**How Facebook Works**

Facebook is a wonderful example of the network effect, in which the value of a network to a user is exponentially proportional to the number of other users that network has.



Facebook’s power derives from what Jeff Rothschild, its vice president of technology, calls the “social graph”–the sum of the wildly various connections between the site’s users and their friends; between people and events; between events and photos; between photos and people; and between a huge number of discrete objects linked by metadata describing them and their connections.

Facebook maintains data centers in Santa Clara, CA; San Francisco; and Northern Virginia. The centers are built on the backs of three tiers of x86 servers loaded up with open-source software, some that Facebook has created itself.

Let’s look at the main facility, in Santa Clara, and then show how it interacts with its siblings.

The top tier of the Facebook network is made up of the Web servers that create the Web pages that users see, most with eight cores running 64-bit Linux and Apache. Many of the social network’s pages and features are created using PHP, a computer scripting language specialized for simple, automated functions. But ­Facebook also develops complex core applications using a variety of full-featured computer languages, including C++, Java, Python, and Ruby. To manage the complexity of this approach, the company created Thrift, an application framework that lets programs compiled from different languages work together.

The bottom tier consists of eight-core Linux servers running MySQL, an open-source database server application. Rothschild estimates that Facebook has about 800 such servers distributing about 40 terabytes of user data. This tier stores all the metadata about every object in the database, such as a person, photo, or event.

The middle tier consists of caching servers. Even 800 database servers can’t serve up all the needed data: Facebook receives 15 million requests per second for both data and connections. Bulked-up cache servers, running Linux and the open-source Memcache software, fill the gap. About 95 percent of data queries can be filled from the cache servers’ 15 terabytes of RAM, so that only 500,000 queries per second have to be passed to the MySQL databases and their relatively slow hard drives.

Photos, videos, and other objects that populate the Web tier are stored in separate ­filers within the data center.

The San Francisco ­facility replicates the Web and cache tiers, as well as the filers with the database objects, but it uses the Santa Clara MySQL database tier.

The Virginia data center is too far away to share MySQL databases: 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 database tiers in sync.

What’s next for Facebook’s technology? For one thing, says ­Rothschild, the company has discovered that interrupts on the servers’ Ethernet controllers–which let the servers process myriad requests arriving at the same time–are a bottleneck, since they’re generally handled by only one core. So Facebook rewrote the controllers’ drivers to scale on multicore systems. Facebook is also experimenting with solid-state drives, which could speed the performance of the MySQL database tier by a factor of 100.

Given that Facebook is growing–and that connections grow exponentially–the site is going to need that performance soon.

**How Facebook News Feed Works**

**This is the** ultimate guide to how Facebook chooses what to show in your News Feed, and how you can get your content seen by more people.

Understanding how the News Feed works is tough because the algorithm is always changing. So TechCrunch launched this research project for today’s 10th anniversary of News Feed, interviewing [**Facebook’s**](https://www.crunchbase.com/organization/facebook/) team members, compiling the company’s announcements, and reviewing a decade of our coverage. The result is this helpful explainer, which we’ll keep updated as new changes roll out so it’s always accurate.

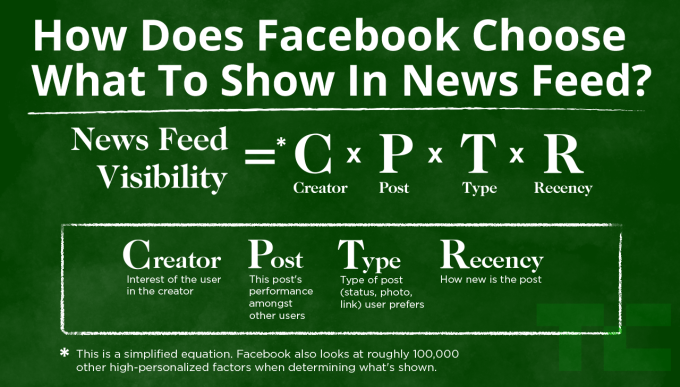
## The Goal Of News Feed

Facebook’s objective is to select the most relevant and engaging stories to show in the News Feed. It wants to choose the best content out of several thousand potential stories that could appear in your News Feed each day, and put those in the first few dozen slots that you’ll actually browse through.

These stories get ranked and shown in order of importance, from big stuff like your sibling getting married or a news article that 10 of your friends have shared, to the average links shared by brands to their websites, to boring stories like a distant acquaintance RSVPing for an event.

Facebook prioritizes stories you’ll Like, comment on, share, click, and spend time reading, which we’ll refer to as “engagement”. Facebook also runs both online surveys and offline focus groups to get more feedback about what stories people think should appear.

The more engaging the content, the more you’ll come back to Facebook, and the better it can accomplish its mission of connecting people while also earning revenue from ads shown in News Feed.



## The Natural Decline Of Reach

Over time as more people and Pages join Facebook and each shares more content, there’s more competition for the limited available space in the News Feed. While people have increased the amount of time they spend on the News Feed over the past 10 years, viewership hasn’t grown as fast as the amount of stories shared.

This causes a natural decline in the reach of what’s posted to the News Feed, in terms of the percentage of people who see a story out of everyone eligible to see it. This is why Facebook Pages see the percentage of their followers who see their content shrinking over time. It’s an inevitable result of people sharing more frequently, rather than some conspiracy of Facebook’s to force businesses to buy ads.

The best way for you to counteract this decline of reach for your own content is to learn what Facebook’s algorithm prefers.

## The Main Factors Influencing What You See

So how does Facebook’s algorithm choose what appears in what order? It assigns each story a personalized relevancy score that’s different for each person that sees it, and puts the most relevant stories first. The algorithm takes into account thousands of different signals. But here are the [four main factors](https://newsroom.fb.com/news/2016/04/news-feed-fyi-from-f8-how-news-feed-works/) that decide a story’s personalized relevancy score, and therefore its visibility to that user.

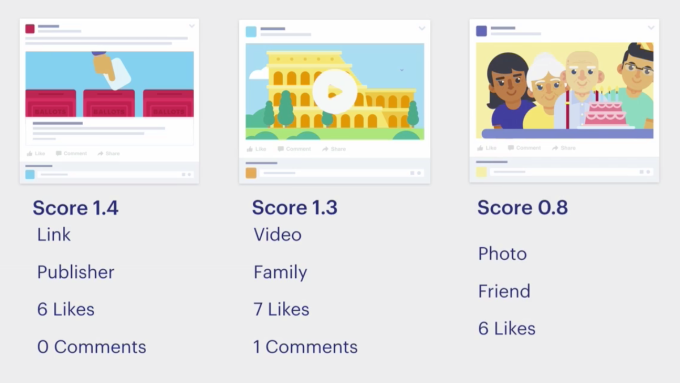


**Who Posted It** – The more you’ve interacted with a post’s author in the past, the more interested Facebook thinks you’ll be in their future posts. This interaction could be engagement such as Liking or commenting, but also clicking or slowing down to read their posts, visiting their Page or profile, tagging them or being tagged together in posts or photos, and many other actions on Facebook. This is why you don’t see post from old friends or Pages you haven’t interacted with in years.

**How Other People Engaged With The Post**– The more that other people have engaged with a particular post, the more likely that Facebook will show it to you too. Sometimes people and Pages post boring things few interact with, so it lets them sink into obscurity. But if a high percentage of people who do see a post at first do engage, Facebook knows it’s interesting and keeps showing it to more people.

**What Type Of Post Is it** – The more that you typically engage with a certain kind of post (status, link, photo, video, event, job change, content from another app), the more Facebook will show you posts of that type. Different people enjoy different kinds of posts. I might love reading news articles, you might love watching videos. Facebook matches people to post types so if you never watch videos, you won’t see as many.

**When It Was Posted** – The more recently a story was posted, the more likely you are to see it. However, Facebook also detects when you last checked the NewsFeed, and will rank older, good posts higher if you haven’t logged in since they were posted and haven’t seen them. Check every few minutes or hours, and Facebook will prioritize very recent posts. Go offline for a week and Facebook might surface a big story like your best friend having a baby even if it was posted 5 days ago.

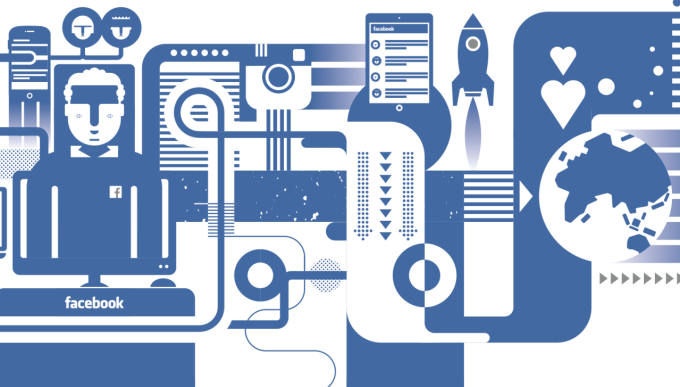


Combined, these factors have a heavy influence on how high up in the News Feed a post appears. Over time as you interact with the News Feed, Facebook learns what you care about and evolves that understanding if your behavior changes.

There are also several other significant factors that determine what you see in the News Feed, though not as heavily as those above

**How Many Others Have Posted the Same Thing**– Facebook creates highly-ranked aggregations of posts when lots of friends or Pages post the same thing. If a dozen friends have all posted the same news article or video, Facebook assumes it’s a big deal and displays a “Josh Constine and 11 other friends shared…” story higher up in the News Feed.

**New Facebook Products**– When Facebook releases a new product such as Live video or Slideshows, the company needs to test how much people want to interact with it. It may initially show too many or too few News Feed stories about the product until it receives enough feedback to learn the appropriate level of visibility.



## How Ads Get Inserted

Facebook also injects ads into the News Feed. These don’t replace any naturally visible post but instead just get injected in between them, pushing down the ones that come after.

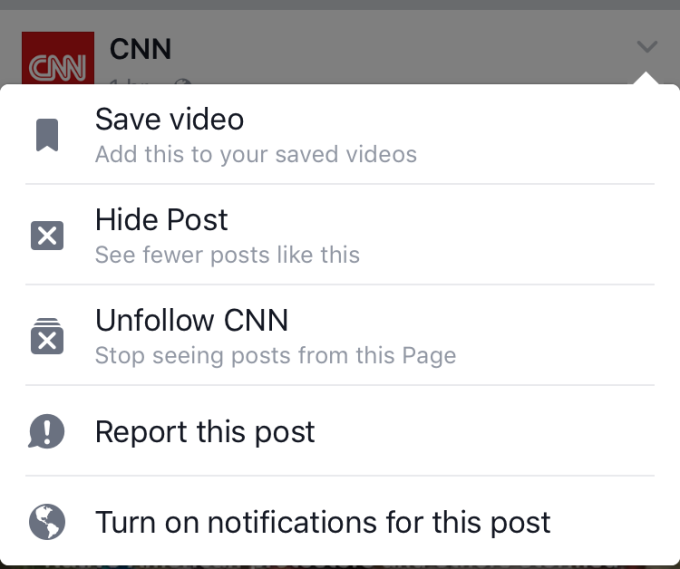
Facebook uses a similar but separate ranking algorithm to determine whether you’re likely to be interested in a Page or business’ ads. Facebook limits the number of ads you see, and therefore wants to maximize the likelihood that the ones it shows you will resonate with you or get you to click, since that’s how it earns more money.

The more Facebook knows about you, the more relevant the ads will be. If you fill out your profile and Like the Pages of things you care about, Facebook’s ads will become more personalized and relevant, informing you about products, apps, events, and more that you’re truly interested in.

## Controlling Your News Feed

Facebook gives you both implicit and explicit ways to teach the News Feed what you want to see.

Implicit signals come from your normal behavior on Facebook. If you keep Liking stories from a certain friend or about a particular topic, you’ll see more of them. If you always skim past someone’s posts or never click on the stories shared by a Page, you’ll see less of them. That’s why it’s important to actually Like things you like, and not pity-Like things you don’t really care about just to be nice to someone.



Facebook also provides explicit tools for directly telling News Feed what you do and don’t want to see. Every story has a little drop-down arrow in the top right corner that lets you:

* Hide a post so you stop seeing it and see fewer stories similar to it
* Unfollow the author so you don’t see any more of their posts
* Save the story for later, which tells Facebook to show you more posts similar to it
* Turn on notifications, so you’re alerted of future posts by that author

There’s also a See First option available in the News Feed settings. This lets you pick people or Pages whose posts you always want to see at the top of your News Feed, which can be useful for staying informed about a loved one, best friend, favorite brand, or your own business.

## An Updated List Of News Feed Algorithm Changes

Facebook is constantly tweaking the News Feed. It adapts to prevent people from gaming or tricking the system, embrace new media types, and correct flaws that lead people to see things they don’t care about. Facebook publishes News Feed FYI [blog posts](https://newsroom.fb.com/news/category/news-feed-fyi/) to be transparent about the changes.



This article will adapt too. As Facebook publishes more FYIs, we’ll add them to this list with a short summary of what each means. That way you can keep referencing this article and share it with friends or colleagues that don’t understand how News Feed works. Here are all the News Feed change announcements so far:

[**High Quality Posts From Pages**](https://newsroom.fb.com/news/2013/08/news-feed-fyi-showing-more-high-quality-content/) – Timely, relevant, from a source you trust, you would share it or recommend it, genuinely interesting and not trying to game the News Feed, not low quality or a meme, wouldn’t complain, doesn’t get hidden, complete Page profile, fan base overlaps with other high quality Pages.

[**More Relevant Ads**](https://newsroom.fb.com/news/2013/09/news-feed-fyi-more-relevant-ads-in-news-feed/) – Fewer ads that other people hide, Fewer ads similar to ones people have already hidden

[**Higher Quality News**](https://newsroom.fb.com/news/2013/12/news-feed-fyi-helping-you-find-more-news-to-talk-about/) – More links to high quality articles, Fewer links to meme photos, related articles to ones you clicked, highlighting stories with new comments

[**More Status Updates From Friends**](https://newsroom.fb.com/news/2014/01/news-feed-fyi-what-happens-when-you-see-more-updates-from-friends/)– More text status updates from friends, fewer text status updates from Pages, more link share stories from Pages, fewer text updates with embedded links from Pages

[**MoreStories About Topics You Like**](https://newsroom.fb.com/news/2014/02/news-feed-fyi-showing-stories-about-topics-you-like/)– Page posts that tag another Page may be shown to followers of the tagged Page

[**Cleaning Up News Feed Spam**](https://newsroom.fb.com/news/2014/04/news-feed-fyi-cleaning-up-news-feed-spam/) – Fewer Page posts that explicitly ask people to Like, comment, or Share. Fewer Page posts that have already been shared by that Page, fewer spammy links that use inaccurate language or formatting to trick people into clicking

[**Focusing On Explicitly Shared Stories**](https://newsroom.fb.com/news/2014/05/news-feed-fyi-giving-people-more-control-over-when-they-share-from-apps/)– More explicitly shared stories from third-party apps and fewer implicitly or automatically shared stories

[**Showing Better Videos**](https://newsroom.fb.com/news/2014/06/news-feed-fyi-showing-better-videos/)– More videos people watch and watch for a long time, more videos to people who watch videos and fewer videos to people who skip videos

[**Fighting Click Bait**](https://newsroom.fb.com/news/2014/08/news-feed-fyi-click-baiting/)– Fewer links that don’t tell people much about what they’re clicking to, fewer links to web pages where people don’t spend much time and come right back to Facebook, more links to web pages where people spend a lot of time, more links to web pages people talk about after visiting and fewer links to web pages people don’t talk about after visiting, more stories with links shared with the link format and fewer stories with links in the description or caption of a photo or video.

[**Incorporating Feedback About Ads**](https://newsroom.fb.com/news/2014/09/news-feed-fyi-listening-to-peoples-feedback-to-show-better-ads/)– Using surveys about why people hide ads, fewer similar ads to ads someone hid because it wasn’t relevant to them, fewer ads shown to anyone that people hide because they were offensive, more heavily weighting the hides by people who infrequently hide ads

[**More Timely Stories**](https://newsroom.fb.com/news/2014/09/news-feed-fyi-showing-more-timely-stories-from-friends-and-pages/) – More stories that reference current Trending Topics, more stories shown soon after they’re posted if people Like them soon after they’re posted but Like them less later

[**More Control Over What You See**](https://newsroom.fb.com/news/2014/11/news-feed-fyi-more-ways-to-control-what-you-see-in-your-news-feed/)– When you hide someone’s story, you can select to see less from that person in the future without completely unfollowing them

[**Reducing Promotional Page Posts**](https://newsroom.fb.com/news/2014/11/news-feed-fyi-reducing-overly-promotional-page-posts-in-news-feed/) – Fewer posts that solely push people to buy a product, install an app, enter a sweepstakes, or that reuse the exact same content from ads.

[**Minimizing Hoaxes**](https://newsroom.fb.com/news/2015/01/news-feed-fyi-showing-fewer-hoaxes/) – Fewer posts that people flag as hoaxes or delete after posting because they are scams or deliberately false news

[**Showing More Content From Friends**](https://newsroom.fb.com/news/2015/04/news-feed-fyi-balancing-content-from-friends-and-pages/) More posts from friends instead of Pages, fewer stories about friends Liking or commenting on a post, more posts from the same sources for new users without much content in their News Feed

[**More Stories You Spend Time Reading**](https://newsroom.fb.com/news/2015/06/news-feed-fyi-taking-into-account-time-spent-on-stories/) – More stories that other people spend significantly more time looking at in their News Feed than other stories.

[**The See First Feature**](https://newsroom.fb.com/news/2015/07/updated-controls-for-news-feed/) – A new feature lets you choose friends or Pages whose stories you want to see first at the top of your News Feed

[**Accounting For Differences In How People Hide Stories**](https://newsroom.fb.com/news/2015/07/news-feed-fyi-a-better-understanding-of-hide/) – People who hide an extremely high number of stories in their feeds including ones they’ve Liked and commented on will have their hides taken less into account by the News Feed algorithm

[**Incorporating Actions Taken On Videos**](https://newsroom.fb.com/news/2015/06/news-feed-fyi-taking-into-account-more-actions-on-videos/)– More videos that people turn on the sound for, watch full screen, or watch in high definition.

[**Improving News Feed For Slow Connectivity**](https://newsroom.fb.com/news/2015/10/news-feed-fyi-building-for-all-connectivity/)– Fewer videos and more status updates and links shown to people with slow Internet connections, re-showing stories you’ve already loaded if you have no Internet connection

[**Incorporating Reactions**](https://newsroom.fb.com/news/2015/10/news-feed-fyi-how-the-reactions-test-will-impact-ranking/) – More stories similar to ones you react (just as with Likes)

[**Surveys To Reduce Low Quality Viral Stories**](https://newsroom.fb.com/news/2015/12/news-feed-fyi-using-surveys-to-better-understand-viral-stories/) – Fewer viral stories that surveys say people would rather not see

[**Offline News Feed**](https://code.facebook.com/posts/1535185823471329) – When someone has slow connectivity, Facebook will re-rank previously downloaded stories by relevance and display them instead of a loading symbol

[**Incorporating Qualitative Feedback**](https://newsroom.fb.com/news/2016/02/news-feed-fyi-using-qualitative-feedback-to-show-relevant-stories/) – More stories that surveys and qualitative research show people would be likely to both rate highly and engage with

[**Matching Reactions And Stories**](https://newsroom.fb.com/news/2016/02/news-feed-fyi-what-the-reactions-launch-means-for-news-feed/) – Over time, Facebook hopes to show people more stories similar to the ones they React to in a certain way, so people who often use the “Haha” Reaction see more funny stories

[**Showing Live Videos When They’re Live**](https://newsroom.fb.com/news/2016/03/news-feed-fyi-taking-into-account-live-video-when-ranking-feed/)– More Live videos shown while they’re currently Live

[**Incorporating Time Spent Viewing Sites**](https://newsroom.fb.com/news/2016/04/news-feed-fyi-more-articles-you-want-to-spend-time-viewing/) – More links to Instant Articles and mobile web pages loaded inside of Facebook that people spend more time viewing, fewer posts in a row from the same Page

[**Prioritizing Friends And Family Over Pages**](https://newsroom.fb.com/news/2016/06/news-feed-fyi-helping-make-sure-you-dont-miss-stories-from-friends/)– More stories from humans you care about, and fewer stories by businesses and news outlets

[**Punishing Clickbait Headlines That Mislead Or Withhold Information**](https://newsroom.fb.com/news/2016/08/news-feed-fyi-further-reducing-clickbait-in-feed/) – Fewer news stories purposefully trick people into clicking by omitting or exaggerating core details

[**Promoting Personally Informative Stories**](https://newsroom.fb.com/news/2016/08/news-feed-fyi-showing-you-more-personally-informative-stories/)– More stories Facebook predicts will be relevant to you personally because they’re related to your interest, engage you in broader discussions, and contain news

[**Re-ranking Stories Client-Side On Slow Connections**](https://code.facebook.com/posts/663139850520576/client-side-ranking-to-more-efficiently-show-people-stories-in-feed/)– More relevant stories shown when you have a weak internet connection by [re-ranking](https://beta.techcrunch.com/2016/08/11/facebook-news-feed-algorithm/) cached and new stories to prioritize the best ones whose photos and videos have already downloaded

[**Addressing Hoaxes and Fake News**](http://newsroom.fb.com/news/2016/12/news-feed-fyi-addressing-hoaxes-and-fake-news/)– Popular, viral fake news stories flagged by users will be [sent to third-party fact checkers](https://beta.techcrunch.com/2016/12/15/facebook-now-flags-and-down-ranks-fake-news-with-help-from-outside-fact-checkers/) for evaluation. Stories confirmed as fake will be down-ranked in News Feed, and be labeled with warnings that they’re disputed. Fake news sites masquerading as legitimate publishers through domain trickery, and articles people share significantly less often after reading will be down-ranked in Feed as well

[**Re-weighting Ranking For Long Video Completion**](https://newsroom.fb.com/news/2017/01/news-feed-fyi-updating-how-we-account-for-video-completion-rates/) – Facebook uses percentage of a video watched as a signal for quality, but will [now weight completion of longer videos more heavily](https://beta.techcrunch.com/2017/01/26/facebook-will-give-some-longer-videos-a-boost-in-the-news-feed/) since watching 75% of a 10 minute video is more actual watch time than 100% of a 1 minute video

[**Promoting Authentic and Real-Time Content**](http://newsroom.fb.com/news/2017/01/news-feed-fyi-new-signals-to-show-you-more-authentic-and-timely-stories/) – Posts that are [“authentic”, and not misleading, sensational, spammy, or frequently hidden](https://beta.techcrunch.com/2017/01/31/facebook-authentic-news/) will be promoted in News Feed. Posts related to a topic that is currently going viral in real-time because lots of people are posting about it or a Page post about it is receiving lots of engagement will be shown higher in the feed while the topic is still hot

[**Reducing Links To Low-Quality Sites**](https://beta.techcrunch.com/2017/05/10/facebook-fights-foot-fungus-ads/)– Websites that contain little substantive content and are covered in pop-ups; interstitials; and aggressive, shocking, sexual, or misleading ads will be downranked in News Feed and will not be allowed to Facebook buy ads

[**Downranking Clickbait Post-By-Post In 9 More Languages**](https://beta.techcrunch.com/2017/05/17/facebook-anti-clickbait/)– Link posts with clickbait headlines that withhold or exaggerate information will appear lower in the News Feed, and Facebook can now detect these in 9 languages beyond English.

[**Demoting Low-Quality Links Overshared By Spammers**](https://newsroom.fb.com/news/2017/06/news-feed-fyi-showing-more-informative-links-in-news-feed/) – Links shared by [people who aggressively share links, often more than 50 per day, will be shown less prominently in the News Feed](https://beta.techcrunch.com/2017/06/30/facebook-spammers/) as these links are often clickbait, false news, or sensational.

[**Displaying Related Articles To Combat Hoaxes And Polarization**](https://beta.techcrunch.com/2017/08/03/facebook-related-articles/) – Links that a lot of people on Facebook are talking about or that have been reviewed by third-party fact checkers will show Related Articles before they’re clicked. These Related Articles will show alternative takes on the same topic by different news sources, or truthiness reports from fact checkers.

[**Prioritizing Links To Faster-Loading Mobile Websites**](https://beta.techcrunch.com/2017/08/02/facebook-will-soon-start-ranking-faster-loading-webpages-higher-in-news-feed/)– Links to mobile sites that load quickly, including Facebook’s Instant Articles, will appear more prominently in the News Feed.

[**Banning Cloaked Spam Sites**](https://beta.techcrunch.com/2017/08/09/facebook-cloaking/) – Pages and accounts will be deactivated if they share links or ads that point to landing pages that use “cloaking” to show Facebook’s content moderators an innocent version of a site while showing everyone else spam, scams, and porn.

[**Readability To Reduce Hoax Sites**](https://beta.techcrunch.com/2017/08/15/facebook-instagram-comments/)– News Feed will now have a more lightweight, readable design with bigger link previews that put the URL domain above the headline so people are more likely to notice and avoid spoofed URLs.

[**Downranking Fake Videos And Play Buttons**](https://beta.techcrunch.com/2017/08/17/facebook-fake-play-buttons/) – Pages that publisher links with play buttons in the preview thumbnail that make the post look like a native video or static images uploaded as videos will have their posts shown to a lot fewer people.

[**Blocking Ad Buys From Pages That Share Fake News**](https://beta.techcrunch.com/2017/08/28/facebook-fake-news-ads/) – Pages that repeatedly share links to articles labeled as false news by Facebook’s outside fact checkers will no longer be able to buy any Facebook ads.

[**Showing Information Buttons About Publishers**](https://beta.techcrunch.com/2017/10/05/facebook-article-information-button/) – Links to news article will now include an information button that can be clicked to see the start of the Wikipedia entry about the publisher to help people avoid sharing stories from phony or disreputable outlets.

[**Releasing Publisher Guidelines**](https://beta.techcrunch.com/2017/10/24/facebook-publisher-guidelines/)–News outlets can now look to Facebook’s publisher guidelines to learn how to succeed in News Feed, which include sharing accurate, meaningful, and safe content.

[**Upranking Repeatedly-Viewed Watch Shows**](https://beta.techcrunch.com/2017/12/14/facebook-pre-roll-ads/) – Facebook’s original programming Watch shows will get better feed placement if they maintain consistent viewership, or if a user follows the creator’s Page but not the Show’s page.

[**Adding A Snooze Button**](https://beta.techcrunch.com/2017/12/15/facebook-adds-a-snooze-button-for-muting-people-groups-and-pages-for-30-days/) – Users can now tap the Snooze button in the drop-down arrow on a News Feed story to hide the story’s author from their feed for 30 days, as an alternative to permanently unfollowing them.

[**Demoting Engagement Bait**](https://beta.techcrunch.com/2017/12/18/facebook-is-clamping-down-on-posts-that-shamelessly-beg-for-your-engagement/)– Posts that shamelessly beg for clicks with calls to action such as “Share with friends to win a free trip” or “Like if you’re an Aries” will get less visibility in News Feed.

[**Replacing Disputed Flags With Related Articles**](https://beta.techcrunch.com/2017/12/20/facebook-will-ditch-disputed-flags-on-fake-news-and-display-links-to-trustworthy-articles-instead/) – News articles that are labeled false by third-party fact checkers will no longer show red flags that can actually entrench people’s misguided beliefs, and instead show Related Articles from other sources to give users more accurate alternate perspectives.

[**Encouraging Time Well Spent**](https://beta.techcrunch.com/2018/01/11/facebook-time-well-spent/)– In a massive overhaul to News Feed’s algorithm, Facebook will now show more content people actively discuss through comments and shares such as posts from friends, and show less content people passively consume including publicly shared publisher content, news links, and videos. Facebook’s goal is to increase time well spent and well-being, even at the cost of total time spent and ad views in the short term.

[**Promoting Trusted Sources**](https://newsroom.fb.com/news/2018/01/trusted-sources/) – Facebook will survey users to find out which news outlets are the most broadly considered reputable, and show more links to these publishers and less from those people consider inaccurate or distrustful.

—

We’ll add more News Feed FYIs as they’re published to keep this list up to date.

## News Feed Strategy

The best tactics for appearing prominently in the News Feed end up being quite straight-forward: share things that are interesting, authentic, and resonate with your audience. That typically means visually compelling media, funny or emotional content, and important news that’s fascinating to a wide audience.

Avoid overly self-promotional spam, dry or long-winded content, and boring looking media that only appeals to a fraction of your audience.

So before you post, ask yourself, is this actually interesting or entertaining to other people? Or are you just vainly bragging about your life or greedily marketing your business?

Facebook’s News Feed algorithm is complex, but the humans it serves are still pretty simple. We all just want to be stimulated. Do that, and Facebook will share what you have to say.

[**How Does Facebook Work? The Nuts and Bolts [Technology Explained]**](https://www.makeuseof.com/tag/facebook-work-nuts-bolts-technology-explained/)

Social networking is the art of connecting with those who share common interests. Your “˜network’ is a community that helps keep you united with others and offers many benefits. Networking via social media sites has revolutionized how we use the Internet and is at the forefront of what we now call Web 2.0.

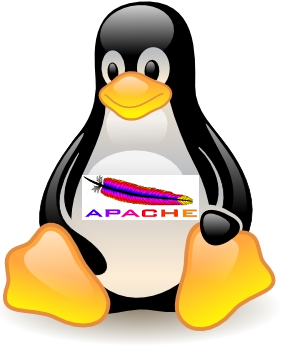
Facebook is social networking. People have been “facebooking” each other for about 6 years now, making [**Facebook**](http://facebook.com/) the most used social network with over 350 million users worldwide. But how does Facebook work?

In this article, I will discuss Facebook’s inner workings, covering its [**architecture**](http://www.infoq.com/presentations/Facebook-Software-Stack) and frontend/backend infrastructure””the nuts and bolts that hold Facebook together.

## **How Does Facebook Work?””The 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. Not a computer science expert? Let’s take a look at exactly what that means.

### **Linux & Apache**



This part is pretty self-explanatory. [**Linux**](http://www.linux.org/) is a Unix-like computer operating system kernel. It’s open source, very customizable, and good for security. Facebook runs the Linux operating system on Apache HTTP Servers. [**Apache**](http://www.apache.org/) is also free and is the most popular open source web server in use.

### **MySQL**



For the database, Facebook utilizes MySQL because of its speed and reliability. [**MySQL**](http://www.mysql.com/) is used primarily as a key-value store as data is randomly distributed amongst a large set of logical instances. These logical instances are spread out across physical nodes and load balancing is done at the physical node level.

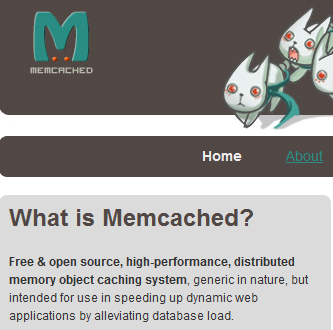
As far as customizations are concerned, Facebook has developed a custom partitioning scheme in which a global ID is assigned to all data. They also have a custom archiving scheme that is based on how frequent and recent data is on a per-user basis. Most data is distributed randomly.

### **PHP**



Facebook uses PHP because it is a good web programming language with extensive support and an active developer community and it is good for rapid iteration. [**PHP**](http://www.php.net/) is a dynamically typed/interpreted scripting language.

### **Memcache**



[**Memcache**](http://memcached.org/) is a memory caching system that is used to speed up dynamic database-driven websites (like Facebook) by caching data and objects in RAM to reduce reading time. Memcache is Facebook’s primary form of caching and helps alleviate the database load.

Having a caching system allows Facebook to be as fast as it is at recalling your data. If it doesn’t have to go to the database it will just fetch your data from the cache based on your user ID.

### **Downsides to Using LAMP**

Facebook has realized that there are downsides to using the LAMP stack. Notably, PHP is not necessarily optimized for large websites and therefore hard to scale. Also, it is not the fastest executing language and the extension framework is difficult to use.



[**Mike Schroepfer**](http://www.facebook.com/schrep), Facebook’s Vice President of Engineering, recently did an [**interview**](https://www.technologyreview.com/emtech/videos/video.aspx?SkipLog=T) at [**EmTech@MIT**](http://www.technologyreview.com/emtech/09/) concerning this. “Scaling any website is a challenge,” Schroepfer said, “but scaling a social network has unique challenges.”

He went on to say that unlike other websites, you can’t just add more servers to solve the problem because of Facebook’s “huge interconnected dataset.” New connections are created all the time due to user activity.

Facebook has grown so quickly that they are often faced with issues regarding database queries, caching, and storage of data. Their database is huge and [**largely complex**](http://www.flickr.com/photos/ikhnaton2/533233247/). To account for this, Facebook has started a lot of open source projects and backend services.

## **How Does Facebook Work?””The Back End**

Facebook’s backend services are written in a variety of different programming languages including C++, Java, Python, and Erlang. Their philosophy for the creation of services is as follows:

1. Create a service **if needed**

2. Create a framework/toolset for easier creation of services

3. Use the right programming language for the task

A list of all of Facebook’s open source developments [**can be found here**](http://developers.facebook.com/opensource.php). I will discuss a few of the essential tools that Facebook has developed.

### **Thrift (protocol)**

[**Thrift**](http://incubator.apache.org/thrift/) is a lightweight remote procedure call framework for scalable cross-language services development. Thrift supports C++, PHP, Python, Perl, Java, Ruby, Erlang, and others. It’s quick, saves development time, and provides a division of labor of work on high-performance servers and applications.

### **Scribe (log server)**

Scribe is a server for aggregating log data streamed in real-time from many other servers. It is a scalable framework useful for logging a wide array of data. It is built on top of Thrift.

### **Cassandra (database)**



[**Cassandra**](http://incubator.apache.org/cassandra/) is a database management system designed to handle large amounts of data spread out across many servers. It powers Facebook’s Inbox Search feature and provides a structured key-value store with eventual consistency.

### **HipHop for PHP**

HipHop for PHP is a source code transformer for PHP script code and was created to save server resources. HipHop transforms PHP source code into optimized C++. After doing this, it uses g++ to compile it to machine code.

## **Conclusion**

In a nutshell, that’s Facebook. This article could easily be 37 pages longer if I were to go into more detail, but to answer the question “How does Facebook work?” I think this will suffice. If you 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.

Did you find this article useful? Leave your thoughts, comments, and ideas below!

## **Exploring the software behind Facebook, the world’s largest site**

At the scale that [Facebook](http://facebook.com/) operates, a lot of traditional approaches to serving web content break down or simply aren’t practical.

The challenge for Facebook’s engineers has been to keep the site up and running smoothly in spite of handling close to **half a billion active users**. This article takes a look at some of the software and techniques they use to accomplish that.

### **Facebook’s scaling challenge**

Before we get into the details, here are a few factoids to give you an idea of the scaling challenge that Facebook has to deal with:

* 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.
* Facebook has more than **30,000 servers** (and this number is from last year!)

***Update:*** Check out [*this blog post*](https://royal.pingdom.com/2017/05/10/social-media-in-2017/) for the State of Social Media in 2017, a lot has happened since 2010 to say the least…

### **Software that helps Facebook scale**

In some ways Facebook is still a LAMP site (kind of), but it has had to change and extend its operation to incorporate a lot of other elements and services, and modify the approach to existing ones.

For example:

* Facebook still uses PHP, but it has built a compiler for it so it can be turned into native code on its web servers, thus boosting performance.
* Facebook uses Linux, but has optimized it for its own purposes (especially in terms of network throughput).
* Facebook uses MySQL, but primarily as a key-value persistent storage, moving joins and logic onto the web servers since optimizations are easier to perform there (on the “other side” of the Memcached layer).

Then there are the custom-written systems, like Haystack, a highly scalable object store used to serve Facebook’s immense amount of photos, or Scribe, a logging system that can operate at the scale of Facebook (which is far from trivial).

But enough of that. Let’s present (some of) the software that Facebook uses to provide us all with the world’s largest social network site.

#### **MEMCACHED**

[Memcached](http://memcached.org/) is by now one of the most famous pieces of software on the internet. It’s a distributed memory caching system which Facebook (and a ton of other sites) use as a caching layer between the web servers and MySQL servers (since database access is relatively slow). Through the years, Facebook has made a ton of optimizations to Memcached and the surrounding software (like optimizing the network stack).

Facebook runs thousands of Memcached servers with tens of terabytes of cached data at any one point in time. It is likely the world’s largest Memcached installation.

#### **HIPHOP FOR PHP**

PHP, being a scripting language, is relatively slow when compared to code that runs natively on a server. [HipHop](http://wiki.github.com/facebook/hiphop-php/) converts PHP into C++ code which can then be compiled for better performance. This has allowed Facebook to get much more out of its web servers since Facebook relies heavily on PHP to serve content.

A small team of engineers (initially just three of them) at Facebook spent 18 months developing HipHop, and it is now live in production.

#### **HAYSTACK**

[Haystack](http://www.facebook.com/note.php?note_id=76191543919) is Facebook’s high-performance photo storage/retrieval system (strictly speaking, Haystack is an object store, so it doesn’t necessarily have to store photos). It has a ton of work to do; there are more than 20 billion uploaded photos on Facebook, and each one is saved in four different resolutions, resulting in more than 80 billion photos.

And it’s not just about being able to handle billions of photos, performance is critical. As we mentioned previously, Facebook serves around 1.2 million photos per second, a number which doesn’t include images served by Facebook’s CDN. That’s a staggering number.

#### **BIGPIPE**

[BigPipe](http://www.facebook.com/notes/facebook-engineering/bigpipe-pipelining-web-pages-for-high-performance/389414033919) is a dynamic web page serving system that Facebook has developed. Facebook uses it to serve each web page in sections (called “pagelets”) for optimal performance.

For example, the chat window is retrieved separately, the news feed is retrieved separately, and so on. These pagelets can be retrieved in parallel, which is where the performance gain comes in, and it also gives users a site that works even if some part of it would be deactivated or broken.

#### **CASSANDRA**

Cassandra[Cassandra](http://cassandra.apache.org/) is a distributed storage system with no single point of failure. It’s one of the poster children for the NoSQL movement and has been made open source (it’s even become an Apache project). Facebook uses it for its Inbox search.

Other than Facebook, a number of other services use it, for example Digg. We’re even considering some uses for it here at Pingdom.

#### **SCRIBE**

[Scribe](http://github.com/facebook/scribe) is a flexible logging system that Facebook uses for a multitude of purposes internally. It’s been built to be able to handle logging at the scale of Facebook, and automatically handles new logging categories as they show up (Facebook has hundreds).

#### **HADOOP AND HIVE**

Hadoop[Hadoop](http://hadoop.apache.org/) is an open source map-reduce implementation that makes it possible to perform calculations on massive amounts of data. Facebook uses this for data analysis (and as we all know, Facebook has massive amounts of data). [Hive](http://hadoop.apache.org/hive/) originated from within Facebook, and makes it possible to use SQL queries against Hadoop, making it easier for non-programmers to use.

Both Hadoop and Hive are open source (Apache projects) and are used by a number of big services, for example Yahoo and Twitter.

#### **THRIFT**

Facebook uses several different languages for its different services. PHP is used for the front-end, Erlang is used for Chat, Java and C++ are also used in several places (and perhaps other languages as well). [Thrift](http://incubator.apache.org/thrift/) is an internally developed cross-language framework that ties all of these different languages together, making it possible for them to talk to each other. This has made it much easier for Facebook to keep up its cross-language development.

Facebook has made Thrift open source and support for even more languages has been added.

#### **VARNISH**

Varnish[Varnish](http://varnish-cache.org/) is an HTTP accelerator which can act as a load balancer and also cache content which can then be served lightning-fast.

Facebook uses Varnish to serve photos and profile pictures, handling billions of requests every day. Like almost everything Facebook uses, Varnish is open source.

### **Other things that help Facebook run smoothly**

We have mentioned some of the software that makes up Facebook’s system(s) and helps the service scale properly. But handling such a large system is a complex task, so we thought we would list a few more things that Facebook does to keep its service running smoothly.

#### **GRADUAL RELEASES AND DARK LAUNCHES**

Facebook has a system they called Gatekeeper that lets them run different code for different sets of users (it basically introduces different conditions in the code base). This lets Facebook do gradual releases of new features, A/B testing, activate certain features only for Facebook employees, etc.

Gatekeeper also lets Facebook do something called “dark launches”, which is to activate elements of a certain feature behind the scenes before it goes live (without users noticing since there will be no corresponding UI elements). This acts as a real-world stress test and helps expose bottlenecks and other problem areas before a feature is officially launched. Dark launches are usually done two weeks before the actual launch.

#### **PROFILING OF THE LIVE SYSTEM**

Facebook carefully monitors its systems (something we here at Pingdom of course approve of), and interestingly enough it also monitors the performance of every single PHP function in the live production environment. This profiling of the live PHP environment is done using an open source tool called [XHProf](http://pecl.php.net/package/xhprof).

#### **GRADUAL FEATURE DISABLING FOR ADDED PERFORMANCE**

If Facebook runs into performance issues, there are a large number of levers that let them gradually disable less important features to boost performance of Facebook’s core features.

#### **THE THINGS WE DIDN’T MENTION**

We didn’t go much into the hardware side in this article, but of course that is also an important aspect when it comes to scalability. For example, like many other big sites, Facebook uses a CDN to help serve static content. And then of course there is [the huge data center](http://www.facebook.com/prinevilledatacenter) Facebook is building in Oregon to help it scale out with even more servers.

And aside from what we have already mentioned, there is of course a ton of other software involved. However, we hope we were able to highlight some of the more interesting choices Facebook has made.

### **Facebook’s love affair with open source**

We can’t complete this article without mentioning how much Facebook likes open source. Or perhaps we should say, “loves”.

Not only is Facebook using (and contributing to) open source software such as Linux, Memcached, MySQL, Hadoop, and many others, it has also made much of its internally developed software available as open source.

Examples of open source projects that originated from inside Facebook include HipHop, Cassandra, Thrift and Scribe. Facebook has also open-sourced Tornado, a high-performance web server framework developed by the team behind FriendFeed (which Facebook bought in August 2009).

(A list of open source software that Facebook is involved with can be found on [Facebook’s Open Source page](http://facebook.com/opensource).)

### **More scaling challenges to come**

Facebook has been growing at an incredible pace. Its user base is increasing almost exponentially and is now close to half a billion active users, and who knows what it will be by the end of the year. The site seems to be growing with about 100 million users every six months or so.

Facebook even has a dedicated “growth team” that constantly tries to figure out how to make people use and interact with the site even more.

This rapid growth means that Facebook will keep running into various 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.

But this is just a fact of life for a service like Facebook. Facebook’s engineers will keep iterating and coming up with new ways to scale (it’s not just about adding more servers). For example, Facebook’s photo storage system has already been completely rewritten several times as the site has grown.

So, we’ll see what the engineers at Facebook come up with next. We bet it’s something interesting. After all, they are scaling a mountain that most of us can only dream of; a site with more users than most countries. When you do that, you better get creative.