Evaluation of paper: An Analysis of Internet Content Delivery Systems
Familiarity: Some knowledge

Recommendation: Strongly Accept

Background and context

This paper was published in a Usenix conference - 5th Symposium on Operating Systems Design and Implementation. This is a highly respected single track systems conference with an acceptance ratio of around 20% when this paper was published in 2002.

This was published a year after the 1st generation peer to peer application, Napster was shut down for helping distribute copyrighted material and a year after the birth of the 2nd and 3rd generation of peer to peer clients emerged like Gnutella and Kazaa which made use of a more distributed architecture making it harder to shut down. The Akamai content distribution system had been running for about 3 years when this paper was published.

Overview

The goal of the paper was to characterize incoming and outgoing Internet traffic at the University of Washington (UW) and quantify the impact of new content delivery systems (CDNs) such as peer-to-peer networks and Akamai. It also sought to predict the benefit of using caching for these systems to decrease the load on the University Internet connection.

Up until this point, there had been some published work on the benefits of CDN's in reducing average download response times but no study had been done of the work-loads of CDNs. The key startling conclusion was that peer-to-peer traffic accounted for almost three quarters of the the HTTP traffic but because of the large object sizes was being generated by a small number of P2P users. For example 200 users accounted for 27% of all HTTP bytes received by UW. It was shown that the University is a net provider rather than consumer of HTTP data despite its large client population, exporting 16.65TB and only importing 3.44TB in a 9 day measurement window. This was partly understandable as UW was a large publisher of web documents but the largest contributor to this was the volume of outbound P2P traffic, for example there was 1.78TB of inbound Kazaa traffic as opposed to 13.57TB of outbound data. This spurred a discussion on the need for a reverse P2P cache at the University's ISP which could save the University over 120 megabits per second.

Detailed discussion

The methodology to collect data was to use passive network monitoring on a dedicated monitoring port of a switch which connected two border routers handling incoming and outgoing traffic to the Internet. Traffic types were broken up into HTTP and non-HTTP traffic TCP traffic. HTTP traffic was broken up into Akamai, Gnutella, Kazaa and Akamai traffic by looking at port numbers and servers being accessed in the case of Akamai traffic. The capture and analysis software was able to capture P2P downloads by extracting HTTP headers and other metadata from flows.

At a high level the trace data revealed that Akamai only consumed 0.2% of observed TCP traffic, Gnutella consumed 6.04% followed by WWW traffic at 14.3% and then the largest contributor , Kazaa, consumed 36.9%. Other TCP-based network protocols such as streaming media, news and mail etc. consumed the remaining 43% with TCP traffic representing 97% of all network traffic. As far as content types, the main consumers were GIF and JPEG (16.3% of bytes), MPG and AVI (29.3%), HTML (14.6%) and hashed data from Kazaa transfers that couldn't be identified (9.9%).

The majority of requests, through, were for GIF images (31.8%) and HTML pages (17.8%). There was a marked increase (400%) in AVI/MPG traffic and MP3 traffic (300%) from a 1999 study.

Analysing the objects themselves, it was discovered that the median object size for P2P systems was 4MB which is 3 orders of magnitude higher than the average web document size. For Akamai and P2P traffic, a relatively small number of objects account for a large portion of the bytes transferred. For example the top 1000 Kazaa objects were responsible for 50% of bytes transferred. For Akamai the top 34 objects already account for 20% of the bytes transferred. The top bandwidth consuming objects for Kazaa were mostly video files in the order of 700MB consumed by a few clients. For Akamai, there are a few large unpopular objects (8 out of the top 10) but the worst offenders (2 out of top 3) were small and popular.

Peer-to-peer systems claim that one of the advantages is distributing the load uniformly across its peers but the results from this study show that of the 3888 Kazaa peers that were serving data, 170 were serving 50% of the data. A heavy-tailed (Zipf) distribution was evident for the WWW which is logical as the curve would rise steeply for naturally popular servers and then level off. In the case of a peer-to-peer system where only a file is requested which may exists on any number of peers, the expectation is for this curve to be much flatter. Part of the reason for this heavy-tailed behaviour in Kazaa is that less than 20% of the P2P servers result in a successful transaction suggesting that P2P servers are often saturated. The result is that a small number of bandwidth capable servers are taking the majority of the burden.

The paper closes with a discussion on caching for CDN and P2P networks using the same 9-day workload. The Akamai traffic achieved a 50% practical hit rate as opposed to a 36% hit rate by WWW requests. The practical hit rate takes into account objects which are not cachable such as HTTP headers with no-cache pragmas and suffixes naming dynamic objects. Akamai requests are skewed towards more popular documents (normally static image and video) compared to WWW requests which makes it ideal to cache locally. They conclude that a local web proxy cache could achieve nearly the same hit rate as an Akamai replica and eliminate the need for a separate content delivery network. Simulating a cache for Kazaa traffic revealed that the outbound hit rate stabilizes at 85% over the 9 day window and would result in a peak bandwidth saving of over 120 megabits per second if deployed in the University's ISP.

Paper analysis

Overall the paper exposed the key consumers of Internet traffic at this particular location, further investigation would be necessary to extrapolate these results to the general developed world or regions such as North America or Eastern Europe. The paper focused on P2P traffic as Akamai and others made little impact on the Campus Network. Drilling down into the P2P data and categorizing the type of content, the size of objects and the number of users within the P2P traffic streams was a useful contribution. This helped expose that a small set of users was transferring large data objects consuming a disproportionately high fraction of bandwidth. The disproportionally high amount of data being exported by the University and the suggestion of mitigation using a reverse P2P cache was also a useful contribution.

One of the weaknesses in the paper was the lack of analysis of the user base e.g. break down of students and faculty or possible artefacts that may be present in the data due to the specific 9 day window that was used between May 28th and June 6th 2002. For example students may download more P2P data during weekends and holiday periods or during periods where there are no mid-term or final exams. 7 or 14 days may be a better period of as it includes a balance of 5 day weeks and weekends.

Another weakness is the assumption that the small set of P2P users that consumed the disproportionately high fraction of bandwidth were always the same set of users. Over a 1 year period a larger total set of users than noticed here may be consuming bandwidth in bursts of obsessive P2P usage.

No discussion was given on how University policy could effect the large amount of P2P traffic on the border routers. There are many reports that show that a large portion of P2P traffic is illegal content. The policy rules could specify that using University infrastructure to download or store unauthorized copyrighted material could lead to criminal prosecution or expulsion. Perhaps the Digital Millennium Copyright Act (DMCA) was not being fully enforced at that stage and lawyers where still busy suing large companies like Napster rather than focus on Universities and individuals like students.

The paper would have had a lot more impact if a set of traffic traces were recorded at Universities and ISP's around the country to at least get a sense of traffic trends in North America. A worldwide study would obviously have had the strongest impact but would require significant resources probably beyond the scope of a University department.

The paper organisation was fairly logical but the amount of data was a little overwhelming. For example section 4 discussed too many numbers for the reader to come away with a clear sense of what they all mean. The graphics were often difficult to interpret in black and white – but this was probably because the specific conference allowed for full colour graphics. The summary at the end section 5 tended to overlap with the conclusions at the end of the paper and was largely redundant.