Maxwell Brown
MSU: brown308
Courseroute: MaxwellGBrown
github.com/MaxwellGBrown/tc362

The Technology Stack of Facebook

Facebook is the largest social media with 1.06 billion users. There is a tremendous amount of power put into providing over a billion people a consistent social media network, so I was interested into finding out the inner workings of Facebook. I figured Facebook would use some sort of super-fancy and high cost equipment, but was surprised that the social media giant uses a large amount of freeware equipment, and only develops new applications when absolutely necessary for the gigantic website.

On the front end, Facebook utilizes the LAMP structure which Stands for Linux, Apache, MySQL, PHP, all being integral to Facebook's architecture. Facebook runs a 64-bit Linux system along with Apache for its web servers. For its database storage, Facebook runs MySQL on over 10,000 Linux servers, distributing around 40 TB of data. These servers are part of Facebook's Santa Clara's data center, and are completely duplicated in Facebook's Virginia datacenter. MySQL is useful for Facebook because it is reliable in holding and not losing data, which is crucial as Facebook has to store an extremely large amount of data pertaining to each user. MySQL assigns a key to each item on Facebook such as photos, users, groups, and much more. Facebook uses some MySQL customizations too, most notably so that Facebook can access data from multiple data centers, which is necessary because of the large amount of content Facebook receives.

At the end of the lamp architecture is PHP, which Facebook uses to build the pages that users access. However, PHP isn't optimized for websites with a scale as large as Facebook. Along with the LAMP architecture, Facebook utilizes Memcache to retrieve user-specific information at a fast pace. Memcache reduces the load on Facebook's databases by caching data into RAM. When loading information, if the items can be found in Facebook's RAM then the information can be quickly accessed without ever going into the database. About 95 percent of data queries are filled by the cache servers RAM, so only 500,000 of the 15 million queries (numbers from 2008, but I think they illustrate the power of Memcache in Facebook's technology stack) are passed to the MySQL databases. Each of Facebook's datacenters (the last being in San Francisco) cache and file data objects, but the San Francisco datacenter uses the Santa Clara's MySQL databases. Facebook has other datacenters in Oregon, North Carolina, and Sweden.

On the back end, Facebook uses several languages such as Java, Python, C++, and Erlang. The Facebook engineers try to use open source developments whenever possible, but occasionally had to develop some tools for their specific needs. Thrift is a protocol that Facebook uses for scalable cross-language services development. It is an Apache add on that stores datatypes and service interfaces into a language-neutral file for seamless use between the multiple programming languages. Built on top of Thrift is Scribe, a log server that collects log data from a number of servers. For database use, Facebook uses Cassandra, which is a database management system designed for large data amounts spread across many servers. Cassandra is also very durable, which is important for Facebook because the loss of data

Maxwell Brown
MSU: brown308
Courseroute: MaxwellGBrown
github.com/MaxwellGBrown/tc362

would be catastrophic. Facebook also uses Hiphop for PHP, which takes PHP code and changes it into C++, and then compiles it with g++ for machine use.

All in all, Facebook is composed mostly of open source programs and codes. They make very little special applications, and try to use as much open source applications as possible, even when it might not be 100% optimal. While the main building blocks of Facebook are LAMP, Memcache, Thrift, and Scribe, I'm sure there is much more going on behind the scenes. But these pieces effectively illustrate how Facebook's technology stack keeps it working.

Sources:

http://www.makeuseof.com/tag/facebook-work-nuts-bolts-technology-explained/

http://www.infoq.com/presentations/Facebook-Software-Stack

http://www.technologyreview.com/featuredstory/410312/how-facebook-works/

http://www.dailymail.co.uk/sciencetech/article-2054168/Facebook-unveils-massive-data-center-Lulea-Sweden.html