RADIUS Server Load Balancing

The RADIUS Server Load Balancing feature distributes authentication, authorization, and accounting (AAA) authentication and accounting transactions across RADIUS servers in a server group. These servers can share the AAA transaction load and thereby respond faster to incoming requests.

This module describes the RADIUS Server Load Balancing feature.

Finding Feature Information

Your software release may not support all the features documented in this module. For the latest caveats and feature information, see Bug Search Tool and the release notes for your platform and software release. To find information about the features documented in this module, and to see a list of the releases in which each feature is supported, see the feature information table at the end of this module.

Use Cisco Feature Navigator to find information about platform support and Cisco software image support. To access Cisco Feature Navigator, go to www.cisco.com/go/cfn. An account on Cisco.com is not required.

Prerequisites for RADIUS Server Load Balancing

- Authentication, authorization, and accounting (AAA) must be configured on the RADIUS server.
- AAA RADIUS server groups must be configured.
- RADIUS must be configured for functions such as authentication, accounting, or static route download.

Restrictions for RADIUS Server Load Balancing

- Incoming RADIUS requests, such as Packet of Disconnect (POD) requests, are not supported.

- Load balancing is not supported on proxy RADIUS servers and for private server groups.

Information About RADIUS Server Load Balancing

RADIUS Server Load Balancing Overview

Load balancing distributes batches of transactions to RADIUS servers within a server group. Load balancing assigns each batch of transactions to the server with the lowest number of outstanding transactions in its queue. The process of assigning a batch of transactions is as follows:

1 The first transaction is received for a new batch.

2 All server transaction queues are checked.

3 The server with the lowest number of outstanding transactions is identified.

4 The identified server is assigned the next batch of transactions.

The batch size is a user-configured parameter. Changes in the batch size may impact CPU load and network throughput. As batch size increases, CPU load decreases and network throughput increases. However, if a large batch size is used, all available server resources may not be fully utilized. As batch size decreases, CPU load increases and network throughput decreases.

Note There is no set number for large or small batch sizes. A batch with more than 50 transactions is considered large and a batch with fewer than 25 transactions is considered small.

Note If a server group contains ten or more servers, we recommend that you set a high batch size to reduce CPU load.

Transaction Load Balancing Across RADIUS Server Groups

You can configure load balancing either per-named RADIUS server group or for the global RADIUS server group. The load balancing server group must be referred to as "radius" in the authentication, authorization, and accounting (AAA) method lists. All public servers that are part of the RADIUS server group are then load balanced.

You can configure authentication and accounting to use the same RADIUS server or different servers. In some cases, the same server can be used for preauthentication, authentication, or accounting transactions for a session. The preferred server, which is an internal setting and is set as the default, informs AAA to use the same server for the start and stop record for a session regardless of the server cost. When using the preferred

server setting, ensure that the server that is used for the initial transaction (for example, authentication), the preferred server, is part of any other server group that is used for a subsequent transaction (for example, accounting).

The preferred server is not used if one of the following criteria is true:

- The **load-balance method least-outstanding ignore-preferred-server** command is used.

- The preferred server is dead.

- The preferred server is in quarantine.

- The want server flag has been set, overriding the preferred server setting.

The want server flag, an internal setting, is used when the same server must be used for all stages of a multistage transaction regardless of the server cost. If the want server is not available, the transaction fails.

You can use the **load-balance method least-outstanding ignore-preferred-server** command if you have either of the following configurations:

- Dedicated authentication server and a separate dedicated accounting server

- Network where you can track all call record statistics and call record details, including start and stop records and records that are stored on separate servers

If you have a configuration where authentication servers are a superset of accounting servers, the preferred server is not used.

RADIUS Server Status and Automated Testing

The RADIUS Server Load Balancing feature considers the server status when assigning batches. Transaction batches are sent only to live servers. We recommend that you test the status of all RADIUS load-balanced servers, including low usage servers (for example, backup servers).

Transactions are not sent to a server that is marked dead. A server is marked dead until its timer expires, at which time it moves to quarantine state. A server is in quarantine until it is verified alive by the RADIUS automated tester functionality.

To determine if a server is alive and available to process transactions, the RADIUS automated tester sends a request periodically to the server for a test user ID. If the server returns an Access-Reject message, the server is alive; otherwise the server is either dead or quarantined.

A transaction sent to an unresponsive server is failed over to the next available server before the unresponsive server is marked dead. We recommend that you use the retry reorder mode for failed transactions.

When using the RADIUS automated tester, verify that the authentication, authorization, and accounting (AAA) servers are responding to the test packets that are sent by the network access server (NAS). If the servers are not configured correctly, packets may be dropped and the server erroneously marked dead.

⚠
Caution We recommend that you use a test user that is not defined on the RADIUS server for the RADIUS server automated testing to protect against security issues that may arise if the test user is not correctly configured.

Note	Use the **test aaa group** command to check load-balancing transactions.

How to Configure RADIUS Server Load Balancing

Enabling Load Balancing for a Named RADIUS Server Group

SUMMARY STEPS

1. **enable**
2. **configure terminal**
3. **radius-server host** {*hostname* | *ip-address*} [**test username** *name*] [**auth-port** *number*] [**ignore-auth-port**] [**acct-port** *number*] [**ignore-acct-port**] [**idle-time** *seconds*]
4. **aaa group server radius** *group-name*
5. **load-balance method least-outstanding** [**batch-size** *number*] [**ignore-preferred-server**]
6. **end**

DETAILED STEPS

	Command or Action	Purpose
Step 1	**enable** **Example:** `Device> enable`	Enables privileged EXEC mode. • Enter your password if prompted.
Step 2	**configure terminal** **Example:** `Device# configure terminal`	Enters global configuration mode.
Step 3	**radius-server host** {*hostname*	*ip-address*} [**test username** *name*] [**auth-port** *number*] [**ignore-auth-port**] [**acct-port** *number*] [**ignore-acct-port**] [**idle-time** *seconds*] **Example:** `Device(config)# radius-server host 192.0.2.1 test username test1 idle-time 1`	Enables RADIUS automated testing.
Step 4	**aaa group server radius** *group-name* **Example:** `Device(config)# aaa group server radius rad-sg`	Enters server group configuration mode.

	Command or Action	Purpose
Step 5	**load-balance method least-outstanding** [**batch-size** *number*] [**ignore-preferred-server**] **Example:** `Device(config-sg)# load-balance method least-outstanding batch-size 30`	Enables the least-outstanding load balancing for a named server group.
Step 6	**end** **Example:** `Device(config-sg)# end`	Exits server group configuration mode and enters privileged EXEC mode.

Enabling Load Balancing for a Global RADIUS Server Group

The global RADIUS server group is referred to as "radius" in the authentication, authorization, and accounting (AAA) method lists.

SUMMARY STEPS

1. **enable**
2. **configure terminal**
3. **radius-server host** {*hostname* | *ip-address*} [**test username** *name*] [**auth-port** *number*] [**ignore-auth-port**] [**acct-port** *number*] [**ignore-acct-port**] [**idle-time** *seconds*]
4. **radius-server load-balance method least-outstanding** [**batch-size** *number*] [**ignore-preferred-server**]
5. **load-balance method least-outstanding** [**batch-size** *number*] [**ignore-preferred-server**]
6. **end**

DETAILED STEPS

	Command or Action	Purpose
Step 1	**enable** **Example:** `Device> enable`	Enables privileged EXEC mode. • Enter your password if prompted.
Step 2	**configure terminal** **Example:** `Device# configure terminal`	Enters global configuration mode.

	Command or Action	Purpose
Step 3	**radius-server host** {*hostname*	*ip-address*} [**test username** *name*] [**auth-port** *number*] [**ignore-auth-port**] [**acct-port** *number*] [**ignore-acct-port**] [**idle-time** *seconds*] **Example:** Device(config)# radius-server host 192.0.2.1 test username test1 idle-time 1	Enables RADIUS automated testing.
Step 4	**radius-server load-balance method least-outstanding** [**batch-size** *number*] [**ignore-preferred-server**] **Example:** Device(config)# radius-server load-balance method least-outstanding	Enables the least-outstanding load balancing for the global RADIUS server group and enters server group configuration mode. • The default batch size is 25. The batch size range is from 1 to 2147483647.
Step 5	**load-balance method least-outstanding** [**batch-size** *number*] [**ignore-preferred-server**] **Example:** Device(config-sg)# load-balance method least-outstanding batch-size 5	Enables least-outstanding load balancing for a global named server group.
Step 6	**end** **Example:** Device(config-sg)# end	Exits server group configuration mode and enters privileged EXEC mode.

Troubleshooting RADIUS Server Load Balancing

After configuring the RADIUS Server Load Balancing feature, you can monitor the idle timer, dead timer, and load balancing server selection or verify the server status by using a manual test command.

SUMMARY STEPS

1. Use the **debug aaa test** command to determine when an idle timer or dead timer has expired, when test packets are sent, the status of the server, or to verify the server state.
2. Use the **debug aaa sg-server selection** command to determine the server that is selected for load balancing.
3. Use the **test aaa group** command to manually verify the RADIUS load-balanced server status.

DETAILED STEPS

Step 1 Use the **debug aaa test** command to determine when an idle timer or dead timer has expired, when test packets are sent, the status of the server, or to verify the server state.

The idle timer is used to check the server status and is updated with or without any incoming requests. Monitoring the idle timer helps to determine if there are nonresponsive servers and to keep the RADIUS server status updated to efficiently utilize available resources. For instance, an updated idle timer would help ensure that incoming requests are sent to servers that are alive.

The dead timer is used either to determine that a server is dead or to update a dead server's status appropriately.

Monitoring server selection helps to determine how often the server selection changes. Server selection is effective in analyzing if there are any bottlenecks, a large number of queued requests, or if only specific servers are processing incoming requests.

The following sample output from the **debug aaa test** command shows when the idle timer expired:

Example:
```
Device# debug aaa test

Jul 16 00:07:01: AAA/SG/TEST: Server (192.0.2.245:1700,1701) quarantined.
Jul 16 00:07:01: AAA/SG/TEST: Sending test request(s) to server (192.0.2.245:1700,1701)
Jul 16 00:07:01: AAA/SG/TEST: Sending 1 Access-Requests, 1 Accounting-Requests in current batch.
Jul 16 00:07:01: AAA/SG/TEST(Req#: 1): Sending test AAA Access-Request.
Jul 16 00:07:01: AAA/SG/TEST(Req#: 1): Sending test AAA Accounting-Request.
Jul 16 00:07:01: AAA/SG/TEST: Obtained Test response from server (192.0.2.245:1700,1701)
Jul 16 00:07:01: AAA/SG/TEST: Obtained Test response from server (192.0.2.245:1700,1701)
Jul 16 00:07:01: AAA/SG/TEST: Necessary responses received from server (192.0.2.245:1700,1701)
Jul 16 00:07:01: AAA/SG/TEST: Server (192.0.2.245:1700,1701) marked ALIVE. Idle timer set for 60
sec(s).
Jul 16 00:07:01: AAA/SG/TEST: Server (192.0.2.245:1700,1701) removed from quarantine.
```

Step 2   Use the **debug aaa sg-server selection** command to determine the server that is selected for load balancing.

The following sample output from the **debug aaa sg-server selection** command shows five access requests being sent to a server group with a batch size of three:

Example:
```
Device# debug aaa sg-server selection

Jul 16 03:15:05: AAA/SG/SERVER_SELECT: Obtaining least loaded server.
Jul 16 03:15:05: AAA/SG/SERVER_SELECT: [3] transactions remaining in batch. Reusing server.
Jul 16 03:15:05: AAA/SG/SERVER_SELECT: Obtaining least loaded server.
Jul 16 03:15:05: AAA/SG/SERVER_SELECT: [2] transactions remaining in batch. Reusing server.
Jul 16 03:15:05: AAA/SG/SERVER_SELECT: Obtaining least loaded server.
Jul 16 03:15:05: AAA/SG/SERVER_SELECT: [1] transactions remaining in batch. Reusing server.
Jul 16 03:15:05: AAA/SG/SERVER_SELECT: Obtaining least loaded server.
Jul 16 03:15:05: AAA/SG/SERVER_SELECT: No more transactions in batch. Obtaining a new server.
Jul 16 03:15:05: AAA/SG/SERVER_SELECT: Obtaining a new least loaded server.
Jul 16 03:15:05: AAA/SG/SERVER_SELECT: Server[0] load: 3
Jul 16 03:15:05: AAA/SG/SERVER_SELECT: Server[1] load: 0
Jul 16 03:15:05: AAA/SG/SERVER_SELECT: Server[2] load: 0
Jul 16 03:15:05: AAA/SG/SERVER_SELECT: Selected Server[1] with load 0
Jul 16 03:15:05: AAA/SG/SERVER_SELECT: [3] transactions remaining in batch.
Jul 16 03:15:05: AAA/SG/SERVER_SELECT: Obtaining least loaded server.
Jul 16 03:15:05: AAA/SG/SERVER_SELECT: [2] transactions remaining in batch. Reusing server.
```

Step 3   Use the **test aaa group** command to manually verify the RADIUS load-balanced server status.

The following sample output shows the response from a load-balanced RADIUS server that is alive when the username "test" does not match a user profile. The server is verified alive when it issues an Access-Reject response to an authentication, authorization, and accounting (AAA) packet generated using the **test aaa group** command.

Example:
```
Device# test aaa group SG1 test lab new-code

00:06:07: RADIUS/ENCODE(00000000):Orig. component type = INVALID
00:06:07: RADIUS/ENCODE(00000000): dropping service type, "radius-server attribute 6 on-for-login-auth"
 is off
00:06:07: RADIUS(00000000): Config NAS IP: 192.0.2.4
00:06:07: RADIUS(00000000): sending
00:06:07: RADIUS/ENCODE: Best Local IP-Address 192.0.2.141 for Radius-Server 192.0.2.176
00:06:07: RADIUS(00000000): Send Access-Request to 192.0.2.176:1645 id 1645/1, len 50
00:06:07: RADIUS:  authenticator CA DB F4 9B 7B 66 C8 A9 - D1 99 4E 8E A4 46 99 B4
00:06:07: RADIUS:  User-Password      [2]   18   *
00:06:07: RADIUS:  User-Name          [1]   6    "test"
00:06:07: RADIUS:  NAS-IP-Address     [4]   6    192.0.2.141
00:06:07: RADIUS: Received from id 1645/1 192.0.2.176:1645, Access-Reject, len 44
00:06:07: RADIUS:  authenticator 2F 69 84 3E F0 4E F1 62 - AB B8 75 5B 38 82 49 C3
00:06:07: RADIUS:  Reply-Message      [18]  24
00:06:07: RADIUS:   41 75 74 68 65 6E 74 69 63 61 74 69 6F 6E 20 66  [Authentication f]
00:06:07: RADIUS:   61 69 6C 75 72 65                                [failure]
00:06:07: RADIUS(00000000): Received from id 1645/1
00:06:07: RADIUS/DECODE: Reply-Message fragments, 22, total 22 bytes
```

Configuration Examples for RADIUS Server Load Balancing

Example: Enabling Load Balancing for a Named RADIUS Server Group

The following examples show load balancing enabled for a named RADIUS server group. These examples are shown in three parts: the current configuration of the RADIUS command output, debug output, and authentication, authorization, and accounting (AAA) server status information.

The following sample output shows the relevant RADIUS configuration:

```
Device# show running-config
.
.
.
aaa group server radius server-group1
 server 192.0.2.238 auth-port 2095 acct-port 2096
 server 192.0.2.238 auth-port 2015 acct-port 2016
 load-balance method least-outstanding batch-size 5
!
aaa authentication ppp default group server-group1
aaa accounting network default start-stop group server-group1
.
.
.
```

The lines in the current configuration of the preceding RADIUS command output are defined as follows:

- The **aaa group server radius** command shows the configuration of a server group with two member servers.

- The **load-balance** command enables load balancing for global RADIUS server groups with the batch size specified.

- The **aaa authentication ppp** command authenticates all PPP users using RADIUS.

- The **aaa accounting** command enables sending of all accounting requests to the AAA server when the client is authenticated and then disconnected using the **start-stop** keyword.

The show debug sample output below shows the selection of the preferred server and the processing of requests for the preceding configuration:

```
Device# show debug

*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT(0000002C):No preferred server available.
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT:Obtaining least loaded server.
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT:No more transactions in batch. Obtaining a new
server.
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT:Obtaining a new least loaded server.
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT:Server[0] load:0
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT:Server[1] load:0
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT:Selected Server[0] with load 0
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT:[5] transactions remaining in batch.
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT(0000002C):Server (192.0.2.238:2095,2096) now being
 used as preferred server
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT(0000002D):No preferred server available.
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT:Obtaining least loaded server.
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT:[4] transactions remaining in batch. Reusing
server.
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT(0000002D):Server (192.0.2.238:2095,2096) now being
 used as preferred server
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT(0000002E):No preferred server available.
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT:Obtaining least loaded server.
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT:[3] transactions remaining in batch. Reusing
server.
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT(0000002E):Server (192.0.2.238:2095,2096) now being
 used as preferred server
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT(0000002F):No preferred server available.
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT:Obtaining least loaded server.
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT:[2] transactions remaining in batch. Reusing
server.
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT(0000002F):Server (192.0.2.238:2095,2096) now being
 used as preferred server
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT(00000030):No preferred server available.
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT:Obtaining least loaded server.
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT:[1] transactions remaining in batch. Reusing
server.
*Feb 28 13:51:16.019:AAA/SG/SERVER_SELECT(00000030):Server (192.0.2.238:2095,2096) now being
 used as preferred server
*Feb 28 13:51:16.023:AAA/SG/SERVER_SELECT(00000031):No preferred server available.
*Feb 28 13:51:16.023:AAA/SG/SERVER_SELECT:Obtaining least loaded server.
*Feb 28 13:51:16.023:AAA/SG/SERVER_SELECT:No more transactions in batch. Obtaining a new
server.
*Feb 28 13:51:16.023:AAA/SG/SERVER_SELECT:Obtaining a new least loaded server.
*Feb 28 13:51:16.023:AAA/SG/SERVER_SELECT:Server[1] load:0
*Feb 28 13:51:16.023:AAA/SG/SERVER_SELECT:Server[0] load:5
*Feb 28 13:51:16.023:AAA/SG/SERVER_SELECT:Selected Server[1] with load 0
*Feb 28 13:51:16.023:AAA/SG/SERVER_SELECT:[5] transactions remaining in batch.
*Feb 28 13:51:16.023:AAA/SG/SERVER_SELECT(00000031):Server (192.0.2.238:2015,2016) now being
 used as preferred server
*Feb 28 13:51:16.023:AAA/SG/SERVER_SELECT(00000032):No preferred server available.
*Feb 28 13:51:16.023:AAA/SG/SERVER_SELECT:Obtaining least loaded server.
*Feb 28 13:51:16.023:AAA/SG/SERVER_SELECT:[4] transactions remaining in batch. Reusing
server.
.
.
.
```

The following sample output from the **show aaa servers** command shows the AAA server status for the named RADIUS server group configuration:

The sample output shows the status of two RADIUS servers. Both servers are alive, and no requests have been processed since the counters were cleared 0 minutes ago.

```
Device# show aaa servers

RADIUS:id 8, priority 1, host 192.0.2.238, auth-port 2095, acct-port 2096
```

```
       State:current UP, duration 3781s, previous duration 0s
       Dead:total time 0s, count 0
       Quarantined:No
       Authen:request 0, timeouts 0
               Response:unexpected 0, server error 0, incorrect 0, time 0ms
               Transaction:success 0, failure 0
       Author:request 0, timeouts 0
               Response:unexpected 0, server error 0, incorrect 0, time 0ms
               Transaction:success 0, failure 0
       Account:request 0, timeouts 0
               Response:unexpected 0, server error 0, incorrect 0, time 0ms
               Transaction:success 0, failure 0
       Elapsed time since counters last cleared:0m
RADIUS:id 9, priority 2, host 192.0.2.238, auth-port 2015, acct-port 2016
       State:current UP, duration 3781s, previous duration 0s
       Dead:total time 0s, count 0
       Quarantined:No
       Authen:request 0, timeouts 0
               Response:unexpected 0, server error 0, incorrect 0, time 0ms
               Transaction:success 0, failure 0
       Author:request 0, timeouts 0
               Response:unexpected 0, server error 0, incorrect 0, time 0ms
               Transaction:success 0, failure 0
       Account:request 0, timeouts 0
               Response:unexpected 0, server error 0, incorrect 0, time 0ms
               Transaction:success 0, failure 0
       Elapsed time since counters last cleared:0m
```

Example: Enabling Load Balancing for a Global RADIUS Server Group

The following examples show how to enable load balancing for global RADIUS server groups. These examples are shown in three parts: the current configuration of the RADIUS command output, debug output, and authentication, authorization, and accounting (AAA) server status information. You can use delimiting characters to display relevant parts of the configuration.

The following example shows the relevant RADIUS configuration:

```
Device# show running-config | include radius

aaa authentication ppp default group radius
aaa accounting network default start-stop group radius
radius-server host 192.0.2.238 auth-port 2095 acct-port 2096 key cisco
radius-server host 192.0.2.238 auth-port 2015 acct-port 2016 key cisco
radius-server load-balance method least-outstanding batch-size 5
```

Lines in the current configuration of the preceding RADIUS command output are defined as follows:

- The **aaa authentication ppp** command authenticates all PPP users using RADIUS.

- The **aaa accounting** command enables the sending of all accounting requests to an AAA server when the client is authenticated and then disconnected through use of the **start-stop** keyword.

- The **radius-server host** command defines the IP address of the RADIUS server host with the authorization and accounting ports specified and the authentication and encryption keys identified.

- The **radius-server load-balance** command enables load balancing for global RADIUS server groups with the batch size specified.

The **show debug** sample output below shows the selection of the preferred server and the processing of requests for the configuration:

```
Device# show debug

General OS:
  AAA server group server selection debugging is on
#
```

```
<sending 10 pppoe requests>
Device#
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT(00000014):No preferred server available.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:Obtaining least loaded server.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:No more transactions in batch. Obtaining a new
server.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:Obtaining a new least loaded server.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:Server[0] load:0
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:Server[1] load:0
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:Selected Server[0] with load 0
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:[5] transactions remaining in batch.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT(00000014):Server (192.0.2.238:2095,2096) now being
 used as preferred server
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT(00000015):No preferred server available.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:Obtaining least loaded server.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:[4] transactions remaining in batch. Reusing
server.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT(00000015):Server (192.0.2.238:2095,2096) now being
 used as preferred server
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT(00000016):No preferred server available.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:Obtaining least loaded server.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:[3] transactions remaining in batch. Reusing
server.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT(00000016):Server (192.0.2.238:2095,2096) now being
 used as preferred server
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT(00000017):No preferred server available.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:Obtaining least loaded server.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:[2] transactions remaining in batch. Reusing
server.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT(00000017):Server (192.0.2.238:2095,2096) now being
 used as preferred server
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT(00000018):No preferred server available.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:Obtaining least loaded server.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:[1] transactions remaining in batch. Reusing
server.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT(00000018):Server (192.0.2.238:2095,2096) now being
 used as preferred server
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT(00000019):No preferred server available.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:Obtaining least loaded server.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:No more transactions in batch. Obtaining a new
server.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:Obtaining a new least loaded server.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:Server[1] load:0
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:Server[0] load:5
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:Selected Server[1] with load 0
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT:[5] transactions remaining in batch.
*Feb 28 13:40:32.199:AAA/SG/SERVER_SELECT(00000019):Server (192.0.2.238:2015,2016) now being
 used as preferred server.
```

The following sample output from the **show aaa servers** command shows the AAA server status for the global RADIUS server group configuration:

The sample output shows the status of two RADIUS servers. Both servers are up and successfully processed in the last 2 minutes:

- Five out of six authentication requests

- Five out of five accounting requests

```
Device# show aaa servers

RADIUS:id 4, priority 1, host 192.0.2.238, auth-port 2095, acct-port 2096
    State:current UP, duration 3175s, previous duration 0s
    Dead:total time 0s, count 0
    Quarantined:No
    Authen:request 6, timeouts 1
            Response:unexpected 1, server error 0, incorrect 0, time 1841ms
            Transaction:success 5, failure 0
    Author:request 0, timeouts 0
            Response:unexpected 0, server error 0, incorrect 0, time 0ms
            Transaction:success 0, failure 0
```

```
        Account:request 5, timeouts 0
                Response:unexpected 0, server error 0, incorrect 0, time 3303ms
                Transaction:success 5, failure 0
        Elapsed time since counters last cleared:2m
RADIUS:id 5, priority 2, host 192.0.2.238, auth-port 2015, acct-port 2016
        State:current UP, duration 3175s, previous duration 0s
        Dead:total time 0s, count 0
        Quarantined:No
        Authen:request 6, timeouts 1
                Response:unexpected 1, server error 0, incorrect 0, time 1955ms
                Transaction:success 5, failure 0
        Author:request 0, timeouts 0
                Response:unexpected 0, server error 0, incorrect 0, time 0ms
                Transaction:success 0, failure 0
        Account:request 5, timeouts 0
                Response:unexpected 0, server error 0, incorrect 0, time 3247ms
                Transaction:success 5, failure 0
        Elapsed time since counters last cleared:2m
```

Example: Monitoring Idle Timer

The following example shows idle timer and related server state for load balancing enabled for a named RADIUS server group. The current configuration of the RADIUS command output and debug command output are also displayed.

The following sample output shows the relevant RADIUS configuration:

```
Device# show running-config | include radius

aaa group server radius server-group1
radius-server host 192.0.2.238 auth-port 2095 acct-port 2096 test username junk1 idle-time
 1 key cisco
radius-server host 192.0.2.238 auth-port 2015 acct-port 2016 test username junk1 idle-time
 1 key cisco
radius-server load-balance method least-outstanding batch-size 5
```

The lines in the current configuration of the preceding RADIUS command output are defined as follows:

- The **aaa group server radius** command shows the configuration of a server group.

- The **radius-server host** command defines the IP address of the RADIUS server host with authorization and accounting ports specified and the authentication and encryption key identified.

- The **radius-server load-balance** command enables load balancing for the RADIUS server with the batch size specified.

The **show debug** sample output below shows test requests being sent to servers. The response to the test request sent to the server is received, the server is removed from quarantine as appropriate, the server is marked alive, and then the idle timer is reset.

```
Device# show debug

*Feb 28 13:52:20.835:AAA/SG/TEST:Server (192.0.2.238:2015,2016) quarantined.
*Feb 28 13:52:20.835:AAA/SG/TEST:Sending test request(s) to server (192.0.2.238:2015,2016)
*Feb 28 13:52:20.835:AAA/SG/TEST:Sending 1 Access-Requests, 1 Accounting-Requests in current
 batch.
*Feb 28 13:52:20.835:AAA/SG/TEST(Req#:1):Sending test AAA Access-Request.
*Feb 28 13:52:20.835:AAA/SG/TEST(Req#:1):Sending test AAA Accounting-Request.
*Feb 28 13:52:21.087:AAA/SG/TEST:Obtained Test response from server (192.0.2.238:2015,2016)
*Feb 28 13:52:22.651:AAA/SG/TEST:Obtained Test response from server (192.0.2.238:2015,2016)
*Feb 28 13:52:22.651:AAA/SG/TEST:Necessary responses received from server
(192.0.2.238:2015,2016)
*Feb 28 13:52:22.651:AAA/SG/TEST:Server (192.0.2.238:2015,2016) marked ALIVE. Idle timer
set for 60 secs(s).
*Feb 28 13:52:22.651:AAA/SG/TEST:Server (192.0.2.238:2015,2016) removed from quarantine.
.
```

.
.

Example: Configuring the Preferred Server with the Same Authentication and Authorization Server

The following example shows an authentication server group and an authorization server group that use the same servers 209.165.200.225 and 209.165.200.226. Both server groups have the preferred server flag enabled.

```
aaa group server radius authentication-group
   server 209.165.200.225 key radkey1
   server 209.165.200.226 key radkey2
aaa group server radius accounting-group
   server 209.165.200.225 key radkey1
   server 209.165.200.226 key radkey2
```

When a preferred server is selected for a session, all transactions for that session will continue to use the original preferred server. The servers 209.165.200.225 and 209.165.200.226 are load balanced based on sessions rather than transactions.

Example: Configuring the Preferred Server with Different Authentication and Authorization Servers

The following example shows an authentication server group that uses servers 209.165.200.225 and 209.165.200.226 and an authorization server group that uses servers 209.165.201.1 and 209.165.201.2. Both server groups have the preferred server flag enabled.

```
aaa group server radius authentication-group
   server 209.165.200.225 key radkey1
   server 209.165.200.226 key radkey2
aaa group server radius accounting-group
   server 209.165.201.1 key radkey3
   server 209.165.201.2 key radkey4
```

The authentication server group and the accounting server group do not share any common servers. A preferred server is never found for accounting transactions; therefore, authentication and accounting servers are load-balanced based on transactions. Start and stop records are sent to the same server for a session.

Example: Configuring the Preferred Server with Overlapping Authentication and Authorization Servers

The following example shows an authentication server group that uses servers 209.165.200.225, 209.165.200.226, and 209.165.201.1 and an accounting server group that uses servers 209.165.201.1 and 209.165.201.2. Both server groups have the preferred server flag enabled.

```
aaa group server radius authentication-group
   server 209.165.200.225 key radkey1
   server 209.165.200.226 key radkey2
   server 209.165.201.1 key radkey3
aaa group server radius accounting-group
   server 209.165.201.1 key radkey3
   server 209.165.201.2 key radkey4
```

If all servers have equal transaction processing capability, one-third of all authentication transactions are directed toward the server 209.165.201.1. Therefore, one-third of all accounting transactions are also directed toward the server 209.165.201.1. The remaining two-third of accounting transactions are load balanced equally

between servers 209.165.201.1 and 209.165.201.2. The server 209.165.201.1 receives fewer authentication transactions because the server 209.165.201.1 has outstanding accounting transactions.

Example: Configuring the Preferred Server with Authentication Servers As a Subset of Authorization Servers

The following example shows an authentication server group that uses servers 209.165.200.225 and 209.165.200.226 and an authorization server group that uses servers 209.165.200.225, 209.165.200.226, and 209.165.201.1. Both server groups have the preferred server flag enabled.

```
aaa group server radius authentication-group
   server 209.165.200.225 key radkey1
   server 209.165.200.226 key radkey2
aaa group server radius accounting-group
  server 209.165.200.225 key radkey1
   server 209.165.200.226 key radkey2
   server 209.165.201.1 key radkey3
```

One-half of all authentication transactions are sent to the server 209.165.200.225 and the other half to the server 209.165.200.226. Servers 209.165.200.225 and 209.165.200.226 are preferred servers for authentication and accounting transaction. Therefore, there is an equal distribution of authentication and accounting transactions across servers 209.165.200.225 and 209.165.200.226. The server 209.165.201.1 is relatively unused.

Example: Configuring the Preferred Server with Authentication Servers As a Superset of Authorization Servers

The following example shows an authentication server group that uses servers 209.165.200.225, 209.165.200.226, and 209.165.201.1 and an authorization server group that uses servers 209.165.200.225 and 209.165.200.226. Both server groups have the preferred server flag enabled.

```
aaa group server radius authentication-group
   server 209.165.200.225 key radkey1
   server 209.165.200.226 key radkey2
   server 209.165.201.1 key radkey3
aaa group server radius accounting-group
  server 209.165.200.225 key radkey1
   server 209.165.200.226 key radkey2
```

Initially, one-third of authentication transactions are assigned to each server in the authorization server group. As accounting transactions are generated for more sessions, accounting transactions are sent to servers 209.165.200.225 and 209.165.200.226 because the preferred server flag is on. As servers 209.165.200.225 and 209.165.200.226 begin to process more transactions, authentication transactions will start to be sent to server 209.165.201.1. Transaction requests authenticated by server 209.165.201.1 do not have any preferred server setting and are split between servers 209.165.200.225 and 209.165.200.226, which negates the use of the preferred server flag. This configuration should be used cautiously.

Additional References for RADIUS Server Load Balancing

Related Documents

Related Topic	Document Title
Cisco IOS commands	Cisco IOS Master Command List, All Releases
Security commands	• Security Command Reference: Commands A to C • Security Command Reference: Commands D to L • Security Command Reference: Commands M to R • Security Command Reference: Commands S to Z
AAA and RADIUS	*Authentication, Authorization, and Accounting Configuration Guide*
AAA server groups and RADIUS configuration	"Configuring RADIUS" module in the *RADIUS Configuration Guide*
Failover retry reorder mode	"RADIUS Server Reorder on Failure" module in the *RADIUS Configuration Guide*

Technical Assistance

Description	Link
The Cisco Support website provides extensive online resources, including documentation and tools for troubleshooting and resolving technical issues with Cisco products and technologies. To receive security and technical information about your products, you can subscribe to various services, such as the Product Alert Tool (accessed from Field Notices), the Cisco Technical Services Newsletter, and Really Simple Syndication (RSS) Feeds. Access to most tools on the Cisco Support website requires a Cisco.com user ID and password.	http://www.cisco.com/techsupport

Feature Information for RADIUS Server Load Balancing

The following table provides release information about the feature or features described in this module. This table lists only the software release that introduced support for a given feature in a given software release train. Unless noted otherwise, subsequent releases of that software release train also support that feature.

Use Cisco Feature Navigator to find information about platform support and Cisco software image support. To access Cisco Feature Navigator, go to www.cisco.com/go/cfn. An account on Cisco.com is not required.

Table 1: Feature Information for RADIUS Server Load Balancing

Feature Name	Releases	Feature Information
RADIUS Server Load Balancing	Cisco IOS 15.2(1)E	The RADIUS Server Load Balancing feature distributes authentication, authorization, and accounting (AAA) authentication and accounting transactions across servers in a server group. These servers can share the AAA transaction load and thereby respond faster to incoming requests. The following commands were introduced or modified: **debug aaa sg-server selection**, **debug aaa test**, **load-balance (server-group)**, **radius-server host**, **radius-server load-balance**, and **test aaa group**.