

Consolidated Review of

Analyzing the Potential Benefits of CDN Augmentation Strategies for Internet Video Workloads

1. Strengths

This paper taps a useful dataset for Internet video workloads and analyzes how Telco-CDN federation and CDN-P2P would perform with those workloads.

It is a wonderful data set for video downloads, even if the results are typically unsurprising. Great dataset.

Interesting observations.

Well written paper. Straightforward writing

The Telco-CDN federation analysis is convincing. It outlines a systematic approach and in case, clearly explains their hypothesis, implications of observed results, and how they perform the evaluation. Nice to see it manages to debunk some conventional wisdom. Simple models seem sufficient for their purpose.

2. Weaknesses

The CDN-P2P analysis felt weak, in part because I never felt I understood what the P2P model was for this case.

The definition of fairness (which ISP gain most) in Section 4.4 is misleading.

This paper does not show any results from experiments or simulations to quantify the benefits of using two video delivery strategies. This paper mostly focuses on analyzing the benefits of using two strategies, and sheds little light on their performance cost.

3. Comments

I really enjoyed reading the paper - it flows very well, and it's very easy to follow the thought process. I appreciate the systematic presentation of their methodology, empirical findings, insights gained (or implications), which then led to the evaluation of the CDN augmentation strategies. The benefits for federated Telco-CDN may have been over-estimated. The biggest concern here is that the paper completely ignores the ISPs' need to satisfy their respective own customers. While it's good to pool resources together for more efficient allocation, how does one decide whose customers to drop when available resources cannot meet the demand?

This paper focuses on data-sets from two content providers among the top 500 sites in overall popularity rankings. This paper could briefly compare these providers with most popular video sites, e.g., YouTube and Netflix, and discuss whether the observed video access patterns likely hold for other sites as well. The paper observes a heavy tailed Zipf distribution on the overall popularity of video objects for both VOD and live sessions. A graph will be helpful here. This paper analyzes the potential benefits of two video delivery strategies.

The critical problem for me was a sense that neither Telco-CDN nor CDN-P2P was well explained and while I could intuit Telco-CDN, I could not intuit CDN-P2P.

Section 3: Surprised that there were only 14M views for 4M users in 2 months. Is this expected? I assume you filtered out users who didn't have any views?

Section 4.1.1: I was surprised at how few Comcast clients you saw and how many Verizon clients you saw. Is this expected? Is it a reflection of ISP presence or is it specific to conviva (or something else with your dataset)?

Fig 7a: I was a bit surprised that country-wide federation didn't provide more of a benefit.

The availability we saw on Fig. 7 and 8 are from the view of the federation or set of ISPs at a whole. What about availability at individual ISP level? This leads to the question of how fairness is studied in Section 4.4. The ISP gain shown in Fig. 9 is the ratio, $\text{TotalServedbyOthers}/\text{Capacity}$, which completely ignored the portion of capacity of this ISP that is used to serve other customers in the federation. Unless there is a totally different charging model, other wise, an ISP's priority will be to serve its own customers first. So, why not compare the availability metric for a specific ISP before and after federation with respect to its own customers? I'm also a little skeptical about using P2P to bootstrap serving the content. There is no study of 'latency' involved. What is the condition (e.g., minimum number of users interested in the same content) for P2P bootstrapping to work?

Fig 11: It seems surprising that so few sessions are of the whole or nearly whole event. Do you know how many users watched the entire tail of the event from when they joined?

Section 5.2: I'm concerned at how simple the P2P model is, abstracting away any sense of capacity. Is the model "if there is a user in a region (say) at chunk N of a video, then up to 99 other users in that region can be assumed to get the chunk for free?"

This paper does not present the performance cost of these two strategies. This paper could be much stronger with results from experiments or simulations. This paper assumes the "peering" relationship between ISPs in sharing server capacity.

In practice some ISPs might have much higher server capacities than others. Thus, would a "provider-customer" relationship be more appropriate for ISPs with different capacities? This paper explores the benefits of two video delivery strategies separately. Is that possible for ISPs to explore a combination of both strategies? If not, why? If yes, what are the potential benefits and challenges of employing both strategies at the same time?

"VoD" in "Several pure P2P VoD system" (page 2) should be changed to "VOD" for consistency.

4. Summary from PC Discussion

The PC agreed that the data set was wonderful and the analysis was a little weak, but still insightful.

5. Authors' Response

We thank the reviewers for their feedback and comments. At a high-level, the main concern from the reviewers is that our current analysis does not capture cost or policy concerns from the

perspectives of the ISPs, CDNs, and content providers. We acknowledge that this is a valid concern. Our goal in this paper was to provide a data-driven estimate of the potential benefits of the proposed augmentation approaches such as hybrid P2P-CDN or federation given the recent interest in deploying these in the wild. Incorporating more fine-grained policy or cost models is an interesting direction of future work.

In response to specific reviewer concerns, we made the following changes from the submission version:

1. We added a discussion on the performance costs caused by both hybrid P2P-CDN and federation.
2. We added more details about the hybrid P2P-CDN model and the telco-CDN federation.
3. We stressed that the hybrid-P2P model that we are using is simple since our goal is just to estimate the potential benefits of the approach.
4. We clarified various technical details – e.g., users leaving early, few views compared to number of people.
5. We fixed various typos and issues with figure fonts.