DHCP Capacity and Performance Guidelines

This section provides capacity and performance guidelines for Cisco Prime Network Registrar 9.0 and later, and also for 64-bit versions of Cisco Prime Network Registrar 8.3.2 and later.

The goal of this section is to provide an understanding of what influences the capacity and performance of the servers to help in planning how to deploy the product and what to consider when purchasing hardware for these systems. These recommendations primarily apply to the Linux releases.

When multiple clusters are running on virtual machines, the underlying physical hardware needs to be at least the sum of the individual virtual machine requirements. Also, it should be noted that high availability solutions (that is, HA-DNS or DHCP failover) should not have both partners located on the same physical machine in virtual environments, as that makes the hardware a single point of failure.

Note These are just guidelines, as actual performance may vary based on variances in the live deployment.

Local Cluster DHCP Considerations

There are two common questions concerning DHCP capacity:

1. How many leases can I put on a single server?

2. If I want to put *n* leases on a server, what sort of server should I purchase or virtual machine should I configure?

Number of Leases Allowed on a Single Server

When discussing about the capacity of a server, the number of DHCP operations per second that the server can support is the most important issue. There are two regimes that affect the operations per second that the server will be required to support:

- **Steady state**: Made up of existing DHCP clients renewing their leases and the arrival of DHCP clients not previously seen by the server.

- **Avalanche**: Made up of a large (possibly vast) quantity of existing DHCP clients, all contending at the DHCP server to get an address. This situation can occur with restoration of power after a failure or perhaps a blanket reset of many customer devices. This can often consist of tens of thousands of DHCP clients all trying to get an IP address from the DHCP server at the same time. It can even be hundreds of thousands of DHCP clients trying to get an IP address.

For the steady state situation, the number of DHCP clients and the lease times of the leases they are granted will dominate the load.

The operations per second required by a DHCP client population is largely driven by the size of that client population coupled with the lease times (both expiration and renewal times) that are granted to that population. These values are all configurable, and so the actual requirements can vary dramatically.

Following table presents a range of these data points showing the operations per second required for various client populations and differing lease times:

Table 1: Client lease Times

Operations per Second						
	Client Lease Times					
Active Leases	30 min	1 hr	1 day	1 week	2 weeks	30 days
1,000	1	1	-	-	-	-
10,000	11	6	-	-	-	-
100,000	111	56	2	-	-	-
500, 000	556	278	12	2	1	-
1,000,000	1,111	556	23	4	2	1
1,500,000	1,667	833	35	5	2	1
2,000,000	2,222	1,111	46	7	3	2
4,000,000	4,444	2,222	93	13	7	3
6,000,000	6,667	3,333	139	20	10	5

The lease times granted to the clients has an overwhelming influence on the steady state operations per second required on the DHCP server. A server's operations likely include a mix of lease times, as lease times for clients without an existing lease are limited by the failover Maximum Client Lead Time (MCLT), and there may be other operations (such as from "bad" clients or lease query requests).

The DHCP server will not collapse under any client load, but it can take seconds to minutes to work through tens or hundreds of thousands of clients. It is for this reason that our recommendations for the operations per second that the server is required to support in steady state tends to be on the lower side; so that the server has plenty of headroom to process the eventual avalanche.

DHCP operations per second

It is difficult to give concrete recommendations regarding the operations per second that the DHCP server can deliver to DHCP clients, since there are many factors that are involved in this aspect of DHCP server performance.

Cisco has measured DHCP server performance in the lab well above 20,000 operations per second. However, that was a DHCP server which was configured specifically for maximal performance (no failover, no logging, no lease history, no extensions, and no LDAP). Almost every feature that you configure in the DHCP server costs some amount of performance; frequently trimming 10 percent or so off of the previous performance. Some features, for instance LDAP lookup or running with the Prime Cable Provisioning (PCP) product, can have a much bigger effect on performance; since the LDAP lookup or PCP interaction with the DPE requires interlocking with a separate server and the round-trip delays that entails, prior to even processing the incoming DHCP request. Failover costs at least 10 percent, basic logging can also cost 10 percent of performance or more. Extensions will cost an unpredictable amount on top of a constant overhead to just call the extension. The time spent in the extension is also synchronous and additive to the time it takes to process every DHCP request.

The upshot of all of this is that there is no way to reasonably predict the operations per second that the DHCP server will be able to supply given a particular load when running on a particular hardware configuration with a particular software configuration.

Also, the operations per second load placed on the DHCP server by the constant requirement to process DHCP RENEW requests from DHCP clients ("steady state") is frequently overshadowed by the requirements to process large "avalanche" loads, where many thousands to tens of thousands of DHCP clients attempt to get service from the DHCP sever in a very short time. These events can be generated by a power outage among the DHCP clients or network element resets that will provoke many thousands of DHCP clients to re-DISCOVER / re-SOLICIT for IP addresses. The DHCP server needs to be able to process these loads, which typically dwarf the loads generated by the steady state RENEWAL traffic.

Cisco recommends that the steady state load on the DHCP server be limited to a few hundred operations per second, in part to ensure that headroom exists to process the avalanche loads presented to the DHCP server in unusual circumstances. We have customers which have high performance hardware and excellent monitoring regimes that run with several hundred operations per second and sometimes more with constant load. They are running successfully, in part because they are careful to ensure that they do not let the avalanche load size get too large; by limiting the number of active leases on each server.

The DHCP server has several features to reduce the load on the server and help it service requests as quickly as possible, especially under avalanche conditions:

- **Defer-lease-extension**

  By default, the server will defer extending a lease to a client if the client "renews" before its expected renewal time. This usually helps out with avalanches if the outage that triggered it was short (less than 1/2 the lease time) as a large number of clients will avoid the need for a disk write (and failover update).

- **Reduced logging when overloaded**

  By default, the server will reduce the logging when the request buffers in use exceeds 67 percent of the configured buffers. As logging can be costly, this allows the server to handle additional capacity when very busy. This feature can be disabled. Note that the server dropping requests under avalanche conditions should be expected, as that is the only way that the server can shed load, and the client will re-transmit the request. Under steady state conditions, if a server is frequently dropping requests, that is probably an indication that it is unable to handle the load.

- **Chatty Client Filter**

Use of this provided extension is highly recommended in all service provider networks. This extension monitors client activity and blocks those clients that are considered to be "chatty". Once a client is blocked, it is unblocked if it quiets down. In many service provider networks, the Chatty Client Filter can reduce the requests to the server by about 50 percent. However, the Chatty Client Filter requires careful tuning and requires reviewing that tuning periodically to assure traffic patterns have not changed. For more details, see the *"Preventing Chatty Clients by Using an Extension" section in the Cisco Prime Network Registrar 10.1 DHCP User Guide*.

- **Discriminating Rate-Limiter**

    The Discriminating Rate-Limiter reduces downtime after an outage in service networks by restricting the rate of DISCOVER and SOLICIT requests while still honoring all RENEW requests. The basic concept is to assure a client that was offered a lease is able to complete getting that lease. For more details, see the *"Setting Advanced DHCP Server Attributes" section in the Cisco Prime Network Registrar 10.1 DHCP User Guide*.

Number of leases you want on a server

If the only thing that mattered was the steady state operations per second load, then looking at the table above and with a one week lease time, you could imagine 12 million or even 24 million leases would pose no problem. However, there are other factors as follows:

- **Avalanche load**: Which may or may not scale with the total leases on a server.

- **Reload time**: The server needs to refresh its in-memory cache whenever it is reloaded, and the reload time scales linearly with the number of active leases in the server.

- **Service interruption impact**: If you have millions of leases to start with, then there is probably a relationship between DHCP clients and customers of some sort. You probably want to avoid having a DHCP server have so many leases that having an entire DHCP failover pair out of service for a few hours would cause an unacceptable risk to your business. While DHCP failover will prevent almost all service interruptions and you probably have no single points of failure, sometimes two things do fail at the same time. It is possible that both servers in a DHCP failover pair will fail for a while, and in the unlikely event that this should happen, the difference between having 2 million DHCP clients on a server and 10 million DHCP clients on a server could be very important. With the reasonable DHCP lease times, only some small percentage of DHCP clients will have their leases expire every hour that a failover pair is out of service.

Recommendations

Cisco strongly recommends that you limit the total active leases on a single DHCP server (or server failover pair) to 6 million leases. In addition, Cisco strongly recommends that you limit the steady-state operations per second requirement to 500 operations per second, in order to have sufficient bandwidth to handle avalanche and other exceptional conditions.

Scale out, not up, beyond some point!

Instead of loading vast quantities of leases into a single DHCP server or failover pair, consider keeping the number of leases to a more modest number, say 3 to 5 million leases. Cisco resource limits set the warning level to be 6 million leases, and it is wise to configure more like 4 million leases per server to allow for growth in the future. While managing multiple failover pairs is more work than just managing one failover pair, the ease of management of a server that is more modestly loaded with 3 to 4 million leases will pay long term

dividends, to say nothing of the impact on your business in the unlikely event that an entire server pair should fail for a couple of hours.

Request Latency

It should be noted that the DHCP server's design is optimized to respond to large numbers of requests quickly – it is not optimized to have the lowest latency for each request. This often complicates testing for scale as the server's performance with a few simultaneous requests may not show its true processing power.

Server Considerations

If you do not need a lot of operations per second and do not have a lot of leases on the server, pretty much any server will do. For the purpose of this discussion, we will assume that you want to get the maximum performance possible.

For DHCP, the general recommendations in terms of physical or virtual server considerations are as follows:

1. Disk write performance is the primary consideration. SAN storage, SSD, or 15K RPM HDD disks are recommended. The DHCP server is disk write performance limited, because it must commit to disk any changes to leases (primarily assigning a lease to a new client and extending the lease times on a lease) before responding to a client. Configuration options, such as failover, lease history, and DNS updates also increase the disk write load on the server, as each of these require additional write operations. There are up to 4 writes for a lease on the server that grants, extends (renew/rebind), releases, or expires a lease plus 1 more write on the failover partner as follows:

   • The lease itself (before responding to the client). Generally, this also results in a failover binding update if failover is used.

   • A history record (this only occurs if lease history is enabled and the lease was leased but is no longer).

   • The partner writes the lease when it receives a failover binding update (if failover used).

   • The lease after the receipt of the failover binding update acknowledgement (if failover used).

   • The lease after the DNS update completes (if configured and initiated for the lease).

   A server may also initiate writes at other times for a lease, such as for failover state transitions for the lease, when balancing failover pools, and because of user action (such as to force a lease available). The DHCP server lease state database disk space requirements are generally as follows:

   • 1 KB for each configured or active lease, and

   • If lease history is enabled, 1 KB for each historical record.

   These numbers can be reduced about 30 percent if the lease record compression is enabled (see the DHCP server's *server-flags* attribute).

   Note These numbers need to be multiplied by 3 to accommodate the shadow backups. These numbers just reflect the lease state database and no other system requirements.

2. Memory (RAM) is secondary, with 64-bit support, memory limits are not generally a concern provided the system has sufficient memory. It is important to have sufficient "free" memory for the file system to

be able to keep the entire DHCP lease state database in memory to avoid the need for disk reads. A rough rule of thumb is to assume:

- 1 KB for each configured or active lease for the DHCP server's memory usage. Configuration options, such as DNS update and the length of host and domain names and the amount of option-82 (DHCPv4) or Relay-forward message (DHCPv6) data can influence this rule of thumb.

- 1 KB of "free" memory for the file system cache for each lease (configured or active) and,

- If lease history is enabled, 1 KB of "free" memory for the file system cache for each history record (this will be more difficult to judge as it depends on how frequently leases expire or are released).

3. CPU performance is the least significant as the processing required to service requests is generally low. On the other hand, avalanche processing is largely handled with just CPU cycles and minimal disk writes. So, if you have a large avalanche possibility, invest in a system with good CPU capability and fast network interfaces. Most modern multi-processor systems should be sufficient for modest avalanche loads. For higher capacity/performance applications, both the CPU speed and number of effective processors should be higher. The DHCP server is highly multi-threaded, so that, additional CPU cores will help DHCP server performance up to a point. Due to the requirements for some minimal amount of locking inside the DHCP server, performance will improve when adding up to 12 CPU cores. Beyond 12 CPU cores, there is not much of any performance improvement due to the requirements for synchronization.

Regional Cluster DHCP Considerations

The regional cluster disk space requirements are dictated by several factors for DHCP:

1. **Lease history**—When lease history is enabled at the local clusters, by default, the regional cluster collects this history from the local clusters for longer term storage (the default is to retain these records for 24 weeks, see the CCM server's *trim-lease-hist-age* attribute). As mentioned above for the DHCP server, each lease record (active and historic) should be assumed to require about 1KB, but this should be multiplied by 3 to accommodate backup requirements – thus, 3 KB/lease record. The regional cluster disk space needed will depend on the total number of lease history records, which depends on the number of servers, their lease counts and client activity levels, and the period of time over which the history is to be retained. In very large service provider networks, this can easily be 100 GB or more.

> **Note** These disk space requirements can be reduced to 30 percent for the lease history data by enabling lease record compression in Cisco Prime Network Registrar 9.0 and later (see the CCM server's *lease-hist-compression* attribute).

2. **Network utilization**—The regional cluster also collects subnet and prefix utilization data from the local clusters (by default, every hour and retained for 24 weeks; see the CCM server's *addrutil-poll-interval* and *addrutil-trim-age* attributes). While each record is about 1/2 KB (the scope/prefix names, owner, region, selection tags, and other data cause the size to vary), this can add up if there are many subnets and prefixes, a 10,000 scope/prefix deployment can use 10 GB over a 24 week period (not considering the backup requirements, which make this 30 GB).