

Big Data Initiative

Atabak Hafeez

September 27, 2015

1 Big Data Initiative: Why effective caching on online videos is important relative to Big Data?

Caching is very important when it comes to streaming videos online. According to an article (1), in 2007, 50% traffic came from several thousand sites and by 2009, 50% traffic came from 150 sites. Furthermore, by 2013, the 50% of all internet traffic came from 35 or more sites. From this we can see that the traffic of websites that are popular is being aggravated. A lot of this increase in traffic to a smaller number of websites is due to online video streaming.

Before 2012, video streaming was done using Real Time Messaging Protocol (RTMP)-based servers. This requires a streaming server and a near-continuous connection between the server and player. Requiring a streaming server can increase implementation cost, while RTMP-based packets can be blocked by firewalls. In 2012, this was replaced by an (Hypertext Transfer Protocol) HTTP-based servers known as MPEG DASH (Dynamic Adaptive Streaming over HTTP). By using this technology, the servers were able to use HTTP-Caching. This latter capability should both decrease total bandwidth costs associated with delivering the video, since more data can be served from web-based caches rather than the origin server, and improve quality of service, since cached data is generally closer to the viewer and more easily retrievable.

(1) <http://www.streamingmedia.com/Articles/Editorial/Featured-Articles/Online-Video-Jumps-on-the-Big-Data-Bandwagon-91353.aspx>

(2) <http://www.streamingmedia.com/Articles/Editorial/What-Is-.../What-is-MPEG-DASH-79041.aspx>