Squid

CS174

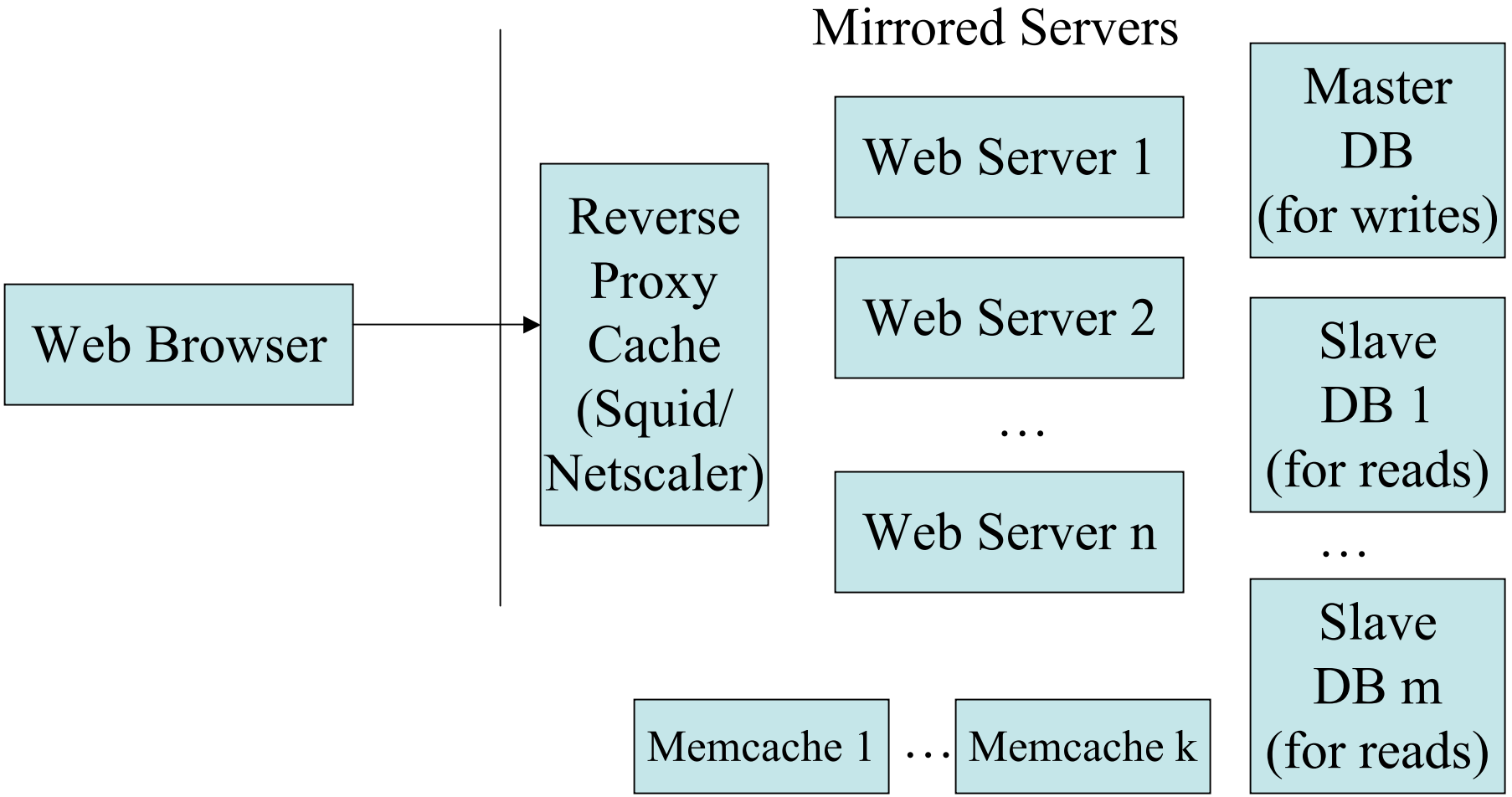
Chris Pollett

Nov. 19, 2008.

Outline

- More Caching

Sample Website Architecture



Squid

- Is a caching proxy available at:
  http://www.squid-cache.org/
- It originally developed from Harvest Cache Daemon project at UCSD. The latter project started in the 1990s.
- We are interested in using squids as a reverse proxies:
- The idea is clients (browsers), connect to our site and hit our squid, the squid determines if it already has non-expired copy of the page requested in its cache, if it does it returns it; otherwise, it forwards the request to one of a collection of webservers who runs the requested script. The returned result in this second case is then cached in the squid.

Setting up squid

- We first need to download and compile the squid src's. (./configure; make; sudo make install)
- Next we need to edit the squid.conf file in the squid ${squid}/etc/squid.conf
- Then one can type the command:

/usr/local/squid/sbin/squid -z

- Then one can run squid with:

/usr/local/squid/sbin/squid -NCd1 #for debugging

without -NCd1 if don't want debugging.

Testing Squid

- To just test out squid, we will set it up first as a proxy cache.
- The idea of a proxy cache is we want a cache for our LAN of any websites off of the LAN.
- The squid.conf has the ability to finely control which requests through it will be allowed.
- You can set up acl (access control lists). The default acl in squid.conf looks like:

  acl localhost src 127.0.0.1/32
- We next say what http stuff localhost is allowed to do by adding the following directive:

  http_access allow localhost
- Next you can start your squid server, and either request web pages using the squidclient or set up a proxy of localhost and port 3128 in your browser.
- To see the actual cache hits and misses you can look in ${squid}/var/logs/access.log

Configuring Squid as a Reverse Proxy

- To match the sample website architecture we want the squid to be on the main machine people come into our website at.
- Then we want it to cache pages that will be coming from our web servers.
- We need to edit the lines:

  http_port (probably set to port 80 from default 3128)

  httpd_accel_host (set to virtual if have more than one web server; otherwise, to ip address of the web server)

  httpd_accel_port (port number of web servers 80)

  httpd_accel_single_host (on or off depending on how many webservers you have,)

  httpd_accel_uses_host_header off

Load Balancing Machines

- You can also set up squid to do the load balancing between your web servers using directives like:

```
cache_peer ip.of.server1 parent 80 0 no-query round-robin
acl sites_server_1 dstdomain www.mysite.com
http_access allow sites_server_1

cache_peer ip.of.server2 parent 80 0 no-query round-robin
acl sites_server_2 dstdomain www.mysite.com
http_access allow sites_server_2
```