DHCP Technical Overview

As its name implies, Dynamic Host Configuration Protocol (DHCP) is the networking service used to dynamically assign IP addresses and configuration information. To communicate on a TCP/IP network, every machine needs a TCP/IP address (not just any address; every machine needs a unique IP address) and a subnet mask. Additionally, other TCP/IP configuration information might be necessary. To communicate outside their local networks, clients need a default gateway. To find resources across the network, they need to find name resolution services such as DNS servers, WINS servers, and so on. You can manually configure every machine on your network with this information (TCP/IP address, subnet mask default gateway, DNS servers, WINS servers, and the like) and then maintain a list of who has what address. This works fine if you have only a few machines in a single location. After you start getting more machines and more locations, the administrative burden becomes overwhelming. This is where the DHCP protocol comes in. DHCP is an Internet standard protocol (RFC 2131) for dynamically assigning TCP/IP configuration information.

When a machine that is configured as a DHCP client boots up on the network, it sends out a broadcast request for a DHCP address. When the DHCP server receives the request, it sends an acknowledgement, and the client then requests the IP address information. The DHCP server maintains a database of available IP addresses and issues the client an available address and any other IP configuration that is specified for that address range (called a *scope*). The client is then configured with the appropriate information and can now communicate on the network.

DHCP Server Configuration

DHCP not only provides a centralized mechanism for issuing addresses, but it also maintains those addresses. The DHCP database contains all the IP addresses issued and information on how long (each address is leased to the client for a specified interval). This serves two purposes. First, by having a centralized database, DHCP can minimize IP address conflicts because it already knows which addresses are in use—the ones in its database. DHCP can also be configured to check whether an IP address is available before issuing it to a client. This helps prevent conflicts with machines that are statically configured. For example, according to the DHCP database, a particular address should be available because it hasn't been assigned by DHCP. However, in reality that particular address is configured on a static device, such as a printer, server, or router. Rather than issue the address to a client and cause an IP address conflict, the DHCP server can be configured to test the address first (by pinging it). Because the address is being used, the DHCP server gets a response, marks that IP address in its database as used, and picks another address to issue to the client (testing this new address as well). Secondly, by issuing addresses for a specified interval, it more efficiently utilizes the available IP addresses. For example, if your organization contains a sales force, it probably has laptops that are constantly connected and disconnected from the network. Sometimes they might be disconnected for months. By issuing IP addresses with leases of 8 days (the default), DHCP can reclaim the IP addresses used by the machines that haven't connected in a while.

DHCP servers are configured with scopes. Each scope is a range of IP addresses to assign to a particular segment. For each scope, you can configure different IP configuration options. For example, for each segment, you can configure a DHCP scope with the appropriate range of addresses and the default gateway for that segment. In addition to configuring scope-specific options, you can configure global options. Global options provide IP configuration information for all scopes on the DHCP server—for example, DNS servers, WINS servers, and NetBIOS node type.

DHCP Relay Agents

As mentioned earlier, DHCP clients broadcast DHCP requests. Does this mean you need a DHCP server on every segment of your network? Not necessarily. RFC 1542 provides a specification for a DHCP relay agent. Some routers are RFC 1542 compliant, meaning that they can be configured to forward DHCP requests. Depending on the router, they can be configured to forward the request to one or more DHCP servers (additional servers for fault tolerance). Even if you can't forward requests with your routers, don't despair. You can configure just about any Windows NT-based machine (Windows NT, 2000, XP, or .NET) as a DHCP relay agent.

Tip An alternative to using DHCP relay agents is *multihoming* your DHCP server(s). By having multiple NICs in your DHCP server (one for each segment), you do not need DHCP relay agents and you also improve performance by having the DHCP server local to the clients. Simply configure multiple scopes on the DHCP server and configure it to listen to each of the NICs.

Scopes and DHCP relay agents can work together to provide fault tolerance in your DHCP topology. It is generally recommended that you have multiple DHCP servers, each configured with overlapping scopes. For example, configure one server with the scope for a particular segment, but use only about 80% of the scope. Then, configure a DHCP relay agent to forward to a second DHCP server and give that server the other 20% of the scope. That way, if a problem occurs with the first server, the second can at least issue a few addresses.

DHCP gives administrators a quick and easy way to provide clients with TCP/IP configuration information. By simply configuring scopes, you can provide any machine that boots on the network (provided it's a DHCP client) with an IP address and the appropriate configuration information. Not only does it ease the administrative burden, but it also makes it easier on the end user, particularly in today's mobile workforce. No longer do you have to reconfigure when moving from one office to another, or from the office to home. DHCP just handles it.