LAN-Awareness: Improved P2P Live Streaming

Zhijie Shen
School of Computing
National University of Singapore
Singapore 117417
z-shen@comp.nus.edu.sg

Roger Zimmermann
School of Computing
National University of Singapore
Singapore 117417
rogerz@comp.nus.edu.sg

ABSTRACT

The popularity of P2P streaming systems has rapidly created extensive, far-reaching Internet traffic. Recent studies have demonstrated that localizing cross-ISP (Internet service provider) traffic can mitigate this challenge. Another trend shows that households own an increasing number of devices, which are sharing a LAN of 2 or more peers. To this date, however, no study has investigated the potential of localizing traffic within LANs. In our presented work, we propose the concept of *LAN-awareness* and introduce its threefold benefits: 1) reducing Internet streaming traffic, 2) lowering stream server workload, and 3) improving streaming quality. First we conduct a large-scale measurement on PPLive, confirming that a considerable number of peers (up to 21%) are connected to the LANs having 2 or more peers. Recognizing the opportunity of localizing traffic within LANs, we discuss the principles to construct a LAN-aware overlay and propose a heuristic. The results of our trace-driven simulations confirm the benefits outlined above.

Categories and Subject Descriptors

C.2.4 [**Computer-Communication Networks**]: Distributed Systems; C.2.5 [**Computer-Communication Networks**]: Local and Wide-Area Networks

General Terms

Measurement, Design, Performance

Keywords

P2P, streaming, LAN-aware, traffic locality

1. INTRODUCTION

A number of P2P live streaming systems, such as PPLive, PPStream and UUSee, have attracted large user communities in recent years. Such marvelous popularity has already resulted in heavy workload for both Internet service providers (ISPs) and content providers. Moreover, some forecasts predict that by 2014 Internet video will account for 57% of all consumer IP traffic [2]. Meanwhile, households possess increasingly more devices on which they consume entertainment content: desktops, laptops, tablets (*e.g.*, Apple iPad), *etc.* These devices are connected through a wired or wireless router, form a home area network (HAN) and reach the Internet through a shared link. This link tends to be the bottleneck when multiple devices are accessing the Internet simultaneously.

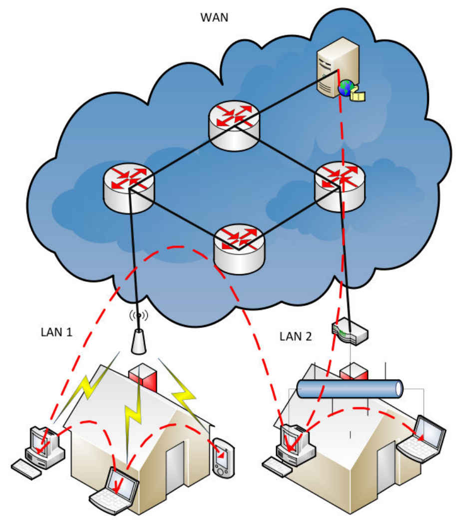

Figure 1: Exhibition of a LAN-aware overlay.

Recently we conducted a comprehensive measurement study on PPLive, and one of the interesting observations was that a significant number of local area networks (LANs) accommodated multiple peers watching the same live channel. Surprisingly, up to 21% peers belonged to such LANs where they had local neighbors. This phenomenon is likely to be more prominent in the future, since the population of Internet subscribers is expected to dramatically increase [8]. Video streaming systems are bandwidth-intensive applications. Multiple peers further raise the bandwidth consumption on the link connected to the Internet (last-hop link), affecting other network-based applications or even disturbing the streaming quality of each other. Fortunately, the use of P2P architectures opens the opportunity for solving this problem. Here we define a new concept, *LAN-awareness*, indicating that *the overlay construction is optimized to exploit local peer resources within a LAN*. Fig. 1 exhibits a sample LAN-aware overlay, where only one peer in each LAN (or HAN) receives a stream from outside and then forwards it to its local neighbors. The benefits of LAN-awareness are threefold.

- Video delivery has become one of the most problematic traf-

NOSSDAV'11, June 1–3, 2011, Vancouver, British Columbia, Canada.

fic generators, heavily stressing ISPs. Prior studies [7, 9, 10, 12] have proposed to restrict traffic within autonomous systems (ASes) to lower ISPs' operating costs. LAN-awareness restricts the traffic entering the wide area network (WAN), thus reducing both intra-AS and inter-AS traffic.

- Recent measurement studies show that the bandwidth contribution of peers is limited [3, 5]. A stream server has to compensate for the upload deficit. In contrast, the intra-LAN bandwidth capacity is usually abundant. Hence if some peers can satisfy their video demand within the LAN, content providers need to purchase less bandwidth.

- For most Internet subscribers, their download bandwidth is still not enough to allow several concurrent streaming sessions while the trend in video rates is growing higher (more videos encoded at 700 Kbps and 2 Mbps are available on PPLive). LAN-awareness avoids the last-hop bandwidth competition, thus preventing a possible streaming quality decline. Even if the bandwidth is not the bottleneck, LAN-awareness is still helpful to leave more spare bandwidth for other applications.

Current P2P live streaming systems are not LAN-aware and often use random partner selection. This seldom results in peers connecting to neighbors in the same LAN. To exploit the potential benefits from LAN-awareness, we investigate the principles of designing a LAN-aware overlay, propose a lightweight solution and evaluate it with trace-driven simulations. Overall our work provides the following contributions.

- To the best of our knowledge, this is the first study that proposes the concept of LAN-awareness and discusses its benefits of reducing Internet streaming traffic, lowering stream server workload and improving streaming quality.

- We have performed large-scale measurements with PPLive. Our analysis of the collected data demonstrates that there exist a considerable number of LANs containing multiple peers when a channel is popular. This opens a significant opportunity to exploit LAN-awareness.

- We introduce the principles of designing a LAN-aware overlay and propose a lightweight solution. The results of our trace-driven simulations show that the LAN-aware solution can achieve the three benefits outlined earlier.

The rest of the paper are organized as follows. Section 2 surveys the related work. We explain our measurements and analyze the collected traces in Section 3. The LAN-aware solution is introduced in Section 4 and evaluated in Section 5. Concluding remarks are presented in Section 6.

2. RELATED WORK

There exist several recent studies that focus on a closely related topic, *i.e.*, AS-aware P2P streaming. One measurement study [6] has shown that PPLive naturally exhibits a certain traffic locality caused by the skewed ISP-size distribution. With some proactive solutions, the locality can be further improved. Picconi *et al.* [9] propose an ISP-friendly scheduling strategy where each peer requests most data from its topologically-close partners and only resorts to distant peers when data is not available locally. Similarly, a two-tier scheduling scheme is proposed by Magharei *et al.* [7]. Tomozei *et al.* [12] present an intriguing, fully decentralized algorithm based on flow control. The merit of decentralization is unfortunately offset by a slow adaptation to any load dynamics. Our previous work [10] proposes a self-adapting, QoS-protected peer selection algorithm.

There are only two other studies that are similar to our work. Tan *et al.* [11] propose to deploy caching agents at the access points (APs) of wireless networks to avoid duplicate traffic traversing the APs. Meanwhile, Lai *et al.* [4] also suggest to add agents at the APs, in order to leverage the broadcast protocol of WLAN to optimize the bandwidth usage of the APs. However, these studies have not provided detailed insights of whether multiple users are connected to the same AP, so that the potential improvement for real systems is not clear. Moreover, the infrastructure-based solutions will meet the same obstacle as IP multicast did when they are put into practice. Here we conducted a comprehensive measurement study on a real system, confirming the existence of multiple peers within the same LAN. Furthermore, we propose an LAN-aware solution that restricts the modification efforts to just the application layer. Thus, it can be easily deployed.

3. MEASUREMENT OF REAL-WORLD SYSTEM

The potential benefits of LAN-awareness naturally depends on the number of peers inside each LAN, *i.e.*, the peer density of a LAN. If there were just one peer in each LAN, it would be useless to design a LAN-aware solution. Therefore, as a first step, we require a better understanding of how peers are distributed among LANs in real-world systems. However, there is no such information available from prior related work [4,11] or recent measurement studies [3,5]. Thus, we conducted our own measurements and analysis.

3.1 Overview of PPLive Partnership Protocol

Our measurements target PPLive. We start by briefly introducing the partnership protocol of the live streaming functionality. Once a PPLive client (or peer) joins the system, it contacts the bootstrap server to retrieve the list of active channels. After the client chooses a program to watch, it informs the tracker server of its participation in the corresponding channel and obtains a list of active clients in this channel as the initial partners. Then, the client expands its partner list by querying the partners' neighborhoods. During the session, the client keeps updating the partnership information at regular intervals. Notably, PPLive constructs the overlay on a per-channel basis, containing the peers that are watching the channel.

Currently PPLive adopts UDP as the transport protocol by default. A client randomly chooses a unique port, through which all the IP packets related to PPLive are transmitted. Therefore, a UDP socket $\langle$IP, port$\rangle$ can exclusively represent a client. In our measurements, we observed a considerable number of sockets that share some common IPs. Since their ports were different, these sockets represented different clients. This phenomenon occurs because of network address translation (NAT), which is usually deployed at the gateway between WAN and LAN and translates private IPs of multiple clients within the same LAN into the same public IP.

3.2 Measurement Methodology

In our measurements, we are interested in the topology-related information. To acquire this information, we need to know all the concurrent peers in a channel, *i.e.*, obtain an *active peer population snapshot*. Thus, we implemented a PPLive crawler in Python, inspired by the method of Vu *et al.* [13]. The crawler imitates the

Name	Type	Event	Rate (Kbps)	Population (daily basis)
Hunan Satellite TV	Comprehensive	N/A	400	481 – 17,949
SiTV-Sports	Sports	3^{rd} match of NBA Final	400	47 – 11,160
CCTV-5	Sports	4^{th} match of NBA Final & 1^{st} match of World Cup	700	789 – 183,109

Table 1: Summary of the studied channels. The one-day traces of these three channels were respectively collected on May 19^{th}, June 9^{th} and June 11^{th}, 2010. The time zone is GTM+8.

behavior of a PPLive client to probe peers in the channel from both the tracker server and the clients. Probing peers from the tracker server is straightforward. The crawler repeatedly requests peers from the tracker server until it can no longer get any new peers. In contrast, probing peers from the clients is complex. Whenever the crawler acquires a new peer, it will query this peer for the peers in its local view. This operation is repeated until no new peer is discovered by the crawler. When both operations are complete, we have obtained the snapshot. Formally, let $\mathcal{S}$ denote the set of the captured peers and $\mathcal{T}$ denote the set of the peers returned by the tracker server. $Neighbor(p)$ represents the peers in the local view of peer p and $Neighbor(\mathcal{S}) = \bigcup_{p \in \mathcal{S}} Neighbor(p)$. Then, $\mathcal{S}$ can be defined as

$$\mathcal{S} = \{p | p \in \mathcal{T} \vee p \in Neighbor(\mathcal{S})\}. \quad (1)$$

Actually, the crawler cannot probe all the concurrent peers at the same time due to the recursive nature of Eq. 1. $\mathcal{S}$ is an approximation of the snapshot and converges to it as the execution time decreases. Thus, we parallelized the independent probes so that the crawler could complete its work in just a few minutes. Additionally, to enable the communication with the tracker server and the clients, we have collected a considerable amount of PPLive traffic with Wireshark to recognize the packet format of the partnership protocol.

To avoid that the captured peers might be biased to the crawler's location, we deployed the copies on 44 PlanetLab nodes spread over 15 countries and regions and executed them simultaneously at the regular interval of 15 minutes. Our selection of the PlanetLab node number and execution interval is moderate to restrict the overhead brought to PPLive. The snapshot collection lasted from May 17^{th} to June 13^{th}, 2010 and we obtained more than 70 GB of raw log data. Afterwards, we merged the concurrent snapshots from different PlanetLab nodes to construct a complete snapshot. Furthermore, we utilized the Team Cymru service (whois.cymru.com) to retrieve the AS ID and the country information of each peer. While the data can be leveraged to infer many overlay properties such as churn behavior and AS-level peer distribution, in this paper, we concentrate on the peer distribution among LANs and analyze the data of three representative PPLive channels, which are summarized in Table 1.

3.3 Trace Analysis

From the data, we analyze the peer distribution among LANs. Here we assume that the recorded peers sharing the same IP are connected to the same LAN. This assumption may cause underestimation to some extent, because peers from some large LANs, such as a campus network, connect into the WAN through multiple gateways so that the peers are recorded with different IPs. However, even with this underestimation, we still observed abundant LANs hosting multiple peers.

Figs. 2(a) – (c) show the one-day evolutions of the channel population, the percentage of the IPs supplying at least 2 peers and the percentage of the peers within these IPs. While the evolution of Hunan Satellite TV (HunanTV) follows a diurnal pattern, those of the other two sports channels are event-related (peaks occur in the morning of the NBA Final and at the night of the World Cup). At peak times, 2% to 7% of IPs supply at least 2 peers, and 5% to 21% of the total number of peers are associated with these IPs. Specifically, 21% of the population of CCTV-5 represents about 39,000 peers, the number of which is much more than the population of a whole channel at regular times. As we know, most ISPs charge content providers with the 95^{th} percentile rule, implying that the monetary cost is determined by the peak bandwidth usage. Thus, both the percentage and the number illustrate a significant prospect for lowering the monetary cost for content providers if traffic is localized within LANs.

One important observation from the evolutions is that the incidence rate of multiple peers within a LAN is correlated with the channel population. Figs. 2(d) and 2(e) present a clearer picture. Let $\mathcal{X}(i)$ denote the number of IPs who supply at least i peers and $\mathcal{N}(i)$ denote the total number of peers within these IPs. For simplicity, the two terminologies are respectively reduced to $\mathcal{X}$ and $\mathcal{N}$ when $i = 1$. According to the traces, the correlations approximate

$$\mathcal{X}(2) \propto \mathcal{N} \log \mathcal{N}, \quad (2)$$

$$\mathcal{N}(2) \propto \mathcal{N} \log \mathcal{N}. \quad (3)$$

Eqs. 2 and 3 imply that the incidence rate of multiple peers within the same LAN increases $\log \mathcal{N}$ times faster than the channel population. Originally, Fig. 2(d) infers $\mathcal{X}(2) \propto \mathcal{X} \log \mathcal{N}$. However, we derive Eq. 2 as we find a linear dependence between $\mathcal{X}$ and $\mathcal{N}$ as well as between $\mathcal{X}(2)$ and $\mathcal{N}(2)$, shown in Figs. 2(h) and 2(i). Here we conducted curve fitting and obtained

$$\mathcal{N} = 1.16\mathcal{X} - 543, \quad (4)$$

$$\mathcal{N}(2) = 3.4\mathcal{X}(2) - 112. \quad (5)$$

Then, let $\Delta = \mathcal{N} - \mathcal{X} = \mathcal{N}(2) - \mathcal{X}(2)$, which indicates the number of peers that can satisfy their streaming demand within LANs if one peer in each LAN imports a stream from outside. As $\mathcal{N}$ is very large, Δ approximates $0.16\mathcal{X}$ and $0.14\mathcal{N}$. Therefore, if a LAN-aware solution is provided, about 14% of the traffic can be eliminated from the WAN.

Next, we would like to understand the peer distribution among the LANs accommodating multiple peers. Fig. 2(g) shows that the number of peers associated with one IP versus the IP's rank follows a *Zipf-like* distribution. Only a small fraction of IPs host tens of or even more peers. These IPs may be from some campuses or companies, where there are numerous users. In contrast, the remaining fraction of IPs come from households, where the members are few. Furthermore, an additional question is how crowded a LAN will be. In our traces, the most crowded IP supplied 214 peers. Fig. 2(f) shows the correlation between the number of peers in the top ranking IPs and the channel population, which can be approximately represented by

$$\mathcal{N}(i_{top}) \propto \mathcal{N} \quad (6)$$

where i_{top} satisfies $\mathcal{X}(i_{top}) > 0$ and $\mathcal{X}(i_{top} + 1) = 0$.

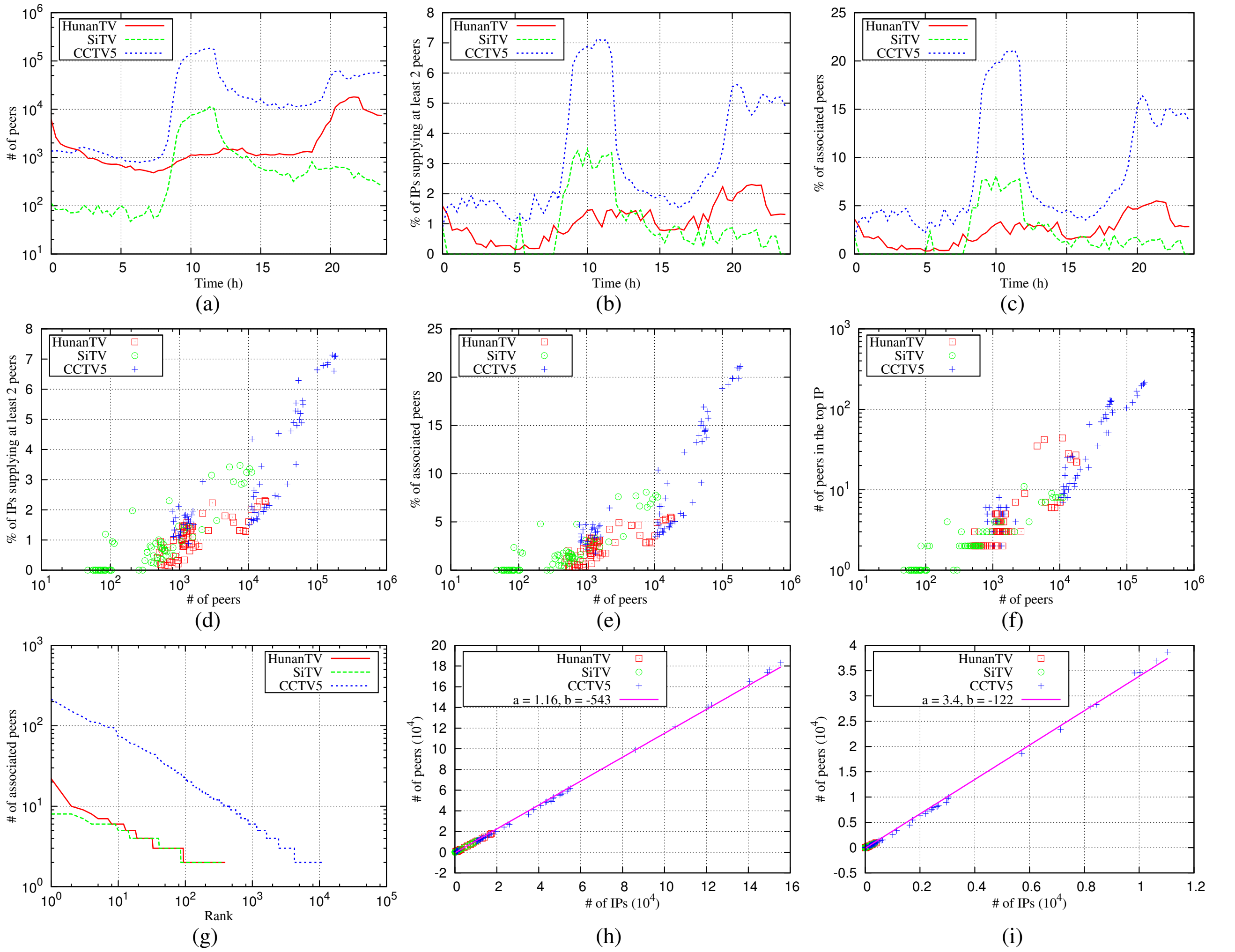

Figure 2: One-day evolutions of (a) the channel population, (b) the percentage of the IPs that supply at least 2 peers and (c) the percentage of the peers from the IPs that supply at least 2 peers; correlations between the channel population and (d) the percentage of the IPs that supply at least 2 peers, and (e) the percentage of the peers from the IPs that supply at least 2 peers, and (f) the number of peers in the top ranked IP, respectively; (g) peer distribution in the IPs that supply at least 2 peers at peak; (h) correlation between the number of IPs and the channel population; and (i) correlation between the number of the IPs that supply at least 2 peers and the number of the peers from them. The time zone is GTM+8.

4. CONSTRUCTING A LAN-AWARE OVERLAY TOPOLOGY

The prior section affirmed the existence of multiple peers in the same LAN, establishing the basis for our LAN-aware solution. Because of limited space, we concisely discuss the principles of designing such a solution and propose a heuristic. LAN-awareness aims to exploit local peer resources, thus minimizing the ingress traffic but still preserving streaming quality. There are two challenges that must be addressed to achieve LAN-awareness: 1) importing a complete, uninterrupted stream into a LAN with the least redundancy, and 2) disseminating the stream to all the peers in this LAN.

4.1 Maximizing Stream Integrity with Minimal Bandwidth

We first analyze the former challenge. The whole video stream is quantized as several sub-streams of equal rate, $s_1, s_2, \ldots, s_n$. They are equivalent to the sliding window containing n active chunks of equal size in a mesh-based pull scheme. k peers are assumed to simultaneously stay in a LAN. To assign the tasks of importing a complete stream from the WAN for these peers, we leverage a scoreboard

$$\begin{array}{c} \\ p_1 \\ p_2 \\ \vdots \\ p_k \end{array} \begin{array}{c} \begin{array}{cccc} s_1 & s_2 & \cdots & s_n \end{array} \\ \begin{bmatrix} \clubsuit & \clubsuit & \cdots & \clubsuit \\ \clubsuit & \clubsuit & \cdots & \clubsuit \\ \vdots & \vdots & \ddots & \vdots \\ \clubsuit & \clubsuit & \cdots & \clubsuit \end{bmatrix} \end{array} \quad (7)$$

where $\clubsuit$ can be replaced with 0 or 1. If p_i is designated to download s_j, $\langle p_i, s_j \rangle = 1$. Otherwise, $\langle p_i, s_j \rangle = 0$. In addition, we define the function

$$sch(\vec{s_j}) = \begin{cases} 1, & \sum_{i=1}^{k} \langle p_i, s_j \rangle \geq 1 \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

Link type	Latency (ms)
Inter-AS	normal$(50, 150)$
Intra-AS	normal$(10, 20)$
Intra-LAN	normal$(1, 2)$

Table 2: Latency configuration.

Type	$\binom{\text{download bandwidth}}{\text{upload bandwidth}}$ combination (Kbps)
Homog.	$\binom{1{,}500}{384} \times 100\%$
Heterog.	$\binom{768}{128} \times 30\% + \binom{1{,}500}{384} \times 60\% + \binom{3{,}000}{768} \times 10\%$

Table 3: Bandwidth configuration.

$sch(\vec{s_j}) = 1$ means that the j^{th} sub-stream is scheduled by at least one peer. Hence, the larger $\sum_{j=1}^{n} sch(\vec{s_j})$ is, the more of a complete video stream is imported. Conversely, it is better if $\sum_{i=1}^{k}\sum_{j=1}^{n}\langle p_i, s_j\rangle$ can be kept small, saving last-hop download bandwidth. Furthermore, this download bandwidth is always limited. For simplicity, we also quantize the bandwidth as the number of sub-streams that can be downloaded simultaneously, denoted by $\mathcal{C}$. Thus, the whole problem can be formulated as finding the optimal solution for the following linear program

$$\begin{array}{ll} \textbf{min} & \alpha\sum_{h=1}^{n} sch(\vec{s_h}) + \beta\sum_{i=1}^{k}\sum_{j=1}^{n}\langle p_i, s_j\rangle \\ \textbf{s.t.} & \sum_{i=1}^{k}\sum_{j=1}^{n}\langle p_i, s_j\rangle \le \mathcal{C}. \end{array} \qquad (9)$$

where α and β are weight variables ($\alpha < 0$, $\beta > 0$). This linear program has two goals. We prioritize streaming quality over bandwidth usage so that $|\alpha| \gg |\beta|$.

A heuristic is to nominate a peer as the agent of the LAN having multiple peers. The agent is the only peer that can receive stream data from outside. The video consumption requirement obligates it to bring a complete copy of the video into the LAN. As the network resources of the peers from the same LAN are usually similar, the agent nomination prefers the earliest-arriving peer, which is likely to provide long-term stability. The nomination algorithm can be implemented on the tracker. Assuming n peers participate in the channel in succession, the algorithm is only invoked less than n times when an agent leaves, causing little overhead. Certainly, the nomination could be improved by considering device heterogeneity and using a distributed protocol.

This heuristic is a special solution of the presented linear program, where $\vec{p_i}^{\top} = \vec{1}$ and $\forall j \neq i, \vec{p_j}^{\top} = \vec{0}$ as p_i is the agent. It may not be advisable to rely on the single peer to import a stream. Nevertheless, in this paper, we focus on demonstrating the feasibility and the potential benefits of such a solution. For the future work, we are interested in solving this linear program in a distributed and autonomic way since some other scenarios share a similar problem statement.

4.2 Intra-LAN Stream Dissemination

The other challenge that needs to be addressed is the intra-LAN stream dissemination. The first design choice is whether to utilize IP multicast or application-layer multicast. Usually, LANs are based on Ethernet or WiFi, which supports broadcast functionality. However, we still choose application-layer multicast to provide the broadcast compatibility. For example, some LAN administrators disable the broadcast function for security reasons, and some big LANs are divided into several subnets. Therefore, IP multicast may not always be available.

With application-layer multicast specified, there is another design choice: using a tree-based push scheme or a mesh-based pull scheme. The decision is made based on the network conditions. Intra-LAN bandwidth capacity is usually high, enough to support several peers. Therefore, the tree-based push scheme will not suffer the problem often encountered in the access links of WANs, *i.e.*, the wasting of insufficient upload bandwidth that is not enough to deliver a stream. Moreover, both latency and loss rate are quite low in LAN environments so that intra-LAN connections are relatively more reliable. Additionally, a tree-based overlay is controllable as the intra-LAN population is usually small. Therefore, the tree-based push scheme which introduces less signaling overhead is considered as a better choice. In our heuristic, we build an overlay similar to *ESM* [1] for intra-LAN connections. Since the number of peers in most LANs is small (less than 10), the tree depth is usually just 2. Nevertheless, existing systems, such as PPLive, can keep using a mesh-based pull scheme to minimize the system changes.

5. EVALUATION OF BENEFITS FROM LAN-AWARENESS

Now we evaluate the benefits of our LAN-aware solution through trace-driven simulations.

5.1 Simulation Setup

We built a mesh-based prototype on OMNeT++. The prototype incorporates a heuristic chunk scheduler similar to that of a real system [3]. The scheduler gives priority to the rarest chunks in the partners and to the recently played-back chunks. A stream server supplies peers with the chunks that are played back immediately but are inaccessible from their partners.

The overall simulation settings are as follows. The length of the simulation time is set to 2 hours. The number of concurrent peers is kept to around 10,000. We assume two levels of peer density in the LAN: 1) a low level (L) of about 2,400 peers staying in 1,100 LANs containing more than 1 peer, and 2) a high level (H) of about 4,200 peers in 1,300 such LANs. Next, the underlay topology is assumed to consist of 10 ASes, and peers non-uniformly spread among these ASes (our measurement reflects that 90% of peers gather in the top 10 ASes). The latency setting of different links is shown in Table 2. As to churn, we let the length of peer life follow a log-normal distribution (Log-$\mathcal{N}(3.5, 0.85)$ min). Furthermore, we consider both homogeneous (Hm) and heterogeneous (Htr) bandwidth settings of the last-hop link to the Internet, as Table 3 shows. Meanwhile, the intranet bandwidth is symmetrically set to 10 Mbps, and the video rate is set to 400 Kbps. Finally, an evaluation scenario can be represented by a triple $\langle$LAN-aware (A) or LAN-unaware (U), peer density level, bandwidth type$\rangle$.

5.2 Experimental Results

Generally, the experimental results, shown in Figs. 3(a) – (d), demonstrate the improvement of traffic efficiency with our LAN-aware solution, benefiting ISPs, content providers and users. Moreover, we observe the trend that the improvement is more prominent if the peer density in LANs is higher.

5.2.1 Reducing Internet Streaming Traffic

One obvious benefit is that less streaming traffic will enter the Internet if the LAN-aware solution is applied, shown in Fig. 3(a), because part of the peers ($\mathcal{N} - \mathcal{X}$) satisfy their streaming demand within the LAN. Specifically, there is always a certain fraction of traffic traversing through ASes due to the AS-level peer distribution. Thus, the problematic inter-AS traffic also decreases, indicating that our solution is ISP-friendly as well.

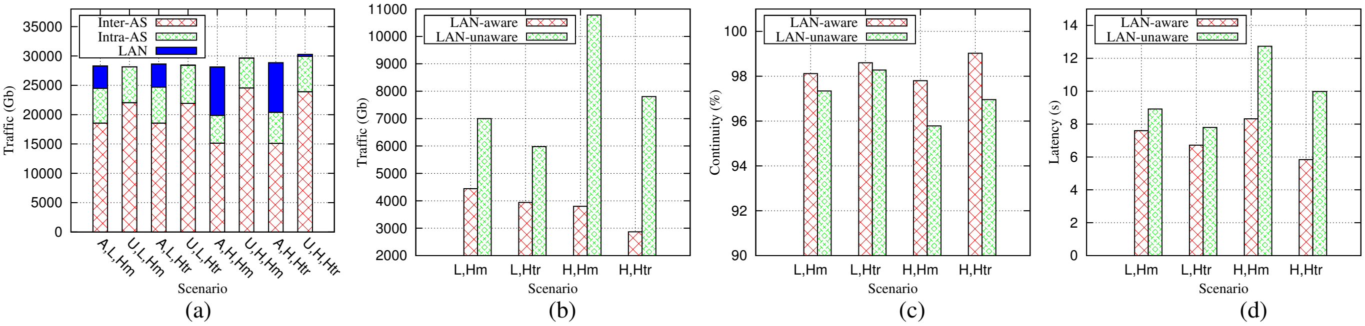

Figure 3: (a) Accumulated traffic of different types during 2-hour simulation time; (b) accumulated traffic delivered by the stream server; (c) playback continuity of the peers (inferring streaming quality); and (d) average latency of receiving the stream.

5.2.2 Lowering Stream Server Workload

Another benefit is for content providers. Usually, the last-hop upload bandwidth is not enough to make the system self-scaling. In particular, multiple peers sharing a single last-hop link intensifies the deficit in peers' contributions. To make up for the deficit, the stream server is required. The amount of compensation is proportional to the population. However, the LAN-aware solution addresses part of peers' streaming demand without any traffic coming from the Internet, equivalent to reducing the population of peers requesting the stream server's help. Hence, as is shown in Fig. 3(b), the amount of the stream server's upload traffic decreases with the LAN-aware solution.

5.2.3 Improving Streaming Quality

We are also interested in the impact on streaming quality. Fig. 3(c) presents the average playback continuity of the peers. The playback continuity equals the percentage of requested chunks that are received by the peers before their playback deadlines. With the LAN-aware solution, the concurrent peers are prevented from competing for the limited last-hop download bandwidth. The single agent can monopolize abundant bandwidth to make up for the sub-optimal chunk scheduling. Meanwhile, the stream dissemination to other local peers is lossless. Therefore, the LAN-aware solution improves the playback continuity. Moreover, the good continuity achieved by our solution also implies a high chunk availability on peers, amplifying the benefits of lowering the stream server's workload.

Next, Fig. 3(d) shows that the latency drops when the LAN-aware solution is applied. The latency is proportional to the diameter of the mesh overlay. With the LAN-aware solution, a fraction of peers are removed from this overlay and participate in their intra-LAN overlays, thus decreasing the diameter and the latency over WAN as a consequence. In contrast, a LAN is so fast that the stream dissemination latency within a LAN is negligible against the latency over WAN. From another point of view, the LAN-aware solution effectively achieves that part of the distant links are replaced with short ones.

6. CONCLUSION

We presented the concept and the benefits of LAN-awareness. To confirm the potential of LAN-aware solutions, we conducted a large-scale measurement on PPLive. We then discussed the principles to construct a LAN-aware overlay and proposed a heuristic. The trace-driven simulations show that LAN-awareness helps to reduce Internet streaming traffic, lower stream server workload and improve streaming quality.

Acknowledgments

This research has been funded in part by A*Star grant 082 101 0028.

7. REFERENCES

[1] Y.-h. Chu, S. G. Rao, S. Seshan, and H. Zhang. A case for end system multicast. *IEEE Journal on Selected Areas in Communications*, 2002.

[2] Cisco Systems, Inc. Cisco Visual Networking Index: Forecast and Methodology, 2009-2014. White Paper, 2010.

[3] X. Hei, C. Liang, J. Liang, Y. Liu, and K. W. Ross. A Measurement Study of a Large-Scale P2P IPTV System. *IEEE Transactions on Multimedia*, 2007.

[4] C.-Y. Lai, C.-S. Hsu, and T.-J. Shang. Improved P2P Streaming in Wireless Networks Utilizing Access Point P2P Agents. In *IEEE International Symposium on Multimedia Workshops*, 2007.

[5] B. Li, S. Xie, Y. Qu, G. Y. Keung, C. Lin, J. Liu, and X. Zhang. Inside the New Coolstreaming: Principles, Measurements and Performance Implications. In *INFOCOM*, 2008.

[6] Y. Liu, L. Guo, F. Li, and S. Chen. A Case Study of Traffic Locality in Internet P2P Live Streaming Systems. In *ICDCS*, 2009.

[7] N. Magharei, R. Rejaie, V. Hilt, I. Rimac, and M. Hofmann. ISP-Friendly Live P2P Streaming. Technical report, University of Oregon, 2009.

[8] Miniwatts Marketing Group. Internet Growth Statistics. http://www.internetworldstats.com/emarketing.htm, 2010.

[9] F. Picconi and L. Massoulie. ISP Friend or Foe? Making P2P Live Streaming ISP-Aware. In *ICDCS*, 2009.

[10] Z. Shen and R. Zimmermann. ISP-friendly peer selection in P2P networks. In *ACM Multimedia*, 2009.

[11] E. Tan, L. Guo, S. Chen, and X. Zhang. SCAP: Smart Caching in Wireless Access Points to Improve P2P Streaming. In *ICDCS*, 2007.

[12] D.-C. Tomozei and L. Massoulie. Flow Control for Cost-Efficient Peer-to-Peer Streaming. In *INFOCOM*, 2010.

[13] L. Vu, I. Gupta, J. Liang, and K. Nahrstedt. Measurement of a large-scale overlay for multimedia streaming. In *International Symposium on High Performance Distributed Computing*, 2007.