Unreeling Netflix: Understanding and Improving Multi-CDN Movie Delivery

Submitted by: Shrikrishna Bhat

This paper deals with the study of the architecture of Netflix - a Pioneer in providing On-Demand content not only in the US but all over the world. Netflix's architecture is considered to be one of the best architectures in the networking community. This paper unfolds how Netflix designed its architecture to provide fast, reliable, and uninterrupted streaming services. The main goal of the paper was to propose a new strategy, Multiple Content Delivery Network(CDN) based video delivery and measurement-based adaptive CDN detection which increases user's bandwidth.

This paper stands among the rest in describing the atomic-level architecture of Netflix. Multiple authors have collaborated and presented the paper beautifully such that even a layman understands the basics of Netflix. Despite the popularity of Netflix, there have been few studies to understand the working of the streaming industry. Primarily, extensive research is done before proposing this paper and the data gathered by the authors made way for the new strategy which strengthens the core feature of Netflix.

The paper starts with an overview of Netflix's architecture. Authors describe that Netflix has 4 primary components - Netflix Data Centers, Amazon Cloud, CDNs, and players. Data centers handle the top layer things such as data entered through UI/UX interfaces such as registration of new users and recommending movies. Netflix primarily uses Amazon Cloud except for its domain. Netflix primarily uses Amazon Web Services like EC2 and S3, even DRM is done on the amazon cloud. CDNs deliver the content to the users, point to be noted here is that DRM and encoding do not take place in CDNs but they take place in Amazon Cloud, CDN's job is to copy the data and it to the users. The final part is the players, Netflix uses an environment called Silverlight to decode the movies and play them. The authors highlight that Netflix uses Dynamic Streaming Over HTTP (DASH) protocol for streaming.

One key aspect to be noted of the Silverlight is that it checks the user system for which type of video encoding can the user system rendered. Since Silverlight is a runtime application, it checks whether the user system is capable of rendering h.264 video format and if it does, the video is decoded for that format and the video will be played. One of the key aspects is the manifest files which are the XML files, which contain information like CDN list, server names, bitrate, download URL expiration, etc. Manifest files play a key part in seamless streaming, they are transferred over SSL and they cannot be read using Wireshark.

Authors have worked hard on exploring and studying the various CDNs data, they have conducted various experiments to find the overall performance of CDNs, they have collected data of bandwidth variation at different locations over different timestamps and the data is promising. They have also conducted experiments where bandwidth is varied instantaneously. In this case, there is a significant pattern in bandwidth change when 3 systems are considered, and the average difference of variation for 2 hr levels was between 0.01-0.03. One key observation made by the authors is quite impressive, that if the users were tied to a bad CDN choice, their viewing quality was affected even though other Nearby CDNs could give them a greater video quality. This observation made by the authors gave a thought of improvement in the existing Netflix architecture how the bandwidth allocation must be improved and if the highest quality videos like 3D videos were likely to be streamed, bandwidth would be hogged by 3D videos because of their size and some users would likely be unable to watch normal quality videos, and if

multiple users in the same household are watching movies concurrently, there is a requirement of increasing the performance of bandwidth and all the people have the equal share of the bandwidth.

Authors later described the room for improvement in providing equal bandwidth opportunities. Firstly they have selected the strategy called top-level CDN is chosen every time. Practically this is not possible because bandwidth is allocated dynamically. Theoretically, the highest bandwidth allocated by the client is calculated and this will be considered as an average bandwidth and each client will receive the next CDN respective to that highest bandwidth called Upper Bound Average Bandwidth.
The second method is the measurement-based CDN selection. Here the player will instantaneously measure the bandwidth and users will be assigned the best-performing CDN until the streaming is stopped.
The third method is to use multiple CDNs in a parallel fashion. Silverlight player can run the processes in parallel fashion and it can obtain 3 different CDNs and can get larger bandwidth.

Overall this paper gives a detailed explanation of how the Netflix architecture works and how it utilizes multiple CDNs for seamless data streaming. Authors have worked hard in collecting the data and analyzing it to find out that usage of Multi CDNs improves the overall bandwidth by 12%-50% which can pave the way for streaming services to evolve to stream 3D movies.