Traffic pattern on different Internet technologies

Andersson, Jens A; Bonnedahl, Tomas; Höst, Stefan; Kihl, Maria; Aurelius, Andreas

Published in:
Proceedings of SNCNW+Adhoc 2009

2009

Traffic Patterns on Different Internet Access Technologies

Jens A Andersson, Tomas Bonnedahl, Stefan Höst and Maria Kihl
Dept. of Electrical and Information Technology
Lund University
Box 118, SE-221 00 Lund
Sweden
Email: jens_a.andersson@eit.lth.se

Andreas Aurelius
Acreo AB
Electrum 236
SE-164 40 Kista
Sweden

***Abstract*—The incentives for traffic measurements and traffic pattern analysis are increasing. New technologies for accessing the Internet like high speed packet data services in mobile networks are deployed, which imply in many cases that users share transfer channel resources. Applications and user behaviour from a wired environment may have a negative impact on these resources. This paper discusses the need for deeper knowledge of application and user behaviour in new access technologies and also suggests measuring and analysis methods.**

I. Introduction

Traffic measurement and traffic pattern analysis is growing more and more vital. In wired access networks we can expect that the increasing usage may lead to constraints. Utilization of high speed packet data services in mobile networks has grown from approximately 10% of total mobile network traffic in September 2006 to approximately 75% in July 2008. During the same time the total amount of mobile network traffic increased 10 times [1]. The way to deal with this in fixed wired environments up to now has been to "throw bandwidth on the problem". But we are getting closer to the point where this is no longer economically or technically possible, and in mobile networks it is not an applicable solution. Therefore other methods have to be deployed, which rely on thorough knowledge of the traffic and user and application behaviour. This is the aim of Traffic Measurements and Models in Multi-Service Networks (TRAMMS) project[1].

Internet usage a decade ago was mostly devoted to traditional WWW browsing. This has since shifted to triple-play, where households use their broadband access for TV, telephony and data including file sharing, online gaming and streaming video. The challenge this evolution will lead to is the design of IP access networks that can deliver services with strict Quality of Service (QoS) demands, *e.g.* IPTV, but also for mobile network techniques to take this user behaviour, as well as traffic pattern from individual applications, in consideration.

The increasing availability to Internet access with equivalent to wired capacity even in mobile networks raises the question how the radio network will cope with user traffic demands. It can be foreseen that new users of high speed packet data services simply move their computer from a wired or WLAN access network to a mobile network, and thus expect applications to behave in the same manner.

Most users have applications installed on their computers that load themselves in the background when the computer is booted. This is often done unknowingly to the user, and hence, if the applications use the Internet, it is possibly without the users intention and knowledge[2]. Examples of such applications are Peer-to-Peer (P2P) client applications or instant messaging client software like MSN messenger and Skype .

Even when not actively used, these background applications may still send and receive traffic. On a wired access network this has little or no impact, but a radio network will have to create and allocate a channel for that user in order to send and receive this background traffic. If only a small amount of packets are sent or received, it is unnecessary and a waste of resources to allocate more than needed for these packets[3]. It must be noted that user behaviour cannot be based on mere port identification. Several applications assign ports randomly, which calls for identification of data flows per application and establish current ports and protocols used for that application. To identify applications Deep Packet Inspection (DPI) must be used.

The fundamental difference between fixed Internet access and a connection to the Internet through the radio network is the sharing of resources with other users in the latter case. Latency, round trip time (RTT) and jitter are other characteristics that differ between the wired and mobile networks. Users affect each other in terms of bandwidth to a greater extent when connected using the radio network since the capacity in fixed Internet access is often very much higher. Interestingly, with higher capacity in upcoming DSL techniques comes also the fact that bundles of pairs can have to be considered as one common resource [2].

Bringing the fixed Internet access users to the mobile broadband world, the most important effect will probably be

[1]TRAMMS is a project within Celtic, a EUREKA cluster. Notable Swedish partners of the TRAMMS project include Acreo AB, Ericsson AB, Lund University and Procera Networks. http://projects.celtic-initiative.org/tramms/

[2]These applications are usually called daemons in UNIX and Linux environments and background services in Windows.

[3]Handheld mobile devices often lack these applications that run in the background and create network traffic that the user is unaware of.

seen when looking at traffic volume differences. Previously, almost the only device to connect to the mobile broadband was a mobile phone, often not capable of transferring large amounts of data. This will of course change when computers more and more are connected to Internet via mobile networks with the possibility to send and receive large amounts of data at high rates. Since several subscriptions to mobile broadband come with a flat rate, it is fair to assume that users will not be too concerned with bandwidth utilization.

The Internet Protocol [3] and the Transmission Control Protocol [4] were developed for a wired environment. TCP's Automatic Repeat Request (ARQ) is aimed at congestion control. The challenge in mobile network technologies deploying high speed packet datat services is therefore to adapt to this fact. Bit error in the physical and link layers must be avoided to impact on the congestion control mechanisms in higher layers. It is, as an example, important for the Long Term Evolution's (LTE) scheduling mechanism [5] to both adapt to and exploit the channel as well as variations of the traffic carried by the channel. This calls for well-founded knowledge of traffic patterns in this environment, as well as in application behaviour in constrained environments.

Different approaches have been used to perform traffic measurements. [6] investigates the implications of media downloading and video streaming which is increasing in popularity. The first IP packets in sessions with media content are captured at an aggregated level in the backbone and then analyzed.Research in the area of 3G radio networks include [7] which presents how the different channels in a Wideband Code Division Multiple Access (WCDMA) air interface operates internally.

This paper highlights incentives for new traffic analysing methods within this area for the use in fixed and mobile access networks. Some findings from traffic measurements [8] are discussed. We also introduce a methodology to measure a user's traffic and separate different application categories. For new access technologies where the physical layer channel is a shared resource, *e.g.* LTE and upcoming DSL technologies, it is of great importance to use this channel in the most efficient way. This includes not only control of the physical and link layers but also an adaption to traffic patterns and user and application behaviour.

In section II the networks and techniques used for the measurements are discribed. The analysis is presented in section III, where the distribution of applications per user is discussed in III-A and the clustering technique is presented, and section III-B compares packet inter-arrival times and packet size in the DSL and HSPA networks. In Section IV the work is concluded.

II. Measurement networks and measuring techniques

Data in this study is collected from three different access networks, based on three different access technologies, Fibre to the Home (FTTH), Digital Subscriber Line (DSL), and High Speed Packet Access (HSPA), respectively. Extensive traffic measurements on application level and packet level have been performed using the PacketLogic network appliance [9]. PacketLogic is a network router and firewall both used in networks operated by Small and Medium Enterprises (SME) and Internet service providers. It performs deep packet analysis and identifies applications. Some applications are descriminated into subapplications; Skype is an example of an application that is reported as several subapplications by the PacketLogic.

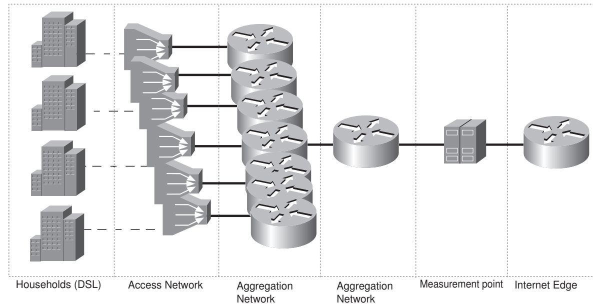

Fig. 1. Conceptual network diagram showing the DSL network.

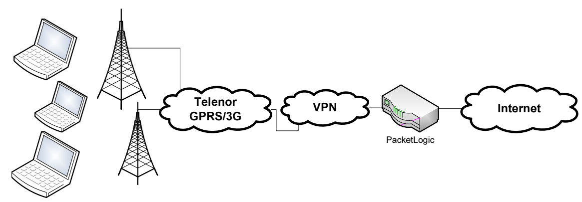

Fig. 2. Network diagram showing the HSPA network.

The measuring principle for both the FTTH and the DSL network is shown in Figure 1; All packets are fed to the PacketLogic device. In practise the deployment of the PacketLogic differ slightly between the two networks. The number of households on the FTTH and DSL networks are 2081 and 104, respectively. The households using DSL access consists of privately owned houses in a residential district, each with an asymmetric 24 Mbps connection. The FTTH households have a wide geographical spread over the area and represents different demographic groups and household constellations. Users in this latter network are given options on service provider and symmetrical connection speed, up to 100 Mbps.

The HSPA network in Figure 2 is more complex compared to the fixed networks. Obviously this network offers geographic freedom for the users. The USB modem will connect to the HSPA base station closest to the user's current position. The geographic placement of the measuring device, which is vital since we want the traffic to go through it, is therefore problematic. The solution is to give the users a special SIM card to use, which forces the user's traffic internally in the operator's network via a VPN tunnel to the PacketLogic which enable us to analyze the traffic. The number of HSPA users in tris study are 20.

III. Analysis and results

Two examples of traffic analysis are presented in this section. Clustering the end users into different groups is a

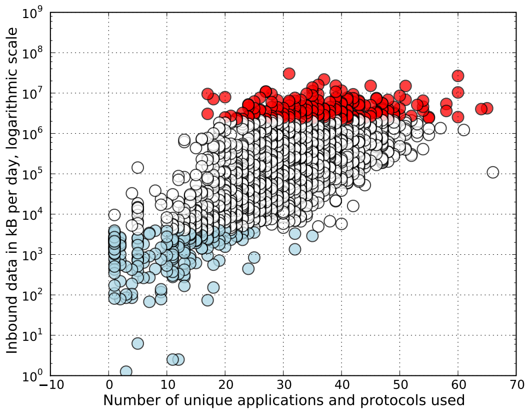

Fig. 3. Applications used versus consumed inbound bandwidth per household for the FTTH access network user group.

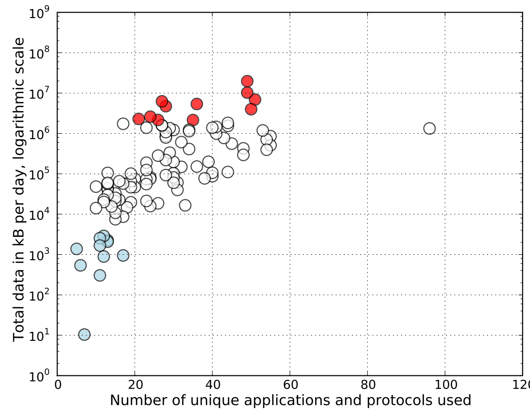

Fig. 5. Applications used versus consumed total bandwidth per household for the DSL access network user group.

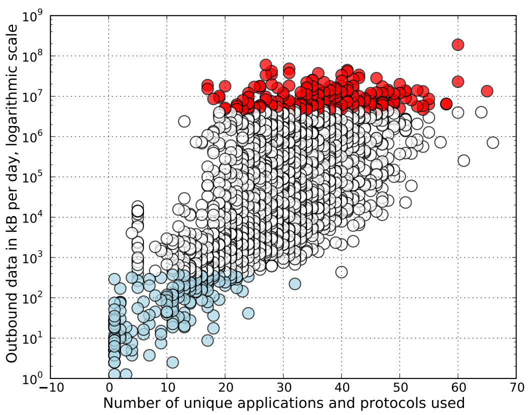

Fig. 4. Applications used versus consumed outbound bandwidth per household for the FTTH access network user group.

way to analyze user types as described in section III-A. In III-B Cumulative distribution function (CDF) graphs are used to present application behaviour on packet level. The data analysis presented in here is not final or complete; further and deeper analysis is necessary.

A. Applications per user

The unique number of applications per user, combined with the amount of data transferred, are the chosen parameters for the cluster analysis. The goal is to divide users into groups based on their habits. We identify three user categories. The top utilizers are relativley few but consume the majority of the aggregated capacity. The bulk of the users use the same amount of applications as the top utilizers but consume less bandwidth. The third user group use few applications and do not contribute to the capacity utilization in greater extent.

Figures 3 and 4 show plots of the number of applications as classified by the PacketLogic versus the consumed bandwidth for 2081 households with FTTH access. The corresponding plots for the 104 DSL customers are shown in Figure 5. In the plots users have been separated according to their network utilization. The limit for the 90th percentile in Figure 3 corresponds to 2GB per day, while the 10th percentile corresponds to 4MB. The corresponding limits for outbound FTTH traffic, Figure 4, are 4.3GB and 0.5MB, and for DSL traffic, Figure 5, 2GB and 5MB. The outbound traffic volume, Figure 3, in the FTTH network is higher than the inbound traffic volume, Figure 4. To a large extent, traffic dominating this is p2p file sharing from computers left on around the clock. No significant difference between inbound and outbound traffic were seen in the DSL case, thus the inbound and outbound traffic per user is not presented separately. Instead the total traffic per user is shown in Figure 4.

Comparing Figures 3 and 4 with Figure 5 it can be seen that users in the DSL network use fewer applications and protocols than the users in the FTTH network. It is also clear that there is a relation between the number of applications and the amount of data produced. This behaviour may not only be dependent on access technology, but can also reflect the demographic differences between the two networks. Comparing penetration of peer-to-peer applications in Sweden [8][10] and Spain [11] also supports the importance of demographic knowledge. In Sweden the dominating p2p application is BitTorrent, while e-donkey is the dominate in Spain.[4]

Figures 3, 4, and 5 also show that a few number of users contribute with the majority of the capacity utilization. This is also seen in [12], and it was shown practically already in 2000 when the p2p application Napster overloaded the Lund University campus network (LUNET) [13], and especially in 2003 when LUNET network management had to prohibit all p2p usage [14]. There are too few data points in the available measurements from the HSPA network to prove if this is true for HSPA. But since the radio channel is a shared resource this possible behaviour has an impact on the channel planning.

[4] [12] intimates a genus perspective on the user behaviour.

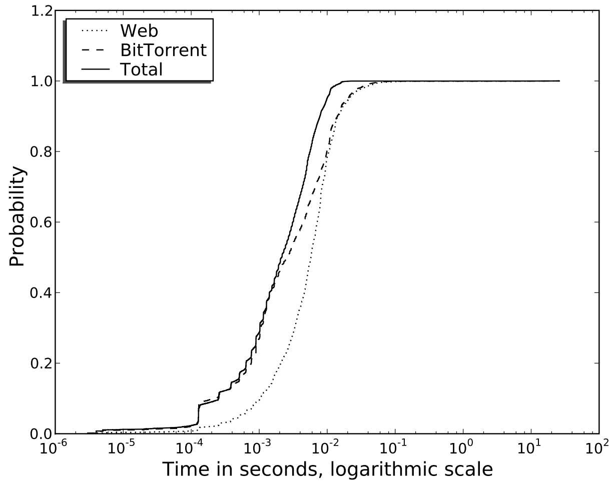

Fig. 6. Inter-arrival times of packets from one user in the DSL network.

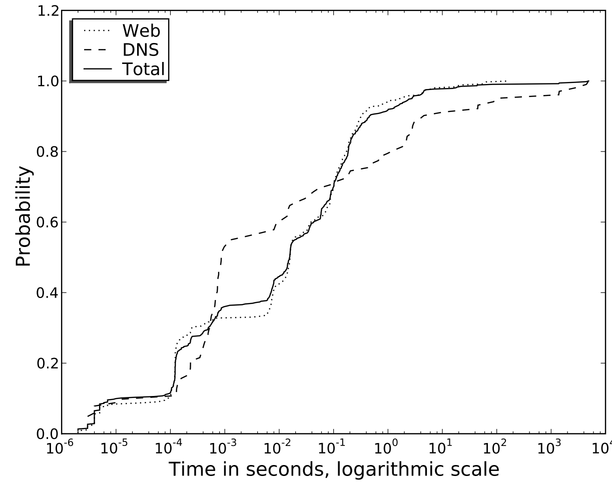

Fig. 8. Inter-arrival times of packets from one user in the HSPA network.

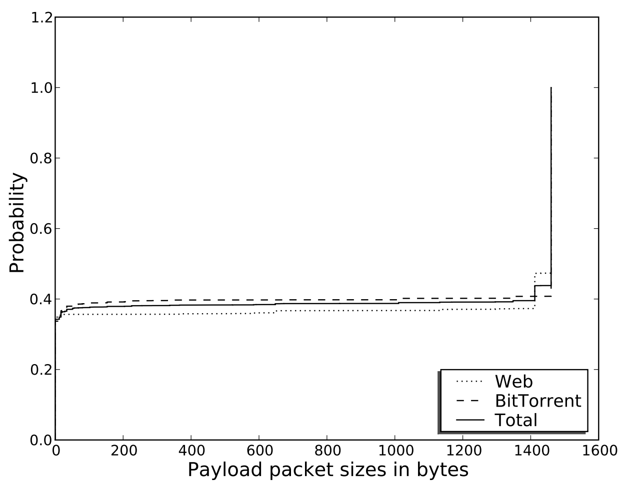

Fig. 7. Payload packet sizes of traffic from one user in the DSL network.

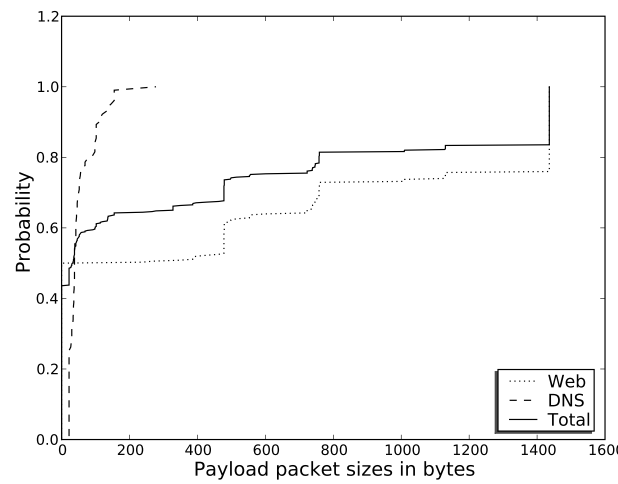

Fig. 9. Payload packet sizes of traffic from one user in the HSPA network.

B. Application behaviour on packet level

We have studied network traffic from one user in the DSL network and one in the HSPA network. Figures 6 – 9 show typical inter-arrival times and packet size for DSL and HSPA access. In the DSL case, Figures 6 and 7, inter-arrival times within in a well defined time interval are exponentially distributed. 40% of the packets have the minimum size and the rest have the maximum size. In the HSPA case however, Figures Figure 8 and Figure 9, the inter-arrival times are unevenly spread over a much wider time interval and the packet sizes are differentiated. Even if it is web applications that are examined in both cases we cannot be sure that the access method is the explanation behind the differences; Web applications are defined by the protocols HTTP and HTTP media stream. The more unreliable link in the HSPA case can be an explanation for the more uneven curves in the HSPA case. Since the exact situation the user was in when the data was collected is not known, we cannot be sure of how much of the user's own behaviour is influencing the measure.

IV. Conclusion

This paper presents incentives for evolvement and deployment of traffic measurement methods in access networks with different technologies. In a situation with shared resources it is of great importance to have deep knowledge of both user behaviour as well as the behaviour of applications and the traffic pattern these behaviours create. With the introduction of mobile broadband for computers, radio networks need to be better prepared for, amongst other, background traffic. Traffic measurements and traffic pattern analysis on the aggregated level is not sufficient for deployment in for example mobile networks; per user and per application analysis is needed for a full analysis and understanding. Detailed knowledge of the traffic patterns including, but not limited to, packet sizes and inter-arrival times is the key to better resource allocation and increased quality experience for end users in radio networks. The resource allocation in the radio interfaces of the mobile network make decisions for each user individually. Thus, when modelling and simulating traffic, it is necessary to recreate one user's actual traffic instead of aggregating over a large population. Differences in traffic pattern and application behaviour between different access network technologies can be identified, and we can describe them. The causes that explain the differences are yet to be analysed.

REFERENCES

[1] S. Parkvall: *Broadband Wireless Access - HSPA and LTE*, http://www.s3.kth.se/signal/edu/s3_seminar/2009/talks/sem2.pdf (last visited 2009-03-12)

[2] P. Ödling, T. Magesacher, M. Berg, E. A. Sanchez, S. Höst and P. Börjesson: *The Fourth Generation Broadband Concept*, IEEE Communications Magazine, Vol. 47, No. 1, pp. 63-69, IEEE Communication Society, 2009

[3] *RFC 791 Internet Protocol*, IETF, 1981

[4] *RFC 793 Transmission Control Protocol*, IETF, 1981

[5] E. Dahlman, S. Parkvall, J. Skold and P. Beming: *3g Evolution*, Elsevier Science & Technology, 2007

[6] L. Guo, E. Tan, S. Chen, Z. Xiao, O. Spatscheck and X. Zhang: *Delving into Internet Streaming Media Delivery: A Quality and Resource Utilization Perspective*, IMC, October 25-27, 2006, Rio de Janeiro, Brazil.

[7] T. Sung: *3rd Generation WCDMA / UMTS Wireless Network*, MC Lab, IE CUHK, 2003

[8] T. Bonnedahl: *Traffic Measurement and Analysis in Fixed and Mobile Broadband Access Networks*, Dept. of Electrical and Information Technology, Lund Universtity, Master Thesis, 2009

[9] *PacketLogic Products Overview*, Procera Networks, http://www.proceranetworks.com/products-overview.html (last visited 2009-04-20)

[10] *TRAMMS Deliverable D3.2 Traffic Characterization*, 2009

[11] *TRAMMS Deliverable D3.1 Traffic Characterization*, 2008

[12] H. Selg: *SUNET användarstudier: Användning av fildelningstjänster*, SUNET, 2003, http://basun.sunet.se/aktuellt/fildelning.pdf (last visited 2009-03-13)

[13] M. Persson: *Avstängning av Napstertrafik*, LDC, Lund Universtity, 2000, http://aron.ldc.lu.se/externwebb/security/meddelande-20001219.shtml (last visited 2009-03-10)

[14] M. Persson: *Trafikbelastning på studentnätet*, LDC, Lund University, 2003, http://aron.ldc.lu.se/externwebb/security/meddelande-20030124.shtml (last visited 2009-03-10)