Analysis of World of Warcraft Traffic patterns and User behavior

Maria Kihl[1], Andreas Aurelius[2], and Christina Lagerstedt[2]
[1]Dept. of Electrical and Information Technology, Lund University, Sweden
[2]Acreo AB, Kista, Sweden

***Abstract*—Internet traffic measurements are necessary in order to both understand how the Internet is used and how new applications may affect the networks. World of Warcraft (WoW) is currently the most popular massively multiplayer online role-playing game (MMORPG) with more than 11 million subscribers worldwide. In this paper, we present WoW results from traffic measurements performed in a Swedish commercial broadband access network. The data was obtained with deep packet and flow inspections using a commercial traffic monitoring tool. Data was collected during a 4 months period. The objective of the measurements was to analyze World of Warcraft traffic patterns and user behavior. We show weekly traffic patterns, session duration analyses, and traffic rate distributions. Also, we show how World of Warcraft usage varies depending on the households' broadband access rates.**

I. INTRODUCTION

Massively Multiplayer Online Role-playing Games (MMORPGs) are attracting numerous players [11], and World of Warcraft (WoW) is currently the largest MMORPG, with more than 11 million subscribers worldwide [1]. In order to understand how WoW players affect the Internet, traffic measurements are necessary. Internet traffic and applications have been analyzed since long, see for example, [15][16]. Some of these papers report on gaming traffic, however, mainly on an aggregated level. Analyzing WoW traffic patterns and user behavior requires traffic measurements on the application layer, using deep packet and flow inspection. Also, the measurements need to be performed close to the users, for example in the broadband access network.

Some papers have investigated user behavior and social interaction in WoW, see for example [11]. Also, traffic characteristics depending on player actions have been investigated, see for example, [3] [6] [8]. Other papers have analyzed data traces and derived game session duration distributions [2][4][6]. One large study on gaming sessions was performed in [4], where game-play histories of about 7,000 players during a two-year period were analyzed. The data traces were obtained by logging into a game server and using the "who" command in WoW. The results showed that WoW players are very active with long session durations and that they often play every day. A similar measurement methodology was used in [13], where data from more than 12,000 players during a 5-weeks period was analyzed. However, none of these studies have analyzed WoW data obtained by traffic measurements close to the users in the broadband access network.

In this paper, we analyze traffic characteristics and user behavior for WoW, with data obtained during the time period spanning from May to August 2009 from a municipal broadband access network in Sweden. The measurements were performed with a commercial monitoring tool called PacketLogic [10], using deep packet and flow inspections. A total of 435 households with active WoW players in this network are analyzed. We derive average traffic rates, daily traffic patterns, session durations and analyze how the broadband access subscription affects the user behavior. Also, we investigate the general Internet usage of the households in the network.

We show that the application group generating most traffic is P2P file sharing, mainly BitTorrent. Also, households with higher access bandwidths generate more traffic and are online more than households with lower bandwidths. About 20% of the households in the network had active WoW players. Our results show that the average WoW session was 2.3 hours, and that the longest session was 16 hours. Further, WoW generates very low traffic rates, in average only 21 kbps. Active households played in average almost one hour per day.

II. TARGET NETWORK AND MEASUREMENTS

The data presented in this paper was collected from a municipal broadband access network in Sweden. The network offers broadband Internet access as well as other services such as IPTV to its customers. The network is fiber-based, and customers can freely choose among the services offered by the different providers connected to the network. The network uses dynamic IP-address allocation with DHCP. The lease time for the IP-addresses varies with the service provider with the shortest lease time being 20 minutes. It should be noted, however, that we during the investigation did not have access to the actual IP addresses. All data was anonymized, where each IP address was coded with a unique identification number, using a hash function. This is necessary in order to comply with the Swedish laws on personal integrity.

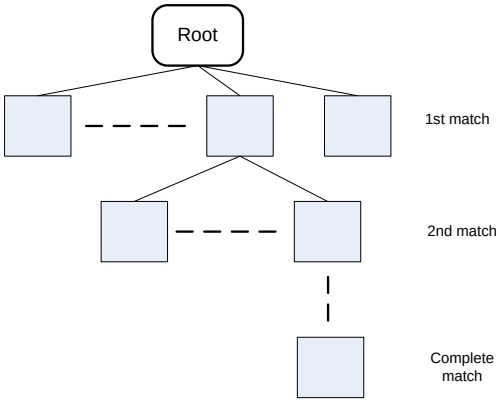

Fig. 1 A schematic view of the identification algorithm in PL.

A. Measurement tool

The measurements have been performed using PacketLogic (PL) [9], a commercial real-time hardware/software solution used mainly for traffic surveillance, traffic shaping or as a firewall. The PL is 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. PL uses the self-developed Datastream Recognition Definition Language (DRDL) [10] to identify different application protocols. Currently, the PL can identify more than 1000 Internet application protocols, and the signature database is continuously updated when new applications are deployed.

Since the PL is a commercial product, the details of its functions are proprietary. However, 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, usually detected by some hand shaking procedure, the identification of this connection begins. The identification algorithm searches for specific patterns, called application signatures, in the connection. The patterns are found in the IP header and application payload. The identification algorithm uses a tree-like structure of patterns, see Fig. 1. When the identification starts, the algorithm is in the root of the tree. When certain patterns occur in the traffic, the algorithm moves down in the tree. When it reaches one of the leaves, the full identification of the connection is completed.

Most connections are bidirectional. The PL uses the traffic in both directions in the identification process. Also, different types of connections need to be tracked for different periods of time. For example, a Bit Torrent connection only needs to be tracked when it is established. Once it is identified as a Bit Torrent connection, the tracking of the connection can stop. However, an HTTP connection needs to be tracked until it is finished, since HTTP may be used to tunnel other application protocols, e.g. some file-sharing applications and VPNs.

The PL can track and identify several hundred thousand simultaneous connections, storing statistics in a database. It records the short-time average amount of traffic in the inbound and outbound directions as well as the total traffic for all nodes in the network. The data is averaged over 5 minute periods.

B. Measurements

The studied households have Fiber-To-The-Home (FTTH) broadband access. The FTTH households are spread all over the town, representing many demographic groups and household constellations. Internet speeds range from 1 Mb/s to 100 Mb/s, depending on which service the households have chosen.

The measurement equipment was connected to the network via optical 50/50 splitters, see Fig. 2. The optical splitters merely split the optical signal into 2 exact copies, so that the traffic in the network is not affected by the measurement device. The measurement point is the Internet Edge (IE) aggregation point, where the service providers are connected to the network. The measurements were performed during 4 months, from May 1 until August 31, 2009.

The data logged by the PL is based on hosts (IP addresses). Since one household can have several users, and thereby, several hosts, it is not possible to separate households in these analyses. Therefore, we also registered data from the DHCP server used to provide network addresses for users. The DHCP server logs data concerning date and time, IP address, service provider, access switch and access port. This data was combined with the host based data from the PL in a MySQL database, which meant that households with difference access speeds could be separated. It is worth noting that the online time of a household is measured in 5-minute periods. Thus, if the household has sent or received traffic during a 5 minute period, it is classified as active for 5 minutes. Also, data is truncated at 100 kbps average for this 5-minute period. If below this threshold, data is truncated to zero, and thus not included in the statistics.

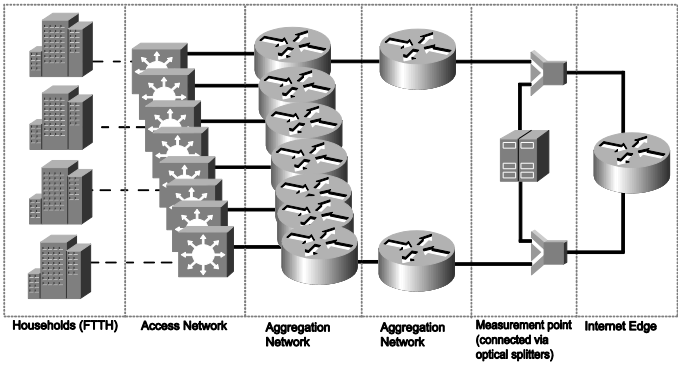

Fig. 2 Conceptual overview of the Municipal network architecture for FTTH. The dashed lines denote 100 MB/s links and the solid lines denote 1 Gb/s links.

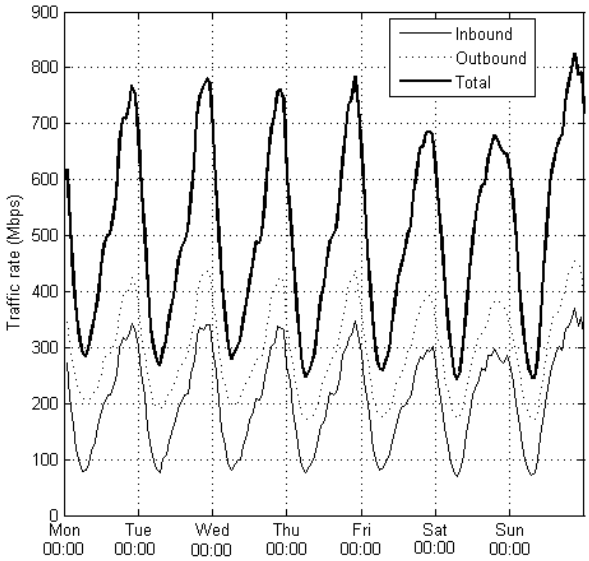

Fig. 3 Weekly Internet traffic pattern.

Table 1 Traffic volume ratios

Application group	*Total*	*Inbound*	*Outbound*
Web browsing	5.6%	13%	0.9%
P2P File sharing	80%	68%	89%
Streaming	2.2%	5.0%	0.6%
Messaging and Collaboration	0.8%	0.9%	0.8%
File Transfer	3.4%	3.1%	3.5%
Online Gaming	0.4%	0.9%	0.1%
Other traffic	7.1%	9.7%	5.5%

Table 2 General Internet usage behavior

Broadband access (in/out Mb)	*Number of households*	*Online minutes per day*	*MB per day*
1/1	159	71	113
5/5	116	162	613
10/10	1187	196	1330
30/30	170	250	3420
100/10	165	243	3350
100/100	82	291	5170

III. RESULTS

In this section we present the results of the measurements. During the measurement period (May-Aug 2009), 435 households had active WoW players. To be registered as an active player, it is necessary that a host in the household played WoW during at least one measurement interval of 5 minutes.

A. General Internet behavior in the network

As background information, we here show some aggregated traffic analysis for the network. Fig. 3 shows the average weekly traffic pattern for all Internet applications and hosts in the network during 17 weeks, May 4 until Aug 30 2009. As can be seen, the traffic is asymmetric, with more outbound traffic (to the Internet) than inbound traffic (from the Internet). As expected, the peak time is in the evenings, and the peak time starts earlier on weekends than on weekdays. Also, the peak traffic rate is lower on Fridays and Saturdays than the other days. However, the average accumulated traffic volume is about the same each day except Sundays, about 2.5Tb. On Sundays, the average accumulated traffic volume is slightly larger, 2.8Tb.

Table 1 shows the traffic volume ratios for the most used application groups. The most commonly used P2P file sharing application is Bit Torrent. Other applications in this group are, for example, Direct Connect (DC), eDonkey, Kazaa, Gnutella and PeerEnabler. The Streaming group includes all applications used for multimedia streaming, such as RTSP, HTTP media stream (HTTP ms), and RTP. Messaging and collaboration includes applications for email, instant messaging, and Voice over IP. File transfer consists of applications such as FTP, used for raw download of files. Online gaming includes multiplayer online games, on all platforms. Many games are included in this category, where two popular examples are World of Warcraft and Half-Life. The group "Other" includes many applications with low volumes, for example, SSH and SSL. Also, applications related to network management belong to this category.

However, even if online games generate very little traffic compared to other application groups, gaming is not an uncommon Internet activity. Since online games generate low traffic rates compared to downloads of for example film and music files with Bit Torrent, online games will not be a major source of Internet traffic. In our network, the most popular online game is World of Warcraft, followed by various Xbox games and Counterstrike.

Table 2 shows the general Internet usage behavior for households with different broadband access subscriptions. The results are average values for the measurement period. All Internet applications are included in this analysis. As can be seen in Table 2, households with more bandwidth are more active Internet users.

B. Traffic rates for WoW

Fig. 4 shows the average weekly traffic pattern for WoW, calculated from measurements during 17 weeks, from May 4 until Aug 30 2009. As can be seen, the peak time starts in the afternoon when the school day ends, and peak traffic occurs at around 10pm. WoW generates very little outbound traffic due to the game architecture.

Also, WoW shows a very distinctive weekly pattern, where Fridays and Saturdays only generate about 50% of the peak traffic rate compared to other days. On the other hand, the peak time starts earlier on these days. Therefore, when analyzing the actual traffic volume, we found only a small difference between the days. On Mondays-Thursdays, the total WoW traffic volume is about 3Gb per day. On Fridays and Saturdays, the total volume is about 2.5Gb per day. On Sundays, there is a slightly larger traffic volume of 3.3 Gb.

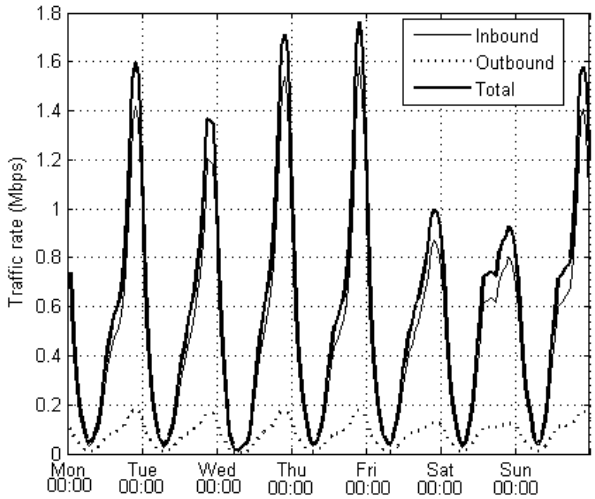

Fig. 4 Weekly traffic pattern for WoW.

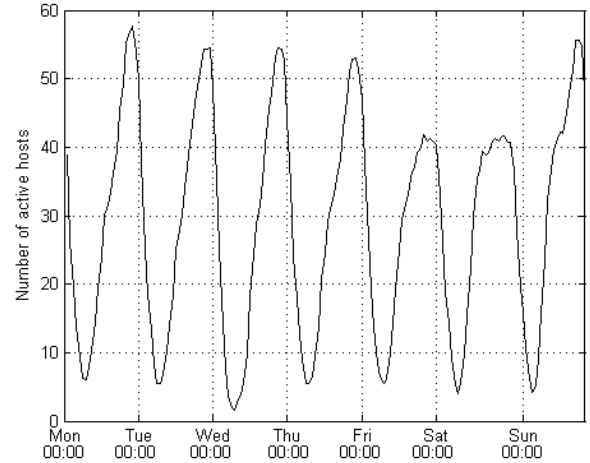

Fig. 5 Average number of active hosts during a week

This means that even if the peak rate is lower on Fridays and Saturdays, people play almost as much as the other days, since they instead start earlier in the day. Fig. 5 shows the average number of active hosts during a week. As assumed, the number of WoW hosts is partly related to the generated traffic.

None of the other papers studying WoW traffic and usage have performed the same type of measurements as we have done. However, in [13], it was shown that the number of WoW players was about the same for all days in the week. In [5] it was shown that the playtime was longer on weekends than weekdays, however, no significant differences could be found. It would, therefore, be interesting to see more measurements, in order to further analyze country specific user patterns.

C. Usage patterns

Table 3 shows the average WoW usage pattern for the households, measured in average playing time per day and average number of generated Mbytes per day. Also, we show the results for some of the households grouped according to their broadband access rate. The other households belong to other ISPs with varying bandwidths. If the group with 1/1 Mbps broadband access rate is ignored, more than 20% of the households in the network have active WoW players.

As can be seen, households that play WoW are active in average almost one hour per day. The most active households played in average 13 hours per day! However, we cannot see how many accounts there are in a household. WoW only allows one active user per account at a time, and a household may of course pay for more than one account. Households with 30/30 Mbps broadband access appear to be very active WoW players. However, since we have no demographic information about the subscribers, it is not possible to make any "social" conclusions regarding this behavior.

Table 3 Households' usage of WoW

Broadband access (in/out Mb)	*Number of WoW households*	*Minutes per day*	*MB per day*
All households in network	435	48	9.8
1/1	16 (10%)	5	1.1
5/5	26 (22%)	31	7.5
10/10	240 (20%)	44	9.0
30/30	47 (28%)	100	21
100/10	47 (28%)	52	9.2
100/100	17 (21%)	56	13

D. Session activity

We performed a detailed analysis of user sessions during May 2009. A user session is here defined as the number of consecutive measurement intervals when a host plays WoW. A total of 7823 sessions were analyzed.

Fig. 6 shows the cumulative density function of the session length, measured in 5 minutes intervals. The average session length was 2.3 hours, whereas the median length was 1.4 hours. The longest recorded session was 15 hours! These results are rather consistent with the results in [2] and [5], which means that long playing sessions is not only a Swedish behavior. Also, in the graph we have fitted the data with a Weibull distribution using a maximum likelihood estimation of the parameters computed from the data samples. The Weibull distribution is used in many areas due to its flexibility. The cumulative distribution function for a Weibull distributed stochastic variable X, $F(x) = P(X \leq x)$, is given by

$$F(x) = 1 - e^{-(x/\lambda)^k}$$

for x>0. The distribution contains two parameters, the shape, $k > 0$, and the scale, $\lambda > 0$. The fit shown in has λ=137 and

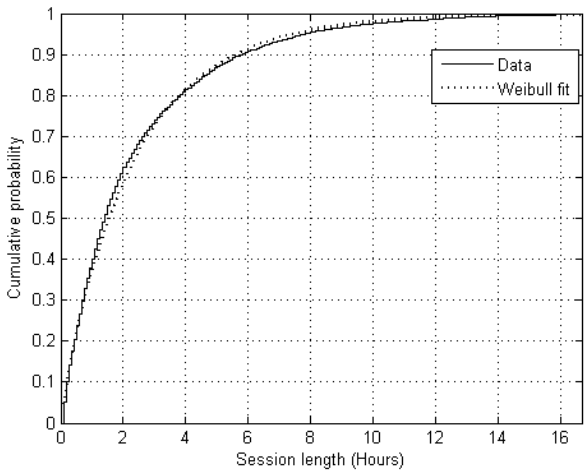

Fig. 6 Cumulative density function for session lengths. Solid line = Measurement data; Dotted line = Weibull distribution with parameters λ = 137 and k= 0.93.

k=0.93. As can be seen, *k<1*, which indicates a so called decreased failure rate. In our case, this means that there is a high probability for low session durations, but there are also some very long sessions.

Fig. 7 shows the cumulative density function for the traffic rate, measured as averages during 5 minutes periods. The average traffic rate was only 21 kbps, which means that WoW generates very low traffic**.** One reason is that WoW stores most of the game information on the hosts' computers, which means that the Internet communication can be reduced.

IV. Conclusions

World of Warcraft (WoW) is currently the most played MMORPG in the world. In this paper, we have analyzed WoW usage and traffic patterns for households in a Swedish commercial broadband access network. The data was obtained with deep packet and flow inspection, using the commercial traffic monitoring tool PacketLogic. In comparison with other studies on WoW, our measurements are long term and close to the users, and we have access to household usage patterns and application layer traffic data.

Our results show that about 20% of the households in the network had active WoW players during the 4 months measurement period (May – Aug 2009). Also, the WoW households were very active, with an average game session length of 2.3 hours, and an average playing time of about 1 hour per day. The peak time started in the afternoon and continued all evening with a highest traffic rate at about 10pm. On Fridays and Saturdays, the peak rate was much lower than other days. However, instead the gamers played more during the afternoon.

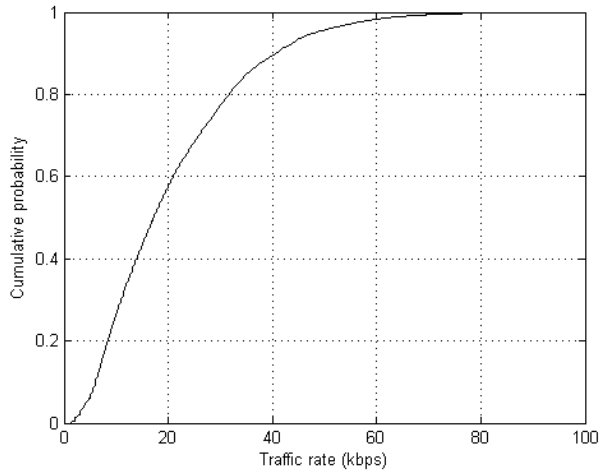

Fig. 7 Cumulative density function for traffic rates.

V. Acknowledgments

The work has partly been financed by the CELTIC project TRAMMS, with the Swedish Governmental Agency for Innovation Systems (VINNOVA) supporting the Swedish contribution. Maria Kihl is financed in the VINNMER program at VINNOVA. Andreas Aurelius and Christina Lagerstedt are partly financed by the VINNOVA project Broadband Behavior.

References

[1] Blizzard Entertainment Inc. World of Warcraft, http://www.worldofwarcraft.com.

[2] P. Svoboda, W. Kamer, and M. Rupp, "Traffic analysis and modeling for World of Warcraft", Proc. of the IEEE International Conference on Communications, 2007.

[3] M. Suznjevic, M. Matijasevic, and O. Dobrijevic, "Action specific massive multiplayer online role playing games traffic analysis: Case study of World of Warcraft", Proc. of the 7th ACM SIGCOMM Workshop on Network and System Support for Games, 2008.

[4] HyoJoo Park, C. Film, and TaeYong Kim, "Network traffic analysis and modeling for games", Proc. of first International Workshop on Internet and Network Economics, Lecture Notes in Computer Science, Vol. 3828, 2005.

[5] P. Tarng, K. Chen, and P. Huang, "An analysis of WoW players' game hours", Proc. of the 7th ACM SIGCOMM Workshop on Network and System Support for Games, 2008.

[6] M. Suznjevic, O. Dobrijevic, and M. Matijasevic, "MMORPG Player actions: Network performance, session patterns and latency requirements analysis", *Multimedia Tools and Applications*, v 45, n 1-3, p 191-214, Oct. 2009.

[7] K. Chen, P. Huang, and C. Lei, "Game traffic analysis: An MMORPG perspective", *Computer Networks*, v 50, n 16, p 3002-3023, Nov. 2006.

[8] YoungTae Han and HongShik Park, "Distinctive traffic characteristics of pure and game P2P applications", Proc. of 10th International Conference on Advanced Communication Technology, p 405-8, 2008.

[9] G. Szabo, A. Veres, and S. Molnar, "Effects of user behavior on MMORPG traffic", Proc. of IEEE International Conference on Communications, 2009.

[10] "Procera networks." http://www.proceranetworks.com.

[11] "Packetlogic DRDL signatures and properties 10340 (beta)," tech. rep., 2007.

[12] L. Achterbosch, R. Pierce, and G. Simmons, "Massively multiplayer online role-playing games: the past, present, and future", *Computers in Entertainment*, Vol. 5, No. 4, 2008.

[13] D. Pittman and C. GauthierDickey, “A measurement study of virtual populations in massively multiplayer online games”, Proc. of the 6th ACM SIGCOMM Workshop on Network and System Support for Games, 2007.

[14] M. van Lent, "The Future Is Virtually Here," *Computer*, Vol. 41, No.8, pp. 87-89, Aug. 2008.

[15] V. Paxson, “Empirically derived analytic models of wide-area TCP connections,” *IEEE/ACM Transactionson Networking*, vol. 2, no. 4, 1994.

[16] G. Maier, A. Feldmann, V. Paxson, and M. Allman, “On dominant characteristics of residential broadband Internet traffic”, Proc. of ACM Internet Measurement Conference, 2009.