Characteristics of the Traffic on Serbian Open Exchange

Nenad Krajnović, *Member, IEEE*

***Abstract*—Internet is constantly changing and the same is true for its traffic. Because of that, it is important to constantly measure and analyses Internet traffic. The results of Internet traffic analysis is valuable source of information for network traffic modeling and defining network design strategies. This paper presents characteristics of Internet traffic measured at Serbian Open Exchange, which is the only Internet exchange in Serbia. Presented results shows change in Internet traffic structure, comparing with the Internet traffic in the past.**

***Index Terms*—Internet traffic; measurement; SOX.**

I. Introduction

INTERNET services and traffic are changing everyday. Few years ago, video streaming was at the beginning and today we have Web TV as one of the main service for end users. All typical telecommunication services, like audio communication and television, are migrating to the Internet. All those changes have high impact on Internet traffic characteristics. The only valid method for characterization of Internet traffic is analysis of huge amount of Internet traffic. The appropriate position for traffic measurement is the backbone part of the network, because of very high number of traffic flow that exists on it. Statistical analysis of huge amount of traffic with very high number of traffic flow produced valid average characteristics of Internet traffic. Looking on modern Internet architecture, one of the main elements is Internet Exchange Point (IXP). The majority of big Internet Service Providers (ISP) are feeding there networks with traffic from IXPs. Therefore, IXP is very good point for traffic measurement.

In Serbia, the only public, carrier neutral IXP is Serbian Open Exchange (SOX). The backbone of the SOX is located in Belgrade, with extensions to Vienna Internet Exchange (VIX), NET-IX and B-IX in Sofia (Bulgaria) and Amsterdam Internet Exchange (AMS-IX). SOX network topology is presented on Fig. 1. Major SOX customers are presented on Fig. 2. As it can be seen on Fig. 2, all major Serbian ISPs are connected to SOX. Besides them, world major CDNs (Content Distribution Network) are also connected to SOX. Such a number of SOX customers guarantee representative Internet traffic. Besides already mentioned SOX customers, SOX is hosting three root DNS (Domain Name Service) servers (J, K and L root DNS servers). Total daily traffic on SOX network is presented on Fig. 3. As it can be seen on Fig. 3, total traffic in peak is over 170Gb/s. This amount of traffic is representative for the purpose of obtaining general conclusions about today Internet traffic. All SOX users can be classified in three groups: ISPs, CDNs and root DNS servers. Therefore, three types of traffic can be identified and analyze: ISP traffic, CDN traffic and root DNS servers' traffic. Mentioned three types of Internet members are representative for the conclusions about Internet traffic.

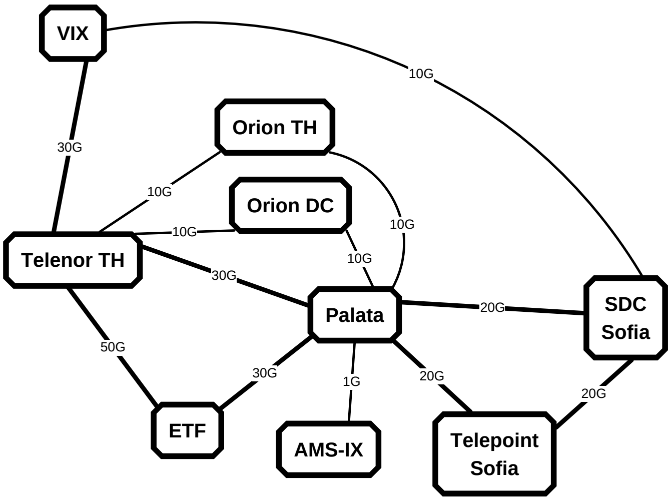

Fig. 1. SOX network topology

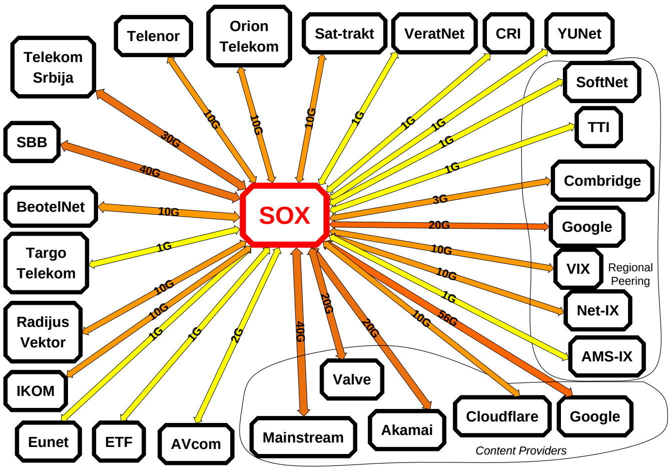

Fig. 2. SOX customers

Second section of the paper contains description of measuring methodology used for collecting the traffic data. Measured results and the results of analysis are presented in

Nenad Krajnović is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: krajko@ etf.bg.ac.rs).

the third section. Comparison of measured results presented in this paper with traffic analysis results published in the past are in forth section. Conclusions are presented in fifth section.

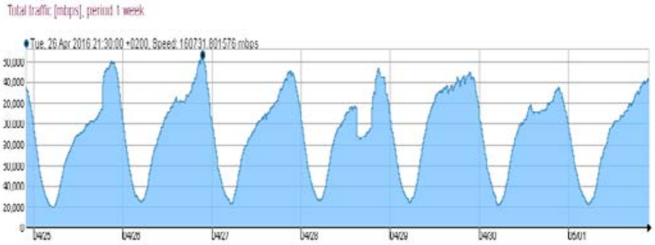

Fig. 3. Total traffic on SOX network in one week

II. Measuring of Internet Traffic

SOX is implemented as a Layer 2 Internet Exchange [1]. It means that SOX network consists mainly of Layer 2 Ethernet switches and the routers belong to SOX's end users. It means that the traffic can be measured only on those Ethernet switches. Generally, two methods for measuring traffic on Ethernet switches are known. First one is using measuring tools implemented in Ethernet switch by the vendor. Those tools are typically:

- RMON (Remote MONitoring) statistics [2],
- Netflow/Sflow/Jflow statistics ([3], [4], [5]),
- Per port traffic analysis on packet length, number of bytes, number of packets, destination MAC addresses (broadcast, unicast, multicast).

Second one is to distribute probes to the network and to mirror traffic on the switches to the ports with probes. Depending on capabilities of probes, more or less information about traffic can be collected. By using the probes, two approaches can be done. First one is with intelligent probes which do the analysis of traffic and generate statistical parameters ([6]). The second one is with dumb probes which have to save all traffic which will be latter analyze on some bigger computer, like it was done in [7]. With 50Gb/s of average traffic in the network, this second approach would collect 540 TB of data per 24 hours that would need to be analyzed! But, off line analysis of the captured traffic allows the researchers detailed analysis such as: lifetime of sessions, traffic distribution during the session, flow characteristics, number of retransmissions etc. Since the permanent traffic measurement and analysis is not an easy task, Hoogesteger et al. presented the idea ([8]) to establish Internet traffic statistics archive. The purpose of this archive would be to help researchers to understand the trends in Internet traffic.

For traffic analysis on SOX network, tools implemented on network devices were used. SOX network consists of Ethernet switches manufactured by Cisco Systems, Extreme Networks, Dell, Supermicro and DCN. Majority of devices have implemented sflow traffic analysis and all of them have statistical data about number of packets, number of bytes, packet lengths and type of MAC addresses per port.

III. Characteristics of Measured Internet Traffic

Daily traffic distribution is presented on Fig. 3. It is typical traffic distribution for the network which is closely linked with human activity. The lowest traffic is between 4:00 and 6:00 hour in the morning and the highest traffic is between 21:00 and 22:30 in the evening. The lowest traffic is about 20Gb/s and the highest traffic is more than 170Gb/s.

Traffic distribution per service is obtained from sflow data collected on network devices and 928TB of traffic and it is presented in table I.

TABLE I
Traffic Distribution per service

Internet Service	share in total traffic
web	76.33%
unknown	23.16%
rtsp	0.26%
dns	0.08%
ntp	0.03%
rsync	0.02%
openvpn	0.02%
rtp	0.02%
pop3	0.01%
ftp	0.01%
mysql	0.01%
ssh	0.01%
imap	0.01%
Total:	**99.98%**

It is obvious, from results presented in table I, that web traffic is dominant traffic in today Internet. It should be noted that web traffic is not necessary typical web service. Web TV and video streaming is one of the most popular service and they are using HTTP/HTTPS protocols. Since sflow measuring system implemented on Ethernet switches can not provide deeper information about HTTP/HTTPS traffic, detailed analysis should be done by traffic capturing and off-line analysis. Web traffic (76.33%) consists of “clear text” content based on HTTP (56,73%) and encrypted content based on HTTPS (19.6%). Those figures show that more and more traffic everyday becomes encrypted, but, still significant part of web traffic is unprotected. Since YouTube, one of the most known video streaming services is using HTTPS, conclusion is that large part of Internet web space is unprotected. This fact clearly shows that web space is very vulnerable, like classical phishing attack and traffic interception.

Large portion of analyzes traffic (23.16%) could not be easily identified because it was not using standard TCP/UDP port numbers. For detailed analysis of this traffic, it is necessary to capture traffic and do the detailed analysis off-line. Everything else is less than 0.5% of total Internet traffic covered with this research. In table I some of those services is identified by name. It is important to notice existence of POP3 protocol which is well known unsecure protocol for e-mail retrieval. All best current practice documents regarding e-mail service are stating that POP3 protocol should be obsolete, but this traffic analysis shows it is still significantly used in everyday life.

Second analysis was done based on statistical data obtained from directly for network devices. This analysis covers distribution of packet length. When we are talking about performance of network devices, two parameters are of major importance, number of bytes and number of packets that can be forwarded in one second. Number of packets is, generally, more restricted parameter. Not all network devices on the market can forward theoretical maximum number of packets that can occur on their Ethernet ports. Because of that, it is important to know the distribution of packet length per user type. Since the traffic measurement was done on SOX IXP, four Internet user types were identified:

- ISP,
- CDN,
- Root DNS servers,
- AMRES – Serbian Academic and Research Network.

In table II, average packet length to and from those SOX users are presented. Average packet lengths presented in table II clearly shows significant differences between traffic of different users.

TABLE II
AVERAGE PACKET LENGTH PER SOX USER

SOX user type	average packet length to user [B]	average packet length from user [B]
ISP	**794**	**581**
CDN	**345**	**1081**
Root-DNS servers	**129**	**625**
AMRES	**992**	**416**

The main purpose of CDNs is to host and provide data to the end users. Basically, that service is very similar to file transfer. To achieve high efficiency, CDNs are, in general, using longer packets. The result is average packet length from CDN of 1081B. On the opposite, requests to CDNs are mostly very short which is presented by average packet length of 345B to CDNs. ISPs are different story. They have end users and they are also hosting content on their own servers or the content is hosted on networks of ISP’s customers. That’s the reason why the packet length to and from ISP network is more similar, comparing to CDNs.

Root DNS servers are of crucial importance for functioning of Internet. Practically, every Internet service is starting by request to root DNS servers. Since three root DNS servers are hosted on SOX, it was very convenient to analyze their traffic. It is well known that DNS queries are very short and the answers are much longer. Classical DNS response is limited to 512B but, using extension mechanisms for DNS (EDNS0 – [9]), this limit is overcome. The results are that average packet length to root DNS server is 129B and the response is 625B.

AMRES represents very specific user of SOX. Since AMRES has its own links to the Internet, only traffic with domestic part of the Internet is exchanged over the SOX. Because of that, AMRES can be seen as a SOX user who is mainly retrieving data from other SOX users, like CDNs and ISPs. That’s the reason why the average packet length to AMRES is 992B and from AMRES is only 416B.

Based on obtained results, average packet length on today Internet is 609B. This number is very important for estimation of Internet traffic characteristics and for dimensioning network devices. Packet length distribution for four different types of SOX users is presented on Fig. 4.

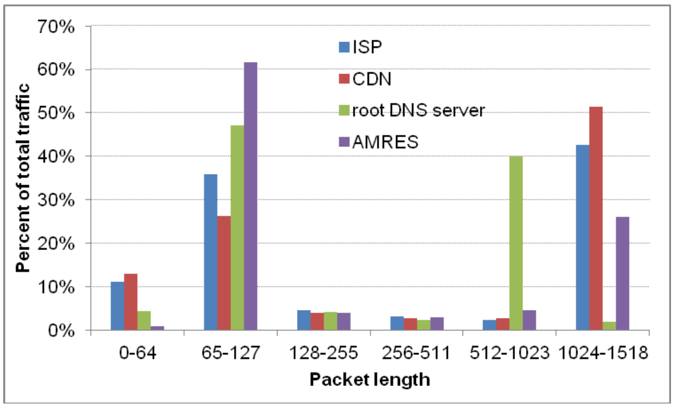

Fig. 4. Packet length distribution per SOX user type

Results presented on Fig. 4 shows that network designer should now the type of network user to be in position to correctly estimate future Internet traffic. There is significant different in the characteristics of the traffic depending on its nature.

IV. CHANGING OF THE NATURE OF INTERNET TRAFFIC

Internet traffic and services are changing every day. It is also truth for human habits. To illustrate that constant change, measuring results presented in this paper is compared with the results from 1997. presented in [10]. At 1997, average packet length varied between 200B and 400B. Today, it is 609B. Since the average packet length is significantly longer, it is obvious that today Internet is more efficient in transport data, with less overhead.

Packet length distribution also changed significantly. In the past, shortest packet (64B) accounted for more than 40% of total traffic and today is almost unnoticeable.

In the past, web traffic was about 60% of total traffic and today is 73%. This is increase of 13%. On the other side, services like Telnet and NNTP (News service) practically disappeared from the traffic statistics. Unknown services were less than 20% and today are 23%. All those figures illustrate constant change of the Internet.

On the opposite, daily traffic distribution practically didn't change, only absolute amount of traffic increase few order of magnitude. It means that the daily habits of humans didn't change significantly, looking from the Internet activity side.

V. Conclusion

Like it was said by Brian Carpenter in RFC 1958, "The principle of constant change is perhaps the only principle of the Internet that should survive indefinitely ...", Internet is changing every day. Comparison of Internet traffic characteristics in 1997 and today clearly illustrate that principle. Today Internet become more efficient, with longer average packet length. The structure of packet length distribution also changed because of change of Internet services usage. Web service become dominant Internet service today and other services become almost unnoticeable in total Internet traffic. Besides all previous mention changes in Internet traffic, human habits stayed the same. The graph of daily traffic change keeps the same shape with increase the absolute values for few orders of magnitude.

References

[1] N. Krajnović, Z. Perović, "Realization of SOX *exchange*", Proceedings of 18th Telecommunications Forum (TELFOR 2010), 23-25 November 2010, Belgrade, Serbia

[2] S. Waldbusser, R. Cole, C. Kalbfleisch, D. Romascanu, "RFC 3577 - Introduction to the Remote Monitoring (RMON) Family of MIB Modules", August 2003.

[3] Cisco NetFlow Ver. 9: http://www.cisco.com/c/en/us/products/ios-nx-os-software/netflow-version-9/index.html, April 2016.

[4] Sflow Industry standard technology: http://www.sflow.org/, April 2016.

[5] Jflow - Juniper Flow Monitoring, http://www.juniper.net/us/en/local/pdf/app-notes/3500204-en.pdf, April 2016.

[6] B. Trammell, P. Casas, D. Rossi, A. Bär, Z. B. Houidi, I. Leontiadis, T. Szemethy, M. Mellia, "mPlane: An Intelligent Measurement Plane for the Internet", *IEEE Communications Magazine*, vol. 52, Issue 5, pp. 148-156, May 2014, DOI: 10.1109/MCOM.2014.6815906

[7] J. S. Park, J. Y. Lee, S. B. Lee, "Internet Traffic Measurement and Analysis in a High Speed Network Environment: Workload and Flow Characteristics", *Journal of Communications and Networks*, vol. 2, no. 3, pp. 287-296, September 2000, DOI: 10.1109/JCN.2000.6596720

[8] M. Hoogesteger, R. O. Schmidt, A. Pras, "ITSA: Internet Traffic Statistics Archive", IEEE/IFIP Network Operations and Management Symposium (IEEE NOMS 2016), 25-29 Apr 2016, Istanbul, Turkey

[9] J. Damas, M. Graff, P. Vixie, "RFC 6891 - Extension Mechanisms for DNS (EDNS(0))", April 2013.

[10] K. Thompson, G.J. Miller, R. Wilder, "Wide-Area Internet Traffic Patterns and Characteristics", *IEEE Network*, vol. 11, Issue 6, pp. 10-23, December 1997, DOI: 10.1109/65.642356