Characterization of a campus Internet workload

Adam H. Villa
Department of Computer Science
University of New Hampshire
Durham, NH 03824

Elizabeth Varki
Department of Computer Science
University of New Hampshire
Durham, NH 03824

November 27, 2011

Abstract

The campus network is a microcosm of the entire Internet, containing various types of users utilizing a variety of Internet applications. This paper characterizes the campus Internet workload for a mid-size university consisting of over 10,000 users. The paper characterizes bandwidth usage, top bandwidth consuming applications and users' Internet usage patterns on our campus network. The total amount of data transferred between campus and the Internet for an entire year, as well as the data amounts for individual application classes are presented. The paper also distinguishes the significant differences between students and faculty usage patterns. Over the past few years, there has been a major shift in application usage on campus networks. Real-time and bandwidth intensive applications dominate the workload of academic networks. The workload characterization in this paper identifies the degree of impact that these applications have on system resources and user workloads.

1 Introduction

This paper presents the first recent characterization of Internet workloads for a campus network that supports over 10,000 users. Until now, there have been no other papers that describe specific bandwidth usage rates, data amounts or top bandwidth consuming applications for campus networks. Our workload characterization will be beneficial to system administrators, application developers and researchers.

Campus networks are a microcosm of the Internet. The university campus is the workplace for researchers, faculty, staff, and students. Unlike commercial networks, the campus is also a home for the majority of the student body. The campus network must therefore support both academic and non-academic workloads in order to keep all users on campus content. The system must support a wide variety of application classes, such as: email, web browsing, streaming multimedia, gaming, video conferencing, voice over IP, cloud/grid workloads and file transfers. Each of these application classes has its own demands and requirements for bandwidth. In this paper, we characterize bandwidth utilization rates, users' access patterns and data consumption amounts for these application classes on the campus network.

Over the past few years there has been a major shift in Internet applications used on campus networks. Users have progressed from low-bandwidth, best-effort applications to real-time and bandwidth intensive applications. One such application is streaming multimedia, which is capable of consuming extraordinary amounts of bandwidth [1, 2, 3, 4, 5, 6]. Each year the bandwidth utilization rates are increasing for these types of applications. Since these applications can dynamically adjust their output quality based on bandwidth availability, they have unbounded demand for Internet resources. Users continually want better quality and high-definition viewing, which places extreme strain on system resources, especially on campus networks. A workload characterization is needed to determine the degree of impact these types of applications have on the campus network and how users are using these applications.

We realize that our network data represents only one possible network configuration used by academic institutions. Obtaining the following detailed data about bandwidth usage and user workloads required several rounds of authorization and working with network administrators to access live, mission critical hardware devices. Attempting to obtain similar in-depth data from other institutions and corporations proved impossible due to security concerns and confidentiality issues. We realize that some of the specifics from our analysis might only relate to our network, but the trends that we observe are definitely present at universities throughout the country [7, 8, 9], as well enterprise networks [10] and the Internet in general [11, 12].

The rest of this paper is organized as follows: first, we explain the configuration of our campus network in Section 2. We identify bandwidth usage information in Section 3. We then present our workload characterization for the Internet applications used on campus in Section 4. Finally, we summarize our findings in Section 5.

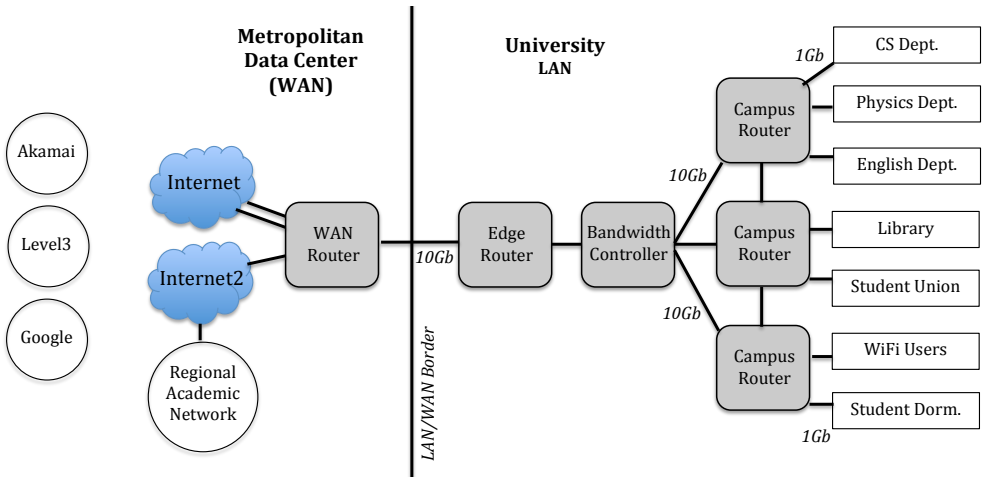

Figure 1: Network layout for the university network and its connection to the shared data center in a nearby metropolitan area.

2 Campus network

In this section, we describe the general setup and configuration of our campus network. Our network, illustrated in Figure 1, is designed to support over 10,000 students with an average of 6000 concurrent connections. Users consist of students, faculty and staff. These users connect to the network through Ethernet or WiFi connections and are distributed across multiple subnets around campus that are connected to the campus core via 1 Gbps links. The core of the campus network consists of 10 Gbps connections. The university utilizes three different wide area network connections for both Internet and Internet2 traffic, which have a total bandwidth capacity of 2.0 Gb/s for the entire university network.

Since our mid-size university supports over 10,000 users, the campus network has to ensure that each user has equal and fair access to the shared Internet connections. In order to accomplish this task, the university employs a bandwidth management device that is located at the edge or border of the LAN network. Each user device is limited to 8 Mb/s. As demand for bandwidth increases, the per device bandwidth allowance will be further restricted.

In the following sections, we characterize the Internet workload for the campus network. We examine the total amount of traffic flowing into and out of the shared Internet connections. We also examine the applications that are transferring the most amounts of data over the campus network. In order to gain access to this information, we utilize network monitoring devices that are placed throughout the network. We gather live data from the network and perform off-line data analysis of all traffic flows. We also use the bandwidth manager to gather data regarding users' workloads. Due to page limit constraints, only a portion of our findings are presented.

3 Bandwidth usage

We begin our characterization of campus Internet workloads by examining the total bandwidth usage of the shared Internet connections for the campus network. We monitor and examine bandwidth consumption on the campus network for an entire academic year. Figure 2 illustrates the variations in the daily maximum bandwidth consumption for this 12 month period. We find that there is very high demand during academic semesters and reduced demand during breaks. Since students are the main consumers of bandwidth on campus, changes in consumption correlate to their leaving and returning to campus. There is however a constant level of usage throughout the year regardless of the month, as the university hosts multiple government run projects that continually transfer data. Internal services that connect to satellite and regional campus networks also conduct data transfers on regular schedules.

Figure 2 demonstrates that several times during the Fall 2010 and Spring 2011 semesters the maximum bandwidth usage rates reached the bandwidth limits of the shared Internet connections for the entire campus network. Multiple times throughout the semester users consumed their entire bandwidth allotments and were forced to utilize less than their maximum rate of 8 Mbps.

Bandwidth demand changes throughout the year, as illustrated by the peaks and valleys on the graph. In order to understand of these shifts in demand, we examine the bandwidth usage from a weekly perspective. Figure 3A shows the maximum, average and minimum bandwidth usage for a typical week during the Spring 2011 semester. We find that network usage is the highest between Sunday evening and Friday afternoon. This correlates with classes starting and ending for a given a week. Between Friday night and Sunday afternoon, the network utilization is generally at its lowest. Even the maximum bandwidth rates during this pe-

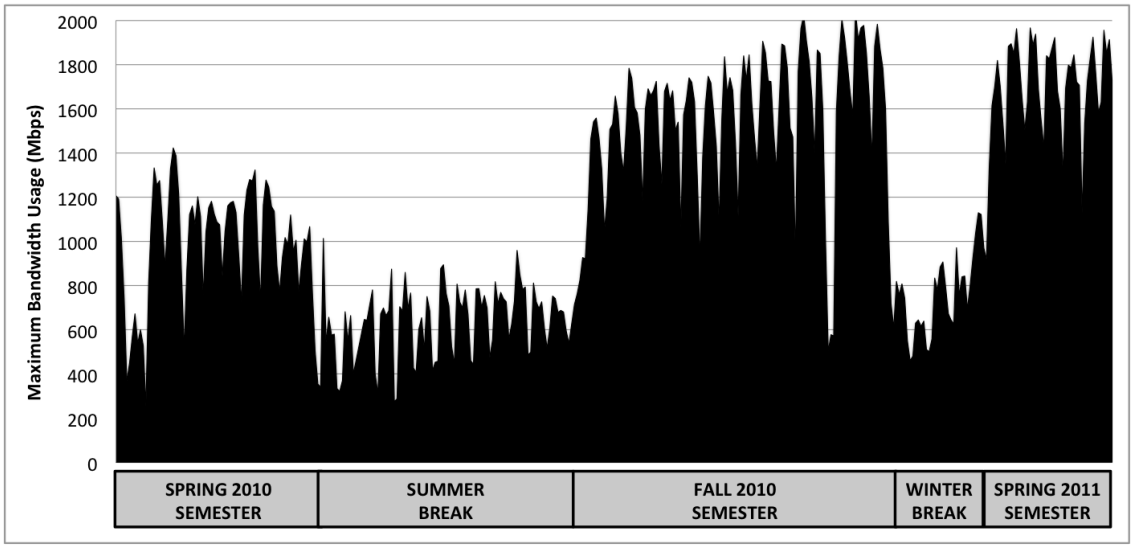

Figure 2: Changes in maximum bandwidth consumption for the past 12 months for all data passing through all of the university's shared Internet connections. Each semester user demand and bandwidth consumption increases.

riod are much lower than during the rest of the week. We attribute this occurrence to the fact that many students and staff leave campus or reduce their network usage on the weekends.

As we observe that the network utilization changes from day to day, we also find that it changes from hour to hour. In Figure 3B, we examine the maximum, average and minimum bandwidth usage for each hour in a typical day during the Spring 2011 semester. We find that peak usage occurs between noon and midnight. There is a slight dip around dinnertime and then usage increases until 1AM when demand starts to drop off. The lowest usage point occurs between 4 and 7 AM and then demand increases as faculty return to campus and students prepare for the start of classes. Throughout the 24-hour period, there is always some amount of bandwidth utilized as indicated by the minimum values on the graph. We observe very large differences between the minimum and maximum values, which indicates that users' workloads are dynamically adapting to changing bandwidth availability.

Overall, we find that a significant amount of data is transferred between the campus network and the Internet daily. On an average day during an academic semester, about 7 TB of data is transferred through the shared Internet connections. 5.5 TB of outgoing data is sent to the Internet and 2.5 TB of data is transferred into the campus network. The maximum amount of data ever transferred in a single day is roughly 10 TB. As in Figure 2, we also observe usage patterns that correlate to the academic calendar. More data is transferred during the Fall and Spring semesters than any other time. As previously discussed, there is a constant workload for the shared Internet connections and they are never completely idle. The minimum amount of data transferred on any day in the year is 870 GB, which occurred on Christmas day.

Bandwidth Summary: Overall, we find that most amount of data transferred between the campus network and the Internet occurs during academic semester in the Fall and Spring. There is a continual amount of traffic regardless to the time of year, which is created by special projects and internal services on campus. The peak usage time for the campus network is between noon and midnight from Sunday to Friday. We see decreased usage during the early morning hours (4AM to 10AM) and on the weekends. On a typical day the campus network is transferring roughly 7 TB of data to and from the Internet.

4 Internet application workloads

In the previous section, we characterize the amount of data being transferred to and from the Internet on the campus network. The next component of our characterization is to identify the applications that are transferring these large amounts of data.

Working with network management devices on campus, we are able to obtain usage profiles for users on campus. Using the bandwidth management device, which categorizes traffic by application using packet header and payload information, we gather user traffic statistics for a period of 35 days during the Spring 2011 semester. We examine the traffic workload to identify the top applications consuming Internet bandwidth during this time period. Figure 4A illustrates the applications that consume the most amount of bandwidth on a typical day for all users. We find that the applications utilizing the most amount of In-

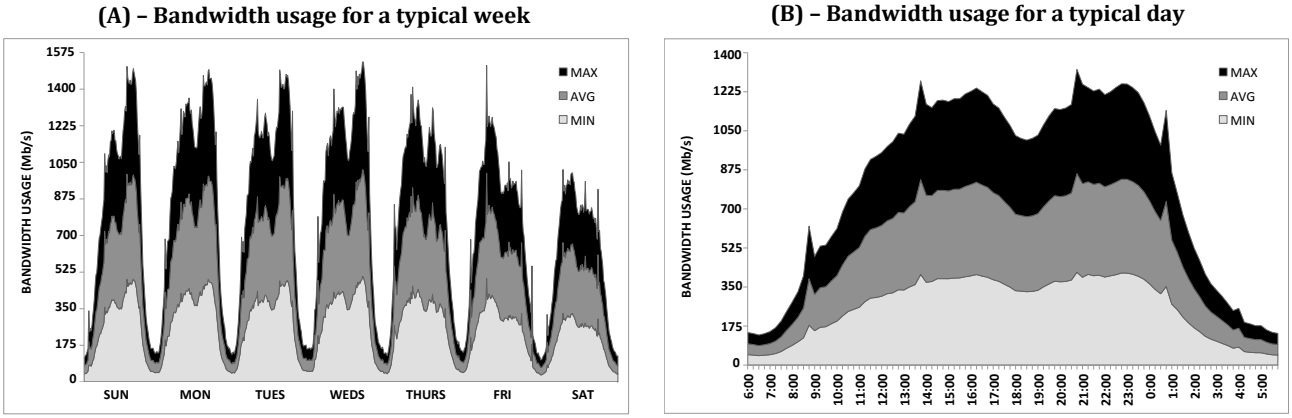

Figure 3: Changes in the minimum, average and maximum bandwidth usage (all receiving and transmitting traffic) for a typical week (A) and typical day (B) during the Spring 2011 semester.

ternet bandwidth are streaming multimedia applications, such as Netflix, HTTP Streaming and YouTube. On a typical day, these three application classes consume more than triple the bandwidth of general web browsing. This is the case on many campus networks, as well as the entire Internet [7, 8, 9]. Netflix currently consumes the most amount of bandwidth for the entire Internet [13, 14]. We also find that Skype and file transfers register in the top ten application classes utilized by users. Popular applications such as Facebook and iTunes rank in the top 15 user applications.

We continue our workload characterization by examining application usage by user type. We begin by comparing the usage patterns for students and faculty staff. In Figure 4B, we identify the top bandwidth consuming applications for faculty and staff users. We find that their workload is dominated by web browsing and file transfers. The applications with next highest levels of bandwidth consumption are streaming video and YouTube. Netflix is very low on the list of applications for the staff users. The bandwidth used by web browsing for the faculty is double that of any streaming application for their user group, very unlike the student users.

Figures 4C and 4D illustrate the top applications for the student users on campus. We separate the applications by daytime and nighttime usage. During the day, the applications utilized by the students are mainly streaming multimedia (NetFlix, YouTube, HTTP Streaming). At nighttime, the same streaming applications are still high in the list of applications consuming the most Internet bandwidth, however the bandwidth usage for these applications increases in the evening time. The major difference between daytime and nighttime periods is that Skype utilization increasing dramatically. Skype is the top application for bandwidth consumption during nighttime hours, with usage rates doubled in comparison to daytime hours.

Our user workload characterization also examines the changes in application usage based on the time of day. We have already compared student usage during the day to nighttime. We continue our characterization by looking at all users for specific hourly periods over the course of 24 hours. The top bandwidth consuming application, Netflix, is used to the greatest extent between 6PM and midnight. Netflix utilization is double during this time period in comparison to other parts of the day. Skype also has a significant increase in utilization during the evening time. Skype bandwidth consumption increases by 300% at night. Web browsing, YouTube viewing and HTTP streaming applications have the highest usage levels between noon and midnight. All applications see decreased usage between 6AM and noon. Skype and Netflix have the most noticeable decreases when compared to their peak periods. Web browsing is the only application to have usage levels during the 6AM to noon period that are comparable to normal daytime rates. The SSH application class has a fairly consistent level of usage regardless of the time of day. Many internal services (data backups and replicated data sets) utilize SSH for automatic file transfers throughout the day. The average daytime (6AM-6PM) rates is almost equal to the average nighttime rates (6PM-6AM) for the SSH application class.

In addition to characterizing the bandwidth usage rates for the application classes that make up the Internet workload on campus, we also identify the total amount of data being utilized by each application class. In Figure 5, we display the applications that received and transmitted the greatest amount of data between October 2010 and May 2011. Since this time period includes Winter break, the data essentially display usage information for six months. We identify the top five applications for both sending and receiving. We find that users utilizing the Netflix application were able to receive over 25,000 GB of data during the six month period. Both HTTP streaming and YouTube received over 44,000 GB combined. Web browsing and

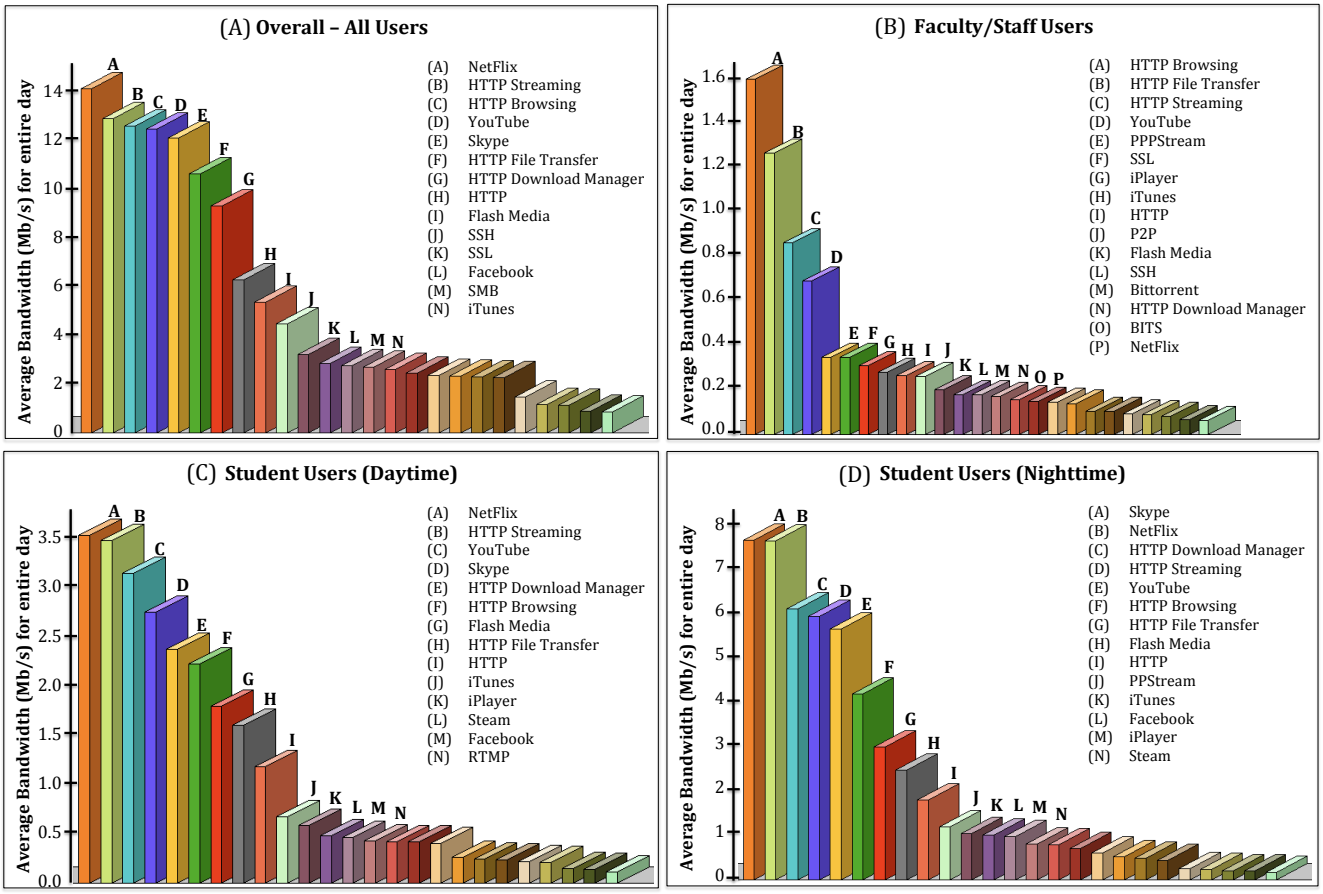

Figure 4: Most active protocols utilized on an average day. Protocol usage by types of users are shown: A) all users, B) faculty/staff users, C) student users during the daytime and D) student users during the nighttime.

HTTP file transfers each consumed roughly 15,000 GB of data individually.

We also examine the applications sending the most amount of data from campus to the Internet. The amount of data leaving the campus network for the Internet is considerably lower than the amount of data being received. Skype sent the largest amount of data during the six month time period, almost 10,000 GB. Both the sending and receiving amounts for Skype were almost identical. The next two applications that sent the largest quantities of data out of our network were secure communications (IPSEC-ESP and SSH). Each of these application classes transferred over 7000 GB of data out of the campus network. File transfers and web browsing also sent about 6000 GB. Web browsing had a bandwidth usage ratio of 2:1. The amount of data being received by web browsing users was double that of the data being sent by the same users. A full table of the data amounts by application is displayed in Figure 6 at the end of the paper.

Application Summary: Overall, we find that the real-time, bandwidth-intensive applications dominate the Internet workload on campus. Users are utilizing interactive applications that are sensitive to changes in latency and network congestion. Netflix consumes the maximum bandwidth and receives the largest amount of data in comparison to all other applications on campus. Skype transmits the largest amount of data to the Internet. Web browsing and SSH communications have fairly stable usage patterns in comparison to other applications.

5 Summary

Our campus network supports over 10,000 users and allows each user's device to utilize up to 8 Mbps. During peak periods, the bandwidth limit per device decreases automatically based on demand. The entire campus network currently shares multiple connections to both the Internet and Internet2. Given this configuration, we examine the bandwidth and application usage for all users. The following points represent the main findings of our workload characterization.

- Internet demand varies throughout the academic year,

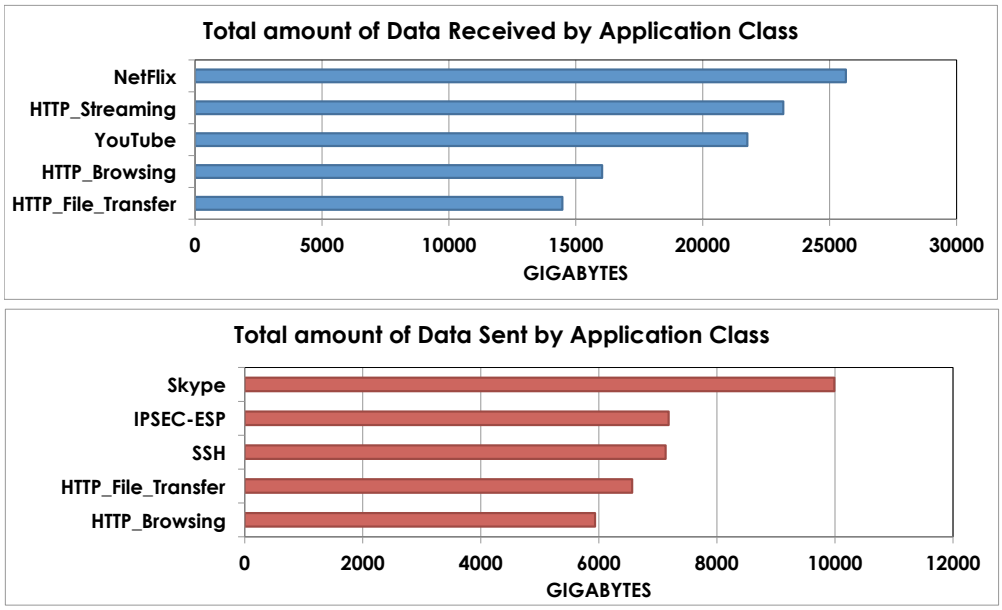

Figure 5: Total amount of data transferred by each application class between October 2010 and May 2011 for all users on campus.

however each semester more and more bandwidth is consumed by users as demand grows.

- User demand and bandwidth usage is greatest between Sunday evening and Friday afternoon. On average, bandwidth usage reaches a high-load condition between noon and midnight each day.
- On a typical academic day, the campus network transfers 7 TB of data. The network has frequently transferred up to 10 TB during peak periods. The minimum amount of data transferred on a given day is 0.8 TB.
- The applications consuming the most bandwidth on an average day are streaming multimedia (Netflix, YouTube), web browsing and Skype.
- The top applications for student users are Netflix, streaming web videos and Skype. Faculty and staff users' workloads are dominated by web browsing and file transfers.
- During the daytime hours (6AM-6PM), Internet traffic is mostly web browsing, file transfers, SSH and streaming multimedia. At nighttime, Skype, Netflix, YouTube and other streaming multimedia take over as the applications demanding the most bandwidth.
- During a six month period, the top applications transferred tens of thousands of gigabytes of data. Netflix (25,000 GB), HTTP streaming (23,000 GB), YouTube (21,000 GB) and web browsing (16,000 GB) had the largest amounts of received data. Skype sends the most amount of data on a given day (10,000 GB) and receives roughly the same amount of data.
- When bandwidth is increased, users quickly utilize any new capacity made available to them and the data transfer rates for the top bandwidth consuming applications greatly increase.

We recognize that these values and findings are specific to our campus network, however the same trends are found at many universities and networks around the world [7, 9]. Real time, bandwidth intensive applications have taken over the Internet as the most dominating workloads.

Currently, users are restricted to a combined bandwidth limit of 2 Gbps. Each year additional bandwidth is acquired by the university and the users' limits are increased. The university is scheduled to increase the overall bandwidth for campus and allow student users access to a 3 Gbps bandwidth partition. If the students' usage patterns remain the same, the campus network is estimated to transfer over 12 TB of data daily. Given that applications like Netflix dynamically adjust to available bandwidth and attempt to utilize as much as possible to achieve high definition viewing, we expect that the daily data consumption amount will be even higher [15, 16, 17, 18, 19, 20, 21, 22, 23].

Campus networks are important indicators of overall trends. Since these networks contain a large variety of users each using their own chosen Internet applications, we are able to gain insight into the workload of a typical

user. Commercial entities and Internet service providers do not publish information about the bandwidth usage rates and workloads of their users. Thus, it is difficult to determine general usage patterns for all users around the world. Our characterization of a campus Internet workload will therefore be valuable to administrators, developers and researchers.

The Internet was designed to be a best effort network and today's applications require guaranteed, real-time performance. We find that these kinds of applications dominate bandwidth consumption on the campus network. It will be necessary to further examine how the demands of these applications can be met given the best effort nature of the Internet and the increasing contention for shared bandwidth on campus. For future work, we intend to expand our campus workload characterization by examining user workloads from different viewpoints and compare usage rates as bandwidth limitations are modified.

References

[1] B. Ben Moshe, A. Dvir, and A. Solomon, "Analysis and optimization of live streaming for over the top video," in *IEEE CCNC*, 2011.

[2] M. Cha and et. al, "I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system," in *Internet Measurement Conference*, 2007.

[3] P. Gill, M. Arlitt, Z. Li, and A. Mahanti, "Youtube traffic characterization: a view from the edge," in *Internet Measurement Conference*, 2007.

[4] N.-F. Huang and et. al, "A novel bandwidth management scheme for video streaming service on public-shared network," in *IEEE ICC*, 2008.

[5] N. Hur and et. al, "3dtv broadcasting and distribution systems," *IEEE Transactions on Broadcasting*, 2011.

[6] M. Zink and et al., "Characteristics of youtube network traffic at a campus network - measurements, models, and implications," *Comput. Netw.*, vol. 53, pp. 501–514, March 2009. [Online]. Available: http://dx.doi.org/10.1016/j.comnet.2008.09.022

[7] M. McNierney, "College increases web speed for trial period," *The Dartmouth*, 2011. [Online]. Available: http://thedartmouth.com/2011/02/25/news/internet

[8] S. O'Malley, "Flood watch: Ohio's internet connection overflows," *http://www.ohio.edu/oit/news/ohio-internet-connection-overflows.cfm*, March 2011.

[9] J. Roettgers, "Ohio unversity blocks netflix, backpedals," *http://gigaom.com/video/ohio-university-blocks-netflix*, March 2011.

[10] R. Pang and et. al, "A first look at modern enterprise traffic," in *ACM SIGCOMM conference on Internet Measurement (IMC)*, 2005.

[11] C. Labovitz and et. al, "Internet inter-domain traffic," *SIGCOMM Comput. Commun. Rev.*, vol. 41, pp. 75–86, August 2010.

[12] G. Maier and et al., "On dominant characteristics of residential broadband internet traffic," in *ACM SIGCOMM conference on Internet measurement conference (IMC)*, 2009, pp. 90–102.

[13] R. Singel, "Most content online is now paid for, thanks to netflix," *Wired*, vol. http://www.cnn.com/2011/TECH/web/05/18/, May 2011.

[14] J. Yarrow, "Netflix is eating up more of north america's bandwidth than any other company," *Business Insider*, May 2011.

[15] V. K. Adhikari and et. al, "Youtube traffic dynamics and its interplay with a tier-1 isp: an isp perspective," in *Internet Measurement (IMC)*, 2010.

[16] L. De Cicco and et al., "Skype video responsiveness to bandwidth variations," in *NOSSDAV*, 2008.

[17] N. Hakken, "Netflix and akamai reports show sustained broadband speeds falter in u.s." [Online]. Available: http://broadbandbreakfast.com/2011/02

[18] N. Hunt, "Netflix lowers data usage by 2/3 for members in canada," Tech. Rep. http://blog.netflix.com/2011/03/netflix-lowers-data-usage-by-23-for.html, 2011.

[19] D. Levin and et. al, "Bittorrent is an auction: analyzing and improving bittorrent's incentives," *SIGCOMM Comput. Commun. Rev.*, 2008.

[20] J. Qi and et al., "Analyzing bittorrent traffic across large network," in *Cyberworlds, 2008 International Conference on*, 2008, pp. 759 –764.

[21] D. Rossi and et. al, "Evidences behind skype outage," in *IEEE ICC*, Piscataway, NJ, USA, 2009.

[22] D. Rossi and et al., "Understanding skype signaling," *Comput. Netw.*, vol. 53, pp. 130–140, February 2009.

[23] S. Sen and J. Wang, "Analyzing peer-to-peer traffic across large networks," *Networking, IEEE/ACM Transactions on*, vol. 12, no. 2, pp. 219 – 232, april 2004.

Protocol	Total Bandwidth (GB)	% Total Bandwidth	In Bandwidth (GB)	Out Bandwidth (GB)
NetFlix	26221.97	5.7	25640.919	581.052
HTTP_Streaming	24088.721	5.2	23172.675	916.045
YouTube	22511.961	4.9	21757.611	754.35
HTTP_Browsing	21971.461	4.8	16035.051	5936.41
HTTP_File_Transfer	21036.135	4.6	14472.353	6563.782
HTTP DownloadManager	13678.501	3	13108.437	570.064
Skype	19041.435	4.1	9052.676	9988.759
Flash Media	8758.3	1.9	8524.36	233.94
HTTP	9052.033	2	7681.611	1370.422
SSH	12884.821	2.8	5753.53	7131.291
iPlayer	4914.076	1.1	4642.528	271.548
iTunes	4890.349	1.1	4579.231	311.118
Facebook	4502.95	1	3866.792	636.158
SSL	7287.619	1.6	3552.22	3735.399
FTP-DATA	4662.748	1	3434.913	1227.834
PPStream	7643.39	1.7	2947.715	4695.676
Steam	2604.319	0.6	2541.917	62.402
RTMP	2640.165	0.6	2375.654	264.51
Other P2P	4473.183	1	2303.007	2170.176
YouTube-HD	1793.975	0.4	1754.968	39.007
BITS	1679.611	0.4	1640.144	39.466
MegaUpload	1439.513	0.3	1298.464	141.05
HTTPS	1170.22	0.3	990.882	179.338
BitTorrent Enc	2299.763	0.5	990.672	1309.091
Microsoft Live	1758.971	0.4	979.049	779.923
STUN	1801.377	0.4	865.712	935.665
IPSEC-ESP	8019.213	1.7	836.415	7182.799
HTTP_Audio	968.473	0.2	834.739	133.734
BitTorrent	977.689	0.2	496.192	481.497
SMB	3814.965	0.8	170.655	3644.309
All Others	211854.784	45.7	176674.118	35180.667
TOTAL	**460442.691**	**100**	**362975.21**	**97467.482**

Figure 6: This table lists the total amount of data transferred by application class between October 2010 and May 2011 for all users on campus.