Uses Supported by Higher Education Computer Networks and an Analysis of Application Traffic

Mark Pisano
Seidenberg School of CSIS, Pace University, Pleasantville, New York
Mp34549w@pace.edu

***Abstract*—This paper shows, a summarized view of higher education networks, through data analysis what applications are utilizing the network by evaluating the unique application signatures in the network packets. This is then categorized with the intent to offer insight as to what the network is supporting. This paper brings current knowledge, offers insight, and confirms the expectations of network administrators.**

Keywords—network; higher education; application traffic; network; packet signature

I. Introduction

Why have universities spent so much keeping their network infrastructure so up to date? Is it to support some large demand that is being put onto them? Or is it simply a selling point used to attract students and faculty from all over? It could be as simple as planning for the future. A common thought is that these networks are built out to support all of the technology that is scattered across campuses that is used to conduct research. An additional view, which is equally as common, is that the students need to have access to these high end networks in order to study and prepare themselves for their academics. Is the reality that these expensive networks are being used very heavily as entertainment networks? A better understanding of what these networks are supporting needs to take place not just how much bandwidth is being used or how much wireless coverage there is.

II. Methodology

The report that will be created is made up of data that has been collected in monthly phases. At the end of each month a report was run. This report included the total bytes, bytes sent, bytes received, application category, app technology, and number of sessions for the top fifty used applications. The monthly data is run through a consolidation process leaving a single data set to work with.

The study is conducted using a firewall appliance. This is a piece of computer networking hardware that was specifically engineered for its purpose. It was engineered to be able to process large volumes of data packets. Not only do they have to handle large volumes of traffic but they need to be able to process all the traffic, by comparing the traffic to the configured rules, and pass it along without disturbing the flow and speed of the traffic. This is a capability that could not be handled by a software firewall like the ones installed on a desktop computer. The hardware resources alone limit the capabilities of these desktop firewalls. This is why the firewall appliances are used and why they are a great place to collect and analyze data.

Applications are determined by their unique signature. Each application has its own signature that is used to identify it by the firewall. Signatures are determined by a detailed analysis of the data packet and crafted around content that is unique and yet the same for all data packets generated by the applications. There is further discussion about the application signatures and their generation in this dissertation.

Many firewall appliance manufacturers maintain a database of application signatures. These database are updated on a frequent bases as new application traffic is discovered. In addition the manufacturer creating and maintaining the signature database, customers have the ability to create their own signatures and update their local copy of the signature database.

The top applications are then sorted and organized by the application categories, the summarized data is analyzed. Pivot tables have been compiled to help create a visual representation of the collected data. The report is then created showing the pivot tables and break down of the categories that make up the data.

III. Literature Review

When researching in the area around Higher Education, which includes colleges and universities, and how that industry uses technology, the current trend of discussion is toward the bandwidth usage of students in the residence halls along with what devices the students are trying to connect to the network. Higher Education is more concerned in the speeds, consumption, and devices on its network rather than evaluation of what applications are running across it and what that could mean.

A. ResNet

Since 2012 ACUTA has been the primary group to provide data surrounding the residential network at Higher Education institutes. Even at this early point ACUTA recognized that schools are responsible for providing a fast and robust network to the students living on campus. "Each fall, a new generation of technology-savvy residents arrives on campus with an expanding set of mobile devices, greater demand for pervasive wireless coverage, and complex needs for connectivity support.[1]" This growth puts schools in a very unique position as an Internet Service Provider ensuring an internet connection that is used for more than just academic use. "Colleges and universities today are among the largest Internet Service Providers (ISPs) in the nation.[1]" Being an ISP means that you must keep up with the customer demands for connectivity. Evaluating the individual studies will show the historical growth of bandwidth demand, wireless growth, and the increase in devices being put on the network.

In the 2012 study 249 universities or colleges where represented. In terms of bandwidth 42.5% of the schools shared an internet connection with the residence halls which

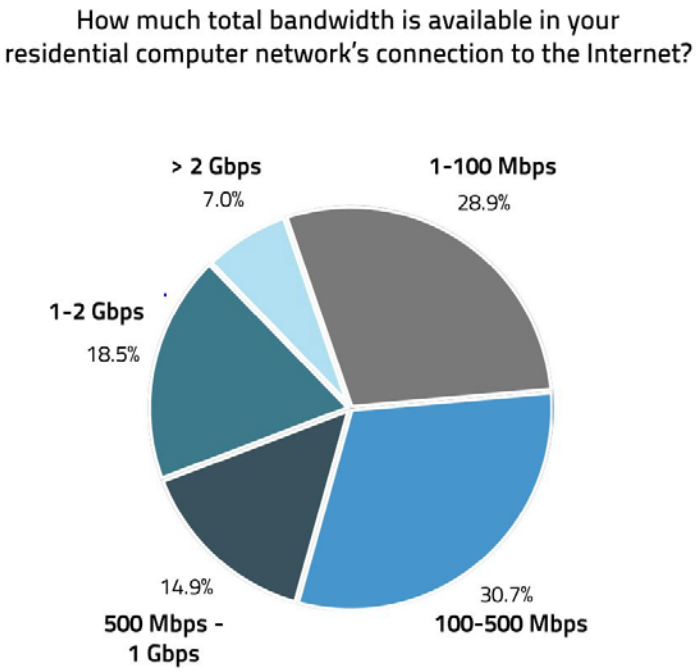

Fig. 1 - How Much Total Bandwidth is Available in Your Residential Computer Network's Connection to the Internet

was portioned out. However, a whopping 41.3% of the schools share an internet connection with the residence halls which was not capped. The remaining 8.7% of schools outsourced their residence hall networks. At this time the following chart shows the breakdown of allotted bandwidth to residence hall networks with the largest chunk being 100 to 500 Mbps.

In addition to just bandwidth, the study makes mention of the available wireless footprint in the residential locations at the schools. At this time, 56.4% of schools are offering 81-100% coverage. The breakdown of offering can be seen in the following figure.

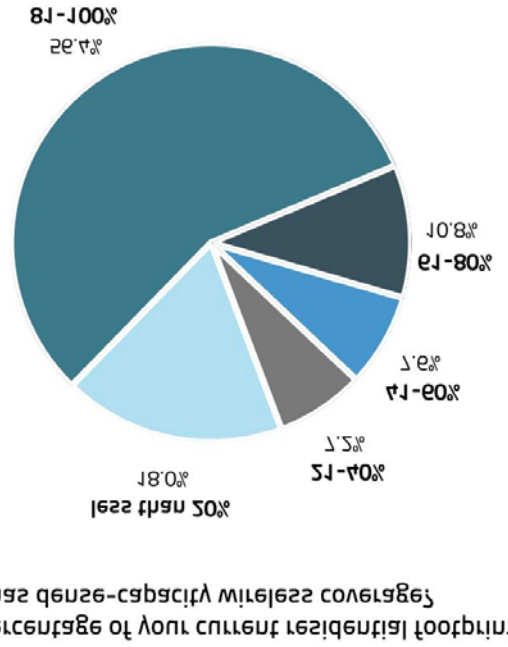

Fig. 2 – What Percentage of Your Current Residential Footprint has Dense-Capacity Wireless Coverage

The 2013 study represented 251 university and colleges. In terms of bandwidth this study does not breakdown how much of the overall schools bandwidth is shared with the residence halls or what might be outsourced. However it still gives us the results of what bandwidth is available. It is of no surprise that there has been an increase in the bandwidth offerings with the most schools still offering 100-500 Mbps, 36% of the schools. However, 1 GB or more is a close second with 33% of the schools offering the higher capacity.

By analyzing the State of ResNet Report for each year from 2012 through 2016 there has been a clear focus on the overall bandwidth made available to students in residence halls. There is a clear trend over the years to meet a growing demand for more bandwidth by the students to increase the amount available to them. When the study started there was minimal amount of bandwidth needed by the students as they only had one or two devices. As time moves forward the era of Bring Your Own Device has students increasing the number of devices they bring to campus who eat up bandwidth. This BYOD then shifts into the current state of Bring Your Own Everything which has the students demanding even more bandwidth. Not only do these studies show a drive for increased bandwidth but also an increasing demand for wireless access and a trend of schools growing their coverage to meet this demand. These reports offer some very significant facts to keep in mind.

- "Schools are bolstering their residential networks with more bandwidth for two reasons: To meet the BYOE (bring your own everything) demands

of the millennial residents and to support increasingly sophisticated educational applications and techniques.[2]"

- "More than 64% of institutions now offer 1 Gb or more of bandwidth per student – a more than two-fold increase from 25.5% in 2012. Twenty-one percent of campuses offer as much as 7Gb or more to accommodate student needs.[2]"
- "Eighty percent of colleges that have in-house Internet are implementing bandwidth management practices, such as shaping by protocol, compared with just 11% of those that have outsourced ResNet.[2]"
- "Today, 83% of campuses provide a robust strong wireless connection, almost doubling from 45% in 2013.[2]"

B. Bring Your Own Device

The use of bandwidth consumed in residence halls is not the only topic being discussed with regards to higher education computer networks. Just as import, is the concern for Bring Your Own Device (BYOD) movement or the Bring Your Own Everything Era. These are terms used to describe the vast amount of connected devices that students are leveraging for entertainment, lifestyle, and education. Pearson publishes the "Student Mobile Devices Survey," one of a few publications that is trying to get a grasp on the overall volume of devices and how students might be using them. The first publication was is 2011 and for the purposes of being current years 2013, 2014, and 2015 have been evaluated. These studies evaluate both two and four degree schools with a sampling of about 1,200 students[3]. What is illustrated in the below diagram is that for the three past years of the study almost 90% of students have both a laptop and a smartphone. In addition to this we also see that in 2015 52% of students also own a tablet. This means that over half the student population have three devices that they would like to connect to the network with another 40% having two devices that they would like to connect.

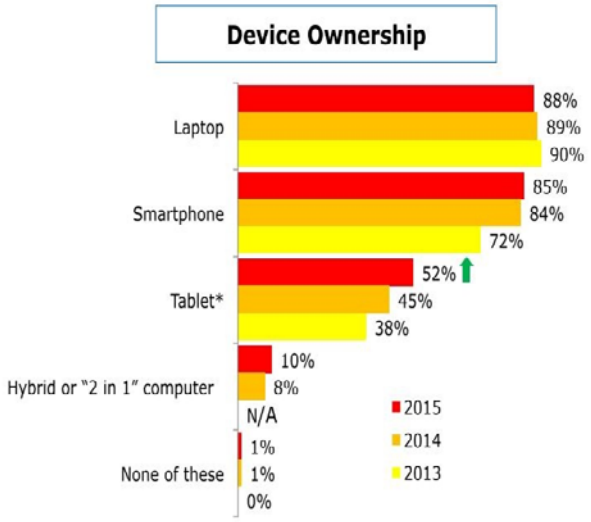

Fig. 3 – Device Ownership

IV. Data and Data Analysis

A. Data

The data is collected in monthly phases, this is due to the configuration of the device. The recommended best practice for the appliance is to only store logs for thirty days. At the end of each month a report was run. This report included categories which include the application, application category, application technology. It also included component data that is associated with the categories which includes bytes, bytes sent, bytes received and the number of sessions. This was all collected for the top fifty used applications. These applications are determined by packet signatures.

The packet signatures are unique to each application. For example, traffic being exchanged with YouTube is uniquely identified and cataloged based off of the signature. The volume of bytes sent and received along with sessions to YouTube is then able to be recorded.

B. Analysis Process

The data that was exported from the firewall was placed into Microsoft Excel 2013 spreadsheets, then all of the spreadsheets are then consolidated down to a single spreadsheet.

Each month had a limit of the top fifty applications and the final consolidated list had a total of 140 different applications. This clearly shows that the top fifty applications every month changes frequently and drastically. The top applications are then sorted and organized by the application categories, the summarized data is analyzed. In addition to the summarized data each application category and application technology can be analyzed. Pivot table, charts, and graphs have been compiled to help create a visual representation of the collected data. This is very helpful for gaining insight, understanding, and overall examination of the collected and grouped data.

C. *Application Categories*

There are a total of six application categories being used to analyze the traffic. The following is a list of the application categories and a description of what types of applications are associated with the category.

- business-systems: This category is made up of applications that focus on cloud based business services, software or hardware update services, databases, and applications that provide overall support to business processes and operations.
- collaboration: Applications that make up the category are focused on communication and consist of social media, conferencing, messaging, and email.
- general-internet: These applications focus on http traffic, web browser based applications, and supportive applications.
- media: This category is made up of applications that are primarily used for entertainment. They include gamming, video streaming, and photo sharing or viewing.
- networking: These applications take on a supportive role to all of the other applications that process across the network while also maintaining and keeping the network communicating.
- unknown: There is not enough information available to help fully classify the traffic.

D. *Summarized Application Categories*

Before proceeding with any type of analysis, it is important to define what entertainment is. The author is defining entertainment applications as applications that are used to view video, images, gaming, and any supporting applications that might be needed to help operate the entertainment applications. For example a portion of the DNS usage would have to be associated with entertainment since DNS is needed to help resolve names for most video streaming applications.

This pivot table shows the summarization of the applications within a category. The categories offer a high level organized view to what the volume or use of applications are across the network. We can see that applications that fall within the networking category are used the most. This is followed by applications that fall into the media category. In addition, we see that other categories fall to about half of what the networking and media categories are consuming. This offers some insight showing that business-systems and collaboration tools are not used as often as one might think.

Row Labels	Sum of Bytes	Sum of Mega bytes	Sum of Giga bytes	Sum of Sessions	Percentage of Sessions	Percentage of Bytes
business-systems	1.25093E+12	1.E+06	1165	9826955	18%	15%
collaboration	3.85919E+11	4.E+05	359	2323740	4%	5%
general-internet	1.33389E+12	1.E+06	1242	8199522	15%	16%
media	2.38697E+12	2.E+06	2223	2436554	5%	29%
networking	2.83931E+12	3.E+06	2644	30623092	57%	34%
unknown	1.15381E+11	1.E+05	107	656171	1%	1%
Grand Total	8.31245E+12	8.E+06	7742	54066034	100%	100%

Table 1 – Summarized Application Categories

V. Observations and Discussion

What is the summarization of the categories table describing?

- Business-systems utilize almost the same amount of traffic as general-internet at 15.05%, this equates to about 1,165 GB.
- Business-systems are consuming traffic at almost an identical volume as general-internet usage and this is most likely due to the ever growing push to cloud hosted solutions.
- It comes as a surprise that collaboration applications are only moving 4.65% or about 360 GB worth of data. This category includes applications that are heavily used by students for social interaction such as snapchat, Facebook, and twitter.
- General-internet usage is a small 16.05% which is about 1,242 GB. This is about half of either networking or the media categories. That is no small gap in traffic consumption.
- Media is a close second consuming 28.72% or nearly 2,223 GB.
- Media applications although making up about 29% of the traffic are using very few sessions at only about 5% of them which is about 2,436,554.
- Networking applications are utilizing 34.16% of the traffic or approximately 2,644 GB.
- Networking applications use the most sessions out of all the categories with a whopping 56.64% or 30,623,092 with business-systems being the next

closest with just 18.18% of the sessions which equals 9,826,955.

The current major observation is that the largest use of the network involves supporting itself. This is discussed further in the Networking section. Closely behind the Networking applications are the Media applications. More discussion follows about what is being used in the Media section. However it would appear that the overall use of the network is to provide entertainment applications.

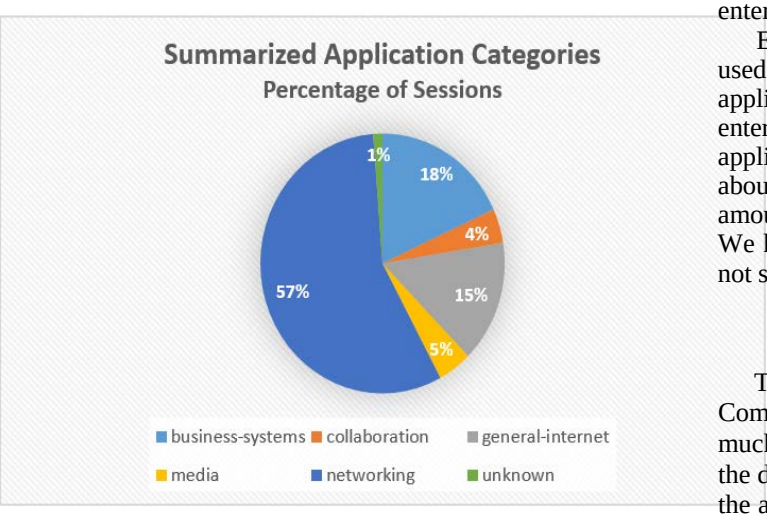

Fig. 4 – Summarized Application Categories Percentage of Sessions

As mentioned earlier the network application category consumes the most network resources and visually, in figure 16, we can see it consumes 57% of the used sessions. This would consume more sessions easily since a single session would be generated every time a DNS request is made. Overall, media applications would only open a single session and then consume more bandwidth over the life of the session.

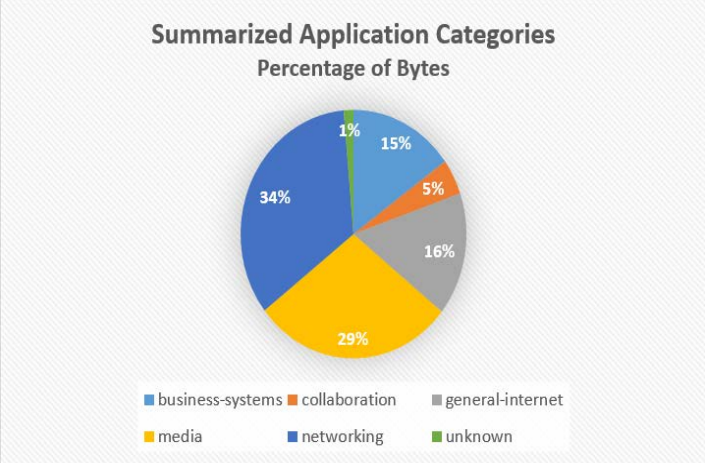

Fig. 5 – Summarized Application Categories Percentage of Bytes

Now when looking visually at the consumed bandwidth in figure 17, we see that the media application category is a very close second to networking application category. Now when factoring in that a portion of the networking applications are used to directly support the media applications and even portions of the general internet applications and collaboration applications are functioning as entertainment, we can see how the overall usage is really just supporting users' entertainment needs and demands.

It has been made clear that applications that make up the network application category are the largest users of network resources. Again, applications in the media application category are a close second. Due to the fact that network applications play a supportive role to all other applications, the overall usage of the network can be viewed as an entertainment network.

Entertainment has been defined as applications that are used to view video, images, gaming, and any supporting applications that might be needed to help operate the entertainment applications. Applications in the media application category without the supportive applications use about 29% of the overall bandwidth. This is a staggering amount, with Netflix and YouTube being at the top of the list. We have clearly seen that the academic computing network is not supporting much academics.

VI. Conclusion

The main focus has been to show what Higher Education Computing Networks are being used for. There has been much discussion in the bandwidth use in the residence halls, the deployment and coverage of wireless networks, and finally the amount of devices that are being brought onto campus and will need to have adequate network connectivity.

This research adds to the current and existing body of knowledge which focuses on Higher Education Technology by providing a summary and analysis of the application traffic traversing the network. Currently there is no snapshot or writing describing the applications that are being used across the network. It is the goal of this paper to provide such information.

A. Contributions

Through the use of pivot tables, figures, and tables this dissertation is able to produce a snapshot of application traffic that is being used across a universities computing network. It is this type of insight which has value and can be used for a number of purposes such as network planning, residence improvements, campus improvements, enhancement of the student experience, and much more.

By having the applications organized into categories, the author is able to illustrate the overall uses of the network. It is far easier to understand categories such as collaboration, media, and business systems than it is to take in and process all the individual applications.

This paper confirms, based on analysis of the categories that higher education computing networks are being heavily used as entertainment networks. It is the hope that others will have an interest and now an understanding of how to gather, collect, and analyze their own networks in an effort to gain their own insights.

B. Limitations

There are a few limitations to the study worth mentioning. These limitations are around the demographics and the collection the data. There may be other less significant limitations but it is the author's belief that the two following have merit worth mentioning.

The overall size of the demographic could be considered a limitation. The study was limited to one institution of a defined size. Universities and colleges come in different sizes. It is very possible that a larger institution would generate different results as different factions of the population could gravitate to different applications which would change the weight of their usage. This could also be said for smaller institutions.

A second limitation is the collection of data for the study, it has been limited to a period at the end of a spring semester and the beginning of a fall semester when students are actively living on campus and classes are in session, along with summer break when fewer students are on campus. It should also be considered that not all higher education institutions function under the dual semester model. There are many institutions that operated under trimesters or a year round model.

As with most research there are limitations. Even with the limitations of this dissertation it is believed that this research is able to offer a valuable result and insight into the usage of academic computer networks in higher education.

C. Future Work

As it has been mentioned, this paper had pointed out a few limitations of the study. Below are some areas of future research that are recommended to continue the work.

In order to compensate for the demographic limitation it is recommended to expand the study by including a large set of universities of varying sizes. This would allow for a more complex and diverse population of network users.

In addition, the amount of data analyzed could be expanded to fill the duration of a semester or full academic year. This would hopefully give a better overall picture of network usage and could be helpful when dealing with institutions that operate under trimesters or full year models. A comparison study as to when classes are in session and when universities are on break would also be an area of interest. This type of study could shed light onto the uses of the network when students are using it and when the network is being used for just administrative purposes.

REFERENCES

1 '2012 STATE OF RESNET REPORT ResNet Trends & Practices Across Higher Education', in Editor (Ed.)^(Eds.): 'Book 2012 STATE OF RESNET REPORT ResNet Trends & Practices Across Higher Education' (The Association of Information Communications Technology Professionals in Higher Education, 2012, edn.), pp.

2 '2016 STATE OF RESNET REPORT ResNet Trends & Practices Across Higher Education', in Editor (Ed.)^(Eds.): 'Book 2016 STATE OF RESNET REPORT ResNet Trends & Practices Across Higher Education' (he Association of Information Communications Technology Professionals in Higher Education
Association of College and University Housing Officers International, 2016, edn.), pp.

3 Poll, H.: 'Pearson Student Mobile Device Survey 2015 National Report: College Students', in Editor (Ed.)^(Eds.): 'Book Pearson Student Mobile Device Survey 2015 National Report: College Students' (PEARSON, 2015, edn.), pp.