimension, performance statistics, and some information about the internal architecture of the router itself are provided for each model.

3.1.1 M5 and M10

The M5 and M10 routers were introduced in September 2000 as the latest additions to the router family. With their introduction, Juniper Networks hoped to gain a larger marketshare by appealing to networks needing a smaller footprint router. Due to its minimal physical requirements— $5.25 \times 17.4 \times 24$ in., or $13.33 \times 44.2 \times 60.96$ cm—single rack can hold 15 M5s, which creates a bandwidth-to-footprint-to-price ratio that is hard to beat.

The M5 and M10 were released at the same time because they had similar architectures with two different throughput capabilities (5Gbps on the M5 and 10Gbps on the M10). Both routers employ the Internet Processor II ASIC, providing forwarding table lookups at 40Mpps. The M10's chassis looks the same as the M5's; however, there are two forwarding engine boards (FEBs) in the M10, allowing for a maximum of eight physical interface cards (PICs) to be used.

3.1.2 M20

The second router introduced by Juniper Networks was the M20, released in December 1999. The M20 also uses the Internet Processor II ASIC and is capable of throughput in excess of 20Gbps.

With physical dimensions of $14 \times 19 \times 21$ in., or $35.56 \times 48.26 \times 53.34$ cm, a network administrator can stack five chassis in a single equipment rack. The M20 was the first Juniper Networks router available with redundancy (power supply, routing engine, and system and switch board [SSB]). This greatly increased the appeal of the Juniper Networks routers to the marketplace. Component failure in an operational network can be disastrous. By addressing the need for component redundancy, Juniper Networks was able to allay this fear in the minds of potential customers.

3.1.3 M40

The M40 router was the first product launched by Juniper Networks. With a chassis size of $35 \times 19 \times 23.5$ in., or $88.9 \times 48.26 \times 59.69$ cm, deployment is limited to two chassis per equipment rack. However, the router's architecture provides over 40Gbps throughput. The M40 supports the same PICs as the M20. The PICs are compatible between both platforms.

Although initially deployed with the Internet Processor I and without ACL capability, the M40 now runs on the Internet Processor II and has addressed the need for filtering. This platform, however, does not provide for the same component redundancy as the M20 and M160 models, an important distinction for most customers.

3.1.4 M40e

To answer the need for the throughput of the M40 coupled with redundant-component capability, Juniper Networks introduced the M40e platform in February 2002. The M40e router has the same footprint and port density as the M40, but it provides the optional redundancy that the M40 does not. This model is compatible with most of the PICs from the M20, M40, and M160 models.

3.1.5 M160

The M160 was introduced in March 2000 as the third box in the M-Series. It is a formidable router, both in size and capacity. The M160 chassis is $35 \times 19 \times 29$ in., or $88.9 \times 48.26 \times 73.66$ cm. This allows for two per equipment rack.

The M160, to date, is the highest-rated core router on the market. Independent testing has shown that the M160 outperforms the competition in areas of BGP table capacity, MPLS LSP capacity, route flapping recovery at OC-192 speeds, convergence at both OC-192 and OC-48 speeds, and filtering at both OC-192 and OC-48 speeds. In additional tests, the M160 has matched or exceeded the competition in the areas of CoS at OC-48 and OC-192 speeds and IP and MPLS baseline testing at OC-48 and OC-192 speeds.

The M160 platform provides the maximum throughput and port density necessary for the next generation of Internet architectures.

3.1.6 G10

In November 2001, Juniper Networks announced its intent to acquire Pacific Broadband Communications and its CMTS. Subsequently, Juniper Networks rereleased that CMTS as its G10 router. This product is aimed at the growing broadband-remote-access-service (BRAS) market that delivers Internet service into private homes and small businesses primarily through cable modems.

A chief complaint of cable Internet subscribers is that as more subscribers join in a given area, the amount of bandwidth available to each end user can drop dramatically. The G10 uses a custom-built ASIC that has the capability of 20 legacy CMTS chips. The end result is that this device is capable of supporting greater numbers of subscribers using less bandwidth.

3.2 Architecture Overview

When Internet routers were first introduced, the amount of traffic processed was much less than it is now. The router's brain, its CPU, was robust enough to handle its own tasks, the tasks of creating and managing the routing tables, and the task of forwarding the packets. Some routers on the market today still use this processor-based technology.

Juniper Networks has created a device that segregates the tasks and assigns them to different parts of the router, sort of like an assembly line. You can see this process in Figure 3-1. Because Juniper Networks routers are designed to serve the busy core of the network, the number of packets processed per second is in the millions. If the router were required to manage everything from its CPU, throughput would suffer, as would the ability to provide service guarantees.

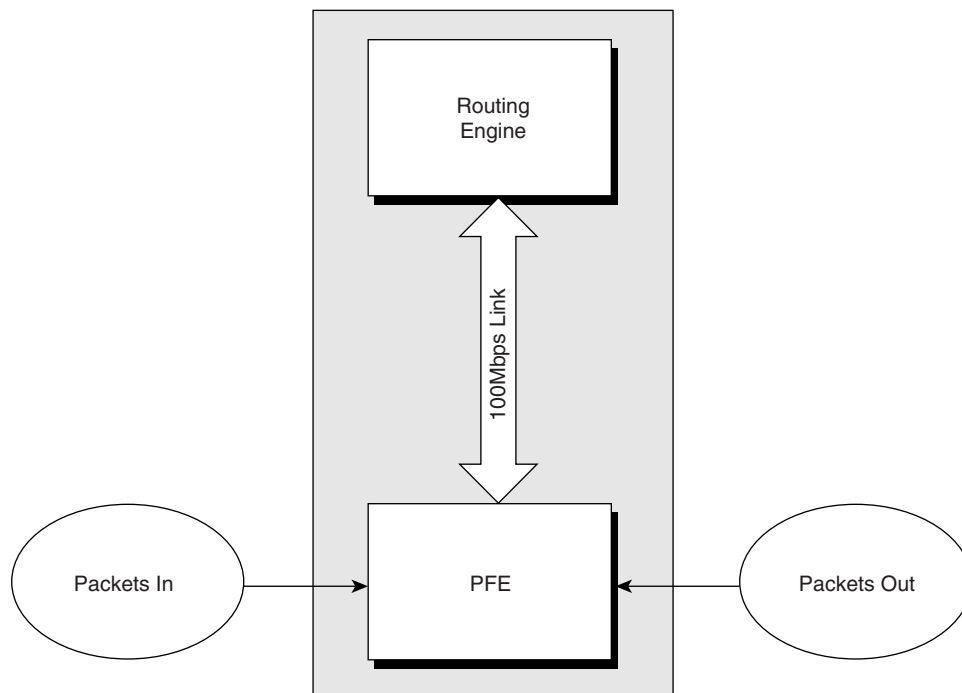


Figure 3-1 Juniper Networks Routing Architecture

Most routers today are beginning to follow this model of separate processes within the router—routing and forwarding. Either through the use of ASICs on the line cards containing the interfaces or through the use of separate processors

within the routing unit containing the CPU, the router manufacturers have all started to turn towards this technology. It simply makes sense. Let's examine the way a Juniper Networks router attacks this issue.

Juniper Networks routers consist of a simple architecture containing two basic components: a routing engine and a PFE. The routing engine handles the more mundane tasks, such as routing protocol calculations, control packet processing, and so on, while the PFE is allowed to move packets out of the router as quickly as possible. After all, the main duty of an Internet router is to handle the packets as little as possible as they move through its realm. To achieve flawless, wirespeed performance, Juniper Networks ensured that certain processes would be physically and logically segregated on their routers. Figure 3-2 shows the segregation of processes between the two main components of a Juniper Networks router. The rest of this chapter explains these components and processes further.

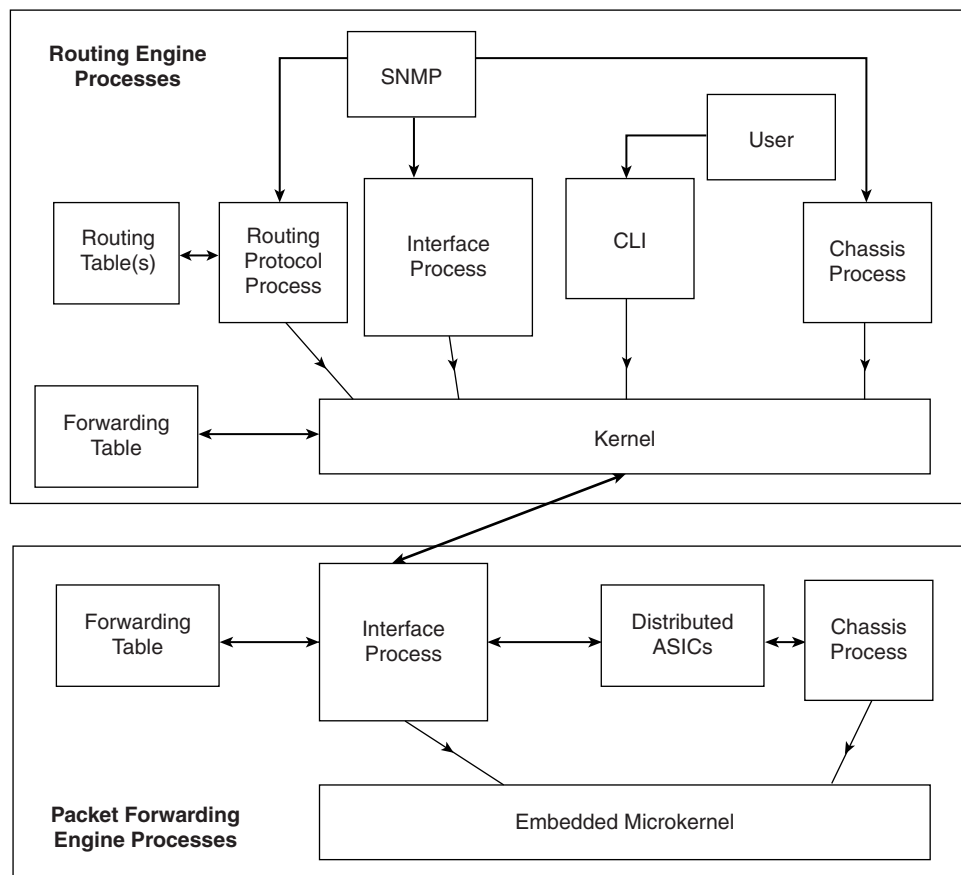


Figure 3-2 Routing Engine and PFE Processes

3.2.1 Routing Engine

The Juniper Networks routing engine is an integral part of the router architecture and provides for all of the central processing and route processing requirements. As you will learn in this section, the routing engine also provides storage for the operating system and provides the CLI through which the operating system is configured.

Although there are differences between the various Juniper Networks router models, which are outlined in the Section 3.2.1.3, the design of the routing engine is similar in all of them. The routing engine is primarily responsible for running the routing protocols, keeping the routing tables up-to-date, sending routing updates to the PFE, and performing system management. The routing engine communicates directly with the PFE through a 100Mbps connection.

3.2.1.1 Function of the Routing Engine

Functions provided by the routing engine include the following:

- Handling of routing protocol packets
- Management interface
- Configuration management
- Accounting and alarms
- Modular software
- Scalability

Routing protocol packets that arrive on the network are not handled on the PFE itself, but are passed directly to the routing engine. This effectively reduces the amount of work that the PFE has to do, enabling it to process packets to be forwarded efficiently. An example of a routing protocol packet would be a link-state advertisement (LSA) from an OSPF router. The LSA would be received on an ingress interface and sent directly to the routing engine. The routing engine would then perform the shortest-path-first (SPF) route calculations and update its OSPF routing table, which, in turn, would send LSAs to its neighbors. (For more information on the functions of OSPF, please refer to Section 8.4.)

The routing engine also provides several ways to manage the router. First, it provides the CLI, which allows the system operator to interact with the JUNOS software, the PFE, and the interfaces through configuration, modifications, and monitoring. The routing engine also runs SNMP, permitting management of the system from a network management station running software, such as Hewlett-Packard's Network Node Manager (HP-NNM), through a framework, such as Hewlett-Packard's OpenView. Finally, the craft interface, discussed in Section 3.2.1.4, provides more information for management of the router.

Accounting functionality and alarms provides further manageability of the router. Alarms, seen via the craft interface, provide information to the system administrator about the condition of the router or functions of the router. Accounting of packets is done in the routing engine, thus negating any impact on wirespeed routing taking place in the PFE.

For change and configuration management, the routing engine allows for the storage of configuration files, microcode, and system images in one primary and two secondary locations. A unique rollback feature, provided by JUNOS, also allows the system administrator to return to a previous configuration quickly, should the new configuration prove problematic.

Finally, Juniper Networks routers use modular software that cleanly separates processes from each other. The problems of one process will not impact other processes that may be running. Additionally, the software is designed to scale well to the needs of tomorrow's Internet routing demands.

3.2.1.2 JUNOS

JUNOS is the operating system currently used on all Juniper Networks routers. JUNOS is not just an operating system providing a CLI for configuration, but also a feature-rich platform providing troubleshooting tools and advanced feature sets. The operating system also incorporates an application programming interface (API) system for external program calls and scripting capabilities. JUNOS, in conjunction with the Internet Processor II, comprises the industry's most advanced BSD-based router operating system in today's marketplace. The routing engine is based on an Intel PCI platform running JUNOS. JUNOS runs in flash memory with an alternate copy stored on the router's hard disk. As you can see in Figure 3-3, the operating-system kernel is layered on the PCI platform and establishes communication between the PCI platform and the system processes. The kernel is also responsible for making sure that the forwarding tables in use by the PFE are in sync with those in the routing engine.

There are five essential functions provided by JUNOS:

1. The routing protocol process provides all routing and routing control functions within the platform. The modularity of this package allows for the addition and removal of protocols and functions, providing both flexibility and scalability.
2. The interface process performs configuration of the physical interfaces and encapsulation.
3. The SNMP and management information base (MIB) II processes allow SNMP-capable systems to communicate with the router platform. This also allows the platform to provide necessary SNMP information to external agents. JUNOS is SNMP I and II compliant.

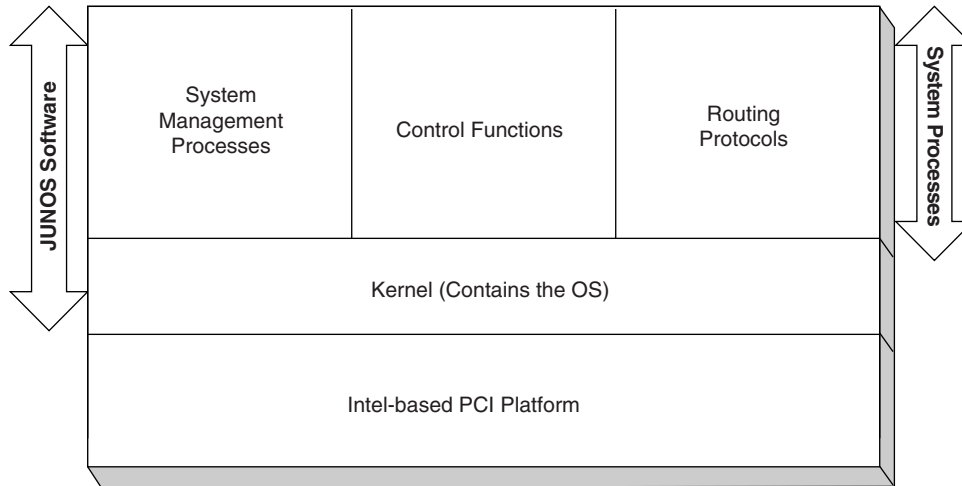


Figure 3-3 *Routing Engine Components*

4. The management process starts and monitors all other software processes in JUNOS. If a particular management function stops, this process will attempt to restart it.
5. The routing kernel process controls everything else. In addition to providing the underlying infrastructure to all JUNOS software processes, the routing kernel process provides the link between the routing engine and the PFE.

One or more routing tables are maintained by the operating system. Routing policy is maintained on JUNOS within the routing engine.

Let's take a closer look at the processes running on the routing engine. Figure 3-4 gives a visual representation of how these processes all work together to carry out the business at hand. You'll notice that nearly all of the processes communicate directly with the kernel. The exceptions are the user process, which must access the kernel through the CLI, and the routing table, which is simply a product of the routing protocol processes. Any communication between the routing engine and PFE originates from the kernel itself.

3.2.1.3 Routing Engine Specifications

Routing engine specifications will depend upon the router model. The differences between the Juniper Networks M-Series router models are listed in Table 3-1. Note that all of the routers provide for out-of-band management, RS-232 DB9 ports for serial console and remote management access, and tertiary storage

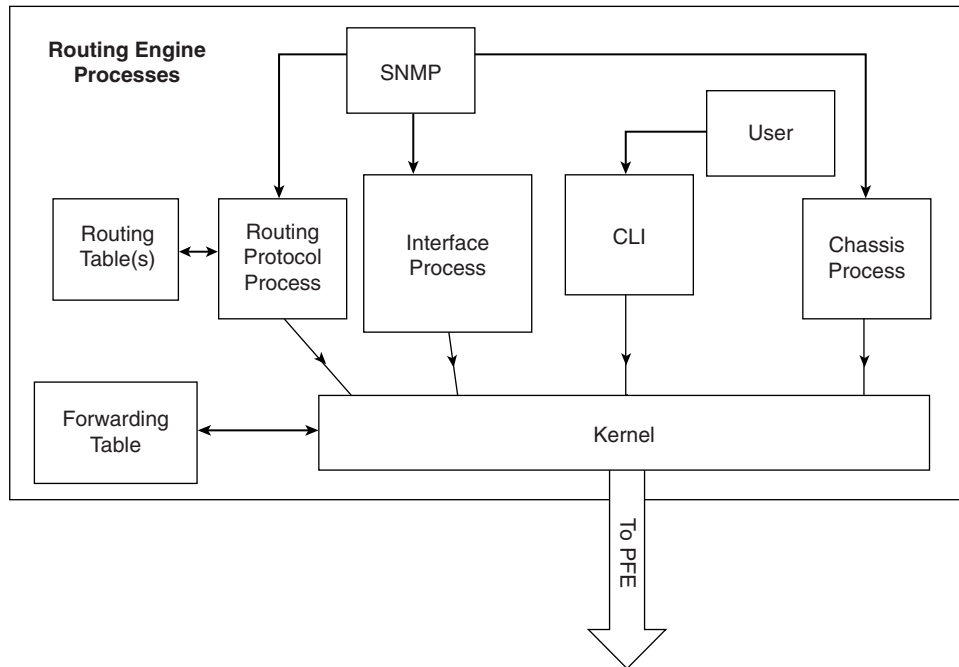


Figure 3-4 Routing Engine Processes

Table 3-1 Routing Engines on Different Models

Model	Platform	Available Redundancy	Compact Flash	SDRAM
M5 and M10	333MHz Pentium II with integrated 256MB level 2 cache	NO	96MB	256, 512, or 768MB
M20, M40, and M40e	333MHz Pentium II with 512MB cache	M20—YES M40—NO	80MB	Up to 768MB
M160	333MHz Pentium II with integrated 256MB level 2 cache	YES	80 or 96MB	768MB

using a removable PC card. The main difference between the models is the amount of flash or SDRAM available. Of course, these differences only apply to the routing engines. There are substantial differences in capacity, throughput and available interfaces between models, as well.

The M160 Miscellaneous Control Subsystem

Only the M160 router uses the routing engine in conjunction with a miscellaneous control subsystem (MCS). The two are installed adjacently in the rear of the chassis and together form a host module. Both components are required to function. If a routing engine is installed without the MCS, the routing engine will not work—and vice versa. The router will accommodate up to two host modules.

The MCS performs the following functions:

- Acts as a middle man between the routing engine and the sensors throughout the system; relays statistical information to the routing engine, and relays control messages and alarms out to the system from the routing engine
- Controls power-up and power-down of system components
- Decides which of any given redundant components will act as master
- Performs reset attempts on flexible PIC concentrators (FPCs), when necessary (the FPC will be discussed later in this chapter in Section 3.2.2)
- Acts as the SONET 19.44MHz clock source and monitors all other system clocks

You have probably noticed that some of these functions are performed by the routing engine itself on other router modules. While this is true, no other M-Series model provides the port density and forwarding speed of the M160. Some of the functions traditionally built into the routing engine have been moved out into this new router component, the MCS, to focus the routing engine on more specific tasks.

The MCS comprises the following:

- A PCI interface to the routing engine
- Two BITS interfaces for external clock sources
- A 100Mbps Ethernet interface to other system modules
- A 19.44MHz stratum 3 SONET clock source
- A controller for monitoring the sensors
- A debugging port (RS-232)
- LEDs
- An offline button

3.2.1.4 The Craft Interface

Positioned on the front of the chassis, the craft interface provides an external look into the internal workings of the router. It can be used as a troubleshooting tool, a monitoring tool, or both. Although the craft interface looks different on each model, the workings are very similar. The example figures in this section are based on the M40 model.

The main features of the craft interface are the following:

- LED indicators
- Alarm indicators
- Routing engine ports
- LCD display screen (on the M40 and M160 only)

The craft interface on an M40 model, shown in Figure 3–5, displays the status of the FPCs, of the routing engine and of general alarm conditions. Each FPC has a corresponding button on the craft interface. LEDs above the button indicate whether the FPC’s status is “OK” or it has failed to initialize.

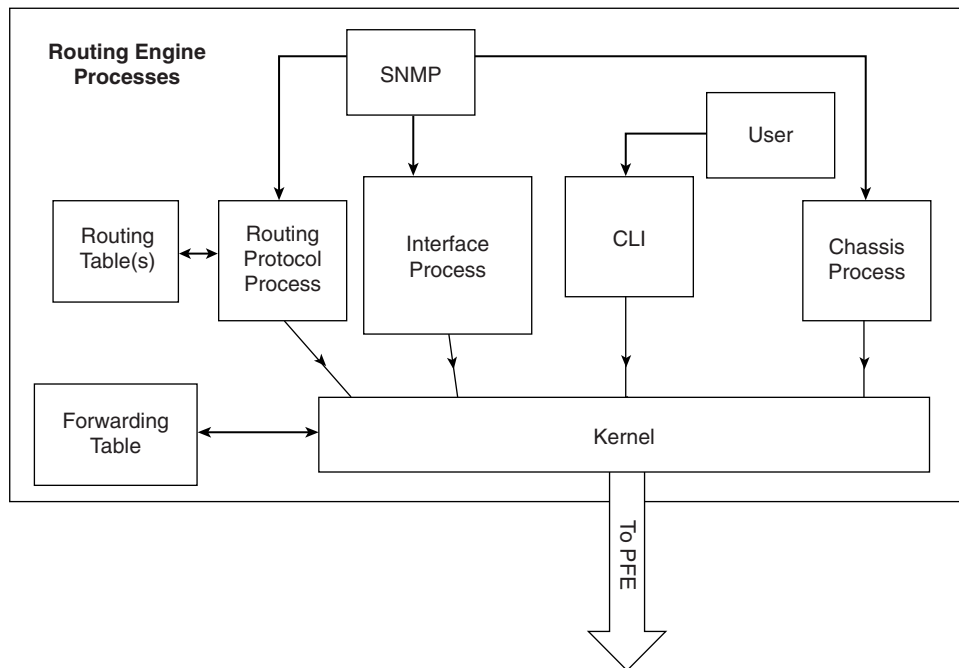


Figure 3–5 M40 Craft Interface

84 Juniper Networks Router Architecture

Alarm LEDs indicate the level of an alarm if one has occurred. On this alarm panel are two alarm relay contacts. These can be used to connect external alarm devices to the craft interface. If a yellow or red alarm occurs, the external alarm device would also be activated. Alarms can be silenced with the alarm cutoff button, but this does not remove the alarm condition.

Red alarms indicate a condition in which a service interruption could occur, such as a component failure. Yellow alarms are generally indicative of recoverable errors or maintenance alerts.

Routing engine access is provided on the right side of the craft interface through a console port, an auxiliary port, and a management Ethernet port. The status of the routing engine is indicated as either OK or Fail. More information about the LED status indicators is provided in Table 3-2.

Table 3-2 *Craft Interface Indicators*

LED	Color/Action	Description
OK	Green/Blinking	Initializing
OK	Green/Solid	Running
Fail	Red/Solid	Offline, owing to failure (In the case of the routing engine, this could mean that the system control board did not detect the routing engine.)
ALARMS		
Red Alarm	Red/Solid	System failure, power supply failure, or system threshold exceeded
Yellow Alarm	Amber/Solid	Maintenance alert or indication of temperature increase

The FPC buttons are used to take an FPC offline, before removing the FPC, for instance. To do this, press and hold the button for three seconds, or until the red Fail LED becomes solid. Then, it is safe to remove the FPC.

The LCD display screen, shown in Figure 3-6, works in either *idle mode* or in *alarm mode*. When in idle mode, the LCD display will show the current system status. When in alarm mode, the LCD display will provide more information about the alarm condition. To interact with the LCD menu, use the buttons and directional arrows to the right of the LCD display screen.

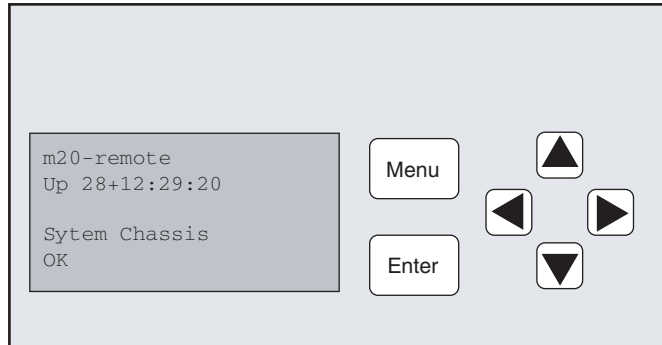


Figure 3-6 *Craft Interface LCD Panel*

Finally, the craft interface provides three ways to interact with the CLI:

1. Console port
2. Auxiliary port
3. Management Ethernet port

Using an RS-232 serial cable, an external console, such as a dumb terminal and keyboard, can be connected to the console port to display system messages and information constantly or to enter the CLI. A laptop computer or modem may be connected to the auxiliary port for quick, portable access to the CLI.

The management Ethernet port can be used to connect the router to any Ethernet LAN through an autosensing 10/100 RJ-45 port. Most network administrators connect this port to a management LAN for out-of-band management of the routers. Unlike with routers from some other vendors, this management port can be controlled via the CLI, but it will not route traffic and, therefore, cannot be used as a spare port.

3.2.1.5 Redundancy and Maintenance Options

In the busy network core, network availability is everything. Juniper Networks has designed its routers to reduce single points of failure. Routing engines are no exception. Most router models can be configured with redundant routing engines, thereby reducing system downtime in the event of a failure. (There will be an interruption of routing services during the failover, however, as there is whenever an routing engine is inserted or removed.) Table 3-3 provides more information on redundant components in each router model. (Note that some of the components described here will be covered in detail later in this chapter.)

86 Juniper Networks Router Architecture

Table 3-3 *Redundancy in Components by Model*

Model	M5 and M10	M20	M40 and M40e	M160
Redundant routing engine functions	none	P/S Cooling Routing engine SSB	P/S Cooling	P/S Cooling Host module PFE clock generator SFM

The routing engine is said to be *hot-pluggable*, which simply means that it may be inserted while the router is powered up. Routing functions will be interrupted whenever a routing engine is removed or inserted, however.

Note: If the router does not have two routing engines, the router will not be operational without the single routing engine inserted.

Maintaining the routing engine requires attention to the LED on the craft interface to check for alarms or other indications of operational problems. The system administrator can also find some information from the CLI by using the following command:

```
system@m20# show chassis routing-engine

Routing Engine status:
  Temperature      28 degrees C / 82 degrees F
  DRAM             768 Mbytes
  CPU utilization:
    User           0 percent
    Background     0 percent
    Kernel         0 percent
    Interrupt      0 percent
    Idle           100 percent
  Start time      2002-03-06 17:23:09 UTC
  Uptime         20 hours, 44 minutes, 41 seconds
  Load averages: 1 minute      5 minute      15 minute
                  0.00          0.00        0.00
```

Notice that you can see items such as router uptime (the amount of time the router has been powered up), temperature of the chassis, and the amount of DRAM installed.

Many parts of the routing engine are *field-serviceable* (also called *field-replaceable*), meaning that replacement or spare parts can be used to get the

routing engine back into operation quickly without having to ship it to Juniper Networks for repair. Replacement of these parts should be done under the guidance of an engineer from the JTAC and is not recommended for inexperienced service personnel.

3.2.2 Packet Forwarding Engine (PFE)

The PFE is the second basic component of the Juniper Networks router. It is the mass-transit system part of the router, so to speak. Whereas the routing engine is the brain of the router, the PFE tends to be more of a workhorse, carrying out the instructions it has been given. The job of the PFE is to move packets as quickly as possible back out of the router. If it can't do that, for instance when there is no entry in the forwarding table for a given destination, it hurries the packets bound for that unknown destination off to the routing engine and goes on about its business.

This section will give you an overview of the design and function of the PFE. It will also show you how the packets move through the router so that you can fully understand the way the whole system works.

3.2.2.1 Design and Operation

On Juniper Networks routers, the PFE is designed to perform Layer 2 and Layer 3 switching, route lookups, and rapid forwarding of packets. Using ASICs, the strategy of the PFE is to divide and conquer the business of forwarding. To that end, the PFE itself is split into several major components:

- Midplane
- PICs
- FPCs
- Control board (switching/forwarding)

The midplane, sometimes referred to as the backplane, is really the back of the cage that holds the line cards. The line cards connect into the midplane when inserted into the chassis from the front. The routing engine plugs into the rear of the midplane from the rear of the chassis. The purpose of the midplane is to carry the electrical signals and power to each line card and to the routing engine.

The PICs are the actual components that contain the interface ports. Each PIC is plugged into a FPC, such as the one shown in Figure 3-7. Each individual PIC contains an ASIC that handles media-specific functions, such as framing or encapsulation, and has its own LED status indicator on the front. PICs are available for SDH/SONET, ATM, Gigabit Ethernet, Fast Ethernet, and DS3/E3.

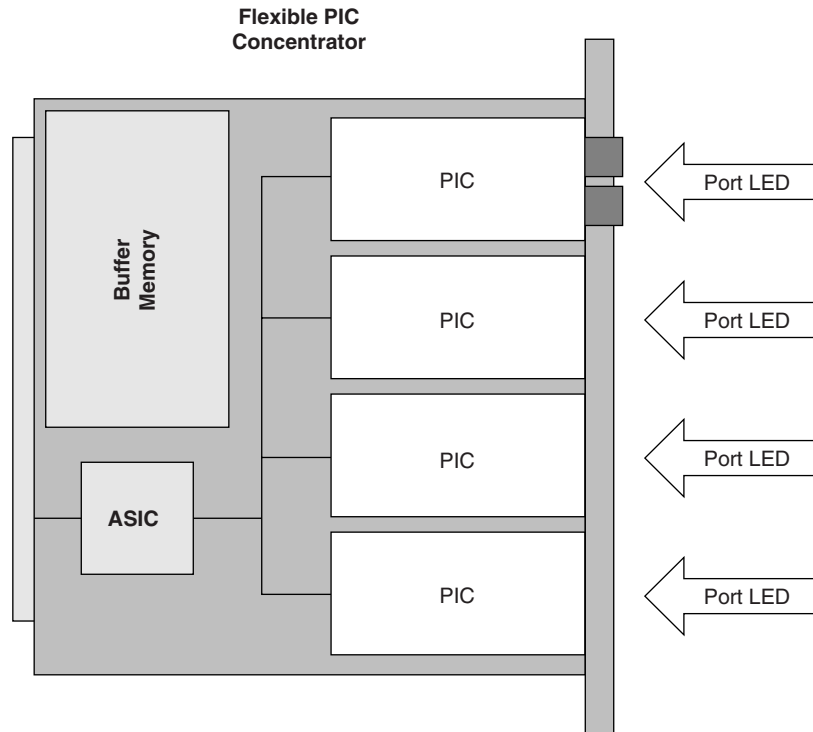


Figure 3-7 *The FPC*

The FPC can contain from one to four PICs in a mix-and-match style. In other words, you could have four different kinds of PICs on a single FPC. This reflects a great deal of flexibility that is welcome in most networks. Installed from the front of the chassis, the FPC carries the signals from the PICs to the midplane. Each FPC has its own input-output (I/O) ASIC and buffer memory.

In the M5 and M10, PICs do not connect to an FPC, but to an FEB. In the M20, M40, and M160, PICs connect to an FPC. There are obviously other significant architectural differences. The PICs for the M5 and M10 are interchangeable; however, due to architectural differences these same PICs cannot be used in the M20, M40, and M160.

The FPC performs the important functions of decapsulating the packet, parsing it, and breaking it up into 64-byte memory blocks, before passing it to the distributed buffer manager (DBM) ASIC. It is at this point that the packet is first written to memory. The DBM ASIC manages and writes packets to the shared memory across all FPCs. While writing the packets to buffer memory, the DBM ASIC is also extracting information on the destination of the packet, as you will see this when we look at packet flow later in this section.

Note: In each router slot, there must either be an FPC or a blank panel installed to ensure adequate cooling and airflow through the router.

The M160 router is the only exception to this overview of the FPC. The M160 actually can use two different types of FPC—the FPC1 and the FPC2.

The control board is an add-on component in the PFE and will be covered in more detail in Section 3.2.2.2. Each control board performs part of the overall function of the PFE, such as communications with the routing engine through an internal interface and with the FPCs through an internal hub.

PFE Processes

As Figure 3–8 shows, the PFE has an embedded microkernel that serves as the brains of the PFE, interacting with the interface process and the chassis process to monitor and control these functions. It is the interface process that has direct communication with the kernel of the routing engine. This communication includes forwarding exception and control packets to the routing engine, receiving packets to be forwarded, receiving the forwarding table updates, providing information about the health of the PFE, and permitting configuration of the interfaces from the user-CLI process on the routing engine.

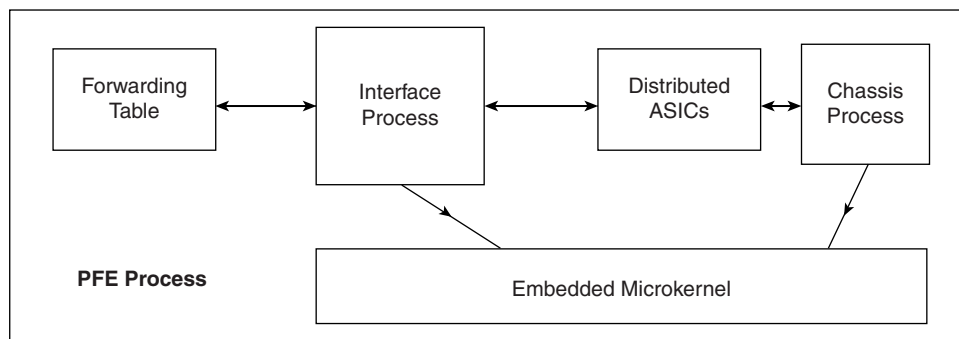


Figure 3–8 Packet Forward Engine Processes

The PFE contains a stored forwarding table, which is static until a new one is received from the routing engine. No dynamic routing protocol processes run on the PFE. The interface process consults the forwarding table to look up next-hop information. The interface process also has direct communication with the ASICs on the PFE, which will be discussed in detail in the next section. Finally, the chassis processes—environment, health, and so on—communicate directly with the microkernel of the PFE and with the ASICs.

ASICs

Now we will take a look at the location of the ASICs involved in packet processing and see how they relate to one another. Figure 3–9 shows a section-by-section view of the positioning and communication.

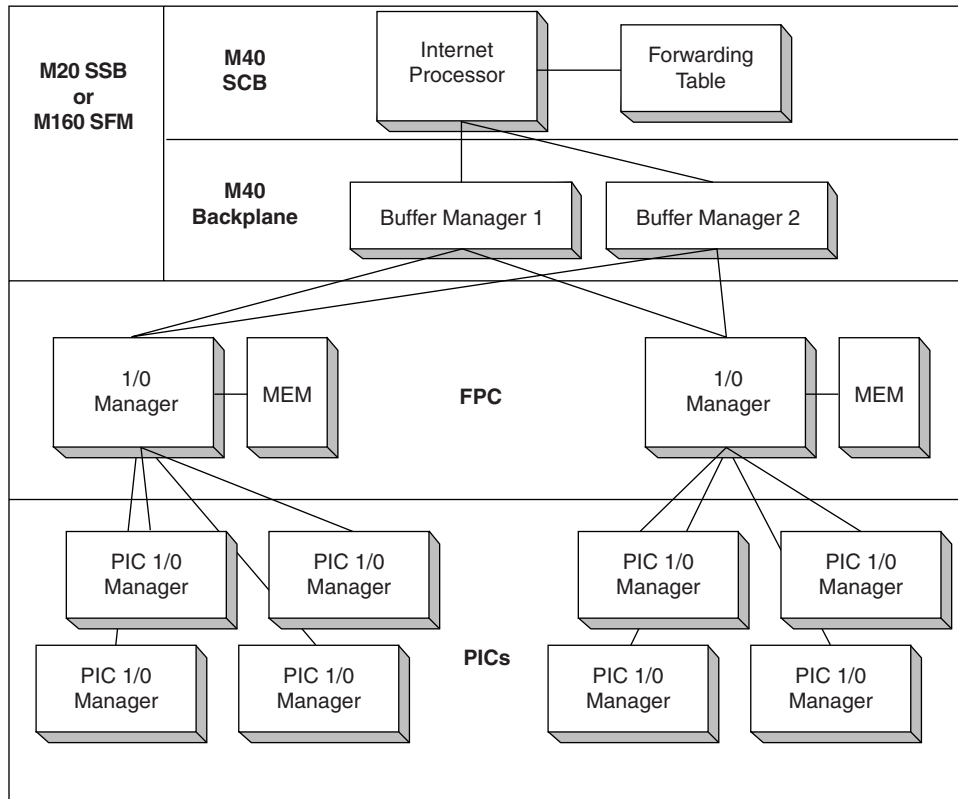


Figure 3–9 System ASICs

Starting from the bottom of Figure 3–9, you can see that each of the PICs contain at least one I/O manager ASIC responsible for media-specific tasks, such as encapsulation. The packets pass through these I/O ASICs on their way into and out of the router. The I/O manager ASIC on the PIC is specifically responsible for the following:

- Managing the connection to the I/O manager ASIC on the FPC
- Managing link-layer framing and creating the bit stream

- Performing cyclical redundancy checks (CRCs)
- Detecting link-layer errors and generating alarms, when necessary

On the FPC is another I/O manager ASIC. This ASIC takes the packets from the PICs and breaks them into 64-byte memory blocks, also known as J-cells, for storage in shared FPC memory. It is at this point that accounting is performed and CoS policies, which define the handling of traffic based upon classification of types of service, are implemented. This ASIC is specifically responsible for the following:

- Breaking incoming packets (as bit streams) into 64-byte blocks, or J-cells
- Sending the J-cells to the first DBM
- Decoding encapsulation and protocol-specific information
- Counting packets and bytes for each logical circuit
- Verifying packet integrity
- Applying CoS rules to packets

The first DBM ASICs encountered are responsible for receiving the J-cells and spreading them across the shared memory. In the M40, it is the backplane that contains the DBM ASICs; on the M5, M10, M20 and M160, the DBM ASICs are on the control boards.

In parallel, the first DBM ASIC passes forwarding-related information extracted from the packets to the Internet processor, which then performs the route lookup and sends the information over to a second DBM ASIC. The Internet processor ASIC also collects exception packets and sends them to the routing engine. This second ASIC then takes this information and the 64-byte blocks and forwards them to the I/O manager ASIC of the egress FPC—or multiple egress FPCs, in the case of multicast—for reassembly.

The DBM ASICs are responsible for the following:

- Managing the packet memory distributed across all FPCs
- Extracting forwarding-related information from packets
- Telling the FPC where to forward packets

The Internet processor ASIC is responsible for the following:

- Extracting next-hop information from the forwarding table
- Passing the next-hop information to the second DBM ASIC
- Collecting exception packets to send to the routing engine

The I/O manager ASIC on the egress FPC can perform some value-added services. In addition to incrementing TTL values and re-encapsulating the packet for handling by the PIC, it can also apply CoS rules. To do this, it may queue a pointer to the packet (never the packet itself) in one of four available queues, each having a share of link bandwidth, before applying the rules to the packet. Queuing can be based on destination address, the random early detection (RED) or weighted RED (WRED) algorithm, the value of precedence bits, and so on. Thus, we can say that the I/O manager ASIC on the FPC is responsible for the following:

- Receiving the J-cells from the second DBM ASIC
- Incrementing TTL values, as needed
- Queuing a pointer to the packet, if necessary, before applying CoS rules
- Re-encapsulating the J-cells
- Sending the encapsulated packets to the PIC I/O manager ASIC

Packet Flow

Now that you have a little background information on the various ASICs, it is helpful to see exactly how a packet moves through the router. Knowing how the packets move through the router can help clarify what you have learned about the architecture of the router. First, take a look at Figure 3–10, and then read the explanations below to see how the forwarding decisions are made.

The router first receives a packet on an *ingress*, or *incoming*, PIC. The PIC I/O manager ASIC performs the type of checksum and frame checks that are required by the type of medium it serves. Once this is done, the packet is passed, as a serial bit stream, to the FPC that houses the PIC.

The I/O manager ASIC on the FPC performs the important functions of decapsulating the packet, parsing it, and breaking it up into 64-byte memory blocks, before passing it to the first DBM ASIC. At this point, the packet is first written to memory. The DBM ASIC writes all packets to packet buffer memory, which is distributed across all FPCs on the router. While writing the packets to buffer memory, the DBM ASIC is also extracting information on the destination of the packet.

Once destination information is determined, it is sent to the Internet processor ASIC, which performs the lookup in the forwarding table. Note that the forwarding table is not omnipotent. It can handle unicast packets that do not have options, such as accounting, set. It can also handle multicast packets for which it already has a cached entry. All other packets must go to the routing engine for advanced lookup and resolution. If the PFE can handle the forwarding of the packet, it finds the next hop and egress interface. The packet is then forwarded to the second DBM ASIC, which passes the packet to the I/O manager ASIC on the FPC of the egress interface.

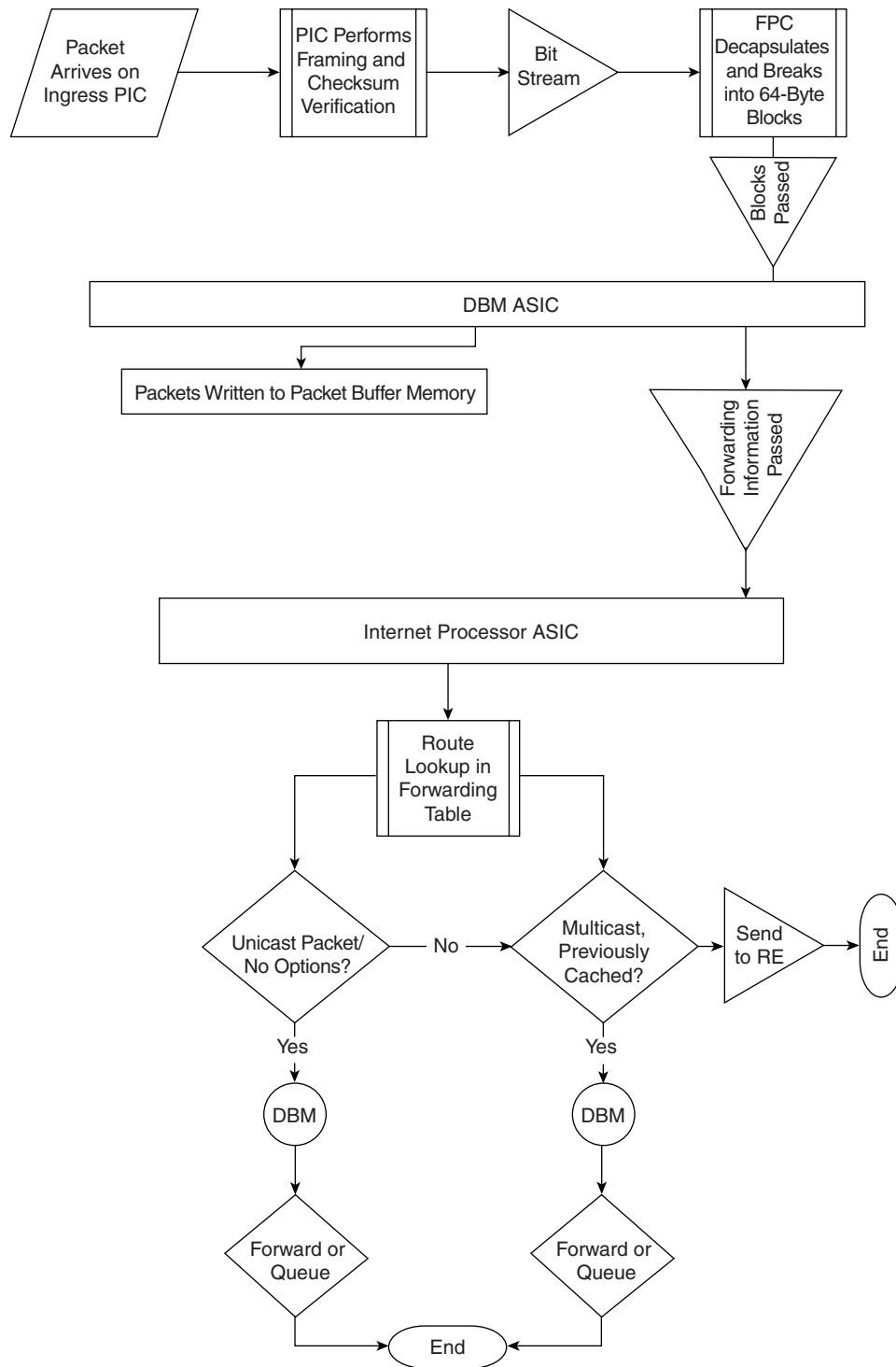


Figure 3-10 Packet Flow Forwarding Decisions

Now the packet may be queued. Actually, as stated earlier it is a pointer to the packet that is queued. The packet itself remains in the shared FPC memory. All queuing decisions and CoS rules are applied in the absence of the actual packet. When the pointer for the packet reaches the front of the line, the I/O manager ASIC sends a request for the packet to the second DBM ASIC. The DBM ASIC reads the J-cells from shared memory and sends them to the I/O manager ASIC on the FPC, which then serializes the bits and sends them to the media-specific ASIC of the *egress interface*.

The I/O manager ASIC on the egress PIC applies the physical-layer framing, performs the CRC, and sends the bit stream out over the wire.

3.2.2.2 Model Differences in Control Boards

Each model of router has its own component that performs part of the overall function of the PFE. Each board will be described in a little more detail, but the operations performed by them are similar in nature. Each board communicates with the routing engine through a 100Mbps internal interface and with the FPCs through 10Mbps interfaces on an internal hub. The primary functions of the control boards are as follows:

- Reset of FPC when abnormal behavior is detected—The board will attempt to reset the FPC up to three times. If unsuccessful, the control process takes the FPC offline and sends a notification to the routing engine.
- Transfer of control and exception packets—The control board handles nearly all exception packets—those packets for which there is no known path to the destination—passed to it from the Internet processor ASIC. The board may then pass exception packets to the routing engine. It may also communicate errors to the routing engine via *syslog* messages.
- Route lookups—A copy of the forwarding table is stored in SSRAM. When packets are received to be processed, the Internet processor ASIC performs the lookup on this table, makes a forwarding decision, sends a message to the midplane about the decision, and forwards the packets to the egress interface.
- System monitoring—The control board keeps tabs on the condition of the router based on information it receives from sensors. If an abnormal condition is detected, the board immediately notifies the routing engine.

On most of the M-Series routers, the control board is hot-pluggable, meaning that the router need not be powered down to install or uninstall the control boards. A brief service interruption will occur, usually about 500 ms. On the M5 and M10

routers, however, the FEB is not hot-pluggable. The system must be powered down for maintenance and replacement of this board.

M5 and M10 FEB

The M5 and M10 routers are the newest of the M-Series line. Despite their small footprint, these powerful routers also use control boards in the PFE. The FEB installs in the rear of the M5 or M10 chassis, just above the power supply. The FEB on the M5 and M10 is neither hot-removable nor hot-insertable! You must power down the chassis before removing or inserting these boards.

The FEB contains the following:

- A processor
- The Internet Processor II ASIC
- Two DBM ASICs
- I/O ASIC with 1MB SRAM—one on the M5, two on the M10
- 33MHz PCI bus connecting the system ASICs

The FEB also has its own storage—four slots of 2MB RAM to store forwarding tables associated with the ASICs, 64MB DRAM for the microkernel, EEPROM for the storage of the serial number and version of the FEB, and 512MB flash EPROM, which is programmable.

M20 SSB

On the Juniper Networks M20 router, the SSB installs from the front of the chassis into the upper-most slot. The SSB contains the Internet Processor II ASIC and two DBM ASICs. The SSB has its own processor and its own storage—four slots of 2MB RAM to store forwarding tables associated with the ASICs, 64MB DRAM for the microkernel, EEPROM for the storage of the serial number and version of the SSB, and 512MB flash EPROM, which is programmable.

M40 System Control Board

On the Juniper Networks M40 router, the system control board (SCB) installs from the front of the chassis into the center slot. The SCB contains its own processor, a PCI bus and the Internet processor ASIC, as well as 1 to 4MB SSRAM (for forwarding tables), 64MB DRAM (for the microkernel), EEPROM (which stores the SCB serial number and version), and 512MB flash EPROM (programmable).

M160 Switching and Forwarding Module

On the Juniper Networks M160 router, up to four interconnected switching and forwarding modules (SFMs) can be configured. Each SFM is a two-board system containing the following components:

- Internet Processor II ASIC for route lookups and forwarding
- Two DBM ASICs, one to send packets to the output buffer and another to communicate notifications to the I/O ASIC on the FPCs
- 8MB parity-protected SSRAM
- A processor subsystem for the handling of exception and control packets
- EEPROM for storage of board serial number and version information
- LEDs and an offline button for use prior to module removal

As stated earlier in this section, the M160 control board may be removed without a complete service interruption. There will, however, be a pause of about 500 ms while the router redistributes the functions to all other SFMs still inserted in the chassis.

3.2.3 PFE Clock Generator

The Juniper Networks M160 router has an additional unique feature—an added board that acts as a clock source. The PFE clock generator (PCG) is located in the rear of the chassis, beside the routing engine. The PCG supplies a 125MHz clock source to the ASICs and modules that are part of the PFE. The M160 has two PCGs installed for redundancy. These PCGs are field replaceable and hot-pluggable.

The PCG has three LEDs—one to indicate an OK state, one to indicate a Fail condition, and one that will illuminate if the PCG is the master. In addition, there is an offline button that will permit the user to take the PCG offline before removing it.

3.3 Management and Traffic Interfaces

This section will introduce you to the two types of interfaces available on the routers: management and traffic. A *management* interface is a physical or virtual port through which the router can be configured, maintained, or monitored, but which does not route traffic. A *traffic* interface is one through which routable network conversations are forwarded.

Several methods are provided that permit for the management and administration of the routers. These interfaces to the router include the following:

- SNMP—Network engineering staff or administrators can not only learn about the health and activity of the router through SNMP, but can also con-

figure it from a network management workstation using any popular SNMP tool. The benefit of this is that it makes configuration management simple. Past configurations can be archived (and dated) on the management station. It also means that many remote routers can be managed and configured from a central workstation.

- **User CLI**—The user process running on the routing engine permits management of the router through the CLI. The network engineer or administrator can, in this way, configure routing protocols, interface specifics, and systemwide instructions through a console, workstation, or laptop.
- **Craft interface**—The craft interface, as we discussed in Section 3.2.1.4, provides a window into the operations of the router—its health, uptime, and alarms. The craft interface also allows the administrator to take an FPC offline for removal and maintenance.

Table 3-4 shows the types of traffic interfaces that each M-Series model can support:

Table 3-4 *Traffic Interface Types per Model*

PIC Type	M5 and M10	M20 and M40	M160
ATM 4-port DS-3 4-port E3 2-port OC-3/STM-1 MM and SMIR 1-port OC-12/STM-4 MM and SMIR	Uses all ATM types	Uses all ATM types	Uses all ATM types
Channelized DS-3 2-port DS-3 with 28 T1 channels per port 4-port DS-3 with 28 T1 channels per port	Uses both	Uses the 4 port only	Uses the 4 port only
Channelized OC-12 to DS-3 1-port OC-12 SMIR with 12 DS-3 channels per port	Yes	Yes	Yes
Channelized STM-1 to E1 1-port STM-1 SMIR with 63 E1 channels per port	Yes	Yes	Yes
DS-3 2-port 4-port	Uses both	Uses the 4 port only	Uses the 4 port only
E1 4-port	Yes	Yes	Yes

(continued)

98 Juniper Networks Router Architecture

Table 3-4 (continued)

PIC Type	M5 and M10	M20 and M40	M160
E3 2-port 4-port	Uses both	Uses the 4-port only	Uses the 4-port only
Fast Ethernet 4-port 48-port	Uses the 4-port only	Uses the 4-port only	Uses both
Gigabit Ethernet 1-port LH, LX, SX 2-port LX and SX 4-port SX	Uses the 1-port only	Uses the 1-port only	Uses all
SONET/SDH			
2-port OC-3c/STM-1 MM and SMIR	Yes	No	No
4-port OC-3x/STM-1 MM and SMIR	Yes	No	No
1-port OC-12c/STM-1 MM and SMIR both in concatenated and nonconcatenated modes	Yes	No	No
4-port OC-3c/STM-1 MM and SMIR	No	Yes	Yes
1-port OC-12c/STM-4 MM and SMIR both in concatenated and nonconcatenated modes	No	Yes	Yes
4-port OC-12c/STM-4 MM and SMIR	No	No	Yes
1-port OC-48c/STM-16 SMSR concatenated and nonconcatenated modes	No	Yes	Yes
1-port OC-48c/STM-16 SMLR concatenated and nonconcatenated modes	No	Yes	No
1-port OC-192x/STM-64 SR2 and LR both in concatenated and nonconcatenated modes	No	No	Yes
T1 4-port	Yes	Yes	Yes
Tunnel Services PIC	Yes	Yes	Yes

The difference between the models lies primarily in the number of ports supported and in the type of throughput that is available on the backplane. Table 3–5 lists the type of throughput, the number of PICs supported, and the number of ports for each model.

Table 3–5 *Port Density by Model*

Model	M5	M10	M20	M40	M160
Full-Duplex Throughput	6.4Gbps	10Gbps	20Gbps	40Gbps	160Gbps
Target Network Size	Medium to Large	Medium to Large	Medium to Large	Large	Very Large
Number of PICs Supported	4	8	16	32	32
Number of Ports	Up to 16	Up to 32	Up to 64	Up to 128	Up to 32 OC-12 or Up to 32 OC-48 or Up to 8 OC-192

It is important to note that, by default, all physical interfaces on the Juniper Networks routers use PPP, but can be configured to use other Layer 2 encapsulation types. If an interface is of a type that does not support PPP, you must configure the appropriate encapsulation type. For specific information about interface encapsulation options, including configuration examples, please refer to Chapter 8.

3.4 Cooling Systems

With the exception of the M5 and M10 routers, each Juniper Networks router incorporates redundancy in the cooling system. This section provides a general familiarity with the airflow through each model of the M-Series routers.

Caution: Never operate the router for more than one minute without the fan assembly installed! The function of the fans is to keep all components cooled to an optimal operating level. Running the unit without fans in place could void your warranty and limit any maintenance available to you. Check with your Juniper Networks representative for more information.

3.4.1 M5 and M10

The single fan assembly on either an M5 or M10 is on the left side of the router if you are facing the front of the chassis. The fan assembly contains four fans that pull air from the left to the right across the PICs. Juniper Networks suggests leaving six inches on either side of any installed M-Series router, as shown in Figure 3-11, to ensure adequate airflow.

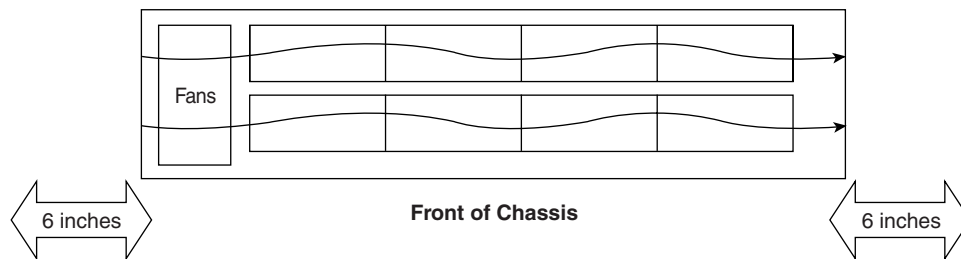


Figure 3-11 M5/M10 Airflow

3.4.2 M20

The M20 router has four fan assemblies—three on the front of the chassis and one in the rear. The three fans on the front left side of the chassis provide side-to-side cooling for the FPCs and for the SSB. The rear fan assembly is located to the right of the routing engine and provides cooling directly to the routing engine. All four fan assemblies plug directly into the router midplane. In addition, the power supplies on the M20 router have integrated fans, providing built-in power-supply cooling. The airflow model is shown in Figure 3-12.

3.4.3 M40

The cooling system for the M40 router is a little more complex, consisting of three separate subsystems: the impellers, the integrated power-supply fans, and the triple fan assemblies.

The router contains two sets of redundant impellers, located at the top of the chassis and at the bottom of the card cage. These impellers pull air in through the front of the card cage and across the PFE, forcing the exhaust out the back of the chassis, thus cooling the PFE. You can see this process in Figure 3-13. The impellers are designed to run at less than full capacity unless a condition is detected, such as a rise in temperature, which would increase the cooling needs. At that point, the impellers can adjust the fan speed to meet the new requirements. An air filter protects the impeller from drawing in foreign objects that could

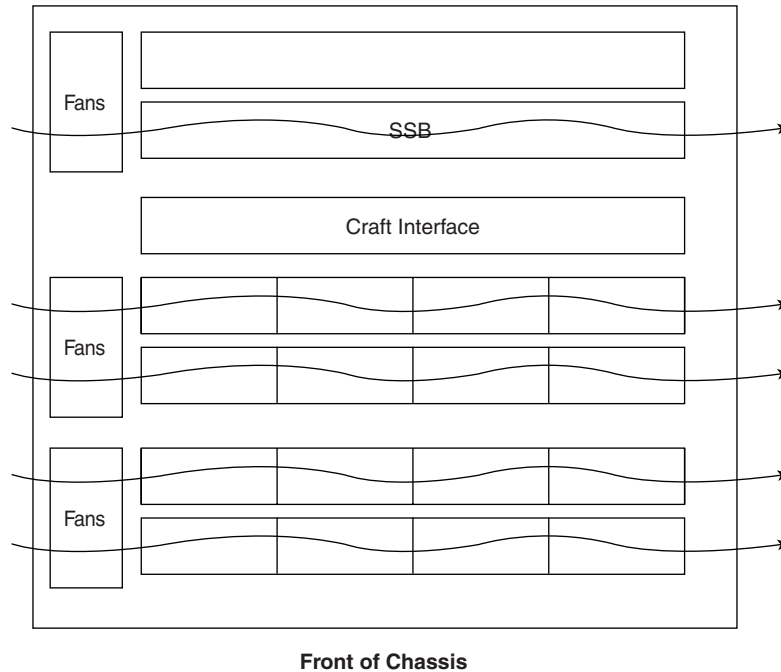


Figure 3-12 *M20 Airflow*

damage the fans. It is a good idea to keep the air filter in place when the router is operational.

A set of three load-sharing fan trays, located at the upper rear of the chassis, pull in air through a filter and intake at the front of the chassis to keep the routing engine cool. Because the fan trays are load-sharing, if one fan tray is removed, the others remaining will adjust to meet the current cooling requirement.

Finally, the power supplies on the M40 router have integrated cooling fans, just like those on the M20.

3.4.4 M160

The M160 router also uses impellers and fan trays to keep the system cool. A front cooling system uses an upper impeller that works with the fan tray (installed in front) to pull air through the front of the chassis, up through the card cage, and then sends the exhaust out the rear of the chassis.

The rear cooling system uses two impellers in the upper and lower part of the chassis to pull air in over the routing engine, SFM, MCS, and PCGs. As you can see in Figure 3-14, the air is drawn in through the front of the chassis, through the air

102 Juniper Networks Router Architecture

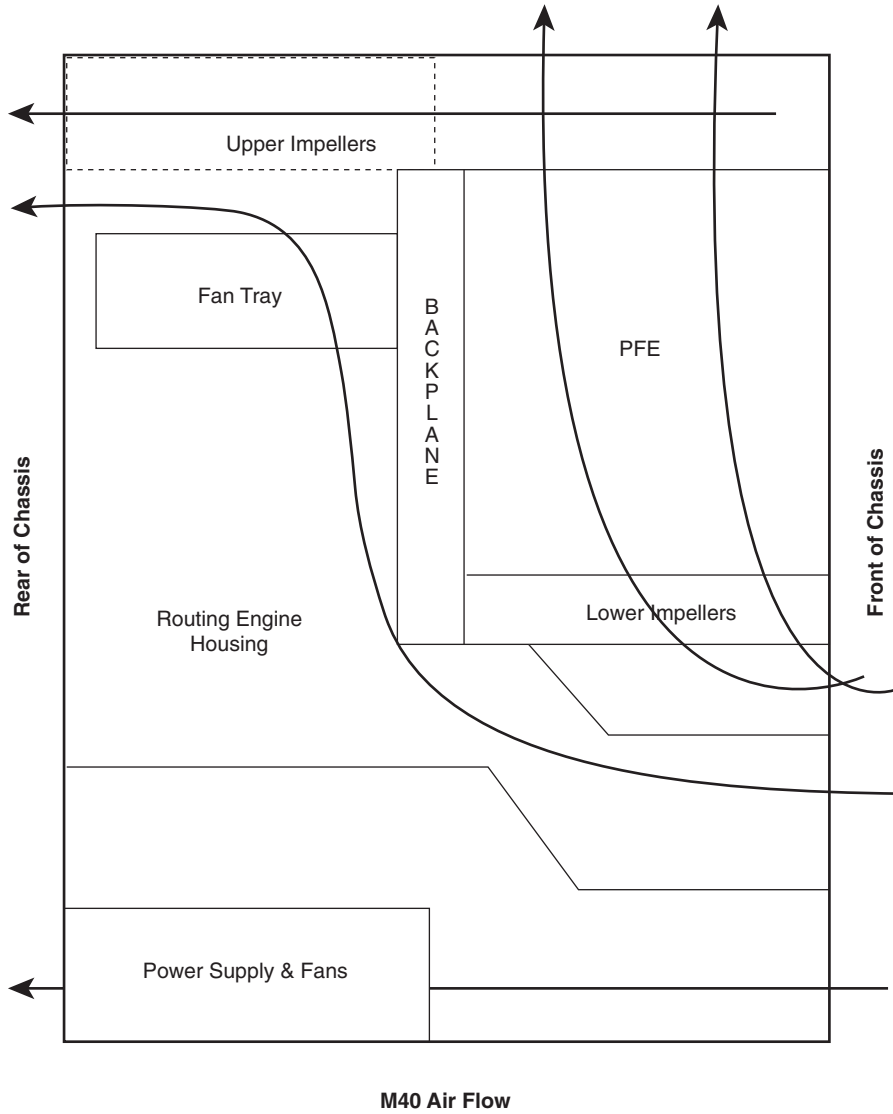


Figure 3-13 M40 Airflow

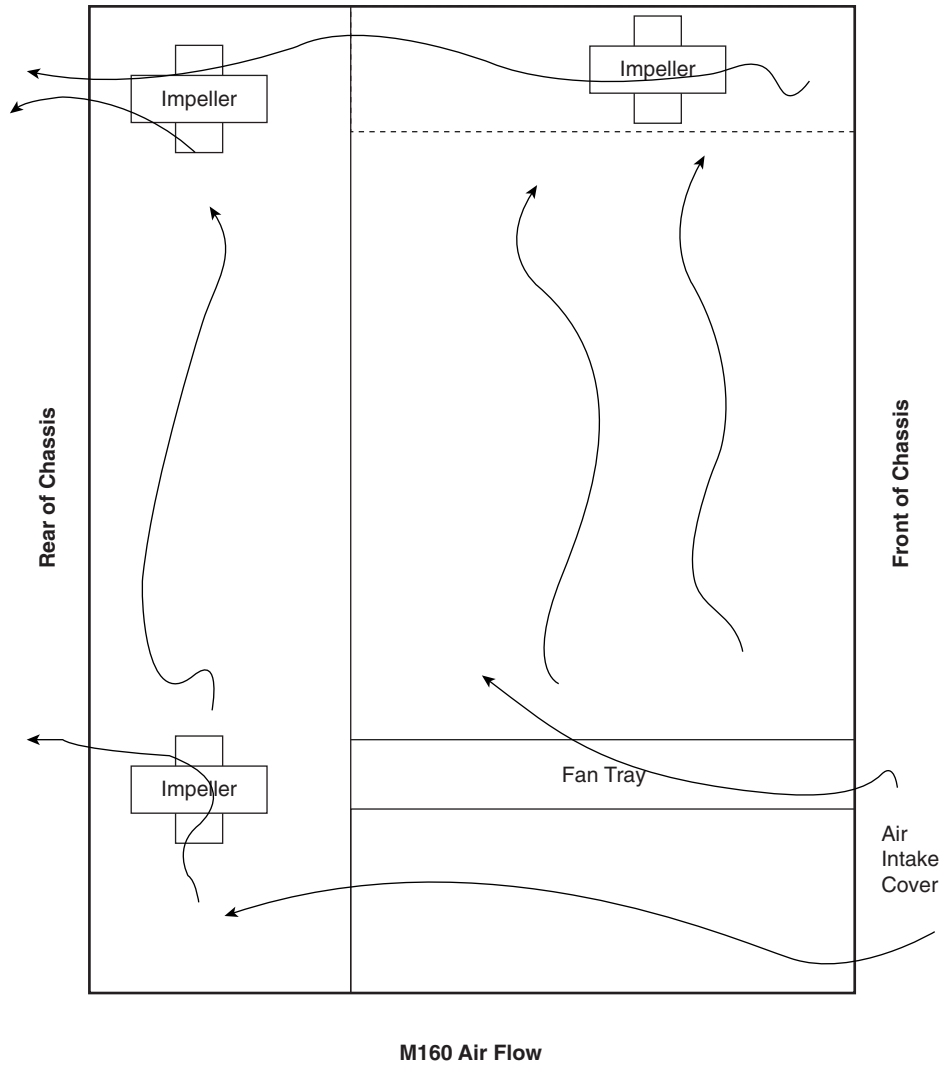


Figure 3-14 *M160 Airflow*

intake cover, and the exhaust is sent out the rear of the chassis. The power supplies are cooled by the front and rear cooling systems and do not have integrated fans.

3.5 Router Power-up and Boot Process

When the router is shipped, it comes with the latest version of JUNOS software loaded on nonrotating flash memory. Additional copies are provided on the hard disk and on a PC card that is shipped with the router. When the router is powered-up for the first time, it looks for the software in the following sequence:

1. PC card, if installed
2. Flash memory
3. Hard disk

The first copy that it encounters is the one it will use. Therefore, it is important to not insert a PC card if you want to boot from flash memory, for example.

Note: It is a good idea to keep the original PC card (as well as any copies of future releases) in a safe place in case the file system becomes corrupted or unstable. This will allow you to revert to the original version by rebooting the router with the PC card in place.

To power-up the router, perform the following steps:

1. Power-up the management device that is connected to the console, auxiliary, or management Ethernet port on the routing engine.
2. Turn the power on for each power supply.
3. Check the power supply OK LEDs and the output to the management device to ensure that the router booted properly.

Note: If nothing is plugged into the management Ethernet interface, a RED alarm will be generated. Plug in any device cable with an RJ-45 connector to prevent this alarm.

Other visible activity to check at startup includes the following:

- Craft interface displays, such as Starting routing engine, Starting PFE, and Starting Card
- FPC LEDs, which blink green until testing is complete (if all tests pass, the lights should be solid green)
- Alarm LEDs, as appropriate
- PIC LEDs, which remain off unless the interfaces have been configured

3.5.1 Configuring the Router

During the initial configuration of the router, you will need to configure the following:

- A password for user ROOT. This can be set in three different ways:
 1. Plain text (not logged, if logging is enabled):

```
root@router# set system root-authentication plain-text-password password (password is prompted on a separate line)
```
 2. Pre-encrypted:

```
root@router# set system root-authentication encrypted-password password
```
 3. Secure shell (SSH) key (on domestic U.S. systems only):

```
root@router# set system root-authentication ssh-rsa key
```
- A hostname for the router: [edit]

```
root@router# set system host-name name
```
- The domain name:

```
[edit] root@router# set system domain-name domain
```
- The IP address and subnet mask of the management Ethernet port:

```
[edit] root@router# set interfaces fxp0 unit 0 family inet address ip-address/prefix-length
```
- The default route:

```
[edit] root@router# set system backup-router gateway-address root@router# set routing-options static route default nexthop gateway-address retain no-advertise
```
- The domain name server (DNS) IP address:

```
[edit] root@router# set system name-server dns-address
```
- At least one NON-ROOT user:

```
[edit] root@router# set system login user username class class authentication plain-text-password
```

106 Juniper Networks Router Architecture

Note: It is vitally important to configure a NON-ROOT user, as ROOT cannot Telnet into the router!

After configuring the preceding, save your changes:

```
[edit]
root@router# commit
```

3.6 JUNOS Software Upgrade Procedure

Juniper Networks releases several new versions of JUNOS software each year, as needed. This section will prepare you to perform a JUNOS software upgrade to the router.

Every JUNOS software release is actually a group of files bundled together. These files can be installed all at once or individually. Table 3–6 lists the files contained in the release.

Table 3–6 *JUNOS Upgrade Software Release Files*

Filename	Contents
jbase	Additions to JUNOS
jkernel	The operating system package
jroute	The routing engine software
jpfe	The PFE software
jdocs	Updated online reference documentation
jcrypto	Security software (U.S. domestic only)
jbundle	All of the files combined

To see the software that is currently installed on the router, use the following command:

```
root@router> show system software

Information for jbase:
Comment:
JUNOS Base OS Software Suite [5.0R3.3]

Information for jcrypto:
Comment:
JUNOS Crypto Software Suite [5.0R3.3]
```

Information for jdocs:

Comment:

JUNOS Online Documentation [5.0R3.3]

Information for jkernel:

Comment:

JUNOS Kernel Software Suite [5.0R3.3]

Information for jpfe:

Comment:

JUNOS Packet Forwarding Engine Support [5.0R3.3]

Information for jroute:

Comment:

JUNOS Routing Software Suite [5.0R3.3]

Information for junos:

Comment:

JUNOS Base OS boot [5.0R3.3]

To upgrade your software, there are several simple steps (following this list, we will explain each item in more detail):

1. Download the software package(s) you need from the Juniper Networks Web site.
2. Perform a system backup on the router you wish to upgrade.
3. Copy the software package(s) to the router you wish to upgrade.
4. Add the new software package(s) to the router you wish to upgrade (this will effectively delete the old packages).
5. Reboot the router.

3.6.1 Downloading the Software

When you download JUNOS software from the Juniper Networks Web site at www.juniper.net/support using your authorized username and password, you will notice that the packages use a standard naming convention. The format is *package-x.yZnumber.tgz*:

- *Package* is the name of the file (such as *jbundle*).
- *x.y* is the software version.
- *Z* is the type of release. For most customers, this will be an R for released version. For some customers, it may be A (alpha release), B (beta release), or I (internal release).
- *number* is the software release number and build, if applicable.

108 Juniper Networks Router Architecture

3.6.2 Backing Up the System

You can create a recoverable snapshot of the current system, if it is stable, before proceeding. Using this command, however, will make the running and backup versions of the software identical and will mean that you cannot revert back to the original version that shipped with the router. To make the snapshot, use the following command:

```
root@router# request system snapshot
```

This will back up the /root file system to /altroot and the /config file system to /altconfig on the hard disk.

3.6.3 Copying the Package(s) to the Router

After you have backed up your system files, copy the new software bundle to the router's hard disk, using a command such as the following:

```
root@router# file copy
ftp://username:prompt@ftp.hostname.net/ filename
/var/tmp/filename
```

This will copy the file from an FTP server to the /var/tmp directory on the router's hard disk. This is simply an example of one way to copy the file. The M40 also has an LS-120 floppy drive that can be used for the storage and transfer of files. When installing a new software version, you should do so from an out-of-band management source, such as the console. An in-band source, such as Telnet, could be lost while you are upgrading.

3.6.4 Adding the Package(s)

Once the files have been copied to the hard disk, upgrade the software using the following command:

```
root@router> request system software add package-name
```

Checking available free disk space...11600k available, 6076k suggested...

```
Installing package '/var/tmp/jbundle-package-name' ...
Auto-deleting old jroute...
```

```
Auto-deleting old jdocs...
Auto-deleting old jpfe...
Auto-deleting old jkernel...
Adding JUNOS base software release-number...
Adding jkernel...
Adding jpfe...
Adding jdocs...
Adding jroute...
NOTICE: uncommitted changes have been saved in
/var/db/config/juniper.conf.pre-install
Saving package file in /var/sw/pkg/jbundle-package-name ...
```

As you can see in the example, once you begin to install the new jbundle, the system deletes the old information and explodes (expands) the zipped files contained in the new bundle into the `/var/sw/pkg/jbundle-package-name` file.

3.6.5 Finishing the Upgrade

Once the software upgrade has completed, you should perform a system reboot. This is the last step in upgrading the JUNOS software package(s). Use the following command:

```
root@router> request system reboot
```

Note: Special instructions for upgrading to version 5.0 or reverting to an earlier release from version 5.0 are available on the Juniper Networks Web site at www.juniper.net/techpubs/software/junos50/swconfig50-getting-started/html/getting-started-upgrade50.html#1017395.

3.7 JUNOScript

As of JUNOSv4.3B2, the JUNOScript feature is available. This feature provides an alternate method for communicating information to and from the router. This method is an extensible-markup-language (XML) interface for communicating with the router. Because the JUNOScript API is an XML interface, it enables you to build client applications for monitoring and communicating information to and from the router using Perl. JUNOScript can also support HTML through the use of CGI scripts.

110 Juniper Networks Router Architecture

JUNOScript needs to be configured on a UNIX machine. The process to enable a client application can be found on the Juniper Networks Web site at www.juniper.net/support/junoscript. This link lists the required software components needed for the installation, as well as links to instructions and sample Perl scripts for use with JUNOScript.

3.8 Chapter Summary

This chapter covered information about the M-Series routers, their architecture, port density, and design, including the latest models introduced by Juniper Networks—the M40e and the G10. By focusing on the structure, process flow, and components of each router model, the material has given you a solid basis by which to understand and apply the information that you will be given in the rest of this book.

After reading this chapter, you should have a good idea of how routing takes place through a Juniper Networks router, the function of the ASICs, and how fan trays and impellers keep the system components cool. You have also been familiarized with the router startup process and in what order the boot image locations are searched.

In addition to information on architecture and processes, you were also given information on upgrading the system software and some simple system administration. These topics will be covered in detail later in the book.

Bibliography

Ericsson IP Infrastructure, AXI-520/580 course material.

www.juniper.net

www.lightreading.com