DNS server caching

The DNS name resolution process might seem long and complex, but in many cases it isn't necessary for the client's DNS server to send queries to the servers for each domain specified in the requested DNS name. This is because DNS servers are capable of retaining the information they learn about the DNS namespace in the course of their name resolution procedures and storing it in a cache on the local hard drive.

A DNS server that receives requests from clients, for example, caches the IP addresses of the requested systems and the addresses for authoritative servers of particular domains. The next time a client requests the resolution of a previously resolved name, the server can respond immediately with the cached information. In addition, if a client requests another name in one of the same domains, the server can send a query directly to an authoritative server for that domain rather than to a root name server. Thus, the names in commonly accessed domains generally resolve quickly because one of the servers along the line has information about the domain in its cache, whereas names in obscure domains take longer, because the entire request/referral process is needed.

Caching is a vital element of the DNS architecture because it reduces the number of requests sent to the root name and top-level domain servers, which, being at the top of the DNS tree, are the most likely to act as a bottleneck for the whole system. However, caches must be purged eventually, and there is a fine line between effective and ineffective caching. Because DNS servers retain resource records in their caches, it can take hours or even days for changes made in an authoritative server to be propagated around the Internet. During this period, users might receive incorrect information in response to a query. If information remains in server caches too long, then the changes administrators make to the data in their DNS servers take too long to propagate around the Internet. If caches are purged too quickly, then the number of requests sent to the root name and top-level domain servers increases precipitously.

The amount of time that DNS data remains cached on a server is called its *time to live (TTL)*. Unlike most data caches, the TTL is not specified by the administrator of the server

where the cache is stored. Instead, the administrators of each authoritative DNS server specify how long the data for the resource records in their domains or zones should be retained in the servers where it is cached. This enables administrators to specify a TTL value based on the volatility of their server data. On a network where changes in IP addresses or the addition of new resource records is frequent, a lower TTL value increases the likelihood that clients will receive current data. On a network that rarely changes, a longer TTL value minimizes the number of requests sent to the parent servers of your domain or zone.

To modify the TTL value for a zone on a Windows Server 2012 R2 DNS server, right-click the zone, open the Properties sheet, and click the Start Of Authority (SOA) tab.
On this tab, you can modify the TTL for this record setting from its default value of 1 hour.