Managing DHCP Failover

Cisco Prime Network Registrar failover protocol is designed to allow a backup DHCP server to take over for a main server if the main server is taken offline for any reason. Prior to 8.2, this protocol was UDP based, only operated over IPv4, and only supported DHCPv4. Starting with 8.2, this protocol is TCP based, can be configured to use either IPv4 or IPv6, and supports both DHCPv4 and DHCPv6 over a single connection. With 9.0, this protocol now will try both IPv4 and IPv6 transports if configured to use both, and will use whichever connection comes up first.The DHCP failover supports the following features:

- DHCPv4 addresses
- DHCPv6 addresses (non-temporary and temporary)
- DHCPv6 prefix delegation

DHCP failover is not applicable to DHCPv4 subnet allocation (on-demand address pools).

- How DHCP Failover Works, on page 1
- DHCP Simple Failover, on page 2
- DHCPv6 Failover, on page 2
- Setting Up Failover Server Pairs, on page 3
- Configuring Failover Parameters Based on Your Scenario, on page 10
- Recovering from a DHCP Failover, on page 16
- Setting Advanced Failover Attributes, on page 21
- Maintaining Failover Server Pair, on page 22
- Recovering Failover Configuration, on page 22
- Using PARTNER-DOWN State to Allow a Failover Server to Operate for Extended Periods without Its Failover Partner, on page 23
- Restoring a Standalone DHCP Failover Server - Tutorial , on page 24
- Changing Failover Server Roles, on page 30
- Troubleshooting Failover, on page 32
- Supporting BOOTP Clients in Failover, on page 34

How DHCP Failover Works

DHCP failover is based on a server-partner relationship. The partner must have identical DHCPv4 scopes, DHCPv6 prefixes, DHCPv6 links, reservations, policies, and client-classes, as the server. After the servers start up, they contact each other. The main server provides its partner with a DHCPv4 addresses and DHCPv6 delegated prefixes, and updates its partner with every client operation. If the main server fails, then the partner takes over offering and renewing leases, using its DHCPv4 addresses and DHCPv6 delegated prefixes. When

the main server becomes operational again, it re-integrates with its partner without administrative intervention. These servers are in a relationship known as a failover pair.

The failover protocol keeps DHCP operational, if:

- **The main server fails**—The partner takes over services during the time the main server is down. The servers cannot generate duplicate addresses, even if the main server fails before updating its partner.
- **Communication fails**—A partner can operate correctly even though it cannot tell whether it was the other server or the communication with it that failed. The servers cannot issue duplicate addresses, even if they are both running and each can communicate with only a subset of clients.

After a failover pair is configured:

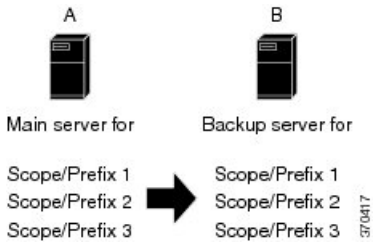
1. The partners connect.
2. The main server supplies data about all existing leases to its partner.
3. The backup server requests a pool of backup addresses from the main server.
4. The main server replies with a percentage of available addresses from each scope or prefix to its partner.
5. The backup server ignores all DHCPDISCOVER and Solicit requests, unless it senses that the main server is down or load balancing is enabled for the failover pair. In normal operation, the backup server handles only some renewal and rebinding requests.
6. The main server updates its partner with the results of all client operations.

You can automatically synchronize the configuration of the servers in a failover pair. The two servers dynamically rebalance the available leases; if the main server hands out a large portion of its available leases, it can reclaim leases from its partner.

DHCP Simple Failover

Starting from release 8.2, Cisco Prime Network Registrar supports only simple failover configuration. Simple failover involves a single main server and a single backup server pair (see the image below). In the example, main server A has three scopes or prefixes that must be configured identically on backup server B.

Figure 1: Simple Failover Example



DHCPv6 Failover

DHCPv6 failover works very similar to the DHCPv4 simple failover configuration. The DHCPv6 failover partners keep each other updated on stateful address and delegated prefix leases that are granted, perform synchronization when communication is restored, and generally follow and adhere to the DHCPv4 failover protocol requirements (except the differences between DHCPv4 and DHCPv6).

The maximum client lead time (MCLT) and lease time restrictions are applied to DHCPv6 leases and both the valid lifetime and preferred lifetime of leases are limited to MCLT defined for the failover pair. Only

when the longest lease time allowed by the failover pair exceeds the configured preferred lifetime and if the configured preferred lifetime is less than the configured valid lifetime, the preferred lifetime and valid lifetime of the lease can be different. The delegated prefixes are managed and balanced similar to DHCPv4 addresses.

The most significant difference is that the DHCPv6 failover servers do not balance the available addresses on each prefix but instead use an algorithm to determine which new addresses each server can lease. The algorithm uses the least significant bit of the address and the main server assigns odd addresses whereas the backup server assigns even addresses. This applies to client requested and randomly generated addresses and is not applicable if:

- A lease is already assigned to the client.
- A reservation exists for the client.
- The allocation-algorithms interface-identifier is set and is used. In this case, the interface-identifier (EUI-64) bit is assumed to be unique, and as the global bit is set, these addresses do not conflict with randomly generated addresses as these never have the global bit set.
- Client reservations are configured on the prefix.
- An extension supplies the address.

Setting Up Failover Server Pairs

You can create and synchronize failover pairs at the local and regional clusters.

A failover pair has two main elements, its configuration and the state information that the servers maintain. The key configuration attributes are the name of the failover pair, the role of the local server (main or backup), and the address of the partner. The failover state is defined when you reload the server and the server processes this state data at startup.

Note Cisco Prime Network Registrar 8.2 or later DHCP failover does not interoperate with Cisco Prime Network Registrar 8.1 or earlier releases DHCP failover. You must upgrade both the main and the backup servers in the same maintenance window.

Related Topics

Adding Failover Pairs, on page 3

Synchronizing Failover Pairs, on page 7

Failover Checklist, on page 9

Adding Failover Pairs

Create the DHCP failover pair based on the cluster of the main and backup servers. Then synchronize the configuration of the failover pair so that the scopes, prefixes, policies, and other DHCP properties match between the servers.

To add a failover pair:

Local and Regional Web UI

Step 1 From the **Deploy** menu, choose **Failover Pairs** under the **DHCP** submenu to open the List/Add DHCP Failover Pairs page.

Step 2 Click the **Add Failover Pair** icon in the **Failover Pairs** pane.

Step 3 On the Add DHCP Failover Pair dialog box, add a failover pair name.

This is required and can be any distinguishing name. (See Changing Failover Pair Names, on page 22.)

Step 4 Choose the cluster for the main DHCP server. This can be localhost or some other cluster you define.

Step 5 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.

Step 6 Click **Add DHCP Failover Pair**. The failover pair is created.

Step 7 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.

You may specify the *main-server*, *backup-server*, *main-ip6address*, or *backup-ip6address* attributes if you want to override the values that are configured for the main and backup cluster objects, or if you want to disable a specific transport (by specifying 0.0.0.0 or 0::0 for the addresses). If both IPv4 and IPv6 addresses are available, failover will attempt to connect on both transports and use the connection that comes up first. Click **Save** to save these additional changes.

Following attributes can be configured on the Edit DHCP Failover Pair page (in Advanced mode):

Table 1: Failover Pair Attributes

Attribute	Description
Main Server (*main*)	Identifies the cluster with the main server for a failover pair.
Backup Server (*backup*)	Identifies the cluster that contains the backup server for a failover pair.
Scope Template (*scopetemplate*)	Associates a scope template with a specified failover pair.
Failover Settings	
failover	Enables failover configuration. If you disable this attribute, you turn off failover on attached subnets without changing configuration fundamentals.
mclt	Sets the maximum client lead time in seconds. This attribute controls how far ahead of the backup server that you can make the client lease expiration. You must define this value on both the main and backup servers, and make sure the value is identical on both servers.

Attribute	Description
backup-pct	Controls the percentage of available addresses that the main server sends to the backup server. Set this value on the main server. If it is set on a backup server, it is ignored (to enable copying of configurations). Unless you explicitly set this value on a scope and you disable load balancing, the value set here becomes the default value.
dynamic-bootp-backup-pct	Determines the percentage of available addresses that the main server sends to the backup server for scopes on which dynamic BOOTP is enabled. If defined, it must be defined on the main server. If it is defined in a backup server, it is ignored (to enable copying of configurations). If it is not defined at all or the value is 0, the *backup-pct* is used instead. This parameter is separate from "backup-pct" because if dynamic BOOTP is enabled on a scope, a server will never, even in PARTNER-DOWN state, grant leases on addresses that are available to the other server because they can never safely be assumed to be available again. The MCLT has no meaning for dynamic BOOTP leases.
load-balancing	Determines whether load balancing (RFC 3074) is enabled on a failover pair. The default is disabled. When enabled, the *backup-pct* is ignored and the main and backup server evenly split the client load and available leases for all scopes in the failover relationship (that is, as if backup-pct were configured at 50%).
safe-period	Controls the safe period, in seconds. It does not have to be the same on both main and backup servers. It only has meaning if *use-safe-period* is enabled.
use-safe-period	Controls whether a server can enter PARTNER-DOWN state without an operator command. If disabled, a server never enters PARTNER-DOWN without an operator command.
Failover Server Addresses	
main-server	Controls the IPv4 address used for the failover protocol on the main server. If this value is unset, the address specified for the main cluster is used. Cisco recommends setting this value only if the server is configured with different interfaces for configuration management and clients requests. This value may be set to 0.0.0.0 to disable use of IPv4 for failover communication. If both IPv4 and IPv6 addresses are available, the servers will try both transports for the TCP connection and use whichever comes up first.

Attribute	Description
backup-server	Controls the IPv4 address used for the failover protocol on the backup server. If this value is unset, the address specified for the backup cluster is used. Cisco recommends setting this value only if the server is configured with different interfaces for configuration management and clients requests. This value may be set to 0.0.0.0 to disable use of IPv4 for failover communication. If both IPv4 and IPv6 addresses are available, the servers will try both transports for the TCP connection and use whichever comes up first.
main-ip6address	Controls the IPv6 address used for the failover protocol on the main server. If this value is unset, the address specified for the main cluster is used. Cisco recommends setting this value only if the server is configured with different addresses for configuration management and clients requests. This value may be set to 0::0 to disable use of IPv6 for failover communication. If both IPv4 and IPv6 addresses are available, the servers will try both transports for the TCP connection and use whichever comes up first.
backup-ip6address	Controls the IPv6 address used for the failover protocol on the backup server. If this value is unset, the address specified for the backup cluster is used. Cisco recommends setting this value only if the server is configured with different addresses for configuration management and clients requests. This value may be set to 0::0 to disable use of IPv6 for failover communication. If both IPv4 and IPv6 addresses are available, the servers will try both transports for the TCP connection and use whichever comes up first.

CLI Commands

Use **failover-pair** *name* **create** *main-cluster backup-cluster* [*attribute=value* ...]. For example:

```
nrcmd> failover-pair example-fo-pair create Example-cluster Boston-cluster
```

Related Topics

Failover Checklist, on page 9
Changing Failover Pair Names, on page 22
Synchronizing Failover Pairs, on page 7
Restarting the Failover Servers, on page 22

Synchronizing Failover Pairs

Once you create the failover pairs, you must synchronize the failover pair configuration.

Web UI

Step 1 From the **Deploy** menu, choose **Failover Pairs** under the **DHCP** submenu to open the List/Add DHCP Failover Pairs page.

Step 2 Select the failover pair on the Failover pane.

Step 3 On the List/Add DHCP Failover Pairs page, click the **Synchronize DHCP Failover Pair** tab.

For synchronization in the regional web UI, see the *"Managing DHCP Failover Pairs" section in Cisco Prime Network Registrar 9.0 Administrator Guide*.

Step 4 Choose the direction of synchronization. The direction of synchronization can be either from main to backup server or from backup to main server.

Step 5 Choose the synchronization operation, depending on the degree to which you want the main server objects to replace those of the backup server. The following are the basic synchronization operations that can be performed on the servers:

- **Update operation**—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 operation**—This operation is appropriate for all initial synchronizations. It is more complete than an update operation, while still preserving many of the backup server unique properties.
- **Exact operation**—This operation is appropriate for simple failover configuration.

It makes the two servers mirror images of each other, as much as possible, although this operation retains the unique DHCP server, and extension points on the backup server.

Note For initial failover configurations, use the Exact or Complete operation.

For a better understanding of the functions that are performed on the classes of the objects, consider the following example. Here, we have a main server and its backup server with the following objects:

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

Note In this example, we consider failover synchronization from the main server to the backup server.

Each operation performs a different mix of functions on the classes of objects. The following are the four functions that are performed on the objects based on the operation selected.

- no change—Makes no change to the list of properties or their values on the backup server.

  For example, the result would be Name2=B, Name3=D.

- ensure—Ensures that a copy of the main server object exists on the backup. The target server objects with the same name as main server objects are left unchanged, the objects that are not on the target server are added to it, and the objects only on the target server are left unchanged.

  For example, the result would be Name1=A, Name2=B, Name3=D.

- replace—Ensures that the existing object in the target server is replaced by the main server object of the same name. Also the objects that are not on the target server are added to it and the objects only on the target server are left unchanged. The only exceptions to this are for policies and option definition sets, where the option lists are extracted to compare the list entries.

  For example, the result would be Name1=A, Name2=C, Name3=D.

  Note After deleting the client on the main server and performing the failover synchronization Update or Complete operation to remove the entry on the backup, the client is not removed from the backup. The only failover synchronization operation that will delete the client entry on the backup, after it is removed from the main server, is the failover synchronization Exact operation.

- exact—Puts an exact copy of the main server object on the backup server and removes the unique ones. That is, the objects of target server are made identical to the objects of main server.

  For example, the result would be Name1=A, Name2=C.

For more information, see table below. This table provides the information on the functions (no change, ensure, replace, or exact) that are performed on the objects based on the operations (Update, Complete, Exact) you select.

Table 2: Failover Pair Synchronization Functions

Data Description	Update	Complete	Exact
DHCP Server: Client-Class Properties Client Host Name Processing Properties Dynamic DNS Properties Failover Tuning Properties	replace	replace	replace
All other properties	no change	replace	replace
LDAP Remote Server	ensure	replace	exact
Policy: Option List Properties Packet Boot File Properties All other properties	ensure ensure replace	replace replace replace	replace replace replace
Client	replace	replace	exact
Client-Class	replace	replace	exact
Scopes	exact	exact	exact
Links	exact	exact	exact
Prefixes	exact	exact	exact
Reservations	exact	exact	exact
DNS Update Configuration	replace	replace	exact
Trap Configuration	ensure	replace	exact

Data Description	Update	Complete	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	ensure	replace	exact
DHCP Listener Configuration	ensure	replace	exact

Step 6 Click **Report** on the Synchronize DHCP Failover Pair page:

- When 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. You have the option to choose the direction of synchronization and also the option to check the desired mode of synchronization operation (**Update**, **Complete**, **Exact)**. Check the desired values and click **Report**.

Step 7 Click **Save** on the View DHCP Failover Pair Sync Report page.

Step 8 On the List/Add DHCP Failover Pairs page, click the Manage Failover Servers tab.

Step 9 Click the **Restart Server** icon to reload the backup server.

Step 10 Try to get a lease.

Step 11 On the Manage Failover Servers tab, look at the health of the servers. Also, click the Logs tab 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.
```

CLI Commands

Use **failover-pair** *name* **sync** {**update** | **complete** | **exact**} [{**main-to-backup** | **backup-to-main**}]:

```
nrcmd> failover-pair example-fo-pair sync exact main-to-backup
```

Failover Checklist

Once you create the failover pair, you must synchronize the configuration of the failover servers. Use this checklist to prepare for an effective failover configuration:

- Duplicate the DHCPv4 scope, DHCPv6 prefix, DHCPv6 links, reservations (IPv4 and IPv6), selection tags, policy, DHCP option, IP addresses, client-classes, dynamic DNS updates, dynamic BOOTP, VPN, DHCP extensions, DHCP extensions, LDAP server, and address configurations on the partner servers by synchronizing a failover server pair for a simple failover scenario.
- 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 use BOOTP (DHCP) relay agents (IP helpers), configure all BOOTP relay agents to point to both partners. Cisco Prime Network Registrar does not automatically detect this.

  You can detect BOOTP configuration errors only 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.

Configuring Failover Parameters Based on Your Scenario

Following are the advanced failover properties that are important to set:

- Backup percentage (see Setting Backup Percentages, on page 10)
- Backup allocation boundaries (see Setting Backup Allocation Boundaries, on page 21)
- Maximum client lead time (MCLT) (see Setting the Maximum Client Lead Time, on page 11)
- Safe period (see Using the Failover Safe Period to Move Servers into PARTNER-DOWN State, on page 13)
- Request and response packet buffers (see Setting DHCP Request and Response Packet Buffers, on page 15)
- Load balancing (see Setting Load Balancing, on page 15)

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 the addresses for a single server. Configure the main server to allocate a percentage of the currently available addresses in each scope and prefix delegation prefixes 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.

If the main server detects that the address pool is significantly out of balance or the server has no leases, then the pool of available or other-available leases are rebalanced even when the failover pair is functioning in the Normal state. The failover pair must be carefully monitored during failover and if the failover partner is down for an extended period then operator intervention may be required to move the failover partner to the PARTNER-DOWN state.

You can set the percentage of currently available addresses by setting the *backup-pct* attribute on the failover pair or DHCPv4 scope (**failover-pair** *name* **set fail backup-pct** or **scope** *name* **set backup-pct** in the CLI). The default backup percentage is 50%. DHCPv6 prefix delegation prefixes are fixed at 50% for the backup-pct equivalent.

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.

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 backup percentage 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, if backup percentage is set to 60%, the main server will reclaim addresses if its address pool falls below 60%.

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, on page 15.

The percentage depends on the new client arrival rate and the network operator reaction time. It depends on the arrival rate of new DHCP clients and the reaction time of your network administration staff. 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, on page 11. When these times expire, the backup server offers:

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

During the working hours, 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 can 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.

Related Topics

BOOTP Backup Percentage, on page 35

Setting the Maximum Client Lead Time

You can set a property for the failover pair 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 later of the partner-down time or the latest lease expiration time communicated with the partner gets over. The latest least expiration time communicated to the partner is typically 1.5 times the lease time from the last client lease request before communication was interrupted.
- 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 for 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 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 or DHCPv6 Solicit to the server, requesting a desired lease period (for example, three days). The server responds with a DHCPOFFER or DHCPv6 Advertise 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 desired period plus half of the client desired period (3 +1.5 = 4.5 days). 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 desired lease period (3 days). The server then updates its partner with the lease expiration as the current time plus the desired lease period (3 days), and the potential expiration time (4.5 days. See the description in **Step 2**). 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 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 preset value 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 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 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. 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. Never set both partners to PARTNER-DOWN.
- **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.

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.

In Cisco Prime Network Registrar 8.2 or later, the use-safe-period attribute is enabled by default for a failover pair and the default safe period is 4 hours. This ensures that if the failover partner is in COMMUNICATIONS-INTERRUPTED state for 4 hours, it will enter PARTNER-DOWN state automatically after the safe period elapses. You may need to review if this setting is appropriate for your network and adjust the safe-period based on your network requirements.

In addition, 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. In order to prevent this problem, it is important that you do not change the default settings for use-safe-period and put operational procedures in place to alert operations personnel when the failover pair loses contact with each other. Especially, in the event of network communications failure, operator intervention is required before the safe period elapses. Either one failover server needs to be taken offline or the use-safe-period attribute needs to be disabled on both the servers before the safe period elapses.

Note In Cisco Prime Network Registrar 8.2 or later, use-safe-period is enabled by default. You may want to review if this is appropriate for your network and you may want to disable the use-safe-period or adjust the safe-period based on your network requirements and monitoring.

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 shortage 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.

In Cisco Prime Network Registrar 8.2 or later, if the failover safe period length is more than the length of the MCLT and the failover server enters into PARTNER-DOWN state because of the safe-period, the server can start allocating its partner other-available leases to DHCP clients immediately. The advantage of this is that the server has additional leases to allocate. However, the disadvantage is that operator intervention is required within the safe period in case of network communications failure. Either the failover server needs to be taken offline or the use-safe-period attribute needs to be disabled on both the servers before the safe period elapses. Without operator intervention, both failover servers will transition to PARTNER-DOWN state and start allocating its partner addresses to new DHCP clients.

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 (4 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.
- If your corporate policy is to avoid conflict under any circumstances, then never let either 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 on the Manage Failover Servers tab of the 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.)

  Use **failover-pair name setPartnerDown** date 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.

  In Cisco Prime Network Registrar 8.2 or later, if you use setPartnerDown in the CLI and specify the date and time when the partner was last known to be operational then the failover server calculates the MCLT from the time specified in the setPartnerDown command. If the date and time is not specified for the setPartnerDown command, then the failover server calculates the MCLT from the time the failover server moved to the COMMUNICATIONS-INTERRUPTED state. In case of network communications failure, it is important that you specify the actual time the partner was last known to be operational in the setPartnerDown command. Otherwise, it can result in duplicate IP addresses.

  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 o'clock midnight on October 31, 2002:

```
nrcmd> failover-pair dhcp2.example.com. setPartnerDown -3d

nrcmd> failover-pair dhcp2.example.com. setPartnerDown Oct 31 00:00:00 2001
```

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

DHCP failover allows a limited number of binding updates to be outstanding (set through the (expert mode) max-unacked-bndupd failover-pair attribute). The default value of max-unacked-bndupd is 1/5th (20%) of the max-dhcp-requests value and also it is at least the min of 100 and max-dhcp-requests. The server allocates additional request buffers to accommodate failover (as it must have these resources available for failover).

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, Cisco Prime Network Registrar introduced the load balancing feature (based on RFC 3074).

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 applies only while the servers are in NORMAL mode; in other states, both servers can respond to clients.

According to RFC 3074, the servers calculate a hash value for each request that the server receives, based on the client identifier option value or hardware address. The request is serviced if the hash value is assigned to that server.

With failover load balancing enabled, the servers split the client load evenly. The main partner processes 50% of the hash values and the backup partner the other 50%.

While the failover partners periodically balance the available leases on the backup server or do so shortly after a scope or prefix is detected to be out of leases, enabling the rebalancing-delta-pct attribute (Expert mode) on the main server to set the percentage difference between the desired and actual available leases on the backup server that will trigger a rebalancing on the scope or prefix.

Each partner responds to all clients whenever a partner is not in NORMAL mode. Each partner responds only to the broadcast DHCPDISCOVER or SOLICIT messages from clients that are in their assigned hash values.

For broadcast DHCPREQUEST or REBIND messages, the server responds only if it is the targeted one (based on the server identifier option); so, if the targeted server is the main server and it is down, the backup does not service the client (unless you release the lease). Broadcast BOOTP, DHCPINFORM, and INFORMATION-REQUEST requests are also load-balanced.

Related Topics

Configuring Load Balancing, on page 16

Configuring Load Balancing

In the web UI, when setting the failover properties for the pair (see the Setting Up Failover Server Pairs, on page 3), 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**.

Note You must restart the DHCP server on both main and backup to apply the changes.

Recovering from a DHCP Failover

During normal operation, the failover partners undergo transition between states. If one of the failover server fails, then the partner takes over offering and renewing leases, using its private pool. When the main server becomes operational again, it re-integrates with its partner without administrative intervention.

The following sections describe how to confirm a DHCP failover, monitor DHCP failover event, what happens when servers enter various states, and how the servers integrate.

Confirming Failover

To confirm the 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/Add DHCP Failover Pairs page, 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 (for IPv4) and BNDUPD6 or BNDACK6 messages (for IPv6) 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**.

Related Topics

State Transitions During Integration, on page 19

Configuring Failover Parameters Based on Your Scenario, on page 10

Monitoring DHCP Failover

When the main failover server goes down, the backup server moves to COMMUNICATIONS-INTERRUPTED state. The backup server cannot determine whether the main server is down or whether it cannot contact with the backup server. Depending upon the nature of outage you should monitor situation and follow the following steps:

1. Monitor both the failover servers and take action immediately when the main server goes down.
2. When the backup server first takes over, attempt to get the main server operational.
3. If you succeed in getting the main server operational within the MCLT, then nothing more needs be done.
4. If the main server is not operational until the MCLT has expired, then move the backup server to PARTNER DOWN state. On the backup server, use failover-pair name setPartnerDown date in the CLI.
5. When the main server is operational, ensure that it can contact the backup server before it is restarted.

For more information, see State Transitions During Integration, on page 19.

Failover States and Transitions

During normal operation, the failover partners undergo transition between states. They stay in their current state until all the actions for the state transition are completed. If communication fails, they stay in their current state until the conditions for the next state are fulfilled. The states and their transitions are described in Table 3: Failover States and Transitions , on page 17.

Table 3: Failover States and Transitions

State	Server Action
STARTUP	Tries to contact its partner to learn its state, then transitions to another state after a short time, typically seconds.
NORMAL	Can communicate with its partner. The main and backup servers act differently in this state: • The main server responds to all client requests using its pool. If its partner requests a backup pool, the main server provides it. • The backup server only responds to renewal and rebinding requests. It requests a backup pool from the main server.

State	Server Action
COMMUNICATIONS-INTERRUPTED	Cannot communicate with its partner, whether it or the communication with it is down. The servers cycle between this state and NORMAL state as the connection fails and recovers, or as they cycle between operational and nonoperational. During this time, the servers cannot give duplicate addresses. During this state, you usually do not need to intervene and move a server into the PARTNER-DOWN state. However, this is not practical in some cases. A server running in this state is not using the available pool efficiently. This can restrict the time a server can effectively service clients. A server is restricted in COMMUNICATIONS-INTERRUPTED state: • It cannot reallocate an expired address to another client. • It cannot offer a lease or renewal beyond the maximum client lead time (MCLT) longer than the current lease time. The MCLT is a small additional time added that controls how much the client lease expiration is ahead of what the backup server thinks it is. • A backup server can run out of addresses to give new clients, because it normally has only a small pool, while the main server has most of them. The server is limited only by the number of addresses allocated to it and the arrival rate of new clients. With a high new client arrival or turnover rate, you may need to move the server into PARTNER-DOWN state more quickly.
PARTNER-DOWN	Acts as if it were the only operating server, based on one of these facts: • The partner notified it during its shutdown. • The administrator put the server into PARTNER-DOWN state. • The safe period expired and the partner automatically went into this state. In this state, the server ignores that the other server might still operate and could service a different set of clients. It can control all its addresses, offer leases and extensions, and reallocate addresses. The same restrictions to servers in COMMUNICATIONS-INTERRUPTED state do not apply. Either server can be in this state, but only one should be in it at a time so that the servers do not issue duplicate addresses and can properly resynchronize later on. Until then, an address is in a pending-available state.
POTENTIAL-CONFLICT	Might be in a situation that does not guarantee automatic reintegration, and is trying to reintegrate with its partner. The server might determine that two clients (who might not be operating) were offered and accepted the same address, and tries to resolve this conflict.
RECOVER	Has no data in its stable storage, or is trying to reintegrate after recovering from PARTNER-DOWN state, from which it tries to refresh its stable storage. A main server in this state does not immediately start serving leases again. Because of this, do not reload a server in this state.
RECOVER-DONE	Can transition from RECOVER or PARTNER-DOWN state, or from COMMUNICATIONS-INTERRUPTED into NORMAL state.

State	Server Action
PAUSED	Can inform its partner that it will be out of service for a short time. The partner then transitions to COMMUNICATIONS-INTERRUPTED state and begins servicing clients.
SHUTDOWN	Can inform its partner that it will be out of service for a long time. The partner then transitions to PARTNER-DOWN state to take over completely.

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. The table below describes what happens when servers enter various states and how they initially integrate and later reintegrate with each other under certain conditions.

Table 4: Failover State Transitions and Integration Processes

Integration	Results
Into NORMAL state, the first time the backup server contacts the main server	**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 *backup-pct* configured.
After COMMUNICATIONS-INTERRUPTED state	**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	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. • If the server finds that it went down and the partner subsequently went into this state: **1.** The returning server moves into RECOVER state and sends an Update Request message to the partner. **2.** The partner returns all the binding data it was unable to send earlier and follows up with an Update Done message. **3.** The returning server moves into RECOVER-DONE state. **4.** Both servers move into NORMAL state.

Integration	Results
	• If the returning server finds that it was still operating when the partner went into PARTNER-DOWN state: 1. The server goes into POTENTIAL-CONFLICT state, which also causes the partner to go into this state. 2. The main server sends an update request to the backup server. 3. The backup server responds with all unacknowledged updates to the main server and finishes off with an Update Done message. 4. The main server moves into NORMAL state. 5. The backup server sends the main server an Update Request message requesting all unacknowledged updates. 6. The main server sends these updates and finishes off with an Update Done message. 7. The backup server goes into NORMAL state.
After the server loses its lease state database	A returning server usually retains its lease state database. However, it can also lose it because of a catastrophic failure or intentional removal. 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, on page 11. 4. Both servers then move into NORMAL state.
After a lease state database backup restoration	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. 1. In this case, you must configure the returning server with the *failover-recover* attribute set to the time the backup occurred. 2. The server moves into RECOVER state and requests all its partner 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, on page 11. 3. Once the server returns to NORMAL state, you must unset its *failover-recover* attribute, or set it to zero. `nrcmd> dhcp set failover-recover=0`

Integration	Results
After the operational server had failover disabled	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 *failover-recover* 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, on page 11. 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 lease data, and goes into RECOVER-DONE state. 3. Both servers go into NORMAL state. At this point, you must unset the backup server *failover-recover* attribute, or set it to zero. `nrcmd> dhcp set failover-recover=0`

Setting Advanced Failover Attributes

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

- Setting backup allocation boundaries (see Setting Backup Allocation Boundaries, on page 21)
- DHCPLEASEQUERY and failover (see DHCPLEASEQUERY and Failover, on page 21

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 no addresses 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.

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.

Maintaining Failover Server Pair

This section describes how to maintain failover server pair and perform the following administrative tasks:

- Changing failover pair names (see Changing Failover Pair Names, on page 22)
- Restarting the failover servers (see Restarting the Failover Servers, on page 22)

Changing Failover Pair Names

Use **failover-pair** *old-name* **set name=***new-name* to change the name of the failover pair. In the web UI, you will have to remove and then create a new object (do this without reloading the DHCP server until the new object is ready).

Note If a cluster role in a failover relationship is changed (main to backup or backup to main), any existing state information for that relationship is discarded.

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/Add DHCP Failover Pairs page, select the failover pair on the Failover pane.

Step 2 On the Manage Failover Servers tab for the main server, select the server you want to restart.

Step 3 Choose Restart Server from the Quick View menu.

Related Topics

Confirming Failover, on page 16

Recovering Failover Configuration

When you upgrade Cisco Prime Network Registrar to the latest version, you can revert to the earlier version, in case the upgrade fails. You can upgrade one partner and when it has recovered to normal state and is working well, then upgrade the other partner.

You may be able to recover from the archive created during the upgrade, but if the upgrade is scheduled during a maintenance window, then you need to:

- Stop Cisco Prime Network Registrar completely using nwreglocal stop.
- Tar up the Cisco Prime Network Registrar DATADIR (/var/nwreg2/local/data) and save it in a safe location.
- Upgrade the server.

If it fails, then you need to:

- Stop Cisco Prime Network Registrar completely using nwreglocal stop.
- Delete the corrupt version of Cisco Prime Network Registrar DATADIR (The location is: /var/nwreg2/local/data).
- Extract the saved Cisco Prime Network Registrar DATADIR tarfile in the path the tarfile came from.
- Install the original version of Cisco Prime Network Registrar, which finds the existing DATADIR and use it.

Using PARTNER-DOWN State to Allow a Failover Server to Operate for Extended Periods without Its Failover Partner

One or both failover partners could potentially move into COMMUNICATIONS-INTERRUPTED state. 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. Never set both partners to PARTNER-DOWN.
- **Failover safe period expires**—When the servers run unattended for longer periods, they need an automatic way to enter PARTNER-DOWN state.

For more information, see Using the Failover Safe Period to Move Servers into PARTNER-DOWN State, on page 13.

Note It is strongly recommended that when one server in a failover pair has been or will be out of service for any extended period that the other server be placed into PARTNER-DOWN state and that the failover relationship remain configured.

The alternative, unconfiguring the failover relationship, will have much the same effect on the server that remains operational, but reintegrating that server and the returning failover partner back into a working failover relationship with no impact on the lease state data is difficult and may be impossible.

When one server in a failover pair will be down for a while, you must place the remaining, operational server into PARTNER-DOWN state. DO NOT unconfigure the failover relationship on the operational server.

Reintegrating the Returning Failover Partner

If the returning server has retained an intact lease state database, it is brought back into service and should make contact with the operational server.

If the returning server has failed catastrophically and could not be returned to service with an intact lease state database, then the situation is a bit more complicated. In this case, a new installation of CPNR is usually

required on the returning server (which may not even be the same physical machine). The returning server should have the same IP address as the failed server and the new CPNR installation must be configured identically to the failed server. Which, typically, is the same as the operational server. Then the new server is brought into service and makes contact with the existing operational server.

Note In both cases, it is vital that the existing, operational server actually be operational at the time that the returning server is brought online, since if the returning server cannot contact the operational server it will think it is the only operational server and start handing out IP addresses without regard or knowledge of what the operational server has done.

When a returning server first comes up it will contact the operational server and they will exchange the times that they last communicated.

There are two possible situations that can arise:

- When a server with an intact lease state database (where CPNR was not re-installed) returns to service, it will determine after contacting its partner that it was out of service for a while, and move into RECOVER state and its partner will send it information about what has happened since it left service. When this update is complete, both servers will move into NORMAL state.
- When server that had CPNR re-installed on it completes this exchange, it will recognize that it has never communicated with the operational server, but the operational server has communicated with it (or with its predecessor), and the newly restored server will realize that it has lost its lease state database. It will move into RECOVER state and then request a complete download of all of the lease state information from the operational server. When this download is complete (which may take minutes or possibly longer, depending on the size of the lease state database and the load on the servers), both servers will move into NORMAL state.

Restoring a Standalone DHCP Failover Server - Tutorial

This section describes how to recreate a DHCP failover relationship between a main and backup server where a backup server was put in standalone mode. This situation does not come up very often.

The proper way to handle a situation where a main server is out of service for any period beyond a few minutes is to set the backup server into PARTNER-DOWN state. For more information, see Using PARTNER-DOWN State to Allow a Failover Server to Operate for Extended Periods without Its Failover Partner, on page 23.

The following procedure is offered to recover from the situation where an administrator has mistakenly believed that the proper approach is to remove the backup server from the failover relationship if the main server is out of service. To reiterate, this is NOT the correct procedure. It is challenging to recover from this mistake, but the following procedure should help.

1. The standalone server assumes the role of the main server.
2. The original main server becomes the backup server.
3. The partners then synchronize.
4. Failover relationship to be intentionally broken to reverse the server roles.
5. Partners to resynchronize in their original failover roles.

Related Topics

Background , on page 25
Repair Procedure , on page 25
Reversing the Failover Role on Backup Server, on page 26
Starting with Server A Powered Off, on page 26
Starting with Server A Replaced, on page 28
Transferring Current Lease State to Server A, on page 28

Background

For the remainder of this section, the main DHCP failover server is identified as Server A (with a cluster object named cluster-A), and the backup server as Server B (with a cluster object named cluster-B). Server A is administratively or otherwise shut down or its Cisco Prime Network Registrar server agent gets stopped. At this point, Server B goes into the Communications-Interrupted mode.

The system administrator may then take one of the following approaches:

- **Continue running backup Server B in Communications-Interrupted mode**—The risk of running the backup server in this mode indefinitely is that it can exhaust the pool of typically 10% of the available addresses with which the backup server is allocated to service new clients.
- **Put Server B into Partner-Down mode without breaking the failover relationship**—One major caveat of giving the backup server full control of the address space, without suspending failover, is that the full transfer of the address space ownership does not occur until after the configured Maximum Client Lead Time (MCLT). The MCLT is an additional time period set on the main server, which controls the duration for which the client lease expiration is ahead of what the backup server detects it to be. The MCLT is typically 60 minutes. Until the MCLT expires, the available address pool of the backup server is limited to its allocated reserve.
- **Put Server B into Partner-Down mode and break the failover relationship**—This approach puts the backup server in standalone mode, and is the approach that the administrator chose in this scenario. The deciding factors were that the main server was expected to be offline for an extended period, and the number of new devices coming online was higher than anticipated. Because the low percentage of available addresses that the backup server could service would soon cause an outage for new devices, the administrator put Server B in standalone mode. The disadvantage of this approach is the care and effort required to preserve the original state of the network when restoring the partners to their original relationship.

The first two approaches have distinct advantages over the third. In most cases, the backup server is expected to have enough addresses to cover newly arrived clients until the MCLT expires. Pursuing the third approach can incur unnecessary administrative burden and risk.

Repair Procedure

The repair procedure is:

1. **Temporarily assign the backup Server B the role of the main failover server**—Reversing the failover partner roles effectively allows Server A to learn the current failover state from Server B.
2. **Migrate Server A and Server B back to their original failover roles**—The goal is for Server A to reacquire its original status as the main DHCP failover server.

The assumptions are:

- The Original main Server A is nonoperational and Cisco Prime Network Registrar is stopped.
- The Original backup Server B is operational.
- Failover between the partners is administratively disabled.
- Decision was made not to permanently reverse the failover roles of the two partners.
- Domain Name Services (DNS) is not running on either of the failover partners.

Note The IP addresses used as examples are for demonstration purposes only.

Reversing the Failover Role on Backup Server

The following steps restore failover by temporarily moving Server B into the main server mode.

On **Server B** (cluster-B):

Step 1 Ensure that failover is disabled. Modify the failover configuration, so that Server B becomes the main and Server A the backup:

```
nrcmd> failover-pair examplepair set failover=false
nrcmd> failover-pair examplepair set main=cluster-B backup=cluster-A
```

Step 2 Save the changes and reload the server:

```
nrcmd> save
nrcmd> dhcp reload
```

Step 3 Re-enable failover and reload the server again:

```
nrcmd> failover-pair examplepair set failover=true
nrcmd> dhcp reload
```

Server B is now the main failover server, ready for its partner to become operational again. Any further action that you take to prevent Server A from beginning to give out addresses in the meantime depends on its current state.

If Server A is:

- **Powered off**—See Starting with Server A Powered Off, on page 26.
- **Powered on with the Cisco Prime Network Registrar DHCP not configured to start**—See Starting with Server A Powered On and DHCP Server Stopped, on page 27.
- **Replaced by another machine**—See Starting with Server A Replaced, on page 28.

Starting with Server A Powered Off

If Server A was powered off, you must power it on again to continue. The next steps ensure that Server A comes online while preventing IP address leakage.

On **Server A** (cluster-A):

Step 1 Before turning on the server, you must take steps to prevent it from communicating with clients. The best way to do this is to manually disconnect the network cable, then boot up the machine. You will require a local console to perform the next step. Other alternatives include reconfiguring the relay agents not to forward packets to the server or otherwise preventing DHCP traffic to be received on the machine (such as by installing a temporary filter for DHCP packets on a firewall).

Note If it is not possible to prevent client traffic from reaching the server, it may provide erroneous information to clients that do attempt to communicate with it, until the DHCP server can be stopped. Therefore, you must be ready to stop the DHCP server as soon as possible after turning the server on, as described in the next steps, to reduce the number of clients that might be provided erroneous information, potentially resulting in duplicated leases.

Step 2 Turn on the server.

Step 3 Stop the DHCP server as quickly as possible:

```
nrcmd> dhcp stop
```

Step 4 Go to the Starting with Server A Powered On and DHCP Server Stopped, on page 27.

Starting with Server A Powered On and DHCP Server Stopped

Starting from a point where Server A is powered on, but the Cisco Prime Network Registrar DHCP server is stopped:

On **Server A** (cluster-A):

Step 1 Modify the failover configuration so that Server A becomes the backup server:

```
nrcmd> failover-pair examplepair set main=cluster-B backup=cluster-A
```

Step 2 Stop Cisco Prime Network Registrar:

- RHEL/CentOS 6.x— **/etc/init.d/nwreglocal stop**
- RHEL/CentOS 7.x— **systemctl stop nwreglocal**
- Windows— **net stop nwreglocal**

Step 3 Examine the DHCP logs to confirm that the DHCP server is not running.

Step 4 Bring Server A back on the network. Reconnect the network cable, reconfigure the relay agents, or remove any firewall filter added in the previous section.

Step 5 Remove the lease state database and event store:

- Linux

```
rm -rf /var/nwreg2/local/data/dhcpeventstore
rm -rf /var/nwreg2/local/data/dhcp/ndb
```

- Windows

```
cd install-path\local\data
rmdir /s dhcpeventstore
rmdir /s dhcp\ndb
```

Step 6 Start Cisco Prime Network Registrar:

- RHEL/CentOS 6.x— **/etc/init.d/nwreglocal start**
- RHEL/CentOS 7.x— **systemctl start nwreglocal**
- Windows— **net start nwreglocal**

Step 7 Set the DHCP service to be enabled on reboot and start the DHCP server:

```
nrcmd> dhcp enable start-on-reboot
nrcmd> dhcp start
```

Step 8 Go to the Transferring Current Lease State to Server A, on page 28.

Starting with Server A Replaced

If Server A was decommissioned and replaced, you must install Cisco Prime Network Registrar and push the failover configuration from Server B to the new machine. Also, you must restore any customer configuration specific to Server A. After these steps, Cisco Prime Network Registrar will start but not give out addresses:

Step 1 On **Server A** (cluster-A), install Cisco Prime Network Registrar.

Step 2 Reconstruct the Cisco Prime Network Registrar operating environment by restoring the accompanying software, such as Cisco Broadband Access Center and its required DHCP extensions. Do not make any administrative changes to the configuration until after pushing the configuration to Server B.

Step 3 On **Server B** (cluster-B), by using the Cisco Prime Network Registrar web UI, push an exact failover configuration to Server A. This effectively makes Server A the backup partner.

Step 4 On **Server A**:

a) If necessary, customize the Cisco Prime Network Registrar configuration as required for the operating environment, which might include making administrative changes.

b) Reload the DHCP server:

```
nrcmd> dhcp reload
```

Step 5 Go to Transferring Current Lease State to Server A, on page 28.

Transferring Current Lease State to Server A

- At this point, the failover partnership reestablishes itself, both servers will resynchronize their states.
- Server A becomes operational as the backup server.
- The operation will pause for the MCLT period (of one hour) and both partners resume their failover operations in normal communication mode.

Note Do not proceed to the Repairing Partners to Their Original Roles, on page 29 until both partners synchronize and report normal communication.

Repairing Partners to Their Original Roles

Assume that both partners are fully synchronized and report normal communication. To ensure that the failover partners can assume their original roles, you should:

Step 1 On **Server A** (cluster-A), stop the DHCP server:

```
nrcmd> dhcp stop
```

Step 2 On **Server B** (cluster-B), stop the DHCP server:

```
nrcmd> dhcp stop
```

Step 3 On **Server A**:

a) Disable failover, then make Server A the main server and Server B the backup:

```
nrcmd> failover-pair examplepair set failover=false

nrcmd> failover-pair examplepair set main=cluster-A
 backup=cluster-B
```

b) Save the changes and reload DHCP:

```
nrcmd> save

nrcmd> dhcp reload
```

c) Ensure that the configuration is in place and currently running. At this point, Server A is the sole operational DHCP server with 100% of the address pool.

d) Re-enable failover:

```
nrcmd> failover-pair examplepair set failover=true
```

e) Reload DHCP and double-check the configuration changes:

```
nrcmd> dhcp reload
```

Server A is now the failover main server awaiting Server B to become operational.

Step 4 On **Server B:**

a) Make Server A the main server and Server B the backup, then enable failover:

```
nrcmd> failover-pair examplepair set main=cluster-A
 backup=cluster-B

nrcmd> failover-pair examplepair set failover=true
```

b) Save the new configuration, but do not reload the server:

```
nrcmd> save
```

c) Restart the DHCP server on Server B:

```
nrcmd> dhcp reload
```

At this point, the failover partnership reestablishes itself in its original roles, both servers will resynchronize their states, and Server B becomes operational as the backup server. The operation will pause for the MCLT period (of one hour) and both partners resume their failover operations in normal communication mode.

Step 5 On **Server A** and **Server B**:

a) Validate whether both partners are in normal failover state:

```
nrcmd> dhcp getRelatedservers
```

b) Run a report and ensure that the results match on both partners, allowing a bit of skew for the difference in running times between the partners.

Changing Failover Server Roles

Caution Be careful when you change the role of a failover server. Remember that all address states in a DHCPv4 scope or DHCPv6 prefix are lost from a server if it is ever reloaded without that scope or prefix in its configuration.

Related Topics

Establishing Failover Using Standalone Server as Main, on page 30
Replacing Servers Having Defective Storage, on page 31
Removing Backup Servers and Halting Failover Operation, on page 32
Adding Main Servers to Existing Backup Servers, on page 32
Configuring Failover on Multiple Interface Hosts, on page 32

Establishing Failover Using Standalone Server as Main

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

Step 1 Install Cisco Prime Network Registrar on the machine that is to be the backup server. Note the IP address of the backup server.

Step 2 Configure the cluster. Enable failover on the standalone server, configure it to be the main server and recently installed as the backup.

To configure the cluster, use:

```
cluster name create address | ipv6-address scp-port=value admin=value password=value
```

For example:

```
cluster backup create 10.65.201.23 scp-port=1234 admin=admin password=changeme
```

Step 3 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 4 To sync the configuration use failover synchronization and do a exact sync from Main to Backup.

Step 5 Reconfigure all operational BOOTP relays to forward broadcast packets to the main server and backup server.

Step 6 Reload the backup server.

What to do next

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 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 sends a pool request to get its pool of address.
6. After allocating these addresses, the main server allocates the IP address to backup based on backup percentage.

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 **failover-pair** *name* **setPartnerDown** [*date*] 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 Cisco Prime Network Registrar.

Step 4 Sync the server configuration from its partner configuration using failover synchronization. However, do not recover any lease databases from an earlier backup or the partner system.

Step 5 Reload the replacement server.

What to do next

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 or prefixes that were designated as a backup to the main server.

Step 2 On the main server, remove the failover capability from those scopes or prefixes 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 Sync the main server scopes, policies, and other configurations on the backup server using failover synchronization.

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. Cisco Prime Network Registrar performs the same steps as those described in the Establishing Failover Using Standalone Server as Main, on page 30.

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 name or address. This requires an additional command. For example, if you have a host with two interfaces, server A and server B, and you want to make server A the a main failover server, you must define server A as the failover-main-server before you set the backup server name (external server B). 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 hostname that points to only one address or a record. You cannot set up your servers for round-robin support.

Troubleshooting Failover

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

Related Topics

Monitoring Failover Operations, on page 33

Detecting and Handling Network Failures, on page 33

Things to Avoid When Troubleshooting Issues Related to Failover, on page 34

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/Add DHCP Failover Pairs page, or by using **dhcp getRelatedServers** in the CLI.

Detecting and Handling Network Failures

The table below describes some symptoms, causes, and solutions for failover problems.

Table 5: 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 agent (ip-helper) was not configured to point at both servers; see the Configuring BOOTP Relays, on page 35.	• 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.

Symptom	Cause	Solution
SNMP trap: dhcp failover configuration mismatch	Mismatched scope configurations between partners	Reconfigure your servers.
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.

Things to Avoid When Troubleshooting Issues Related to Failover

When using failover, here are some things NOT to do when troubleshooting issues:

- Removing the failover configuration. It is far better to set the remaining server into PARTNER-DOWN state. There are cases where this will require a longer wait to reuse the lease, but it is far safer to keep failover configured and operate in PARTNER-DOWN.
- Never copy the DHCP lease database (.../data/dhcp/ndb) from one failover partner to the other. See the *Restoring DHCP Data from a Failover Server section in Cisco Prime Network Registrar 9.0 Administration Guide* for how to recover the lease data from the failover partner. If this is done, you MUST use the leaseadmin tool to remove the server-duid after copying the database (see Moving Leases Between Servers for more details on the leaseadmin tool). Any time the lease databases are copied, the server-duid must be removed from the copy.

Note If the server-duid is not deleted, you can end up with two servers having the same server-id and hence DHCPv6 will not work as intended; this can have serious consequences for regional lease history data.

Supporting BOOTP Clients in Failover

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

Related Topics

Static BOOTP, on page 34

Dynamic BOOTP, on page 35

Configuring BOOTP Relays, on page 35

Static BOOTP

You can support static BOOTP clients using DHCP reservations. When you enable failover, remember to configure both the main and backup servers 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 Cisco Prime 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 server fails, clients are not serviced, because the backup server 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.

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. Cisco Prime 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, Cisco Prime Network Registrar ignores it (to enable duplicating configuration through scripts). If you do not define it, Cisco Prime 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