



Configuring DHCP Failover

DHCP failover is a protocol designed to allow a backup DHCP server to take over for a main server if the main server is taken off the network for any reason.

Failover Scenarios

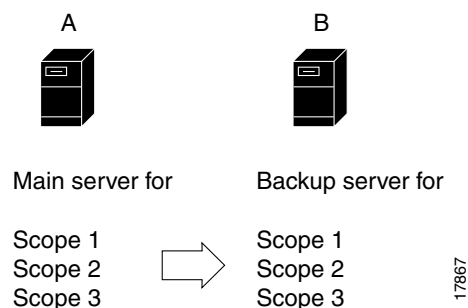
There are three basic failover scenarios:

- Simple failover (recommended)—One server acting as main and its partner acting as backup.
- Back office failover—Two mains having the same backup server.
- Symmetrical failover—Two servers acting as main and backup for each other.

Simple Failover

Simple failover involves a main server and a single backup server pair (see [Figure 26-1](#)). In the example, main server A has three scopes that must be configured identically on backup server B.

Figure 26-1 Simple Failover Example



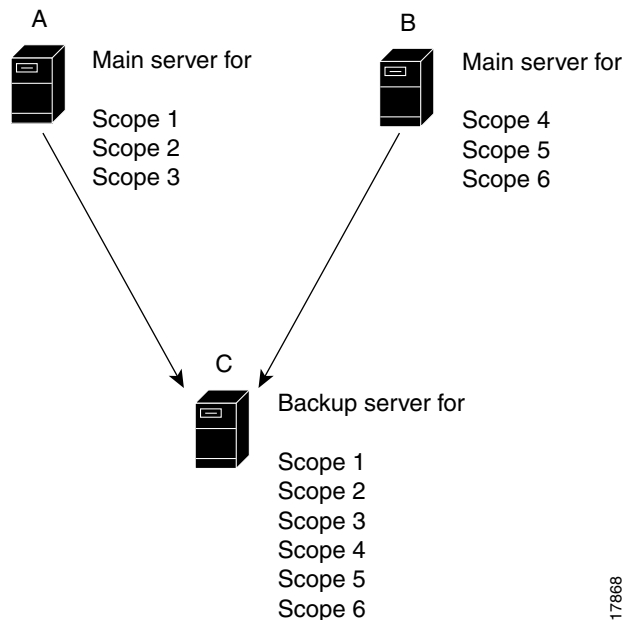
The advantages of simple failover over the other scenarios are:

- It is the easiest to manage as the network changes—It is fully supported by the Web UI so that changes to the main server configuration are automatically propagated to the backup server.
- Provides the greatest performance benefits.
- Having the additional load balancing feature in effect eliminates the need for a back office or symmetrical scenario (see the [“Setting Load Balancing”](#) section on page 26-17).

Back Office Failover

Back office failover involves two (or more) main servers that share the same backup server (see [Figure 26-2](#)). In the example, main servers A and B have different scopes, and backup server C must include all these scopes. This scenario is appropriate for scopes on the same LAN segment, which require the same main and backup servers, but with the sets of scopes on different LAN segments.

Figure 26-2 Back Office Failover Example



An advantage of back office failover over the other scenarios is that it reduces the number of servers managed. However, simple failover is still recommended, because in back office failover:

- The backup server must be sized to handle the sum of the configurations.
- Changes to any of the main servers must be duplicated on the backup server.
- The increased complexity can substantially reduce the actual availability.

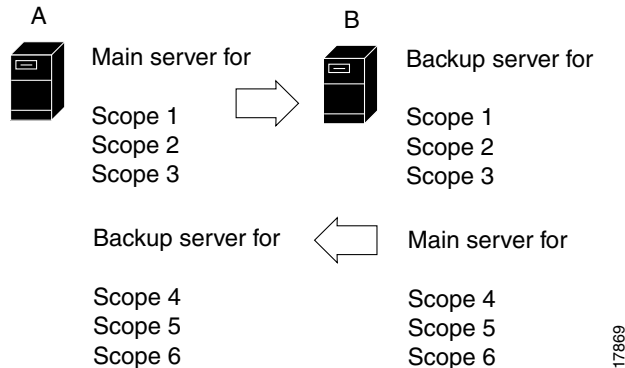
Symmetrical Failover

Symmetrical failover involves servers that act as backups for each other (see [Figure 26-3](#)). This scenario is extremely tricky in that there can be no variance in scope attribute values between the servers, or the relationship will not work properly.

Symmetrical failover used to be a way of load balancing the servers, although this is now more effectively done in a simple failover scenario using the load balancing feature (see the [“Setting Load Balancing”](#) section on page 26-17).

Unfortunately, symmetrical failover provides little to no performance benefit over the simple or back office scenarios. A backup server operates at about 40% of the main server to keep its lease database synchronized. If the servers back each other up, a portion of their processing capacity goes to this task, with less capacity available to service clients. Moreover, because each scope must be configured individually, symmetrical failover is more prone to configuration errors.

Figure 26-3 Symmetrical Failover Example



Failover Checklist

Use this checklist to prepare for an effective failover configuration:

- Duplicate the scope, policy, DHCP option, and address configurations on the partner servers—The Network Registrar Web UI provides a way to automate this process.
- Ensure that both partners are configured with a wide enough range of addresses so that the backup server can provide leases for a reasonable amount of time while the main server is down.
- If you change any of the following configurations on the main server, also change them on the backup server:
 - Scopes, including ensuring identical scope-selection tags
 - Policies
 - IP addresses
 - Reservations
 - Clients
 - Client-classes
 - Dynamic DNS updates
 - Dynamic BOOTP
 - Virtual private networks (VPNs)
 - DHCP extensions
- If you use LDAP, direct the partner servers to the same LDAP server.
- If you use BOOTP relay (IP helpers), configure all BOOTP relay agents to point to both partners. Network Registrar does not automatically detect this. You can only detect BOOTP configuration errors by performing live tests in which you periodically take the main server out of service to verify that the backup server is available to DHCP clients.

Creating and Synchronizing Failover Server Pairs

A failover pair consists of a main and backup DHCP server. You can create failover pairs in the local and regional cluster Web UIs, and you can synchronize the main and backup servers.

Adding Failover Pairs

You first create the DHCP failover pair based on cluster main and backup servers. You then synchronize the failover pair so that the scopes, policies, clients, extensions, and other DHCP properties match between the servers.

Step 1 In the Web UI, click **DHCP**, then **Failover** to open the List DHCP Failover Pairs page.

Step 2 Click **Add DHCP Failover Pair**.

Step 3 On the Add DHCP Failover Pair page, add a failover pair name. This can be any distinguishing name and is required.

In the CLI, use **failover-pair name create main-server-address backup-server-address**. If you want to synchronize between the servers, you must also specify the main and backup clusters by setting the *main* and *backup* attributes. For example:


```
nrcmd> failover-pair example-fo-pair create 192.168.50.1 192.168.60.1 main=Example-cluster backup=Boston-cluster
```

Step 4 In the Web UI, proceed stepwise:

- a. Choose the cluster for the main DHCP server. This can be localhost or some other cluster you define. Whatever you select here becomes the IP address value for the *main-server* attribute once you add the failover pair.
 - b. Choose the cluster for the backup DHCP server. This cannot be the same as the main server cluster, but it must be localhost if the main cluster is not localhost. Whatever you select here becomes the IP address value for the *backup-server* attribute once you add the failover pair.
 - c. For each address block to include in the failover configuration, enter its IP address and mask, then click **Add Address Block**. (To create address blocks, see the [“Adding Address Blocks” section on page 8-5](#).)
 - d. You should not change the IP address values of the *main-server* and *backup-server* attributes unless you have multihomed hosts.
 - e. You can set additional attributes, such as the maximum client lead time (*mclt*) or backup percentage (*backup-pct*). Most of the default values are optimized. Leave the *failover* attribute enabled by default unless you want to temporarily disable failover for the pair.
 - f. Click **Add Failover Pair**.
-

Synchronizing Failover Pairs

Once you create the failover pairs, you then synchronize the servers.

- Step 1** In the Web UI, click the Report icon () to open the Report Synchronize Failover Pair page.
- Step 2** Choose the synchronization operation, depending on the degree to which you want the main server's property values to replace those of the backup server. There are three basic operations:
- **Update**—This is the default and least radical operation. It is appropriate for update synchronizations in that it has the least effect on the unique properties of the backup server.
 - **Complete**—This operation is appropriate for all initial synchronizations. It is more complete than an update operation, while still preserving many of the backup server's unique properties, such as are required for back office failover configurations.
 - **Exact**—This operation is appropriate for initial simple and symmetrical failover configurations, and is not appropriate for back office configurations. It makes the two servers as much as possible mirror images of each other, although it retains unique DHCP server, LDAP event services, and extension points on the backup server. (For initial failover configurations, use the Exact or Complete operation.)

Each operation performs a different mix of functions on the failover properties, as described in [Table 26-1](#). There are four functions, with examples based on these property name-value pairs:

On the main server:	On the backup server:
Name1=A	Name2=B
Name2=C	Name3=D

- **no change**—Makes no change to the list of properties or their values on the backup server. For the example, the result would be Name2=B, Name3=D.
- **ensure**—Ensures that a copy of the main server property exists on the backup server, but does not replace its value. For the example, the result would be Name1=A, Name2=B, Name3=D.
- **replace**—Replaces the value of a property that the two servers have in common with that of the main server. For the example, the result would be Name1=A, Name2=C, Name3=D.
- **exact**—Puts an exact copy of the main server's list of properties and values on the backup server and removes the unique ones. For the example, the result would be Name1=A, Name2=C.

Table 26-1 Synchronization Functions Based on Update, Complete, or Exact Operations

Data Description	Update	Complete	Exact
DHCP Server (server level failover pair):	replace	replace	replace
Client-Class Properties			
Client Hostname Properties			
DNS Update Properties			
Failover Tuning Properties			
(See Table 26-2 on page 26-6 for a list of the failover pair attributes affected by failover synchronization)			
All other properties	no change	replace	replace
LDAP Event Service	no change	replace	replace

Table 26-1 Synchronization Functions Based on Update, Complete, or Exact Operations (continued)

Data Description	Update	Complete	Exact
Policy:			
Option List Properties	ensure	replace	exact
Packet Boot File Properties	ensure	replace	exact
All other properties	replace	replace	exact
Client	replace	replace	exact
Client-Class	replace	replace	exact
Scopes (related to failover pair)	exact	exact	exact
DNS Update Configuration	replace	replace	exact
Trap Configuration	ensure	replace	exact
VPN	replace	replace	exact
Key	replace	replace	exact
Extensions (You must copy extension files.)	ensure	replace	exact
Extension Point	replace	replace	replace
Option Information:			
Custom options list			
Vendor options list			

Table 26-2 Failover Pair Attributes Affected by Failover Synchronization

Affected Failover Pair Attributes
<i>failover-bulking</i>
<i>failover-poll-interval</i>
<i>failover-poll-timeout</i>
<i>recover</i>

Step 3 Click **Run** on the Run Synchronize Failover page, or **Report** on the Report Synchronize Failover page:

- If you click **Run** and if the connection was accepted, the resulting View DHCP Failover Pair Sync Report page shows what change entries the synchronization added.
- If you click **Report**, the resulting View DHCP Failover Pair Sync Report page shows what change entries the synchronization will apply if you run the synchronization. A **Run Update**, **Run Complete**, or **Run Exact** button indicates what kind of synchronization you want to perform. Click the applicable button to run the synchronization.



In the CLI, use **failover-pair name sync {update | complete | exact}**:

```
nrcmd> failover-pair example-fo-pair sync exact
```

Step 4 On the List DHCP Failover Pairs page, click the View icon (🔍) in the Manage Servers column to open the Manage DHCP Failover Servers page.

Step 5 Click the Reload icon (🔄) next to the backup server to reload the backup server.







Step 6 Try to get a lease.

- Step 7** On the Manage DHCP Failover Servers page, look at the health of the servers (they should show as ). Also, click the Logs icon () to view the log entries on the Log for Server page, and ensure that the servers are in NORMAL failover mode. The log file should contain an item similar to the following:


```
06/19/2003 9:41:19 name/dhcp/1 Info Configuration 0 04092 Failover is enabled
server-wide. Main server name: '192.168.0.1', backup server name: '192.168.0.110', mclt =
3600, backup-pct = 10, dynamic-bootp-backup-pct = 0, use-safe-period: disabled,
safe-period = 0.
```

Restarting the Failover Servers

For any failover synchronization to take effect, you must first connect to, and restart, both the main and backup failover servers.

- Step 1** On the List Failover Pairs page, click the Go Local icon () in the Main DHCP Server column.
- Step 2** On the local Manage DHCP Server page for the main server, click the Reload icon () on the right-hand side of the page.
- Step 3** Click the Go Regional icon () at the top-right corner of the page.
- Step 4** On the regional List Failover Pairs page, click the Go Local icon () in the Backup DHCP Server column.
- Step 5** On the local Manage DHCP Server page for the backup server, click the Reload icon () on the right-hand side of the page.
- Step 6** Click the Go Regional icon () at the top-right corner of the page.

Confirming Failover

- Step 1** Ping from one server to the other to verify TCP/IP connectivity. Make sure that routers are configured to forward clients to both servers.
- Step 2** Check that the server is in NORMAL mode by clicking the Related Servers icon () on the Manage DHCP Server or List DHCP Failover Pairs page in the Web UI, or use `dhcp getRelatedServers` in the CLI.
- Step 3** After startup, have a client attempt to get a lease.
- Step 4** Set the log settings on the main server to include at least *failover-detail*.
- Step 5** Confirm that the name_dhcp_1_log log file (in *install-path/logs*) on the main server contains DHCPBNDACK or DHCPBNDUPD messages from each server.
- Step 6** Confirm that the name_dhcp_1_log log file on the backup server contains messages that the backup server is dropping requests because failover is in NORMAL state.
- Step 7** Repeat [Step 2](#).

State Transitions During Integration

During normal operation, the failover partners transition between states. They stay in their current state until all the actions for the state transition are completed and, if communication fails, until the conditions for the next state are fulfilled. Table 26-3 describes what happens when servers enter various states and how they initially integrate and later reintegrate with each other under certain conditions.

Table 26-3 Failover State Transitions and Integration Processes

Integration	Results
Into NORMAL state, the first time the backup server contacts the main server	<ol style="list-style-type: none"> 1. The newly configured backup server contacts the main server, which starts in PARTNER-DOWN state. 2. Because the backup server is a new partner, it goes into RECOVER state and sends a Binding Request message to the main server. 3. The main server replies with Binding Update messages that include the leases in its lease state database. 4. After the backup server acknowledges these messages, the main server responds with a Binding Complete message. 5. The backup server goes into RECOVER-DONE state. 6. Both servers go into NORMAL state. 7. The backup server sends Pool Request messages. 8. The main server responds with the leases to allocate to the backup server based on the <i>backup-pct</i> configured.
After COMMUNICATIONS-INTERRUPTED state	<ol style="list-style-type: none"> 1. When a server comes back up and connects with a partner in this state, the returning server moves into the same state and then immediately into NORMAL state. 2. The partner also moves into NORMAL state.
After PARTNER-DOWN state	<p>When a server comes back up and connects with a partner in this state, the server compares the time it went down with the time the partner went into this state.</p> <ul style="list-style-type: none"> • If the server finds that it went down and the partner subsequently went into this state: <ol style="list-style-type: none"> a. The returning server moves into RECOVER state and sends an Update Request message to the partner. b. The partner returns all the binding data it was unable to send earlier and follows up with an Update Done message. c. The returning server moves into RECOVER-DONE state. d. Both servers move into NORMAL state.

(continued)

Table 26-3 Failover State Transitions and Integration Processes (continued)

Integration	Results
After PARTNER-DOWN state (continued)	<ul style="list-style-type: none"> • If the returning server finds that it was still operating when the partner went into PARTNER-DOWN state: <ol style="list-style-type: none"> a. The server goes into POTENTIAL-CONFLICT state, which also causes the partner to go into this state. b. The main server sends an update request to the backup server. c. The backup server responds with all unacknowledged updates to the main server and finishes off with an Update Done message. d. The main server moves into NORMAL state. e. The backup server sends the main server an Update Request message requesting all unacknowledged updates. (continued) f. The main server sends these updates and finishes off with an Update Done message. g. The backup server goes into NORMAL state.
After the server loses its lease state database	<p>A returning server usually retains its lease state database. However, it can also lose it because of a catastrophic failure or intentional removal.</p> <ol style="list-style-type: none"> 1. When a server with a missing lease database returns with a partner that is in PARTNER-DOWN or COMMUNICATIONS-INTERRUPTED state, the server determines whether the partner ever communicated with it. If not, it assumes to have lost its database, moves into RECOVER state, and sends an Update Request All message to its partner. 2. The partner responds with binding data about every lease in its database and follows up with an Update Done message. 3. The returning server waits the maximum client lead time (MCLT) period, typically one hour, and moves into RECOVER-DONE state. For details on the MCLT, see the “Setting the Maximum Client Lead Time” section on page 26-13. 4. Both servers then move into NORMAL state.
After a lease state database backup restoration	<p>When a returning server has its lease state database restored from backup, and if it reconnects with its partner without additional data, it only requests lease binding data that it has not yet seen. This data may be different from what it expects.</p> <p>(continued)</p>

Table 26-3 Failover State Transitions and Integration Processes (continued)

Integration	Results
After a lease state database backup restoration (continued)	<ol style="list-style-type: none"> 1. In this case, you must configure the returning server with the <i>failover-recover</i> attribute set to the time the backup occurred. 2. The server moves into RECOVER state and requests all its partner's data. The server waits the MCLT period, typically one hour, from when the backup occurred and goes into RECOVER-DONE state. For details on the MCLT, see the “Setting the Maximum Client Lead Time” section on page 26-13. 3. Once the server returns to NORMAL state, you must unset its <i>failover-recover</i> attribute, or set it to zero. <pre>nrcmd> dhcp set failover-recover=0</pre>
After the operational server had failover disabled	<p>If the operating server had failover enabled, disabled, and subsequently reenabled, you must use special considerations when bringing a newly configured backup server into play. The backup server must have no lease state data and must have the <i>failover-recover</i> attribute set to the current time minus the MCLT interval, typically one hour. For details on the MCLT, see the “Setting the Maximum Client Lead Time” section on page 26-13.</p> <ol style="list-style-type: none"> 1. The backup server then knows to request all the lease state data from the main server. Unlike what is described in “After the server loses its lease state database” section of this table, the backup server cannot request this data automatically because it has no record of having ever communicated with the main server. 2. After reconnecting, the backup server goes into RECOVER state, requests all the main server's lease data, and goes into RECOVER-DONE state. 3. Both servers go into NORMAL state. At this point, you must unset the backup server's <i>failover-recover</i> attribute, or set it to zero. <pre>nrcmd> dhcp set failover-recover=0</pre>

Setting Advanced Failover Attributes

The advanced failover properties that are important to set are the following:

- Backup percentage
- Maximum client lead time (MCLT)
- Safe period
- Request and response packet buffers
- Polling attributes
- Network discovery
- Load balancing

Setting Backup Percentages

To keep failover partners operating despite a network partition (when both servers can communicate with clients, but not with each other), allocate more addresses than for a single server. Configure the main server to allocate a percentage of the currently available addresses in each scope to the backup server. This makes these addresses unavailable to the main server. The backup server uses these addresses when it cannot talk to the main server and cannot tell if it is down.



Note

If a Network Registrar failover server receives an update from a Network Registrar DHCP server running prior to Network Registrar 6.0, the unavailable leases do not have a timeout value. In this case, the Network Registrar 6.2 server uses the *unavailable-timeout* value configured in the scope policy or **system_default_policy** policy as the timeout for the unavailable lease. When the lease times out, the policy causes the lease to transition to available in both failover partners.

You can set the percentage of currently available addresses by setting the *backup-pct* attribute on the failover pair or scope (**failover-pair name set backup-pct** or **scope name set backup-pct** in the CLI). Note that setting the backup percentage on the failover pair level sets the value for all scopes not set with that attribute. However, if set at the scope level, the backup percentage overrides the one at the failover pair level. If the *load-balancing* attribute is enabled for the failover pair (**failover-pair name enable load-balancing** in the CLI), the backup percentage is fixed at 50% and any of the backup percentage attributes (on a failover pair or scope) are ignored. (See the “[Load Balancing Compatibility with Earlier Network Registrar Versions](#)” section on page 26-17.)

The backup percentage should be set large enough to allow the backup server to continue serving new clients in the event that the main server fails. The backup percentage is calculated based on the number of available addresses. The default backup percentage is 10%. However, this number can safely be set to a larger value, if extended outages are expected, because the main server periodically reclaims addresses (once per hour) if, in the course of normal leasing activity, the main server's available address pool drops below its predefined percentage. For example, with the default 10% backup percentage, the main server will reclaim addresses if its address pool falls below 90%.

The percentage depends on the new client arrival rate and the network operator's reaction time. The backup server needs enough addresses from each scope to satisfy all new clients requests arriving during the time it does not know if the main server is down. Even during PARTNER-DOWN state, the backup server waits for the maximum client lead time (MCLT) and lease time to expire before reallocating leases. See the “[Setting the Maximum Client Lead Time](#)” section on page 26-13. When these times expire, the backup server offers:

- Leases from its private pool.
- Leases from the main server's pool.
- Expired leases to new clients.

During the day, an operator likely responds within two hours to COMMUNICATIONS-INTERRUPTED state to determine if the main server is working. The backup server then needs enough addresses to support a reasonable upper bound on the number of new clients that could arrive during those two hours.

During off-hours, the arrival rate of previously unknown clients is likely to be less. The operator can usually respond within 12 hours to the same situation. The backup server then needs enough addresses to support a reasonable upper bound on the number of clients that could arrive during those 12 hours.

The number of addresses over which the backup server requires sole control is the greater of the two numbers. You would express this number as a percentage of the currently available (unassigned) addresses in each scope. If you use client-classes, remember that some clients can only use some sets of scopes and not others.

**Note**

During failover, clients can sometimes obtain leases whose expiration times are shorter than the amount configured. This is a normal part of keeping the server partners synchronized. Typically this happens only for the first lease period, or during COMMUNICATIONS-INTERRUPTED state.

Server and Scope Backup Percentages

For all servers or scopes for which you enable failover, you must set the *backup-pct* attribute. This is the number of currently available (unreserved) leases that the backup server can use for allocations to new DHCP clients when the main server is down. You can use the default, which is 10 percent, or specify another value.

**Note**

When failover load balancing is in effect, the main and backup servers actively move available leases between them to maintain the backup percentage of available leases. See the [“Setting Load Balancing” section on page 26-17](#).

BOOTP Backup Percentage

For scopes for which you enable dynamic BOOTP, use the *dynamic-bootp-backup-pct* attribute rather than the *backup-pct* attribute for the failover pair. The *dynamic-bootp-backup-pct* is the percentage of available addresses that the main server should send to the backup server for use with BOOTP clients.

The *dynamic-bootp-backup-pct* is distinct from the *backup-pct* attribute, because if you enable BOOTP on a scope, a server, even in PARTNER-DOWN state, never grants leases on addresses that are available to the other server. Network Registrar does not grant leases because the partner might give them out using dynamic BOOTP, and you can never safely assume that they are available again.

**Note**

You must define the dynamic BOOTP backup percentage on the main server. If you define it on the backup server, Network Registrar ignores it (to enable duplicating configuration through scripts). If you do not define it, Network Registrar uses the default *backup-pct* for the failover pair or scope.

To properly support dynamic BOOTP while using the failover protocol, do this on every LAN segment in which you want BOOTP support:

- Create one scope for dynamic BOOTP.
- Enable BOOTP and dynamic BOOTP.
- Disable DHCP for that scope.

Setting Backup Allocation Boundaries

You can be more specific as to which addresses to allocate to the backup server by using the *failover-backup-allocation-boundary* attribute on the scope. The IP address set as this value is the upper boundary of addresses from which to allocate addresses to a backup server. Only addressees below this boundary are allocated to the backup. If there are none available below this boundary, then the addresses above it, if any, are allocated to the backup. The actual allocation works down from this address, while the normal allocation for DHCP clients works up from the lowest address in the scope.

If you set *failover-backup-allocation-boundary* for the scope, you must also enable the *allocate-first-available* attribute. If *failover-backup-allocation-boundary* is unset or set to zero, then the boundary used is halfway between the first and last addresses in the scope ranges. If there are no available addresses below this boundary, then the first available address is used.

Setting the Maximum Client Lead Time

You can set a property for failover that controls an adjustment to the lease period, the maximum client lead time (MCLT). The MCLT adjusts for a potential period of uncertain connectivity between the servers. It is the maximum time one server can grant (or extend) a lease to a client without first negotiating a longer time with its partner. This time has the following implications:

- Clients may initially (or if the partners are not communicating) only receive leases of the MCLT length. This means that they need to renew leases sooner than they might otherwise without failover. At this renewal, the client should get a full lease time (unless the partners are not communicating).
- If a server enters PARTNER-DOWN state, it must wait until the MCLT after the down time before it can fully take over all leases.
- If a failover recovery occurs where there is uncertainty about what one partner did (such as when it loses its lease database), the partners may have to restrict leasing activity for the MCLT period after they synchronize before they can resume normal failover operations.

The default MCLT is one hour, the optimum for most configurations. As defined by the failover protocol, the lease period given a client can never be more than the MCLT plus the most recently received potential expiration time from the failover partner, or the current time, whichever is later. That is why you sometimes see the initial lease period as only an hour, or an hour longer than expected for renewals. The actual lease time is recalculated when the main server comes back.

The MCLT is necessary because of failover's use of lazy updates. Using lazy updates, the server can issue or renew leases to clients before updating its partner, which it can then do in batches of updates. If the server goes down and cannot communicate the lease information to its partner, the partner may try to reoffer the lease to another client based on what it last knew the expiration to be. The MCLT guarantees that there is an added window of opportunity for the client to renew. The way that a lease offer and renewal works with the MCLT is:

1. The client sends a DHCPDISCOVER to the server, requesting a desired lease period (say, three days). The server responds with a DHCPOFFER with an initial lease period of only the MCLT (one hour by default). The client then requests the MCLT lease period and the server acknowledges it.
2. The server sends its partner a bind update containing the lease expiration for the client as the current time plus the MCLT. The update also includes the potential expiration time as the current time plus the client's desired period plus the MCLT (three days plus an hour). The partner acknowledges the potential expiration, thereby guaranteeing the transaction.
3. When the client sends a renewal request halfway through its lease (in one-half hour), the server acknowledges with the client's desired lease period (three days). The server then updates its partner with the lease expiration as the current time plus the desired lease period (three days), and the potential expiration as the current time plus the desired period and another half of this period ($3 + 1.5 = 4.5$ days). The partner acknowledges this potential expiration of 4.5 days. In this way, the main server tries to have its partner always lead the client in its understanding of the client's lease period so that it can always offer it to the client.

There is no one correct value for the MCLT. There is an explicit trade-off between various factors in choosing one. Most people use the default of one hour effectively and it works well in almost all environments. Here are some of the trade-offs between a short and long MCLT:

- **Short MCLT**—A short MCLT value means that after entering PARTNER-DOWN state, a server only has to wait a short time before it can start allocating its partner's IP addresses to DHCP clients. Furthermore, it only has to wait a short time after a lease expires before it can reallocate that address to another DHCP client. However, the down side is that the initial lease interval that is offered to every new DHCP client will be short, which causes increased traffic, because those clients need to send their first renewal in a half of a short MCLT time. Also, the lease extensions that a server in COMMUNICATIONS-INTERRUPTED state can give is the MCLT only after the server has been in that state for around the desired client lease period. If a server stays in that state for that long, then the leases it hands out will be short, increasing the load on that server, possibly causing difficulty.
- **Long MCLT**—A long MCLT value means that the initial lease period will be longer and the time that a server in COMMUNICATIONS-INTERRUPTED state can extend leases (after it being in that state for around the desired client lease period) will be longer. However, a server entering PARTNER-DOWN state must wait the longer MCLT before being able to allocate its partner's addresses to new DHCP clients. This may mean that additional addresses are required to cover this time period. Also, the server in PARTNER-DOWN state must wait the longer MCLT from every lease expiration before it can reallocate an address to a different DHCP client.

Using the Failover Safe Period to Move Servers into PARTNER-DOWN State

One or both failover partners could potentially move into COMMUNICATIONS-INTERRUPTED state. Fortunately, they cannot issue duplicate addresses while in this state. However, having a server in this state over longer periods is not a good idea, because there are restrictions on what a server can do. The main server cannot reallocate expired leases and the backup server can run out of addresses from its pool. COMMUNICATIONS-INTERRUPTED state was designed for servers to easily survive transient communication failures of a few minutes to a few days. A server might function effectively in this state for only a short time, depending on the client arrival and departure rate. After that, it would be better to move a server into PARTNER-DOWN state so it can completely take over the lease functions until the servers resynchronize.

There are two ways a server can move into PARTNER-DOWN state:

- **User action**—An administrator sets a server into PARTNER-DOWN state based on an accurate assessment of reality. The failover protocol handles this correctly.
- **The failover safe period expires**—When the servers run unattended for longer periods, they need an automatic way to enter PARTNER-DOWN state.

Network operators might not sense in time that a server is down or uncommunicative. Hence, the failover safe period, which provides network operators some time to react to a server moving into COMMUNICATIONS-INTERRUPTED state. During the safe period, the only requirement is that the operators determine that both servers are still running and, if so, fix the network communications failure or take one of the servers down before the safe period expires.

During this safe period, either server allows renewals from any existing client, but there is a major risk of possibly issuing duplicate addresses. This is because one server can suddenly enter PARTNER-DOWN state while the other is still operating. Because of this risk, the failover safe period is disabled by default. That is why it is best to enable the safe period only if, during a server failure, it is more important to get an address than risk receiving a duplicate one.

The length of the safe period is installation-specific, and depends on the number of unallocated addresses in the pool and the expected arrival rate of previously unknown clients requiring addresses. The safe period is typically 24 hours, although many environments can support periods of several days.

The number of extra addresses required for the safe period should be the same as the expected total of new clients a server encounters. This depends on the arrival rate of new clients, not the total outstanding leases. Even if you can only afford a short safe period, because of a dearth of addresses or a high arrival rate of new clients, you can benefit substantially by allowing DHCP to ride through minor problems that are fixable in an hour. There is minimum chance of duplicate address allocation, and reintegration after the solved failure is automatic and requires no operator intervention.

Here are some guidelines to follow, to help you decide whether to use manual intervention or the safe period for transitioning to PARTNER-DOWN state:

- If your corporate policy is to have minimal manual intervention, set the safe period. Enable the failover pair attribute *use-safe-period* to enable the safe period. Then, set the DHCP attribute *safe-period* to set the duration (86400 seconds, or 24 hours, by default). Set this duration long enough so that operations personnel can explore the cause of the communication failure and assure that the partner is truly down. At least 12 hours is recommended.
- If your corporate policy is to avoid conflict under any circumstances, then never let the backup server go into PARTNER-DOWN state unless by explicit command. Allocate sufficient addresses to the backup server so that it can handle new client arrivals during periods when there is no administrative coverage. You can set PARTNER-DOWN in two ways in Network Registrar:
 - On the View Failover Related Server page of the regional cluster Web UI, if the partner is in the Communications-interrupted failover state, you can click **Set Partner Down** in association with an input field for the PARTNER-DOWN date setting. This setting is initialized to the value of the *start-of-communications-interrupted* attribute. (In Normal Web UI mode, you cannot set this date to be an earlier value than the initialized date. In Expert Web UI mode, you can set this value to any date.) After clicking **Set Partner Down**, you return to the List Related Servers for DHCP Server page to view the result of the PARTNER-DOWN action.
 - Use **dhcp setPartnerDown** in the CLI, specifying the name of the partner server. This moves all the scopes running failover with the partner into PARTNER-DOWN state immediately, unless you specify a date and time with the command. This date and time should be when the partner was last known to be operational.

There are two conventions for specifying the date:

- *-num unit* (a time in the past), where *num* is a decimal number and *unit* is *s*, *m*, *h*, *d*, or *w* for seconds, minutes, hours, days or weeks respectively. For example, specify *-3d* for three days.
- Month (name or its first three letters), day, hour (24-hour convention), year (fully specified year or last two digits). This example notifies the backup server that its main server went down at 12 midnight on October 31, 2002:

```
nrcmd> server dhcp setPartnerDown dhcp2.example.com. -3d
nrcmd> server dhcp setPartnerDown dhcp2.example.com. Oct 31 00:00:00 2001
nrcmd> dhcp reload
```



Note

Wherever you specify a date and time in the CLI, enter the time that is local to the **nrcmd** process. If the server is running in a different time zone than this process, disregard the time zone where the server is running and use local time instead.

Setting DHCP Request and Response Packet Buffers

The number of request buffers (set through the *max-dhcp-requests* DHCP attribute) sets the maximum number of simultaneous requests that the server can accept. The default value is 500, which is suitable for most deployments, but can be tuned to the capacity of the server. This capacity relates to the leasing rate and average latency of the leasing transaction. For example, if clients receive new leases every 250 milliseconds, then a request buffer value of 500 is sufficient for the server to respond to 2000 clients per second. (This assumes sufficient processing capacity to service clients at that rate.) A lower value would throttle the server's performance below this capacity. A higher value allows servicing a greater number of clients without retries during periods of burst load, but results in a higher average latency to each client. Average latency under two seconds is sufficient to properly service clients.

The number of response buffers (set through the *max-dhcp-responses* DHCP attribute) sets the maximum number of simultaneous requests that the server can complete by issuing a client response. When the network operates at a steady state, responses should track with the number of requests accepted. Because the same pool of response buffers serves for both lease and failover activity, when failover is enabled, the server adjusts the response buffer value to be at least four times the request buffers. This ensures that sufficient resources are available to process all pending client and failover activity simultaneously.

Changing Polling Attributes

You can change some system defaults, such as the number of leases that the main server should send to the backup server, or the MCLT. See the “[Setting the Maximum Client Lead Time](#)” section on [page 26-13](#). However, you need to change them on both servers.

On each server:

- Change the poll interval (DHCP attribute *failover-poll-interval*)—The interval that partners contact each other to confirm network connectivity. The default is 10 seconds.
- Change the poll timeout (DHCP attribute *failover-poll-timeout*)—Failover partners who cannot communicate for failover polling timeout seconds will conclude that they lost network connectivity, and change their operational states appropriately. The default is 60 seconds.

Generally, you should not have to change the *failover-poll-timeout*. It is intimately linked to the *failover-poll-interval* and is based on real world experience.



Note

To collect subnet utilization history for the failover pair, if you are configuring simple failover, disable individual polling of the main and backup DHCP servers, but enable failover pair polling by setting the failover pair attribute *poll-subnet-util-interval*, so as to collect one set of data from both servers.

Setting the Network Discovery Attribute

If you enable failover on a UNIX system, you could set the *sms-network-discovery* attribute to enable the computing client os-type for leased addresses, which can help if you have a Windows partner server and want to use **dhcp updateSms** in the CLI on it.

Setting Load Balancing

In normal failover mode, the main DHCP server bears most of the burden of servicing clients when the failover partners are in normal communication mode. The main server not only services all new client requests, but has to handle renewal and rebinding requests and expired leases from the backup partner. To distribute the load more evenly between the two servers in a simple failover configuration scenario, Network Registrar introduced the load balancing feature in release 6.2. This feature is based on RFC 3074, “DHCP Load Balancing.”

Failover load balancing allows both servers to actively service clients and determine which unique clients each will serve without running the risk of both servicing the same ones. Failover load balancing only applies while the servers are in NORMAL state; in other states, both servers can respond to clients. Per RFC 3074, the servers calculate a hash value for each request that the server receives based on the client’s identifier option value or hardware address, and the request is serviced if the hash value is assigned to that server. When failover load balancing is enabled for a failover relationship, the servers split the client load evenly (the main server processes 50% of the hash values, the backup the other 50%).

With load balancing, each partner responds to all clients whenever it is not in the normal communications state. When they are in a normal communications state, they respond only to broadcast DHCPDISCOVER messages from clients that are in their assigned hash values. For broadcast DHCPREQUESTs, the server responds only if it is the targeted server (based on the server identifier option). Broadcast BOOTP and DHCPINFORM requests are also load balanced.

Load Balancing Compatibility with Earlier Network Registrar Versions

Failover load balancing is disabled by default to ensure backward compatibility with earlier Network Registrar releases, and is used only if both servers support load balancing. Hence, the failover pair load balancing is unset and the default of disabled applies until you explicitly enable it. If you enable load balancing, each server services about 50% of the clients, and the free leases given to the backup will be 50%, regardless of the configured percentage (see the [“Setting Backup Percentages” section on page 26-11](#)).

Configuring Load Balancing

In the Web UI, when setting the failover properties for the pair (see the [“Creating and Synchronizing Failover Server Pairs” section on page 26-4](#)), enable or disable the *load-balancing* attribute in the Failover Settings attributes as desired to enable or disable failover load balancing. In the CLI, use **failover-pair *name* set load-balancing**.

Changing Failover Server Roles



Caution

Be careful when you change the role of a failover server. Remember that all address states in a scope are lost from a server if it is ever reloaded without that scope in its configuration.

Making Nonfailover Servers Failover Mains

You can update an existing installation and increase the availability of the DHCP service it offers. You can use this procedure only if the original server never participated in failover.

-
- Step 1** Install Network Registrar on the original server and ensure that it operates correctly after the installation.
 - Step 2** Install Network Registrar on the machine that is to be the backup server. Note the machine's DNS name.
 - Step 3** Enable failover on the original server. Use the DNS name of the recently installed backup server. See the [“Simple Failover” section on page 26-1](#).
 - Step 4** Reload the main server. It should go into PARTNER-DOWN state. It cannot locate the backup server, because it is not yet configured. There should be no change in main server operation at this point.
 - Step 5** Duplicate the main server's configuration on the backup server, including scopes (including secondary), policies, and client-classes. If you use client-classes, make sure the clients are entered into each cluster or that each server can access an LDAP database with the client data.
 - Step 6** Enable failover on the backup server. Be sure to define the main server.
 - Step 7** Reconfigure all operational BOOTP relays to forward broadcast packets to the main and backup server.
 - Step 8** Reload the backup server.
-

After you complete these steps:

1. The backup server detects the main server and moves into RECOVER state.
2. The backup server refreshes its stable storage with the main server's lease data and, when complete, moves into RECOVER-DONE state.
3. The main server moves into NORMAL state.
4. The backup server moves into NORMAL state.
5. The backup server uses a pool request to ask the main server for addresses to allocate if communication is interrupted.
6. After allocating these addresses, the main server sends this data to the backup server.

Replacing Servers Having Defective Storage

If a failover server loses its stable storage (hard disk), you can replace the server and have it recover its state information from its partner.

-
- Step 1** Determine which server lost its stable storage.
 - Step 2** Use `dhcp setPartnerDown` in the CLI to tell the other server that its partner is down. If you do not specify a time, the current time is used.
 - Step 3** When the server is again operational, reinstall Network Registrar.
 - Step 4** Duplicate the server configuration from its partner. However, do not recover any lease databases from an earlier backup or the partner's system.
 - Step 5** Reload the replacement server.
-

After you complete these steps:

1. The recovered server moves into RECOVER state.
2. Its partner sends it all its data.
3. The server moves into RECOVER-DONE state when it reaches its maximum client lead time (and any time set for *failover-recover*).
4. Its partner moves into NORMAL state.
5. The recovered server moves into NORMAL state. It can request addresses, but can allocate few new ones, because its partner already sent it all its previously allocated addresses.

Removing Backup Servers and Halting Failover Operation

Sometimes you might need to remove the backup server and halt all failover operations.

-
- Step 1** On the backup server, remove all the scopes that were designated as a backup to the main server.
- Step 2** On the main server, remove the failover capability from those scopes that were main for the backup server, or disable failover server-wide if that is how it was configured.
- Step 3** Reload both servers.
-

Adding Main Servers to Existing Backup Servers

You can use an existing backup server for a main server.

-
- Step 1** Duplicate the main server's scopes, policies, and other configurations on the backup server.
- Step 2** Configure the main server to enable failover and point to the backup server.
- Step 3** Configure the backup server to enable failover for the new scopes that point to the new main server.
- Step 4** Reload both servers. Network Registrar performs the same steps as those described in the [“Making Nonfailover Servers Failover Mains”](#) section on page 26-18.
-

Configuring Failover on Multiple Interface Hosts

If you plan to use failover on a server host with multiple interfaces, you must explicitly configure the local server's name or address. This requires an additional command. For example, if you have a host with two interfaces, serverA and serverB, and you want to make serverA the a main failover server, you must define serverA as the failover-main-server before you set the backup server name (external serverB). If you do not do this, failover might not initialize correctly and tries to use the wrong interface.

Set the DHCP server properties *failover-main-server* and *failover-backup-server*.

With multiple interfaces on one host, you must specify a host name that points to only one address or A record. You cannot set up your servers for round-robin support.

Supporting BOOTP Clients in Failover

You can configure scopes to support two types of BOOTP clients—static and dynamic.

Static BOOTP

You can support static BOOTP clients using DHCP reservations. When you enable failover, remember to configure both the main and the backup server with identical reservations.

Dynamic BOOTP

You can enable dynamic BOOTP clients by enabling the *dynamic-bootp* attribute on a scope. When using failover, however, there are additional restrictions on address usage in such scopes, because BOOTP clients get permanent addresses and leases that never expire.

When a server whose scope does not have the *dynamic-bootp* option enabled goes to PARTNER-DOWN state, it can allocate any available (unassigned) address from that scope, whether or not it was initially available to any partner. However, when the *dynamic-bootp* option is set, each partner can only allocate its own addresses. Consequently, scopes that enable the *dynamic-bootp* option require more addresses to support failover.

When using dynamic BOOTP:

- Segregate dynamic BOOTP clients to a single scope. Disable DHCP clients from using that scope by disabling the *dhcp* attribute on the scope.
- Set the *dynamic-bootp-backup-pct* failover pair attribute to allocate a greater percentage of addresses to the backup server for this scope, as much as 50 percent higher than a regular backup percentage.

Configuring BOOTP Relays

The Network Registrar failover protocol works with BOOTP relay (also called IP helper), a router capability that supports DHCP clients that are not locally connected to a server.

If you use BOOTP relay, ensure that the implementations point to both the main and backup servers. If they do not and the main fails, clients are not serviced, because the backup cannot see the required packets. If you cannot configure BOOTP relay to forward broadcast packets to two different servers, configure the router to forward the packets to a subnet-local broadcast address for a LAN segment, which could contain both the main and backup servers. Then, ensure that both the main and backup servers are on the same LAN segment.

DHCPLEASEQUERY and Failover

To accommodate DHCPLEASEQUERY messages sent to a DHCP failover backup server when the master server is down, the master server must communicate the *relay-agent-info* (82) option values to its partner server. To accomplish this, the master server uses DHCP failover update messages.


Troubleshooting Failover

This section describes how to avoid failover configuration mistakes, monitor failover operations, and detect and handle network problems.

Monitoring Failover Operations

You can examine the DHCP server log files on both partner servers to verify your failover configuration.

You can make a few important log and debug settings to troubleshoot failover. Set the DHCP log settings to *failover-detail* to track the number and details of failover messages logged. To ensure that previous messages do not get overwritten, add the *failover-detail* attribute to the end of the list. Use the *no-failover-conflict* attribute to inhibit logging server failover conflicts, or the *no-failover-activity* attribute to inhibit logging normal server failover activity. Then, reload the server.

You can also isolate misconfigurations more easily by clicking the Related Servers icon () on the Manage DHCP Server or List DHCP Failover Pairs page in the Web UI, or by using `dhcp getRelatedServers` in the CLI.

Detecting and Handling Network Failures

Table 26-4 describes some symptoms, causes, and solutions for failover problems.

Table 26-4 Detecting and Handling Failures

Symptom	Cause	Solution
New clients cannot get addresses	A backup server is in COMMUNICATIONS-INTERRUPTED state with too few addresses	Increase the backup percentage on the main server.
Error messages about mismatched scopes	There are mismatched scope configurations between partners	Reconfigure your servers.
Log messages about failure to communicate with partner	Server cannot communicate with its partner	Check the status of the server.
Main server fails. Some clients cannot renew or rebind leases. The leases expire even when the backup server is up and possibly processing some client requests.	Some BOOTP relay (ip-helper) was not configured to point at both servers; see the “Configuring BOOTP Relays” section on page 26-20 .	<ul style="list-style-type: none"> Reconfigure BOOTP relays to point at both main and backup server Run a fire drill test—Take the main server down for a day or so and see if your user community can get and renew leases
SNMP trap: other server not responding	Server cannot communicate with its partner	Check the status of the server.
SNMP trap: dhcp failover configuration mismatch	Mismatched scope configurations between partners	Reconfigure your servers.

Table 26-4 *Detecting and Handling Failures (continued)*

Symptom	Cause	Solution
Users complain that they cannot use services or system as expected	Mismatched policies and client-classes between partners	Reconfigure partners to have identical policies; possibly use LDAP for client registration if currently registering clients directly in partners.