INTRODUCTION

- Scalability is essential for web applications with massive numbers of users.
- The high growth rate of web based systems having hundreds of thousands of users accessing it continuously led to response time problems for users who are trying to receive information at the same time.
- The current architecture of Facebook is very large and consists of many technologies and thousands of servers.
- In this presentation Facebook’s architecture and how it handles scalability issues is going to be described.

HOW FACEBOOK WORKS

- Facebook needs to handle large amounts of information every second.
- Facebook continues to be a LAMP (Linux-Apache-Memcached-PHP) website, but had to change and expand its operation to incorporate other elements and services.
- Personalized and implemented systems also exist inside Facebook, such as Haystack.

FRONT-END

- Facebook uses a variety of services, tools, and programming languages to make up its core infrastructure.
- At the front-end, their servers run a LAMP (Linux, Apache, MySQL, and PHP) stack with Memcache.

LINUX & APACHE

- Linux is a kernel operating system.
- It is open source, highly customizable and good for safety.
- Facebook uses Linux with the Apache HTTP server.
- Apache is also open source and is the most popular web server.
PHP & BigPipe

- **PHP**
  - fast iterations
  - easy to use
  - Simple to learn
  - good web programming language with extensive community support.

- **BigPipe** is a dynamic web page system developed by Facebook.
  - The general idea is to perform pipelining of sections through the implementation of various stages within web browsers and servers.

HIP-HOP

- Developed by Facebook
- PHP compiler.
- Converts PHP scripts into optimized C++ code.
- Created to minimize server resources.
- Some key features of HipHop are:
  - Easy to implement extensions
  - Reduces CPU and Memory usage significantly

MySQL

- Fast and reliable
- Thousands of MySQL servers
- Users randomly distributed across these servers
- Relational aspect of DB is not used
  - No joins. Logically difficult (Data is distributed randomly)
  - Primarily key-value store

SCRIBE

- Scalable distributed logging framework
- Useful for logging a wide array of data
- Simple data model
- Built on top of Thrift

Back-End

- At the heart of the application back end are the application servers.
- Application servers are responsible for answering all queries and take all the writes into the system.
- They also interact with a number of services to achieve this.

**SCRIBE**
- Frontend servers
- Aggregator nodes
- Hadoop HDFS

**THRIFT**
- Lightweight software framework for cross-language development
- Very quick
- Swaps information between various applications developed in different languages.
- Thrift protocol offers serialization between several languages.

**MEMCACHED**
- Key-value memory storage system for arbitrary pieces of information that result either from database research or page rendering.
- System uses client-server architecture.
- Server maintains a key-value vector association
- Client populates this vector and performs research on it.
- The server then stores these values in memory therefore decreases reading time.
- If a server consumes the available memory, it removes older values.

**HADOOP & HIVE**
- Hadoop is ideal for scalable systems with an enormous need to store and process large amounts of information.
- Scales horizontally.
- Instead of replacing existing storage systems, Hadoop complements them.
- This process allows existing systems to serve data in real time or provide transactional interactive business intelligence.

**Haystack**
- Efficient storage of billions of photos.
- High performance.
- Highly scalable.
- Uses extensive caching in its main memory.

**Facebook’s Architecture**

**Conclusions**
- Facebook has gained a tremendous amount of popularity in the last years and has become the most successful social networking website ever created.
- If we look past all of the features and innovations, the main idea behind Facebook is really very basic “keeping people connected”.
- Facebook realizes the power of social networking and is constantly innovating to keep their service the best in the business.