Student Name: Taylor Tang
SID: 1323782

Introduction

In the past few years, the video streaming business has grown exponentially due to COVID-19 restrictions, more people choose to stay home and watch tv shows on Netflix, YouTube, etc. This boom happened so quickly that Netflix had to limit its video streaming quality from 4k to 1080p in some areas, to ensure it can cope with the network load from the burst of demand. This is the short side of traditional video-on-demand streaming services where the main media server (or servers) is responsible for sending data to each end-user. When there are more end-users than exceed the capacity, the server is under heavy load when it hits the limitation of the server's network bandwidth. Compared to traditional methods, peer-to-peer video streaming has several advantages including low set-up cost and low difficulties when scaling. Each end-user (a peer) not only downloads and consumes the video content but also uploads it to other users (peers). In this way, the main server only needs to send the video data to some peers, while most of the data transferring to other peers is done between peers and peers (hence, so-called P2P video streaming) (Shah, et al., 2018). A P2P system is good at up-scaling since with more peers, the better overall service quality it gets. However, a P2P video streaming service needs to overcome a few difficulties. First, it is difficult to maintain a stable performance due to the uncertainty of a peer's performance, such as the peer's network condition (upload speed) and stability (peer come and go). Second, P2P streaming is usually delayed for seconds, even minutes. This is caused by the video content data may be transferred a few times before it reaches a peer's machine (Karayer & Sayit, 2015). This survey will discuss the three P2P network topologies, namely a push-based system, pull-based system, and hybrid system, and their advantages and disadvantages. Then, this survey will touch on the optimized peer selection problem by studying approaches proposed by a few reports and compare their pros and cons. Last, this survey will briefly discuss the future directions of P2P video streaming.

Related works

Push-based p2p network (tree topology)

A push-based P2P network has a tree topology, where the peers establish a clear and explicit relationship with each other. Source (main media server) located at the root and sending data to its child nodes (peers). Within the relationship of two peers, one peer has to be the parent while the other has to be a child. Once the parent peer receives data (from its parent), it will forward data to its child's peers. The data transmission policy is simple, once a peer receives data from the above structure, sending data to the below structure. So, the delay is low since there is not much overhead in terms of requesting or synchronizing. However, the main drawback is that if one peer leaves the system, all its child peers (and their child peers) will lose the connection (Lo, Lin, Chen, & Yu, 2012).

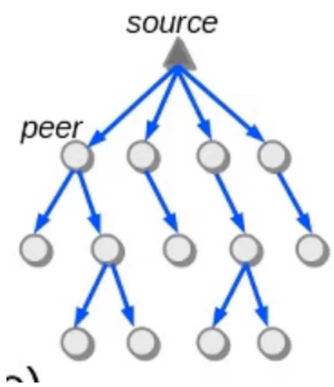

Figure 2.1 Push-based

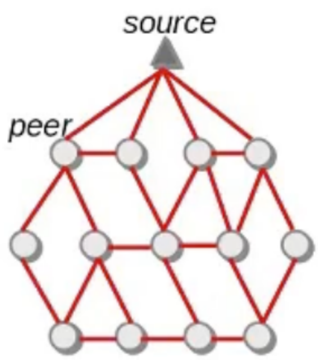

Figure 2.2 Pull-based

Pull-based p2p network (mesh topology)

A pull-based p2p network has a mesh topology embedded. A few peers establish a direct connection to the source, while other peers form a mesh-like system among themselves. Therefore, there is no clear parent-child relationship but rather a partnership relationship. If one peer decides to leave the system, other peers can seek new partners among other peers. In other words, there are many paths for one peer to another peer. Data streams are divided into data chunks. A peer sends a request to its connected peers for the data chunk it needed, then receives the data chunk from several peers. However, this relationship requests peers to constantly exchange information on data chunk they need and data chunk they own, which causes overhead and delay.

Hybrid p2p network (super node selection)

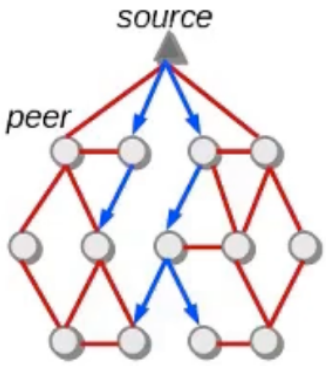

Figure 2.3 Hybrid peer-to-peer network

A hybrid system inherited both pull-based & push-based systems' advantages. In the hybrid system, the server establishes a direct data connection with a few selected peers (super-peer). Those peers usually perform better than others in terms of upload bandwidth, stability, etc. Then, the rest of the peers along with the selected peers forms a mesh-based system. In this case, the system now has a low delay & overhead and increased robustness. However, those selected peers are the backbone of the system, if they decide to leave, it will cause the whole system unstable.

	Delay	**Overhead**	**Robustness**
Push-based	Low	Low	Low
Pull-based	High	High	High
Hybrid	Low	Low	Medium

Figure 2.4 Three P2P network topology ranking

Best path & best node selection

The main advantage of a p2p streaming service is that it can significantly reduce the load of the main server, shifting the load of data transferring from a one-to-many system to a many-to-many system. Each user is not only the consumer but also the provider to other peers. Also, the system is scalable as the more peers join the more paths, they can get the data from the main server (Wu, et al., 2013).

Therefore, consider a p2p video streaming service with multiple kinds of users. How to wisely choose a path and best node becomes the foundation of p2p streaming quality. The focus of this part is on the methodology of each approach to selecting peers, hence we will not discuss the detailed algorithms and processes.

Super node selection algorithm combining reputation and capability model in P2P streaming media network.

Based on the previous discussion, a hybrid pull-push-based p2p structure depends on a few backbone nodes. Those nodes usually have higher bandwidth capacity, relatively close distance, and a reputation of reliability in providing service continuously. Rongfei, M.A. (2019) proposed a super node selection method to promote nodes with the best performance as super nodes.

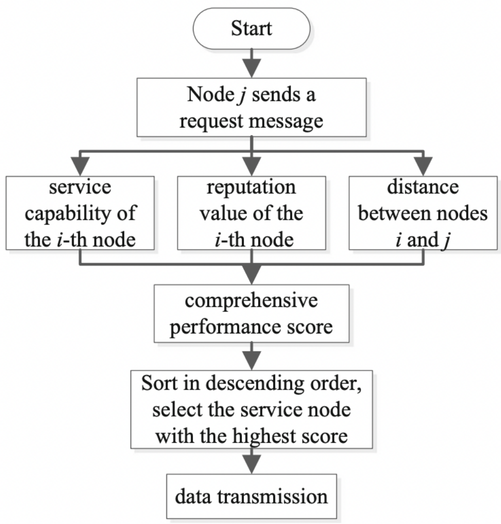

Figure 2.5 Rongfei, M.A.'s (2019) super node selection model

The selection of super nodes starts from a series of assessments of capacity, reputation & distance. As shown in figure 2.5, when a node j sends data request messages to all its partners peers, it assesses the above key factors of its partners. For capacity assessment, this approach considered CPU, storage, memory, and bandwidth then comes with a capacity coefficient number. For distance assessment, this approach uses ICANN (The Internet Corporation for Assigned Names and Numbers) to check the registered IP address and its approximate physical location. For reputation, this approach provided a detailed and complex calculation. There are two scenarios: 1) In a direct interaction situation, where nodes and nodes have interacted with each other for a period, this approach uses the base case calculation of "direct reputation", which is a calculated coefficient number between one node, and another based on the success rate of transmitting files. 2) In an indirect interaction situation, where certain nodes haven't established connections with others or had limited connections, this approach uses a third-party node (common node) to obtain the reputation factor of the success rate of transmitting between them.

After obtaining the capacity, reputation & distance factors, this approach applies weight to each of the factors (a, b & c in which a + b + c = 1). These factors are set based on close observation of a detailed situation. For example, if this system is mostly used by a local group, it will decrease the distance factor weight. Another example, if this system is used to stream ultra-high-definition video, the capacity factor's weight will be increased.

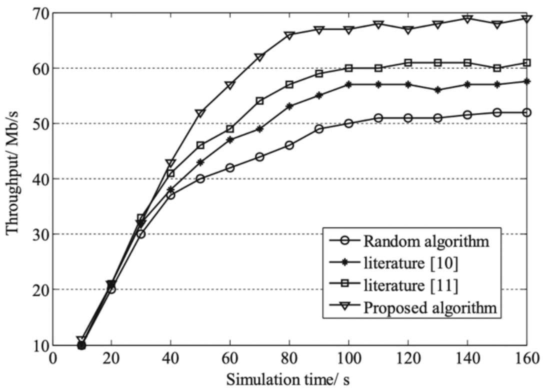

Figure 2.6 Rongfei, M.A.'s (2019) super node selection algorithm performance comparison

In the lab testing, the proposed approach achieves better quality in terms of data transmission rate. It can be observed that the proposed approach provides approximate 40% performance gain comparing to the random node selection algorithm, as per figure 2.6.

Neighbour Selection Method based on Sending Capacity for P2P Live Streaming with Layer Coding

Another approach to selecting the best nodes (super nodes) is sorting. In this article, the authors (Endo, Takayama, Sakata, & Shigeno, 2012) argue that if one low-capacity node is with neighbours of one high-capacity node, it can receive sufficient data from the stream. However, if one high-capacity node is among other low-capacity nodes, it cannot receive enough data to fully utilizes it selves high-capacity to forward data to others. Therefore, the best scenario happens when all nodes are patterners with similar capacity nodes, with higher capacity nodes sitting closer to the media source, as per figure 2.7.

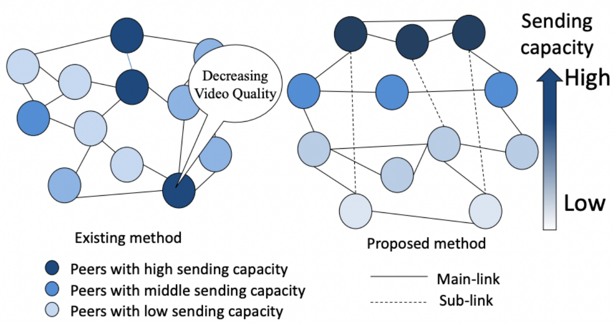

Figure 2.7 Endo, Takayama, Sakata, & Shigeno (2012) Node selection based on sending capacity

To rank each node's sending capacity and sort them in order, each peer needs to perform a self-assessment in terms of capacity (upload-bandwidth & download-bandwidth). A higher capacity node will be placed closer (topology speaking) to the media server, a lower capacity node will be placed farther from. In this way, each node can always receive data from high-capacity nodes and forward data to lower-capacity nodes.

To decide how many peers a node can be partnered with, this approach pre-defines a constant number of C for the total number of communication links allowed for one node. This approach divides the communication links into two categories: main-links and sub-links. Main-links represent neighbours with similar or higher bandwidth for receiving. While the sub-links represent neighbours with lower bandwidth for sending. Each node has the right to terminate the connection with another node.

Once a new node enters the system, it first receives the constant number C and pre-recommended neighbour nodes from the server based on its capacity. Then, it periodically updated its main-links and sub-links by replacing the lowest ranking neighbour with a new neighbour from the main server. After a few rounds of replacing and updating, all nodes will sort themselves in an up-down bandwidth order. Hence, the data flow bottleneck is limited to achieving better overall data transferring.

Efficient neighbour selection through connection switching for P2P live streaming

Kim, Kim & Lee (2018) took another approach in their study of efficient neighbour peers selection. The main argument of this report is that in the conventional peer selection progress, each peer may be connected to a certain number of peers, which will help to build a stable structure and minimize congestion. However, by following this rule, a new user will most likely have a less good experience than the previous user. Therefore, the author proposed a method of adding a new user to a closer topological position from the main server without affecting the existing service performance by fully utilizing the video buffer space. In this way, the new peer can have a better experience.

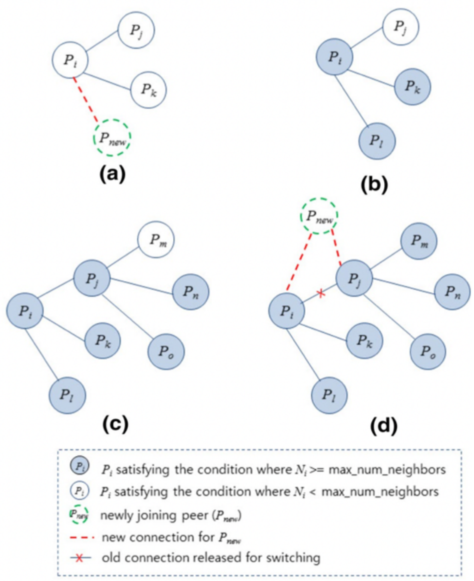

Figure 2.8 Kim, Kim & Lee (2018) New user joining process

As per figure 2.8 (d), when peer P(j) has enough buffered video space, it can keep playing video without receiving new data for some time. Under this situation, it is possible to slide new peer P(new) in between P(i) and P(j). In this way, P(new) would have less playback delay than it in the original placing method in figure 2.8 (c).

Best path & best node selection comparison

In Rongfei, M.A. (2012) proposed approach, peer selection is based on a series of assessments of capacity, reputation, and distance. By weighting these factors based on a certain scenario, an optimized peer is promoted to be the super peer. It is an all-around assessing method, however, the cost of this method is high and the accuracy is not guaranteed. First, it assesses each peer's capacity and reputation, given the size of the P2P network, the cost of this assessment might be high in terms of time spent and information being exchanged. Also, the distance is calculated approximately, instead of using the traceroute function to detect the delay from one peer to another.

In Endo, Takayama, Sakata, & Shigeno (2012) proposed approach, the peer selection is purely based on the bandwidth capacity. Higher bandwidth capacity peers get promoted, while the lower bandwidth capacity peers get demoted. Also, each peer exchange and update its super peer periodically, so the cost of this approach is relatively low. However, the drawback of this approach is that other factors are not considered, such as delay, peer reliabilities, etc.

In Kim, Kim & Lee (2018) proposed approach, they took a different angle on peer selection. Rather than assessing a peer's performance and then choosing its patterner peers, the proposed approach is to fully utilize the buffered video to place the new peer in between two peers with a steady link. In this way, the playback delay of the new user is reduced. However, this approach only performs well under a few assumptions such as enough buffered video and stable connection of peers.

Ranking	**Covered area**			**Evaluation**		
Proposed approaches	Capacity	Reliability	Delay	Cost	Stability	Performance
Rongfei, M.A. (2012)	Yes	Yes	Yes	High	High	High
Endo, Takayama, Sakata, & Shigeno (2012)	Yes	No	No	Low	Low	Medium
Kim, Kim & Lee (2018)	No	No	No	Low	High	Low

Figure 2.9 Three proposed approach on optimized peer selection ranking

Conclusions and Future Directions

P2P video streaming has certain advantages over traditional video streaming, such as low setup cost and better scalability. However, P2P also comes with drawbacks such as low stability and high playback delay. To address these drawbacks, this survey researched a few studies regarding P2P video streaming topology and peer selection. There are three common topologies namely push-based, pull-based, and hybrid. A hybrid P2P topology tends to be a better choice out of these three as it inherited the benefits of both push and pull-based. On the topic of optimized peer selection, this survey included three studies that optimize the peer selection approach based on different factors, namely capacity, reliability, and delay. Overall, Rongfei, M.A.'s approach to selecting the best peer based on capacity, reliability, and distance is the most effective and most costly.

Nowadays, although upload bandwidth is no longer low compared to the download speed from a decade ago, it is still not enough when general consumers to stream 4k UHD video content. As a normal consumer, the importance of their role in a p2p video streaming service is still low. As mentioned in the above discussion about the p2p network topology, both push-based and hybrid systems rely on server super peers. Those peers usually have high bandwidth capacity and stability. Could computing will fit in this position perfectly.

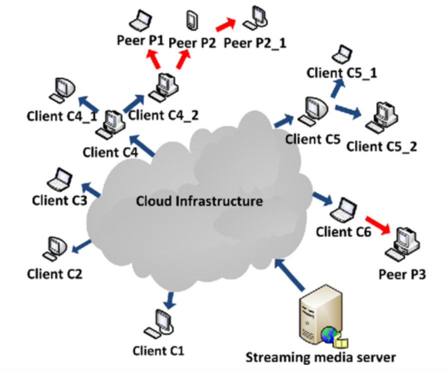

Figure 2.9 P2P streaming with the assistance from cloud computing

We could use cloud computing as the assistant to power up P2P streaming one step more. Due to the dynamic feature of cloud computing, we could utilize it as the backup super-peer when needed. In this way, the huge upload bandwidth and high stability will even the drawbacks of P2P video streaming service. If not needed, we could scale down the cloud computing assistant to minimize the cost (Alghazawy & Satoshi, 2016).

References

Alghazawy, B. A., & Satoshi, F. (2016). Low Cost Cloud-Assisted Peer to Peer Live Streaming. *KSII Transactions on internet and information systems : TIIS. 2016 10(4):1732-1750*.

Endo, R., Takayama, K., Sakata, Y., & Shigeno, H. (2012, October 18). *Neighbor Selection Method Based on Sending Capacity for P2P Live Streaming with Layer Coding.* Retrieved from ieeexplore: https://ieeexplore.ieee.org/document/6332007/authors#authors

Karayer, E., & Sayit, M. (2015, September 5). *A path selection approach with genetic algorithm for P2P video streaming systems.* Retrieved from Springer Link: https://link.springer.com/article/10.1007/s11042-015-2912-y

Kim, E., Kim, J., & Lee, C. (2018, January 29). *Efficient neighbor selection through connection switching for P2P live streaming.* Retrieved from SpringerLink: https://link.springer.com/article/10.1007/s12652-018-0691-9

Lo, C.-W., Lin, C.-W., Chen, Y.-C., & Yu, J.-Y. (2012, June 08). *Contribution-Guided Peer Selection for Reliable Peer-to-Peer Video Streaming Over Mesh Networks.* Retrieved from IEEE Xplore: https://ieeexplore.ieee.org/document/6213524?arnumber=6213524

Rongfei, M. (2019, April 13). *Super node selection algorithm combining reputation and capability model in P2P streaming media network.* Retrieved from Springer Link: https://link.springer.com/article/10.1007/s00779-019-01219-y#citeas

Shah, S. S., Said, N., Nayab, A., Khan, W., Shinwari, Z. A., Jawwad, M., & Minallah, N. (2018, December 17). *Performance Comparison of Chunk and Peer Scheduling Algorithms of Peer-to-Peer Streaming Systems.* Retrieved from IEEE Xplore: https://ieeexplore.ieee.org/document/8617019/authors#authors

Wu, H., Liu, J., Jiang, H., Sun, Y., Li, J., & Li, Z. (2013, January 11). *Bandwidth-aware peer selection for P2P live streaming systems under flash crowds.* Retrieved from IEEE Xplore: https://ieeexplore.ieee.org/document/6407778?arnumber=6407778