ADANAC
SOFTWARE

Load Balancing and Round Robin DNS

By Dirk Hart

Round robin DNS is a leading technique for providing a high level of availability of some service (typically http/web site) and for providing load balancing.

Large web sites need to be able to handle a huge number of requests - often more than a single web server can handle. The solution is to hand off incoming requests to a group of servers that each handle a portion of the load. The most convenient - and least expensive - way to do that is to use Round Robin DNS.
cartoon

Round Robin DNSsimply takes advantage of the way most DNS servers provide DNS records. Specifically consecutive queries provide the same results but in a different order. One website using this technique is cnn.com.

Using dig (dig cnn.com) I got the following results:

```
cnn.com. 300 IN A 64.236.16.20
cnn.com. 300 IN A 64.236.16.52
cnn.com. 300 IN A 64.236.16.84
cnn.com. 300 IN A 64.236.16.116
cnn.com. 300 IN A 64.236.24.12
cnn.com. 300 IN A 64.236.24.20
cnn.com. 300 IN A 64.236.24.28
cnn.com. 300 IN A 64.236.29.120
```

And a few seconds later I got this:

```
cnn.com. 266 IN A 64.236.24.28
cnn.com. 266 IN A 64.236.29.120
cnn.com. 266 IN A 64.236.16.20
cnn.com. 266 IN A 64.236.16.52
cnn.com. 266 IN A 64.236.16.84
cnn.com. 266 IN A 64.236.16.116
cnn.com. 266 IN A 64.236.24.12
cnn.com. 266 IN A 64.236.24.20
```

What's interesting to note is that the records are all the same but in a different order. In fact, the sequence is unchanged as well; 64.236.29.120 always follows 64.236.24.28.

So when you browse cnn.com you might get your response from the server with the IP address 64.236.24.28 while the next browser might get a response from 64.236.16.20.

In this fashion the aggregate load being placed on cnn.com is distributed over 8 servers. And, if the load increases you simple add a server and another DNS record. So far this looks like an ideal solution - simple to setup and administer and inexpensive too.

Let's consider the case where one of the servers stops working. Those unlucky folks who get the IP of the failed server just won't get a response. So you call your buddy Mario across town and gasp that the whole of cnn.com is down. Mario tries it and gets a response since his DNS inquiry got the IP of a server that was running. Unfortunately 12.5% (1/8) of the NEW inquiries will get the IP of the failed server. That's no good!

Well, I don't think that the sysadmins at cnn.com are asleep at the wheel - in fact they probably have a klaxon that sounds whenever a server takes a dive. And when a server fails the sysadmin does one of two things: she rushes over to the DNS server and removes the DNS record for the failed server (and synchronizes the other DNS servers). The second thing the sysadmin might do is simply confirm that the program monitoring the failed has already removed the corresponding DNS record. There is, after all, no reason that this task can't be automated. That way the 12.5% of new inquiries is really restricted to *12.5% of new inquiries* while the DNS record for the failed machine is still available.

Also, note that the TTL (Time To Live) of the DNS records is quite short at 300 seconds. This is done to control how long the DNS record survives in each downstream DNS server. It's not reasonable - or even desirable - to expect all the DNS servers in the world to make a new DNS inquiry each time someone wants to browse cnn.com. Neither is it desirable to set the TTL too long, since you wouldn't want a failed server to show up any longer than necessary (TTLs are typically set to 3600 or 7200 seconds or even longer). In short, when a server fails you want it's DNS record to be expired off all the DNS servers in all the world as soon as possible.

Of course, when a server is ready for use you only need to create a DNS record for it and wait for it to propagate, which won't take long since it's peers (other servers) have DNS records with short TTLs as well.

It's not necessary to setup and maintain your own DNS servers at all. AOL is providing this service for cnn.com (they're part of the same company) and UltraDNS provides similar services as well.

There are a few downsides to this arrangement. First, while round robin DNS load balancing provides an elegant and inexpensive solution for load balancing it does in no way provide high availability, Indeed, the very mechanism of it's elegance is beyond the control of the sysadmin! Second, some Internet Service Providers (ISPs) routinely substitute longer TTLs where they feel -arbitrarily- that your published TTL is too short. Finally, many desktops perform their own DNS caching or are behind servers that provide DNS caching.

As an aside, some routers provide for a feature that really works against Round Robin DNS, namely negative caching (RFC 2308 - Negative Caching of DNS Queries). This evil feature records sites from which no answer was received. This means that a 10 minute server outage can be locally multiplied to some arbitrary TTL (negative TTL?) often 30 to 60 minutes.

Also, since Windows 98, MS Windows has been caching DNS locally as well undoing the benefits of Round Robin DNS. Although this feature can be turned off most Windows users are unaware that their machines are caching DNS at all.

In summary, Round Robin DNS Load Balancing is an effect, robust and inexpensive method of load balancing with some warts and bumps. It should not be taken to be a solution for high availability computing.