

Peer-to-Peer Traffic: Another Internet Myth is Born

Net Forecasts – Peter J. Sevcik

BCR Volume 35, Number 11

November 2005

I witnessed an urban legend come to life at the NGN conference in September. Dave Passmore used his traditional opening presentation for the conference to describe important trends that were going to be discussed. Two of the seven trends were video as the next killer application and customers competing with their service providers. Both of these trends relied heavily on a significant fact Dave stated, which was that peer-to-peer traffic is now greater than 60 percent of all Internet traffic.

According to Dave, video will most likely come to the Internet as shared DVD files (legal or not) using peer-to-peer systems over broadband Internet access, and this puts consumers in competition with their broadband suppliers that want to charge for video-on-demand over the same high speed access lines. Furthermore, he stated that since peer-to-peer has many variants which try hard to stay undetected, then peer-to-peer is surely already inside corporate networks.

The conference later featured Bram Cohen, creator and author of BitTorrent, as a keynote speaker, and had a session on the critical impact of peer-to-peer applications on networks. I heard many discussions about peer-to-peer in the halls and over lunch every day of the conference. All of them would start with “since peer-to-peer is 60 percent of all Internet traffic, then....” In the conference’s wrap-up session, the panel mentioned the 60 percent peer-to-peer fact at least 10 times during the hour.

I was skeptical, so I thought it would be interesting to find out how much peer-to-peer traffic really exists on the Internet.

Starting With the Source

The source of the 60 percent claim is Andrew Parker, co-founder and CTO of CacheLogic, a vendor of traffic measurement technology targeted towards the ISP market. It is clear that they are trying to educate their market to the potential impacts of sudden significant traffic growth from unplanned sources. An ISP sizes its infrastructure with a set of traffic assumptions, given its

subscriber base. If those assumptions are wrong and the ISP must handle a lot more traffic than planned, then the ISP has a problem.

CacheLogic has placed its StreamSight analysis boxes within several Tier 1 and Tier 2 ISPs. StreamSight performs protocol-based peer-to-peer recognition using deep-packet inspection at Layer 7. One of their claims is that since peer-to-peer is so hard to find, you have to look for it with a special tool. Interestingly, CacheLogic provides the boxes free of charge in order to build an ecosystem of participating ISPs. The ISPs get to see new traffic reports, and CacheLogic uses the insights to channel its R&D investments into future products.

I have had an informative email dialogue with Andrew Parker since NGN. In fact, during our exchanges, he restated the situation in stronger terms, saying that “60-80 percent of the traffic on an ISP’s network is P2P. We see higher levels in Asia with demand for U.S./UK TV programs via things like BitTorrent.”

I believe that CacheLogic has, in fact, seen 60-80 percent peer-to-peer traffic on some of the ISPs that they are monitoring. It appears that these are primarily carrier access networks of consumer broadband services. Just the kind of place where teenagers are exchanging files.

Furthermore, most of the peer-to-peer traffic observed by CacheLogic is within Asia. One of the reasons that peer-to-peer file exchange can become a significant ratio of offered traffic is that it operates machine-to-machine, with little human intervention once it is initiated. The CacheFlow charts show that peer-to-peer is relatively steady traffic, while other interactions observe the typical peaks and valleys of human activity during the day. Thus the peer-to-peer traffic is naturally more pronounced over a 24-hour period.

Other Data Sources

Clearly, the next step in this brief investigation is to find other sources of data regarding the volume of

peer-to-peer traffic. I pursued an informal non-scientific poll of contacts in the business. The story is very mixed.

Yes, some consumer-oriented broadband ISPs do report that nearly 60 percent of their traffic is peer-to-peer. Again, these are primarily in Asia. One Asian university is experiencing 80 percent peer-to-peer.

The steady symmetrical volume of peer-to-peer is most significant relative to the asymmetric low traffic generated on the uplink of most Internet applications like Web, email and file downloads. This explains why most of the peer-to-peer traffic "impact" is reported on uplinks to the Internet. Furthermore, these very high numbers mostly come from file sharing peer-to-peer traffic, rather than the equally popular Skype service, which generates less traffic per machine.

Carriers are getting serious about this by deploying measurement and control technology from a variety of vendors. So far, CacheLogic has been helping its competitors in this space sell proven products.

Andrew Odlyzko, director of the Digital Technology Center at the University of Minnesota, reported on a 2004 study conducted by David Plonka of the University of Wisconsin/Madison; Plonka's study looked at UW/Madison's access links to the Internet. Odlyzko presented the data in a presentation called "The Paradoxes of Broadband," delivered at the Optical Fiber Conference in February 2004 (www.dtc.umn.edu/~odlyzko/talks/index.html).

My review of the data in Odlyzko's presentation indicates that peer-to-peer represented 10 percent to 20 percent of the traffic handled. This is the same year that CacheLogic started to see 60 percent peer-to-peer traffic. So a different but equally likely venue for students to be running file-sharing software has a much smaller percentage than the CacheFlow result.

On the other end of the carrier spectrum is Internap, a Tier 1 ISP that peers with all of the largest ISPs. They primarily serve e-commerce sites, business-to-business commerce portals, and the financial

industry. Internap states that it sees less than 5 percent P2P traffic.

The observations of carriers from lowest to highest peer-to-peer percentage are an order of magnitude different! All anecdotal sources of information from enterprises show that enterprises are very vigilant about tracking, finding, and eliminating peer-to-peer traffic. They are motivated to do this for two reasons: Keeping their networks open for business traffic and avoiding copyright suits.

It appears that there is enough conflicting data and confusion over how to measure and report the data that no general conclusions should be drawn. Furthermore, the Internet is a very big place, so observations from a few locations cannot be generalized to represent the total Internet. We simply do not know enough.

Don't Base Decisions on Uncorroborated Data

But simple uncertainty is not the end of this story. Many people at the NGN conference were discussing lots of business implications of the 60 percent peer-to-peer "fact." And the real problem with pumping uncorroborated data up into a myth is that it leads to bad business decisions that hurt people.

The first uncorroborated assertion was that large amounts of peer-to-peer have flowed into enterprise networks like floodwaters that simply can't be held back. No one at CacheLogic, NGN, or my small poll sample could provide any evidence of high amounts of peer-to-peer in enterprises. But people were already talking at NGN about how to cope, if it can be controlled at all, the implications for mission critical applications, etc. It sounded like the sky was about to fall.

Second, there was much discussion of how this will kill the cable operators. The thesis is that cable companies sell broadband Internet access as a loss leader in order to get the consumer to buy expensive television programming or video-on-demand over the traditional cable plant. The fear was that cable customers will leverage the cheap Internet access to get movies for free and not buy any programming content from the operator.

Fortunately, I sat at a lunch table where a cable operator provided an informal reply to his foretold demise. He pointed out that people are inherently impatient, and claimed the IPTV content takes a long time to load and is of inferior quality to the cable broadcast. Cable companies plan to stay ahead with better speed and quality.

I must add my observation that I think there will be two markets for TV: The personal fill-the-time show and the group sitting in front of the big screen. The first is for all those times when you are waiting in a line, on a plane or bus, etc. This is best served by the Apple iPod video and its competitors.

The second is served by the high definition monitor with surround sound. Neither of these modes will ever be used in the other setting. You don't take your plasma TV on the bus and you don't invite your friends to a Super Bowl party to watch your iPod. There will always be a difference between the bus seat and the couch.

The cable company can enable and profit from both viewing modes. For example, the cable company can put a file sharing server on the access network to participate in the file sharing program, which will reduce the traffic impact to their network while transparently improving the user experience. They can make legal peer-to-peer copying faster than other means.

For this model to succeed, however, customers must accept the fundamental premises: That lower-quality, highly portable content can be freely obtained and exchanged, but higher-quality content and viewing experiences come with a cost. Carriers should join the movement and help control the outcome rather than fighting it.

Renewed Call for Good Impartial Information

I am sure that there are entrepreneurs giving pitches to VCs, and business executives at large companies already working the angles, on how to make money on the traffic that is at 60 percent of the Internet and growing. Why, at the current growth rate peer-to-peer will be the Internet in a few years!

Do not make the same mistakes of the past by jumping on a bandwagon that has no wheels. I have been here before.

My first BCR column, published in January 1999, was entitled, "The Myth of Internet Growth" (see pp. 12-14). In that column, I challenged WorldCom's John Sidgmore, who was the most vocal proponent of the myth that Internet traffic was doubling every 3 months. I revisited the topic and restated the inconsistencies of the myth in my January 2001 column (see pp. 10-11). The WorldCom bankruptcy filing came in July 2002. End of myth.

The fact is that WorldCom did see Internet traffic double in one three-month period during the early days of the Internet boom. It was astonishing. So astonishing that it had good shock value. But WorldCom decided to milk it as disinformation for a variety of self-serving strategies.

I called for a transparency regarding facts about the Internet in November 2001, in a column entitled "Federal Responsibilities for the Internet" (see pp. 10-12), and again in September 2002 (see pp. 8-9). I think that the federal government should operate a clearinghouse of information regarding the use of the Internet and digital media.

This kind of craziness comes around every few years. Back in 1995, well before the traffic growth myth, Martin Rimm at Carnegie Mellon University was touting a study proving that 83.5 percent of all content on the Internet was pornography. It took hold. Time magazine ran a cover story, a frenzy ensued, and even the U.S. Congress took action over nothing more than a student's calculation. I guess we were due for another good myth.

We do not need another round of misguided investments and bad business decisions based upon an unsubstantiated myth. But who will step up to the task of arbitrating the truth? I again call for an impartial organization to gather detailed traffic information from many reliable sources and integrate the data into a sanitized but accurate view of real trends in the Internet economy.

Companies Mentioned

BitTorrent (www.bittorrent.com)

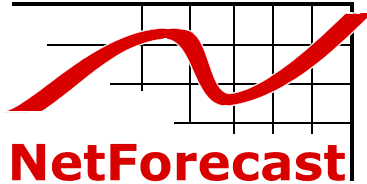
CacheLogic (www.cachelogic.com)

Internap (www.internap.com)

Peter Sevcik is president of NetForecast and is a leading authority on Internet traffic, performance

and technology. Peter has contributed to the design of more than 100 networks, including the Internet, and holds the patent on application response-time prediction. He can be reached at peter@netforecast.com.

NetForecast helps change delivery systems to improve the performance of networked applications. This includes advising enterprises on how to evaluate, improve and manage the performance of business applications, as well as advising vendors about customer requirements, technology issues, and adoption trends.



Smart Strategies from Hard Data

www.netforecast.com