Centralized Control for CDN-based Live Video Delivery – Public Review

David R. Oran
Cisco Systems
Cambridge, MA, USA
oran@cisco.com

Live video delivery on the Internet is challenging for a variety of reasons. The "obvious" approach of using global IP multicast has been a definitive failure and has been superseded by a combination of adaptive streaming based on HTTP, augmented with robust Content Delivery Networks (CDNs). These handle the vast majority of the tens of terabits of video flowing on the Internet today. While CDNs excel at optimizing the high-volume delivery of static content (e.g. video-on-demand, web content), live video is more difficult due to the update rate, the requirement for low latency, and the limited utility of caching to reduce load. The workload mix is also quite varied, from "mega-events" with millions of viewers, through a long tail of limited-interest channels that may or may not have topological locality in viewership.

Current CDNs use load balancing through DNS to manage the load on CDN servers and to choose servers close to the individual consumers. However, using this highly-distributed method runs into problems with responsiveness to load changes and network congestion, the most severe being the need for long DNS TTLs to keep update overhead low. Given the immense amounts of live video traffic and the commercial importance of video on the internet, having an efficient and adaptive scheme for handling live video is of both immediate and long-term interest.

This paper proposes that instead of a purely distributed DNS-based scheme, CDNs would benefit significantly by adopting a hybrid-control approach. The system they have designed, VDN, retains distributed control at the edge in order to rapidly respond to changing network conditions, while introducing a centralized controller that sees all the global traffic and can perform load distribution optimizations across both the clusters of CDN servers in the delivery tree and across the live video workload mix.

Although centralized integer programming optimizers have been employed in many contexts, the authors demonstrate a design that has reasonable computational overhead while keeping up with changing load and global network conditions. This is achieved by retaining the distributed load control at the edge so that the central control need not respond at very short timescales, nor deal with every stream request from every client. The combination represents a meaningful advance over the state of the art.

Since the limits to responsiveness and global optimization of current CDNs is due in part to lengthy DNS TTLs, it is fair to ask if a simpler alternative to VDN would be to just shorten the TTLs. This would considerably increase the DNS query overhead, but avoid the messaging and computational overhead of VDN in coordinating the edge clusters with the central controller.

The authors consider this, and compare the overhead and performance of a baseline CDN approach, the use of reduced DNS TTLs, and the VDN scheme. The evaluation results show substantial advantages in both performance and cost reduction of VDN over either "pure" CDN approach. From this, they successfully demonstrate that hybrid control can achieve superior live adaptive video delivery over a wide range of workloads.