

## Frontend and Backend Web Technologies in Social Networking Sites: Facebook as an Example

Hanin M. Abdullah

College of Information Technology  
University of Bahrain  
Sakhir, Kingdom of Bahrain  
han.eenmoh@hotmail.com

Ahmed M. Zeki

Department of Information Systems  
College of Information Technology  
University of Bahrain  
Sakhir, Kingdom of Bahrain  
amzeki@uob.edu.bh

**Abstract**—Social Networking has been pointed out as an important evolution in the social and Internet history. The leader and the pioneer of this evolution was and still Facebook despite the big and tough rivalry nowadays. Although Facebook is very well known and most of the people who surf the Internet use it and know how to use it and what the different pages are and the modules it has, a few number of technologists know the technology basis upon which this success sustains. This research paper shades a light on the web programming aspect and other supporting technologies used for the frontend and backend of social networks. Facebook is taken as an example.

**Keywords**—Frontend Web Technology; Backend Web Technology; Web Programming; Social Networks; Facebook

### I. INTRODUCTION

Web services are becoming the key technology for the implementation of integrated and distributed applications for B2B, cloud computing, and grid computing. The performance of mobile applications has critical impact on user acceptance and success. Those applications -including mobile applications- require the connectivity to backend services that are hosted in cloud environments [1,2].

In web context, frontend is the part that the user sees and interacts with, such as menus, contact forms, etc. In order to design and develop such frontend web interface, certain tools and technologies are to be used, which are usually a combination of HTML, CSS, and JavaScript all being controlled by the browser.

On the other hand, the backend usually consists of three parts: a server, an application, and a DB. Backend technologies usually consist of languages like PHP, Ruby, Python, etc.

Today's sophisticated web applications can't run without both the frontend and backend services. This paper looks into the technologies used for the frontend and the backend specifically in Social Networks (SNs) websites. Facebook will be taken as an example of the SNs. The rest of this paper is organized as follows: the next section will briefly introduce the SNs including Facebook. Section 3 focusses on the architecture of Facebook, while section 4 and 5 will discuss the Facebook frontend and backend technologies. This will be followed by future challenges and conclusion.

### II. THE LEADING SN: FACEBOOK IN BREIF

The origin of the term SNs was for a different purpose than the one used nowadays. It was originally used to distinguish between a network used for business purposes and another one used for socializing [3]. After the evolution of dynamic social websites, the term has been used to reflect those websites also; websites that have chat rooms, message boards, and the ability to have discussions and leave comments [4].

The social networking services are those online services, platforms and sites that utilize the building blocks of SN and relations between people who share different life aspects like interests, activities, backgrounds, and even connections [4].

Such services usually represent each user in a profile. The profile facilitates social links and a variety of additional services. The social services mentioned here are mostly web-based with different means for the users to interact over the Internet, such as e-mail and instant messaging [4]. However, some of the most recent services utilizes mobile platform applications with and sometimes without web based platform [4].

Some popular examples of social networking websites are: Facebook, Twitter, LinkedIn, MySpace, Pinterest, Google Plus+, DeviantArt, LiveJournal, Meetup, Tagged, Orkut, CafeMom, myLife, and Multiply [5].

Facebook is a social networking service and website intended to connect friends, family members, and business associates. Facebook is the most used SN with over 1 billion registered users worldwide [6]. It is the most common of the networking sites, with several runner ups like twitter and MySpace. It has begun as a networking website for college students and has expanded to include the whole universe [7].

A Harvard student who is very well known as the youngest billionaire in the world, Mark Zuckerberg, has founded Facebook which was originally named "theFacebook" in 2004. It has made a quick successful on Harvard campus and later expanded beyond Harvard into other universities. With the exceptional growing in popularity, the founder, Zuckerberg, added two other students, Dustin Moskovitz and Chris Hughes, to assist. After few months only, theFacebook became a college networking website for the United States of America [7].

Zuckerberg and his partners left Harvard to run theFacebook full time just after taking the site in a national level. In 2005, theFacebook has been renamed Facebook. At that time, it was only available to schools, universities, organizations, and companies with English language only, but has later expanded to include anyone everywhere [7]. An insight in the mechanisms of Facebook would show a start into which users create a profile page that shows their friends and networks information [5]. The network of friends in the profile reflects the ones who can view the profile. The profile includes information, friends, status, photos, notes, groups, and the wall. Users can search for friends by e-mail address, school, university, or just by typing in a name or location for search. Friends are able to see all of each other's profiles including contact information. Messages are allowed between users for free and recently Facebook acquires money to send messages to people who are not in the friends list. Facebook makes money also from advertisements. E-mail notifications are customizable and let users know when new friends have added them to their list or when someone has sent a message to them. A popular feature of Facebook is sharing photographs uploaded from different types of sources such as cameras and phones.

Users have the option to allow only friends to see their pictures and information or to allow anyone. Storage is unlimited, which is a major advantage of Facebook's photograph sharing capabilities as free service.

A major redesign after the first website was in 2008, which was intended to streamline the website and make it easier to see what friends were doing. It was an outstanding expenditure and has boosted its dominance in social networking. Right after that and specifically in 2009, users of Facebook worldwide were double that of its older competitor MySpace [7].

### III. THE ARCHITECTURE OF FACEBOOK

Facebook data centers are in Santa Clara, San Francisco, and Northern Virginia. Those centers were built on the backs of three major tiers of x86 servers which are loaded with open source software [8]. Some of those open source software have been developed by Facebook and made publicly open source after that. In the coming paragraphs we will be exploring the three tiers of the architecture.

The top tier of the Facebook network which is closer to users is made up of the Web servers that create the Web pages seen by users. Those servers are mostly with eight cores based on 64-bit Linux and Apache. Most of the SN's pages and features are created using PHP since the start of the Facebook [8]. PHP is a very common well known computer scripting language specialized for simple and automated functions. Although Facebook mainly utilizes PHP for its powerful capabilities, it has also developed complex core applications using a variety of different computer languages, including C++, Java, Ruby, and Python [8].

The middle tier is mainly for caching servers. Although more than 800 DB servers are available in the bottom tier, they can't serve up all the needed data [8]. To feel the difficulty, Facebook receives 15 million requests per second

for data and connections. For this reason, cache servers, running Linux and the open source Memcache software, fill this gap and receives most of the requests. Nearly 95% of data queries are being filled from the cache servers' 15 TBs of RAM. Approximately 500,000 queries per second only have to be passed to the MySQL DBs [8].

The bottom tier consists of eight-core Linux servers running MySQL as a DB server application. Facebook is estimated to have about 800 servers in this tier distributing nearly 40 TBs of user data. This tier stores all the data about every object in the DB [5]. Figure 1 shows a general overview of Facebook architecture.

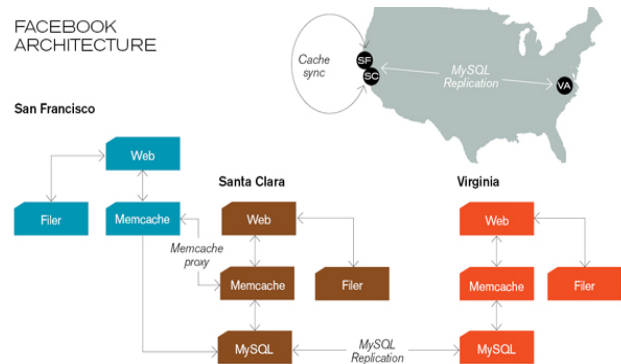


Figure 1. Facebook architecture [8].

The San Francisco facility replicates the Web and cache tiers, as well as the filers with the DB objects, but it uses the Santa Clara MySQL DB tier. The Virginia data center is too far away to share MySQL DBs with 70 milliseconds of Internet delay, give or take, it just won't work. Thus, it completely duplicates the Santa Clara -facility, using MySQL replication to keep the DB tiers in sync [8]. Facebook had more than 30,000 servers in 2012 which may increase overtime.

#### A. How do Facebook Applications Work?

Facebook applications are a little different from other online applications in that they live on two different servers [9]. For instance you host the application on your own server, which is responsible for the business logic of your application, and Facebook's servers work with the user's profile. When a user interacts with your application, Facebook interprets the request and passes it to your server. Your pages are constructed with FBML (Facebook Markup Language), a superset of HTML, to pass back to the Facebook server farm [9]. Facebook interprets the FBML response from your server and generates the resulting page the user sees. The process is illustrated in Figure 2.

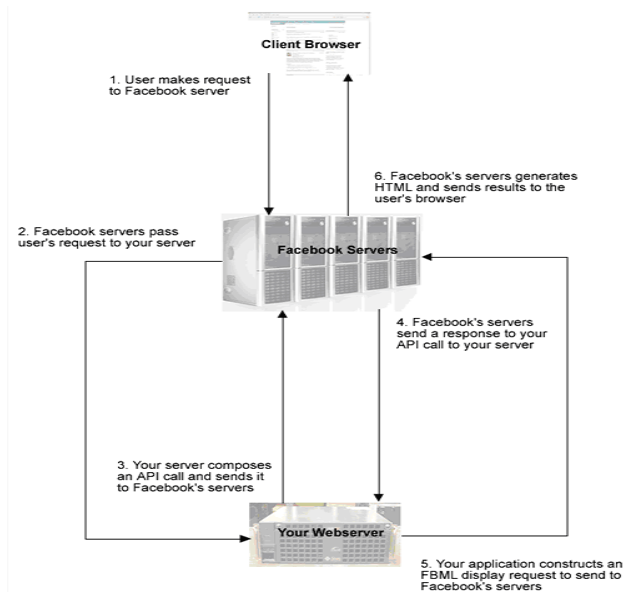


Figure 2. Facebook page processing [9].

#### IV. FACEBOOK FRONTEND TECHNOLOGIES

To make up its infrastructure, Facebook uses various tools and programming languages. At the frontend, the servers are based on LAMP (Linux, Apache, MySQL, and PHP) stack [10]. The following paragraphs explain what that means.

LAMP is a combination of free and open source software. It refers to the first letters of the OS Linux, the HTTP server Apache, the DB software MySQL, and programming languages PHP, Perl, and Python. Those are the main components used to build a viable general purpose web server in the frontend of any scalable large scale servicing website [11].

Although LAMP has the standard components which are mentioned but variations are possible with respect to the web scripting languages for instance where the Perl may be excluded. Other variations exist such as variations of the OS like the WAMP which is based on Windows OS and MAMP which is based on Mac OS although this doesn't apply in the Facebook case in which Linux is the OS [10].

The software combination has become popular between developers although the original designers of these programs did not design them all to work specifically with each other. The development philosophy and tool sets are common among several platforms because it is free of cost and open source which make them easy to adapt [10].

##### A. LINUX & APACHE

Linux is an OS which is based on a Unix kernel. In addition to the fact that it is an open source OS, what makes it favorable is the good security and customization ability of it. Facebook runs the Linux OS on Apache HTTP Servers. Apache is also free, as the rest of the bundle, and is a very popular open source web server in use [10,11,12].

##### B. MYSQL

MySQL is a DB software that has high speed and high reliability, the reason for which Facebook is using it. It is used primarily as a key-value store because data is randomly distributed among a huge set of logical instances. These instances are spread out across physical nodes, and the load balancing is done at the physical node level [13].

For customization purposes, Facebook has developed a custom partitioning scheme in which all data has an assigned global ID. Another customization of MySQL from Facebook is the archiving scheme that is based on the frequency of data for every user [11].

##### C. PHP

Facebook uses PHP as the main programming language. PHP is a web programming language with extensive support and an active development community [14]. PHP is appreciated for its rapid iterations and dynamic interpretation as a scripting language. As mentioned earlier, Facebook doesn't just use PHP of course, but it is heavily invested in the language [14].

A lot of programmers have asked that why Facebook doesn't migrate from PHP to other languages and the best answer came from a Facebook engineer named Wong who has worked at the company in various roles between 2005 and 2010. He stated that it is because PHP has incumbent inertia and Facebook's engineers have managed to work around many of its flaws through a combination of patches at all levels of the stack and excellent internal discipline via code convention and style - the worst attributes of the language are avoided and coding style is rigidly enforced through a fairly tight culture of code review. Engineering management has never had to take a strong hand here; this arose largely due to key internal technical leaders just sort of corralling everyone else along [10].

##### D. MEMCACHE

Memcache is a memory caching system that is used to speed up websites that are driven by DBs, such as Facebook, by caching data and objects in RAM in order to reduce the reading time. Facebook uses Memcache as a primary caching stage which helps in relieving the DB load, if direct access is granted for all requests. By having a caching system, Facebook can be as fast as it is in recalling the data [11].

Facebook has realized that there are downsides to using the LAMP stack. For example, PHP is not inherently optimized for large websites and therefore hard to scale. In addition, it is not the fastest executing language and the extension framework is difficult to use [14]. Keeping in mind that Memcache is considered as the middle tier as has been mentioned before although it is described here under the frontend section.

#### V. FACEBOOK BACKEND

Facebook has used different programming languages for its Backend. A variety of Facebook services uses languages such as Java, C++, Erlang and Python. Facebook doesn't use different programs for the sake of the variety but they initially think of a service which they need to implement and

then they create the framework/toolset for the service and at the end the right programming language for the function is chosen [13]. Facebook uses several different languages for its different services. PHP is used for the frontend as explained earlier, and Erlang is used for Chat, while Java and C++ are used in several places.

Next subsections will briefly go through some of the software Facebook is using in its backend.

#### A. THRIFT

Thrift is a cross-language framework which has been developed in the labs of Facebook. The function of Thrift is to tie all of the different languages Facebook uses together to make it possible for them to talk to each other. Facebook has made Thrift open source [13].

#### B. SCRIBE

Scribe is a server that is built on top of Thrift. Scribe has been designed for the log data streamed from all of the servers in real time. It aggregates the log data and it is a scalable framework useful for logging a wide array of data [15].

#### C. CASSANDRA

Cassandra is another open source system which is used by Facebook. It is a distributed storage system which is used by Facebook for its inbox search. What makes Cassandra special is that it has no single point of failure [16].

#### D. HIPHOP FOR PHP

Although PHP has a lot of benefits which made Facebook and a lot of other websites to use it, it still has drawbacks such as optimization. HipHop is a transformer which transforms the source code of PHP to an optimized C++ code to be ready for a g++ compiler to make the machine code out of it [15].

#### E. HAYSTACK

Haystack is an object store system which is used by Facebook for the storage of photos. It is a high performance system that is handling more than 20 billion uploaded photos on Facebook, and each one is saved in four different resolutions. Haystack stores photo data inside 10 GB bucket with 1 MB of metadata for every GB stored [15].

#### F. HADOOP AND HIVE

Hadoop is a map-reduce system that performs calculations on massive amount of data. Facebook uses this open source system for data analysis. Hive originated from within Facebook, and makes SQL queries with Hadoop, making it easier for non-programmers to use. Both Hadoop and Hive are open source [15].

#### G. VARNISH

Varnish is a load balancer and HTTP accelerator. It also caches the content which can then be served lightning-fast. Serving photos and profile pictures and handling billions of requests every day is done by Varnish [17].

There are a lot of more programming languages like bagpipe for dynamic pages but the scope of this report is for the general overview [15].

## VI. FUTURE CHALLENGES

More scaling challenges will come and the programming languages and technologies Facebook is using will always be revised and evolved. The pace into which social networking is growing led by Facebook is incredible. Its user base is increasing almost exponentially. Facebook is expected to be running into different performance bottlenecks as it's challenged by more and more page views, searches, uploaded images, status messages, and all the other ways that Facebook users interact with the site and each other [15]. Here are a few facts to give an idea of the scaling challenge that Facebook has to deal with [15]:

- Facebook serves 570 billion page views per month (according to Google Ad Planner).
- There are more photos on Facebook than all other photo sites combined (including sites like Flickr).
- More than 3 billion photos are uploaded every month.
- Facebook's systems serve 1.2 million photos per second, (this doesn't include the images served by Facebook's CDN).
- More than 25 billion pieces of content (status updates, comments, etc) are shared every month.

## VII. CONCLUSION

Finally, it is good to say that behind of all the well maintained and operated website that the entire world is using a very robust programs and systems that are all in harmony. Facebook is providing an easy way for friends to keep in touch and for anyone to have a presence on the Web without the need to build a website. People have been "Facebooking" each other for years now making Facebook the most used SN worldwide. Another important outcome out of this report is that Facebook has a lot of self-invented technologies and languages, some of which have been made open source and public to be customized. The reason behind this was the challenges which haven't been looked into before Facebook in terms of the huge extraordinary traffic and users.

## REFERENCES

- [1] M. Debusmann, and K. Geihs, "Towards Dependable Web Services", 10<sup>th</sup> IEEE Pacific Rim International Symposium on Dependable Computing, pp. 5-14, 2004.
- [2] F. Schulz, and W. Theilmann, "Towards Systematic Mobile Cloud Performance Analysis", 6<sup>th</sup> Joint IFIP Wireless and Mobile Networking Conference (WMNC), pp. 1-4, 2013.
- [3] A. Hashmi, F. Zaidi, A. Sallaberry, and T. Mehmood, "Are All Social Networks Structurally Similar?", IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp. 310-314, 2012.
- [4] L. Brown, "The Social Networking Handbook - Everything you Need to Know about Social Networking", Emereo Publishing, 2011.

- [5] Top 15 Most Popular Social Networking Sites, <http://www.ebizmba.com/articles/social-networking-websites>
- [6] Z. Whittaker, "Facebook Hits 1 Billion Active User Milestone", CNET, 2012, [http://news.cnet.com/8301-1023\\_3-57525797-93/facebook-hits-1-billion-active-user-milestone](http://news.cnet.com/8301-1023_3-57525797-93/facebook-hits-1-billion-active-user-milestone)
- [7] L. Li, "Social Network Sites Comparison between the United States and China: Case Study on Facebook and Renren Network", International Conference on Business Management and Electronic Information (BMEI), vol.1, pp. 825-827, 2012.
- [8] A. Zeichick, "How Facebook Works?" <http://www.technologyreview.com/featuredstory/410312/how-facebook-works>
- [9] W. Graham, "Reaching Users Through Facebook: A Guide to Implementing Facebook Athenaeum", Code4lib Journal, Issue 5, 2008.
- [10] D. Dougherty, "LAMP: The Open Source Web Platform", 2001, <http://www.onlamp.com/pub/a/onlamp/2001/01/25/lamp.html>
- [11] J. Lee, and B. Ware, "Open Source Web Development with LAMP: Using Linux, Apache, MySQL, Perl, and PHP", Addison Wesley, 2002.
- [12] Apache Software Foundation, "Axis architecture guide", 2003, Version 1.0, <http://ws.apache.org/axis>
- [13] E. Protalinski, "Why Facebook Hasn't Ditched PHP?" <http://www.zdnet.com/blog/facebook/why-facebook-hasnt-ditched-php/9536>
- [14] S. Campbell, "How Does Facebook Work? The Nuts and Bolts: Technology Explained", <http://www.makeuseof.com/tag/facebook-work-nuts-bolts-technology-explained>
- [15] Pingdom, "Exploring the Software Behind Facebook, The World's Largest Site", <http://royal.pingdom.com/2010/06/18/the-software-behind-facebook>
- [16] G. Wang, and J. Tang, "The NoSQL Principles and Basic Application of Cassandra Model", International Conference on Computer Science & Service System (CSSS), pp.1332-1335, 2012.
- [17] Varnish Cache, <https://www.varnish-cache.org/>