LUND UNIVERSITY

How much of the bandwidth do we really use? An investigation of residential access traffic load

Aurelius, Andreas; Arvidsson, Åke; Heegard, Poul; Villa, Björn; Kihl, Maria; Zhang, Yichi

Published in:
[Host publication title missing]

2012

Link to publication

Citation for published version (APA):
Aurelius, A., Arvidsson, Å., Heegard, P., Villa, B., Kihl, M., & Zhang, Y. (2012). How much of the bandwidth do we really use? An investigation of residential access traffic load. In *[Host publication title missing]* IEEE - Institute of Electrical and Electronics Engineers Inc..

Total number of authors:
6

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

How Much of the Bandwidth Do We Actually Use? An Investigation of Residential Access Traffic Load

Andreas Aurelius[1], Åke Arvidsson[2], Poul Heegard[3], Bjørn Villa[3], Maria Kihl[4], Yichi Zhang[5]
[1]*Networking and Transmission Lab, Electrum 236, SE-164 40 Kista, Sweden*
Tel: +46 8 632 78 02, Fax: +46 8 632 77 10, e-mail: andreas.aurelius@acreo.se
[2]*Ericsson AB, SE-164 80 Stockholm*
[3]*Department of Telematics, Norwegian University of Science and Technology (NTNU), Trondheim, Norway*
[4]*Dept. of Electrical and Information Technology, Lund University, Sweden*
[5]*KTH – Royal Institute of Technology, Stockholm, Sweden*

ABSTRACT
Internet traffic from a fibre based residential access network is investigated concerning traffic volumes and link load. Also the cost of the services is analyzed. We show that 1 Mbps accesses subscribers maintain high loads, and that the price they pay per GB used is five times higher than the one paid by 100 Mbps access subscribers.
Keywords: traffic measurements, monitoring, access networks, internet traffic.

1. INTRODUCTION

The volume of global IP traffic in telecommunications networks is undisputedly growing at reasonably fast pace. Numerous reports show a yearly traffic increase. The common understanding is that the Internet traffic doubles each year or each 1.5 years, see *e.g.* Cisco forecasts and reports [1]. The reasons for traffic growth may be several: increased number of users, increased traffic per user, changed traffic patterns, etc. Recently, a 5 year study of residential IP traffic in a Swedish fibre-to-the-home (FTTH) network [2] showed that the increase over 5 years was as moderate as ~ 33% per IP address. Of course, there were major events that impacted the results, like new legislation hitting P2P traffic, but the results contribute to formulate an important question: Is the annual traffic increase distributed reasonably evenly between networks and access types, or are there major differences? One hypothesis is that when a network (or rather the use of the Internet service in that network) has reached a certain level of maturity, the usage will level out similar to, *e.g.*, the spread of consumer durables [3]. If this hypothesis would be true, services where the access speed has been sufficiently high for close to 10 years already would have a quite moderate growth rate, whereas the reasonably immature mobile broadband networks over 3G/4G would experience a rapid increase of data volumes [4][5]. This paper does not claim to test the hypothesis, but inspired by the indication of "network maturity", we will investigate how much of the available bandwidth that is actually used in a fibre based residential access network. Traffic load and Internet subscription costs are investigated, in order to analyze how much of the available bandwidth that is actually used by the customers.

The paper is organized in the following way: next section reviews relevant work. The following sections describe the networks and the measurements. All results are in the results section, and in the end, the conclusions can be found.

2. RELEVANT WORK

Several papers have investigated IP traffic on aggregated links, from various perspectives. Fraleigh [6] analyzed traffic in high speed links in the Sprint backbone, investigating *e.g.* traffic workload, packet and flow characteristics. Other studies analyzing traffic from large populations are Fukuda [7], who studied traffic in Japanese ISPs, and Maier [8] who studied traffic from 20.000 DSL lines in a European ISP.

IP Traffic analysis on a household level can be found in Kihl [9], who analyzed traffic characteristics in a Swedish fiber based access network, and in Aurelius [10], where traffic from Swedish and Spanish access networks was analyzed.

Several authors have analyzed workload of multimedia streaming traffic, *i.e.* [11]–[13], whereas http traffic was the focus of [14].

3. NETWORK

The data analyzed in this paper comes from a fibre based Swedish municipal access network (MN). This open network offers fibre based, high speed Internet access. Traffic measurements are performed by the network operator, and post processed, anonymous data is made available to the authors for analysis. The data comprises traffic from roughly 1400 customers, with access speeds varying from 1 to 100 megabits per second (Mbps). The measurement probe is placed at the Internet edge of the municipal network. This is the point where the local network operator hands over to the ISP, meaning that all customer traffic towards the Internet has to pass this point. The traffic that is not visible at this point is P2P traffic between peers in the local network.

978-1-4673-2229-4/12/$31.00 ©2012 IEEE

4. MEASUREMENTS

The measurements have been performed using PacketLogic (PL) [17], a commercial traffic management device, used in many commercial broadband access networks all over the world. Traffic is identified based on packet content (deep packet inspection and deep flow inspection) instead of port definitions. The device can identify more than 1000 Internet application protocols, and the signature database [18] is continuously updated.

The identification process is connection-oriented, which means that each established connection between two hosts is matched to a certain application protocol. When a new connection is established, the identification of this connection begins. The identification algorithm searches for specific patterns, signatures, in the connection. The patterns are found in the IP header and the application payload. The PL uses the traffic in both directions in the identification process. The measurement point in the network is the Internet edge, as described in the previous section.

The PL can track and identify several hundred thousand simultaneous connections, storing statistics in large databases. The statistics database records the short-time average amount of traffic in inbound and outbound directions as well as the total traffic for all nodes in the network. The data is averaged over 5 minute periods. The present study is based on data from the period 2010-09-01 – 2010-09-14.

5. RESULTS

When analyzing user behaviour, sub groups of users may be considered in order to reveal specific patterns. Sub groups may be based on geography, age, income, ethnicity, etc. One important sub group is subscription type (ST). This is important, since it is defining the relation to the Internet service provider (ISP). It also sets the limit for the traffic that can be transmitted, thus paramount for dimensioning of links. By combining traffic measurements with knowledge of subscription type, interesting facts may be discovered.

5.1 Bandwidth and price comparisons

In our measurements, we have analyzed traffic and cost per ST. Most STs are symmetric (1 Mbps, 5 Mbps, 10 Mbps, 30 Mbps, and 100 Mbps) but some are asymmetric (100/10 Mbps and 30/10 Mbps), with lower uplink capacity. There are several ISPs to choose from; hence the prices for the services differ slightly. We have chosen to use the average price per ST in our analysis, Table 1. The 10 Mbps ST has the highest number of users, and is shown in bold in the table.

Table 1. Average monthly price and number of subscribers per ST.

Bandwidth (Mbps)	1	5	10	30/10	30	100 /10
Average monthly fee (SEK)	149	189	232	249	284	318
Number of households	171	72	849	18	135	133
Active time, percent	30.5	37.3	41.8	53.7	58.1	56.6

Table 2. Monthly bandwidth consumption per subscription, in Gigabytes (GB).

Bandwidth (Mbps)	1	5	10	30/10	30	100/10
Monthly inbound	12.9	23.5	35.2	84.2	88.4	98.4
Monthly outbound	4.8	23.1	41.9	47.1	159.2	157.8
Price per GB, in	11.6	8.0	6.6	3.0	3.2	3.2

The most expensive subscription is almost 3 times the price compared to the cheapest one, so it can be assumed that price plays an important role in choosing subscription. The most obvious parameter to characterize usage with, is the consumed traffic volume. Both the total number of bytes transmitted and the periodic variations differ considerably between STs. The monthly consumption in bytes is shown in Table 2. Comparing the 1/1 Mbps with 100/100 Mbps, it can be observed that the transmitted volume in the downlink direction is 14 times larger in the 100/100 Mbps subscription while in the uplink direction, the transmitted volume is 90 times larger in the 100/100 Mbps case. Another interesting fact revealed in the table, is related to symmetry. Looking at the symmetric STs, only the 1 Mbps ST yields an asymmetry with more inbound than outbound traffic. The 5 Mbps ST yields a symmetric pattern, whereas the 10, 30 and 100 Mbps STs all are (increasingly) asymmetric in the other direction, *i.e.* with dominating uplink traffic. In the two asymmetric STs, the lowest capacity (30/10) is asymmetric and dominated by downlink traffic, while the highest capacity (100/10) is also asymmetric but dominated by the uplink traffic.

Having pointed out the bandwidth consumptions of the STs, we continue with analyzing the actual price paid for the service in relation to the transmitted volume. The willingness to pay for services is something that is of great interest for operators. In Table 2, it can be seen that 100 Mbps subscribers pay almost a fifth of the price as compared to the 1 Mbps subscribers (2.4 SEK per GB compared to 11.6) for the traffic in the downlink direction. Looking at the uplink direction instead, the corresponding numbers are 1.0 and 30.8 respectively. Even though the subscription is roughly 3 times the price, the cost per byte is thus very low in comparison. Looking at the other STs in table 2, a clear trend in price per bit can be observed. It is gradually decreasing as the subscribed bandwidth is increased.

5.2 Daily variations

Further, the daily variations of the traffic are analyzed in order to investigate the maximum load. The daily variations are shown in Fig. 1 (downlink) and Fig. 2 (uplink). The figures show relative load, *i.e.* traffic volumes normalized by the subscribed bandwidth, so that loads of different link can be compared. We have already shown that high speed STs yield significantly higher volumes, but looking at Figures 1 and 2, another pattern can be observed. For the downlink, (Fig. 1), it can be seen that the 1 Mbps ST achieves a significantly higher load than the 10 and 100/10 Mbps STs. For the uplink, (Fig. 2) however, the 100/10 Mbps ST achieves the highest load.

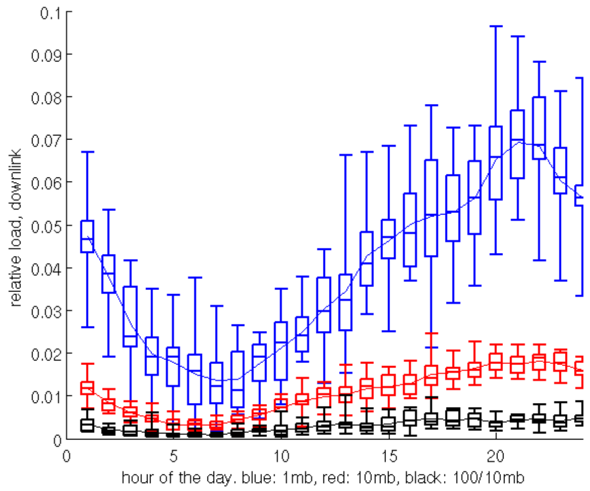

Figure 1. Average daily load per ST in the downlink direction for the 1 Mbps, 10 Mbps and 100/10 Mb STs.

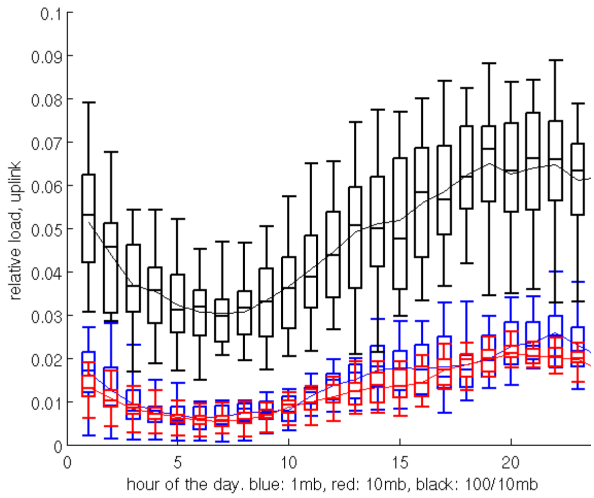

Figure 2. Average daily load per ST in the uplink direction for the 1 Mbps, 10 Mbps and 100/10 Mbsp STs.

The data in the figures is averaged over all days in the measurement period and over all households per ST. Thus, any peaks will not be revealed in the graph, apart from the typical evening peak hour of 20.00-22.00. We have computed the ratio between the maximum load for each ST and its average load (downlink) and all STs have a load in the range of 1.8-2.6. The lowest value belongs to the 10 Mbps ST (the ST with the largest population) and the highest value belongs to the 30/10 Mbps ST (the ST with the smallest population). The same numbers for the uplink traffic are 1.4 (100 Mbps ST) and 2.6 (30 Mbps ST). In order to analyze peaks, the traffic rate per active[1] 5 minute interval is computed for all STs and plotted as a survival function (1 - CDF) in Fig. 3. The graph confirms the information from fig. 1, that the 1 Mbps ST has high relative load. It is actually above its maximum rate[2] in 3% of the active periods. The customers of this ST actually have a load over 80% during 6.1% of their active time. Such a high load is not seen for other STs. The 5 Mbps ST customers are above 80% load during 1.5% of their active periods, which is significantly lower, but still noticeable. Also the 10 Mbps ST shows a portion of high load, 0.8% of the time over 80% load. The conclusion from this graph is that although the average load has been shown to be low, there are several occasions where the traffic peaks at its maximum rate, and that during a significant fraction of the time the load is at a level where there is a risk of service degradation. Also, we can conclude, by means of the active time numbers in table 1, that the 1 Mbps households have a lower average active time (30.5% of the time) than the 100/10 Mbps ST (56.5). In general, the higher the speed, the more active, in terms of time spent using the Internet.

[1] An active interval is an interval during which any traffic is sent. The percentages of such periods are shown in Table 1.
[2] The subscribed bandwidth is enforced by a shaping mechanism which allows for some peaks above the subscribed rate.

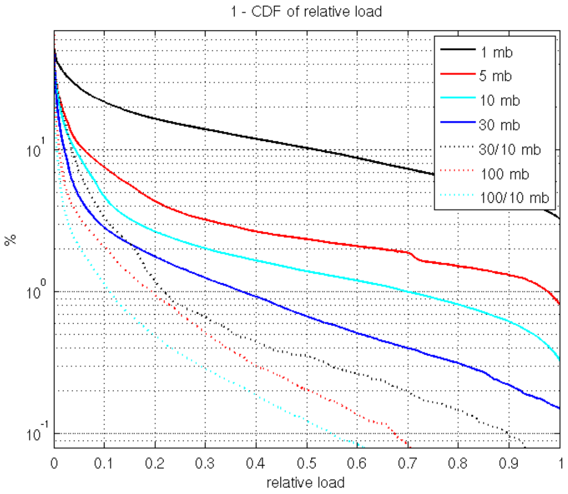

Figure 3. Survival function for relative load (in %) per subscriber type in the downlink direction.

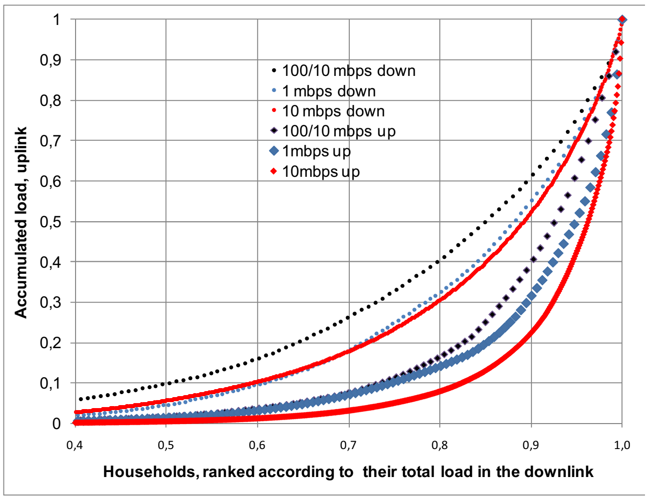

Figure 4. Load per household, ranked according to load, and plotted accumulated. We can see that 20% of the households in 1 Mbps subscriber type generate 70% of the traffic in the subscriber type.

5.3 Subscription types and cost

In Fig. 3, it was shown that households using the 1 Mbps ST on average create a load which exceeds 8% in the downlink direction during 6.1% of their active time. This is clearly an interesting fact, but it does not reveal the actual number of households that contribute to this average. When dealing with Internet traffic and usage, one should be aware of the fact that average numbers may be misleading, since the variations may be very large. In order to further analyze the household behaviour, we have computed the total load per household normalized by the total load of the ST. The households are ordered by their total load (ascending) and their accumulated load is plotted in Fig. 4 (1, 10 and 100/10 Mbps ST). The purpose is to test the validity of the so called 80/20 rule which states that, for many phenomena, roughly 80% of the effects come from 20% of the causes. First, it can be seen that there is a major difference between the downlink and uplink distribution of load. In the downlink case, users are more evenly distributed; the top 20% of the households generate 60% – 70% of the traffic. In the uplink, the curves show sharper "bends", meaning that the top households regarding load are more dominant. Here, the top 20% households generate roughly 82% – 92% of the traffic. We also note that in both uplink and downlink, the 10 Mbps ST is the most prominently "heavy" one on dominant users (the red lines in Fig. 4 are beneath the 1 and 100/10 Mbps ST lines).

So, although Fig. 1 showed that the downlink of 1 Mbps ST households on an average day peaks at around 7% load, there are clearly several households that are running on a high load during a substantial fraction of the active time. Among network providers, dimensioning is a fundamental issue, and a rule of thumb is that when the load is close to 50%, it is time to upgrade. If the residential users would follow the same rule of thumb, a substantial number of the households in the 1 Mbps ST should consider upgrading to a faster service. This is of course true only if the customers really expect their service to deliver a performance that can cater for applications that are sensitive to delay and packet loss, for example gaming, music and video streaming. For example, we have shown in [9], that these kinds of services are commonly used. Upgrading is of course also a cost issue. Upgrading from 1 Mbps to 5 Mbps would increase the monthly cost for the customer by 40 SEK or 27%. However, from the perspective of the ISP, there would be much to gain by attracting these quite active customers to a service of higher speed. A challenge here would be to quantify the experienced quality of the service delivered, and to illustrate how this would improve by proper dimensioning of the access service.

The same kind of reasoning may be applied to the 100 Mbps and 100/10 Mbps STs. However, in these cases, the issue would rather be to consider downgrading as the relative load is below 10% for 97.9% and 98.9% of the time, respectively. On the other hand, the customers may be prepared to pay a higher price, just to be ensured that they never even get close to filling their allocated bandwidth, in order to ensure optimal performance of the link. Performance, such as low delay, is an important parameter for applications such as some online games, or real time video conferencing.

6. CONCLUSIONS

In this paper, we have investigated the actual load of residential users in a fibre based Swedish municipal access network. The investigation is aimed at revealing how much of the allocated bandwidth that was actually used.

The Internet subscriptions we investigated were 1, 5, 10, 30, 30/10, 100, 100/10 Mbps, where the forward slash denotes the split between downlink/uplink access speed. Also the cost for the subscriptions was investigated.

We found that there were huge differences between the generated data volumes between the STs The lowest speed subscriptions (1 Mbps) generated 12.9 GB of data per month in the downlink direction. This makes the cost per GB 11.6 SEK. The same numbers for the 100 Mbps ST were 183.4 GB at the price of 2.4 SEK per GB.

We also found that there were large differences in the actual traffic load. This was shown by viewing traffic volume generated in relation to the subscribed access speed. For the 1 Mbps ST, the load actually exceeded maximum speed at ~ 3% of the active time. In general the 1 Mbps customers maintain a high load when active. In the other extreme are the 100 Mbps subscribers, which rarely (~ 3% of active time) exceed 5% load in the downlink direction.

The uplink traffic volume exceeds that of the downlink for the 10, 30, 100 and 100/10 Mbps ST, while it is the opposite case for the 1, 5 and 30/10 Mbps ST. Concerning differences within the ST, we conclude that 20% of the households in the 1 Mbps ST generate 70% (85%) of the traffic in the downlink (uplink) direction. The same numbers for the 100/10 Mbps ST are 60% (downlink) and 83% (uplink).

ACKNOWLEDGEMENTS
This work has partly been financed by the CELTIC project IPNQSIS, with the Swedish Governmental Agency for Innovation Systems (VINNOVA) supporting the Swedish contribution. Maria Kihl is financed in the VINNMER programme at VINNOVA. Andreas Aurelius is partly financed by the Swedish National Strategic Research Area (SRA) within the program TNG (The Next Generation) and the project eWIN.

REFERENCES

[1] http://www.cisco.com/en/US/netsol/ns827/networking_solutions_sub_solution.html#~overview
[2] Jie Li, A. Aurelius, V. Nordell, Manxing Du, Å. Arvidsson, and M. Kihl: A five year perspective of traffic pattern evolution in a residential broadband access network, *Future Network & Mobile Summit*, paper 6e.2, 2012.
[3] F. Bass: A new product growth model for consumer durables, *Management Science*, vol. 15, pp. 215-227, 1969.
[4] http://www.ericsson.com/news/111012_mobile_data_traffic_244188808_c
[5] http://www.ericsson.com/res/investors/docs/2011/cmd/traffic_and_market_data_report_111107.pdf
[6] Fraleigh, C.; Moon, S.; Lyles, B.; Cotton, C.; Khan, M.; Moll, D.; Rockell, R.; Seely, T. & Diot, C.: Packet-level traffic measurements from the Sprint IP backbone, *IEEE Network*, vol. 17, 6-16, 2003.
[7] Fukuda, K.; Cho, K. & Esaki, H.: The impact of residential broadband traffic on Japanese ISP backbones, *SIGCOMM Comput. Commun. Rev.*, ACM, vol. 35, 15-22, 2005.
[8] Maier, G.; Feldmann, A.; Paxson, V. & Allman, M.: On dominant characteristics of residential broadband internet traffic, *Proceedings of The 9th Acm Sigcomm Conference on Internet Measurement Conference, ACM*, 90-102, 2009.
[9] Kihl, M.; Odling, P.; Lagerstedt, C. & Aurelius, A.: Traffic analysis and characterization of Internet user behavior, *Ultra-Modern Telecommunications and Control Systems and Workshops (ICUMT*), 2010 International Congress on, 224-231, 2010.
[10] Aurelius, A.; Lagerstedt, C.; Sedano, I.; Molnar, S.; Kihl, M. & Mata, F.,:TRAMMS: Monitoring the evolution of residential broadband Internet traffic", *Future Network & MobileSummit 2010 Conference Proceedings*, 2010.
[11] Sripanidkulchai, K.; Maggs, B. & Zhang, H.: An analysis of live streaming workloads on the Internet, *Proceedings of the 2004 ACM SIGCOMM Internet Measurement Conference, IMC 2004*, 41–54, 2004.
[12] Guo, L.; Chen, S.; Xiao, Z. & Zhang, X.,: Analysis of multimedia workloads with implications for Internet streaming, *Proceedings of the 14th International Conference on World Wide Web, ACM*, 519–528, 2005.
[13] Chesire, M.; Wolman, A.; Voelker, G. M. & Levy, H. M.,: Measurement and analysis of a streaming-media workload, *Proceedings of the 3rd Conference on Usenix Symposium on Internet Technologies and Systems, USENIX Association*, vol. 3, 2001.
[14] Bhole, Y. & Popescu, A.: Measurement and analysis of HTTP traffic, *Journal of Network and Systems Management*, Springer New York, 2005, vol. 13, 357-371.
[15] Findahl, O.: The Swedes and the Internet, World Internet Institute, 2008.
[16] World Internet Project, http://www.worldinternetproject.net
[17] Procera Networks, http://www.proceranetworks.com
[18] PacketLogic DRDL Signatures and Properties v12.4