

Chapter 12

PGM Overview

Multicast applications often require real-time operation. These applications cannot take advantage of Transmission Control Protocol (TCP) reliability features such as sequencing, retransmission, and flow control through windowing between sender and receiver. The User Datagram Protocol (UDP), the major transport layer alternative to TCP, leaves much to be desired in its reliability for multicast traffic. Pragmatic General Multicast (PGM) is a special protocol layer for multicast traffic that can be used between the IP layer and the multicast application to add reliability to multicast traffic. PGM allows a receiver to detect missing information in all cases and to request replacement information if the receiver application requires it. PGM is IP protocol number 113.

Although PGM is mainly concerned with the operation of multicast source and receiver, PGM-enabled routers (called PGM network elements) play a *router assistance* role in the initial delivery and potential replacement of multicast traffic. PGM routers are not mandatory in PGM, but they can provide the following benefits when placed anywhere between the source and receivers:

- Reduce the load on the multicast source by aggregating duplicate messages to the source. PGM routers are required to perform this function.
- Limit the flooding of *repair data* (replacement information) to only those downstream receivers that requested the repair data. PGM routers are required to perform this function.
- Act as *designated local repairers* (DLRs) by caching the repair data and resending it to receivers that request it later. DLR functions are a PGM option, and PGM routers are not required to perform this role.

PGM adds reliability to multicast traffic streams. It is not a complete multicast protocol like the Distance Vector Routing Multicast Protocol (DVMRP) or Protocol Independent Multicast (PIM). Adding PGM to a router does not enable the router to perform multicast functions. Instead, a PGM router with multicast capabilities and a preconfigured multicast protocol such as PIM can offer more reliable multicast services to PGM sources and receivers. PGM is not an alternative to multicast routing protocols, but an enhancement of the multicast capabilities already present and configured on the router.

This chapter provides the following information about PGM:

- PGM Architecture and PGM Routers on page 100
- PGM Configuration Statements on page 104
- Summary of PGM Configuration Statements on page 105

For information about supported standards for PGM, see “IP Multicast Standards” on page 28.

PGM Architecture and PGM Routers

PGM is defined in RFC 3208 and forms a reliable transport layer for multicast applications. Almost any multicast application can use PGM. Applications most suitable for PGM include stock market ticker update information, news reports, weather warnings, and other information that must reach multiple listeners in its entirety and in a timely fashion.

The basic PGM architecture consists of a multicast content source, one or more receivers, and zero or more routers between source and receivers. All end devices must be PGM-enabled, although there can be non-PGM routers between source and receiver. If all routers are non-PGM routers, then no routers are capable of the PGM router assistance function, and all PGM functions take place directly between source and receiver.

PGM sources send sequenced content in *sessions* to receivers, using multicast protocols. Other, non-PGM protocols allow receivers to learn about a particular source, its sessions, and its location. PGM receivers listen to multicast *original data (ODATA)*, detect missing content through the sequence numbers, and send unicast *negative acknowledgments (NAKs)* back to the source. NAKs are answered by multicast *NAK confirmations (NCFs)*, which suppress any NAKs from receivers on the same subnet that have not yet sent a NAK upstream. The source sends multicast *repair data (RDATA)* to receivers containing the missing content. PGM routers assist in this process by making sure that the negative acknowledgments follow the same path as the outbound content upstream to the source, and by suppressing duplicate negative acknowledgments and repair information.

PGM sources must maintain a sliding window of retransmittable information. There is no concept of group membership in PGM, so receivers never need to communicate with the source unless they request repair data with a negative acknowledgment. However, this means that the PGM source determines the window size for each receiver, in contrast to almost all other protocols, and requires a certain processing power in each receiver. The absence of positive receiver-to-source acknowledgments also means that PGM scales well and cuts down on control message traffic that can easily overwhelm a multicast network.

PGM receivers can start receiving a PGM session from a PGM source at any time and request any missing previous information that the receiving application needs. If the session is long enough or the transmit window small enough so that the source does not maintain a long session history, the receiver cannot get all required information.

This section describes in more detail the behavior of the three PGM elements in a multicast network:

- PGM-Enabled Source on page 101
- PGM-Enabled Receivers on page 102
- PGM-Enabled Routers on page 102

PGM-Enabled Source

A PGM-enabled source of multicast content generates sequenced packets of ODATA that are multicast to receivers. Interleaved with the content packets are *source path messages (SPMs)*, which tell PGM routers and receivers about their upstream next-hop PGM device—either another PGM router or the PGM source.

ODATA packets and SPMs are multicast from the source. A PGM router always appends its own IP address to the SPM before it is multicast on the downstream interfaces. The SPMs are sent by the source and upstream PGM routers with the router alert option set in the IP headers so that PGM routers do not have to examine every packet in the session for SPM packets.

The PGM source acknowledges a received NAK by multicasting an NCF downstream to the next PGM device on the path to the receiver. NCFs make sure that PGM routers and receivers do not bombard sources with NAKs. Downstream PGM routers suppress all subsequent NAKs that indicate the same missing information once one NCF is received from the upstream device.

The PGM source also responds to NAKs by multicasting RDATA packets with the same sequence number as the one indicated by the NAK. RDATA packets have the router alert option set in the IP header so that PGM routers can distinguish them from ODATA packets.

PGM sources organize their packets in sessions. PGM sources are not required to retain copies of information older than the current session, although they might. Long sessions are not necessarily kept on the source in their entirety.

PGM sources identify themselves through a global source ID (GSID). This globally unique source identifier is formed from the low-order 48 bits of the Message Digest 5 (MD5) signature of the Domain Name System (DNS) name of the source.

PGM-Enabled Receivers

The PGM architecture requires one or more PGM-enabled receivers of the multicast content generated by a PGM source. PGM receivers accept all types of downstream PGM messages: ODATA, SPMs, NCFs, and RDATA.

Receivers process the ODATA packets as they arrive from the source, constantly checking the 32-bit sequence number in the ODATA PGM header for gaps in the sequence. If the receiver detects missing information, it generates a NAK for that sequence number. The NAK is unicast upstream to the PGM next hop, which is a router or the source, as determined by the last address in the received SPM.

A receiver knows that its NAK was received by the PGM next hop when it gets an NCF in response to its NAK. If several receivers on a subnet are missing the same ODATA packet, receivers getting an NCF for the packet before sending a NAK suppress the NAK. If a receiver does not get an NCF in response to a NAK, the receiving application can send a NAK again or continue, with the certainty that information is missing.

After the NCF, PGM receivers get an RDATA packet with the same sequence number indicated in the NAK and a copy of the missing ODATA. NCFs and RDATA can originate from the source or a router acting as a DLR for a subnet. The receiver now has complete information or knows for certain what is missing.

PGM receivers can request almost anything from the PGM source. However, because the source determines the window size, there is no guarantee that older information is available.

PGM-Enabled Routers

Multicast-capable routers can implement the PGM router assistance functions, although not all multicast routers must be PGM-enabled routers. Mandatory PGM router assistance functions include aggregating duplicate NAKs sent to the source to reduce the load on the multicast source, generating NCFs in response to NAKs, and flooding RDATA packets to only those downstream receivers that requested it with a NAK. Optionally, a PGM router can be a DLR, caching PGM information and cutting down on network traffic by resending RDATA packets locally.

There can be zero or more PGM-enabled network elements (routers) between the source and receiver. If there are no PGM routers between the source and receiver, then all PGM messages flow directly between the source and receiver, and no router assistance functions are possible. Both PGM and non-PGM routers can be freely mixed on a network because PGM is a transport layer protocol and is not involved with router multicast functions.

PGM routers also receive SPMs from the source or an upstream PGM router and forward them downstream, inserting the router's own downstream IP interface address into the SPM so that receivers always know their upstream PGM next hop.

When a PGM router receives unicast NAKs from a downstream PGM router or receiver, the router unicasts one NAK for each missing sequence number to the next-hop PGM device upstream toward the source. The address of the PGM next-hop device is determined by received SPMs.

The PGM router multicasts NCFs in response to received NAKs on the downstream interfaces that received the NAKs. NCFs are not multicast on interfaces that have not received NAKs.

PGM routers must multicast all ODATA and RDATA packets they receive from upstream PGM devices. Normal multicast protocols are used to determine downstream interfaces.

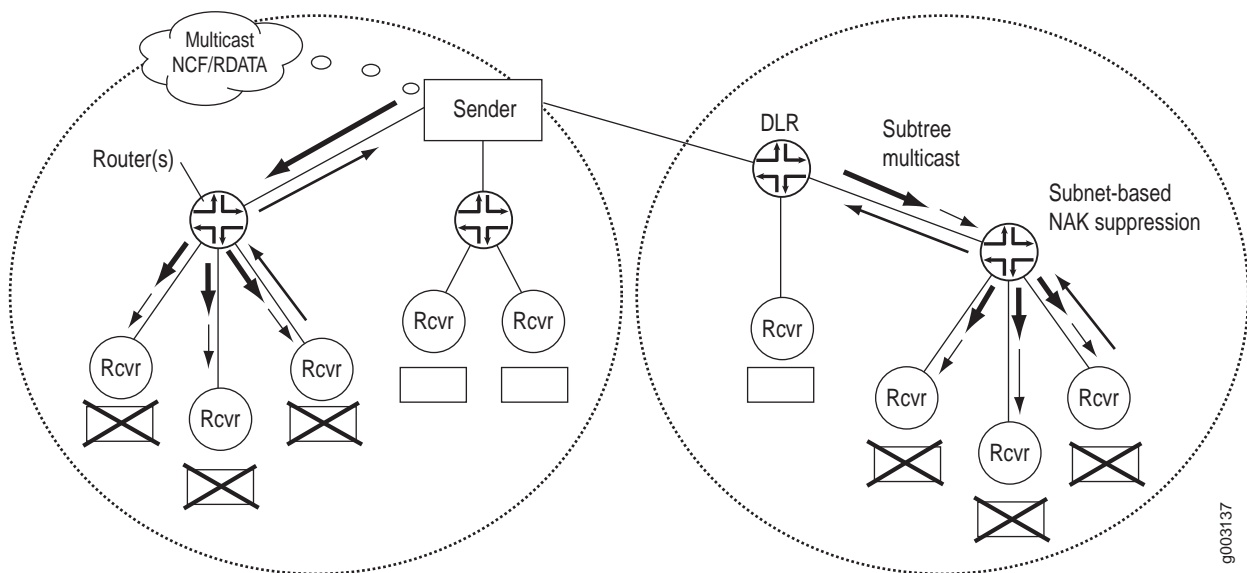
If the PGM router is a DLR, it responds to received NAKs with an NCF and with its own RDATA packet. NAKs are not forwarded upstream from a DLR.

Figure 9 shows the overall PGM architecture and the role of PGM-enabled routers.

Figure 9: PGM Architecture and General Operation

Case 1: RDATA from source in response to a NAK

Case 2: RDATA from DLR in response to a NAK



The figure shows only NAKs, NCFs, and RDATA flows. RDATA can come from either the source (left) or a DLR router (right). In both cases, unicast NAKs from a receiver are forwarded upstream by the routers, and multicast NCFs are generated downstream. Subnet NAK suppression is shown, as well as RDATA from the source or DLR sent only to the portions of the network requesting it.

PGM Configuration Statements

PGM allows the router to participate in defined PGM router assistance functions between PGM-enabled sources and receivers. Although PGM is a transport layer protocol and is not directly concerned with IP packet routing, PGM must be explicitly configured on the router.

To enable PGM globally on the router, include the `pgm` statement:

```
pgm;
```

You can include this statement at the following hierarchy levels:

- [edit protocols]
- [edit logical-routers *logical-router-name* protocols]

To trace the operation of PGM, include the `traceoptions` statement:

```
pgm {
  traceoptions {
    file name <replace> <size size> <files number> <no-stamp>
      <(world-readable | no-world-readable)>;
    flag flag <flag-modifier>;
  }
}
```

You can include this statement at the following hierarchy levels:

- [edit protocols]
- [edit logical-routers *logical-router-name* protocols]

You can specify the following PGM-specific options in the PGM `flag` statement:

- `all`—Trace all PGM packets.
- `init`—Trace all PGM initialization events.
- `packets`—Trace all PGM packet processing.
- `parser`—Trace all PGM parser processing.
- `route-socket`—Trace all PGM route-socket events.
- `show`—Trace all PGM `show` command servicing.
- `state`—Trace all PGM state transitions.

For an overview of logical routers and a detailed example of logical router configuration, see the logical routers chapter of the *JUNOS Feature Guide*.

By default, PGM is enabled on every interface of the router, but global, explicit configuration is required. No options are available for PGM operation.

Summary of PGM Configuration Statements

The following sections explain each PGM configuration statement. The statements are organized alphabetically.

pgm

Syntax `pgm {
 traceoptions {
 file name <replace> <size size> <files number> <no-stamp>
 <(world-readable | no-world-readable)>;
 flag flag <flag-modifier>;
 }
}`

Hierarchy Level [edit logical-routers *logical-router-name* protocols],
[edit protocols]

Release Information Statement introduced before JUNOS Release 7.4.

Description Configure PGM globally and set tracing options.

To specify more than one tracing operation, include multiple **flag** statements.

Default The default PGM trace options are inherited from the routing protocol **traceoptions** statement included at the [edit routing-options] hierarchy level.

Options The remaining statement is explained separately.

traceoptions

Syntax traceoptions {
 file *name* <replace> <size *size*> <files *number*> <no-stamp>
 <(world-readable | no-world-readable)>;
 flag *flag* <flag-modifier>;
 }

Hierarchy Level [edit logical-routers *logical-router-name* protocols *pgm*],
 [edit protocols *pgm*]

Release Information Statement introduced before JUNOS Release 7.4.

Description Configure PGM tracing options.

To specify more than one tracing operation, include multiple **flag** statements.

Default The default PGM trace options are those inherited from the routing protocol **traceoptions** statement included at the [edit routing-options] hierarchy level.

Options **disable**—(Optional) Disable the tracing operation. You can use this option to disable a single operation when you have defined a broad group of tracing operations, such as **all**.

file *name*—Name of the file to receive the output of the tracing operation. Enclose the name within quotation marks. All files are placed in the directory `/var/log`. We recommend that you place tracing output in the file `pgm-log`.

files *number*—(Optional) Maximum number of trace files. When a trace file named *trace-file* reaches its maximum size, it is renamed *trace-file.0*, then *trace-file.1*, and so on, until the maximum number of trace files is reached. Then the oldest trace file is overwritten.

If you specify a maximum number of files, you also must specify a maximum file size with the **size** option.

Range: 2 through 1000 files

Default: 2 files

flag—Tracing operation to perform. To specify more than one tracing operation, include multiple **flag** statements.

PGM Tracing Flags

- **all**—Trace all PGM packets.
- **init**—Trace all PGM initialization events.
- **packets**—Trace all PGM packet processing.
- **parser**—Trace all PGM parser processing.
- **route-socket**—Trace all PGM route-socket events.
- **show**—Trace all PGM **show** command servicing.

- **state**—Trace all PGM state transitions.

Global Tracing Flags

- **all**—All tracing operations
- **general**—A combination of the **normal** and **route** trace operations
- **normal**—All normal operations
Default: If you do not specify this option, only unusual or abnormal operations are traced.
- **policy**—Policy operations and actions
- **route**—Routing table changes
- **state**—State transitions
- **task**—Interface transactions and processing
- **timer**—Timer usage

flag-modifier—(Optional) Modifier for the tracing flag. You can specify one or more of the following modifiers:

- **detail**—Detailed trace information
- **receive**—Packets being received
- **send**—Packets being transmitted

no-stamp—(Optional) Do not place timestamp information at the beginning of each line in the trace file.

Default: If you omit this option, timestamp information is placed at the beginning of each line of the tracing output.

no-world-readable—(Optional) Disallow any user to read the log file.

replace—(Optional) Replace an existing trace file if there is one.

Default: If you do not include this option, tracing output is appended to an existing trace file.

size size—(Optional) Maximum size of each trace file, in kilobytes (KB), megabytes (MB), or gigabytes (GB). When a trace file named *trace-file* reaches this size, it is renamed *trace-file.0*. When *trace-file* again reaches its maximum size, *trace-file.0* is renamed *trace-file.1* and *trace-file* is renamed *trace-file.0*. This renaming scheme continues until the maximum number of trace files is reached. Then the oldest trace file is overwritten.

If you specify a maximum file size, you must also specify a maximum number of trace files with the **files** option.

Syntax: *xk* to specify KB, *xm* to specify MB, or *xg* to specify GB

Range: 10 KB through the maximum file size supported on your system

Default: 1 MB

world-readable—(Optional) Allow any user to read the log file.

Usage Guidelines See “PGM Configuration Statements” on page 104.

Required Privilege Level routing and trace—To view this statement in the configuration.
routing-control and trace-control—To add this statement to the configuration.