Analysis of technical limitations of NAPT with respect to TCP parameters.

Kungliga Tekniska Högskolan
IK1550 Internetworking
Manousis Antonios
Stockholm, Sweden, June 2012.

1. Purpose of this Paper

This purpose of this paper is to provide some insights into the technical limitations of a Network Address Port Translation (NAPT) device with one external global IP address, with respect to TCP. Specifically, by analyzing the overall usage of TCP ports in a browser during typical Internet usage, the number of users that this specific type of NAT can support is defined. Moreover, after taking into consideration the special characteristics of the TCP state machine, we attempt to define the latencies that are caused due to resource exhaustion and possible problems that could be caused because of this resource exhaustion.

2. Introduction

2.1 Few words about NAT- NAPT

Network Address Translation (NAT), was originally presented in RFC 1631 [1] in May 1994. NAT was designed as an *interim* approach to the problem of IPv4 address depletion by defining a set of IP addresses that could be shared or reused by many hosts. These addresses are allocated for private use *within* a home or office network and they have a meaning only for the devices within that subnet, thus creating a realm of private addresses that are transparent to the larger global Internet. Communication with the globally routable Internet relies on a NAT-enabled router which uses (at least) one global IP address and a translation table to map private addresses to globally routable addresses and forward the modified packets to the Internet.

For the purposes of this paper, only the case of Network Address Port Translation (NAPT) will be analyzed and the terms NAT and NAPT will be used interchangeably. Further analysis of the technical characteristics of NAT can be found in RFC 2663 [2] and RFC 3022[3]. In the case of NAPT, instead of using multiple public IP addresses, NAPT uses one global address and distinguishes the sessions being mapped by it based upon a source and destination port number. When a packet from the private realm reaches the NAT, the router replaces the private source address and port number with a global address and a port number that is not currently in its translation table and then forwards the packet to the Internet. Since port numbers are 16 bits long, this implies that the NAT can handle more than 60000 simultaneous connections for a given transport protocol (UDP, TCP). But to how many users does this correspond to assuming each of the users utilizes simple browser traffic?

2.2 About the TCP State machine- time to wait parameter-relevance with NAPT

The Transmission Control Protocol (TCP) is Internet's main transport-layer protocol that provides connection-oriented, reliable transport, and in order delivery of bytes. The TCP connection, from the moment it is established until it is closed, is implemented as a finite state machine. For the purposes of this paper a thorough technical description of TCP is not required, but the reader may find all the relevant information in RFC 793 [4], RFC 1122 [5], RFC 1323 [6], RFC 2018 [7], and RFC 2581 [8]. However, emphasis will be given to the sequence of events that take place when a TCP connection is closed, since the analysis of the NAPT limitations will be done with respect to these parameters.

The termination of a TCP connection follows the three-way handshake model. When one of the two parties that have established a connection wants to close it (the client generally, if we assume a client-server model), this party issues a command, known as *active close,* sends a FIN segment and goes to the FIN-WAIT-1 state of the TCP state machine. When it receives an acknowledgment (ACK) from the other party, the terminating party moves to the FIN-WAIT-2 state and waits until the other party also sends a FIN segment. The ACK and the FIN from the server could also be in the same segment, thus avoiding the transition to TIME-WAIT-2. After the second FIN segment is received, the terminating party acknowledges again and moves in the TIME-WAIT state and sets a timer until the connection is finally closed. According to RFC 793 [4], the TIME-WAIT state duration is twice the MSL (Maximum Segment Lifetime) time, hence 2* 2min or 4 minutes.

The relevance to NAPT of the technical characteristics of TCP regarding closing an established TCP connection becomes important in the case of NAPT resource exhaustion. Assuming that the number of users in a subnet that is connected to the Internet via a NAT-router with only one global IP address has reached its upper limit, what will the behavior of the NAT be, bearing in mind that a TCP port will need to stay in an idle TIME-WAIT period for 4 minutes?

3. Assumptions – characteristics of measured traffic

All the results of this research done for this paper are based on browser traffic. The browser used for this purpose was the latest version of Google Chrome for Linux (Ubuntu). Specifically, the sites that were active while measuring the traffic were an email client (Gmail), Google docs, social network sites (i.e. Facebook and Twitter), online newspaper sites and YouTube. The traffic was measured with the `ss` command.

4. Calculation of the upper bound on the number of users

A good approach to estimate the upper bound on the number of users that a NAPT can support would require estimating the number of TCP ports that are used on average per user and then given the number of available TCP ports by the NAT, calculate the upper bound. As far as the number of available TCP ports is concerned, there are a few observations that need to be made. Since a TCP port number is a 16 bits number, the total number of TCP ports has a maximum of $2^{16} = 65536$. However, port numbers in the range from 0 to 1023, also known as well-known ports are restricted for use by well-known application protocols [12, page 204] and cannot be used or counted in the number of ports that a NAT router can use, thus giving a total of 64512 available TCP ports.

In order to estimate the number of TCP ports which are on average in use due to browser traffic, measurements of the traffic were made every one minute for a period of 20 minutes. The results of these measurements showed that normal browser use, as it was defined in the section 3, requires on average 82 TCP ports. Given this result, the maximum number of such average users that a NAPT can support is 787 users per IP.

5. Problems caused by resource exhaustion

To get some insight into the technical problems caused by TCP ports exhaustion and the CLOSE-WAIT state of the TCP state machine, first of all one needs to assume that the NAT has reached its maximum number of users which, as shown above, is around 787 users. Generally, a NAT must not abandon TCP connections just because they have been in a partially open or partially closed state or idle for a long time. However, when facing resource exhaustion, the NAT has the option to abandon a TCP connection. In this case, the first connections to be closed are those that are in a partially closed state, then those who are in a partially open state, and finally those who have been idle. In any of these cases and in order to avoid unexpected behavior a NAT is obliged to send RST packets to each of the sessions that it wishes to terminate [9]. In this paper we assume that the NAT that is being analyzed does *not* abandon any established sessions and that no abrupt terminations due to unexpected rebooting of the devices sharing a connection occur.

5.1 Simultaneous closing of all established TCP connections

As a first approach, the scenario that will be analyzed is the one where all established connections start their connection termination procedure at the same time. In this case the private realm cannot communicate with the external realm for a period of 4 minutes by establishing new sessions, as all TCP sessions will be in the CLOSE-WAIT state and the corresponding TCP ports will be unavailable. As a result, all SYN segments that will be sent, trying to establish a new connection will eventually be dropped. After 4 minutes have elapsed, if all devices belonging to the private realm are still trying to start new connections, then the result will be an impulse-like increase in the number of TCP ports that are used, thus immediately reaching a state where all resources (TCP ports) are occupied. Fig. 2 shows graphically the transitions described above.

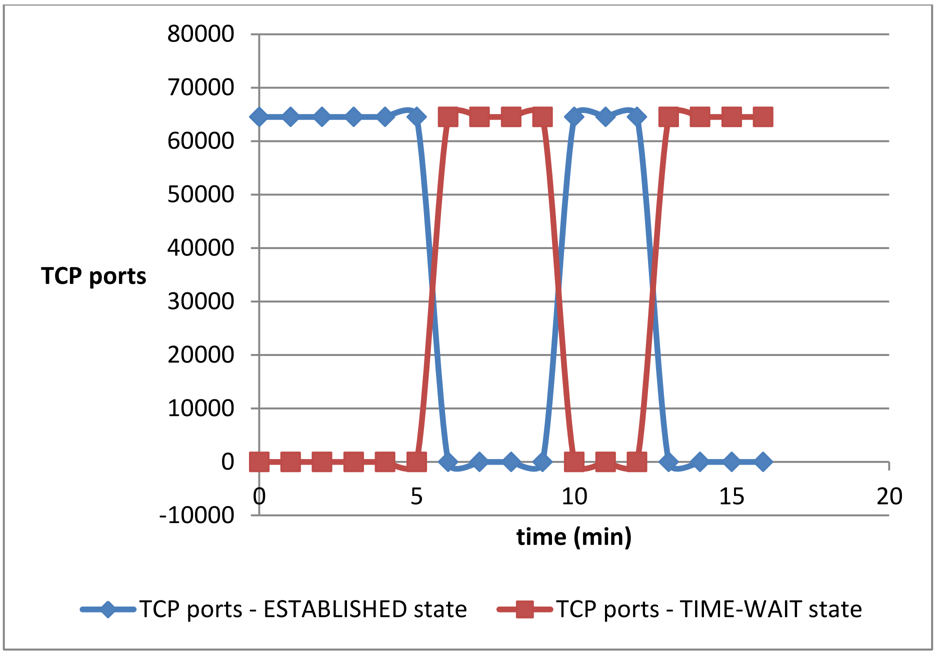

Figure 2.
Plot of Ports status at simultaneous termination of TCP connections

5.2 Asynchronous opening and termination of established TCP connections.

For this more realistic approach to the behavior of a NAT given Internet traffic created by the users of the private realm, we assumed that the number of terminating and opening TCP sessions per minute follows a Poisson distribution. The reasons for selecting a Poisson distribution are because we assume that the establishing or terminating in a given time interval sessions are independent to the number of opening or terminating connections in previous time intervals. The mean and average of these two distributions, which for a Poisson are equal, were based upon the traffic measurements taken, and the results are shown below, in Table 1.

The number of connections that are switching from the ESTABLISHED to the TIME-WAIT state every minute is calculated as the number of connections that already were in a TIME-WAIT state at minute $i-1$ minus the number of connections who were in the TIME-WAIT state at minute i. This approach, though, is not absolutely accurate since a number of ports might have moved from TIME-WAIT to CLOSED state in the time interval between two subsequent measurements of traffic. This means that the values calculated every minute represent a *lower bound* of the connections moving to the TIME-WAIT state every minute. The exact same procedure was followed for newly established connections.

Average of terminating and opening TCP connections per minute

	Result for 1 user	Result scaled for 787 users
Avg. Closing connections per min.	6	4722
Avg. Opening connections per min.	9	7083

Table 1

Given that every minute the probability of having 6 closing and 9 opening connections per user per minute is relatively high (the probability density function of a Poisson distribution reaches its peak for the average value of the random variable) and in order to simplify the rest of the analysis, only these values will be used. As far as the original NAT is concerned, again the analysis starts with the NAT having reached its upper bound of users (787) and every minute it has to process 4722 connections moving to TIME-WAIT state and 7083 requests for new connections. Each new connection would require a TCP port to be allocated. The behavior of the NAT in this case is shown in Fig. 3.

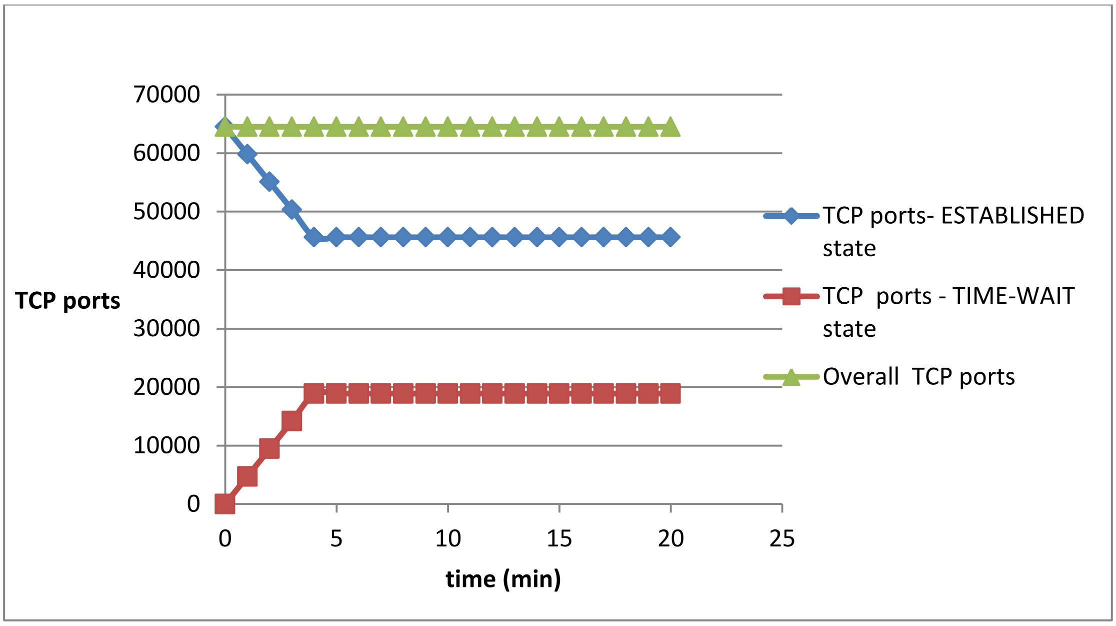

Figure 3
Port status - Asynchronous termination of TCP connections

Fig. 3 provides a lot of information about the behavior of the NAT in the case of port exhaustion. The analysis starts with the NAT having reached its user limit and with all its available TCP ports occupied. Since no new connections can be initialized it can be seen that during the first four minutes the number of TCP ports that are in the ESTABLISHED state decreases linearly by an average of 4722 ports per minute. During this period, an average of 7083 requests for new connections is issued, but due to the lack of available TCP ports and according to the original assumptions the NAT does *not* abort open connections, thus the packets that carry these requests are dropped. After the fourth minute has elapsed, the 4722 ports that first entered the TIME-WAIT stage become available again and serve a portion of the 7083 new requests for connections. The remaining packets are dropped. At the same time 4722 ports switch to the TIME-WAIT stage and thus the total number of ports in the ESTABLISHED and TIME-WAIT stage remains the same. Since only the average values are taken into consideration the graph shows a straight line but in reality the number of opening and closing connections is not fixed but still oscillates around the numbers mentioned above. Of course for every minute after the 5th minute the same series of events occurs.

An obvious difference between the behaviors shown in fig.2 and fig.3 is that in fig.2 after 4 minutes have passed, all ports become available and can serve up to 64512 new TCP connections. In the case of fig. 3 after the 5th minute, every minute only 4722 ports become available on average and can support a much smaller number of connections. In both cases a very large number of packets has to be dropped, hence showing how limited this implementation of NAPT can be in the case of big subnets, both in the number of users and in the latencies that are inevitable in the case of port exhaustion.

6. Conclusions

In this paper an effort has been made to define some special characteristics of NAT. NATs have been viewed as a short term solution to the problem of IPv4 address depletion. However, NATs have found extremely widespread use. The specific implementation that was analyzed was NATP with one global IP address. The focus of the analysis was on the TCP parameters of the connection. The basic assumption that was made was that all measurements and all calculations would be based on browser traffic only and that we have considered an average user who uses popular sites (email client, social network sites, etc.)

The results of the traffic measurements showed that this TCP traffic requires around 82 TCP ports per user. As a consequence, since the number of available ports is known and limited, defining an upper bound for the number of users that a NAPT with one global IP address can support was the next step. As it turned out, the number of users that can be supported is in the vicinity of 780 users, which is a very small number, while this is a limitation for subnets that are set for larger companies or offices. Even for ISPs who decide to hide their users behind a NAT this limitation can be quite severe in practice.

After the upper bound of users was calculated, the next step was to analyze two scenarios that would show some serious problems caused by resource exhaustion, in this case the exhaustion of TCP ports. In the first scenario the assumption was that all ports were in the ESTABLISHED state and then simultaneously switched to the TIME-WAIT state since all the connections were to be terminated. In that case, the observation that could be made was that for a time interval of 2*(Maximum Segment Lifetime) or 4 minutes no new requests could be processed and all packets that were sent requesting a new TCP connection had to be dropped because there were no available ports to send them out to the Internet. After 4 minutes elapsed, all TCP ports were available again and the NAT could support its maximum number of users again.

In the second scenario, the assumption that was made was that the number of terminating or opening TCP connections per minute followed a Poisson distribution with an average value that was calculated based on the traffic measurements. The results showed a behavior quite different compared to the previous scenario. In that case, at the beginning all ports were occupied and after the first four minutes the number of ports in the ESTABLISHED state was abruptly decreasing. After the fourth minute, every minute around 4722 ports are available for new connections and serve some of the 7000 requests for new connections that were on average issued every minute while around 4722 ports were used by connections who entered the TIME-WAIT state. The main difference from the first scenario is that in the second scenario, after the fourth minute the NAT reaches a more stable situation and every minute there is a limited number of requests served; whereas in the first scenario every time the TCP connections close there are 4 minutes during which inevitably no new requests can be processed.

Of course this situation can become quite problematic especially in cases where high traffic is expected, for example on the day when national tax filing is due. In that case or in similar ones, ISPs cannot possibly allow long latencies or high packet losses so as a result if they have their users behind a NAT they need to find more efficient ways to deal with this situation. Some solutions include either having a pool of global IP addresses, thus increasing the number of users that can be supported by the NAT by a factor of around 700 users/IP address, or instead using the TCP protocol use the Stream Control Transmission Protocol (SCTP) which can establish one association instead of one connection per port. More information about the SCTP can be found in [10],[11].

7. References

[1] K. Egevang and P. Francis, "The Network Address Translator", IETF RFC 1631, May 1994, http://www.ietf.org/rfc/rfc1631.txt.

[2] P. Srisuresh, M. Holdrege, "IP Network Address Translator (NAT) Terminology and Considerations", IETF RFC 2663, August 1999, http://www.ietf.org/rfc/rfc2663.txt.

[3] P. Srisuresh, K. Egevang, "Traditional IP Network Address Translator (Traditional NAT)", IETF RFC 3022, January 2001, http://www.ietf.org/rfc/rfc3022.txt.

[4] Information Sciences Institute, University of Southern California, "Transmission Control Protocol",IETF RFC 793, September 1981,http://www.ietf.org/rfc/rfc793.txt

[5] R. Braden, "Requirements for Internet Hosts -- Communication Layers", IETF RFC 1122, October 1989, http://www.ietf.org/rfc/rfc1122.txt.

[6] V. Jacobson, R. Braden, D. Borman," TCP Extensions for High Performance", IETF RFC 1323, May 1992, http://www.ietf.org/rfc/rfc1323.txt.

[7] M. Mathis, J. Mahdavi, S. Floyd, A. Romanow, "TCP Selective Acknowledgment Options", IETF RFC 2018, October 1996, http://www.ietf.org/rfc/rfc2018.txt.

[8] M. Allman, V. Paxson, W. Stevens, “TCP Congestion Control”, IETF RFC 2581, April 1999, http://www.ietf.org/rfc/rfc2581.txt.

[9] P. Hoffman, “NAT Behavioral Requirements for Unicast TCP, draft-hoffman-behave-tcp-00.txt”, June 2005.

[10] Behrouz A. Forouzan, *TCP/IP Protocol Suite*, 3rd edition, McGraw-Hill, publication date January 2005.

[11] David A. Hayes, Jason But, and Grenville Armitage, Issues with Network Address Translation for SCTP , ACM SIGCOMM Computer Communications Review, Volume 39, Number 1, January 2009, pages 24-33.

[12] James F. Kurose and Keith W. Ross, *Computer Networking: A Top-Down Approach*, Fifth Edition, Pearson, 2010.

[13] Cisco, IP Addressing Services, “Network Address Translation (NAT) FAQ”, February 2010, http://www.cisco.com/en/US/tech/tk648/tk361/technologies_q_and_a_item09186a00800e523b.shtml#nat-deploy

[14] L. Phifer, The Internet Protocol Journal, Vol.3 n.4, “The Trouble with NAT”, December 2000,http://www.cisco.com/web/about/ac123/ac147/ac174/ac182/about_cisco_ipj_archive_article09186a00800c83ec.html