

Abstract

According to web100.com, YouTube is the ninth-best website on the list of top 100 websites in the world. It is also the largest online video-sharing platform in the world. Such a large site will require large online storage space and extensive management of the website. The purpose of this report is to highlight the things required for the proper functioning of YouTube. The research method involved secondary data collected from various websites. The problems that YouTube site administrators face in regard to configuration, performance, optimization, fault, and security, as well as social, political, privacy, and ethical and legal concerns were deduced and stated.

Keywords – configuration, ethics, Google, legal concerns, privacy, security, YouTube.

I. PROBLEMS THAT YOUTUBE SITE ADMINISTRATORS FACE

YouTube is the world's most popular destination for video content online, with users watching 4 billion hours' worth of video per month and uploading 72 hours' worth of video per minute (YouTube, 2013, as cited in Alias, Razak, elHada, Kunjambu, & Muniandy, 2013). Since its launch in 2005, the site has gained popularity by enabling users to upload, view, rate, comment, share, report, and subscribe to other users. Content creators can monetize their uploaded videos as a result of running advertisements alongside the uploaded video to be watched by viewers around the world. Such a complex website will require a lot of resources such as lines of code, hardware (e.g. storage), and employees. This paper will highlight the problems that YouTube site administrators face in terms of configuration, performance and optimization, fault and security, as well as discuss the social and political issues, privacy, and ethical and legal concerns.

II. METHODOLOGY

This report exclusively involved secondary data collection. The research method mostly involved the use of Google search engine searches which resulted in links to the various websites referenced including Wikipedia, YouTube, and the Google Data Centers website. The rationale behind using Google to search for resources is simply because Google offers a vast pool of articles on this subject matter and provides open access to significant research in this domain.

III. RESULTS AND FINDINGS

This section is dedicated to the presentation of research findings and analysis. It delineates the key findings regarding the problems that YouTube site administrators face in terms of configuration, performance and optimization, and fault and security.

Configuration

YouTube was originally created by three former PayPal employees – Chad Hurley, Steve Chen, and Jawed Karim – and in November 2006, Google bought YouTube for \$1.65 billion USD (Wikipedia [a], 2020). So, YouTube makes use of the same servers as Google. Google has four data centers in Asia, five in Europe, nine in North America, and two in South America. The exact number of servers at Google's data centers is not known. However, Gartner – a global

research and advisory firm – in a July 2016 report, estimated that Google had about 2.5 million servers around the world (Data Center Knowledge, 2017). Google maintains these servers in different locations to enable load balancing; an efficient distribution of application or network traffic across multiple servers (CITRIX, n.d.). NetScaler is used to achieve this load balancing and to cache static content (High Scalability [a], 2020). Also, storing data in multiple systems help to avoid a single point of failure.

In terms of Operating System (OS), Google's YouTube uses their own version of Linux nicknamed Goobuntu (Vaughan-Nichols, 2012). Goobuntu is used as an OS for their desktops as well as their servers. Even though Windows 10 is the most popular OS for desktop and laptop computers in the world, Linux is the most preferred OS for managing web servers (Computer Hope, 2019). YouTube uses Google File System to provide efficient storage and reliable access to data using large clusters of commodity hardware to get the greatest amount of useful computation at low cost (Wikipedia [b], 2019).

YouTube makes use of Apache as a server and most lines of code are written using python (High Scalability [b], 2012). They make use of MySQL as their database; when a user watches a video, they are getting data from MySQL (High Scalability [b], 2012). They also use a set of software called "Vitess" (written with "Go" code), to enable open-source MySQL databases to work more efficiently in large-scale production environments (Jackson, 2012).

Accounting

Accounting management is concerned with tracking network utilization information in order to determine how to distribute resources among enterprise subscribers (TECH-FAQ, n.d.). About 58% of downstream traffic on the internet is video streaming including on YouTube (Wilson, 2020). YouTube offers its users the ability to conserve bandwidth by offering different video resolutions from the minimum YouTube video size (240p) to the maximum video quality (4k or 2160p). Thus, one way for a user to conserve bandwidth usage is to watch videos at a lower resolution.

Performance and optimization

Web performance refers to the speed in which web pages are downloaded and displayed on the user's web browser while website optimization refers to the field knowledge of increasing web performance (Wikipedia [c], 2019). YouTube allows users to embed YouTube videos on their personal websites. Unfortunately, embedding too many YouTube videos can slow the website down due to multiple HTTP requests and downloads. A test carried out by Website Optimization (2016) showed that multiple videos on a page can slow it down and optimization can help speed up the page by a factor of 4.8. So, one way to optimize the performance of YouTube videos embedded on a website is to display a static image from the video at first, then load the video dynamically once the user hits the play button.

Fault and Security

In October 2013, the Washington Post reported that the United States' (U.S) National Security Agency intercepted communications between Google's data centers, as part of a surveillance program called "MUSCULAR" (Gellman & Soltani, 2013). This data breach was possible because Google did not encrypt the data passed inside its own network. In response, Google began encrypting data sent between data centers in 2013 (Wikipedia [d], 2020). Also, all

the data are distributed across many computers in different locations, then chunked and replicated over multiple systems to avoid a single point of failure (Google [a], n.d.). The data chunks are named randomly as an extra measure of security which makes them unreadable to humans (Google [a], n.d.). Google servers automatically back-up critical user data so that when an accident happens, the data can be recovered (Google [a], n.d.). In addition, if a disruption such as a fire outbreak occurs at a data center, data access is shifted automatically and seamlessly to another center for uninterrupted access (Google [a], n.d.). In the event of power failure, emergency backup generators continue to power Google data centers (Google [a], n.d.).

I.V DISCUSSION

This section is focused on examining the social and political issues, privacy, and ethical and legal concerns of YouTube.

Social and political issues

YouTube has been used to improve the socio-economic well-being of people. Over the years, both private individuals and corporate entities have used YouTube to grow their audiences at very little cost. Nowadays, many traditional media have YouTube accounts and have accrued substantial following than could have been attainable by only using television. YouTube's revenue-sharing "Partner Program" allows video producers to earn a substantial living. It's top five hundred partners each earn more than \$100,000 annually[271] and its ten highest-earning channels gross from \$2.5 million to \$12 million (Wikipedia [a], 2020).

YouTube has been used to broaden awareness of social issues. For example, the anti-bullying "It Gets Better Project" expanded from a single YouTube video directed to depressed or suicidal LGBT teens (Wikipedia [e], 2020). Within weeks, hundreds of "It Gets Better" response videos were uploaded to the project by people of various levels including celebrities and U.S. President Barack Obama, White House staff, and several cabinet secretaries (Wikipedia [e], 2020).

Some YouTube videos have also influenced world events negatively. In 2012, a YouTube video called "Innocence of Muslims", produced privately by an Egyptian residing in the U.S was interpreted by some Muslims to be blasphemous in the depiction of Muhammad (Wikipedia [a], 2020). This spurred protests and lead to the bombing of the U.S embassy in Libya and the killing of the U.S Ambassador to Libya.

Privacy

YouTube and Google's privacy policies are tied together. Google collects information about its users in order to target them with ads or videos that they may be interested in. When the user is not signed in, Google stores the information they collect with unique identifiers tied to the browser, application, or device the user is using (Google [b], n.d.). When the user is signed in, Google collects information that the user stores with their Google account, which is treated as personal information (Google [b], n.d.). The YouTube website allows users to manage their subscriptions and keep their video playlists private. Users can also request personal information or a video of them uploaded by someone else without their consent to be removed.

Ethical and Legal concerns

YouTube has a list of community guidelines regarding who can use YouTube and how to use YouTube. Copyright laws are recognized by YouTube and both users and YouTube are responsible for making sure uploaded videos are copyright compliant. For example, when a user posts someone else's video, they are not allowed to monetize it without authorization from the original owner. The original owner can also request YouTube to take down the video. In 2011, YouTube created an automatic detection software named "Content ID" to detect uploaded videos that infringe copyright (Wikipedia [a], 2020). This creation of Content ID was necessary in order to avoid lawsuits from individuals or companies that could accuse YouTube of profiting from content that it did not have the right to distribute.

YouTube has faced criticism over the handling of offensive content in some of its videos that contained defamation, pornography, or material encouraging criminal conduct. YouTube relies on its users to flag the content of the videos as inappropriate so that they can be reviewed by YouTube for whether they violate YouTube's community guidelines or not (Wikipedia [a], 2020). There are also complaints that YouTube's recommendation tends to suggest more radical videos (Wikipedia [a], 2020). This raises ethical concerns regarding how YouTube's algorithm works. However, YouTube officials have denied that they intentionally promote extremist videos (Wikipedia [a], 2020).

V. CONCLUSION

YouTube has evolved from a video sharing website to a platform for creating user-generated content that can be monetized. Google owns YouTube so; YouTube uses Google Data Centers which are spread around the world. Data is distributed across these data centers so that if one computer or entire data center is down, other data centers can serve as a backup. The data is encrypted before being sent between data systems so that they are not human-readable peradventure the data is intercepted by an unauthorized third party. YouTube as been used for both positive and negative things – positive things like "anti-bullying awareness" and negative things like promoting extremist views. YouTube encourages users to flag down videos or comments that do not correlate with YouTube's community guidelines. Google collects user's browser information or personal information in order to personalize ads. YouTube content ID was created to help automatically detect videos that violate copyright laws. Overall, YouTube and its users are a community that has to work together for the betterment of YouTube's services.

References

- Alias, N., Razak, S., elHadad, G., Kunjambu, N., & Muniandy, P. (2013). A Content Analysis in the Studies of YouTube in Selected Journals. *Procedia - Social and Behavioral Sciences*, 10-18.
doi:<https://doi.org/10.1016/j.sbspro.2013.10.301>
- CITRIX. (n.d.). *What is load balancing?* Retrieved from CITRIX: <https://www.citrix.com/glossary/load-balancing.html>
- Computer Hope. (2019). *What is the most popular operating system?* Retrieved from Computer Hope: <https://www.computerhope.com/issues/ch001777.htm>
- Data Center Knowledge. (2017). *Google Data Center FAQ: How Many Servers Does Google Have?* Retrieved from Data Center Knowledge: <https://www.datacenterknowledge.com/archives/2017/03/16/google-data-center-faq>
- Gellman, B., & Soltani, A. (2013). *NSA Infiltrates Links to Yahoo, Google Data Centers Worldwide, Snowden Documents Say*. Retrieved from The Washington Post: https://www.washingtonpost.com/world/national-security/nsa-infiltrates-links-to-yahoo-google-data-centers-worldwide-snowden-documents-say/2013/10/30/e51d661e-4166-11e3-8b74-d89d714ca4dd_story.html
- Google [a]. (n.d.). *Data and Security*. Retrieved from Google Data Centers: <https://www.google.com/about/datacenters/data-security/>
- Google [b]. (n.d.). *We want you to understand the types of information we collect as you use our services*. Retrieved from Google Privacy & Terms: <https://policies.google.com/privacy?hl=en#infocollect>
- High Scalability [a]. (2020). *YouTube Architecture*. Retrieved from High Scalability: <http://highscalability.com/youtube-architecture>
- High Scalability [b]. (2012). *7 Years Of YouTube Scalability Lessons In 30 Minutes*. Retrieved from High Scalability: <http://highscalability.com/blog/2012/3/26/7-years-of-youtube-scalability-lessons-in-30-minutes.html>
- Jackson, J. (2012). *YouTube scales MySQL with Go code*. Retrieved from ComputerWorld: <https://www.computerworld.com/article/2493815/youtube-scales-mysql-with-go-code.html>
- TECH-FAQ. (n.d.). *FCAPS: Accounting Management*. Retrieved from TECH-FAQ: <https://www.tech-faq.com/fcaps.shtml>
- Vaughan-Nichols, S. (2012). *The Truth About Goobuntu: Google's in-house desktop Ubuntu Linux*. Retrieved from ZDNet: <https://www.zdnet.com/article/the-truth-about-goobuntu-googles-in-house-desktop-ubuntu-linux/>
- WEB100 STAFF. (2009). *Entertainment, Music, and Video: YouTube*. Retrieved from Web100.com: <http://www.web100.com/2009/01/21/web-100/entertainment-music-and-video-web/youtube/>
- Website Optimization. (2016). *Optimizing YouTube Video Loading Speed*. Retrieved from Website Optimization: <http://www.websiteoptimization.com/speed/tweak/youtube-videos/>

Wikipedia [a]. (2020). *YouTube*. Retrieved from Wikipedia: <https://en.wikipedia.org/wiki/YouTube>

Wikipedia [b]. (2019). *Google File System*. Retrieved from Wikipedia:
https://en.wikipedia.org/wiki/Google_File_System

Wikipedia [c]. (2019). *Web Performance*. Retrieved from Wikipedia:
https://en.wikipedia.org/wiki/Web_performance

Wikipedia [d]. (2020). *Google Data Centers: Security*. Retrieved from Wikipedia:
https://en.wikipedia.org/wiki/Google_data_centers

Wikipedia [e]. (2020). *Social Impact of YouTube*. Retrieved from Wikipedia:
https://en.wikipedia.org/wiki/Social_impact_of_YouTube

Wilson, M. (2020). *Top 5 Best Bandwidth Monitoring Software and Tools for Network Traffic Usage*. Retrieved from PC & Network Downloads: <https://www.pcworld.com/bandwidth-monitors-best-free-network-tools>