Load balancer

The distribution between number of read and write is heavy biased on the read side. A quick estimation based on collected queries (SELECT, INSERT, UPDATE, DELETE) shows that approximate 75% is read queries (SELECT). It's therefor important that reading data is fast.

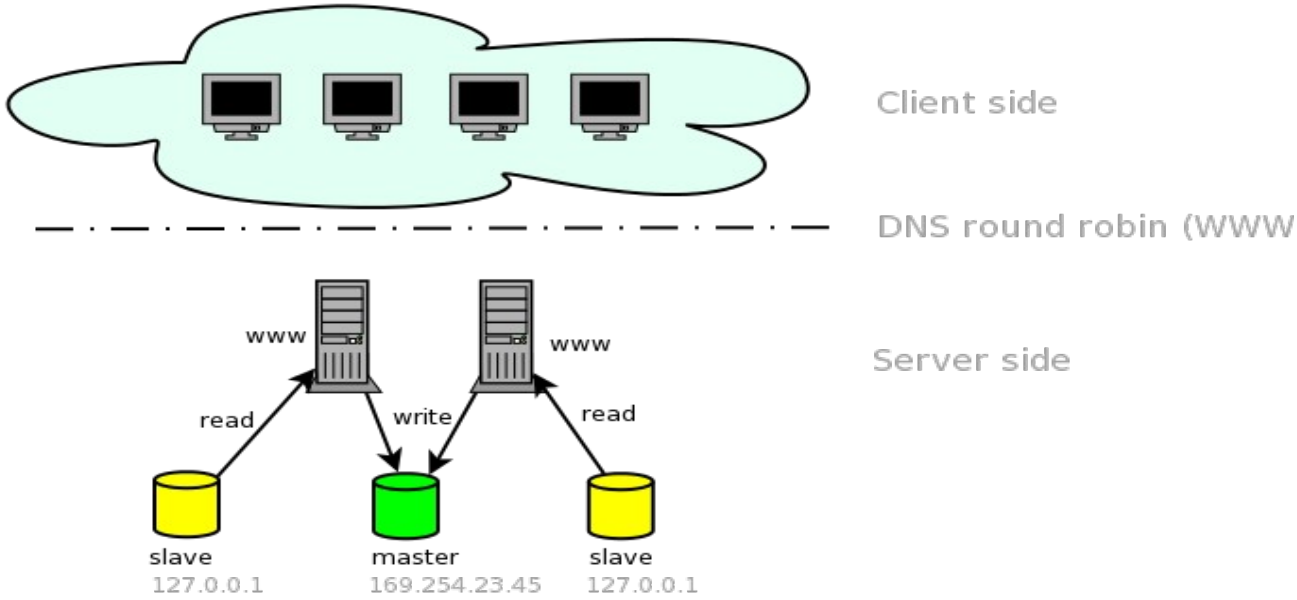
Background:

The idea is to have a read-only (slave) database local attached to each web server. All writes goes to a master database that is replicating data changes to all slave servers. Performance test on the master server shows that data is available on slave servers within one second in 9/10 cases.

The normal flow of an web application (per request) is to read data at start and then modify data at the end. The delay for replication should not be a problem in this case.

Design:

Each www-server has a local attached read-only database listening on localhost. The read connection gets automatic selected by the application model. By this design, all queries are using a UNIX socket to the database slave. The web server frontend is selected (by clients) by using DNS round robin (same hostname/multiple ip-addresses).



High availability:

The proposed solution solves load balancing. For high availability, a virtual server (using the Linux kernel) can transparent distribute the request among the online web servers.

Anders Lövgren, 2014-12-05
Computing Department at BMC, Uppsala University