GRF GateD 1.4.10 Addendum

Ascend Communications

July, 1997

Appendix A:
GateD Configuration Statements

A

Appendix A contains the 1.4.10 updated GateD Statement descriptions, and replaces the information currently found in Chapter 2 of the *GRF Reference Guide 1.4*.

GateD Configuration Statements

2

This release of the Ascend Embedded Operating System supports version 3.5b3 of GateD.

Chapter 2 contains the following GateD configuration topics and statements:

Introduction

Dynamic routing is a mechanism that automatically adjusts to changes in a network topology so that traffic on the network continues to flow without interruption. Without dynamic routing, a network administrator would have to manually update all the necessary routing information by hand if, for example, a new machine is added to a networked system, or part of the network crashes and traffic must be rerouted.

In other words, dynamic routing makes it possible to maintain a network topology automatically, thus making for a more efficient and reliable network. For dynamic routing to accomplish its tasks, it employs an agent. A dynamic routing agent is software that uses dynamic routing protocols to exchange information about the state of the network. It then processes that information to decide how to configure a given router to route traffic through the network efficiently. Dynamic routing is implemented on a networked system via a dynamic routing agent. On the GRF system, the agent is called the Gate Daemon or simply GateD, and is based on the software produced by the Merit GateDaemon Project.

GateD is designed to handle dynamic routing with a routing database built from information exchanged by routing protocols. GateD supports the use of the following dynamic routing protocols: Border Gateway Protocol (BGP), Routing Information Protocol (RIP), Open Shortest Path First (OSPF) protocol, and Intermediate System-to-Intermediate System (IS-IS) protocol. More detailed descriptions of these protocols are in the following subsections.

GateD is a modular software program consisting of core services, a routing database, and protocol modules supporting multiple routing protocols (that are defined in the previous list). GateD allows the network administrator to configure routing policy on the GRF through import/export statements that control learning and advertising (or redistributing) of routing information by individual protocol, by source and destination autonomous system (AS), source and destination interface, previous hop router, and specific destination address.

Ascend's High Performance Networking Division developed the following additions to portions of GateD and particularly to BGP:

- Configurable export-best-BGP (CEEB); send best BGP route when route from another protocol is active.
- Asynchronous multilevel nexthop resolution; nexthopself within IBGP.
- AS path truncate variable has been added.
- BGP enhancements: policy per peer, peer group with separate policy per peer; send original BGP nexthop.
- New AS path length algorithm.
- Increased number of GateD adjacencies / peering sessions.
- The IS-IS protocol is supported.
- GateD monitoring tool (called the GateD State Monitor (GSM)).
- BGP confederations.
- BGP multi-exit discriminator (MED).
- Communities.
- Route reflection.

- Route filtering.
- Routing arbiter interaction.
- Route preference biasing.
- Stability support for BGP-OSPF interaction.
- A weighted route dampening statement.

Syntax

The GateD configuration file consists of a sequence of statements terminated by a semi-colon (**;**).

Statements are composed of tokens separated by white space, that can be any combination of blanks, tabs, and new lines. This structure simplifies identification of the parts of the configuration associated with each other and with specific protocols.

Comments can be specified in either of the two following forms:

- One form begins with a pound sign (**#**) and runs to the end of the line.
- The other form, C style, starts with a **/*** and continues until it reaches ***/**.

Syntax description conventions

Keywords and special characters that the parser expects exactly are shown in **bold** font. Variable parameters are shown in *italic* font. Square brackets ([and]) are used to show optional keywords and parameters. The vertical bar (|) separates optional parameters. Parentheses ( ) group keywords and parameters, when necessary.

For example, in the following syntax description, the square brackets say that the parameters are optional. The keywords are **backbone** and **area**. The vertical bar indicates that either **backbone** or **area** *area* can be specified. Because *area* is in *italic* font, it is a parameter that needs to be provided:

[**backbone** | (**area** *area*)]

Statement grouping

The configuration statements (and the order in which these statements appear) in the GateD configuration file (**/etc/gated.conf)** are as follows: options statements, GSM statements, interface statements, definition statements, protocol statements, static statements, control statements, and aggregate statements. Entering a statement out of order causes an error when parsing the configuration file.

Two types of statements do not fit in these categories: **%directive** statements and **%trace** statements. These statements provide instructions to the parser and control tracing from **/etc/gated.conf.** They do not relate to the configuration of any protocol and can occur anywhere in the **/etc/gated.conf** file.

Statement summary

A summary table of the configuration statements follows this section. The table lists each GateD configuration statement by name, identifies the statement type, and provides a short synopsis of the command's function. More detailed definitions and descriptions of each of the GateD statements are provided in the following subsections.

Statement summary

Table 1 lists each statement name and type, and gives a synopsis of the statement's function.

Table A-1. Summary of GateD configuration statements

Statement	Type	Function
%directory	(directive)	Sets the directory for include files.
%include	(directive)	Includes a file into **/etc/gated.conf**.
traceoptions	(trace)	Specifies which events are traced.
options	(definition)	Defines GateD options.
gsm	(definition)	Defines GSM options.
interfaces	(definition)	Defines GateD interfaces.
autonomoussystem	(definition)	Defines the AS number.
routerid	(definition)	Defines the originating router (BGP, OSPF).
martians	(definition)	Defines invalid destination addresses.
rip	(protocol)	Enables RIP protocol.
isis	(protocol)	Enables ISIS protocol.
kernel	(protocol)	Configures kernel interface options.
ospf	(protocol)	Enables OSPF protocol.
bgp	(protocol)	Enables BGP protocol.
routerdiscovery	(protocol)	Enables Router Discovery protocol.
redirect	(protocol)	Configures the processing of ICMP redirects.
weighted route dampening	(protocol)	Minimizes effects of route flapping
icmp	(protocol)	Configures the processing of general ICMP packets.
snmp	(protocol)	Enables reporting to SNMP.
static	(static)	Defines static routes.
import	(control)	Defines which routes to import.
export	(control)	Defines which routes to export.
aggregate	(control)	Defines which routes to aggregate.
generate	(control)	Defines which routes to generate.

Protocol-precedence and preference

Protocol-precedence is the value GateD uses to order precedence of routes from one protocol over another. Preference is the value GateD uses to order the preference of routes within one protocol. Precedence can be set based on one network interface over another, from one protocol over another, or from one remote gateway over another. Precedence and preference can be set in the GateD configuration file in several different configuration statements.

Preference can not be used to control the selection of routes within an interior gateway protocol (IGP); this is accomplished automatically by the protocol based on metric. Preference can be used to select routes from the same exterior gateway protocol (EGP) learned from different peers or autonomous systems. Each route has only one preference value associated with it, even though preference can be set at many places in the configuration file. That is, the last or most specific preference value set for a route is the value used. (See the Glossary: Preference or Glossary: Precedence entries.) The preference value is an arbitrarily assigned value used to determine the order of routes to the same destination in a single routing database. The active route is chosen by the lowest preference value. The range of preference values is 0 through 65536.

Note: It is recommended that changes to preference be used before changes to protocol-precedence. In other words, if you can resolve the issues using the preference within the routing protocol, do so; this method is preferred over changing the order of protocols via protocol-precedence. Also, note that precedence can be set in all places that preference can be set.

Selecting a route

BGP selects the best path to an AS from all the known paths and propagates the selected path to its neighbors. GateD uses the following criteria, in the order given, to select the best path. If routes are equal at a given point in the selection process, then the next criterion is applied to break the tie:

1 Configured policy: the route with the best (numerically smallest) preference, as determined by the policy defined in **/etc/gated.conf**.

2 Local_Pref: the route with the highest local preference. Local_Pref is used by a BGP speaker to inform other BGP speakers in its own autonomous system of the originating speaker's degree of preference for an advertised route. Local_Pref can be set using the **localpref** option on the **import** or **export** statements.

3 Shortest AS path: the route whose NLRI specifies the smallest set of destinations.

4 Origin IGP < EGP < incomplete: the route with an AS path origin of IGP is preferred. Next in preference is the route with AS path origin of EGP. Least preferred is an AS path that is incomplete.

5 MED (if not ignored): the route with the best multi-exit discriminator (MED) is preferred.

6 Shortest IGP distance: if both routes are from the same protocol and AS, the route with the lower metric is preferred.

7 Source IGP < EBGP < IBGP: prefer first the strictly interior route, then the strictly exterior route, then the exterior route learned from an interior session.

8 Lowest router ID: the route whose nexthop is the lowest numeric IP address.

Assigning protocol-precedences

A default precedence is assigned to each source from which GateD receives routes. Precedence values range from 0 to 255 with the lowest number indicating the most preferred route.

Table 2 summarizes the default precedence values for routes learned in various ways. The table lists the statements (some of these are clauses within statements) that set precedence, and shows the types of routes to which each statement applies. The default precedence for each type of route is listed, and the table notes precedence between protocols.

Table A-2. Precedence values

Precedence of:	Defined by statement:	Default value:
direct connected networks	**interfaces**	`0`
OSPF routes		`10`
IS-IS level 1 routes		`15`
IS-IS level 2 routes		`18`
internally generated default	**gendefault**	`20`
redirects	**redirect**	`30`
routes learned via route socket	**kernel**	`40`
static routes from config	**static**	`60`
RIP routes	**rip**	`100`
point-to-point interface		`110`
routes to interfaces that are down	**interfaces**	`120`
aggregate/generate routes	**aggregate/generate**	`130`
OSPF AS external routes	**ospf**	`150`
BGP routes	**bgp**	`170`

Sample precedence specifications

In the following example, the precedence applied to routes learned via RIP from gateway 138.66.12.1 is 75. RIP routes learned from gateways other than 138.66.12.1 receive a precedence of 90. The precedence set on the **import** statement overrides the precedence set in the **rip** statement. The precedence set on the **interface** statement applies only to the routes that are directly connected to that interface.

```
interfaces {
      interface 138.66.12.2 precedence 10 ;
} ;
rip yes {
       precedence 90 ;
} ;
import proto rip gateway 138.66.12.1 precedence 75 ;
```

Trace statements and global options

Trace statements control tracing options. GateD has global and per-protocol tracing options. Each of the sublayers of options inherit the trace option from the previous layer. For example, you can have trace options on a BGP peer statement, which would inherit any of the following trace options, if specified: BGP group, BGP statement, and the global statement, in that order.

Note: Not all of the trace options described in the following subsections apply to all of the protocols. In some cases, their use does not make sense (for example, RIP does not have a state machine), and in some instances the requested tracing has not been implemented (such as RIP support of the **policy** option).

When protocols inherit their tracing options from the global tracing options, tracing levels that do not make sense (such as **parse**, **adv** and packet tracing options) are masked out.

Global tracing statements have an immediate effect, especially parsing options that affect the parsing of the configuration file. Tracing values inherited by protocols specified in the configuration file are initially inherited from the global options that are currently in effect as the protocol configuration entries are parsed, unless they are overridden by more specific options. After the configuration file is read, tracing options that are not explicitly specified are inherited from the global options in effect at the end of the configuration file.

Note: In a production environment, you do not want to run with tracing on. Tracing should be used for debugging purposes only. Please contact Ascend Technical Support for more information.

Traceoptions syntax

traceoptions [*trace_file* [**replace**] [**size** *size* [**k**|**m**] **files** *files*]]
[*control_options*] *trace_options* [**except** *trace_options*] ;
traceoptions none ;

trace_file
Specifies the file to receive tracing information. If this file name does not begin with a slash (/), the directory where GateD was started is prepended to the name.

replace
Tracing should start by replacing an existing file. The default is to append to an existing file.

size *size* [**k**|**m**] **files** *files*
Limits the maximum size of the trace file to the specified size (minimum 10k). When the trace file reaches the specified size, it is renamed to file.0, then file.1, file.2, up to the maximum number of files (minimum specification is 2).

control_options
Specifies options that control the appearance of tracing. Valid values are as follows:

nostamp
Specifies that a timestamp should not be prepended to all trace lines.

trace_options
See the following subsections for more information on the global tracing options.

except *trace_options*
Used to enable a broad class of tracing and then disable more specific options.

none
Specifies that all tracing should be turned off for this protocol or peer.

Global tracing options

There are two types of global options: those that only affect global operations and those that have potential significance to protocols. The following subsections describe the two types of options.

Global significance only trace options

Note: It is recommended that you do not use the following global trace flags; their primary use is for internal development debugging.

The trace flags that only have global significance are as follows:

parse
Trace the lexical analyzer and parser.

adv
Trace the allocation of and freeing of policy blocks.

symbols
Used to trace symbols read from the kernel at startup.

iflist
Used to trace the reading of the kernel interface list.

Protocol significance only trace options

Note: You should only use the **normal** option flag; use of the remaining flags is not recommended.

The options flags that have potential significance to protocols are as follows:

all
Turn on all of the following.

general
A shorthand notation used for specifying both **normal** and **route**.

state
Trace state machine transitions in the protocols.

normal

Trace normal protocol occurrences; abnormal protocol occurrences are always traced.

policy

Trace application of protocol- and user-specified policy to routes being imported and exported.

task

Trace system interface and processing associated with this protocol or peer.

timer

Trace timer usage by this protocol or peer.

route

Trace routing table changes for routes installed by this protocol or peer.

Packet tracing

Tracing of packets is very flexible. For any given protocol there are one or more options for tracing packets. All protocols allow use of the **packets** keyword for tracing all packets sent and received by the protocol. Most protocols have other options for limiting tracing to a useful subset of packet types. These tracing options can be further controlled with the following modifiers:

detail

detail must be specified before **send** or **recv**. Normally, packets are traced in a terse form of one or two lines. When **detail** is specified, a more verbose format is used to provide additional detail on the contents of the packet.

send or **recv**

These options limit the tracing to packets sent or received. Without these options, both sent and received packets are traced.

Note: **detail**, if specified, must be before **send** or **recv**. If a protocol allows for several different types of packet tracing, modifiers can be applied to each individual type. But be aware that within one tracing specification the trace flags are summed up, so specifying **detail** packets turns on full tracing for all packets.

Directive statement

Directive statements provide direction to the GateD configuration language parser about included files and the directories in which these files reside.

Directive statements are immediately acted upon by the parser. Other statements terminate with a semi-colon (;), but directive statements terminate with a newline. The two directive statements are as follows:

%directory *directory*

Defines the directory where the include files are stored. When it is used, GateD looks in the directory identified by pathname for any included files that do not have a fully qualified filename (that is, do not begin with /). This statement does not actually change the current directory, it just specifies the prefix applied to included file names.

%include *filename*

Identifies an include file. The contents of the file is *included* in the **/etc/gated.conf** file at the point in the **/etc/gated.conf** file where the **%include** directive is encountered. If the filename is not fully qualified (that is, does not begin with "/"), it is considered to be relative to the directory defined in the **%directory** directive. The **%include** directive statement causes the specified file to be parsed completely before resuming with this file. Nesting up to ten levels is supported.

In a complex environment, segmenting a large configuration into smaller, more easily understood segments can be helpful, but one of the great advantages of GateD is that it combines the configuration of several different routing protocols into a single file. Segmenting a small file unnecessarily complicates routing configurations.

Note: These statements must begin in the first character column of the **/etc/gated.conf** file. In other words, there can be no tabs or spaces before the **%**.

Options statement

The options statement allows specification of some global options. If used, **options** must appear before any other type of configuration statement in the **/etc/gated.conf** file.

The options statement syntax is as follows:

```
options
    [ nosend ]
    [ noresolv ]
    [ gendefault [ precedence precedence ] [ gateway gateway ] ]
    [ syslog [ upto ] log_level ]
    [ mark time ]
    ;
```

The **options** list can contain one or more of the following options:

gendefault [**precedence** *precedence*] [**gateway** *gateway*]

When **gendefault** is enabled, and when a BGP neighbor is up, it causes the creation of a default route with the special protocol **default**. This can be disabled per BGP group with the **nogendefault** option. By default, this route has a precedence of 20. This route is normally not installed in the kernel forwarding table; it is only present so it can be announced to other protocols. If a gateway is specified, the default route is installed in the kernel forwarding table with a nexthop of the listed gateway.

Note: the use of the more general option is preferred to the use of the **gendefault** option. The **gendefault** option may go away in future releases. See the section on Route Aggregation for more information on the Generate statement.

nosend

Specifies that no packets are sent. This option makes it possible to run GateD on a live network to test protocol interactions without actually participating in the routing protocols. The packet traces in the GateD log can be examined to verify that GateD is functioning properly. This is most useful for RIP, and possibly the SMUX SNMP interface. This option does not yet apply to BGP and is less than useful with OSPF.

noresolv

By default, GateD tries to resolve symbolic names into IP addresses by using the **gethostbyname**() and **getnetbyname**() library calls. These calls usually use the Domain Name System (DNS) instead of the hosts local host and network tables. If there is insufficient routing information to send DNS queries, GateD deadlocks during start-up. This option can be used to prevent these calls; symbolic names results in configuration file errors.

syslog [**upto**] *log_level*

Controls the amount of data GateD logs via **syslog** on systems where **setlogmask**() is supported. The available logging levels and other terminology are as defined in the **setlogmask**(3) man page. The default is equivalent to **syslog upto** information.

mark *time*

Specifying this option causes GateD to output a message to the trace log at the specified interval. This can be used as a method of determining if GateD is still running.

GSM statement

GateD provides an interactive interface (called the GateD State Monitor (GSM)) from which an operator can interrogate the state of route tables, interfaces, and routing protocols. In the enabled mode, the operator can also control some of the GateD operations: enabling or disabling configured traces, resetting a peering session and adjacencies, and manipulating interface tables. For more information, see **gsm** command in Chapter 1.

From the GateD configuration file (**/etc/gated.conf**), you can control access to the GSM in the following ways:

- The port on which GateD responds to the GSM connections
- The hosts from which GateD can accept the GSM connections
- The users whose passwords provide admission to the GSM

Additionally, the GSM can be turned on or off via the **/etc/gated.conf** file.

Note: For backward compatibility with previous releases, GSM is on by default. In future post-1.4.8 releases, the GSM will be off by default, thus you will have to use the **gsm** statement to enable (that is, turn on) or disable (that is, turn off) the GSM.

The **gsm** statement syntax is:

```
gsm ( yes | no | on | off )
[ {
     [ port port-number ; ]
     [ usernames user-list ; ]
     [ hosts host-set ; ]
} ] ;
```

The **gsm** list can contain one or more of the following options:

port *port-number*

Controls the TCP port to which GSM binds. By default, GSM binds to the port named **gii** in the **/etc/services** file. If no such port is defined, then the final default is port number **616**. In other words, if this option specifies a GSM port, then that port is the one that is used by GSM. If no port is specified with the **port** option, and if **/etc/services** specifies a port number for a service named **gii**, then GateD uses that port for GSM. If neither of the two previous conditions are true, then GateD defaults to using TCP port **616** for GSM.

usernames *user-list*

Lists the users permitted to connect to the GSM; the list is in the format of double-quoted strings (for example, **"netstar"**). Multiple users are separated by white space (for example, **"netstar" "jsmith" "manderson"**). When you open a **telnet** session to the GSM, the GSM prompts for a user name and a password. You can continue with the GSM session if and only if you provide a user name from the *user-list* and the corresponding system password. The default is to permit only the user **netstar**. For backward compatibility with previous releases, the default case is to prompt for password only. In future post-1.4.8 releases, the default will be for GSM to unconditionally prompt for the user name.

hosts *host-set*

Lists the hosts from which GateD can permit **telnet** connections to the GSM. The syntax for the *host-set* is the same as for the prefix matching in the BGP **allow** keyword, which is as follows:

network

network **mask** *mask*

network (**masklen** | **/**) *number*

all

host *host*

Changes to the **port**, **usernames**, and/or **hosts** options requires that you execute a **gdc reconfig** for the changes to take effect. Also, changes to these options do not affect already-established GSM session (regardless of whether you execute the **gdc reconfig** command).

For backward compatibility with previous releases, the default case is to accept connections from any source addresses. In future post-1.4.8 releases, the default will be to permit connections only from the loopback (that is, to accept connections only from the router running this instance of GateD and not from the network).

For example, the following is the current default GSM configuration made explicit in the **/etc/gated.conf** file:

```
gsm on {
    port        616 ;
    usernames   “netstar” ;
    hosts       all ;
} ;
```

Interfaces statement

Syntax

```
interfaces {
    options
        [ strictinterfaces ]
        [ scaninterval time ]
        ;
    interface interface_list
        [ precedence precedence ]
        [ down precedence precedence ]
        [ passive ]
        [ simplex ]
        [ reject ]
        [ blackhole ]
        ;
    define address
        [ broadcast address ] | [ pointtopoint address ]
        [ netmask mask ]
        [ multicast ]
        ;
} ;
```

An interface is the connection between a router and one of its attached networks. A physical interface can be specified by interface name, by IP address, or by domain name (unless the network is an un-numbered point-to-point network).

Multiple levels of reference in the configuration language allow identification of interfaces using wildcard, interface type name, or delete word address. Be careful with the use of interface names as future UNIX operating systems can allow more than one address per interface. The *interface_list* is a list of one or more interface names including wildcard names (names without a number) and names which can specify more than one interface or address, or the token **all** for all interfaces.

options

Allows configuration of some global options related to interfaces. They are as follows:

strictinterfaces

Indicates that it is a fatal error to reference an interface in the configuration file that is not present when GateD is started and not listed in a **define** statement. Without this option, a warning message is issued but GateD continues.

scaninterval *time*

Specifies how often GateD scans the kernel interface list for changes. The default is every 15 seconds on most systems, and 60 seconds on systems that pass interface status changes through the routing socket (for example, BSD 4.4). Note that GateD also scans the interface list on receipt of a **SIGUSR2**.

interface *interface_list*

Sets interface options on the specified interfaces. An interface list is **all** or a list of interface names (see interface warning described above in **options**), domain names, or numeric addresses. Options available on this statement are as follows:

precedence *precedence*

Sets the precedence for directly-connected routes to this interface when it is up and appears to be functioning properly. The default precedence is 0.

down precedence *precedence*

Sets the precedence for directly-connected routes to this interface when GateD does not believe it to be functioning properly, but the kernel does not indicate it is down. The default value is 120.

passive

Prevents GateD from changing the precedence of the route to this interface if it is not believed to be functioning properly due to lack of received routing information. GateD only performs this check if the interface is actively participating in a routing protocol.

simplex

Defines an interface as unable to hear its own broadcast packets. Some systems define an interface as simplex with the **IFF_SIMPLEX** flag. On others it needs to be specified in the configuration file. On simplex interfaces, packets from myself are assumed to have been looped back in software, and are not used as an indication that the interface is functioning properly.

reject

Specifies that the address of the interface that matches these criteria are used as the local address when installing **reject** routes in the kernel. Should only be used with systems based on BSD 4.3 Tahoe or earlier that have installed a **reject**/**blackhole** pseudo interface.

blackhole

Specifies that the address of the interface that matches these criteria are used as the local address when installing **reject** routes in the kernel. Should only be used with systems based on BSD 4.3 Tahoe or earlier that have installed a **reject**/**blackhole** pseudo interface.

define *address*

Defines interfaces that might not be present when GateD is started so they can be referenced in the configuration file when **strictinterfaces** is defined.

Possible **define** keywords are as follows:

broadcast *address*

Defines the interface as broadcast capable (for example, Ethernet or Token Ring) and specifies the broadcast address.

pointtopoint *address*

Defines the interface as a point-to-point interface (for example, SLIP or PPP) and specifies the address on the local side. The first *address* on the **define** statement references the address of the host on the **remote** end of the interface, the *address* specified after this **pointtopoint** keyword defines the address on the **local** side of the interface.

An interface not defined as **broadcast** or **pointtopoint** is assumed to be non-broadcast multi-access (NBMA), such as an X.25 network.

netmask mask

Specifies the subnet mask to be used on this interface. This is ignored on **pointtopoint** interfaces.

multicast

Specifies that the interface is multicast capable.

Interface lists

An interface list is a list of references to interfaces or groups of interfaces. There are four methods available for referring to interfaces. They are listed here from most general to most specific.

all

This refers to all available interfaces.

interface name wildcard

This refers to all the interfaces of the same type. UNIX interfaces consist of the name of the device driver (for example, **ie**), and a unit number (for example, like 0, 5, 22). References to the name contain only alphabetic characters and match any interfaces that have the same alphabetic part.

For example, **ie** on a Sun would refer to all Interlan Ethernet interfaces, **le** would refer to all Lance Ethernet interfaces. But **ie** would not match **ie10**.

Interface name

This refers to a specific interface, usually one physical interface. These are specified as an alphabetic part followed by a numeric part. This matches one specific interface. But be aware that on many systems, there can be more than one protocol (that is, *IP*) address on a given physical interface. For example, **ef1** matches an interface named **ef1**, but not an interface named **ef10**.

Interface address

This matches one specific interface. The reference can be by protocol address (for example, 10.0.0.51), or by symbolic hostname (for example, **nic.dnn.mil**). Note that a symbolic hostname reference is only valid when it resolves to only one address. Use of symbolic hostnames is not recommended.

If many interface lists are present in the configuration file with more than one parameter, these parameters are collected at run-time to create the specific parameter list for a given interface. If the same parameter is specified on more than one list, the parameter with the most specific interface is used.

For example, consider the following system with three interfaces, **le0**, **le1**, and **du0**:

```
rip yes {
    interface all noripin noripout ;
    interface le0 ripin ;
    interface le1 ripout ;
} ;
```

RIP packets would be accepted from interfaces **le0** and **le1**, but not from **du0**. RIP packets would only be sent on interface **le1**.

IP interface addresses and routes

The *BSD 4.3* and later networking implementations allow four types of interfaces.

loopback

This interface must have the address of 127.0.0.1. Packets sent to this interface are sent back to the originator. This interface is also used as a catch-all interface for implementing other features, such as **reject** and **blackhole** routes. Although a netmask is reported on this interface, it is ignored. It is useful to assign an additional address to this interface that is the same as the OSPF or BGP router id. This allows routing to a system based on the router id that works if some interfaces are down.

broadcast

This is a multi-access interface capable of a physical level broadcast such as Ethernet, Token Ring, or FDDI. This interface has an associated subnet mask and broadcast address. The interface route to a **broadcast** network is a route to the complete subnet.

point-to-point

This is a tunnel to another host, usually on some sort of serial link. This interface has a local address, and a remote address. Although it can be possible to specify multiple addresses for a point-to-point interface, there does not seem to be a useful reason for doing so.

The remote address must be unique among all the interface addresses on a given router. The local address can be shared among many point-to-point and up to one non-point-to-point interface. This is technically a form of the router id method for address-less links. This technique conserves subnets as none are required when using this technique.

If a subnet mask is specified on a point-to-point interface, it is only used by RIP version 1 to determine which subnets can be propagated to the router on the other side of this interface.

non-broadcast multi-access or **nbma**

This type of interface is multi-access, but not capable of broadcast. Examples would be frame relay and X.25. This type of interface has a local address and a subnet mask.

GateD ensures that there is a route available to each IP interface that is configured and up. Normally this is done by the **ifconfig** command that configures the interface; GateD does it to ensure consistency.

For point-to-point interfaces, GateD installs some special routes. If the local address on one or more point-to-point interfaces is not shared with a non-point-to-point interface, GateD installs a route to the local address pointing at the loopback interface with a precedence of 110. This ensures that packets originating on this host destined for this local address are handled locally. OSPF prefers to route packets for the local interface across the point-to-point link where they are returned by the router on the remote end. This is used to verify operation of the link. Since OSPF installs routes with a precedence of 10, these routes override the route installed with a precedence of 110.

If the local address of one or more point-to-point interfaces is shared with a non-point-to-point interface, GateD installs a route to the local interface with a precedence of 0 that is not installed in the forwarding table. This is to prevent protocols like OSPF from routing packets to this address across a serial interface when this system could be functioning as a host.

When the status of an interface changes, GateD notifies all the protocols, that then take the appropriate action. GateD assumes that interfaces not marked **UP** do not exist. While this might not be the most correct action, it is the way things currently work.

GateD ignores any interfaces that have invalid data for the local, remote or broadcast addresses, or the subnet mask. Invalid data includes zeros in any field. GateD also ignores any point-to-point interface that has the same local and remote addresses, it assumes it is in some sort of loopback test mode.

Definition statements

Definition statements are general configuration statements that relate to all of GateD or at least to more than one protocol. The five definition statements are **autonomoussystem**, **confederation**, **routerid**, **routing-domain**, and **martians**. When used, these statements must appear before any other type of configuration statement in the **/etc/gated.conf** file.

Also, one must use **confederation** and **routing-domain** together, but one cannot use **confederation**, **routing-domain**, and **autonomoussystem** together. They are mutually exclusive for configuring BGP.

Autonomous system configuration

autonomoussystem *autonomous_system* [**loops** *number*] **;**

Sets the autonomous system number of this router to be *autonomous_system*. This option is required if BGP are in use. The AS number is assigned by the Network Information Center (NIC).

loops is only for protocols supporting AS paths, such as BGP. It controls the number of times this autonomous system can appear in an AS path and defaults to 1 (one).

Equal Cost Multipath Configuration (ECMP)

multipath (**yes** | **no** | **off** | **no**) **;**

The mulitpath parameter enables/disables the installation of multiple gateways for network or host prefixes in the kernel route table. The default is **off**.

The **on** option is the same as the **yes** option, and enables GateD to install multiple routes for a single source with the same destination but different nexthops. The **off** option is the same as the **no** option, and means that each route GateD installs in the kernel has a unique destination and nexthop.

Thus, GateD provides the following services for ECMP:

- The configuration option to enable/disable dynamic ECMP (default is off)
- The ability to insert multiple equal cost routes into the kernel table
- Internal OSPF protocol support

This release supports ECMP prefixes learned via the OSPF and OSPF_ASE protocols, BGP is supported, if the nexthop resolving protocol is OSPF or OSPF_ASE.

Note: This parameter cannot be changed via the **gdc reconfig** command. Changes to this parameter take effect only when Gated is started (**gdc start**) or restarted (**gdc restart**).

For more information on the CRC hash algorithm used by ECMP for selecting destinations per packet, see the Introduction to GRF ECMP section.

Router ID configuration

routerid *host* **;**

Sets the router identifier for use by the BGP and OSPF protocols. The default is the address of the first interface encountered by GateD. The address of a non-point-to-point interface is preferred over the local address of a point-to-point interface. An address on a loopback interface that is not the loopback address (127.0.0.1) is most preferred.

Confederation

confederation *confederation* **;**

Sets the confederation identifier for use by the BGP protocol for the configuration of group type confed. See the BGP statement for more information on configuring confederations. Note that **autonomoussystem** and **confederation** are mutually exclusive; they cannot both be defined within the same `gated.conf`. Also, note that **confederation** and **routing-domain** both must be set within the same **/etc/gated.conf**.

Routing domain identifier

routing-domain *rdi* **;**

Sets the routing domain identifier for use by the BGP protocol for the configuration of group type confed. See the BGP statement for more information on configuring confederations. Note that **confederation** and **routing-domain** both must be set within the same **/etc/gated.conf**.

Martian configuration

```
martians {
        host host [ allow ] ;
        network [ allow ] ;
        network mask mask [ allow ] ;
        network ( masklen | / ) number [ allow ] ;
        default [ allow ] ;
    };
```

Defines a list of **martian** addresses about which all routing information is ignored.

Sometimes a misconfigured system sends out obviously invalid destination addresses. These invalid addresses, called martians, are rejected by the routing software. This command allows additions to the list of martian addresses. See the section on Route Filtering for more information on specifying ranges.

Also, the **allow** parameter can be specified to explicitly allow a subset of a range that was disallowed.

Sample definition statements

The statements in the following example perform the following functions:

- The **options** statement tells the system to generate a default route when it peers with an BGP neighbor.
- The **interface** statement tells GateD not to mark interface 128.66.12.2 as down even if it sees no traffic.
- The **autonomoussytem** statement tells GateD to use AS number 249 when participating in BGP.
- The **martians** statement prevents the host route 10.10.10.26 and the network 10.8.0.0/16 from being installed.

```
options gendefault ;
autonomoussytem 249 ;
interface 128.66.12.2 passive ;
martians {
    host 10.10.10.26 ;
    network 10.8.0 mask 255.255.0.0 ;
};
```

Protocol overview

Routing protocols determine the best route to each destination and distribute routing information among the systems on a network. Routing protocols are divided into two general groups: interior gateway protocols (IGPs) and exterior gateway protocols (EGPs). GateD comgnes the management of all protocols.

Interior Gateway Routing Protocols

Interior gateway protocols (IGPs) are used to exchange reachability information within an autonomous system (AS). There are three interior protocols currently supported by this version of GateD: RIP, OSPF, and IS-IS. The following subsections briefly describe these IGPs.

Routing Information Protocol (RIP)

The Routing Information Protocol, Version 1 and Version 2, is the most commonly used interior protocol. RIP selects the route with the lowest metric as the best route. The metric is a hop count representing the number of gateways through which data must pass to reach its destination. The longest path that RIP accepts is 15 hops. If the metric is greater than 15, a destination is considered unreachable and GateD discards the route. RIP assumes the best route is the one that uses the fewest gateways (that is, the shortest path), not taking into account congestion or delay on route.

The RIP version 1 protocol is described in RFC 1058 and the RIP version 2 protocol is described in RFC 1388.

Open Shortest Path First (OSPF)

Open Shortest Path First is a link-state protocol. OSPF is better suited than RIP for complex networks with many routers in a single AS. OSPF chooses the least cost path as the best path. It provides equal cost multipath routing where packets to a single destination can be sent via more than one interface simultaneously.

OSPF is described in RFC 1583, the MIB is defined in RFC 1253. Other related documents are RFC 1245, RFC 1246, and RFC 1370.

Intermediate System-to-Intermediate System (IS-IS) Protocol

IS-IS is a link state interior routing protocol originally developed for routing ISO packets. The IS-IS version distributed with GateD can route IP packets as well.

In ISO terminology, a router is referred to as an intermediate system. In IS-IS, the network is partitioned into routing domains. IS-IS intradomain routing is configured and organized hierarchically so that a large domain can be administratively divided into smaller areas. It uses level 1 intermediate systems within areas and level 2 intermediate systems between areas. The Route Manager determines the boundaries of routing domains by setting some links to be exterior links. If a link is marked as exterior, no IS-IS routing messages are sent on that link.

Exterior Routing Protocols

Exterior protocols are used to exchange routing information between autonomous systems. Exterior protocols are only required when an autonomous system must exchange routing information with another autonomous system. Routers within an autonomous system run an interior routing protocol like RIP. Only those gateways that connect an autonomous system to another autonomous system need to run an exterior routing protocol. The exterior protocol currently supported by GateD is BGP.

Border Gateway Protocol (BGP)

Border Gateway Protocol (BGP) exchanges reachability information between autonomous systems. BGP uses path attributes to provide more information about each route as an aid in selecting the best route. Path attributes can include, for example, administrative preferences based on political, organizational, or security (policy) considerations in the routing decision. BGP supports non-hierarchical topologies and can be used to implement a network structure of equivalent autonomous systems.

BGP version 1 is described in RFC 1105, version 2 in RFC 1163, version 3 in RFC 1267, and version 4 in RFC 1771. The version 3 MIB is described in RFC 1269. The three documents, RFC 1164, RFC 1268, and RFC 1772, describe the application of versions 2, 3, and 4 in the Internet.

A protocol analysis of and experience with BGP version 3 are available in RFC 1265 and RFC 1266. RFC 1397 talks about advertising a default route in BGP version 2 and 3. And finally, RFC 1403 describes BGP - OSPF interaction.

Link State and Distance Vector Protocols

Another way to categorize routing protocols is by the way route table information is managed and the kind of information that is exchanged. The main types are Link State and Distance Vector.

Link state protocols exchange information about the state of the links in the network. This information is propagated to all routers in the network so that each knows the topology of the entire network. From this information, they can compute the shortest path across the network for each destination and route packets along those paths. By maintaining more complete information, link state protocols can make better routing decisions but must maintain information that may be difficult to manage in large, dynamic networks. The OSPF and IS-IS routing protocols are examples of link state protocols.

In distance vector protocols, each router advertises the shortest path it knows to each destination. As routes are advertised and learned, the router selects the routes with the shortest distance to each destination. The router is not burdened with tracking the state of all of the links in the network, but is making routing decisions with less complete information.

The RIP and BGP routing protocols are examples of distance vector protocols. BGP is a more sophisticated “path vector protocol” that exchanges path information and not just distances. This lets it address some of the limitations of distance vector protocols.

Each protocol has advantages and disadvantages. The choice of the appropriate protocol depends on many factors such as the size, complexity, and stability of the network.

Other Routing Protocols

Router Discovery

The Router Discovery protocol is used to inform hosts of the availability of hosts it can send packets to and is used to supplement a statically-configured default router. This is the preferred protocol for hosts to run, they are discouraged from *wiretapping* routing protocols. Router Discovery is described in RFC 1256. For more information on the Router Discovery protocol, see the Router Discovery Protocol section later in this manual.

Routing information protocol (RIP)

One of the most widely used interior gateway protocols is the Routing Information Protocol (RIP). RIP is an implementation of a distance-vector, or Bellman-Ford routing protocol for local networks. It classifies routers as active and passive (silent). Active routers advertise their routes (reachability information) to others; passive routers listen and update their routes based on advertisements, but do not advertise. Typically, routers run RIP in active mode, while hosts use passive mode.

A router running RIP in active mode broadcasts updates at set intervals. Each update contains paired values where each pair consists of an IP network address and an integer distance to that network. RIP uses a hop count metric to measure the distance to a destination. In the RIP metric, a router advertises directly connected networks at a metric of 1. Networks that are reachable through one other gateway are two hops, etc. Thus, the number of hops or hop count along a path from a given source to a given destination refers to the number of gateways that a datagram would encounter along that path. Using hop counts to calculate shortest paths does not always produce optimal results. For example, a path with hop count 3 that crosses three Ethernets can be substantially faster than a path with a hop count 2 that crosses two slow-speed serial lines. To compensate for differences in technology, many routers advertise artificially high hop counts for slow links.

As delivered with most UNIX systems, RIP is run by the routing daemon, **routed** (pronounced route-"d"). A RIP routing daemon dynamically builds on information received through RIP updates. When started up, it issues a REQUEST for routing information and then listens for responses to the request. If a system configured to supply RIP hears the request, it responds with a RESPONSE packet based on information in its routing database. The RESPONSE packet contains destination network addresses and the routing metric for each destination.

When a RIP RESPONSE packet is received, the routing daemon takes the information and rebuilds the routing database adding new routes and "better" (lower metric) routes to destinations already listed in the database. RIP also deletes routes from the database if the next router to that destination says the route contains more than 15 hops, or if the route is deleted. All routes through a gateway are deleted if no updates are received from that gateway for a specified time period. In general, routing updates are issued every 30 seconds. In many implementations, if a gateway is not heard from for 180 seconds, all routes from that gateway are deleted from the routing database. This 180-second interval also applies to deletion of specific routes.

RIP version 2 (more commonly known as RIP II) add additional capabilities to RIP. Some of these capabilities are compatible with RIP I and some are not. To avoid supplying information to RIP I routes that could be misinterpreted, RIP II can only use non-compatible features when its packets are multicast. On interfaces that are not capable of IP multicast, RIP I-compatible packets are used that do not contain potentially confusing information.

Some of the most notable RIP II enhancements are described in the following subsections.

nexthop

RIP II can advertise a nexthop other than the router supplying the routing update. This is quite useful when advertising a static route to a dumb router that does not run RIP. It avoids having packets destined through the dumb router from having to cross a network twice.

RIP I routers will ignore nexthop information in RIP II packets. This can result in packets crossing a network twice, which is exactly what happens with RIP I. So this information is provided in RIP I-compatible RIP II packets.

Network mask

RIP I assumes that all subnetworks of a given network have the same network mask. It uses this assumption to calculate the network masks for all routes received. This assumption prevents subnets with different netmasks from being included in RIP packets. RIP II adds the ability to specify the network mask with each network in a packet.

While RIP I routers ignores the network mask in RIP II packets, their calculation of the network mask can quite possibly be wrong. For this reason, RIP I-compatible RIP II packets must not contain networks that would be misinterpreted. These networks must only be provided in native RIP II packets that are multicast.

Authentication

RIP II packets can also contain one of two types of authentication strings that can be used to verify the validity of the supplied routing data. Authentication can be used in RIP I-compatible RIP II packets, but be aware that RIP I routers ignore it.

The first method is a simple password in which an authentication key of up to 16 characters is included in the packet. If this does not match what is expected, the packet is discarded. This method provides very little security as it is possible to learn the authentication key by watching RIP packets.

The second method is still experimental and can change in incompatible ways in future releases. This method uses the MD5 algorithm to create a crypto-checksum of a RIP packet and an authentication key of up to 16 characters. The transmitted packet does not contain the authentication key itself, instead it contains a crypto-checksum, called the *digest*. The receiving router performs a calculation using the correct authentication key and discard the packet if the digest does not match. In addition, a sequence number is maintained to prevent the replay of older packets. This method provides a much stronger assurance that routing data originated from a router with a valid authentication key.

Two authentication methods can be specified per interface. Packets are always sent using the primary method, but received packets are checked with both the primary and secondary methods before being discarded. In addition, a separate authentication key is used for non-router queries.

RIP I and network masks

RIP I derives the network mask of received networks and hosts from the network mask of the interface via the packet that was received. If a received network or host is on the same natural network as the interface over which it was received and that network is subnetted (the specified mask is more specific than the *natural* netmask), the subnet mask is applied to the destination. If bits outside the mask are set, it is assumed to be a host, otherwise it is assumed to be a subnet.

On point-to-point interfaces, the netmask is applied to the remote address. The netmask on these interfaces is ignored if it matches the natural network of the remote address or is all ones.

Unlike in previous releases, the zero subnet mask (a network that matches the natural network of the interface, but has a more specific, or longer, network mask) is ignored. If this is not desirable, a route filter can be used to reject it.

The RIP statement

```
rip ( yes | no | on | off ) [ {
      broadcast ;
      nobroadcast ;
      nocheckzero ;
      precedence precedence ;
      defaultmetric metric ;
      query authentication [none | ([simple|md5] password)] ;
      interface interface_list
            [noripin] | [ripin]
            [noripout] | [ripout]
            [metricin metric]
            [metricout metric]
            [version 1]|[version 2 [multicast|broadcast]]
            [[secondary] authentication [none| ([simple|md5] password)]] ;
      trustedgateways gateway_list ;
      sourcegateways gateway_list ;
      traceoptions trace_options ;
} ] ;
```

The **rip** statement enables or disables RIP. If the **rip** statement is not specified, the default is **rip off ;**. If enabled, RIP assumes **nobroadcast** when there is only one interface and **broadcast** when there is more than one.

The RIP options are as follows:

broadcast

Specifies that RIP packets are broadcast regardless of the number of interfaces present. This is useful when propagating static routes or routes learned from another protocol into RIP. In some cases, the use of **broadcast** when only one network interface is present can cause data packets to traverse a single network twice.

nobroadcast

Specifies that RIP packets is not broadcast on attached interfaces, even if there are more than one. If a **sourcegateways** clause is present, routes is still unicast directly to that gateway.

nocheckzero

Specifies that RIP should not make sure that reserved fields in incoming version 1 RIP packets are zero. Normally RIP rejects packets where the reserved fields are zero.

precedence *precedence*

Sets the precedence for routes learned from RIP. The default precedence is 100. This precedence can be overridden by a precedence specified in the import policy. This keyword can also be stated as **protocol-precedence** or **proto-precendence**.

defaultmetric *metric*

Defines the metric used when advertising routes via RIP that were learned from other protocols. If not specified, the default value is 16 (unreachable). This choice of values requires you to explicitly specify a metric in order to export routes from other protocols into RIP. This metric can be overridden by a metric specified in export policy.

query authentication [**none** | ([**simple**|**md5**] *password*)]

Specifies the authentication required of query packets that do not originate from routers. The default is **none**.

interface *interface_list*

Controls various attributes of sending RIP on specific interfaces. See the section on interface list specification for the description of the *interface_list* (page A-19).

Note that if there are multiple interfaces configured on the same subnet, RIP updates are only be sent from the secondary interfaces if they are declared on the Interface statement by IP number (*not by logical interface name*). Also, note that IP aliases for the **lo0** (loopback) interface must also be declared by IP number. Under no circumstances is the export or designation of **lo0** or 127.0.0.1 allowed.

The possible parameters are as follows:

noripin

Specifies that RIP packets received via the specified interface is ignored. The default is to listen to RIP packets on all non-loopback interfaces.

ripin

This is the default. This argument can be necessary when **noripin** is used on a wildcard interface descriptor.

noripout

Specifies that no RIP packets are sent on the specified interfaces. The default is to send RIP on all broadcast and non-broadcast interfaces when in **broadcast** mode. The sending of RIP on point-to-point interfaces must be manually configured.

ripout

This is the default. This argument is necessary when it is desired to send RIP on point-to-point interfaces and can be necessary when **noripin** is used on a wildcard interface descriptor.

metricin *metric*

Specifies the RIP metric to add to incoming routes before they are installed in the routing table. The default is the kernel interface metric plus 1 (which is the default RIP hop count). If this value is specified, it is used as the absolute value, the kernel metric is not added. This option is used to make this router prefer RIP routes learned via the specified interface(s) less than RIP routes from other interfaces.

metricout *metric*

Specifies the RIP metric to be added to routes that are sent via the specified interface(s). The default is zero. This option is used to make other routers prefer other sources of RIP routes over this router.

version 1

Specifies that the RIP packets sent on the specified interface(s) is version 1 packets. This is the default.

version 2

Specifies that RIP version 2 packets are sent on the specified interfaces(s). If IP multicast support is available on the specified interface(s), the default is to send full version 2 packets. If multicast support is not available, version 1 compatible version 2 packets are sent.

multicast

Specifies that RIP version 2 packets should be multicast on this interface. This is the default.

broadcast

Specifies that RIP 1-compatible RIP version 2 packets should be broadcast on this interface, even if IP multicast is available.

[**secondary**] **authentication** [**none** | ([**simple**|**md5**] *password*)]

This defines the authentication type to use. It applies only to RIP version 2 and is ignored for RIP 1 packets. The default authentication type is **none**. If a password is specified, the authentication type defaults to **simple**. The password should be a quoted string with between 0 and 16 characters.

If **secondary** is specified, this defines the secondary authentication. If omitted, the primary authentication is specified. The default is primary authentication of **none** and no secondary authentication.

trustedgateways *gateway_list*

Defines the list of gateways from which RIP accepts updates. The *gateway_list* is simply a list of host names or IP addresses. By default, all routers on the shared network are trusted to supply routing information. But, if the **trustedgateways** clause is specified, only updates from the gateways in the list are accepted.

sourcegateways *gateway_list*

Defines a list of routers to which RIP sends packets directly, not through multicast or broadcast. This can be used to send different routing information to specific gateways. Updates to gateways in this list are not affected by **noripout** on the interface.

traceoptions *trace_options*

Specifies the tracing options for RIP. (See Trace Statements (page A-12) and the RIP-specific tracing options in the following subsection for more information.)

Tracing options

The **policy** option logs information whenever a new route is announced, or the metric being announced changes, or a route goes or leaves holddown.

Packet tracing options (that can be modified with **detail**, **send** or **recv**):

packets

All RIP packets.

request

RIP information request packets, such as **REQUEST**, **POLL** and **POLLENTRY**

response

RIP **RESPONSE** packets, that are the type of packet that actually contains routing information.

other

Any other type of packet. The only valid ones are **TRACE_ON** and **TRACE_OFF**, both of which are ignored.

The OSPF protocol

Open Shortest Path Routing (OSPF) is a shortest path first or link-state protocol. OSPF is an interior gateway protocol that distributes routing information between routers in a single autonomous system (AS). OSPF chooses the least cost path as the best path. Suitable for complex networks with a large number of routers, OSPF provides equal cost multipath routing where packets to a single destination can be sent via more than one interface simultaneously. In a link-state protocol, each router maintains a database describing the entire AS topology, that it builds out of the collected link state advertisements of all routers. Each participating router distributes its local state (that is, the router's usable interfaces and reachable neighbors) throughout the AS by flooding. Each multi-access network that has at least two attached routers has a *designated router* and a *backup designated router.* The designated router floods a link state advertisement for the multi-access network and has other special responsibilities. The designated router concept reduces the number of adjacencies required on a multi-access network.

OSPF allows networks to be grouped into areas. Routing information passed between areas is abstracted, potentially allowing a significant reduction in routing traffic. OSPF uses four different types of routes, listed in order of preference:

- intra-area
- inter-area
- type 1 external
- type 2 external

Intra-area paths have destinations within the same area, inter-area paths have destinations in other OSPF areas and Autonomous System External (ASE) routes are routes to destinations external to the AS. Routes imported into OSPF as type 1 routes are supposed to be from IGPs whose external metrics are directly comparable to OSPF metrics. When a routing decision is being made, OSPF adds the internal cost to the AS Border router to the external metric. Type 2 ASEs are used for EGPs whose metrics are not comparable to OSPF metrics. In this case, only the internal OSPF cost to the AS Border router is used in the routing decision.

From the topology database, each router constructs a tree of the shortest paths with itself as the root. This shortest-path tree gives the route to each destination in the AS. Externally derived routing information appears on the tree as leaves. The link-state advertisement format distinguishes between information acquired from external sources and information acquired from internal routers, so there is no ambiguity about the source or reliability of routes. Externally derived routing information (for example, routes learned from BGP) is passed transparently through the autonomous system and is kept separate from OSPF's internally derived data. Each external route can also be tagged by the advertising router, enabling a passing of additional information between routers on the borders of the autonomous system.

OSPF optionally includes *type of service* (TOS) routing and allows administrators to install multiple routes to a given destination for each type of service (for example, low delay or high throughput.) A router running OSPF uses the destination address and the type of service to choose the best route to the destination.

OSPF intra- and inter-area routes are always imported into the GateD routing database with a precedence of 10. It would be a violation of the protocol if an OSPF router did not participate fully in the area's OSPF, so it is not possible to override this.

Although it is possible to give other routes lower precedence values explicitly, it is ill-advised to do so.

Hardware multicast capabilities are also used where possible to deliver link-status messages. OSPF areas are connected by the *backbone* area, the area with identifier 0.0.0.0. All areas must be logically contiguous and the backbone is no exception. To permit maximum flexibility, OSPF allows the configuration of *virtual* links to enable the backbone area to appear contiguous despite the physical reality.

All routers in an area must agree on that area's parameters. A separate copy of the link-state algorithm is run for each area. Because of this, most configuration parameters are defined on a per area basis. All routers belonging to an area must agree on that area's configuration. Mis-configuration leads to adjacencies not forming between neighbors, and routing information might loop or not flow.

Authentication

All OSPF protocol exchanges are authenticated. Authentication guarantees that routing information is only imported from trusted routers, to protect the Internet and its users. A variety of authentication schemes can be used, but a single scheme must be configured for each area. This enables some areas to use much stricter authentication than others. OSPF protocol exchanges can be authenticated. Authentication guarantees that routing information is imported only from trusted routers, to protect the Internet and its users. There are two authentication schemes available. The first uses a simple authentication key of up to 8 characters and is standardized. The second is still experimental and uses the MD5 algorithm and an authentication key of up to 16 characters.

The simple password provides very little protection because in many cases it is possible to easily capture packets from the network and learn the authentication key. The experimental MD5 algorithm provides much more protection as it does not include the authentication key in the packet.

The OSPF specification currently specifies that the authentication type be configured per area with the ability to configure separate passwords per interface. This has been extended to allow the configuration of different authentication types and keys per interface. In addition, it is possible to specify both a *primary* and a *secondary* authentication type and key on each interface. Outgoing packets use the primary authentication type, but incoming packets can match either the primary or secondary authentication type and key.

The OSPF statement

```
ospf ( yes | no | on | off ) [ {
    defaults {
        precedence precedence ;
        cost cost ;
        tag [ as ] tag ;
        type 1 | 2 ;
        inherit-metric ;
    } ;
    exportlimit routes ;
    exportinterval time ;
    traceoptions trace_options ;
    monitorauthkey authkey ;
    monitorauth none | ( [ simple | md5 ] authkey ) ;
    backbone | ( area area ) {
        authtype  none | simple ;
        stub [ cost cost] ;
        networks {
            network [ restrict ] ;
            network mask mask [ restrict ] ;
            network ( masklen | / ) number [ restrict ] ;
            host host [ restrict ] ;
        } ;
        stubhosts {
            host  cost cost ;
        } ;
        interface interface_list ; [cost cost ] {
            interface_parameters
        } ;
        interface interface_list nonbroadcast [cost cost ] {
            pollinterval time ;
            routers {
                gateway [ eligible ] ;
            } ;
            interface_parameters
        } ;
        Backbone only:
        virtuallink neighborid router_id  transitarea area {
            interface_parameters
        } ;
    } ;
} ] ;
```

The following are the *interface_parameters* referred to previously. These can be specified on any class of interface and are described under the **interface** clause.

enable | disable | passive ;

retransmitinterval *time* **;**

transitdelay *time* ;

priority *priority* ;

hellointerval *time* ;

routerdeadinterval *time* ;

authkey *auth_key* ;

defaults

These parameters specify the defaults used when importing OSPF ASE routes into the GateD routing table and exporting routes from the GateD routing table into OSPF ASEs.

precedence *precedence*

Sets the global precedence for OSPF ASE incoming routes. This precedence can be overridden by a precedence specified by the import policy. The default precedence for OSPF ASE routes is 150. This keyword can also be stated as **protocol-precedence** or **proto-precedence**.

cost *cost*

The cost is used when exporting a non-OSPF route from the GateD routing table into OSPF as an ASE. This can be explicitly overridden in export policy by the **metric** keyword. The default value is 1.

Note: This parameter has no effect on the **cost** keyword on the interface command.

tag [**as**] *tag*

OSPF ASE routes have a 32-bit tag field that is not used by the OSPF protocol, but can be used by export policy to filter routes. When OSPF is interacting with an EGP, the tag field can be used to propagate AS path information, in which case the **as** keyword is specified and the tag is limited to 12 bits of information. If not specified, the tag is set to zero.

type *1* | *2*

Routes imported from BGP into OSPF default to becoming type 2 ASEs. This default can be explicitly changed here and overridden in export policy.

inherit-metric

Inherit-metric allows an OSPF ASE route to inherit the metric of the external route from the BGP MED when no metric is specified on the export. This option maintains compatibility with all the current export functions:

- A metric specified on the export takes precedence.
- The cost specified in the default is used if inherit-metric is not specified.

ASE export rate

Because of the nature of OSPF, the rate at which ASEs are flooded must be limited. These two parameters can be used to adjust those rate limits.

exportinterval *time*

This specifies how often a batch of ASE link state advertisements is generated and flooded into OSPF. Time is specified in seconds. The default is once per second.

exportlimit *routes*

This parameter specifies how many ASEs are generated and flooded in each batch. The default is 100.

traceoptions *trace_options*

Specifies the tracing options for OSPF. (See Trace Statements (page A-12) and the OSPF specific tracing options below.)

monitorauthkey *authkey*

OSPF state can be queried using the **ospf_monitor** (This should be a hyperlink) utility. This utility sends non-standard OSPF packets that generate a text response from OSPF. By default, these requests are not authenticated. If an authentication key is configured, the incoming requests must match the specified authentication key. No OSPF state can be changed by these packets, but the act of querying OSPF can utilize system resources.

backbone
area *area*

Each OSPF router must be configured into at least one OSPF area. If more than one area is configured, at least one must be the backbone. The backbone can only be configured using the **backbone** keyword, it can not be specified as *area* 0. The backbone interface can be a **virtuallink**.

authtype none | **simple**

OSPF specifies an authentication scheme per area. Each interface in the area must use this same authentication scheme although it can use a different **authenticationkey**. The currently valid values are **none** for no authentication, or **simple** for simple password authentication. The default is **none**.

If **authkey simple** is specified but there is no **authkey** ***xxx*** in the **/etc/gated.conf** file, this is equivalent to specifying **authkey none**. In both cases, authentication is disabled.

stub [**cost** *cost*]

A **stub** area is one in which there are no ASE routes. Specifying **stub** prevents participating routers from advertising ASE routes into that area. If a **cost** is specified, this is used to inject a default route into the area with the specified *cost*. This is not valid for backbone area. The **stub** keyword must be specified on all routers within a given area.

networks

The **networks** list describes the scope of an area. Intra-area LSAs that fall within the specified ranges are not advertised into other areas as inter-area routes. Instead, the specified ranges are advertised as *summary network* LSAs. If **restrict** is specified, the summary network LSAs are not advertised. Intra-area LSAs that do not fall into any range are also advertised as summary network LSAs. This option is very useful on well-designed networks in reducing the amount of routing information propagated between areas. The entries in this list are either networks, or a subnetwork/mask pair. See the section on Route Filtering for more detail about specifying ranges (page A-86).

stubhosts

This list specifies directly-attached hosts that should be advertised as reachable from this router and the costs they should be advertised with. Point-to-point interfaces on which it is not desirable to run OSPF should be specified here.

It is also useful to assign a additional address to the loopback interface (one not on the 127 network) and advertise it as a stubhosts. If this address is the same one used as the `router-id`, it enables routing to OSPF routers by **router-id**, instead of by interface address. This is more reliable than routing to one of the router's interface addresses that can not always be reachable.

interface *interface_list* [**cost** *cost*]

This form of the interface clause is used to configure a **broadcast** (that requires IP multicast support) or a **point-to-point** interface. See the section on Interface list specification (page A-19) for the description of the *interface_list*.

Each interface has a ***cost***. The costs of all interfaces a packet must cross to reach a destination are summed to get the cost to that destination. The default cost is one, but another non-zero value can be specified.

Interface parameters common to all types of interfaces are:

enable | **disable** | **passive**

Enable or disable OSPF on the interface. If **passive** is used, OSPF behaves as though it is enabled on the interface (the interface's routes are included in LSAs issued by the router), but OSPF packets are not sent or accepted on the interface. This can be useful for advertising DMZs or other interconnect networks into OSPF, as an alternative to advertising them into ospfase.

retransmitinterval *time*

The number of seconds between link state advertisement retransmissions for adjacencies belonging to this interface.

transitdelay *time*

The estimated number of seconds required to transmit a link state update over this interface. Transitdelay takes into account transmission and propagation delays and must be greater than 0.

priority *priority*

A number between 0 and 255 specifying the priority for becoming the designated router on this interface. When two routers attached to a network both attempt to become the designated router, the one with the highest priority wins. A router whose router priority is set to 0 is ineligible to become designated router.

OSPF supports both NBMA and P2P interfaces. The priority for these interfaces must be manually configured to elect the designated router.

If no priority keyword is present, priority defaults to 0. At least one router on a network must have a non-zero priority if OSPF is to work at all.

hellointerval *time*

The length of time, in seconds, between Hello packets that the router sends on the interface.

routerdeadinterval *time*

The number of seconds not hearing a router's Hello packets before the router's neighbors declares it down.

authkey *auth_key*

Used by OSPF authentication to generate and verify the authentication field in the OSPF header. The authentication key can be configured on a per-interface basis. It is specified by one to eight decimal digits separated by periods, or a one- to eight-character string in double quotes. The password “01.02.03” is equivalent to “1.2.3” since the fields are treated as numeric values, not just as an ASCII string. Similarly, the password “0” is equivalent to “0.0.0.0”.

If no **authkey** keyword is present, authentication is disabled exactly as if **authtype none** has been specified.

Point-to-point interfaces also support this additional parameter:

nomulticast

By default, OSPF packets to neighbors on point-to-point interfaces are sent via the IP multicast mechanism. Although some implementations of IP multicasting for UNIX have a bug that precludes the use of IP multicasting on these interfaces. GateD detects this condition and fall back to sending unicast OSPF packets to this point-to-point neighbor.

If the use of IP multicasting is not desired because the remote neighbor does not support it, the **nomulticast** parameter can be specified to force the use of unicast OSPF packets. This option can also be used to eliminate warnings when GateD detects the bug mentioned above.

This option can only be used on interfaces configured with a point-to-point flag using **ifconfig**.

interface *interface_list* **nonbroadcast** [**cost** *cost*]

This form of the interface clause is used to specify a **nonbroadcast** interface on a **non-broadcast multi-access** (NBMA) medium. Since an OSPF **broadcast** medium must support IP multicasting, a broadcast-capable medium that does not support IP multicasting must be configured as a non-broadcast interface. This may be needed on interfaces that use Frame Relay or on point-to-point interfaces that do not support multicasting.

A non-broadcast interface supports any of the standard **interface** clauses listed above, plus the following two that are specific to non-broadcast interfaces:

pollinterval *time*

Before adjacency is established with a neighbor, OSPF packets are sent periodically at the specified **pollinterval**.

routers

By definition, it is not possible to send broadcast packets to discover OSPF neighbors on a non-broadcast medium, so all neighbors must be configured. The list includes one or more neighbors and an indication of their eligibility to become a designated router.

virtuallink neighborid *router_id* **transitarea** *area*

Virtual links are used to establish or increase connectivity of the backbone area. The **neighborid** is the *router_id* of the other end of the virtual link. The transit *area* specified must also be configured on this system. All standard interface parameters defined by the **interface** clause above can be specified on a virtual link.

Tracing options

In addition to the following OSPF specific trace flags, OSPF supports the **state** that traces interface and neighbor state machine transitions.

lsabuild

Link State Advertisement creation

spf

Shortest Path First (SPF) calculations

Packet tracing options (that can be modified with **detail**, **send** and **recv**):

hello
OSPF **HELLO** packets that are used to determine neighbor reachability.

dd
OSPF Database Description packets that are used in synchronizing OSPF databases.

request
OSPF Link State Request packets that are used in synchronizing OSPF databases.

lsu
OSPF Link State Update packets that are used in synchronizing OSPF databases.

ack
OSPF Link State Ack packets that are used in synchronizing OSPF databases.

IS-IS Intra-Domain Protocol

IS-IS is a link state interior gateway protocol (IGP) originally developed for routing ISO/CLNP (International Organization for Standardization/Connectionless Network Protocol) packets. The version distributed with GateD can route IP packets as well.

In ISO terminology, a router is referred to as an “intermediate system” (IS). IS-IS intra-domain routing is organized hierarchically so that a large domain can be administratively divided into smaller areas using level 1 intermediate systems within areas and level 2 intermediate systems between areas.

Routing between administrative domains is handled by Border Intermediate Systems (BISs) using IDRP, the inter-domain routing protocol. Level 1 systems route directly to systems within their own area and route toward a Level 2 Intermediate System when the destination system is in a different area. Level 2 Intermediate Systems route between areas and keep track of the paths to destination areas. Systems in the Level 2 subdomain route towards a destination area, or another routing domain. As with any internet routing protocol, IS-IS support for large routing domains can also include many types of individual subnetworks. These subnetworks can include point-to-point links, multipoint links and broadcast subnetworks like ISO 8802 LANs.

In IS-IS, all subnetwork types are treated by the subnetwork independent functions as though they were connectionless subnetworks using subnetwork convergence functions where necessary. Like OSPF, IS-IS uses a “shortest-path first” algorithm to determine routes. GateD configuration syntax allows as much autoconfiguration as possible, reducing the probability of error.

This integration also allows the ability to specify policy for exchanging routing information with other protocols running in GateD.

IS-IS Statement

This statement enables the IS-IS protocol in GateD and configures the interfaces that are to run IS-IS. By default, IS-IS is disabled. The IS-IS statement consists of an initial description of the Intermediate System and a list of statements that determine the configuration of the specific circuits and networks to be managed.

```
isis no | ip {
    level 1|2 ;
    [traceoptions <isis_traceoptions> ;]
    [systemid <string > ;]
    [area <string > ;]
    [set <isis_parm > value ;]
    circuit|interface <interface-name >
        metric [level 1|2] metric
        priority [level 1|2] priority pointopoint ;
    [ipreachability level (1|2) (internal|external|summary)
        ipaddr netmask [metric metric;]]
} ;
```

Statements can appear in any order and include the following:

level

Indicates whether GateD is running on a Level 1 (intra-area) or Level 2 (inter-area) IS. The default is Level 1.

traceoptions

These are covered in the Trace Options section.

systemid *string*

This is a mandatory parameter. If no system identifier is specified, GateD reports an error "`systemID is not set`", and abort. User can specify **systemid** in two ways, as a hex string or as a character string. If the string starts with 0x or 0X, GateD interprets it as hex string. If it does not contain 0x or 0X, it interprets it as a character string. However, the trace file prints all the values as hex strings.

area *string*

This is a mandatory parameter. IS-IS area addresses are configured based on this parameter. If **area** is not specified in the configuration file, GateD reports an error "area address is not set", and aborts. User can specify **area** in two ways - as a hex string or as a character string. If the string starts with 0x or 0X, GateD interprets it as hex string. If it does not contain 0x or 0X, it interprets as character string. However, the trace file prints all the values as hex strings.

circuit | interface *interface-name*

Each **circuit** statement specifies one of the circuits or interfaces the system manages. Circuits normally correspond to UNIX interfaces, with *string* being the interface name. The **circuit** attributes are a list of options that can appear in any order in the **circuit** statement.

metric level [**1** | **2**] *metric*

Allows specifications of Level 1 and Level 2 metrics for each circuit. Only the default metric type is supported. IS-IS metrics must be in the range 1 to 63. If no metric is set for the circuit, the default value is 63. If only level 2 is specified, then this circuit is configured to run level 2 only, not level 1 at all.

priority [**level** [**1** | **2**] *priority*]

Determines designated router election results; higher values give a higher likelihood of becoming the designated router. The level defaults to Level 1. If no priority is specified, priority is set to a random value between 0 and 127.

On a level 2 IS, to configure a circuit with a Level 1 metric of 10 and a Level 2 metric of 20, add two metric options to the circuit statement.

The default Level is 1; the default metric is 63. The default precedence for IS-IS Level 1 is 15. For IS-IS Level 2, the default precedence is 18.

pointopoint

Allows this circuit to be point-to-point type. User has to specify this if they want this circuit to be a point-to-point type for IS-IS.

ipreachability level (1 | 2) internal | external | summary *ipaddr netmask* [**metric** metric]

The **ipreachability** statement is used to specify IP networks that are advertised as reachable by the IS-IS System. There are five parts to the options - the keyword ipreachability, the level at which the route has to be advertised, the type of reachability

(internal, external or summary), the network (specified by IP address and network mask), and the metric associated with the network.

All, with the exception of metric, are mandatory. IP addresses and masks are specified in dot notation. These networks are assumed to be directly attached networks to this router. If this is not a directly-attached network, the network is not advertised in the IS-IS packets. **ipreachability** with the **summary** keyword is used to summarize the networks.

set isis_parm *integer*

The following list shows the names and default values of the variables that can be changed using the **set** statement. These can appear in any order in the **isis** statement.

origL1LSPBufSize *integer*

Allows you to set originating Level 1 LSP Buffer size to given number of bytes. The default value is 1497 octets. This size must be less than or smallest MTU of any link the network. Normally 1497 is sufficient unless the MTU of a link is lowered to be less than 1497. All the routers in the network has to be specified the same size of LSP. Otherwise, the results are not predictable.

origL2LSPBufSize *integer*

Allows you to set originating Level 2 LSP Buffer size to given number of bytes. The default value is 1497 octets. This size must be less than or smallest MTU of any link the network. Normally 1497 is sufficient unless the MTU of a link is lowered to be less than 1497. All the routers in the network has to be specified the same size of LSP. Otherwise, the results are not predictable.

dataLinkBlocksize *integer*

Allows you to set the maximum size of Hello packets to given number of bytes. The default value is 1497 octets. This size must be less than or smallest MTU of any link in the network. Normally 1497 is sufficient unless the MTU of a link is lowered to be less than 1497. All the routers in the network has to be specified the same size of LSP. Otherwise, the results are not predictable.

sysHoldingTimer *integer*

Specifies the multiplier for the hold timer in Hello packets. The value specified by **sysHoldingTimer** multiplied by the **sysIIHInterval** gives the holding time for Hello packets. The default value for **sysHoldingTimer** is 3. If the receiving system does not receive a hello packet by the expiration of the holding time, it drops the adjacency formed with the transmitting system.

sysISHInterval *integer*

Specifies the transmit frequency of ISH PDUs on a point-to-point circuit. The default value is 10 seconds; the range is from 5 to 30 seconds.

sysIIHInterval *integer*

Specifies the transmit frequency of IIH PDUs on a broadcast circuit. The default value is 3 seconds; the range is 1 to 10,000 seconds.

minLSPGenInterval *integer*

Specifies the minimum interval between successive LSP generations. The default value is 30 seconds.

maxLSPGenInterval *integer*

Specifies the maximum interval between successive LSP generations. The default value is 900 seconds.

minLSPXmitInterval *integer*

Specifies the minimum LSP transmission interval on a point to point circuit. The default value is 2 seconds.

minBLSPXmitinterval *integer*

Specifies the minimum LSP transmission interval on a broadcast circuit. The default value is 5 seconds.

BLSPThrottle *integer*

Specifies the maximum number of LSPs sent per **minBLSPXmitInterval**. The default value is 50 seconds.

maximumAge *integer*

Specifies the initial (and maximum) life time of any LSP. This is stored in the lifetime field in the LSP. The default value is 1200 seconds.

completeSNPInterval *integer*

Specifies the transmit frequency of CSNPs. The default value is 10 seconds.

partialSNPInterval *integer*

Specifies the transmit frequency of PSNPs. The default value is 2 seconds.

zeroAgeLifetime *integer*

Specifies the time a LSP with zero lifetime is held before purging it from the LSP database. The default value is 60 seconds.

dumpDBinterval *integer*

Specifies the frequency with which LSP database is traced. The default value is 30 seconds.

Tracing options

IS-IS has its own internal tracing flags that are distinct from the flags maintained globally by GateD. The mechanism to set tracing is via the **isis traceoptions** statement, in the form:

traceoptions ["*trace_file*" [**replace**][**size** *size* [**k**|**m**] **files** *files*]]
isis_trace_options;

isis_trace_options can include one or more of the following:

all	-	everything below
iih	-	IIHs sent and received
lanadj	-	lan adjacency updates
p2padj	-	point to point adjacency updates
lspdb	-	signatures in the LSP database
lspcontent	-	contents of LSPs in the database
lspinput	-	input processing of LSPs
flooding	-	flooding of all LSPs
buildlsp	-	generation of local LSPs
csnp	-	processing and construction of CSNPs
psnp	-	processing and construction of PSNPs
route	-	route changes
update	-	individual routes changed
paths	-	route paths as calculated by spf algorithm
spf	-	running of spf algorithm
events	-	interesting protocol events

IS-IS configuration task lists

To configure IS-IS, the user must complete the following tasks. These tasks are mandatory. Other parameters are optional, and GateD uses default values. Users might want to set the other parameters depending on their network topology.

- configure ISO address for the interfaces (in `grifconfig.conf`)
- configure PPP, ATM and Frame Relay for IS-IS
- enable IS-IS in `gated.conf`
- configure **systemid** in `gated.conf`
- configure **area** in `gated.conf`
- configure **interface** in `gated.conf`

Configure ISO address for the interface

This is a mandatory task.

Assign ISO address for the interface. You can come up with ISO address by appending systemid to the **area**. For example, if the **area** address is 49.0000.80 and **systemid** you assigned for the interface is 3260.3260.3260, ISO address becomes 49.0000.80.3260.3260.3260.00. You have to include the last 00 byte when you are configuring ISO address. This is NSEL parameter specified in the ISO addresses. However, the NSEL parameter is not part of the **systemid** in **/etc/gated.conf**. Each system should have a unique **systemid**. If you configure two systems with the same **systemid**, the results are unpredictable. You can add this address in `grifconfig.conf` file.

The syntax for adding the address in `grifconfig.conf` is

```
<interface-name>  <iso-address> <iso-area> - iso
```

Example:

```
gf030 49.0000.80.3260.3260.3260.00 49.0000.80 - iso
```

Configure PPP, ATM and Frame Relay for IS-IS

If you are using PPP, Frame Relay or ATM interfaces, you need to follow this task.

This is a mandatory task. For PPP, you have to add "`enable osinlcp`" in `/etc/grppp.conf` file in order to make IS-IS work over PPP. This is just like existing "`enable ipcp`" line in `/etc/grppp.conf`.

For Frame Relay mode, you have to edit `/etc/grfr.conf` file and edit the PVC line to add `ISIS=Y`. This enables the PVC to start accepting IS-IS packets.

For ATM interfaces, you have to edit `/etc/gratm.conf` file and add `proto=isis` or `proto=isis_ip` for the desired PVC to accept IS-IS packets. The following examples show the additional configurations necessary to enable IS-IS over the media.

Example for grppp.conf:

```
enable osinlcp
```

Example for grfr.conf:

```
pvc gs030 101 192.0.2.99 Enabled=Y Name="Router2" ISIS=Y
```

Example for gratm.conf:

```
pvc ga030 0/40 proto=isis traffic_shape=high_speed_high_quality
pvc ga030 0/41 proto=isis_ip traffic_shape=high_speed_high_quality
```

Enable IS-IS in etc/gated.conf:

This is a mandatory task. IS-IS is enabled by using "**isis yes**" or "**isis ip**" in GateD configuration file.

Example:

```
isis yes {
} ;
```

This statement enables the IS-IS protocol in GateD. By default IS-IS is disabled. The above statement alone does not bring up IS-IS. GateD aborts with "`systemid is not set`" message. **systemid** and **area** are mandatory parameters for IS-IS to run.

Configure area in /etc/gated.conf

This is a mandatory task. **area** can be specified as a hex string or as a character string in GateD configuration. Two routers in Level-1 mode communicate with each other only if they belong to the same area. Two level-2 routers can communicate across area boundaries.

Example 1:

```
isis yes {
        area "0x49000080";
} ;
```

Example 2:

```
isis yes {
         area "aaaa";
         # The above string is printed in the trace file as 61.6161.61
} ;
```

Configure systemid in /etc/gated.conf

This is a mandatory task. **systemid** is specified as a hex string or as a character string in GateD configuration. Each router in the network should have a unique **systemid** and should be 6 characters long. **systemid** has to be specified as 12 hex digits in the hex string format or as 6 characters in the string format. As with **area**, the character string is printed in the hex format in trace file.

Example:

```
isis yes {
         area "0x49000080";
```

```
            systemid "0x326032603260";
    } ;
```

Example:

```
    isis yes {
            area "0x49000080";
            systemid "aaaaaa";
    } ;
```

Configure interface in /etc/gated.conf

This is a mandatory task. If user does not configure **interface** in GateD configuration file, the interface is not transmitting/receiving IS-IS packets. **interface** can be enabled by using either `interface` or `circuit` keyword. User can specify any of the optional `circuit` parameters allowed on this statement.

Example:

```
    isis yes {
           area "49000080";
           systemid "326032603260";
           interface "gf030" metric 10 priority 60;
    } ;
```

The BGP protocol

The Border Gateway Protocol (BGP) is an exterior routing protocol used for exchanging routing information between autonomous systems. BGP is used for exchange of routing information between multiple transit autonomous systems as well as between transit and stub autonomous systems. BGP is related to EGP, but has more capability, greater flexibility, and less required bandwidth. BGP uses path attributes to provide more information about each route, and in particular to maintain an AS path, that includes the AS number of each autonomous system the route has transited, providing information sufficient to prevent routing loops in an arbitrary topology. Path attributes can also be used to distinguish between groups of routes to determine administrative preferences, allowing greater flexibility in determining route preference to achieve a variety of administrative ends.

BGP supports two basic types of sessions between neighbors, internal (sometimes referred to as IBGP) and external. Internal sessions are run between routers in the same autonomous system, while external sessions run between routers in different autonomous systems. When sending routes to an external peer, the local AS number is prepended to the AS path. This means that routes received from an external peer are guaranteed to have the AS number of that peer at the start of the path. In general, routes received from an internal neighbor do not have the local AS number prepended to the AS path, and hence has the same AS path that the route had when the originating internal neighbor received the route from an external peer. Routes with no AS numbers in the path can be legitimately received from internal neighbors. These routes should be considered internal to the receiver's own AS.

The BGP implementation supports three versions of the BGP protocol, versions 2, 3 and 4. BGP versions 2 and 3 are quite similar in capability and function. They only propagate classed network routes, and the AS path is a simple array of AS numbers. BGP 4 propagates fully general address-and-mask routes, and the AS path structure can represent the results of aggregating dissimilar routes.

External BGP sessions may or may not include a single metric, that BGP calls the multi-exit discriminator (MED), among the path attributes. For BGP versions 2 and 3, this metric is a 16-bit unsigned integer. For BGP version 4, it is a 32-bit unsigned integer. Smaller values of the MED are preferred. Currently this metric is only used to break ties between routes with equal preference from the same neighboring AS.

Internal BGP sessions carry at least one metric in the path attributes, that BGP calls the LocalPref. The range of LocalPref is identical to the range of the MED. For BGP versions 2 and 3, a route is preferred if its value for LocalPref is smaller. For BGP version 4, a route is preferred if its value for this metric is larger. BGP version 4 internal sessions can optionally include a second metric, the MED, carried in from external sessions. The use of these metrics is dependent on the type of internal protocol processing that is specified.

BGP collapses routes with similar path attributes into a single update for advertisement. Routes that are received in a single update are readvertised in a single update. The churn caused by the loss of a neighbor is minimized and the initial advertisement sent during peer establishment are maximally compressed. BGP does not read information from the kernel message-by-message, but fills the input buffer. It processes all complete messages in the buffer before reading again. BGP also does multiple reads to clear all incoming data queued on the socket. This feature can cause other protocols to be blocked for prolonged intervals by a busy peer connection.

All unreachable messages are collected into a single message and sent prior to reachable routes during a flash update. For these unreachable announcements, the nexthop is set to the local address on the connection, no metric is sent, and the path origin is set to incomplete. On external connections the AS path in unreachable announcements is set to the local AS. On internal connections, the AS path is set to length zero.

The BGP implementation expects external peers to be directly attached to a shared subnet, and expects those peers to advertise nexthops that are host addresses on that subnet (though this constraint can be relaxed by configuration for testing). For groups of internal peers, however, there are several alternatives that can be selected from by specifying the group type and route reflection options. Type **internal** groups expect all peers to be directly attached to a shared subnet so that, like external peers, the nexthops received in BGP advertisements can be used directly for forwarding. Type **routing** groups instead determine the immediate nexthops for routes by using the nexthop received with a route from a peer as a forwarding address, and by using this to look up an immediate nexthop in an IGP's routes. Such groups support distant peers, but need to be informed of the IGP or IGPs whose routes they are using to determine immediate nexthops.

For internal BGP group types (and for test groups), where possible, a single outgoing message is built for all group peers based on the common policy. A copy of the message is sent to every peer in the group, with possible adjustments to the next-hop field as appropriate to each peer. This minimizes the computational load of running large numbers of peers in these types of groups. BGP allows unconfigured peers to connect if an appropriate group has been configured with an **allow** clause.

Route reflection

Generally, all border routers in a single AS need to be internal peers of each other, and in fact all non-border routers frequently need to be internal peers of all border routers. While this is usually acceptable in small networks, it can lead to unacceptably large internal peer groups in large networks. To help address this problem, BGP supports route reflection for internal peer groups (with BGP version 4 only). When using route reflection, the rule that a router can not readvertise routes from internal peers to other internal peers is relaxed for some routers, called route reflectors. A typical use of route reflection might involve a "core" backbone of fully meshed routers ("fully meshed" means all the routers in the fully meshed group peer directly with all other routers in the group), some of which act as route reflectors for routers that are not part of the core group.

Two types of route reflection are supported. By default, all routes received by the route reflector from a client are sent to all internal peers (including the client's group but not the client itself). If the **no-client-reflect** option is enabled, routes received from a route reflection client are sent only to internal peers that are not members of the client's group. In this case, the client's group must itself be fully meshed. In either case, all routes received from a non-client internal peer are sent to all route reflection clients.

Typically, a single router acts as the reflector for a set, or cluster, of clients. However, for redundancy two or more can also be configured to be reflectors for the same cluster. In this case, a cluster ID should be selected to identify all reflectors serving the cluster, using the **clusterid** keyword. Gratuitous use of multiple redundant reflectors is not advised, as it can lead to an increase in the memory required to store routes on the redundant reflectors' peers.

No special configuration is required on the route reflection clients. From a client's perspective, a route reflector is simply a normal IBGP peer. Any BGP version 4 speaker should be able to

be a reflector client. (Note however that GateD versions 3.5B3 and earlier, and 3.6A1 and earlier, contain a bug that prevents them from acting as route reflection clients.)

Readers are referred to the route reflection specification document (RFC 1965 as of this writing) for further details.

Confederations

In addition to improvements in routing policy control, current techniques for deploying BGP among speakers in the same autonomous system generally require a full mesh of TCP connections among all speakers for the purpose of exchanging exterior routing information. In autonomous systems the number of intra-domain connections that need to be maintained by each border router can become significant. While this is usually acceptable in small networks, it can lead to unacceptably large internal peer groups in large networks. To help address this problem, BGP supports confederations for internal peer groups (with BGP version 4 only). It can be useful to subdivide an autonomous system into smaller domains for purposes of controlling routing policy via information contained in the BGP **AS_PATH** attribute (similar to EBGP). For example, one can chose to consider all BGP speakers in a geographic region as a single entity (confederation).

Subdividing a large autonomous system allows a significant reduction in the total number of intra-domain BGP connections as the connectivity requirements simplify to the model used for inter-domain connections.

There is usually no need to expose the internal topology of this divided autonomous system, that means it is possible to regard a collection of autonomous systems under a common administration as a single entity or autonomous system when viewed from outside the confines of the confederation of autonomous systems itself.

Operationally, a member of a BGP confederation uses its confederation identifier in all transactions with peers that are not members of its confederation. This confederation identifier is considered to be the "externally visible" AS number, and this number is used in **OPEN** messages and advertised in the **AS_PATH** attribute.

A member of a BGP confederation uses its routing domain identifier (the internally visible AS number) in all transactions with peers that are members of the same confederation as the given router.

A BGP speaker receiving an **AS_PATH** attribute containing a confederation ID matching its own confederation shall treat the path in the same fashion as if it had received a path containing its own AS number.

Thus, BGP peering sessions as the confederation border are analogous to external BGP transactions within an AS; and BGP peering sessions within the confederation are analogous to internal BGP transactions.

Readers are referred to the confederation specification document (RFC 1966) for further details.

Communities

The communities attribute allows the administrator of a Routing Domain to tag groups of routes with a community tag. The tag consists of 2 octets of autonomous system (AS) and 2 octets of community ID. The community attribute is passed from routing domain to routing domain to maintain the grouping of these routes. A set of routes can have more than one community tag in its community attribute.

Communities import and export policy is configured using the **aspath-opt** clause (or **mod-aspath** clause) to the **group**, **import** and **export** statements.

Please refer to the Communities specification and its accompanying usage documents (RFC 1997 and RFC 1998 as of this writing) for further details on BGP communities.

Subgroups: differing policy for peers in the same AS/RDI

The **subgroups** variable is used to designate different sub-groups of BGP peers within the same AS (designated by the **group** statement). It works in conjunction with the subgroup identifier on the import and/or export policy configuration using the export or import clause. Essentially, the designation of the mathematical group (the group can have, at minimum, a membership of one element [peer statement]) allows for the extensibility and flexibility for "policy per peer."

Multi-exit discriminator (MED)

The Multi-exit discriminator (MED) allows the administrator of a Routing Domain to choose between various exits from a neighboring AS. This attribute is used only for decision making in choosing the best route to the neighboring AS. If all the other factors for a path to a given AS are equal, the path with the lower MED value takes preference over other paths.

This attribute is not propagated to other neighboring ASs. However, this attribute can be propagated to other BGP speakers within the same AS.

The MED attribute, for BGP version 4, is a four-octet unsigned integer.

MED is originated using the **metricout** option of the export, group and/or peer statement. It is imported using the **med** keyword on the BGP group statement. A MED can also be created from IGP metrics when exporting one protocol into another (for example, a redistribution of OSPF into BGP).

Weighted route dampening

The basic idea of weighted route dampening is to treat routes that are being announced and withdrawn (flapping) at a rapid rate as unreachable.

If a route flaps at a low rate, it should not be suppressed at all, or suppressed only for a brief period of time. With weighted route dampening, the suppression of a route or routes occurs in a manner that adapts to the frequency and duration that a particular route appears to be flapping. The more a route flaps during a period of time, the longer it is suppressed. The adaptive characteristics of weighted route dampening are controlled by a few configurable parameters.

For syntax and configuration information, refer to the Weighted Route Dampening statement (page A-70).

Local_Pref

Routes propagated by IBGP must include a **Local_Pref** attribute. **Local_Pref** can be used by a BGP speaker to inform other BGP speakers in its own autonomous system of the originating speaker's degree of preference for an advertised route. Unless the **localpref** option has been set on the **import** or **export** statements, BGP sends the **Local_Pref** path attribute as 100.

GateD always uses the received **Local_Pref** to select between BGP routes. BGP routes with a larger **Local_Pref** are preferred. The range of values for local preference are 0 - 2**31.

Note: All routers in the same network that are running GateD and participating in IBGP should use **localpref** uniformly. That is, if one router has **localpref** set, all should set it, and all should use the same value.

Configurable Export-Best-BGP (CEBB)

The configurable export-best-BGP (CEBB) functionality allows GateD to export BGP routes that are not actually installed because there is an IGP route that takes precedence over the BGP route. Essentially, if a route prefix is known from a protocol other than BGP and is more preferred than BGP (for example, an IGP), CEBB allows the most preferred BGP route to be exported without having to redistribute the IGP into the BGP. This is especially useful for external BGP peering relationships where IGP routes are not generally redistributed.This allows for simpler policy configuration because explicit protocol redistribution is no longer required.

Additionally, CEBB increases BGP stability both internally and externally. For example, if an IGP is redistributed into BGP, and there is some level of instability within the IGP, it would appear as a BGP flap to external peers.In this case, CEBB can increase stability because the external BGP peer always receives the “best” BGP route regardless of whether a matching IGP route was installed.

This functionality is on by default; to disable it, you must add the **disable export best** clause as a global statement in the BGP portion of the **/etc/gated.conf** file (see the BGP statement subsection for more information (page A-57).

Route selection

BGP selects the best path to an AS from all the known paths and propagates the selected path to its neighbors. GateD uses the following criteria, in the following order, to select the best path. If routes are equal at a given point in the selection process, then the next criterion is applied to break the tie:

1. Configured policy: the route with smallest preference, as determined by the policy defined in **/etc/gated.conf**.
2. **Local_Pref**: the route with the highest BGP local preference.
3. Shortest AS Path: the route with the fewest ASs listed in its AS path.

4 Origin IGP < EGP < Incomplete: the route with an AS path origin of IGP is preferred. Next in precedence is the route with AS path origin of EGP. Least preferred is an AS path that is incomplete.

5 MED (if not ignored): the route with the best multi-exit discriminator (MED) is preferred. MEDs are only compared between routes that are received from the same neighbor AS. (That is, this test is only applied if the local AS has two or more connections to a given neighbor AS.)

6 Shortest IGP distance: the route whose nexthop is closer (with respect to the IGP distance) is preferred.

7 Source IGP < EBGP < IBGP: first, prefer the strictly interior route, then the strictly exterior route, then the exterior route learned from an interior session.

8 Lowest Router ID: the route whose nexthop IP address is numerically lowest.

The BGP statement

```
bgp ( yes | no | on | off )
{
    [ precedence precedence ; ]
    [ preference preference ; ]
    [ allow bad community ; ]
    [ defaultmetric metric ; ]
    [ traceoptions trace_options ; ]
    [ clusterid host ; ]
    [ disable export best ; ]
 [ group type
        ( external peeras autonomous_system
            [ ignorefirstashop ][ subgroup integer ]
            [ med ]  [ nexthopself ] )

       | ( internal peeras autonomous_system
            [ reflector-client [ no-client-reflect ] ]
            [ ignorefirstashop ]
            [ lcladdr local_address ]
            [ outdelay time ]
            [ metricout metric ]
            [ subgroup integer ]
            [ nexthopself ] )

      | ( routing peeras autonomous_system proto proto_list
            interface interface_list
            [ reflector-client [ no-client-reflect ] ]
            [ ignorefirstashop ]
            [ lcladdr local_address ]
            [ outdelay time ]
            [ metricout metric ]
            [ reflector-client [ no-client-reflect ] ]
            [ subgroup integer ]
            [ nexthopself ] )

      | ( confed peeras autonomous_system proto proto_list
            interface interface_list
            [ reflector-client [ no-client-reflect ] ]
            [ ignorefirstashop ]
```

```
                [ lcladdr local_address ]
                [ outdelay time ]
                [ metricout metric ]
                [ reflector-client [ no-client-reflect ] ]
                [ subgroup integer ]
                [ nexthopself ] )

        | ( test peeras autonomous_system ) ]
          {
            [ allow {
                  network
                  network mask mask
                  network ( masklen | / ) number
                  all
                  host host ]
              } ;
              peer host
                  [ metricout metric ]
                  [ ignorefirstashop ]
                  [ nogendefault ]
                  [ gateway gateway ]
                  [ precedence precedence ]
                  [ preference preference ]
                  [ lcladdr local_address ]
                  [ holdtime time ]
                  [ version number ]
                  [ passive ]
                  [ sendbuffer number ]
                  [ recvbuffer number ]
                  [ outdelay time ]
                  [ keep [ all | none ] ]
                  [ show-warnings ]
                  [ noaggregatorid ]
                  [ keepalivesalways ]
                  [ v3asloopokay ]
                  [ nov4asloop ]
                  [ ascount count ]
                  [ throttle count ]
                  [ allow bad routerid ]
```

```
                    [ logupdown ]
                    [ ttl ttl ]
                    [ traceoptions trace_options ]
                    [ nexthopself ]
                    ;
        } ;
    } ;
```

The **bgp** statement enables or disables BGP. By default, BGP is disabled. The default metric for announcing routes via BGP is no metric.

precedence *precedence*

Sets the global precedence for BGP incoming routes. The default precedence is 170. This precedence can be overridden by a precedence specified on the **peer** statement, or by the import policy. This keyword can also be stated as **protocol-precedence** or **proto-precedence**.

preference *preference*

Sets the preference of all BGP peers in the following groups. This preference can be used to prefer routes learned from one group of peers over others. This preference can be overridden by preference set for a particular peer or by the import policy. This keyword can also be stated as **pref**.

allow bad community

BGP communities are specified in RFC 1997 as being (logically) a 16-bit AS number followed by an arbitrary 16-bit integer; these are referred to these as the "AS part" and the "tag part", respectively. The specification explicitly reserves communities with AS part 0 or 65535 for use as well-known communities or other future uses.

Accordingly, GateD ordinarily does not permit use of 0 or 65535 in the AS part of a community. However, some BGP implementations have been found to allow this behavior. In order to permit interoperation with such implementations, **allow bad community** can be used within the BGP clause. The preferred solution is to configure the other routers in use to conform to RFC 1997 by using a valid AS number (normally, an AS number assigned to you by the InterNIC) in the AS part.

On a router that configures communities as 32-bit integers rather than as an AS part and a tag part, the reserved communities are 0 through 65535 and 4294901760 through 4294967295. Use of these communities should be avoided. Any other community (65536 through 4294901759) is legal, although it is advisable to use one's own AS number in the AS part (for example, communities 80871424 through 80936959 have AS 1234 in the AS part).

defaultmetric *metric*

Defines the metric used when advertising routes via BGP. If not specified, no metric is propagated. This metric can be overridden by a metric specified on the neighbor or group statements or in export policy.

traceoptions *trace_options*

Specifies the tracing options for BGP. By default, these are inherited from the global trace options. These values can be overridden on a group or neighbor basis. (See the Trace Statement and the BGP-specific tracing options for more information.)

clusterid *host*

Specifies the route reflection cluster ID for BGP. The cluster ID defaults to be the same as the router ID. If a router is to be a route reflector, then a single cluster ID should be selected and configured on all route reflectors in the cluster. The only constraints on the choice of cluster ID is that IDs of clusters within an AS must be unique within that AS, and the cluster ID must not be 0.0.0.0. Choosing the cluster ID to be the router ID of one router in the cluster always fulfills these criteria. If there is only one route reflector in the cluster, the **clusterid** setting can be omitted, as the default is sufficient.

disable export best

Disables the configurable export-best BGP (CEBB) functionality, which is on by default. CEBB allows GateD to export BGP routes that are not actually installed because there is an IGP route that has taken precedence over the BGP route.

Group parameters

The BGP statement has **group** clauses and **peer** subclauses. A **group** clause usually defines default parameters for a group of peers. These parameters apply to all subsidiary **peer** subclauses. Any number of groups can be specified, but each must have a unique combination of **type**, **peeras**, and **nexthopself** options.

BGP peers are grouped by type and the autonomous system of the peers. Any number of **peer** subclauses can be specified within a group. Some of the parameters from the **peer** subclause can be specified on the **group** clause to provide defaults for the whole group (which can be overridden for individual peers).

The general rule for using the **group** and **peer** options is that any option can appear in either a **group** or **peer** statement. Options in the group statement apply to all of the peering sessions for that group. Options in a **peer** subclause have no effect on the other peering sessions. The following constraints apply:

- Do not specify both the **gateway** and **lcladdr** options.
- Do not specify a **subgroup** on a peer; this can only be done on a group basis.
-Do not specify a **holdtime** of less than six seconds.

For groups of **internal**, **routing**, **confed**, and **igp** types, the following constraints apply:

- Do not specify **metricout** on the peer statement.
- Do not specify **outdelay** on the peer statement.
- Do not specify **nexthopself** on the peer statement.
- Do not specify **ascount** on these group types (it applies only to peers in external groups.

Specifying peers

Within a group, BGP peers can be configured in one of two ways. They can be explicitly configured with a **peer** statement, or implicitly configured with the **allow** statement. Both are described as follows:

allow

The **allow** clause allows **peer** connections from any addresses in the specified range of network and mask pairs. All parameters for these peers must be configured on the group clause. The internal peer structures are created when an incoming open request is received

and destroyed when the connection is broken. For more detail on specifying the network/mask pairs, see the section on Route Filtering.

peer *host*

A **peer** clause configures an individual peer. Each peer inherits all parameters specified on a group as defaults. Many defaults can be overridden by parameters explicitly specified on the peer subclause.

Within each **group** clause, individual peers can be specified or a group of *potential* peers can be specified using **allow**. **allow** is used to specify a set of address masks. If GateD receives a BGP connection request from any address in the set specified, it accepts it and set up a peer relationship.

The following **group** types are allowed:

group type external peeras *autonomous_system* [**med**] [**ignorefirstashop**] [**subgroup** *integer*] [**nexthopself**]

In the classic external BGP group, full policy checking is applied to all incoming and outgoing advertisements. The external neighbors must be directly reachable through one of the machine's local interfaces. The nexthop transmitted is computed with respect to the shared interface.

If the **gateway** parameter is present on peer statement, EBGP peers can be on different networks. Refer to the **gateway** keyword statement described later in this BGP section.

med

By default, any metric (Multi_Exit_Disc, or MED) received on a BGP connection is ignored. If it is desired to use MEDs in routing computations, the **med** option must be specified on the group. By default, MEDs are not sent on external connections. To send MEDs, use the **metric** option of the export statement or the **metricout** peer/group parameter.

ignorefirstashop

Routers known as *Route Servers*, are capable of propagating routes without appending their own AS to the AS Path. By default, GateD drops such routes. Specifying **ignorefirstashop** on either the group statement or peer clause disables this feature. This option should only be used if it is positively known that the peer is a route server and not a normal router.

subgroup *integer*

The subgroup variable is to designate different subgroups of BGP peers within the same AS (designated by the group statement). It works in conjunction with the subgroup identifier on the import and/or export policy configuration using the export or import clause. Essentially, the designation of the mathematical group (the group can have at minimum a membership of one element [peer statement]) allows for the extensibility and flexibility for "policy per peer."

nexthopself

Causes the nexthop in route advertisements sent to this peer or group of peers to be set to the address that the peer is using for its peering session with this GRF (even if it is normally possible to send a third-party nexthop). If this option is used, the result can be that inefficient routes are followed, but this can be necessary to deal with broken bridged interconnect media (for example, in cases where the routers on the “shared” medium do not really have full connectivity to each other).

group type routing peeras *autonomous_system* **proto** *proto* [**interface** *interface_list*] [**reflector-client** [**no-client-reflect**]] [**ignorefirstashop**] [**lcladdr** *local_address*] [**outdelay** *time*] [**metricout** *metric*] [**subgroup** *integer*] [**nexthopself**]

An internal group that uses the routes of an interior protocol to resolve forwarding addresses. A **type routing** group propagates external routes between routers that are not directly connected, and computes immediate nexthops for these routes by using the BGP nexthop that arrived with the route as a forwarding address to be resolved via an internal protocol's routing information. In essence, internal BGP is used to carry AS external routes, while the IGP is expected to only carry AS internal routes, and the latter is used to find immediate nexthops for the former.

The *proto* names the interior protocol to be used to resolve BGP route nexthops, and can be the name of one or more IGPs in the configuration, including **static**, **isis**, **ospf**, **direct**, **bgp**, and **rip**. The **all** keyword can also be used, though its use is discouraged. By default, the nexthop in BGP routes advertised to type routing peers are the original nexthop that was specified in the arriving BGP UPDATE packet. The *interface_list* provides a list of interfaces whose routes can be used to resolve BGP nexthops. At a minium, this list should include the interfaces(s) over which the peering sessions(s) may be active.

lcladdr, **outdelay**, and **metricout** must be set in the **group** statement, not on a per-peer basis, for the group types **internal** and **routing**. If these options are set on the **peer** statement, they must equal the values set on the corresponding group statement. **lcladdr** is required if you want to use loopback alias as the `router_id`.

The **reflector-client** option specifies that GateD acts as a route reflector for this group. All routes received from any group member are sent to all other internal neighbors, and all routes received from any other internal neighbors are sent to the reflector clients. Since the route reflector forwards routes in this way, the reflector-client group need not be fully meshed. If the **no-client-reflect** option is specified, routes received from reflector clients are only sent to internal neighbors that are *not* in the same group as the sending reflector client. In this case, the reflector-client group should be fully meshed. In all cases, routes received from normal internal peers are sent to all reflector clients.

It is necessary to export routes from the local AS *into* the local AS when acting as a route reflector. For example, assume that the local AS number is 2. An export statement like the following would suffice to make reflection work correctly:

```
export proto bgp as 2 {
      proto bgp as 2 {all;}; # for reflection
      # other exports
};
```

If the cluster ID is changed and GateD is reconfigured with a SIGHUP, all BGP sessions with reflector clients are dropped and restarted.

subgroup *integer*

The subgroup variable is to designate different subgroups of BGP peers within the same AS (designated by the group statement). It works in conjunction with the subgroup identifier on the import and/or export policy configuration using the export or import clause. Essentially, the designation of the mathematical group (the group can have at minimum a membership of one element [peer statement]) allows for the extensibility and flexibility for "policy per peer."

nexthopself

Causes the nexthop in route advertisements sent to this peer or group of peers to be set to the address that the peer is using for its peering session with this GRF (even if it is normally possible to send a third-party nexthop). If this option is used, the result can be that inefficient routes are followed, but this can be necessary to deal with broken bridged interconnect media (for example, in cases where the routers on the “shared” medium do not really have full connectivity to each other).

group type confed peeras *autonomous_system* **proto** *proto_list* **interface** *interface_list* [**reflector-client** [**no-client-reflect**]] [**ignorefirstashop**] [**lcladdr** *local_address*] [**outdelay** *time*] [**metricout** *metric*] [**subgroup** *integer*] [**nexthopself**]

An internal group that uses the routes of an interior protocol to resolve forwarding addresses. A **type confed** group propagates external routes between routers that are not directly connected, and computes immediate nexthops for these routes by using the BGP nexthop, that arrived with the route as a forwarding address to be resolved via an internal protocol's routing information. In essence, internal BGP is used to carry AS external routes, while the IGP is expected to only carry AS internal routes, and the latter is used to find immediate nexthops for the former.

The *proto_list* names the interior protocol to be used to resolve BGP route nexthops, and can be the name of one or more IGPs in the configuration, including **static**, **ISIS**, **OSPF**, **direct**, **BGP**, and **RIP**. **All** can also be used, though its use is discouraged. By default, the nexthop in BGP routes advertised to type routing peers are the original nexthop that was specified in the arriving BGP UPDATE packet. The *interface_list* provides a list of interfaces whose routes can be used to resolve BGP nexthops. At a minium, this list should include the interfaces(s) over which the peering sessions(s) may be active.

The **lcladdr**, **outdelay**, and **metricout** must be set in the **group** statement, not on a per-peer basis, for the group types **internal** and **routing**. If these options are set on the peer statement, they must equal the values set on the corresponding group statement.

The **reflector-client** option specifies that GateD acts as a route reflector for this group. All routes received from any group member are sent to all other internal neighbors, and all routes received from any other internal neighbors are sent to the reflector clients. Since the route reflector forwards routes in this way, the reflector-client group need not be fully meshed. If the **no-client-reflect** option is specified, routes received from reflector clients are only sent to internal neighbors, that are not in the same group as the sending reflector client. In this case, the reflector-client group should be fully meshed. In all cases, routes received from normal internal peers are sent to all reflector clients. It is necessary to export routes from the local AS into the local AS when acting as a route reflector.

For example, assume that the local AS number is 2. An export statement like the following would suffice to make reflection work correctly:

```
export proto bgp as 2 {
      proto bgp as 2 {all;}; # for reflection
      # other exports
};
```

If the cluster ID is changed and GateD is reconfigured with a SIGHUP, all BGP sessions with reflector clients are dropped and restarted.

subgroup *integer*

The subgroup variable is to designate different subgroups of BGP peers within the same AS (designated by the group statement). It works in conjunction with the subgroup identifier on the import and/or export policy configuration using the export or import clause. Essentially, the designation of the mathematical group (the group can have at minimum a membership of one element [peer statement]) allows for the extensibility and flexibility for "policy per peer."

nexthopself

Causes the nexthop in route advertisements sent to this peer or group of peers to be set to the address that the peer is using for its peering session with this GRF (even if it is normally possible to send a third-party nexthop). If this option is used, the result can be that inefficient routes are followed, but this can be necessary to deal with broken bridged interconnect media (for example, in cases where the routers on the “shared” medium do not really have full connectivity to each other).

group type internal peeras *autonomous_system* [**reflector-client** [**no-client-reflect**]] [**ignorefirstashop**] [**lcladdr** *local_address*] [**outdelay** *time*] [**metricout** *metric*] [**subgroup** *integer*] [**nexthopself**]

An internal group operating where there is no IP-level IGP (for example, an SMDS network or MILNET). All neighbors in this group are required to be directly reachable via a single interface. All next-hop information is computed with respect to this interface. Import and export policy can be applied to group advertisements. Routes received from external BGP neighbors are by default readvertised with the received metric.

The **lcladdr**, **outdelay**, **metricout**, **reflector-client** and **no-client-reflect** options are described under the **routing** group description. **lcladdr** is required if you want to use loopback alias as **router_id**.

subgroup *integer*

The subgroup variable is to designate different subgroups of BGP peers within the same AS (designated by the group statement). It works in conjunction with the subgroup identifier on the import and/or export policy configuration using the export or import clause. Essentially, the designation of the mathematical group (the group can have at minimum a membership of one element [peer statement]) allows for the extensibility and flexibility for "policy per peer."

nexthopself

Causes the nexthop in route advertisements sent to this peer or group of peers to be set to the address that the peer is using for its peering session with this GRF (even if it is normally possible to send a third-party nexthop). If this option is used, the result can be that inefficient routes are followed, but this can be necessary to deal with broken bridged interconnect media (for example, in cases where the routers on the “shared” medium do not really have full connectivity to each other).

group type test peeras *autonomous_system*

An extension to external BGP that implements a fixed policy using test peers. Fixed policy and *special case* code make test peers relatively inexpensive to maintain. Test peers do not need to be on a directly-attached network. If GateD and the peer are on the same (directly-attached) subnet, the advertised nexthop is computed with respect to that network. Otherwise, the nexthop is the local machine's current nexthop. All routing information advertised by and received from a test peer is discarded, and all BGP-advertiseable routes are sent back to the test peer. Metrics from BGP-derived routes are forwarded in the advertisement; otherwise, no metric is included.

Peer parameters

The BGP peer subclause allows the following parameters, that can also be specified on the **group** clause. All are optional.

metricout *metric*

If specified, this metric can be used on all routes sent to the specified peer(s). The metric hierarchy is as follows, starting from the most preferred:
1) Metric specified by export policy
2) Peer-level metricout
3) Group-level metricout
4) Default metric

For group types **internal** and **routing**, set this option on the **group** clause instead of on the **peer** clause.

localas *autonomous_system*

Identifies the autonomous system that GateD is representing to this group of peers. The default is that which has been set globally in the **autonomoussystem** statement.

nogendefault

Prevents GateD from generating a default route when BGP receives a valid update from its neighbor. The default route is only generated when the **gendefault** option is enabled.

ignorefirstashop

Disable dropping of routes from peers that do not insert their own AS number into the AS Path. This option should only be used if it is positively known that the peer is a route server and not a normal router.

gateway *gateway*

If a network is not shared with a peer, **gateway** specifies a router on an attached network to be used as the nexthop router for routes received from this neighbor. The **gateway** parameter can also be used to specify a nexthop for peers that are on shared networks. This can be used to ensure that third-party nexthops are never accepted from a given peer, by specifying that peer's address as its own gateway. This parameter is not needed in most cases.

precedence *precedence*

Specifies the precedence of routes learned from this particular BGP peer. This precedence overrides the precedence set on the BGP statement. This precedence can be overridden by the import policy. This keyword can also be stated as **protocol-precedence** or **proto-precedence**.

preference *preference*

Specifies the preference used for routes learned from these peers. This can differ from the default BGP preference set in the **bgp** statement, so that GateD can prefer routes from one peer, or group of peers, over others. This preference can be explicitly overridden by import policy. This keyword can also be stated as **pref**.

lcladdr *local_address*

Specifies the address to be used on the local end of the TCP connection with the peer. For *external* peers the local address must be on an interface that is shared with the peer or with the peer's *gateway* when the **gateway** parameter is used. A session with an external peer

can only be opened when an interface with the appropriate local address (through which the peer or gateway address is directly reachable) is operating. For other types of peers, a peer session is maintained when any interface with the specified local address is operating. In either case, incoming connections are only recognized as matching a configured peer if they are addressed to the configured local address. For group types **internal** and **routing**, set this option on the **group** clause. For group type **routing**, it is advisable to set the **lcladdr** to a non-physical interface, such as a *loopback* interface.

holdtime *time*

Specifies the BGP holdtime value, in seconds, to use when negotiating the connection with this peer. If GateD does not receive a keepalive, update, or notification message within the number of seconds specified in the Hold Time field of the BGP Open message, then the BGP connection is closed. The value must be at least 3 minutes (that is, 180 seconds).

version *version*

Specifies the version of the BGP protocol to use with this peer. If not specified, the highest supported version is used first and version negotiation is attempted. If it is specified, only the specified version is offered during negotiation. Currently supported versions are 2, 3 and 4.

passive

If this option is used, GateD never tries to open a BGP connection with this peer (or group). Instead, it waits for the peer to initiate a connection. This option was introduced to handle a problem in BGP3 and earlier, where two peers can both attempt to initiate a connection at the same time. This problem is fixed in the BGP4 protocol, so **passive** is not needed with BGP4 sessions. If it is applied to both sides of a peering session, **passive** prevents the session from ever being established. For this reason, and because it is generally not needed, the use of **passive** is discouraged.

sendbuffer *buffer_size*
recvbuffer *buffer_size*

Controls the amount of send and receive buffering asked of the kernel. The maximum supported is 65535 bytes, although many kernels have a lower limit. By default, GateD configures the maximum supported. These parameters are not needed on normally-functioning systems.

outdelay *time*

Used to dampen route fluctuations. **outdelay** is the number of seconds a route must be present in the GateD routing database before it is exported to BGP. The default value for each is 0, meaning that these features are disabled. For group types Internal and Routing, set this option on the **group** clause.

keep all

Used to retain routes learned from a peer even if the routes' AS paths contain one of our exported AS numbers.

show-warnings

Causes GateD to issue warning messages when receiving questionable BGP updates such as duplicate routes and/or deletions of non-existing routes. Normally these events are silently ignored.

noaggregatorid

Causes GateD to specify the routerid in the aggregator attribute as zero (instead of its routerid) in order to prevent different routers in an AS from creating aggregate routes with different AS paths.

keepalivesalways

Causes GateD to always send keepalives, even when an update could have correctly substituted for one. This allows interoperability with routers that do not completely obey the protocol specifications on this point.

v3asloopokay

By default GateD does not advertise routes whose AS path is looped (that is, with an AS appearing more than once in the path) to version 3 external peers. Setting this flag removes this constraint. Ignored when set on internal groups or peers.

nov4asloop

Prevents routes with looped AS paths from being advertised to version 4 external peers. This can be useful to avoid advertising such routes to peer that would incorrectly forward the routes on to version 3 neighbors.

ascount *count*

count is the number of times we insert our own AS number when we send the AS path to an external neighbor. Legal values are 1..25, inclusive, the default is 1.

Higher values are typically used to bias upstream neighbors' route selection. (All things being equal, most routers prefer to use routes with shorter AS Paths. Using **ascount**, the AS Path we send can be artificially lengthened.) Note that **ascount** supersedes the **nov4asloop** option -- regardless of whether **nov4asloop** is set, we still send multiple copies of our own AS if the ascount option is set to something greater than one. Also, note that if the value of ascount is changed and GateD is reconfigured, routes are **not** sent to reflect the new setting. If this is desired, it is necessary to restart the peer session (by commenting out the peer or group, reconfiguring, and then uncommenting and reconfiguring again, or by restarting GateD).

throttle *count*

Limit the number of UPDATE packets to *count* per second. It was found that non-GRF routers, when heavily overburdened; would "timeout" BGP peering sessions because they could not keep up with the processing of packets forwarded by the GRF.

allow bad routerid

The BGP specification specifically requires that a BGP router choose a reasonable value (one of its IP addresses) as its router ID. If a BGP OPEN message is received with an unreasonable router ID, the specification requires that an error message be sent, and the BGP connection closed (see RFC 1771, section 6.2). GateD obeys this requirement.

Apparently, some non-GateD BGP implementations sometimes decide to send 0.0.0.0 as their router ID. This is bad, not only because it violates the BGP specification, but because some elements of the BGP protocol assume that all router IDs are globally unique. Proper router ID assignment assures this, but if two routers in the same AS go insane and are allowed to send router ID 0.0.0.0, routing problems results, especially if route reflection is in use. In particular, two routers using identical router IDs (whether 0.0.0.0 or otherwise) do not have each other's routes reflected to them.

Despite this, GateD provides a way to disable router ID sanity checking, through the use of the **allow bad routerid** peer option. Use of this option is STRONGLY

DISCOURAGED, and is provided *only* as a means of allowing interoperation with a faulty router for a short period of time until the faulty router can be fixed or replaced.

logupdown

Causes a message to be logged via the **syslog** mechanism whenever a BGP peer enters or leaves the **ESTABLISHED** state.

ttl *ttl*

By default, GateD sets the IP TTL for local peers to *one* and the TTL for non-local peers to 255. This option mainly is provided when attempting to communicate with improperly functioning routers that ignore packets sent with a TTL of one. Not all kernels allow the TTL to be specified for TCP connections.

traceoptions *trace_options*

Specifies the tracing options for this BGP neighbor. By default these are inherited from group or BGP global trace options. (See the Trace Statements subsection and the BGP-specific tracing options below.)

nexthopself

Causes the nexthop in route advertisements sent to this peer or group of peers to be set to the address that the peer is using for its peering session with this GRF (even if it would normally be possible to send a third-party nexthop). If this option is used, the result can be that inefficient routes are followed, but this can be necessary to deal with broken bridged interconnect media (for example, in cases where the routers on the "shared" medium do not really have full connectivity to each other.

Note: You can only specify **nexthopself** for **group type external** peers.

Tracing options

Note that the **state** option works with BGP, but does not provide true state transition information.

Packet tracing options (that can be modified with **detail**, **send**, and **recv**) are as follows:

packets

All BGP packets.

open

BGP **OPEN** packets that are used to establish a peer relationship.

update

BGP **UPDATE** packets that are used to pass network reachability information.

keepalive

BGP **KEEPALIVE** packets that are used to verify peer reachability.

all

Additional useful information, including additions/changes/deletions to the GateD routing table.

Limitations

If an interface goes down and comes back up and there is policy associated with that interface, GateD must be restarted because the interface policy is not recovered.

The policy can be recovered by manually executing the **gdc refconfig** command, which rereads the **/etc/gated.conf** file and does not stop or restart GateD.

Weighted route dampening statement

Currently, only routes learned via BGP are subject to weighted route dampening although no protocols announce suppressed routes.

Note: The weighted route dampening configuration statement is not within the BGP statement, but is a separate and distinct configuration conceptually, much like interface or kernel statements.

The syntax for weighted route damping in GateD is:

```
dampen-flap {
[suppress-above metric ;
reuse-below metric ;
max-flap metric ;
unreach-decaytime ;
reach-decaytime ;
keep-historytime ; ]
};
```

suppress-above *metric*

This is the value of the instability metric at which route suppression takes place. A route is not installed in the forwarding information base (FIB), or announced even if it is reachable during the period that it is suppressed.

reuse-below *metric*

This is the value of the instability metric at which a suppressed route becomes unsuppressed if it is reachable but currently suppressed. The value assigned to **reuse-below** must be less than **suppress-above.**

max-flap *metric*

This is the upper limit of the instability metric. This value must be greater than the larger of 1 and **suppress_above**.

reach-decay *time*

This value specifies the time desired for the instability metric value to reach one-half of its current value when the route is reachable. This half-life value determines the rate at which the metric value is decayed. A smaller half-life value makes a suppressed route reusable sooner than a larger value. Specify this value in seconds.

unreach-decay *time*

This value acts the same as **reach-decay** except that it specifies the rate at which the instability metric is decayed when a route is unreachable. It should have a value greater than or equal to **reach-decay**. Specify this value in seconds.

keep-history *time*

This value specifies the period over which the route flapping history is to be maintained for a given route. The size of the configuration arrays described below is directly affected by this value. Specify this value in seconds.

When only **dampen-flap {};** is specified in the route dampening statement, then the following default values are currently used:

– **suppress-above** = 3.0;

- **reuse-below** = 2.0;
- **max-flap** = 16.0;
- **unreach-decay** = 900;
- **reach-decay** = 300;
- **keep-history** = 1800;

The ICMP statement

On systems without the BSD routing socket, GateD listens to ICMP messages received by the system. Currently, GateD does processing only on ICMP redirect packets; more functionality can be added in the future, such as support for the router discovery messages. Processing of ICMP redirect messages is handled by the **redirect** statement.

Currently, the only reason to specify the **icmp** statement is to be able to trace the ICMP messages that GateD receives.

ICMP Statement

```
icmp {
    traceoptions trace_options ;
}
```

traceoptions *trace_options*

Specifies the tracing options for ICMP. (See the Trace Statements subsection and the ICMP-specific tracing options for more information.)

Tracing options

Packet tracing options (that can be modified with **detail** and **recv**) are as follows:

packets

All ICMP packets received.

redirect

Only ICMP redirect packets received.

routerdiscovery

Only ICMP Router Discovery packets received.

info

Only ICMP informational packets, which include mask request/response, info request/response, echo request/response and time stamp request/response.

error

Only ICMP error packets, which include time exceeded, parameter problem, unreachable and source quench.

The Router Discovery Protocol

The Router Discovery Protocol is an IETF standard protocol, RFC 1256, used to inform hosts of the existence of routers. It is intended to be used instead of having hosts *wiretap* routing protocols such as RIP. It is used in place of, or in addition to, statically-configured default routes in hosts.

The protocol is split into two portions, the *server* portion that runs on routers, and the *client* portion that runs on hosts. GateD treats these much like two separate protocols, only one of that can be enabled at a time.

The Router Discovery server

The Router Discovery Server runs on routers and announces their existence to hosts. It does this by periodically multicasting or broadcasting a **Router Advertisement** to each interface on that it is enabled. These Router Advertisements contain a list of all the router's addresses on a given interface and the preference of each address for use as the default router on that interface.

Initially these Router Advertisements occur every few seconds, then fall back to every few minutes. In addition, a host can send a **Router Solicitation** to which the router will respond with a unicast Router Advertisement (unless a multicast or broadcast advertisement is due momentarily).

Each Router Advertisement contains an *Advertisement Lifetime* field indicating for how long the advertised addresses are valid. This lifetime is configured such that another Router Advertisement are sent before the lifetime has expired. A lifetime of zero is used to indicate that one or more addresses are no longer valid.

On systems supporting IP multicasting, the Router Advertisements are, by default, sent to the all-hosts multicast address 224.0.0.1. However, the use of broadcast can be specified. When Router Advertisements are being sent to the all-hosts multicast address, or an interface is configured for the limited-broadcast address 255.255.255.255, all IP addresses configured on the physical interface are included in the Router Advertisement. When the Router Advertisements are being sent to a net or subnet broadcast, only the address associated with that net or subnet is included.

The Router Discovery client

A host listens for Router Advertisements via the all-hosts multicast address (224.0.0.1), if IP multicasting is available and enabled, or on the interface's broadcast address. When starting up, or when reconfigured, a host can send a few Router Solicitations to the all-routers multicast address, 224.0.0.2, or the interface's broadcast address.

When a Router Advertisement with non-zero lifetime is received, the host installs a default route to each of the advertised addresses. If the preference is **ineligible**, or the address is not on an attached interface, the route is marked unusable but retained. If the preference is usable, the metric is set as a function of the preference such that the route with the best preference is used. If more than one address with the same preference is received, the one with the lowest IP address is used. These default routes are not exportable to other protocols.

When a Router Advertisement with a zero lifetime is received, the host deletes all routes with next-hop addresses learned from that router. In addition, any routers learned from ICMP

redirects pointing to these addresses are deleted. The same happens when a Router Advertisement is not received to refresh these routes before the lifetime expires.

The Router Discovery server statement

```
routerdiscovery server ( yes | no | on | off ) [ {
    traceoptions trace_options ;
    interface interface_list
        [ minadvinterval time ] |
        [ maxadvinterval time ] |
        [ lifetime time ]
        ;
    address interface_list
        [ advertise ] | [ ignore ] |
        [ broadcast ] | [ multicast ] |
        [ ineligible ] | [ preference preference ]
        ;
} ] ;
```

traceoptions *trace_options*

Specifies the Router Discovery tracing options. (See the Trace Statements and the Router Discovery specific tracing options for more information.)

interface *interface_list*

Specifies the parameters that apply to physical interfaces. Note a slight difference in convention from the rest of GateD; **interface** specifies a list of physical interfaces (such as **le0**, **ef0**, and **en1**), while **address** specifies a list of IP addresses. One or more of the following parameters must be provided for the interface:

maxadvinterval *time*

The maximum time allowed between sending broadcast or multicast Router Advertisements from the interface. Must be no less than 4 seconds and no more than 30 minutes (or 1800 seconds). The default is 10 minutes (or 600 seconds).

minadvinterval *time*

The minimum time allowed between sending unsolicited broadcast or multicast Router Advertisements from the interface. Must be no less than 3 seconds and no greater than **maxadvinterval**. The default is **0.75 * maxadvinterval**.

lifetime *time*

The lifetime of addresses in a Router Advertisement. Must be no less than **maxadvinterval** and no greater than 2:30:00 (two hours, thirty minutes or 9000 seconds). The default is **3 * maxadvinterval**.

address *interface_list*

Specifies the parameters that apply to the specified set of addresses on this physical interface. Note a slight difference in convention from the rest of GateD; **interface** specifies a list of physical interfaces (such as **le0**, **ef0,** and **en1**), while **address** specifies a list of IP addresses.

One or more of the following parameters must be provided for the address:

advertise

Specifies that the specified address(es) should be included in Router Advertisements. This is the default.

ignore

Specifies that the specified address(es) should not be included in Router Advertisements.

broadcast

Specifies that the given address(es) should be included in a broadcast Router Advertisement because this system does not support IP multicasting, or some hosts on attached network do not support IP multicasting. It is possible to mix addresses on a physical interface such that some are included in a broadcast Router Advertisement and some are included in a multicast Router Advertisement. This is the default if the router does not support IP multicasting.

multicast

Specifies that the given address(es) should only be included in a multicast Router Advertisement. If the system does not support IP multicasting the address(es) are not be included. If the system supports IP multicasting, the default is to include the address(es) in a multicast Router Advertisement if the given interface supports IP multicasting. If not, the address(es) are included in a broadcast Router Advertisement.

preference *preference*

The degree of preference of the address(es) as a default router address, relative to other router addresses on the same subnet. This is a 32-bit, signed, two's-complement integer, with higher values meaning more preferable. Note that hex **`80000000`** can only be specified as **ineligible**. The default is **0**.

ineligible

Specifies that the given address(es) are assigned a preference of hex **`80000000`** that means that it is not eligible to be the default route for any hosts.

This is useful when the address(es) should not be used as a default route, but are given as the nexthop in an ICMP redirect. This allows the hosts to verify that the given addresses are up and available.

The Router Discovery client statement

```
routerdiscovery client ( yes | no | on | off ) [ {
     traceoptions trace_options ;
     precedence precedence ;
     interface interface_list
          [ enable ] | [ disable ]
          [ multicast ] | [ broadcast ]
          [ quiet ] | [ solicit ]
          ;
} ] ;
```

traceoptions *trace_options*

Specifies the tracing options for Router Discovery. (See the Trace Statements and the Router Discovery specific tracing options for more information.)

precedence *precedence*

Specifies the precedence of all Router Discovery default routes. The default is **55**.

interface *interface_list*

Specifies the parameters that apply to physical interfaces. Note a slight difference in convention from the rest of GateD; **interface** specifies just physical interfaces (such as **le0**, **ef0**, and **en1**). The Router Discovery Client has no parameters that apply only to interface addresses.

enable

Specifies that Router Discovery should be performed on the specified interface(s). This is the default.

disable

Specifies that Router Discovery should not be performed on the specified interface(s).

multicast

Specifies that Router Solicitations should be multicast on the specified interface(s). If IP multicast is not available on this host and interface, no solicitation is performed. The default is to multicast Router Solicitations if the host and interface support it; otherwise Router Solicitations are broadcast.

broadcast

Specifies that Router Solicitations should be broadcast on the specified interface(s) because the interface or remote system does not support IP multicasting. This is the default if the local interface does not support IP multicasting.

quiet

Specifies that no Router Solicitations are sent on this interface, even though Router Discovery is performed.

solicit

Specifies that initial Router Solicitations is sent on this interface. This is default.

Tracing options

The Router Discovery Client and Server support the **state** trace flag, that traces various protocol occurrences.

state

State transitions.

The Router Discovery Client and Server do not directly support any packet tracing options. Tracing of router discovery packets is enabled via the ICMP statement.

The kernel statement

While the kernel interface is not technically a routing protocol, it has many characteristics of one, and GateD handles it similarly to one. The routes GateD chooses to install in the kernel forwarding table are those that are actually used by the kernel to forward packets.

The add, delete and change operations that GateD must use to update the typical kernel forwarding table take a non-trivial amount of time. This does not present a problem for older routing protocols (for example, RIP), that are not particularly time critical and do not easily handle very large numbers of routes anyway. The newer routing protocols (for example, OSPF, BGP) have stricter timing requirements and are often used to process many more routes. The speed of the kernel interface becomes critical when these protocols are used.

To prevent GateD from locking up for significant periods of time installing large numbers of routes (up to a minute or more has been observed on real networks), the processing of these routes is now done in batches. The size of these batches can be controlled by the tuning parameters described in the following subsection, but normally the default parameters provide the proper functionality.

During normal shutdown processing, GateD normally deletes all the routes it has installed in the kernel forwarding table, except for those marked with **retain**. Optionally, GateD can leave all routes in the kernel forwarding table by not deleting any routes. In this case, changes are made to ensure that routes with a **retain** indication are installed in the table. This is useful on systems with large numbers of routes, as it prevents the need to re-install the routes when GateD restarts. This can greatly reduce the time it takes to recover from a restart.

Forwarding tables and routing tables

The table in the kernel that controls the forwarding of packets is a forwarding table, also known in ISO terminology as a forwarding information base (FIB). The table that GateD uses internally to store routing information it learns from routing protocols is a routing table, known in ISO terminology as a routing information base (RIB). The routing table is used to collect and store routes from various protocols. For each unique combination of network and mask, an active route is chosen. This route is the one with the best (numerically smallest) preference and precedence. All the active routes are installed in the kernel forwarding table. The entries in this table are what the kernel actually uses to forward packets.

Kernel statement

```
kernel {
    options
        [ nochange ]
        [ noflushatexit ]
        ;
    routes number ;
    flash
        [ limit number ]
        [ type interface | interior | all ]
        ;
    background
        [ limit number ]
        [ priority flash | higher | lower ]
        ;
    traceoptions trace_options ;
} ;
```

options *option_list*

Configure kernel options. The valid options are as follows:

nochange

On systems supporting the routing socket this ensures that changes operations are not performed, only deletes and adds. This is useful on early versions of the routing socket code where the change operation was broken.

noflushatexit

During normal shutdown processing, GateD deletes all routes from the kernel forwarding table that do not have a **retain** indication. The **noflushatexit** option prevents route deletions at shutdown. Instead, routes are changed and added to make sure that all the routes marked with **retain** get installed.

This is useful on systems with thousands of routes. Upon start-up, GateD notices which routes are in the kernel forwarding table and does not add them back.

routes *number*

On some systems kernel memory is at a premium. With this parameter a limit can be placed on the maximum number of routes GateD installs in the kernel. Normally, GateD adds/changes/deletes routes in interface/internal/external order. That is, it queues interface routes first, followed by internal routes, followed by external routes, and processes the queue from the beginning. If a this parameter is specified and the limit is hit, GateD does two scans of the list instead. On the first scan it does deletes, and also deletes all changed routes, turning the queued changes into adds. It then rescans the list doing adds in interface/internal/external order until it hits the limit again. This tends to favor internal routes over external routes. The default is not to limit the number of routes in the kernel forwarding table.

flash

When routes change, the process of notifying the protocols is called a flash update. The kernel forwarding table interface is the first to be notified. Normally, a maximum of 20 interface routes can be processed during one flash update. The **flash** command allows tuning of the following parameters.

limit *number*

Specifies the maximum number of routes that can be processed during one flash update. The default is **20**. A value of **-l** causes all pending route changes of the specified type to be processed during the flash update.

type interface | **interior** | **all**

Specifies the type of routes that are processed during a flash update. The **interior** keyword specifies that interior routes are also installed. See the definition of interior gateway protocol for more information. The **all** keyword specifies the inclusion of exterior routes as well. See the definition of exterior gateway protocols for more information. The default is **interface,** that specifies that only interface routes are installed during a flash update.

Specifying **flash limit -1 all** causes all routes to be installed during the flash update; this mimics the behavior of previous versions of GateD.

background

Since only interface routes are normally installed during a flash update, the remaining routes are processed in batches in the background; that is, when no routing protocol traffic is being received. Normally, 120 routes are installed at a time to allow other tasks to be performed and this background processing is done at lower priority than flash updates the following parameters allow tuning of these parameters:

limit *number*

Specifies the number of route that can be processed at during one batch. The default is 120.

priority flash | **higher** | **lower**

Specifies the priority of the processing of batches of kernel updates in relationship to the flash update processing. The default is **lower,** which means that flash updates are processed first. To process kernel updates at the same priority as flash updates, specify **flash**; to process them at a lower priority, use **lower**.

Tracing options

While the kernel interface is not technically a routing protocol, in many cases it is handled as one. The following two options make sense when entered from the command line since the code that uses them is executed before the trace file is parsed:

symbols

Symbols read from the kernel, by **nlist( )** or similar interface.

iflist

Interface list scan. This option is useful when entered from the command line as the first interface list scan is performed before the configuration file is parsed.

The following tracing options can only be specified in the configuration file. They are not valid from the command line:

remnants

Routes read from the kernel when GateD starts.

request

Requests by GateD to Add/Delete/Change routes in the kernel forwarding table.

The following general option and packet-tracing options only apply on systems that use the routing socket to exchange routing information with the kernel. They do not apply on systems that use the old BSD4.3 **ioctl ( )** interface to the kernel.

info

Informational messages received from the routing socket, such as TCP loss, routing lookup failure, and route resolution requests. GateD does not currently do processing on these messages; rather it just logs the information if requested.

Packet tracing options (that can be modified with **detail**, **send** and **recv**):

routes

Routes exchanged with the kernel, including Add/Delete/Change messages and Add/Delete/ Change messages received from other processes.

redirect

Redirect messages received from the kernel.

interface

Interface status messages received from the kernel. These are only supported on systems with networking code derived from BSD 4.4.

other

Other messages received from the kernel, including those mentioned in the **info** keyword explanation previously.

Static statements

Static statements define the static routes used by GateD. A single **static** statement can specify any number of routes. The **static** statements occur after protocol statements and before control statements in the **/etc/gated.conf** file. Any number of **static** statements can be specified, each containing any number of static route definitions. These routes can be overridden by routes with better precedence values.

```
static {
    ( host host ) | default |
    ( network [ ( mask mask ) | ( ( masklen | / ) number ) ] )
        gateway gateway_list
        [ interface interface_list ]
        [ precedence precedence ]
        [ retain ]
        [ reject ]
        [ blackhole ]
        [ noinstall ] ;
    ( network [ ( mask mask ) | ( ( masklen | / ) number ) ] )
        interface interface
        [ precedence precedence ]
        [ retain ]
        [ reject ]
        [ blackhole ]
        [ noinstall ] ;
} ;
```

host *host* **gateway** *gateway_list*
(*network* [(**mask** *mask*) | ((**masklen** | /) *number*)])
default gateway *gateway_list*

This is the most general form of the static statement. It defines a static route through one or more gateways. Static routes are installed when one or more of the **gateways** listed are available on directly attached interfaces. If more than one eligible gateway is available, these are limited by the number of multipath destinations supported.

Parameters for static routes are:

interface *interface_list*

When this parameter is specified, gateways are only considered valid when they are on one of these interfaces. See the section on interface list specification for the description of the *interface_list*.

precedence *precedence*

This option selects the precedence of this static route. The precedence controls how this route competes with routes from other protocols. The default precedence is 60. This keyword can also be stated as **protocol-precedence** or **proto-precedence**.

retain

Normally GateD removes all routes except interface routes from the kernel forwarding table during a graceful shutdown. The **retain** option can be used to prevent specific static routes from being removed. This is useful to insure that some routing is available when GateD is not running.

reject

Instead of forwarding a packet like a normal route, **reject** routes cause packets to be dropped and **unreachable** messages to be sent to the packet originators. Specifying this option causes this route to be installed as a reject route. Not all kernel forwarding engines support reject routes.

blackhole

A **blackhole** route is the same as a **reject** route except that **unreachable** messages are not supported.

noinstall

Normally the route with the lowest precedence is installed in the kernel forwarding table, and is the route exported to other protocols. When **noinstall** is specified on a route, it is not installed in the kernel forwarding table even if it has the lowest precedence, and is therefore active.

Even if a route is not installed, it is available for other uses by GateD. In particular, the route can be exported to other protocols (for example, OSPF ASE) and can be used as a contributing route in the formation of aggregates. Additionally, an active but not installed route can be used to determine if a BGP route's nexthop is reachable, thus permitting the BGP route itself to be installed.

(*network* [(**mask** *mask*) | ((**masklen** | **/**) *number*)]) **interface** *interface*

This form defines a static interface route that is used for primitive support of multiple network addresses on one interface. The **precedence**, **retain**, **reject**, **blackhole** and **noinstall** options are the same as described above.

Control statements overview

Control statements control routes that are imported from routing peers and routes that are exported to these peers. These are the final statements to be included in the **/etc/gated.conf** file.

The control statements are:

- Routing Filtering
- Matching AS Paths
- The Import Statement
- The Export Statement
- The Aggregate Statement
- The Generate Statement

Route filtering

Routes are filtered by specifying configuration language that match a certain set of routes by destination, or by destination and mask. Among other places, route filters are used on **martians**, **import**, and **export** statements.

There is also a way to define filters and lists such that they can be referenced by a "common name" through a **define** filter statement. This method offers the following advantages:

- Declaration of a global filter once and reference by name
- Single parse cost, meaning that the structure for the filter is created once and the use of a pointer to the structure reduces the CPU cycles to parse the filter or list, and reduces the memory requirements if the filter or list is used more than once.

The action taken when no match is found is dependent on the context. For instance, **import** and **export** route filters assume an **all restrict ;** at the end of a list.

A route matches the most specific filter that applies. Specifying more than one filter with the same destination, mask, and modifiers generates an error.

Filtering syntax

network **mask** *mask* [**exact** | **refines** | **between** *number* **and** *number*]

network **masklen** |/ *number* [**exact** | **refines** | **between** *number* **and** *number*]

all

default

host *host*

These are all the possible formats for a route filter. Not all of these formats are available in all places, for instance, the `host` and `default` formats are not valid for `martians`.

In most cases it is possible to specify additional parameters relevant to the context of the filter. For example, on a **martian** statement it is possible to specify the **allow** keyword, on an **import** statement you can specify a preference, and on an **export** statement you can specify a metric.

network **mask** *mask* [**exact** | **refines** | **between** *number* **and** *number*]
network (**masklen** | /) *number* [**exact** | **refines** | **between** *number* **and** *number*]

Matching requires both an address and a mask, except for **all**, **default**, and **host**. These forms vary in how the mask is specified. In the first, the mask is explicitly specified. In the second, the mask is specified by the number of contiguous one bits.

If no additional parameters are specified, any destination that falls in the range given by the network and mask is matched. The mask of the destination is ignored.

The three following optional modifiers cause the mask of the destination to be considered also:

exact

This parameter specifies that the mask of the destination must match the supplied mask *exactly*. This is used to match a network, but no subnets or hosts of that network.

refines

Specifies that the mask of the destination must be more specified (that is, longer) than the filter mask. This is used to match subnets and/or hosts of a network, but not the network.

between *number* **and** *number*

Specifies that the mask of the destination must be as or more specific (as long as or longer) than the lower limit (the first *number* parameter) and no more specific (as long as or shorter) than the upper limit (the second *number* parameter). Note that **exact** and **refines** are both special cases of **between**.

More than one filter can be specified with a given network and mask, as long as the filters differ in their modifiers. In the case of two otherwise identical filters with the between modifier, the ranges given by the filters must not overlap.

If two filters differ only in terms of their modifiers, the filters are matched in the following order: **exact**, **between**, **refines**, no modifier. For example, a filter with **exact** specified is matched before an eligible filter with **between**.

all

This entry matches anything. It is equivalent to the following:

```
0.0.0.0 mask 0.0.0.0
```

default

Matches the **default** route. To match, the address must be the default address and the mask must be all zeros. This is equivalent to the following:

```
0.0.0.0 mask 0.0.0.0 exact
```

host *host*

Matches a specific host route. To match, the address must exactly match the specified *host* and the network mask must be a host mask (that is, all ones). This is equivalent to the following:

```
host mask 255.255.255 exact
```

Matching AS paths

An AS path is a list of autonomous systems that routing information has passed through to get to this router, and an indicator of the origin of this information. This information can be used to prefer one path to a destination network over another. The primary method for doing this with GateD is to specify a list of patterns to be applied to AS paths when importing and exporting routes.

Each autonomous system through which a route passes prepends its AS number to the beginning of the AS path.

The origin information details the completeness of AS path information. An origin of **igp** indicates the route was learned from an interior routing protocol and is most likely complete. An origin of **egp** indicates the route was learned from an exterior routing protocol that does not support AS paths (EGP, for example) and the path is most likely not complete. When the path information is definitely not complete, an origin of **incomplete** is used.

There is also a way to define filters and lists such that they can be referenced by a "common name" through a **define** filter statement. This method offers the following advantages:

- Declaration of a global filter once and reference by name
- Single parse cost

  This means that the structure for the filter is created once, and the use of a pointer to the structure reduces the CPU cycles to parse the filter or list, and reduces the memory requirements if the filter or list is used more than once.

AS path regular expressions are defined in RFC 1164, section 4.2.

AS path matching syntax

An AS path is matched using the following syntax.

aspath *aspath_regexp* **origin any** | ([**igp**] [**egp**] [**incomplete**])

This specifies that an AS matching the *aspath_regexp* with the specified origin is matched.

AS path regular expressions

Technically, an AS path regular expression is a regular expression with the alphabet being the set of AS numbers. An AS path regular expression is composed of one or more AS-path expressions. An AS path expression is composed of AS path terms and AS path operators.

AS path terms

An AS path term is one of the following three objects:

```
autonomous_system
.
( aspath_regexp )
```

autonomous_system

Is any valid autonomous system number, from one through 65534 inclusive.

.

Matches any autonomous system number.

(*aspath_regexp*)

Parentheses group subexpressions -- an operator such as * or ? works on a single element, or on a regular expression enclosed in parentheses

AS path operators

Note: Under release 1.4.6, the aspath operator had to be included within double quotes. The double quotes are no longer needed with this release, although the use of them is backwards compatible (that is, if you have them in your code, it still parses correctly).

An AS path operator is one of the following:

aspath_term {**m,n**}

aspath_term {**m**}

aspath_term {**m,**}

aspath_term *****

aspath_term **+**

aspath_term **?**

(aspath_term) | (*aspath_term)*

null

[**n1 n2 n3 ...**]

aspath_term {**m,n**}

A regular expression followed by {m,n} (where m and n are both non-negative integers and m <= n) means at least *m* and at most *n* repetitions.

aspath_term {**m**}

A regular expression followed by {m} (where m is a positive integer) means exactly *m* repetitions.

aspath_term {**m,**}

A regular expression followed by {m,} (where m is a positive integer) means *m or more* repetitions.

aspath_term *****

An AS path term followed by * means *zero or more* repetitions. This is shorthand for {0,}.

aspath_term **+**

A regular expression followed by + means *one or more* repetitions. This is shorthand for {1,}.

aspath_term **?**

A regular expression followed by ? means *zero or one* repetition. This is shorthand for {0,1}.

null

Matches all things with a null aspath. Example usage:

import proto bgp aspath (**null**) **origin any** { **all**; };

[**n1 n2 n3 ...**]

Defines a set of numbers defined by n (where n is an AS number) in the brackets, and the input must match one of the numbers in the set. Ranges of numbers cannot be specified.

(aspath_term) | (*aspath_term)*

Matches the AS term on the left, or the AS term on the right. This is identical to a logical OR operation.

Note: Parentheses can be used to improve readability or to group AS path terms into a subexpression. The subexpression can then be used as an aspath term with any of the operators, as shown in the following example

aspath 1 3 4 *

The previous example translates to 1 3 (4*), with the ***** operator applied only to the 4. If you were to place the parentheses around all the numbers (that is, (1 3 4)*, the ***** would apply to the whole subexpression (that is, (1 3 4)).

Examples:

The following examples show how the AS path operators are used with the AS path terms to select an AS path.

Assume the following AS path regular expression is defined:

aspath 6 {2,5}

The **{2,5}** defines the number of times that the aspath number 6 can be repeated. This translates into the following AS path expressions being selected: 6 6, 6 6 6, 6 6 6 6, and 6 6 6 6 6. In other words, the AS number 6 can be repeated at least twice and as many times as five to be selected.

Assume the following AS path regular expression is defined:

aspath 6 {4}

This translates into only the AS path 6 6 6 6 being selected.

Assume the following AS path regular expression is defined:

aspath 4 {3,}

This translates into three or more repetitions of 4 4 4 being selected.

The following AS numbers are used in the next set of examples:

A) 1 2 3 4

B) 5 7 1 5

C) 4 1 3 7

D) 1 3

E) 4 1 3

F) 4 1 3 7 8

G) 3 4 5 1 3

Assume that the following AS path regular expression is defined:

aspath .* 1 3 .*

This is translated as follows: the **.***, that can be no AS number or, one or more AS numbers (that is, zero or more repetitions); followed by 1; followed by a 3; followed by **.*** , that can be no AS number or, one or more AS numbers (that is, with zero or more repetitions). This definition would result in path expression C, D, E, F, and G being selected.

Assume the following AS path regular expression is defined:

aspath .* 1 3 .+

This is translated as follows: the **.***, that can be no AS number or, one or more AS numbers (that is, zero or more repetitions); followed by 1; followed by a 3; followed by **.+** , that can be any AS numbers (that is, one or more repetitions). This definition would result in path expression C and F being selected.

Assume the following AS path regular expression is defined:

aspath .? 1 3 .?

This is translated as follows: the **.?**), which is optional, but can be any AS number; followed by a 1; followed by a 3; followed by **.?**, which is optional, but can be any AS number. This definition would result in path expression C, D, and E being selected. Another way of viewing the use of the **?** is that it makes the AS numbers at the beginning and end optional; combining the **?** with the **.** means that the position is optional, but if the position exists, it can be any AS number.

Assume the following AS path regular expression is defined:

aspath (1 2 3 4) | (1 3)

This is translated as follows: the string of numbers "1 2 3 4 " or "1 3" can be selected. This definition would result in path expressions A and D being selected. Another way of viewing the use of the **|** is that it is simply a logical OR operation.

The **|** operator is the lowest precedence operator; subexpressions are grouped before the **|** operator is applied and subexpressions are grouped from left to right. For example," 7 | 1 6" is equivalent to "(7) | (1 6)". That is, the "1 6" is grouped before the operator is applied. The use of parentheses is recommended to avoid ambugity; such as someone misreading the previous example as "(7 | 1) 6".

Assume the following AS path regular expression is defined:

aspath null

This results in the selection of items with no AS number. The **null** operator is commonly used to match routes that originate in the same AS and thus have no *aspath*. An example aspath

expression is as follows: **import proto bgp aspath (null) origin any {default restrict;};**. This statement translates to mean that if any of my IBGP peers sends a default route via IBGP, then do not install it.

The use of the **null** operator can be combined with other operators, as shown in the following example:

aspath 4 1 3 (7 | null)

This results in path expressions C and E being selected. The following would be equivalent to the previous example:

4 1 3 7?

4 1 3 (7 | null)

((4 1 3 7) | (4 1 3))

4 1 3 [7 null]

The following AS paths are used in the next example:

A) 3 27 54

B) 6 27 54

C) 12 27 54

Assume the following AS path regular expression is defined:

aspath [3 4 5 6] 27 54

The use of the numbers in the brackets means that the first position can be any of the numbers defined in the brackets. This results in path expressions A and B being selected. Note that you cannot define a range of numbers to be used in the brackets. For example, defining the following AS path regular expression would not work:

aspath [3-6] 27 54

AS path attributes

BGP updates carry a number of path attributes. Some of these, like the AS path, are required. Others are optional and may or may not appear in any given BGP update.

The **aspath-opt** option to the group clause, and its variant the **mod-aspath** option, can be used to generate optional path attributes. The **mod-aspath** option is used in exports to set a community value. Currently only the **community** attribute is supported. The **aspath-opt** attribute can also be used on the import clause to allow optional attributes to be considered when determining GateD's preference for the routes in a particular BGP update.

The syntax of **aspath-op**t (and **mod-aspath**) is as follows:

```
aspath-opt {
     [ community autonomous_system : community-id | community-id ]
     [ community no-export | no-advertise | no-export-subconfed | none ]
}
mod-aspath {
     [ community autonomous_system : community-id |community-id ]
     [ community no-export | no-advertise | no-export-subconfed ]
     [ comm-split autonomous_system community-id ]
}
```

Communities

Communities can be specified as an AS and a community ID or as one of the distinguished special communities (with the **community** keyword).

When originating BGP communities, the set of communities that is actually sent is the union of the communities received with the route (if any), those specified in group policy (if any) and those specified in export policy (if any).

When receiving BGP communities, the update is only matched if **all** communities specified in **aspath-opt** are present in the BGP update. (If additional communities are also present in the update, it is still matched.)

There is a limit of 25 communities in any single policy clause. This limit can be increased at compile time by increasing the value of **AS_COMM_MAX**.

community *autonomous_system* : *community_id*

This causes a community "tag" to be added to the transmitted path attributes. The *autonomous_system* part of the community should be set to the local AS unless there is a specific need to do otherwise. This associates an AS with a community.

community *community_id*

If one only specifies the *community_id*, the *autonomous_system* of the advertising router is implicitly prepended.

Note: For backward compatibility, the [**comm-split** *autonomous_system community-id*] is interchangeable.

community no-export

This is a special community that indicates that the routes associated with this attribute must not be advertised outside a BGP AS boundary.

community no-advertise

This is a special community that indicates that the routes associated with this attribute must not be advertised to other BGP peers.

community no-export-subconfed

This is a special community that indicates that the routes associated with this attribute must not be advertised to external BGP peers.

community none

This is not actually a community, but rather a keyword that specifies that a received BGP update is only to be matched if no communities are present. It has no effect when originating communities; therefore can only be used with **aspath-opt**.

The import statement

Importation of routes from routing protocols and installation of the routes in GateD's routing database is controlled by the **import** statement. The format of an **import** statement varies depending on the source protocol.

There is also a way to define filters and lists such that they can be referenced by a "common name" through a **define** filter statement. This method offers the following advantages:

- Declaration of a global filter once and reference by name
- Single parse cost

This means that the structure for the filter is created once, and the use of a pointer to the structure reduces the CPU cycles to parse the filter or list, and reduces the memory requirements if the filter or list is used more than once.

Controlling installation of routes

When importing, the following four keywords can be specified to control how matching routes are compared to determine the route that becomes the active route:

restrict

precedence *precedence*

preference *preference*

localpref *preference*

restrict

Specifies that the routes are not desired in the routing table. In some cases, this means that the routes are not installed in the routing table. In others, it means that they are installed with a negative preference, that prevents them from becoming active, and they are not installed in the forwarding table or exported to other protocols.

precedence *precedence*

Specifies the precedence value used when comparing this route to other routes from other protocols. The route with the lowest precedence available at any given route becomes the active route, is installed in the forwarding table, and is eligible to be exported to other protocols. The precedence value set in the **import** statement overrides any precedence set for individual protocols (for example, RIP, OSPF, BGP, and so on). Also, it is recommended that changes to preference be used before changes to precedence. In other words, if you can resolve the issues using **preference** within a routing protocol, do so. This method is preferred over changing the order of protocols via **precedence**.

preference *preference*

Specifies the preference value used when comparing this route to other routes from the same protocol. The route with the lowest preference available at any given route becomes the active route, is installed in the forwarding table, and is eligible to be exported to other protocols. The preference value set in the import statement overrides any preference set for individual protocols (for example, BGP, RIP, OSPF, and so on).

localpref *preference*

Specifies the local preference value used when comparing this route to other routes known within the BGP protocol. The route with the most preferred local preference available at any given route becomes the active route, is installed in the forwarding table, and is eligible to be exported to other protocols. The default local preference for BGP routes is 100.

Route filters

All the formats allow route filters as shown below. See the section on route filters (page A-86) for a detailed explanation of how they work.

When no route filtering is specified (that is, when **restrict** is specified on the first line of a statement), all routes from the specified source matches that statement. If any filters are specified, only routes that match the specified filters is imported. Put differently, if any filters are specified, an **all restrict** is assumed at the end of the list.

There is also a way to define filters and lists such that they can be referenced by a "common name" through a **define** filter statement. This method offers the following advantages:

- Declaration of a global filter once and reference by name.
- Single parse cost, meaning that the structure for the filter is created once and the use of a pointer to the structure reduces the CPU cycles to parse the filter or list and reduce the memory requirements if the filter or list is used more than once.

network **mask** *mask* [**exact** | **refines** | **between** *number* **and** *number*]
network (**masklen** | /) *number*
[**exact** | **refines** | **between** *number* **and** *number*]
default
host *host*

Importing routes from BGP

import proto bgp as *autonomous_system*

[**subgroup** *integer*] [aspath-opt *community communityid*] **restrict ;**

import proto bgp as *autonomous_system*

[**subgroup** *integer*] [*aspath*-opt *community communityid*]
[**precedence** *precedence*][**preference** *preference*]
[**localpref** *preference*] **{**

route_filter [**restrict** | (**precedence** *precedence*)
| (**preference** *preference*) | (**localpref** *preference*)] **;**

} ;

import proto bgp aspath *aspath_regexp*

origin any | ([**igp**] [**egp**] [**incomplete**])

[*aspath*-opt] **restrict ;**

import proto bgp aspath *aspath_regexp*

origin any | ([**igp**] [**egp**] [**incomplete**])

[aspath-opt] [**precedence** *precedence*] [**preference** *preference*]
[**localpref** *preference*]

route_filter [**restrict** | (**precedence** *precedence*)
| (**preference** *preference*) | (**localpref** *preference*)] **;**

} ;

BGP also supports controlling propagation by the use of AS path regular expressions, that are documented in the section on Matching AS paths (page A-88). Note that BGP versions 2 and 3 only support the propagation of natural networks, so the **host** and **default** route filters are meaningless. BGP version 4 supports the propagation of any destination along with a contiguous network mask.

The **aspath-opt** option allows the specification of import policy based on the path attributes found in the BGP update. If multiple communities are specified in the **aspath-opt** option, only updates carrying all of the specified communities are matched. If none is specified, only updates lacking the community attribute are matched.

It is quite possible for several BGP import clauses to match a given update. If more than one clause matches, the first matching clause is used. All later matching clauses are ignored. For this reason, it is generally desirable to order import clauses from most to least specific. An import clause without an **aspath-opt** option matches any update with any (or no) communities.

BGP stores any routes that were rejected implicitly by not being mentioned in a route filter, or explicitly with the **restrict** keyword in the routing table with a negative preference. A negative preference prevents a route from becoming active, that prevents it from being installed in the forwarding table or exported to other protocols. This alleviates the need to break and reestablish a session upon reconfiguration if import policy is changed.

The **localpref** keyword only has meaning for BGP routes. Unlike the **precedence** and **preference** keywords, that only have meaning internally to the router, **localpref** is propagated along with a BGP route to other routers. Specifying **localpref** on the import statement changes the **localpref** of a BGP route that is received. The default **localpref** for a BGP route is 100. The **localpref** can also be changed when exporting BGP routes as well.

Importing routes from RIP and Redirects

```
import proto rip | redirect
    [ ( interface interface_list ) | (gateway gateway_list ) ]
    restrict ;
import proto rip | redirect
    [ ( interface interface_list ) | (gateway gateway_list ) ] [ precedence
     precedence ] {
    route_filter [ restrict | ( precedence precedence ) ] ;
} ;
```

The importing of RIP and redirect routes can be controlled by protocol, source interface and source gateway. If more than one is specified, they are processed from most general (protocol) to most specific (gateway).

RIP does not support the use of preference to choose between routes of the same protocol. That is left to the protocol metrics. These protocols do not save routes that were rejected because they have short update intervals.

Importing routes from OSPF

```
import proto ospfase [ tag ospf_tag ] restrict ;
import proto ospfase [ tag ospf_tag ]
    [ precedence precedence ] {
    route_filter [ restrict | ( precedence precedence ) ;
} ;
```

Due to the nature of OSPF, only the import of ASE routes can be controlled. OSPF intra-area and inter-area routes are always imported into the GateD routing table with a precedence of 10. If a tag is specified, the import clause only applies to routes with the specified tag.

It is only possible to restrict the importing of OSPF ASE routes when functioning as an AS border router. This is accomplished by specifying an **export ospfase** clause. Specification of an empty export clause can be used to restrict importation of ASEs when no ASEs are being exported.

Like the other interior protocols, preference can not be used to choose between OSPF ASE routes, that is done by the OSPF costs. Routes that are rejected by policy are stored in the table with a negative preference.

The export statement

The **import** statement controls which routes received from other systems are used by GateD, and the **export** statement controls which routes are advertised by GateD to other systems. Like the **import** statement, the syntax of the **export** statement varies slightly per protocol. The syntax of the **export** statement is similar to the syntax of the **import** statement, and the meanings of many of the parameters are identical. The main difference between the two is that while route importation is controlled just by source information, route exportation is controlled by both destination and source information.

The outer portion of a given **export** statement specifies the destination of the routing information you are controlling. The middle portion restricts the sources of importation that you wish to consider. And the innermost portion is a route filter used to select individual routes.

There is also a way to define filters and lists such that they can be referenced by a "common name" through a **define** filter statement. This method offers the following advantages:

- Declaration of a global filter once and reference by name
- Single parse cost

This means that the structure for the filter is only created once, and the use of a pointer to the structure reduces the CPU cycles to parse the filter or list, and reduces the memory requirements if the filter or list is used more than once.

Controlling exportation of routes

The following keywords can be specified to control how (or if) routes are exported. The most specific statement is the one applied to the route being exported (that is, via the **metric** keyword). The values that can be specified for **metric** depend on the destination protocol that is referenced in the following **export** statement:

restrict

metric *metric* | **all metric igp**

localpref *preference*

restrict

Specifies that nothing should be exported. If specified on the destination portion of the **export** statement, it specifies that nothing at all should be exported to this destination. If specified on the source portion, it specifies that nothing from this source should be exported to this destination. If specified as part of a route filter, it specifies that the routes matching that filter should not be exported.

metric *metric*

Specifies the metric to be used when exporting to the specified destination. **all metric igp** is used for exporting (redistributing) the IGP metric as a BGP MED.

Here is an example that exports the OSPF metric as the BGP MED:

```
export proto bgp as foo {
                proto ospf {
                        all metric igp ;
                };
        };
```

Note: If this is used with BGP, a BGP update is issued immediately upon the IGP metric changing.

When specifying **all metric igp**, instabilities within the IGP can cause substantial BGP route flap. This can be minimized by using the **outdelay** variable in the BGP statement.

localpref *preference*

This is used on BGP only for setting the local preference to be exported to other IBGP peers. It can be set anywhere metric can be set in the **export** statement.

In the following example, assume the AS is **bar** (using IBGP); it exports the OSPF routes into BGP with the **localpref** of 150:

```
export proto bgp as bar {
                proto ospf  localpref 150 {
                        all ;
                };
        };
```

Route filters

All the formats allow route filters as shown in the following subsections. See the Routing Filtering section (page A-86) for a detailed explanation of how filters work. When no route filtering is specified (that is, when **restrict** is specified on the first line of a statement), all routes from the specified source match that statement. If any filters are specified, only routes that match the specified filters are exported. Put differently, if any filters are specified, an **all restrict ;** is assumed at the end of the list.

There is also a way to define filters and lists such that they can be referenced by a "common name" through a **define** filter statement. This method offers the following advantages:

- Declaration of a global filter once and reference by name
- Single parse cost

This means that the structure for the filter is created once, and the use of a pointer to the structure reduces the CPU cycles to parse the filter or list, and reduces the memory requirements if the filter or list is used more than once.

network **mask** *mask* [**exact** | **refines** | **between** *number* **and** *number*]
network (**masklen** | /) *number*
[**exact** | **refines** | **between** *number* **and** *number*]
default
host *host*

Specifying the destination

As mentioned previously, the syntax of the **export** statement varies depending on which protocol it is being applied. A requirement that applies in all cases is the specification of a metric. All protocols define a default metric to be used for routes being exported. In most cases, this can be overridden at several levels of the **export** statement.

The specification of the source of the routing information being exported (the *export_list*) is described in the following subsections.

Exporting to BGP

export proto bgp as *autonomous system*
    restrict ;
export proto bgp as *autonomous system* [*mod-aspath community communityid*]
    [**subgroup** *integer*] [**metric** *metric* | **localpref** *preference*] **{**
    export_list **;**
} ;

Exporting to BGP is controlled by autonomous system (AS); the same policy is applied to all routers in the AS or subgroup.

BGP metrics are 16-bit unsigned quantities (that is, they range from 0 to 65535, inclusive, with 0 being the most attractive). While BGP version 4 actually supports 32-bit unsigned quantities, GateD does not yet support this. In BGP version 4, the metric is otherwise known as the multi-exit discriminator (MED).

In BGP, the **mod-aspath** keyword can be used to send the BGP community attribute. Any communities specified with the **mod-aspath** option are sent in addition to any received with the route or specified in the group statement.

If no export policy is specified, only routes to attached interfaces are exported. If any policy is specified, the defaults are overridden; it is necessary to explicitly specify everything that should be exported.

Note that BGP versions 2 and 3 only support the propagation of natural networks, so the **host** and **default** route filters are meaningless. BGP version 4 supports the propagation of any destination along with a contiguous network mask.

The **localpref** keyword only has meaning for BGP routes. Specifying **localpref** on the export statement changes the already existing **localpref** value for a BGP route when it is sent to other networks. The default **localpref** for a BGP route is 100. Note that **localpref** can be changed when importing BGP routes as well. In fact, it is more preferable to change **localpref** when importing rather than exporting BGP routes.

Exporting to RIP

```
export proto rip
    [ ( interface interface_list ) | (gateway gateway_list ) ]
    restrict ;
export proto rip
    [ ( interface interface_list ) | (gateway gateway_list ) ]
    [ metric metric ] {
    export_list ;
} ;
```

Exporting to RIP is controlled by protocol, interface, or gateway. If more than one is specified, they are processed from most general (protocol) to most specific (gateway).

It is not possible to set metrics for exporting RIP routes into RIP. Attempts to do this are silently ignored.

If no export policy is specified, RIP and interface routes are exported into RIP. If any policy is specified, the defaults are overridden. It is necessary to explicitly specify everything that should be exported.

RIP version 1 assumes that all subnets of the shared network have the same subnet mask so they are only able to propagate subnets of that network. RIP version 2 removes that restriction and is capable of propagating all routes when not sending version 1-compatible updates.

Note: To announce routes from other protocols (that is, IP aliases to the loopback interface, static routes, internally generated default routes, and so on) that do not have a **metric** associated with them, via RIP, it is necessary to specify the **metric** at some level in the **export** clause. Just setting a default metric for RIP is not sufficient. This is a safeguard to verify that the announcement is intended.

Exporting to OSPF

```
export proto ospfase [ type 1 | 2 ] [ tag ospf_tag ]
    restrict ;
export proto ospfase [ type 1 | 2 ] [ tag ospf_tag ]
    [ metric metric ] {
    export_list ;
} ;
```

Note: Keep the order of the OSPF export statements as shown; reversing them causes a parser error.

It is not possible to create OSPF intra- or inter-area routes by exporting routes from the GateD routing table into OSPF. It is only possible to export from the GateD routing table into OSPF ASE routes. It is also not possible to control the propagation of OSPF routes within the OSPF protocol.

There are two types of OSPF ASE routes: **type 1** and **type 2**. See the OSPF protocol configuration for a detailed explanation of the two types (page A-36). The default type is specified by the **defaults** subclause of the **ospf** clause. This can be overridden by a specification on the **export** statement.

OSPF ASE routes also have the provision to carry a **tag**. This is an arbitrary 32-bit number that can be used on OSPF routers to filter routing information. See the OSPF protocol configuration for detailed information on OSPF tags (page A-36). The default tag specified by the **ospf** defaults clause can be overridden by a tag specified on the **export** statement.

Specifying the source

The export list specifies export based on the origin of a route; the syntax varies depending on the source.

Exporting BGP routes

```
proto bgp  as  autonomous_system   [ subgroup  integer ]
     restrict  ;
proto bgp as  autonomous_system  [ subgroup  integer ]
     [ metric  metric  | localpref  preference ] {
     route_filter [ restrict | ( metric  metric ) ] ;
} ;
```

BGP routes can be specified by source autonomous system. All routes can be exported by **aspath**. See the section on Exporting by AS Path for more information.

The **localpref** parameter only has meaning for BGP routes. Specifying **localpref** on the export statement changes the already existing **localpref** value for a BGP route when it is sent to other routes. The default **localpref** for a BGP route is 100. Note that the **localpref** can be BGP routes as well. In fact, it is more preferable to change **localpref** when importing rather than exporting BGP routes.

Exporting RIP routes

```
proto rip
    [ ( interface interface_list ) | (gateway gateway_list ) ]
    restrict ;
proto rip
    [ ( interface interface_list ) | (gateway gateway_list ) ]
    [ metric metric ] {
    route_filter [ restrict | ( metric metric ) ] ;
} ;
```

RIP routes can be exported by protocol, source interface and/or source gateway.

Exporting OSPF routes

```
proto ospf  | ospfase restrict ;
proto ospf  | ospfase [ metric metric ] {
    route_filter [ restrict | ( metric metric ) ] ;
} ;
```

Both OSPF and OSPF ASE routes can be exported into other protocols. See the following subsections for information on exporting by **tag**.

Exporting routes from non-routing protocols

Non-routing with interface

```
proto direct | static | kernel
     [ (interface interface_list ) ]
     restrict ;
proto direct | static | kernel
     [ (interface interface_list ) ]
     [ metric metric ] {
     route_filter [ restrict | ( metric metric ) ] ;
} ;
```

These protocols can be exported by protocol, or by the interface of the nexthop. These protocols are:

direct

Routes to directly-attached interfaces.

static

Static routes specified in a **static** clause.

kernel

On systems with the routing socket, routes learned from the routing socket are installed in the GateD routing table with a protocol of **kernel**. These routes can be exported by referencing this protocol. This is useful when it is desirable to have a script install routes with the **route** command and propagate them to other routing protocols.

Non-routing by protocol

```
proto default | aggregate
     restrict ;
proto default | aggregate
     [ metric metric ] {
     route_filter [ restrict | ( metric metric ) ] ;
} ;
```

These protocols can only be referenced by protocol.

default

Refers to routes created by the **gendefault** option. It is recommended that route generation be used instead.

aggregate

Refers to routes synthesized from other routes when the aggregate and generate statements are used. See the section on Route Aggregation (page A-107) for more information.

Exporting by AS path

```
proto bgp | all aspath aspath_regexp
    origin any | ( [ igp ] [ egp ] [ incomplete ] )
    restrict ;
proto bgp | all aspath aspath_regexp
    origin any | ( [ igp ] [ egp ] [ incomplete ] )
    [ metric metric  | localpref preference ] {
    route_filter [ restrict | ( metric metric ) ] ;
} ;
```

When BGP is configured, all routes are assigned an AS path when they are added to the routing table. For all interior routes, this AS path specifies IGP as the origin and no ASes in the AS path (the current AS is added when the route is exported). For BGP routes, the AS path is stored as learned from BGP.

AS path regular expressions are documented in the section on Matching AS Paths (page A-88).

Exporting by route tag

```
proto rip | ospf | all tag tag restrict ;
proto rip | ospf | all tag tag
    [ metric metric ] {
    route_filter [ restrict | ( metric metric ) ] ;
} ;
```

Both OSPF and RIP version 2 currently support **tags**, all other protocols always have a tag of zero. The source of exported routes can be selected based on this tag. This is useful when routes are classified by tag when they are exported into a given routing protocol.

Route aggregation and generation statements

Route aggregation is a method of generating a more general route given the presence of a specific route. It is used, for example, at an autonomous system border to generate a route to a network to be advertised via BGP given the presence of one or more subnets of that network learned via an IGP.

Older versions of GateD automatically performed this function, generating an aggregate route to a natural network (using the old Class A, B and C concept) given an interface to a subnet of that natural network. However, that was not always the correct thing to do, and with the advent of classless inter-domain routing it is even more frequently the wrong thing to do, so aggregation must be explicitly configured. No aggregation is performed unless explicitly requested in an **aggregate** statement.

Route aggregation is also used by regional and national networks to reduce the amount of routing information passed around. With careful allocation of network addresses to clients, regional networks can just announce one route to regional networks instead of hundreds.

Aggregate routes are not actually used for packet forwarding by the originator of the aggregate route, only by the receiver (if it wishes). A router receiving a packet that does not match one of the component routes that led to the generation of an aggregate route is supposed to respond with an ICMP network unreachable message. This is to prevent packets for unknown component routes from following a default route into another network where they would be forwarded back to the border router, and around and around again and again, until their TTL expires. Sending an unreachable message for a missing piece of an aggregate is only possible on systems with support for reject routes.

Route aggregation can be specified via two statements: **aggregate** and **generate**. A route built with the **aggregate** statement and advertised to a peer via BGP is sent with the aggregator ID and with aggregation attributes. A route built with the **generate** statement and advertised to a peer via BGP is sent without the aggregator ID and without aggregation attributes.

The set of routes that form aggregate routes are known as contributing routes. The contributing routes are defined by the **proto** and/or **route_filter** definitions within the **aggregate**/**generate** statement. The contributing routes are ordered according to the precedence and preference that applies to them. If there is more than one contributing route of the same protocol, the route with the lower preference is ordered first. If there are contributing routes from different protocols, those with a lower precedence are ordered first. The resulting aggregate route is assigned the precedence and preference values from the contributing route which is ordered first. It is recommended that preference be used to control the order of the contributing routes that are learned from the same protocol before using precedence to control the order of contributing routes that are learned from different protocols.

A slight variation of aggregation is the generation of a route based on the existence of certain conditions. This is sometimes known as the route of last resort. This route inherits the nexthops and aspath from the contributor specified with the lowest (most favorable) preference. The most common usage for this is to generate a default route based on the presence of a route from a peer on a neighboring backbone.

The use of the **default** parameter is shown in the following example. Assume that the topology of this simple network contains four GRFs and one TNT. GRF3 is advertising route 10.8.17/30 to GRF1, while GRF4 is advertising route 10.8.16/30 to GRF2. GRF1 and GRF2 are then connected to the TNT box via RIP. Assume that the following code is part of the **/etc/gated.conf** file for GRF1:

```
aggregate default {
   proto direct {10.8.17/30;} ;
} ;

export proto rip metric 3 {
  proto aggregate {all} ;
} ;
```

Assume that the following code is part of the **/etc/gated.conf** file for GRF2:

```
aggregate default {
   proto direct {10.8.16/30;} ;
} ;

export proto rip metric 2 {
  proto aggregate {all} ;
} ;
```

Note: The **aggregate** statement must come after the protocol statements in the **/etc/gated.conf** file.

The use of the **default** parameter in the previous example means that if the TNT receives any packet with an unknown destination address (that is, unknown to the TNT), the packet is sent to the appropriate GRF via the **default** route (which the TNT received via the export statements of the GRFs). The default route is built, but is not put in the forwarding table. It is built for export purposes only.

Instead of using the default route, you can aggregate to a specific route (for example, 10/8) by replacing **default** with 10/8 (that is, by using the **masklen** *number* / *number* parameter).

Aggregation and generation syntax

```
aggregate
    default | ( network [ ( mask mask ) | ( ( masklen number | / ) number ) ] )
    [ precedence precedence ][ preference preference ][ brief | truncate ] {
    proto [ all | direct | static | kernel | aggregate | proto ]
        [ ( as autonomous_system ) | ( tag tag )
            | ( aspath aspath_regexp ) ]
        restrict ; | [ precedence precedence ][ preference preference ]
        route_filter [ restrict | ( precedence precedence ) |
        (preference preference ) ] ;
    } ;
} ;
generate
    default | ( network [ ( mask mask ) | ( ( masklen number | / ) number ) ] )
    [ precedence precedence ] [ preference preference ] [ noinstall ] {
     proto [ all | direct | static | kernel | aggregate | proto ]
        [ ( as autonomous_system ) | ( tag tag )
            | ( aspath aspath_regexp ) ]
         restrict ; | [ precedence precedence ] [ preference preference ] {
       route_filter [ restrict | ( precedence precedence ) |
        ( preference preference ) ] ;
    } ;
} ;
```

default

Specifies that the aggregate route is advertised as a default route (for example, 0.0.0.0/0). When this parameter is used, the aggregate route is not installed as the default route on the local router.

precedence *precedence*

Specifies the precedence to assign to the resulting aggregate route. The default precedence is 130. When the precedence is set for a particular **proto** or **route_filter,** then the precedence is used for comparing against other contributing routes learned from different protocols. In this case, the contributing route with the lowest precedence is ordered first and the aggregate route is assigned that precedence value.

preference *preference*

Specifies the preference to assign to the resulting aggregate route for comparison against other aggregate routes. When the preference is set for a particular **proto** or **route_filter**, then the preference is used for comparing against other contributing routes learned from the same protocol. In this case, the contributing route with the lowest preference is ordered first and the aggregate route is assigned that preference value.

brief

Used to specify that the AS path should be truncated to the longest common AS path. The default is to build an AS path consisting of SETs and SEQUENCEs of all contributing AS paths. The **ATOMIC_AGGREGATE** path attribute is set on truncated AS paths.

noinstall

Normally, the route with the lowest preference is installed in the kernel forwarding table and is the route exported to other protocols. When **noinstall** is specified (on the **generate** statement only), the aggregate route is not installed in the kernel forwarding table when it is active, but it is still eligible to be exported to other protocols.

truncate

Used to specify that the AS path should be completely truncated, removing all elements of contributing AS paths. The truncated AS path is used even if there is only one contributing AS path. The **ATOMIC_AGGREGATE** path attribute is set on truncated AS paths.

proto *proto*

In addition to the special protocols listed, the contributing protocol can be chosen from among any of the ones supported by (and currently configured into) GateD, which includes BGP, OSPF, RIP, and IS-IS.

as *autonomous_system*

Restrict selection of routes to those learned from the specified autonomous system.

tag *tag*

Restrict selection of routes to those with the specified tag.

aspath *aspath_regexp*

Restrict selection of routes to those that match the specified AS path.

restrict

Indicates that these routes are not to be considered as contributors of the specified aggregate. The specified protocol can be any of the protocols supported by GateD.

route_filter

See the following subsection.

A route can only contribute to an aggregate route that is more general than itself. It must match the aggregate under its mask. Any given route can only contribute to one aggregate route, that is the most specific configured, but an aggregate route can contribute to a more general aggregate.

Route filters

All the formats allow route filters as shown below. See the section on Route Filters for a detailed explanation of how they work.

When no route filtering is specified (that is, when **restrict** is specified on the first line of a statement), all routes from the specified source match that statement. If any filters are specified, only routes that match the specified filters are considered as contributors. Put differently, if any filters are specified, an **all restrict ;** is assumed at the end of the list.

There is also a way to define filters and lists such that they can be referenced by a "common name" through a **define** filter statement. This method offers the following advantages:

– Declaration of a global filter once and reference by name

– Single parse cost

This means that the structure for the filter is created once, and the use of a pointer to the structure reduces the CPU cycles to parse the filter or list, and reduces the memory requirements if the filter or list is used more than once.

network **mask** *mask* [**exact** | **refines** | **between** *number* **and** *number*]
network (**masklen** | /) *number* [**exact** | **refines** | **between** *number* **and** *number*]
default
host *host*

GateD State Monitor (GSM)

The GateD State Monitor (GSM) provides an interactive interface to a running GateD daemon that can be used to query internal GateD variables. The GSM is implemented like any other protocol in GateD. You can log into it by executing the **gsm** command from the CLI or via **telnet** from a UNIX prompt. After user authentication is successfully accomplished, and the GSM interface is established, it answers any query sent to it as ASCII commands. The commands include such tasks as querying about the memory, routing table, interface list and other internal parameters.

Note: For this release, the GSM can be configured via the GSM statement in the **/etc/gated.conf** file. Through the use of this statement, you can configure the GSM on or off, plus configure the TCP port number to which the GSM binds, the users who are allowed access to the GSM, and the hosts from which GateD permits **telnet** connections to the GSM. For backward compatibility with previous releases, the GSM is on by default. In future post-1.4.8 releases, the GSM will be off by default. See the GSM statement in Chapter 2 for more information.

For complete information on how to establish a GSM interface, including information on using the CLI or UNIX prompt, the user authentication process, and how to query using the ASCII commands, see the **gsm** command in Chapter 1.

Glossary of terms

Terms used in descriptions throughout this document are defined here:

adjacency

A relationship formed between selected neighboring routers for the purpose of exchanging routing information. Not every pair of neighboring routers becomes adjacent.

autonomous system

A set of routers under a single technical administration, using an interior gateway protocol and common metrics to route packets within the AS, and using an exterior gateway protocol to route packets to other ASs.

Since this classic definition was developed, it has become common for a single AS to use several interior gateway protocols and sometimes several sets of metrics within an AS. The use of the term Autonomous System stresses that even when multiple IGPs and metrics are used, the administration of an AS appears to other ASs to have a single coherent interior routing plan and to present a consistent picture of what networks are reachable through it. The AS is represented by a number between 1 and 65534, assigned by the Internet Assigned Numbers Authority.

BGP
Border Gateway Protocol

One of a class of exterior gateway protocols, described in more detail in the BGP section of the Protocol Overview.

cost

An OSPF metric. See **metric**.

delay

A HELLO metric. Valid values are from zero to 30000 inclusive. The value of 30000 is the maximum metric and means unreachable. See **metric**.

designated router

OSPF: Each multi-access network that has at least two attached routers has a designated router. The designated router generates a link state advertisement for the multi-access network and assists in running the protocol. The designated router is elected by the HELLO protocol.

destination

Any network or any host.

distance

An EGP metric. See **metric**. Valid values are from zero to 255 inclusive.

egp
exterior gateway protocol
exterior routing protocol

A class of routing protocols used to exchange routing information within an autonomous system. A detailed explanation of exterior gateway protocols is available in the Protocol Overview.

`FIB`

The table in the kernel that controls the forwarding of packets is a forwarding table, ISO-defined as the forwarding information base.

gateway

An intermediate destination by which packets are delivered to their ultimate destination. A host address of another router that is directly reachable via an attached network. As with any host address it can be specified symbolically.

gateway_list

A list of one or more *`gateways`* separated by white space.

host

The IP address of any host. Usually specified as a dotted quad, four values in the range of 0 to 255 inclusive separated by dots (.). For example `132.236.199.63 or 10.0.0.51.`

It can also be specified as an eight digit hexadecimal string preceded by `0x`. For example `0x????????` or `0x0a000043`. Finally, if `options noresolv` is not specified, a symbolic hostname such as `gated.cornell.edu` or `nic.ddn.mil` is usable. The numeric forms are much preferred over the symbolic form.

interface

The host address of an attached interface. This is the address of a *broadcast*, *nbma* or *loopback* interface and the remote address of a *point-to-point* interface. As with any host address, it can be specified symbolically.

interface

The connection between a router and one of its attached networks.

A physical interface can be specified by a single IP address, domain name, or interface name. (Unless the network is an unnumbered point-to-point network.) Multiple levels of reference in the configuration language allow identification of interfaces using wildcard, interface type name, or delete word address. Be careful with the use of interface names as future Unix operating systems can allow more than one address per interface. Dynamic interfaces can be added or deleted, and indicated as up or down as well as changes to address, netmask and metric parameters.

igp
interior gateway protocol
interior routing protocol

One of a class of routing protocols used to exchange routing information within an autonomous system. A detailed explanation of interior gateway protocols is available in the Protocol Overview.

interface_list

A list of one or more interface names including wildcard names (names without a number) and names that can specify more than one interface or address, or the token "all" for all interfaces. See the section on interface lists for more information (page A-19).

IS-IS

One of a class of interior gateway protocols, described in more detail in the IS-IS section of the Protocol Overview.

local_address

The host address of an attached interface. This is the address of a *broadcast*, *nbma* or *loopback* interface and the local address of a *point-to-point* interface. As with any host address, it can be specified symbolically.

mask

A means of subdividing networks using address modification. A mask is a dotted quad specifying which bits of the destination are significant. Except when used in a route filter, GateD only supports contiguous masks.

mask length

The number of significant bits in the mask.

metric

One of the units used to help a system determine the best route. Metrics can be based on hop count, routing delay, or an arbitrary value set by the administrator depending on the type of routing protocol. Routing metrics can influence the value of assigned internal preferences. (See **preference**.)

This sample table shows the range of possible values for each routing protocol metric and the value used by each protocol (See Protocol Overview, page A-27) to reach a destination.

SAMPLE ROUTING PROTOCOL METRICS

Protocol	Metric Represents	Range	Unreachable
RIP	distance (hop-count)	0-15	16
OSPF	cost of path	0-?????	Delete
ISIS	cost of path	0-254	Delete
BGP	unspecified	0-65534	65535

multiaccess networks

Those physical networks that support the attachment of multiple (more than two) routers. Each pair of routers on such a network is assumed to be able to communicate directly.

natural mask

neighbor

Another router which with implicit or explicit communication is established by a routing protocol. Neighbors are usually on a shared network, but not always. This term is mostly used in OSPF and EGP. Usually synonymous with *peer.*

neighboring routers

Two routers that have interfaces to a common network. On multi-access networks, routers are dynamically discovered by the OSPF HELLO protocol.

network

Any packet-switched network. A network can be specified by its IP address or network name. The host bits in a network specification must be zero. *Default* can be used to specify the default network (0.0.0.0).

network

The IP address of a network. Usually specified as a dotted quad, one to four values, in the range 0 to 255 inclusive, separated by dots (.), for example, `132.236.199, 132.236` or `10`. It can also be specified as a hexadecimal string preceded by `0x` with an even number of digits, of length between two and eight. For example `0xnnnnnn, 0xnnnn,` or `0x0n`.

Also allowed is the symbolic value `default` that has the distinguished value `0.0.0.0`, the default network. If `options noresolv` is not specified, a symbolic network name such as `nr-tech-prod`, `cornellu-net` and `arpanet` is usable. The numeric forms are preferred over the symbolic form.

number

A positive integer.

OSPF
Open Shortest Path First

One of a class of interior gateway protocols, described in more detail in the OSPF section of the Protocol Overview.

ospf_area

OSPF allows collections of contiguous networks and hosts to be grouped together. Such a group, together with the routes having interfaces to any of the included networks, is called an area.

Each area runs a separate copy of the basic link-state algorithm. This means that each area has its own topological database. The topology of an area is invisible from the outside of the area. Conversely, routers internal to a given area know nothing of the detailed topology external to the area. This isolation of knowledge enables OSPF to effect a marked reduction in routing traffic as compared to treating the entire autonomous system as a single link-state domain.

peer

Another router which with implicit or explicit communication is established by a routing protocol. Peers are usually on a shared network, but not always. This term is mostly usedby BGP. Usually synonymous with *neighbor.*

port

A UDP or TCP port number. Valid values are from 1 through 65535 inclusive.

precedence (*protocol-precedence*)

A precedence is a value between 0 (zero) and 255, used to set the predecence of routing protocols. The protcol with the best (numerically lowest) precedence contains the prefixes found in the kernel forwarding table and exported to other protocols. A default precedence is assigned to each source from which GateD receives routes (See **preference**.)

preference

A preference is a value between 0 (zero) and 65536, used to select between many routes to the same destination. The route with the best (numerically lowest) preference is the active route. The active route is the one installed in the kernel forwarding table and exported to

other protocols. A default preference is assigned to each source from which GateD receives routes; it has the value of 0. (See **precedence**.)

prefix

A contiguous mask covering the most significant bits of an address. The prefix length specifies how many bits are covered.

QoS
quality of service

The OSI equivalent of TOS.

RIB

The table that GateD uses internally to store routing information learned from routing protocols is a routing table, ISO-defined as the routing information base.

RIP
Routing Information Protocol

One of a class of interior gateway protocols, described in more detail in the RIP section of the Protocol Overview.

reject route

XXX - define.

route filter

A filter is a configurable entity based on destination address or destination address and mask specified to match a route that is to be filtered out and then ignored when such a route is advertised.

router id

A 32-bit number assigned to each router running the OSPF protocol. This number uniquely identifies the router within the autonomous system.

router_id

An IP address used as unique identifier assigned to represent a specific router.

This is usually the address of an attached interface.

RIB
routing information base
routing database
routing table

The repository of all of GateD's retained routing information, used to make decisions and as a source for routing information that is to be propagated.

simplex

An interface can be marked as simplex either by the kernel, or by interface configuration. A simplex interface is an interface on a broadcast media that is not capable of receiving packets it broadcasts.

GateD takes advantage of interfaces that are capable of receiving their own broadcast packets to monitor whether an interface appears to be functioning properly.

time

A time value, usually a time interval. It can be specified in any one of the following forms:

number

A non-negative decimal number of seconds. For example, `27, 60` or `3600`.

number:number

A non-negative decimal number of minutes followed by a seconds value in the range of zero to 59 inclusive. For example, `0:27`, `1:00` or `60:00`.

number:number:number

A non-negative decimal number of hours followed by a minutes value in the range of zero to 59 inclusive followed by a seconds value in the range of zero to 59 inclusive. For example, `0:00:27`, `0:01:00` or `1:00:00`.

time to live
ttl

The *Time To Live* (TTL) of an IP packet. Valid values are from one (1) through 255, inclusive.

TOS
type of service

The *type of service* is for internet service quality selection. The type of service is specified, along the abstract parameters precedence, delay, throughput, reliability, and cost. These abstract parameters are to be mapped into the actual service parameters of the particular networks the datagram traverses. The vast majority of IP traffic today uses the default type of service.

RFC references

RFC 827

E. Rosen, *Exterior Gateway Protocol EGP*

RFC 891

D. Mills, *Dcn Local-network Protocols*

RFC 904

D. Mills, *Exterior Gateway Protocol Formal Specification*

RFC 1058

C. Hedrick, *Routing Information Protocol*

RFC 1105

K. Lougheed, Y. Rekhter, *Border Gateway Protocol BGP*

RFC 1163

K. Lougheed, Y. Rekhter, *A Border Gateway Protocol (BGP)*

RFC 1164

J. Honig, D. Katz, M. Mathis, Y. Rekhter, J. Yu,
Application of the Border Gateway Protocol in the Internet

RFC 1227

M. Rose, *SNMP MUX Protocol and MIB*

RFC 1245

J. Moy, *OSPF Protocol Analysis*

RFC 1246

J. Moy, *Experience with the OSPF Protocol*

RFC 1253

F. Baker, R. Coltun, *OSPF Version 2 Management Information Base*

RFC 1256

S. Deering, *ICMP Router Discovery Messages*

RFC 1265

Y. Rekhter, *BGP Protocol Analysis*

RFC 1266

Y. Rekhter, *Experience with the BGP Protocol*

RFC 1267

K. Lougheed, Y. Rekhter, *A Border Gateway Protocol 3 (BGP-3)*

RFC 1268

P. Gross, Y. Rekhter, *Application of the Border Gateway Protocol in the Internet*

RFC 1269

J. Burruss, S. Willis,
*Definitions of Managed Objects for the Border Gateway Protocol (*v. 3)

RFC 1321

R. Rivest, *The MD-5 Message Digest Algorithm*

RFC 1370

Internet Architecture Board, *Applicability Statement for OSPF*

RFC 1388

G. Malkin, *RIP Version 2 Carrying Additional Information*

RFC 1397

D. Haskin, *Default Route Advertisement In BGP2 And BGP3 Versions Of The Border Gateway Protocol*

RFC 1403

K. Varadhan, *BGP OSPF Interaction*

RFC 1583

J. Moy, *OSPF Version 2*

Appendix B: gsm command description

B

Appendix B contains the 1.4.10 updated description for the GateD State Monitor command, **gsm**, and replaces the **gsm** command description currently found in Chapter 1 of the *GRF 1.4 Reference Guide*.

gsm
(CLI + shell)

The GateD State Monitor (GSM) provides an interactive interface from which you can interrogate the state of route tables, interfaces, routing protocols, and other internal parameters by using the GSM query commands. The following subsections explain how to establish a GSM interface, and how to query for information.

Note: Refer to Chapter 2 in this manual for more information about the GSM, including how to configure the use of the GSM in the **/etc/gated.conf** file. For backwards compatibility with previous releases, the GSM is on by default. In future post -1.4.8 releases, the GSM will be off by default.

Permission level: system

GSM connection options

There are two methods for establishing a GSM interface: by using the CLI **gsm** command (which establishes a **telnet** connection for you) or through a direct **telnet** connection. The following subsections describe these two methods.

CLI interface

gsm [***hostname***]

When you execute **gsm** from the CLI prompt, it returns information about the GateD running on that local GRF. To obtain GateD statistics for a different GRF system, you can use the ***hostname*** option to establish GSM connection to that GRF, or you can use the **telnet** command as described in the telnet connection subsection.

Note: As shown in the following example, the CLI help function does not provide information on the various GSM query commands, only how to establish a GSM interface. For descriptions of GSM subcommands, you must have actually logged into the GSM:

```
super> ? gsm
Usage: gsm [hostname]
super> gsm ?
usage: telnet [-l user] [-a] host-name [port]
super>
```

To establish a GSM interface from the CLI, you need the password for the administrative **netstar** ID account; the default password for this account is **NetStar**. If you change the administrative password, use the new one during the authentication process.

In the 1.4.8 release, access to the GSM can be configured through the **/etc/gated.conf** file; the most noticeable difference is that you may not be allowed to connect to the GSM, or if the connection is allowed, you may be prompted for a user name and a password. See the GSM statement subsection in Chapter 2 for more information on how to configure the usage of the GSM.

The following example shows how to establish a GSM interface if it is not configured through the **/etc/gated.conf** file:

```
super> gsm
Trying 127.0.0.1...
Connected to 127.0.0.1.
Escape character is '^]'.
Password?-------- (for example, NetStar)
```

The following example shows how to establish a GSM interface if it has been configured to prompt for a user name through the **/etc/gated.conf** file:

```
super> gsm
Trying 127.0.0.1...
Connected to 127.0.0.1.
Escape character is '^]'.
User?-------- (must be a user name as specified in the /etc/gated.conf file)
Password?-------- (must be the password for the user entered in the line above)
```

Because the GSM can also be configured to refuse connections from specific hosts (via the **/etc/gated.conf** file), you may receive the message "telnet: Unable to connect to remote host: Connection refused" after the "trying 127.0.0.1..." message. If this happens, check the **gsm** statement in **/etc/gated.conf** to see if it excludes the loopback address.

After authentication has successfully completed, the GSM interface starts. The **gsm** prompt uses the machine domain name, as shown in the following example:

```
1GRF Gated State Monitor. Version GateD R3_5Beta_3; CVS
Branch:A1_4_1; Path:/
/gated/code/GSM
GateD-router.sitename.com>
```

It is at this point that you can use the **help** command for information on how to query for information. See the Using the help command subsection for more information.

telnet connection

From a UNIX shell, you can establish a GSM interface by opening a **telnet** connection on the TCP port specified in **/etc/gated.conf** or by default, TCP port **616** to the machine running GateD. You can **telnet** from the administrative LAN or from the GateD machine itself. The syntax for the **telnet** command is as follows:

telnet *host-name* [*port*]

host-name Name or IP address of machine running GateD.

port TCP port **616** or configured port to access GSM.

To **telnet** to the GSM interface, you must use the password from the administrative **netstar** ID account; the default password shipped with your system is **NetStar**. If you have changed the administrative password, use the new password.

In this release, access to the GSM can be configured through the **/etc/gated.conf** file; the most noticeable difference is that when you try to connect through the **telnet** command, you may not be allowed to connect to the GSM, or if the connection is allowed, you may be prompted for a user name, along with the administrative password. See the GSM statement subsection in Chapter 2 for more information on how to configure the usage of the GSM.

The following example shows how to use the **telnet** command to log into GateD running on a router at address 10.22.22.22 and connecting to TCP port **616** (the GSM is not configured via the **/etc/gated.conf** file in this example). Note that the **gsm** prompt contains the machine domain name:

```
telnet 10.22.22.22 616
Trying 10.22.22.22...
Connected to 10.22.22.22.
Escape character is '^]'.
Password? -------   (for example, NetStar)
1GRF Gated State Monitor. Version GateD R3_5Beta_3;
Path:
GateD-router.sitename.com>
```

The following example shows how to establish a GSM interface if it has been configured to prompt for a user name through the **/etc/gated.conf** file:

```
telnet 10.22.22.22 616
Trying 10.22.22.22...
Connected to 10.22.22.22.
Escape character is '^]'.
User?-------- (must be a user name as specified in the /etc/gated.conf file)
Password?-------- (must be the password for the user entered in the line above)
1GRF Gated State Monitor. Version GateD R3_5Beta_3;
Path:
GateD-router.sitename.com>
```

Because the GSM can be configured to refuse connections from specific hosts (via the **/etc/gated.conf** file), you can receive the message "telnet: Unable to connect to remote host: Connection refused" after the "trying 10.22.22.22..." message. If this happens, check the **gsm** statement in **/etc/gated.conf** to see if it excludes the loopback address.

Once the GSM interface has been successfully established, you can use the **help** command for information on how to query for information. The following subsection describes the **help** command.

Using the help command

Once the GSM interface connection has been establish, the GSM answers queries submitted through a series of commands. The top-level **help** | **?** command provides the list of available commands. Commands may be abbreviated when no confusion is possible (see the Abbreviating commands subsection for more information). Some of these commands have associated subcommands, which are explained in following subsections. The following example shows how to use the **help** command:

```
GateD-router.sitename.com> help
1GRF HELP: The possible commands are:
1GRF    ?     : Print help messages
1GRF    help  : Print help messages
1GRF    show  : Show internal values
1GRF    quit  : Close the session
1GRF    enable: Enable the session
1GRF    exec  : Execute actions (must be enabled)
1GRF    exit  : Close the session
```

Second-level commands
To view information regarding different GRF components, you can use the second-level **show** subcommand, as shown in the following example:

```
GateD-router.sitename.com> show
1GRF HELP: The possible show subcommands are:
1GRF    version  : Show the current GateD version
1GRF    kernel   : Show the Kernel support
1GRF    iso      : Show the ISO support
1GRF    interface [name|index]: Show interface status
1GRF    memory   : Show the memory allocation
1GRF    ip       : Show info about IP protocol
1GRF    task     : Show list of active tasks
1GRF    ospf     : Show info about OSPF protocol
1GRF    timer    : Show list of timers
1GRF    bgp      : Show info about BGP protocol
1GRF    rip      : Show info about rip protocol
1GRF    isis     : Show info about ISIS protocol
```

Third-level
You can use third-level commands to obtain a finer granularity of information for a specific GRF component, as shown in the following example (which is for the IP protocol):

```
GateD-router.sitename.com> show ip
1GRF HELP: The possible subcommands are:
1GRF    sum       Show info about IP routes
1GRF    exact     [x.x.x/len]: Show info about specific IP routes
1GRF    aggregate [x.x.x/len]: Show info about specific Aggregate
routes
1GRF    all       Show entire FIB
1GRF    consume   [x.x.x/len]: Show consuming route and more spe-
cific route
1GRF    lessspec   Show less specific routes per protocol
1GRF    refines    Show more specific routes per protocol
```

The following example shows the use of the **show** command to display OSPF options:

```
GateD-router.sitename.com> show ospf
1GRF HELP: The possible subcommands are:
1GRF    summary  : Show OSPF Summary
1GRF    errors   : Show OSPF Error Counters
1GRF    interface: Show interface status
1GRF    io stats : Show I/O stats
1GRF    nexthops : Show OSPF nexthops
1GRF    ase      : Show OSPF ASE Database
1GRF    lsdb     : Show OSPF LSDB Database
1GRF    border   : Show OSPF ASB/AB RTR Database
```

The following example shows the use of the **show** command to display BGP options:

```
GateD-router.sitename.com> show bgp
```

```
1GRF HELP: The possible subcommands are:
1GRF    summary: Show BGP summary
1GRF    peeras [AS number]: Show BGP peer info
1GRF    group  [AS number]: Show BGP group summary
1GRF    aspath : Show BGP aspath info
```

The following example shows the use of the **show** command to display IS-IS options:

```
GateD-router.sitename.com> show isis
1GRF HELP: The possible subcommands are:
1GRF    summary: Show IS-IS summary
```

The recommended way to display and view IP route tables is through the GSM. Establish a GSM session and use the **show ip all** command, as shown in the following example:

```
GateD-router.sitename.com> sh ip all
1GRF Sta        78.78.78/24 203.3.1.153     IGP (Id 1)
1GRF Sta        99.99.98/24 202.1.1.153     IGP (Id 1)
1GRF Sta        99.99.99/24 212.1.3.152     IGP (Id 1)
1GRF Sta             127/8  127.0.0.1       IGP (Id 1)
1GRF Sta      198.174.11/24 206.146.160.1   IGP (Id 1)
1GRF Dir         202.1.1/24 202.1.1.151     IGP (Id 1)
1GRF Dir         202.1.2/24 202.1.2.151     IGP (Id 1)
1GRF Dir         202.1.3/24 202.1.3.151     IGP (Id 1)
1GRF Dir         202.5.2/24 202.5.2.151     IGP (Id 1)
1GRF Dir         202.5.4/24 202.5.4.151     IGP (Id 1)
1GRF Dir         203.3.1/24 203.3.1.151     IGP (Id 1)
1GRF Dir        204.10.1/24 204.10.1.151    IGP (Id 1)
1GRF Sta   204.100.1.147/32 208.1.1.152     IGP (Id 1)
1GRF Sta     205.2.4.138/32 212.1.2.134     IGP (Id 1)
1GRF Dir     206.146.160/24 206.146.160.151 IGP (Id 1)
1GRF Dir         208.1.1/24 208.1.1.151     IGP (Id 1)
1GRF Dir         212.1.1/24 212.1.1.151     IGP (Id 1)
1GRF Dir         212.1.2/24 212.1.2.151     IGP (Id 1)
1GRF Dir         212.1.3/24 212.1.3.151     IGP (Id 1)
```

In the following example, the output for **show ip cons** displays different information for a specific route:

```
GateD-router.sitename.com> sh ip cons 202/8
No less specific routes found for IP route 202 mask 255
More specific routes for IP route 202 mask 255...
1GRF Dir         202.1.1/24 202.1.1.151     IGP (Id 1)
1GRF Dir         202.1.2/24 202.1.2.151     IGP (Id 1)
1GRF Dir         202.1.3/24 202.1.3.151     IGP (Id 1)
1GRF Dir         202.5.2/24 202.5.2.151     IGP (Id 1)
1GRF Dir         202.5.4/24 202.5.4.151     IGP (Id 1)
```

Frequently-used commands

The following commands have proven useful in managing and debugging GateD configurations:

```
s ip al                    # all routes to be installed in kernel
s ip ref <protocol> 0      # all routes learned via <protocol>
s ip ref al xx             # all routes starting with "xx"
s ip exact x.x.x/len       # everything we know about a route
```

For example, the following shows how to display all routes with "99" in first octet:

```
GateD-router.sitename.com> s ip re al 99
OSP         99.82.1/24 10.2.1.82       (666) IGP (Id 3)
OSP     99.82.82.82/32 10.2.1.82       (666) IGP (Id 3)
Sta        99.175.1/24 206.146.160.1   IGP (Id 2)
Sta  99.175.175.175/32 206.146.160.1   IGP (Id 2)
```

The following example shows how to display **exact** (host route) information:

```
GateD-router.sitename.com> s ip ex  99.82.82.82/32
Route 99.82.82.82 Mask 255.255.255.255 Entries 1 Announced 1 Depth
0 <>
                      Instability Histories:
Proto  Route    NextHop     Proto-prec/Pref Metric/2   Tag Installed
* OSPF_ASE 99.82.82.82  10.2.1.82   150/-   1/30   C0000000
67:49:48
       Forwarding Interface: 10.2.1.175(gf010)
       Status <Int Ext Active Gateway>
       ASPATH (666) IGP (Id 3)
```

Abbreviating commands

Abbreviating commands is acceptable provided the abbreviation unambiguously identifies an entity, as shown in the following example:

```
GateD-router.sitename.com> sh ip sum
1GRF IP radix tree: 38 nodes, 20 routes
```

The following example shows the shortest abbreviation of a command that works:

```
GateD-router.sitename.com> s ip al
```

Another form of abbreviation is shown in the following example, where the value x.x.x/len can be replaced with 0 to show all:

```
GateD-router.sitename.com> s ip ref b 0    # show all routes
                                           # learned via BGP
```

Appendix C: /etc/gated.conf file

C

Appendix C contains the 1.4.10 updated version of the `/etc/gated.conf` configuration file. It replaces the version of the file currently found in Appendix A of the *GRF Reference Guide 1.4*.

/etc/gated.conf components

The next pages show many of the components available for use in a `/etc/gated.conf` file.

/etc/gated.conf Page 1 of 9

```
# Trace option statement
   traceoptions ["trace_file" [replace] [ size size[k|m] files files ]]
           [control_options] trace_options [except trace_options] ;
   traceoptions none ;

# Option statement
options
    [ nosend ]
    [ noresolv ]
    [ gendefault [ precedence precedence ] [ gateway gateway] ]
    [ syslog [ upto ] log_level ]
    [ mark time ]
    ;

# GSM statement
gsm ( yes | no | on | off )
[ {
    [ port port-number ; ]
    [ usernames user-list ; ]
    [ hosts host-set ; ]
} ] ;
# Interface statement
interfaces {
    options
        [ strictinterfaces ]
        [ scaninterval time ]
        ;
    interface interface_list
        [ precedence precedence ]
        [ down precedence precedence ]
        [ passive ]
        [ simplex ]
        [ reject ]
        [ blackhole ]
        ;
    define address
        [ broadcast address ] | [ pointtopoint address ]
        [ netmask mask ]
        [ multicast ]
        ;
} ;
# Definition statement
autonomoussystem autonomous_system [ loops number ] ;
multipath ( yes | no | off | on ) ;
routerid host ;
confederation confederation ;
routing-domain rdi ;
martians {
        host host [ allow ] ;
        network [ allow ] ;
        network mask mask [ allow ] ;
        network ( masklen | / ) number [ allow ] ;
        default [ allow ] ;
    } ;
```

/etc/gated.conf Page 2 of 9

```
# RIP statement
rip ( yes | no | on | off ) [ {
    broadcast ;
    nobroadcast ;
    nocheckzero ;
    precedence precedence ;
    defaultmetric metric ;
    query authentication [none | ([simple|md5] password)] ;
    interface interface_list
        [noripin] | [ripin]
        [noripout] | [ripout]
        [metricin metric]
        [metricout metric]
        [version 1]|[version 2 [multicast|broadcast]]
        [[secondary] authentication [none| ([simple|md5] password)]]
    trustedgateways gateway_list ;
    sourcegateways gateway_list ;
    traceoptions trace_options ;
} ] ;



# OSPF statement
ospf ( yes | no | on | off ) [ {
    defaults {
        precedence precedence ;
        cost cost ;
        tag [ as ] tag ;
        type 1 | 2 ;
        inherit-metric ;
    } ;
    exportlimit routes ;
    exportinterval time ;
    traceoptions trace_options ;
    monitorauthkey authkey ;
    monitorauth none | ( [ simple | md5 ] authkey ) ;
    backbone | ( area area ) {
        authtype  none | simple ;
        stub [ cost cost] ;
        networks {
            network [ restrict ] ;
            network mask mask [ restrict ] ;
            network ( masklen | / ) number [ restrict ] ;
            host host [ restrict ] ;
        } ;
        stubhosts {
            host  cost cost ;
        } ;
        interface interface_list; [cost cost ] {
            interface_parameters
        } ;
        interface interface_list nonbroadcast [cost cost ] {
            pollinterval time ;
            routers {
                gateway [ eligible ] ;
            } ;
            interface_parameters
        } ;
```

/etc/gated.conf Page 3 of 9

```
# OSPF statement (continued)
        Backbone only:
        virtuallink neighborid router_id transitarea area {
            interface_parameters
        } ;
    } ;
} ;


# IS-IS statement
       isis no | ip {
          level 1|2 ;
          [traceoptions <isis_traceoptions> ;]
          [systemid <string> ;]
          [area <string> ;]
          [set <isis_parm> value ;]
          circuit|interface <interface-name>
              metric [level 1|2] metric
              priority [level 1|2] priority pointopoint ;
          [ipreachability level (1|2) (internal|external|summary)
              ipaddr netmask [metric metric;]]
       } ;


# BGP statement
    bgp ( yes | no | on | off )
    {
       [ precedence precedence ; ]
       [ preference preference ; ]
       [ allow bad community ; ]
       [ defaultmetric metric ; ]
       [ traceoptions trace_options ; ]
       [ clusterid host ; ]
       [ disable export best ; ]
      [ group type
            ( external peeras autonomous_system
                  [ ignorefirstashop ]  [ subgroup integer ]
                  [ med ] [ nexthopself ] )

           | ( internal peeras autonomous_system
                  [ reflector-client [ no-client-reflect ] ]
                  [ ignorefirstashop ]
                  [ lcladdr local_address ]
                  [ outdelay time ]
                  [ metricout metric ]
                  [ subgroup integer ]
                  [ nexthopself ]  )

           | ( routing peeras autonomous_system proto proto_list
                  interface interface_list
                  [ reflector-client [ no-client-reflect ] ]
                  [ ignorefirstashop ]
                  [ lcladdr local_address ]
                  [ outdelay time ]
                  [ metricout metric ]
                  [ subgroup integer ]
                  [ nexthopself ]  )
```

/etc/gated.conf Page 4 of 9

```
   # BGP statement (continued)

               | ( confed peeras autonomous_system proto proto_list
                        interface interface_list
                        [ reflector-client [ no-client-reflect ] ]
                        [ ignorefirstashop ]
                        [ lcladdr local_address ]
                        [ outdelay time ]
                        [ metricout metric ]
                        [ subgroup integer ]
                        [ nexthopself ]  )

              | ( test peeras autonomous_system )  ]
                {
                  [  allow {
                        network
                        network mask mask
                        network ( masklen | / ) number
                        all
                        host host  ]
                   } ;
                      peer host
                        [ metricout metric ]
                        [ ignorefirstashop ]
                        [ nogendefault ]
                        [ gateway gateway ]
                        [ precedence precedence ]
                        [ preference preference ]
                        [ holdtime time ]
                        [ version number ]
                        [ passive ]
                        [ sendbuffer number ]
                        [ recvbuffer number ]
                        [ outdelay time ]
                        [ keep [ all | none ] ]
                        [ show-warnings ]
                        [ noaggregatorid ]
                        [ keepalivesalways
                        [ v3asloopokay ]
                        [ nov4asloop ]
                        [ ascount count ]
                        [ throttle count ]
                        [ allow bad routerid ]
                        [ logupdown ]
                        [ ttl ttl ]
                        [ traceoptions trace_options ]
                        [ nexthopself ]
                        ;
                  } ;
            } ;
```

/etc/gated.conf Page 5 of 9

```
# Weighted route dampening statement

   dampen-flap {
      [suppress-above metric;
      reuse-below metric;
      max-flap metric;
      unreach-decay time;
      reach-decay  time;
      keep-history  time; ]
   };


# ICMP statement
   icmp {
       traceoptions trace_options ;
   }


# Router discovery statement
   routerdiscovery server ( yes | no | on | off ) [ {
       traceoptions trace_options ;
       interface interface_list
           [ minadvinterval time ] |
           [ maxadvinterval time ] |
           [ lifetime time ]
           ;
       address interface_list
           [ advertise ] | [ ignore ] |
           [ broadcast ] | [ multicast ] |
           [ ineligible ] | [ preference preference ]
           ;
   } ] ;


# Router discovery client statement
   routerdiscovery client ( yes | no | on | off ) [ {
       traceoptions trace_options ;
       precedence precedence ;
       interface interface_list
           [ enable ] | [ disable ]
           [ multicast ] | [ broadcast ]
           [ quiet ] | [ solicit ]
           ;
   } ] ;
```

/etc/gated.conf Page 6 of 9

```
# Kernel statement
   kernel {
       options
           [ nochange ]
           [ noflushatexit ]
           ;
       routes number ;
       flash
           [ limit number ]
           [ type interface | interior | all ]
           ;
       background
           [ limit number ]
           [ priority flash | higher | lower ]
          ;
       traceoptions trace_options ;
   } ;

# Static statement
   static {
       ( host host ) | default |
       ( network [ ( mask mask ) | ( ( masklen | / ) number ) ] )
           gateway gateway_list
           [ interface interface_list ]
           [ precedence precedence ]
           [ retain ]
           [ reject ]
           [ blackhole ]
           [ noinstall ] ;
       ( network [ ( mask mask ) | ( ( masklen | / ) number ) ] )
           interface interface
           [ precedence precedence ]
           [ retain ]
           [ reject ]
           [ blackhole ]
           [ noinstall ] ;
   } ;

# Filtering statements
   network mask mask [ exact | refines | between number and number ]
   network masklen |/ number [ exact | refines | between number and number ]
   all
   default
   host host

# Autonomous system matching statement
   aspath aspath_regexp origin any | ( [ igp ] [egp ] [ incomplete ] )

# AS path attribute statement
   aspath-opt {
       [ community autonomous_system : community-id | community-id ]
       [ community no-export | no-advertise | no-export-subconfed | none ]
   }
   mod-aspath {
       [ community autonomous_system : community-id |community-id ]
       [ community no-export | no-advertise | no-export-subconfed ]
       [ comm-split autonomous_system community-id ]
   }
```

/etc/gated.conf Page 7 of 9

```
   # Import from BGP statement
      import proto bgp as autonomous_system
          [ subgroup integer ]  [ aspath-opt ] restrict ;
      import proto bgp as autonomous_system
          [ subgroup integer ] [ aspath-opt ] [ precedence precedence ]
           [ preference preference ]  [ localpref preference ] {
          route_filter  [ restrict | ( precedence  precedence )
           | ( preference preference ) | ( localpref precedence ) ] ;
      } ;

      import proto bgp aspath aspath_regexp
          origin any | ( [ igp ] [egp ] [ incomplete ] )
          [ aspath-opt ] restrict ;
      import proto bgp aspath aspath_regexp
          origin any | ( [ igp ] [egp ] [ incomplete ] [ localpref preference ] )
          [ aspath-opt ] [ precedence precedence ] [ preference preference ]
           [ localpref preference ] {
          route_filter [ restrict | ( precedence precedence ) | ( preference pref-
   erence )
           | ( localpref preference ) ] ;
      } ;

   # Import from RIP and Redirect statements
      import proto rip | redirect
          [ ( interface interface_list ) | (gateway gateway_list ) ]
          restrict ;
      import proto rip | redirect
          [ ( interface interface_list ) | (gateway gateway_list ) ]
          [ precedence precedence ] {
          route_filter  [ restrict | ( precedence precedence ) ;
      } ;

   # Import from OSPF statements
      import proto ospfase [ tag ospf_tag ] restrict ;
      import proto ospfase [ tag ospf_tag ]
          [ precedence precedence ] {
          route_filter  [ restrict | ( precedence  precedence ) ;
      } ;

   # Exporting to BGP
      export proto bgp as autonomous system
          restrict ;
      export proto bgp as autonomous system [ mod-aspath ]
          [ subgroup integer] [ metric metric  | localpref preference ]  {
          export_list ;
      } ;

   # Exporting to RIP
      export proto rip
          [ ( interface interface_list ) | ( gateway gateway_list ) ]
          restrict ;
      export proto rip
          [ ( interface interface_list ) | ( gateway gateway_list ) ]
          [ metric metric ] {
          export_list ;
      } ;
```

/etc/gated.conf Page 8 of 9

```
   # Exporting to OSPF statements
      export proto ospfase [ type 1 | 2 ] [ tag ospf_tag ]
          restrict ;
      export proto ospfase [ type 1 | 2 ] [ tag ospf_tag ]
          [ metric metric ] {
          export_list ;
      } ;

   # Exporting BGP routes

      proto bgp as autonomous_system   [ subgroup integer ]
          restrict ;
      proto bgp as autonomous_system  [ subgroup integer ]
          [ metric metric  | localpref preference ] {
          route_filter [ restrict | ( metric metric ) ] ;
      } ;

   # Exporting RIP

      proto rip
          [ ( interface interface_list ) | (gateway gateway_list ) ]
          restrict ;
      proto rip
          [ ( interface interface_list ) | (gateway gateway_list ) ]
          [ metric metric ] {
          route_filter [ restrict | ( metric metric ) ] ;
      } ;

   # Exporting OSPF routes
      proto ospf | ospfase restrict ;
      proto ospf | ospfase [ metric metric ] {
          route_filter [ restrict | ( metric metric ) ] ;
      } ;

   # Exporting routes from non-routing protocol statement
      proto direct | static | kernel
          [ (interface interface_list ) ]
          restrict ;
      proto direct | static | kernel
          [ (interface interface_list ) ]
          [ metric metric ] {
          route_filter [ restrict | ( metric metric ) ] ;
      } ;

   # Exporting by AS path
      proto bgp | all aspath aspath_regexp
          origin any | ( [ igp ] [egp ] [ incomplete ] )
          restrict ;
      proto bgp | all aspath aspath_regexp
          origin any | ( [ igp ] [egp ] [ incomplete ] )
          [ metric metric  | localpref preference ] {
          route_filter [ restrict | ( metric metric ) ] ;
      } ;
```

/etc/gated.conf Page 9 of 9

```
   # Exporting by route tag
      proto rip | ospf | all tag tag restrict ;
      proto proto | all tag tag
          [ metric metric ] {
          route_filter [ restrict | ( metric metric ) ] ;
      } ;

   # Aggregation and generation statement
      aggregate
          default | ( network [ ( mask mask ) | ( ( masklen number | / ) number ) ] )
          [ precedence precedence ] [ preference preference ] [ brief | truncate ] {
          proto [ all | direct | static | kernel | aggregate | proto ]
               [ ( as autonomous_system ) | ( tag tag )
                | ( aspath aspath_regexp ) ]
               restrict ; | [ precedence precedence ] [ preference preference ]
               route_filter [ restrict | ( preference preference ) ] |
               ( preference preference ) ] ;
          } ;
      } ;
      generate
          default | ( network [ ( mask mask ) | ( ( masklen number | / ) number ) )
   ] )
          [ preference preference ] [ preference preference ] [ noinstall ] {
          proto [ all | direct | static | kernel | aggregate | proto ]
               [ ( as autonomous_system ) | ( tag tag )
                   | ( aspath aspath_regexp ) ]
               restrict ; | [ preference preference ] [ preference preference ] {
               route_filter [ restrict | ( preference preference ) |
               (preference preference ] ;
          } ;
      } ;


   # Directive statement
      %directory "directory"

   # Include statement
      %include "filename"
```

Figure C-1. Components configurable in a /etc/gated.conf file