P2P, CDNs, and Hybrid Networks: 
The Economics of Internet Video Distribution

Imsook Ha*
Steven S. Wildman**
Johannes M. Bauer***

ABSTRACT

One of the challenges of developing a cost effective distribution system that supports the proliferation of video streams is allocating the video files optimally among various means of media storage. In this paper we develop simple mathematical representations of the cost characteristics of two internet video distribution systems: Content Delivery Networks (CDNs) and Peer-to-Peer (P2P) networks and identify conditions under which each or a blend of the two proves most cost-effective. Our analysis shows that in many cases a hybrid system is likely to have lower costs than either a pure CDN or a pure P2P network. We identify the personal computers that form the backbones of P2P networks as a largely untapped network resource and suggest that potentially very large economic gains might be realized if policy makers can develop policies that effectively address some of the security and transaction cost problems that discourage PC owners from contributing capacity to P2P networks.

Key words: Internet video distribution, Network architecture, Internet policy

JEL codes: L1, L82
I. INTRODUCTION

As the bandwidth available to broadband subscribers has increased and the household penetration of broadband has grown, the importance of the internet as a vehicle for watching video programming has grown apace. Based on a survey of online households in the US, ABI Research (2008) estimated that the fraction of US television viewers who watched video streamed through a browser in 2008 was twice its level in 2007 (63 percent vs. 32 percent). This is undoubtedly both a response and a contributor to the rapid growth in the amount of rich content available through a growing list of online services that includes video hosting services like YouTube and Hulu, websites maintained by television networks to directly stream their programs to viewers, and online movie and program streaming services offered by video rental companies like Netflix and Amazon.com. eMarketer (2008) predicted that by 2011 the number of online video viewers in the United States would grow to 183 million and that 86.6 percent of US Internet users would be watching video online, a trend that by itself should spur further growth in online video services. Business models are still evolving however, as online video services search for better revenue models and less costly ways to deliver content. Attempts to work through the profit implications of online delivery have mostly focused on the revenue side of the profit equation. This paper focuses on the cost side by examining the cost characteristics of different online video delivery systems.

The simplest internet video distribution system employs a single origin server to supply content to all clients. Although easy to implement, this approach is vulnerable to service interruptions as the load on the server and surrounding network elements increases with traffic volume. In addition, bandwidth costs escalate rapidly as traffic grows (Thouin & Coates, 2007). The Content Delivery Network (CDN) was introduced to solve these problems. In the CDN configuration, replica servers (also called edge servers) are installed at strategic locations close to clients to reduce the load on the origin server and to reduce bandwidth costs by
reducing the average distance between server and client. More recently peer-to-peer (P2P) systems have emerged as an alternative used by some prominent online video services. For example, the VeohTV, Babelgum TV and Joost online video services all launched using P2P distribution (Joost subsequently switched to CDN.) and CNN.com was recently reported to rely on a P2P service to supply as much as 30 percent of the bandwidth used for its live video stream (Chartier, 2009). An example outside the United States is the British Broadcasting Corporation’s iPlayer (Chartier, 2009), which is reported to be the most popular video-on-demand streaming service in the United Kingdom (Yoskowitz, 2010).

The cost savings claimed by advocates for P2P systems stem mainly from the substitution of client/peer storage capacity and processing power for edge server capacity in a CDN, and, to some extent, the utilization of unused portions of bandwidth contracted for by owners of peer PCs. In this paper we model and compare the economic characteristics of the P2P and CDN distribution systems under conditions that allow us to ignore bandwidth costs and focus on differences in the costs of procuring storage and processing capacity for the two systems to identify conditions under which each or a blend of the two is most cost-effective.

This article is organized as follows. In the next section, we describe the architectures for the three principal types of content delivery systems and introduce the elements of our video distribution cost models. CDNs and P2P systems are both employed because they have significant cost advantages over simple client server systems. In Section III we summarize the main cost characteristics and develop simple algebraic cost models for a CDN, a P2P, and a hybrid network that utilizes both edge servers and peer-supplied capacity. Using these models, we identify conditions for which a CDN has lower costs than a P2P system and vice versa as well as situations for which costs are minimized by a hybrid CDN-P2P delivery system. Section IV presents a numerical example to illustrate the analysis and conclusions presented in Section III. We summarize our findings in Section V, which also includes a brief discussion of opportunities to use policy to encourage the development of more efficient online delivery systems.
II. INTERNET VIDEO DISTRIBUTION SYSTEMS

1. System Architectures

The simplest distribution system architecture is the basic client-server architecture illustrated in Figure 1. With this architecture content resides in an origin server and a unique stream is sent from the origin server to each requesting client. Simple client-server systems like the one illustrated suffer from congestion and failure at the origin server along with escalating bandwidth costs as the volume of requests increases. Longer transits can also make distant clients substantially more costly to serve than those located near the origin server.

![Figure 1] Centralized architecture

Content delivery networks were developed as a response to these problems. A CDN reduces the load on an origin server by placing copies of its content on edge servers that are geographically closer to clients. A client’s request is directed to the nearest edge server with the capacity to respond. Compared to the simple client-
server architecture, the CDN architecture has a number of advantages, including reduced origin server load, reduced delay for end-users, more efficient usage of bandwidth, and improved scalability through content replication (Buyya, Pathan, Broberg, & Tari, 2006; Pallis & Vakali, 2006). A CDN system is depicted in Figure 2.

P2P networks are a different, although in many ways architecturally similar, response to the limitations of the simple client-server model. Figure 3 describes a P2P system. A critical difference between the video delivery systems depicted in Figure 2 and in Figure 3 is the replacement of CDN servers by PCs owned by network participants (peers), who (in this case) use them to watch video via broadband internet connections. When a peer watches a video program, short segments of the video are saved to her PC’s hard drive. When another user requests the same video, this video comes not from the origin server but from the network of peers. A fragment of the video is then stored on the new client’s hard disk. In this way, the entire program is seeded throughout the network, which is why the origin server for a P2P network is often referred to as a seed server.
2. CDN versus P2P or hybrid networks

The Content Delivery Network (CDN) is by far the most widely-used internet content distribution system. Most Internet video services contract with third party CDN operators rather creating their own distribution systems. According to RESEARCHANDMARKETS (2010), in 2009 CDNs served 22.5 billion professional video views to internet users in the United States, with the market leaders Akamai, Limelight and Level 3 accounting for 31.9 percent, 12 percent, and 11.2 percent of this traffic, respectively. Limelight delivered videos for YouTube prior to its acquisition by Google.

CDN prices have been falling due to capacity expansion, entry by new suppliers, and increasing commoditization of the basic CDN service. Nevertheless, distribution costs are still high relative to revenues for most content delivery services, creating strong incentives to find lower cost delivery options. P2P networks have been intensively studied and are seen as a long-run attractive alternative to CDNs by some researchers (Rowstron & Druschel, 2001; Crowcroft & Pratt, 2002; Miller, Konstan, & Riedl, 2004; Saroiu, Gummadi, & Gribble, 2002).
While the original P2P systems were user-based filesharing services such as Napster, eDonkey, Gnutella, and BitTorrent, today P2P is also used by commercial internet content providers.

Unlike P2P filesharing systems for which a complete copy of a desired unit of content (like a TV program) resides on at least one peer computer, with P2P streaming content units (e.g., programs or other videos) are broken up into many smaller subunits that are then distributed among participating peer computers. Each unit of content thus makes a relatively light demand on the memory and processing capacity of a participating peer’s computer. A P2P service manages access to the distributed fragments to create an appropriately time-ordered sequence that is streamed to a requesting user (Zhou, Li, & Liu, 2008). These systems also economize on bandwidth by initiating streams from the online peers closest to the users requesting content. (Qiu & Srikant, 2004).

Hybrid media streaming systems that combine elements of the CDN and peer-to-peer architectures have been introduced and there is research suggesting that these hybrid systems may be able to deliver media content faster, respond more quickly to requests, and show better system performance than standalone P2P networks or CDNs (Xu, Kulkarni, & Rosenberg, 2006; Tu, Sun, Hefeeda, & Prabhakar, 2005; Hefeeda, Habib, Botev, & Xu, 2003; Padmanabhan, Wang, & Sprapanidkulchai, 2002).

### III. ECONOMIC COST MODEL

In this section we model the cost characteristics of CDN, P2P and hybrid networks to identify the comparative advantages of each. We show that there are circumstances in which each approach has lowest costs. In the following four subsections, we provide an overview of basic video streaming service cost characteristics, present a simple algebraic cost model for a CDN system, develop a cost model for a P2P system using similar terminology, and then present a cost model for a hybrid system and compare cost of service for the three types of video delivery systems.
1. Basic Cost Characteristics

The cost of using a CDN is the sum of the costs of storage and processing at the origin server, the corresponding costs for the network’s edge servers, and what we will call the network’s communication costs, which is the cost of transporting data from the origin server to edge servers and from edge servers to client PCs. We assume that origin server costs are purely a function of content storage and processing capacity and these costs are assumed to be the same for CDNs and P2P systems. Purely for analytical convenience, we assume the cost of an edge server is the same as the cost of an origin server for a CDN, which is different from a P2P system where the cost of an origin server is many times the cost of the smallest amount of capacity that might be accessed through a participating peer’s personal computer.

A PC owner who makes some of her PC’s capacity available to a P2P service may incur some costs in doing so. Processing capacity shared with the service may not be available when needed for other processing-intensive applications, such as video games, and participation as a contributing peer may also increase vulnerability to privacy violations, viruses, and malware. There is thus an opportunity cost associated with participating in a P2P video distribution system and broadband subscribers can be expected to vary in the perceived magnitude of this cost. In our model, the comparative advantages of the CDN and P2P approaches depend on the cost of server capacity compared to opportunity cost to peers of participating in a P2P network.

Our cost models assume a video service provider constructs and manages its own distribution system, as Google does using its own distribution network to distribute videos for YouTube, which it owns, and as the BBC does with its P2P iPlayer service. While contracting with third party service providers is more common for CDNs and is also done for P2P-based services (e.g., CNN.com pays for a P2P distribution service provided by Octoshape), the cost models presented below are not sensitive to the question of whether economies of scale are sufficiently extensive to justify creating independent distribution systems to aggregate demands for bandwidth over multiple video services.
In this study we abstract from the communication costs of the two systems to focus on a comparison of their storage and processing costs. Some articles on P2P systems identify substitution of otherwise unused bandwidth supplied peers by their ISPs for bandwidth purchased from backbone suppliers as a cost saving realized by using a P2P system instead of a CDN. In the long run, we cannot expect ISPs to continue this subsidy to video content suppliers if such services begin to utilize substantial amounts of bandwidth, a point already articulated by UK ISPs in a public dispute with the BBC over handling of traffic generated by the BBC’s popular iPlayer P2P video streaming service. (Anderson, 2007; Chartier, 2009) Thus our model takes the long-run view and assumes that peers do not contribute free bandwidth to a video delivery service and that the cost of bandwidth is the same for both systems.

2. The CDN cost model

Because origin server requirements are assumed to be the same for a CDN and a P2P system, cost differences reflect differences in the cost of procuring edge storage and processing capacity for the two types of networks, and that is what we focus on when comparing costs. To simplify the CDN cost model (but with no loss in generality), we assume each edge server has the same cost and storage capacity as the origin server. New servers are added when customer demand exceeds the capacity of existing servers. Therefore, total edge storage cost for a CDN is a step function that increases by the cost of a new server at intervals equal to a server’s capacity.

Denote by $f(q)$ the total storage capacity cost of a CDN serving $q$ customers and let $C_s$ be the cost of a server that has the capacity to serve a maximum of $Q_s$ customers. Equation (1) and Figure 4 describe the total server capacity cost function (hereafter total cost function) for a CDN system.

$$f(q) = C_s \left\lfloor \frac{q}{Q_s} \right\rfloor,$$  \hspace{1cm} (1)
where $\left\lfloor \frac{q}{Q_s} \right\rfloor$ is a ceiling function that returns the smallest integer not less than $\frac{q}{Q_s}$.

**Figure 4** Total cost of storage capacity for a CDN

3. The P2P cost model

Potential contributing peers are all individuals with home computers and broadband connections to the internet. While contributing peers in P2P music file sharing services contribute music they acquire for their own enjoyment, there is no reason in principle that contributing peers have to consume the same content they store for redistribution. There is also no reason that peers cannot be compensated for the capacity they supply. For example, under its original owners the Kazaa music filesharing service rewarded contributing peers with various perks and in-kind services and CNN.com apparently rewards contributing peers with a higher quality video feed (Chartier, 2009). One could also imagine the development of market institutions that would utilize cash payments as incentives to supply peer capacity.
When peers are compensated for their capacity contributions, it is appropriate to treat broadband subscribers’ roles as content consumers and storage capacity suppliers separately, which is the approach used here. We assume that each contributing peer supplies a fixed amount of capacity, $q_s$, that is small relative to $Q_s$. As long as potential peer suppliers differ in the minimum compensation or reward they require to contribute processing capacity to a P2P service, the supply curve for peer capacity will rise from left to right. We define the reservation price of a potential contributing peer as the minimum that individual is willing to accept to contribute $q_s$ to a P2P system’s storage capacity.

Let $r(q_c)$ be the price that elicits total peer capacity of $q_c$. Equation (2) and Figure 5 describe a linear peer capacity supply schedule with slope $b$ and minimum reservation supply price of $a$. (As long as $q_s$ is very small relative to $Q_s$, it is reasonable to treat the supply of peer capacity as continuous while modeling CDN capacity as added in discrete lumps.)

$$r(q_c) = a + bq_c. \quad (2)$$
Assuming a P2P service cannot price discriminate among suppliers of peer capacity, so each contributing peer is paid the same price as the marginal peer, \( g(q_c) \) is the total cost of acquiring \( q_c \) units of peer capacity.

\[
g(q_c) = r(q_c) q_c. \tag{3}
\]

\( g(q_c) \) is convex as illustrated in Figure 6 for a reservation price schedule for which the first peer is willing to contribute capacity for free. The fact that the total cost of acquiring edge capacity by adding peers is convex while the cost of adding edge capacity by installing more edge servers is a stepwise linear function is critical to our comparison of costs for the two types of video distribution networks and our analysis of the merits of a hybrid network.

Equation (3’) is the formula for \( g(q_c) \) corresponding to the linear peer capacity supply function described by equation (2). This representation of \( g(q_c) \) will be employed in the numerical example presented below.
4. Hybrid networks

The fact that the amount of capacity contributed by a single peer is tiny compared to the capacity of a dedicated CDN edge server (i.e., \( q_s \) is small relative to \( Q_s \)) means that so long as the cost of acquiring the first peer is less than the cost of an edge server there will be situations for which content distribution costs will be lower with a hybrid CDN-P2P network than with either of the other two types of networks alone. This section identifies conditions for which each of the three types of networks is the most cost effective option. We do this in two steps: first determining the number of edge servers required to serve a fixed number of customers, \( q \), beyond the \( Q_s \) that could be served by a single origin server alone, and then solving for the cost-minimizing amount of CDN edge server capacity that could replaced by peer capacity.

We begin by calculating the number of edge servers required for a CDN system to serve \( Q_s + \bar{q} \) customers. First define \( m \) and \( h \) such that

\[
q Q_s = m + h Q_s, \quad (4)
\]

where \( m \) is a non-negative integer and \( 0 \leq h < Q_s \). If \( h > 0 \), a CDN system needs \( m + 1 \) edge servers to serve \( \bar{q} \) customers, but if \( h = 0 \), \( m \) edge servers are sufficient. \( m \) is the maximum number of edge servers the CDN can keep fully-utilized and \( h \) is the minimum number of the CDN’s customers that might be served by less than fully-utilized edge servers.

\[
\frac{\bar{q}}{Q_s} = m + \frac{h}{Q_s}, \quad (4)
\]

where \( m \) is a non-negative integer and \( 0 \leq h < Q_s \). If \( h > 0 \), a CDN system needs \( m + 1 \) edge servers to serve \( \bar{q} \) customers, but if \( h = 0 \), \( m \) edge servers are sufficient. \( m \) is the maximum number of edge servers the CDN can keep fully-utilized and \( h \) is the minimum number of the CDN’s customers that might be served by less than fully-utilized edge servers.

\[
\frac{C_s}{Q_s}
\]

is the cost per customer of edge servers if edge servers are used at full capacity. Define \( \bar{q} \) as the largest amount of peer-supplied capacity for which a straight line through the origin with slope \( \frac{C_s}{Q_s} \) intersects the total peer capacity cost function for a P2P distribution system, as shown in Figure 7. As long as the reservation price for peer capacity starts out sufficiently low and does not rise too
rapidly early on as new peers are added (i.e., \( g' < \frac{C_s}{Q_s} \)), \( g(q_c) \) must lie below the \( \frac{C_s}{Q_s} \) sloped line for some portion of its range and the convexity of \( g(q_c) \) guarantees that it will at some point intersect this line from below. If there is no such intersection, \( \dot{q} = 0 \). If \( \dot{q} > 0 \), \( \dot{q} \) is determined implicitly by equation (5).

\[
g(q) = \frac{C_s}{Q_s} \cdot \dot{q}.
\]  

(5)

For \( \dot{q} > 0 \) define \( z \) a nonnegative integer, and \( 0 \leq t < Q_s \), with values determined by equation (6).

\[
\frac{\dot{q}}{Q_s} = z + \frac{t}{Q_s}.
\]  

(6)

<Figure 7> Determining the mix of edge servers and peer-supplied capacity
At a minimum, a CDN serving $z$ or more customers can reduce costs by replacing $z$ edge servers with peer capacity. Let $x - zQ_s$ be the number of peer-served customers beyond $zQ_s$ required to raise the cost of peer-supplied capacity by an additional $C_s$, as illustrated in Figure 7 and define $s = x - zQ_s$. The value of $x$ is determined implicitly by equation (7).

$$g(x) - g(zQ_s) = C_s$$

What is the optimal mix of edge server and peer capacity?

- If $m \geq z$ and $h > s$, the network should use $m - z + 1$ edge servers and the ratio of edge server capacity to peer-supplied capacity will be $\frac{m+1}{z}$. This will be a pure CDN if $z = 0$.
- If $m \geq z$ and $h < s$, the network should use $m - z$ edge servers and the ratio of edge server capacity to peer-supplied capacity will be $\frac{mQ_s}{zQ_s + s}$.
- If $m < z$ and $h > s$, the network should use one edge server and the ratio of edge server capacity to peer-supplied capacity will be $\frac{1}{z}$.
- If $m < z$ and $h < s$, the network should be a pure P2P network with no edge servers and the ratio of edge server capacity to peer-supplied capacity will be zero.

These peer-supplied versus edge server capacity relationships and the conditions that cause them to vary are illustrated diagrammatically in Figure 8. Networks occupying the lower left quadrant are pure P2P networks while networks in the upper left quadrant have one edge server and the relative amount of network edge capacity supplied by peers increases as we move to the left. The supply of edge capacity is dominated increasingly by edge servers in both right hand side quadrants and networks in the upper right quadrant are pure CDNs when $z$ is zero.
If the cost of edge servers is fixed, then the optimal mix of edge server and peer-supplied capacity is determined by the steepness of the peer capacity supply function, because the steeper is $r(q_e)$ the smaller is $z$. More generally, factors that reduce the cost of edge servers relative to the cost of acquiring peer capacity increase $m$ relative to $z$ and push content suppliers toward greater reliance on edge server capacity in their internet delivery systems. Factors that reduce the cost of acquiring peer capacity relative to the cost of edge servers promote increased reliance on P2P delivery. These relationships are illustrated with a numerical example in the next subsection.

While technology costs will influence the costs of both peer and server-supplied capacity, peer supply is also influenced by the costs of transacting with large numbers of individual computer owners and peers’ perceptions of threats to the security of their personal information and the integrity of software residing in their PCs. The significance of this distinction is discussed in the final section of this article.

5. A Numerical Example

To illustrate how the cost-minimizing mix of peer-supplied and edge server capacity varies with changes in the relative costs of acquiring the two types of
capacity, we present a numerical example using the linear peer capacity supply function described in equation (2). Equation (3′) gives the corresponding functional form for the peer capacity cost function, \( g(q) \). Substituting from equation (3′) into equation (5), we have:

\[ a\bar{q} + bq^2 = \frac{C_s}{Q_s}. \]  

(5′)

Solving for \( \bar{q} \) yields

\[ \bar{q} = \frac{1}{b} \left( \frac{C_s}{Q_s} - a \right). \]  

(8)

Equation (7′), which is the version of equation (7) that corresponds to equation (5′), is constructed using equations (3′) and (5).

\[ g(x) = C_s + axQ_s + bz^2Q_s^2. \]  

(7′)

Using the quadratic formula to solve for \( x \) as the inverse of \( g(x) \) and taking the positive root because \( x \) cannot be negative, we get

\[ x = -\frac{a}{2b} + \sqrt{\frac{1}{b} \left( C_s + axQ_s + bz^2Q_s^2 + \frac{a^2}{4b} \right)}. \]  

(9)

This is the final formula needed to calculate the cost-minimizing mix of peer-supplied and edge server capacity for an online video distribution system.

\( b \), which determines the steepness of \( g(x) \), is the parameter that varies in this example. The other parameters are set at: \( a = 0.1 \), \( C_s = 100 \), \( \bar{q} = 130 \), and \( Q_s = 20 \). Given the values of \( C_s \) and \( Q_s \), \( m = 6 \) and \( h = 10 \). For these fixed parameter values, we plot the relation between the number of edge servers and \( b \) in Figure 9. The relative contributions of peers and edge servers to network capacity are plotted in
Figure 10. Both show that edge server capacity is substituted for peer-supplied capacity the more rapidly the cost of peer-capacity rises and, conversely, that reducing the relative cost of peer-supplied capacity increases reliance on peers for edge of network storage and processing.

**Figure 9** The relation between the number of edge servers and the slope of the peer capacity supply function

**Figure 10** The allocation of network edge capacity between peers and edge servers
V. SUMMARY AND DISCUSSION OF POLICY IMPLICATIONS

The analysis presented in this paper compared the cost of delivering internet video using Content Distribution Networks (CDNs) and Peer-to-Peer (P2P) networks, along with a hybrid combination of the two. P2P networks and CDNs are alternative remedies for cost and performance problems encountered with the basic client-server architecture as client demands increase. While CDNs currently dominate in delivery of internet video, P2P networks are also employed and CNN.com is an example of a service using the two approaches in tandem to deliver video over the internet.

We showed that the comparative advantages of CDNs and P2P networks depend critically on how the cost of purchasing edge servers for a CND compares to the cost to a P2P network of securing access to the mostly unused storage and processing capacity residing in internet users’ PCs. Should over-the-top distribution of internet video direct to television sets catch on with consumers, the processors in the peripheral devices that will be employed to make this possible for most television sets currently in viewers homes and in the emerging generation of internet-ready television sets could become additional sources of peer-supplied storage and processing capacity. Importantly, our analysis shows that in many cases a hybrid system will have lower costs than either a pure CDN or a pure P2P network, with the mix of server and peer-supplied capacity shifting with changes in the relative costs of the two types of capacity.

The owners of most PCs connected to the internet by broadband do not participate in P2P networks. The fact that much of the storage and processing capacity represented by these many millions of PCs is unused much, if not most of the time, raises the possibility that innovations in law and policy might free up this enormous unused resource for socially productive uses. In this regard, it is worth noting it is not the cost of technology that limits potential peer’s participation in P2P networks. Many millions of PCs that might be incorporated in these networks are already sitting in users’ homes and offices. Rather, in addition to the costs to
PC owners and peer network operators of finding each other and contracting for network access to consumer PCs, participation in P2P networks is discouraged by threats, perceived and real, to the security of data, often of a personal nature, residing in personal computers. To date policy interest in internet security has focused largely on limiting damage to individuals and organizations caused by unauthorized access to data stored on their computers. Here we point out that because data and privacy risks discourage participation in P2P networks, network security is also a communications infrastructure issue that deserves attention. In addition, policy makers should look for ways to make it easier for suppliers of peer-based distribution services to find and contract with potential suppliers of peer capacity.

We close by pointing out that as video places increasing demands on networks, the debate on compensation between players will gain new momentum, especially if asymmetric traffic flows exceed thresholds established in existing peering agreements between ISPs. The present model can be extended by incorporating communications costs explicitly to take these emerging issues into account.

REFERENCES


http://www.afterdawn.com/news/article.cfm/2010/12/04/global_iplayer_will_be_subscription-based_ipad-only