Research Article

Bandwidth Impacts of Localizing Peer-to-Peer IP Video Traffic in Access and Aggregation Networks

Kenneth Kerpez

Telcordia Technologies Applied Research, One Telcordia Drive, Piscataway, NJ 08854-4151, USA

Correspondence should be addressed to Kenneth Kerpez, kkerpez@telcordia.com

This paper examines the burgeoning impact of peer-to-peer (P2P) traffic IP video traffic. High-quality IPTV or Internet TV has high-bandwidth requirements, and P2P IP video could severely strain broadband networks. A model for the popularity of video titles is given, showing that some titles are very popular and will often be available locally; making localized P2P attractive for video titles. The bandwidth impacts of localizing P2P video to try and keep traffic within a broadband access network area or within a broadband access aggregation network area are examined. Results indicate that such highly localized P2P video can greatly lower core bandwidth usage.

1. INTRODUCTION

Localized P2P ensures that files are preferably delivered from nearby subscribers, in order to limit backbone bandwidth. Many new IPTV rollouts envision deploying a personal video recorder (PVR) at each subscriber's location. These PVRs could potentially be used to offload traffic from network video servers, by allowing the PVRs to serve titles via P2P. In this way, localized P2P could be implemented by network and service providers. Copyright protection as well as a number of management challenges would need to be addressed to make this practical. Digital rights management (DRM) systems for IP video are emerging to handle copyright protection.

Statistics on current P2P usage are presented in Section 2, along with current techniques for handling high-bandwidth P2P traffic. Then the model of P2P IP video is presented in detail followed by results of simulations.

Analyses here model IP video, particularly P2P delivery and cases where P2P IP video is localized. A model of the demand for video titles is developed based on TV and movie ratings. This model is used to determine frequency of requests for video titles, which in turn determines which titles are stored and subsequently made available to others via P2P. A typical broadband IPTV network aggregation hierarchy is assumed. Models of serving area sizes are

developed and used here that are derived from telephone serving area statistics.

These models are used in simulations that determine IP video bandwidth utilizations at different levels of aggregation. Multicast (linear broadcast), network server-based video on demand (VOD), and P2P IP video traffic is compared at the different levels of network aggregation from access to core network. P2P is localized and P2P usage is varied to determine how it impacts bandwidth. It is shown that backbone bandwidth usage could be greatly lowered by localized P2P for IP video. However, P2P traffic increases upstream access bandwidth utilization.

IP video has two distinct flavors:

(i) IPTV—a service provided by a network operator, perhaps in conjunction with other providers for content and service.

(ii) Internet TV—a service that is generally uncoordinated with the network provider. Called "over-the-top" since it runs over a broadband network without the network provider ever being aware of it.

While P2P IP video is currently associated with Internet TV, it is conceivable that at least some IPTV services would also be provided by P2P.

2. PEER-TO-PEER (P2P) TRAFFIC

It is well known that peer-to-peer (P2P) traffic (primarily file-sharing) currently uses large amounts of the bandwidth on the Internet; approximately 60% of all Internet traffic is currently P2P. P2P can particularly overload upstream access network bandwidth. Unlike downloads from large server sites, P2P data travels upstream across broadband access networks as well as downstream.

Studies of residential Internet access traffic in Japan [1, 2] during 2004 and 2005 found that much bandwidth is used by P2P. Specifically, this study found that 62% of the total volume is user-to-user traffic. A small segment of users dictate the overall behavior; 4% of heavy-hitters account for 75% of the inbound traffic volume. Some of this traffic was probably from small businesses.

Ellacoya Networks studied European Internet traffic in 2005 [3], and found that usage by the top most active 5% of subscribers represented approximately 56% of total bandwidth, while the top 20% of active subscribers consumed more than 97% of total bandwidth. P2P was by far the largest consumer of bandwidth with 65.5% of traffic on the network being P2P applications. Web surfing (HTTP) consumed 27.5% of Internet bandwidth. In numbers of subscribers, web surfing (HTTP) was the most popular application with average daily peaks at 50% of subscribers. Instant messaging (IM) was the second with average daily peaks at 25% of subscribers; while P2P, with average daily peaks at 18% of subscribers, was the third. E-mail had 12% and other applications 10%. There are relatively few P2P users, but they use large amounts of bandwidth.

CacheLogic conducted direct packet monitoring of Internet backbones and ISPs data streams via Layer 7 packet analysis [4]. This study found 61.4% of current peer-to-peer traffic to be video, 11.4% audio, and 27.2% other traffic. On a global scale, 46% of P2P traffic was video in Microsoft formats. 65% of all audio files by volume of traffic were still traded in the MP3 format, and 12.3% were in the open-source OGG file format used by BitTorrent.

Video P2P poses unique bandwidth challenges. P2P systems are emerging for streaming linear broadcast TV via application-layer multicast [5].

2.1. Limiting the impact of P2P

Technical solutions to P2P bandwidth usage include adding more network capacity (particularly upstream broadband access), localizing content, and caching with network servers. There are also ways to either limit P2P traffic [6]; or to have users pay for high-bandwidth usage. P2P traffic could simply be bandwidth limited in an effort to impose fairness, or price mechanisms could be imposed. These limitations might only be implemented during times of heavy usage. Such limits are controversial.

Premium service levels, particularly offering high throughput, could be offered for an increased fee. Bandwidth could be metered, by charging some cost per gigabyte, or via subscriptions to a certain number of gigabytes per month, which could be similar to charging for cellular minutes.

2.2. Network-coordinated P2P

P2P can be combined advantageously combined with traditional network-based delivery techniques. References [7, 8] both focus on methods that allow partial control of P2P traffic by network providers; with [7] focusing on TV P2P and [8] focusing mainly on optimizing current file transfers. Multicast infrastructure is used by network providers to limit core bandwidth, and multicast can be combined with P2P in interesting ways, for example, allowing "VCR" controls even for linear broadcast service [9].

2.3. Localized P2P

P2P systems that prefer to get content from the most local sources should have traffic that traverses fewer links. A recent article [10] showed that 99.5% of current P2P traffic (using "eDonkey" in France) traversed national or international networks. It further showed that 41% to 42% of this long distance traffic could be made local if a preference for local content was built into the protocol.

PVRs that are deployed by network or service providers could host P2P video in order to offload traffic from network video servers. The possibility emerges that P2P could actually save network bandwidth by delivering titles from a source that is closer than the nearest network-owned video server. A number of control, oversight, and copy protection issues would need to be addressed to make this practical. These are beyond the scope of this paper but have been considered elsewhere [11]

Localizing P2P has been discussed previously. It is not uncommon to use the IP hop count or TTL value to localize somewhat when choosing peering sources. Current routing schemes in P2P networks such as Chord [12] work by correcting a certain number of bits at each routing step. Reference [13] used the IP number to localize, and found that the first octet in the IP number provided localization to roughly a national level, improving over global.

3. PEER-TO-PEER IP VIDEO MODEL AND ANALYSES

A detailed model for analyzing peer-to-peer (P2P) traffic was created. This model emphasizes broadband access networks and their aggregation networks. The model analyzes localized P2P; where a requested title is delivered by a local source if possible, rather than a more distant source. This may be accomplished with network-provided and controlled personal video recorders (PVRs), which are becoming popular. Results show how much bandwidth in each network segment P2P IP video would need to use.

Today's P2P traffic is generally not localized and traffic flows anywhere so there is a little difference between aggregate backbone P2P bandwidth and aggregate access network P2P bandwidth. This modus operandi is already wasting much of the bandwidth of the Internet, and if high-quality video becomes the P2P norm, then solutions such as localized P2P will become very desirable.

The focus here is IP video because it is emerging as a potentially huge bandwidth hog. P2P is tinged with

copyright infringement issues. Emerging digital rights management (DRM) systems claim to work with P2P, by encrypting content at the source, allowing free distribution of encrypted content, but only allowing playback on authorized devices after the individual user obtains the right keys which are controlled by the content owner. It may be the case that a number of top-run titles are not released for VOD or P2P, and so this scenario is also modeled.

3.1. *Peer-to-peer video streams*

The usual model of IPTV shows traffic originating at a headend or video hub office (VHO) which is owned by the network operator, and then flowing purely downstream to subscribers. Ignoring production feeds, this is how cable TV works. Figure 1 shows the basic network aggregation hierarchy: a super headend (SHE) feeds video hub offices (VHOs), which feed video serving offices (VSOs), which then feed optical line terminals (OLTs) or digital subscriber line access multiplexers (DSLAMs).

Figure 1 shows that P2P does not only add new upstream loads, but it can also displace downstream bandwidth since it need not travel all the way down from the headend or VHO.

4. VIDEO DEMAND MODEL

The most popular video titles are viewed far more than the least popular. This is particularly true for TV shows, first-run movies, and recently released DVD rentals. TV, movie, and video rental rating statistics were examined, and a demand model of video titles was created by matching these statistics. The model first rank orders all video titles, from most popular to least popular. The title number increases with decreasing popularity. A probability density model is assigned to the title numbers. Analyses here assume that a few of the most popular titles are not available for VOD or P2P; these would be either new movie releases or new broadcast TV shows that are multicast. The model here was built from TV and movie ratings data allyourtv.com, http://www.hollywoodreporter.com/hr/index.jsp

It was found that the popularity of linear broadcast TV channels and new movie releases is modeled closely by an exponential probability density function, with a few titles very popular and the less popular content rapidly dropping off with increasing title number. These include weekend movie gross. Other types of video have a longer tail. For example, video rentals are better modeled with a long-tailed probability density, since there is still some use of even the least popular titles, and these are better matched by a power function or hyperbolic density. The statistics and model are shown in Figure 2 for low-title numbers.

There are few statistics for higher-title numbers, out in the very long tail. Some DVD mail-order rental services purportedly have over 65 000 different titles. It can be expected that the long tail is somewhat supply-driven. As more titles are offered, someone will eventually view them. Such a long tail is well modeled by the hyperbolic density, assigning even the least used titles a probability significantly above zero.

A combined video demand model was used here, this model is a mixture of exponential and hyperbolic probability densities. The density is truncated to limit to a finite number of available titles, title #n, such that $n\min \leq n \leq n\max$. The mathematical definition of the model is

$$\text{video demand probability model}$$
$$= (\text{In})\text{hy}(n) + (1 - \text{In})\text{ex}(n) \tag{1}$$

with

$$\text{hy}(n) = \text{truncated hyperbolic density},$$
$$\text{ex}(n) = \text{truncated exponential density}, \tag{2}$$

where In is an indicator for a Bernoulli random variable; $\Pr(\text{In} = 1) = p_1$, $\Pr(\text{In} = 0) = p_2$; and satisfying $p_1 + p_2 = 1$.

Further, the truncated hyperbolic density is

$$\text{hy}(n) = C1 * n^{-A}, \quad n\min \leq n \leq n\max, \tag{3}$$

where A is a constant, and

$$C1 = \frac{1 - a}{n\max^{(1-A)} - n\min^{(1-A)}}. \tag{4}$$

The mean of the truncated hyperbolic probability density is

$$E(\text{hy}(n)) = \frac{C1 * (n\max^{(2-A)} - n\min^{(2-A)})}{2 - A}. \tag{5}$$

The truncated exponential density is

$$\text{ex}(n) = C2 * e^{(-B*x)}, \quad n\min \leq n \leq n\max, \tag{6}$$

where B is a constant, and

$$C2 = \frac{B}{e^{-(B*n\min)} - e^{-(B*n\max)}}. \tag{7}$$

The mean of the truncated exponential density is

$$E(\text{ex}(n))$$
$$= \left(\frac{C2}{B}\right) * \left(n\min * e^{-(B*n\min)} - n\max * e^{-(B*n\max)}\right) + \left(\frac{1}{B}\right). \tag{8}$$

The overall mean is

$$E(\text{Video demand probability model})$$
$$= p_1 * E(\text{hy}(n)) + p_2 * E(\text{ex}(n)). \tag{9}$$

The TV title demand model parameters used in numerical evaluations here are

$$A = 0.4, \qquad B = 0.09, \qquad p_1 = p_2 = 0.5. \tag{10}$$

5. NUMBER OF TITLES PER SUBSCRIBER

Another aspect is the number of video streams that are viewed by each subscriber. Statistics http://www.nielsen-media.com/nc/portal/site/Public/ show that the maximum
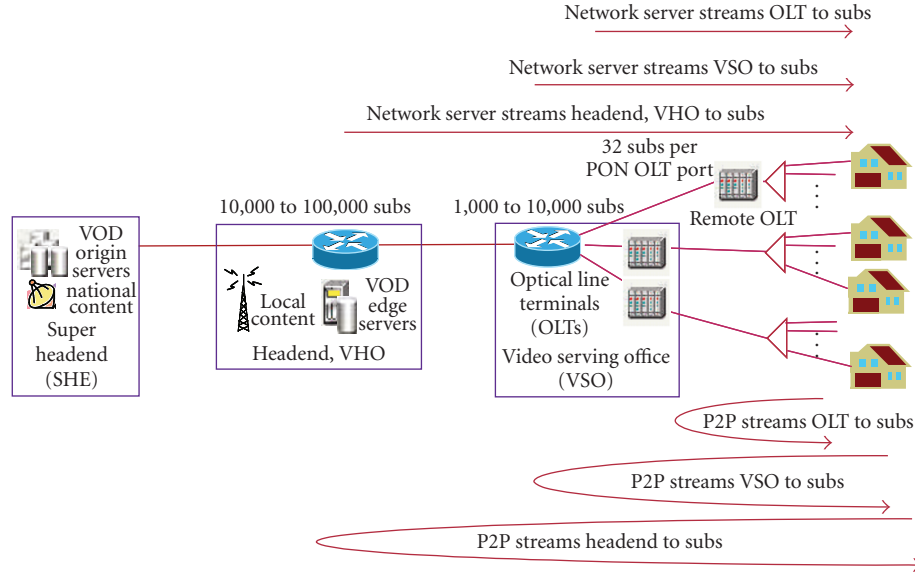
FIGURE 1: Localized Peer-to-Peer traffic. The optical line terminal (OLT) is logically similar to a DSL access multiplexer (DSLAM), except that the DSLAM typically serves several hundred subscribers.
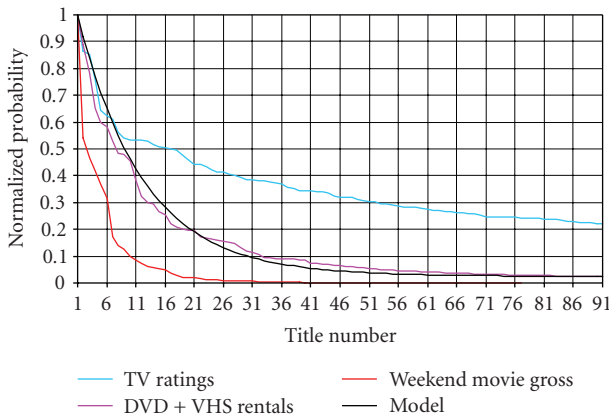

FIGURE 2: Video demand statistics and model. Statistics and probabilities are normalized so that the probability density of title number 1 = 1.

TABLE 1: TV viewing hours per day in 2003, FCC data.

Men	Women	Teens	Children	Average
4:29	5:05	3:07	3:14	3:58

busy hour (prime-time peak) has about 66% of all households watching TV. Also, each home averages a little less than 2 simultaneous TV viewings per home in the busy hour.

According to FCC statistics in 2005, 90% of US homes had TVs, and 67.5% subscribed to cable TV. The average number of TV sets per household was 2.62. Average TV viewing hours per day is listed in Table 1.

IP video subscribers may watch a little more video than the average person. A simple discrete and independent model

TABLE 2: Model used in analyses here of probability of number of video streams per IP video subscriber in the busy hour (prime time). Average = 1.8 streams per sub.

Number of titles	0	1	2	3	4	>4
Probability	0.1	0.35	0.3	0.15	0.1	0

for the number of video viewings per home suffices here, an example of which is in Table 2.

The average total number of streams to each subscriber is 1.8. When considering P2P or VOD only, the video demand model of Section 4 is evaluated for some number n min to determine the proportion of all video titles that can be P2P. The result of this then multiplies the probabilities in Table 2 to determine the probabilities of numbers of P2P or VOD titles requested by each subscriber.

6. SERVING AREA MODEL

Previous work [14], seen in Figure 3, showed that a gamma probability can closely model statistics of telecom serving area sizes. Distance from CO to subscribers or from serving remote terminal to subscribers can be closely fit to a gamma probability. The gamma probability density function (pdf), defined as

$$\Pr(\text{radius} \leq x) = \int_0^x \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} \, dx$$

$$\text{with } \Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} \, dx.$$
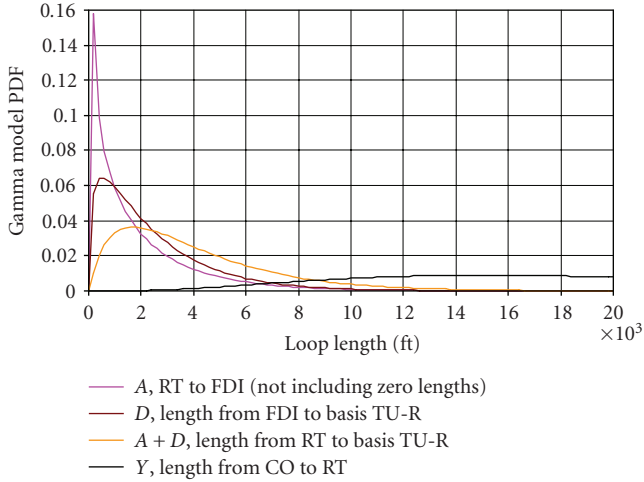
(11)

FIGURE 3: Gamma models of current telephone plant serving area radii [14]. Terminology: central office (CO), remote terminal (RT), and feeder-distribution interface (FDI).

TABLE 3: Simulated serving area sizes for P2P results.

Serving area	Avg. radius (miles)	Avg. no. of subs	Max. no. of subs
OLT or DSLAM	0.4	15	34
VSO or CO	3.0	848	1909
Head end or VHO	10.0	9425	21206

Define the mean of the gamma to be μ, and the standard deviation to be σ. Then

$$\mu = \alpha\beta, \qquad \sigma^2 = \alpha(\beta^2), \qquad \alpha = \frac{(\mu^2)}{(\sigma^2)}, \qquad \beta = \frac{(\sigma^2)}{\mu}. \tag{12}$$

Here, the gamma model determines serving area radii according to the averages in Table 3. Given the radius, r, of the serving area, the number of subscribers in the serving area is simply the average subscriber density multiplied by the serving area size, πr^2.

The gamma model parameters used for modeling serving area sizes here are as follows:

$$\alpha = 25, \qquad \frac{\sigma}{\mu} = 0.2. \tag{13}$$

Average number of subscribers per square mile (subscriber density) = 30.

7. TRAFFIC SIMULATIONS

Monte-Carlo simulations here repeatedly randomly generate serving areas, video demand, P2P supply, and so on, and collect statistics on P2P bandwidth usage. Recall that there is a nested hierarchy of aggregation serving areas: headend/VHO > VSO > OLT/DSLAM. Serving area sizes (number of subscribers per aggregation level) are randomly generated using the model of Section 6 as follows. First,

the total number of subscribers in a VHO or headend is generated. Then the number of subscribers in each individual VSO serving area is generated, until the sum number of subscribers in all VSOs in this VHO equals or exceeds the total number of subscribers in the VHO, then no more VSOs are generated and the size of the VHO is recalculated if need be. Then OLT/DSLAM serving area sizes are generated until the number in each VSO is reached similarly.

Each subscriber is randomly assigned some number of simultaneously demanded video streams according to the model of Section 4. A title is assigned to each demanded video stream using the model of Section 5.

Each subscriber is assumed to store some number of P2P titles. The identity of these stored titles is determined by the same model as used for video demand in Section 4. Each of these stored titles can be delivered if demanded by other subscribers, from the closest source available. The current model only calculates this distance as the height in the aggregation tree (i.e., same OLT/DSLAM < same VSO < same headend), although this could be easily generalized. The simulation first searches for the demanded title in the local OLT/DSLAM area; if there is none then the VSO area is searched, if there is none then the title is delivered through the VHO/headend.

Statistics count up the number of video streams in each segment of the network. Simulations regenerate all serving area sizes, demanded video titles, stored P2P titles, and so forth, 200 times. The total available number of titles is selectable, and is chosen here to be $n\max = 4000$.

The most popular titles are highly likely to be demanded, and they are highly likely to be stored for P2P availability. These results in the localized P2P system here very often deliver content from nearby sources. If only very long-tailed content was available for P2P, then much less traffic would be localized.

7.1. Traffic simulation results

The first set of results ignores P2P effects, and just shows the difference between multicast and unicast traffic. Unicast traffic could be P2P or VOD. Here, it is simply assumed that a number of the top titles, $n = 1, \ldots, n\min -1$ are multicast from the VHO/headend, using the model of Section 4. The remaining less popular titles are unicast from headend servers. Unicast traffic has the same load per subscriber at any point in the network (one stream per subscriber). Multicast is aggregated; there is only one multicast stream for each title between a VHO and a VSO, and between a VSO and an OLT or DSLAM. Figure 4 shows that multicasting popular titles can save large amounts of backbone and aggregation network bandwidth. However, there are broadcast and VOD services which are fundamentally different in many ways, and so a full comparison is not as simple as Figure 4.

Figure 5 shows P2P traffic at various points in the aggregation network, as a function of the number of titles stored available for distribution from each subscriber. This assumes that the 100 most popular titles are not available for P2P. With no localization, aggregation and backbone traffic would be about the same as access network band-
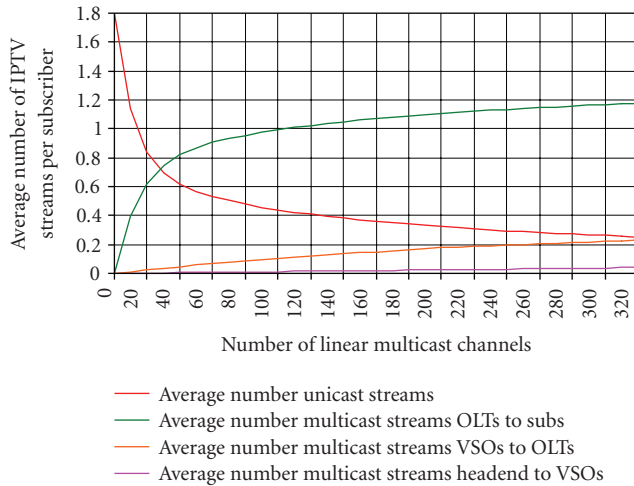
Figure 4: Unicast versus multicast bandwidth. Unicast is considered to be VOD, delivered from the headend or VHO, not P2P.
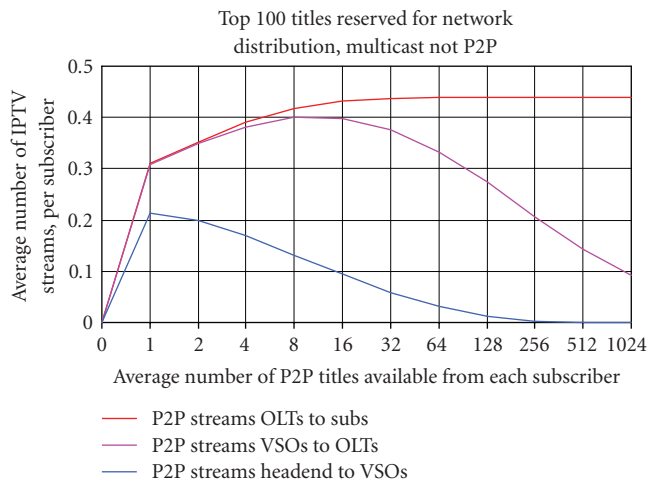

Figure 5: P2P network traffic at various network aggregation points.

width. Figure 5 shows that aggregation (VSOs to OLTs) and backbone bandwidth (headend to VSOs) is much lower than the access network bandwidth (OLTs to subs) as a result of P2P localization.

8.   CONCLUSIONS

P2P already uses a large amount of Internet bandwidth. IP video is now emerging, and for broadcast quality digital video, multiple megabits of data are used per stream. Combine the two, and P2P IP video could overwhelm the network if not properly anticipated and managed.

Localizing IP video should lower core bandwidth usage, and this lowering was quantified by simulations here. Localized P2P video can often be delivered from nearby, within a local serving area, without impacting long-haul network bandwidth. Results show that core network bandwidth could

be greatly decreased by localization. These results show more pronounced bandwidth savings by localized P2P as traffic moves further into the core, and as more titles are stored locally by the subscriber.

Allowing users to serve even a small number of P2P video titles can save bandwidth, mainly because a small number of popular titles should account for most usage. Actually, this has a double effect for P2P video; the most popular videos are not only frequently requested, but they are also frequently available from nearby sources since they are so common.

Besides lowering core and aggregation network bandwidth, P2P should also lower network server usage. However, P2P is far from being purely virtuous. Unlike VOD or multicast, P2P streams will all need to traverse upstream access links and could easily overwhelm access networks with limited upstream bandwidth. Moreover, many issues could impact reliability and availability of P2P-based video service; ranging from copyright issues, to P2P sources being turned off in midstream, to enabling QoS on P2P.

REFERENCES

[1] K. Cho, K. Fukuda, H. Esaki, and A. Kato, "The impact and implications of the growth in residential user-to-user traffic," *Computer Communication Review*, vol. 36, no. 4, pp. 207–218, 2006.

[2] Draft Report of the Study Group on a Framework for Competition Rules to Address the Transition to IP-Based Networks, "New competition promotion program 2010," MIC, Japan, July 2006.

[3] M. Burke, "Ellacoya networks releases data on broadband subscriber and application usage in Europe," September 2005, www.ellacoya.com/news/pdf/2005/EllacoyaEuropeData.pdf.

[4] A. Parker, "P2P in 2005," http://www.cachelogic.com/.

[5] J. Liu, S. G. Rao, B. Li, and H. Zhang, "Opportunities and challenges of peer-to-peer internet video broadcast," *Proceedings of the IEEE*, vol. 96, no. 1, pp. 11–24, 2008.

[6] A. Spognardi, A. Lucarelli, and R. Di Pietro, "A methodology for P2P file-sharing traffic detection," in *Proceedings of the 2nd International Workshop on Hot Topics in Peer-to-Peer Systems (HOT-P2P '05)*, pp. 52–61, San Diego, Calif, USA, July 2005.

[7] M. Cha, P. Rodriguez, S. Moon, and J. Crowcroft, "On next-generation telco-managed P2P TV architectures," in *Proceedings of the 7th International Workshop on Peer-to-Peer Systems (IPTPS '08)*, Tampa Bay, Fla, USA, February 2008.

[8] H. Xie, A. Krishnamurthy, A. Silberschatz, and Y. R. Yang, "P4P: explicit communications for cooperative control between P2P and network providers," http://www.dcia.info/documents/P4P_Overview.pdf.

[9] X. Y. Yang, P. Hernandez, F. Cores, et al., "DVoDP^{2}P: distributed P2P assisted multicast VoD architecture," in *Proceedings of the 20th International Parallel and Distributed Processing Symposium (IPDPS '06)*, p. 10, Rhodes Island, Greece, April 2006.

[10] L. Plissonneau, J.-L. Costeux, and P. Brown, "Detailed analysis of eDonkey transfers on ADSL," in *Proceedings of the 2nd Conference on Next Generation Internet Design and Engineering (NGI '06)*, pp. 255–262, Valencia, Spain, April 2006.

[11] V. Janardhan and H. Schulzrinne, "Peer assisted VoD for set-top box based IP network," in *Proceedings of the Workshop on*

Peer-to-Peer Streaming and IP-TV (P2P-TV '07), pp. 335–339, ACM SIGCOMM, Kyoto, Japan, August 2007.

[12] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan, "Chord: a scalable peer-to-peer lookup service for internet applications," in *Proceedings of the ACM SIGCOMM Conference on Applications, Technologies, Architectures, and Protocols for Computers Communications*, pp. 149–160, San Diego, Calif, USA, August 2001.

[13] R. A. Ferreira, A. Grama, and S. Jagannathan, "An IP address based caching scheme for peer-to-peer networks," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '03)*, vol. 7, pp. 3845–3850, San Francisco, Calif, USA, December 2003.

[14] K. Kerpez and C. Valenti, "Statistical variables for evaluating compatibility of remote deployments," T1E1.4/2001-132, May 2001.