Measuring and Analyzing the Burst Ratio in IP Traffic

Dominik Samociuk, Marek Barczyk, and Andrzej Chydzinski(✉)

Institute of Informatics, Silesian University of Technology, Gliwice, Poland
{dominik.samociuk,marek.barczyk,andrzej.chydzinski}@polsl.pl

Abstract. The burst ratio is a parameter of the packet loss process, characterizing the tendency of losses to group together, in long series. Such series of losses are especially unwelcome in multimedia transmissions, which constitute a large fraction of contemporary traffic. In this paper, we first present and discuss results of measurements of the burst ratio in IP traffic, at a bottleneck link of our university campus. The measurements were conducted in various network conditions, i.e. various loads, ports/applications used and packet size distributions. Secondly, we present theoretical values of the burst ratio, computed using a queueing model, and compare them with the values obtained in the measurements.

Keywords: Packet loss measurement · Burst ratio · Performance evaluation

1 Introduction

Packet loss is an inherent property of networks based on the TCP/IP protocol stack. Their "best effort" design is on one hand flexible and resistant to large-scale failures, but on the other it does not guarantee the delivery of packets. An occasional loss of an individual packet is usually not a problem. It is retransmitted using the TCP mechanism, if completeness of data is required. Or, it can be simply ignored, if it is a part a flow, which allows for some data loss. However, the problem is much more complicated when packet losses tend to group together, i.e. occur one after another, in series. In multimedia transmissions, such series of losses may cause for instance a video image to freeze, which is the main impediment in video Quality of Experience, [1].

The best way to describe the tendency of packet losses to cluster together is by using the burst ratio parameter B, [2]. Formally, B is the ratio of the observed average length of the series of lost packets, $\overline{G}$, to the theoretical average length of the series of losses in the case of independent packet loss, $\overline{K}$. In other words, if $\overline{K}$ denotes the average length of the series in the Bernoulli process, then

$$B = \frac{\overline{G}}{\overline{K}}. \tag{1}$$

This work was conducted within project 2017/25/B/ST6/00110, founded by National Science Centre, Poland.

Q. Li et al. (Eds.): BROADNETS 2019, LNICST 303, pp. 86–101, 2019.
https://doi.org/10.1007/978-3-030-36442-7_6

If L denotes the loss ratio (loss probability), then it is easy to check that in the Bernoulli process it holds

$$\overline{K} = \frac{1}{1-L}. \tag{2}$$

This means, that B is in fact a function of $\overline{G}$ and L, i.e.

$$B = \overline{G}(1-L). \tag{3}$$

It is very important not to confuse the burst ratio with the traffic burstiness, which has been studied earlier and has a vast literature. The burst ratio characterizes the loss process, rather than the arrival process. It has got more attention recently, due to its connection with the quality of multimedia transmissions.

To illustrate the burst ratio, we may consider a stream of 16 packets at a network node's buffer:

$$FFFFDDFFFFDDDFFF$$

Here F denotes a packet successfully transmitted/forwarded and D - a packet that was lost/dropped. There are two series of lost packets of length 2 and 3, respectively. Thus $\overline{G} = \frac{2+3}{2} = 2.5$. Moreover, there are 5 lost packets out of 16, therefore $L = \frac{5}{16}$. Thus $\overline{K} = \frac{1}{1-\frac{5}{16}} = 1.45$ and $B = \frac{\overline{G}}{\overline{K}} = 1.72$. Obviously, if $B > 1$, then the losses tend to group together - the higher the B value, the stronger this tendency is. When $B = 1$, the losses look purely random and independent of each other. Finally, if $B < 1$, then the losses tend to occur separately.

The most important parameter of the packet loss process is, of course, the loss ratio, L. It has been studied for decades now, using measurements, simulations and mathematical modeling. The burst ratio is considered to be the second important loss parameter, especially in the real-time multimedia transmissions. In such applications it makes a difference whether one single packet is lost after every 20 delivered packets, or 3 packets are lost in a row after every 60 delivered packets. This effect can even be captured in strict formulas. For instance, in [3] the impairment factor I for the digitized voice has been proposed as a function of L and B, in the following form:

$$I = I_e + \frac{100(95 - I_e)L}{100L/B + R}, \tag{4}$$

where I_e and R are some constants, unimportant here, with default values $I_e = 0$ and $R = 4.3$. In the latter example with $L = \frac{1}{21} = \frac{3}{63}$, we have $I = 49.9$ for $B = 1$ and $I = 76.8$ for $B = 3$. In other words, the quality may drop by far when B increases and L remains unaltered. The reader can easily find examples, in which the overall voice quality is worse, due to increased B, even though L is decreased.

In this paper, measurements of the burst ratio at the bottleneck link of our university campus are described first. The link under study was the 1 Gbps link,

which connects the dormitory network, with the rest of the campus network. At different times of the day, 18 samples of traffic were collected, each consisting of one milion packets. The experiment started early in the morning and ended late in the evening, which enabled us to observe variable network conditions, depending on end users activities. For each of the collected traffic sample, the burst ratio was computed and discussed, as well as the accompanying loss parameters (L and $\overline{G}$). This is presented in the first part of the paper.

In the second part of the paper, theoretical values of the burst ratio are computed using a queuing model, and compared with the measured B values. The main reason of the packet loss at a bottleneck link is an overflow of the buffer at the output interface. Such interface is typically organized as a simple tail-drop queue. Therefore, it seems reasonable to use a properly parameterized tail-drop queueing model for theoretical prediction of the burst ratio.

The rest of the paper is organized as follows. In Sect. 2, the literature related to the burst ratio and loss ratio is recalled. In Sect. 3, the network topology and devices, as well as the hardware and software used in the experiment, are described. Then, the measurement process is presented in detail. Section 4 is devoted to presentation and discussion of the experiment results. Firstly, observed traffic is characterized in terms of the load, loss ratio, TCP/UDP fractions, port usage and packet size distribution. Then the burst ratio values are shown and discussed, with an accent on their relations with the observed traffic characteristics. In Sect. 5, the theoretical values of the burst ratio, computed via a queueing model, are shown and compared with the measurement results. Finally, conclusions and future work directions are gathered in Sect. 6.

2 Related Work

To the best of the authors' knowledge, there are no published measurement results, nor its analysis, of the burst ratio in a real, operating network.

The closest paper one can find is [4], where the burst ratio measurements carried out in a networking lab, using traffic from a hardware generator, are presented. Such results, obtained for artificial traffic in a local lab, have several drawbacks, when compared to the real IP traffic measurements. Firstly, propagation times are negligible in a local lab, so multiple-flow traffic with different round trip times (RTTs) in different flows cannot be mimicked. (It is well known that RTT has a great impact on the TCP behaviour.) Secondly, due to the limitations of the generator, only very simple distribution of the packet size can be used, far from the very rich structure in real IP traffic. The same concerns diversified duration of flows, port usage and link loads induced by human users and the applications they exploit.

Analytical studies of the burst ratio were based initially on Markovian loss models (see e.g. [5–9] for more information on such models). In particular, in [2] a formula for the burst ratio in a single link, represented by a two-state Markovian loss model, was derived. In [10,11], the exact and simplified formulas for B in a few concatenated links, each represented by the two-state Markovian loss

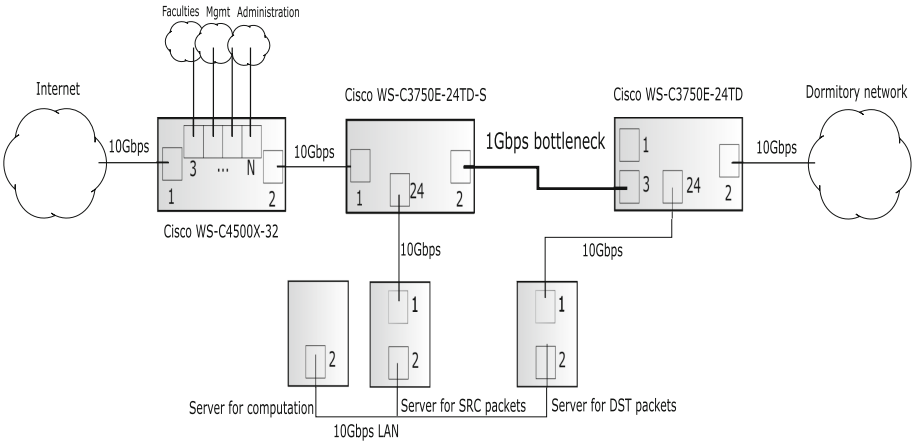

Fig. 1. Network topology and devices.

model, were obtained. In these studies, the finite-buffer queue, which constitutes the main reason of loss in the wired networking, was not modeled. Instead, it was assumed that the loss process follows some Markov chain, which was then parameterized and solved. Such approach gives the fundamental insight into the behaviour of the burst ratio, but its depth is limited by known limitations of Markov modeling – e.g. the series of losses cannot have a large variance. Recently, the burst ratio has been derived for two classic queueing models with single and batch arrivals, see [12,13], respectively. Moreover, it has been simulated for an active queue management scheme, in which the probability of dropping a packet is some function of the queue size, [14]. The best results were obtained for a convex, exponential dropping function, perhaps due to some general properties of convex functions (see e.g. [15]).

The main loss characteristic, the loss ratio, has been studied for many years now using measurements (see e.g. [16–20] and the references given there) and analysis (e.g. [21–23]).

Finally, there are also papers [24–26], in which the statistical structure of packet losses is studied using other characteristics, like the probability that in a block of m packets exactly n packets are lost, but not necessarily one after another.

3 Experiment Setup and Course

3.1 Topology, Devices and Software

The scheme of the campus network is depicted in Fig. 1.

On the left, there is an Internet connection to Cisco WS-C4500X-32. This device routes packets to different campus network segments, i.e. faculties, administration, management (on the top) and dormitories (on the right).

Then, Cisco WS-C3750E-24TD-S and Cisco WS-C3750E-24TD devices perform the distribution function in the part of the network under study.

Traffic was studied at the 1 Gbps bottleneck link depicted in the middle of Fig. 1, bound to the dormitory network on the right. The measured packet losses occurred at the buffer of output port 2 of the Cisco WS-C3750E-24TD-S device.

In order to capture, anonymize, store and analyze traffic, three high-performance servers, HP Enterprise ProLiant DL380 Gen9, were used (on the bottom of Fig. 1).

The servers were equipped with the DPDK network cards (Data Plane Development Kit), which together with the DPDK-dump application [27], enabled very efficient capturing of traffic, for further analysis.

To store the captured data in the servers, the well-known SQLite database was chosen. The data was stored in 3 separate files: source packets, destination packets and combined packets. The last one was used to compute the loss characteristics, i.e. the loss ratio and the burst ratio. Each database maintained a table with all integer fields: timestamps, IP source and destination addresses, IP IDs, IP lengths, IP protocol numbers, Layer 4 source and destination ports and sequence and acknowledgement numbers of TCP segments.

In order to take care of users' privacy, three data anonymizing mechanisms were used when collecting network traffic: Tcpmkpub [28] for anonymization of packet headers in trace files, Tcpdpriv [29] for eliminating confidential information from packets and Crypto-PAn [30–32] for anonymizing IP addresses in their traces in a prefix-preserving manner.

3.2 Course of the Experiment

Internet traffic bound to dormitory network was obtained on Cisco WS-C4500X-32 device on port 1. Then it was routed to port 1 of WS-C3750E-24TD-S device. In this device, it was duplicated using SPAN function to port 24, and from that port forwarded to the second server in Fig. 1.

At the same time, original traffic from port 1 of WS-C3750E-24TD-S was routed to output port 2, and through the 1Gbps link passed to input port 3 of WS-C3750E-25TD device. On that device, arriving traffic was duplicated again using SPAN to port 24, and then forwarded to the third server. At the same time, original traffic from port 3 of WS-C3750E-25TD was routed to destinations in the dormitory network through the output port 2.

Therefore, the second server collected traffic before losses could occur at the buffer of the bottleneck link, while the third server collected traffic thinned by these losses.

Pre-processed, anonymized traces from servers 2 and 3 were then copied to the first server, where source and destination packets were combined, what enabled calculations of all interesting characteristics, including the loss and burst ratio.

The buffer size of the interface of interest, i.e. port 2 of WS-C3750E-24TD-S device, was set to 100 packets.

The experiment was carried out in one day in the middle of the week (Tuesday), from 6:30 to 23:30. During this period, at different times of the day, 18 traces were collected, each containing one milion of packets.

4 Results

4.1 Traffic Characteristics

Two basic characteristics of collected traffic in time are depicted in Figs. 2 and 3. They are the link load, ρ, and the packet loss ratio, L, respectively. As can be seen, the load was small in the morning, 5% at 6:30, then it was growing until

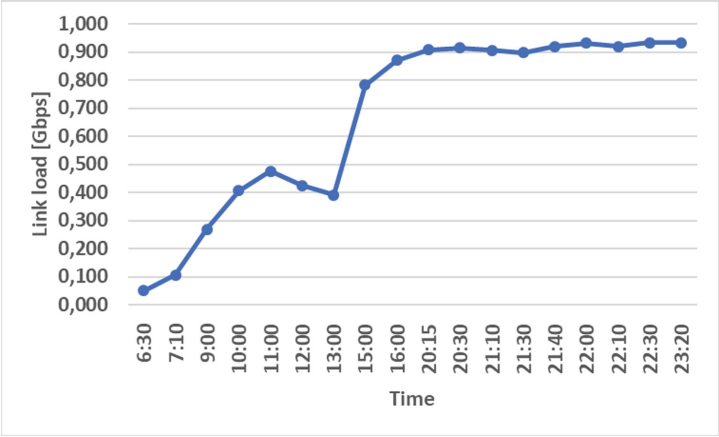

Fig. 2. Link load in time.

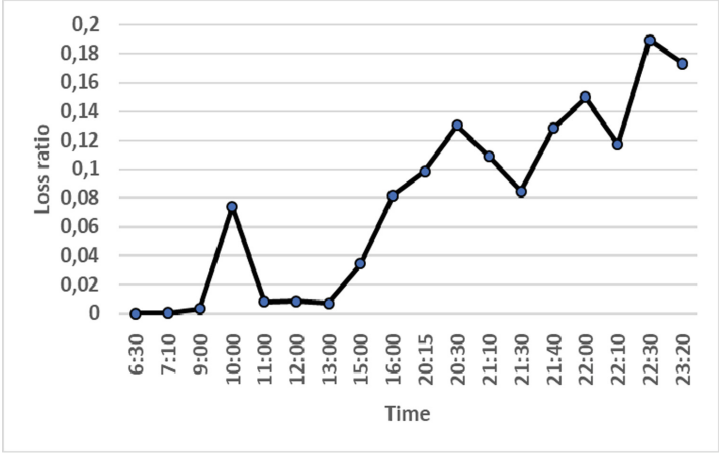

Fig. 3. Loss ratio in time.

Table 1. TOP 5 ports in TCP connections.

Time	Port1	%	Port2	%	Port3	%	Siz4	%	Port5	%
6:30	443	52.32	80	12.91	1185	1.21	8000	0.51	993	0.10
7:10	443	58.31	80	19.39	8000	0.24	9212	0.12	46792	0.04
9:00	443	55.00	80	12.43	50536	0.47	8081	0.43	1935	0.24
10:00	443	61.74	80	17.16	440	1.01	8080	0.10	8000	0.06
11:00	443	37.23	80	19.56	440	1.40	282	0.55	18331	0.07
12:00	443	52.27	80	17.66	440	0.67	8081	0.29	1935	0.13
13:00	443	54.02	80	12.33	1935	1.44	440	0.41	8000	0.07
15:00	443	43.05	80	26.70	440	0.58	1935	0.26	282	0.11
16:00	443	38.41	80	18.18	440	0.21	282	0.15	1935	0.07
20:15	443	43.01	80	20.49	440	0.17	20108	0.14	1935	0.11
20:30	443	37.32	80	29.02	8080	0.69	18073	0.11	8999	0.10
21:10	443	38.97	80	19.06	53238	0.66	1935	0.14	440	0.14
21:30	443	45.76	80	18.61	53238	2.36	8777	0.13	8080	0.10
21:40	443	42.29	80	18.06	9992	4.12	1935	0.17	8999	0.10
22:00	443	42.79	80	21.89	6905	0.24	6906	0.22	6907	0.21
22:10	443	45.12	80	19.64	81	0.25	1935	0.25	8777	0.11
22:30	443	31.32	80	29.06	1935	0.12	8080	0.09	8777	0.09
23:20	443	31.74	80	31.65	81	0.39	8080	0.30	440	0.17

20:00, when it reached about 90%, and then it remained high until the end of the experiment. Similarly, the loss ratio varied from a very small 0.0025% in the morning, up to high 18.9% at 22:30 in the evening. Such characteristics were to be expected - they follow the typical human activity pattern.

In Tables 1 and 2, the five mostly used ports are presented for TCP and UDP transmissions. The largest volume of traffic was generated between the Web browsers and servers, using encrypted (port 443) or plain-text (port 80) communication.

There were also their backup or non-standard versions (port 8080) and the onion routing service for anonymous communication (port 81). The Web browsing would not be possible without the Domain Name System queries (port 53). A group of ports was devoted to direct multimedia transmissions, namely ports 8081, 8777, 8999 – for Apple's Web Service and iTunes streams, port 440 – SGP protocol used in VoIP transmissions, port 1071 – BSQUARE-VoIP, port 5100 – for Mac OS X camera and scanner sharing. Other part of traffic consisted of connections to databases/databases management systems (ports 8000 and 9212) and filesystem access (ports 50536 and 53238 – Apple's Xsan filesystem access). There were also games involved (port 1119 – Battle.net chat/game protocol) and usage of Adobe Flash or similar application with Real Time Messaging Protocol, on port 1935. A part of the throughput was used for downloading files with

Table 2. TOP 5 ports in UDP flows.

Time	Port1	%	Port2	%	Port3	%	Siz4	%	Port5	%
6:30	443	31.76	55743	0.23	24874	0.17	5100	0.14	53	0.05
7:10	443	20.82	24874	0.39	55743	0.10	5100	0.07	1119	0.07
9:00	443	28.86	22203	1.40	6881	0.18	27050	0.12	27053	0.09
10:00	443	17.99	22203	0.96	6881	0.33	27005	0.09	52437	0.05
11:00	443	24.32	47692	1.86	20588	1.16	22535	1.01	61266	0.97
12:00	443	26.72	16617	0.35	6881	0.32	22203	0.21	24934	0.19
13:00	443	28.75	16617	0.30	50667	0.21	6881	0.21	24934	0.20
15:00	443	27.64	11088	0.87	10001	0.03	1071	0.03	1119	0.03
16:00	443	42.10	5749	0.15	24874	0.04	26507	0.03	26523	0.03
20:15	443	30.01	8999	0.92	17501	0.51	53139	0.29	11088	0.25
20:30	443	29.67	17501	0.51	8999	0.40	30808	0.27	25231	0.09
21:10	443	37.51	8999	0.46	25231	0.31	51193	0.20	16394	0.12
21:30	443	29.53	8999	0.48	25231	0.33	51193	0.24	16394	0.16
21:40	443	32.39	8999	0.33	52909	0.30	51193	0.16	51896	0.08
22:00	443	31.17	8999	0.23	25231	0.17	6881	0.13	51193	0.09
22:10	443	30.08	8999	0.43	25231	0.26	51193	0.11	51413	0.10
22:30	443	35.74	8999	0.14	56143	0.14	51896	0.08	25231	0.08
23:20	443	32.28	12345	1.11	16393	0.14	23952	0.09	33480	0.09

BitTorrent applications (ports 6881, 6905, 6906, 6907). Finally, a fraction of traffic consisted of email transfers (port 993).

In Table 3, the distribution of the packet size in every collected trace is shown. As can be observed, large packets (over 1400 bytes) predominated at every time of the day – their share was always above 50%. The second important group consisted of packets in range 1201–1400 bytes, while the third – of small packets, up to 200 bytes. It is worth noticing, that the size distribution varied significantly during the day. For instance, the fraction of small packets ($\leq$200 bytes) varied from 17% at 6:30 to 6% at 18:00. The fraction of large packets (>1400 bytes) varied from 72% at 10:00 to 50% at 21:10.

During the measurements, a surprisingly high percentage of UDP traffic was observed, from 20% up to 43%, at different times of the day. One possible explanation of this phenomenon is that a part of UDP traffic might have been in fact HTTPS traffic on port 443. This is possible when replacing multiple TCP connections with one multiplexed UDP flow by means of QUIC protocol.

4.2 Measured Burst Ratio

The final results of measurements are shown in Table 4. The burst ratio, B, and the average length of the series of losses ($\overline{G}$) are given for every collected trace.

Table 3. Distributions of the packet size [%]. Packet sizes are grouped in ranges: 0–200 bytes, 201–400 bytes, 401–600 bytes, etc.

Time	0–200	201–400	401–600	601–800	801–1000	1001–1200	1201–1400	1401–1600
06:30	16.65	1.48	0.49	0.37	0.41	0.30	27.63	52.67
07:10	16.01	1.26	0.52	0.32	0.37	0.27	21.32	59.94
09:00	12.06	1.02	1.24	0.38	0.48	0.37	27.32	57.13
10:00	7.94	1.14	0.58	0.28	0.30	0.25	17.50	72.01
11:00	8.70	1.25	1.02	0.36	0.39	0.30	23.22	64.75
12:00	8.94	1.59	1.04	0.60	0.63	0.47	25.17	61.55
13:00	16.82	1.62	0.91	0.50	0.60	0.53	26.27	52.75
15:00	6.88	1.00	0.77	0.28	0.34	0.23	27.13	63.36
16:00	6.46	0.97	0.77	0.33	0.42	0.27	40.68	50.11
20:15	10.20	1.56	0.72	0.66	0.60	0.43	28.63	57.20
20:30	9.38	1.51	0.79	0.47	0.46	0.58	28.37	58.43
21:10	10.10	1.63	1.24	0.55	0.57	0.56	34.90	50.45
21:30	12.55	1.67	0.92	0.57	0.56	0.52	27.10	56.11
21:40	11.17	1.81	1.01	0.55	0.51	0.41	30.04	54.51
22:00	10.48	1.57	0.94	0.46	0.47	0.46	30.83	54.78
22:10	11.92	1.67	1.49	0.59	0.57	0.46	28.32	54.98
22:30	10.35	1.44	0.98	0.46	0.48	0.58	33.52	52.19
23:20	9.99	1.24	1.07	0.52	0.48	0.40	30.68	55.62

They are accompanied by the corresponding link load and loss ratio. Additionally, the burst ratio evolution in time is depicted in Fig. 4.

Firstly, the fact that the burst ratio is always significantly greater than 1 can be observed. This means that the losses in real IP traffic do have tendency to group together.

This result is consistent with the previous theoretical results of [12], based on a queueing model, and the results of [4], obtained via artificially generated traffic.

It is important to note, that greater than 1 values of B were observed for a wide variety of loss ratios, from 0.0025% to 18.9% and for a wide variety of loads, from 5% to 93%.

It is rather hard to single out one traffic parameter, which has the greatest influence on B. At the first look at Table 4, it seems that B grows with ρ. But this is not really the truth – almost the same B was obtained at 9:00 and 22:30, but the load was very different at those times: 27% versus 93%. The loss ratio is also not a good candidate - almost the same B was obtained for very different values of L: 0.3% and 18.9%.

These results are also consistent with Theorem 1 of the next section, in which the dependence of B on ρ and L is weak. From the same theorem it follows that the burst ratio depends strongly on the distribution of the service time (here the packet size), which in our experiment varied significantly in time (see Table 3).

Table 4. The measured burst ratio versus time, the link load and the loss ratio.

Time	ρ	L	$\overline{G}$	B
6:30	0.050	0.00251×10^{-2}	1.30769	**1.20921**
7:10	0.107	0.01664×10^{-2}	1.20750	**1.20730**
9:00	0.269	0.31523×10^{-2}	1.31117	**1.30703**
10:00	0.407	7.40980×10^{-2}	1.46823	**1.35944**
11:00	0.476	0.78170×10^{-2}	1.40475	**1.39377**
12:00	0.426	0.82825×10^{-2}	1.44072	**1.42879**
13:00	0.392	0.67354×10^{-2}	1.46910	**1.45921**
15:00	0.782	3.42986×10^{-2}	1.50438	**1.45278**
16:00	0.871	8.16498×10^{-2}	1.52892	**1.40408**
20:15	0.908	9.83570×10^{-2}	1.56778	**1.41358**
20:30	0.915	13.0220×10^{-2}	1.58598	**1.37945**
21:10	0.907	10.8499×10^{-2}	1.58065	**1.40915**
21:30	0.897	8.45893×10^{-2}	1.59269	**1.45796**
21:40	0.920	12.8154×10^{-2}	1.60651	**1.40063**
22:00	0.932	14.9776×10^{-2}	1.60316	**1.36304**
22:10	0.919	11.7032×10^{-2}	1.61177	**1.42314**
22:30	0.934	18.9270×10^{-2}	1.61415	**1.30864**
23:20	0.933	17.3317×10^{-2}	1.59776	**1.32084**

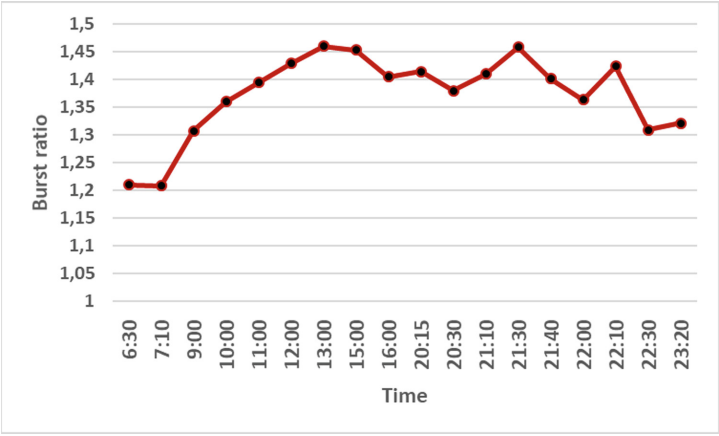

Fig. 4. Burst ratio in time.

From the measurements it follows also that B may depend on other characteristics of IP traffic. For instance, at 22:00, a smaller B than at other times with similar L and ρ were observed. A the same time, an increase in BitTorrent

application usage (ports 6905, 6906, 6907) was noticed. On the other hand, during the measurement at 21:10, a high B was observed. At this time, plenty of filesystem accesses (on port 53238) were performed by end users. At 9:30 and 22:10, traffic on port 443 was consuming a large fraction of the total bandwidth and B assumed very high values. It is likely that usage of these and other specific applications has an impact on the traffic autocorrelation and the packet size distribution, which on the other hand influence the burst ratio.

5 Comparison with the Theoretical Burst Ratio

It is intuitively clear, that the simple drop-tail buffering mechanism, used commonly in networking devices, can cause clustering of packet losses. Namely, when the buffer for packets is full, a few newly arriving packets may be lost in a row, before some space in the buffer becomes available again. This effect can be described precisely using a solution of the queueing model of the buffering mechanism in a device.

Herein, the model of the queue with general distribution of the service time, given by distribution function $F(t)$, finite buffer of size N and Poisson arrivals of rate λ, is used. The greatest advantage of this model is that it enables calculations for an arbitrary service time distribution (the packet size distribution), which has a deep impact on the value of B. Therefore, using this model the burst ratios for all packet size distributions given in Table 3 can be computed. Obviously, all the loads from Table 4 can be incorporated as well. In [12], the following two theorems on the burst ratio in the considered model were proven. (They were obtained using the potential method, used previously in analysis of the loss ratio and other loss characteristics of queueing systems, see e.g. [22,33]). The first theorem enables calculation of the exact value of B as a function of the parameters of the queueing model.

Theorem 1. *In the queueing system with the general service time distribution the burst ratio equals:*

$$B = \frac{\sum_{k=0}^{N-1} r_k}{[1-g(0)](1+\rho\sum_{k=0}^{N-1} r_k)} \sum_{l=1}^{\infty} lg(l), \tag{5}$$

where

$$g(l) = \frac{\sum_{k=1}^{N} R_{N-k}a_{k+l} - \sum_{k=1}^{N-1} R_{N-1-k}a_{k+l}}{\sum_{k=0}^{N} R_{N-k}a_k - \sum_{k=0}^{N-1} R_{N-1-k}a_k}, \tag{6}$$

$$R_0 = 0, \qquad R_1 = \frac{1}{a_0}, \tag{7}$$

$$R_{k+1} = R_1(R_k - \sum_{i=0}^{k} a_{i+1}R_{k-i}), \quad k = 1, 2, \ldots, \tag{8}$$

$$r_0 = 1, \qquad r_1 = \frac{1}{a_0} - 1, \tag{9}$$

$$r_{k+1} = \frac{1}{a_0}(r_k - \sum_{i=0}^{k-1} a_{i+1} r_{k-i} - a_k), \quad k = 1, 2, \ldots, \tag{10}$$

and

$$a_k = \int_0^\infty \frac{e^{-\lambda u}(\lambda u)^k}{k!} dF(u), \quad k = 0, 1, 2, \ldots. \tag{11}$$

Here ρ denotes the load of the queueing system, i.e.:

$$\rho = \lambda \int_0^\infty t dF(t). \tag{12}$$

The second theorem enables calculation of the limiting value of B, as the buffer size grows to infinity. As the convergence is rather quick, it can be also used to obtain approximate values of B, even for relatively small buffer sizes.

Theorem 2. *In the queueing system with the general service time distribution the limiting burst ratio equals:*

$$\lim_{N\to\infty} B(N) = \frac{\min\{1, \rho^{-1}\}}{1 - h(0)} \sum_{l=1}^{\infty} l h(l), \tag{13}$$

where

$$h(l) = \frac{1}{x_0^{l+1}} f(\lambda - \lambda x_0) - \frac{1}{x_0^{l+1}} \sum_{j=0}^{l} x_0^j a_j, \tag{14}$$

$$f(s) = \int_0^\infty e^{-st} dF(t), \tag{15}$$

and x_0 is a positive and not equal to 1 solution of the equation

$$f(\lambda - \lambda x) = x. \tag{16}$$

In the third column of Table 5, the theoretical burst ratios are presented, computed for $N = 100$, ρ values as in Table 4 and the packet size distributions as in Table 3. They were obtained using Theorem 1, but they could have been obtained as well from Theorem 2 (for $N = 100$, Theorem 2 gives results very close to Theorem 1). For comparison, in the second column of Table 5, the measured values of the burst ratio are recalled. Finally, in the last column of Table 5, the relative error is calculated as a percentage of the measured burst ratio.

As can be seen, after 15:00, when the load is high, the theoretical burst ratio agrees very well with its measured counterpart – the relative error is around a few percent. The results for earlier traces, when the load is low or moderate, are worse – the relative error is in the range of 11–22%. This error is most likely caused by the much more complicated statistical structure of real IP traffic, compared with the Poisson traffic in the model. Apparently, this difference is not very important, when the load is high, but gets the more important, the lower the load is.

Table 5. Measured burst ratio, B_m, theoretical burst ratio, B_t, and the relative error.

Time	B_m	B_t	Error, $\frac{\lvert Bt-B_m\rvert}{B_m}$
06:30	1.20921	1.01191	16.3%
07:10	1.20730	1.02895	14.7%
09:00	1.30703	1.08371	17.0%
10:00	1.35944	1.13116	16.7%
11:00	1.39377	1.16085	16.7%
12:00	1.42879	1.14167	20.1%
13:00	1.45921	1.13955	21.9%
15:00	1.45278	1.28959	11.2%
16:00	1.40408	1.33042	5.2%
20:15	1.41358	1.36465	3.4%
20:30	1.37945	1.36449	1.0%
21:10	1.40915	1.36529	3.1%
21:30	1.45796	1.36938	6.0%
21:40	1.40063	1.37616	1.7%
22:00	1.36304	1.37827	1.1%
22:10	1.42314	1.37968	3.0%
22:30	1.30864	1.37855	5.3%
23:20	1.32084	1.37601	4.1%

It is also worth noticing, that the theoretical value of B is typically smaller than its measured counterpart. Therefore, it can be used as a lower bound for the real value.

As discussed before, a higher usage of some specific applications may influence the traffic autocorrelation and, eventually, B. Therefore, to further decrease the error of the queueing model, it might be necessary to incorporate the traffic autocorrelation into the model.

This supposition can be further strengthened by studying the autocorrelation function. In Fig. 5, the autocorrelation function of packet interarrival times is presented for the trace collected at 15:00. (Similar results were obtained for other times). As can be seen in the figure, the autocorrelation function decreases very slowly, and is still non-negligible for the lag of 10^4. Its is well known, that such a long-range autocorrelation of the arrival process may significantly influence the queueing performance. This constitutes, of course, a strong motivation for deriving the burst ratio in models with more complicated traffic, incorporating the autocorrelation of interarrival times.

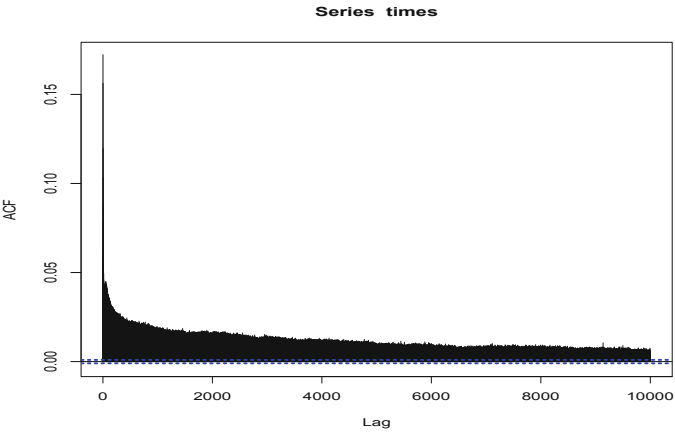

Fig. 5. Autocorrelation function of interarrival times for the trace collected at 15:00.

6 Conclusions and Future Work

In this paper, the results of measurements of the burst ratio, a parameter expressing the tendency of packet losses to occur in series, rather than as separate objects, were presented. Among other things, the study was motivated by the fact that clustering of packets losses is especially bad for multimedia transmissions. The measurements were conducted in real IP traffic, in an operational network.

Firstly, the obtained burst ratio was significantly higher than 1 in all measurements. This means that losses in real IP traffic do have tendency to group together. This observation is consistent with the previous results obtained in a queueing model and in a networking lab, on artificially generated traffic.

Secondly, the measurements showed that the dependence of the burst ratio on the load and the loss ratio is small. On the other hand, they demonstrated the possible dependence of the burst ratio on other traffic characteristics, including the application usage and the resulting traffic autocorrelation.

Thirdly, when comparing the measured burst ratio with its theoretical counterpart computed for the queueing model with Poisson arrivals, a relatively high accuracy of the model was observed for high load values. For a low or moderate load, the relative error was higher, with the maximum of 22%. As observed, the theoretical value of B for the Poisson traffic model is typically smaller than its real counterpart, thus it can be used as a lower bound of the real value.

To decrease further the discrepancy between the queueing model and measurements, the autocorrelation of the interarrival times can be incorporated into the model, e.g. using MMPP [34] or BMAP [35,36] processes.

References

1. Mongay Batalla, J., et al.: Adaptive video streaming: rate and buffer on the track of minimum re-buffering. IEEE J. Sel. Areas Commun. **34**(8), 2154–2167 (2016)
2. McGowan, J.W.: Burst ratio: a measure of bursty loss on packet-based networks, 16 2005. US Patent (2005)
3. ITU-T Recommendation G.113: Transmission impairments due to speech processing. Technical report (2007)
4. Samociuk, D., Chydzinski, A., Barczyk, M.: Experimental measurements of the packet burst ratio parameter. In: Kozielski, S., Mrozek, D., Kasprowski, P., Małysiak-Mrozek, B., Kostrzewa, D. (eds.) BDAS 2018. CCIS, vol. 928, pp. 455–466. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-99987-6_35
5. Sanneck H.A., Carle, G.: Framework model for packet loss metrics based on loss runlengths. In: Proceedings of Multimedia Computing and Networking, pp. 1–11 (2000)
6. Jiang, W., Schulzrinne, H.: Modeling of packet loss and delay and their effect on real-time multimedia service quality. In: Proceedings of NOSSDAV, pp. 1–10 (2000)
7. Veeraraghavan, M., Cocker, N., Moors, T.: Support of voice services in IEEE 802.11 wireless LANs. In: Proceedings of IEEE INFOCOM, vol. 1, pp. 488–497 (2001)
8. Hasslinger, G., Hohlfeld, O.: The Gilbert-Elliott model for packet loss in real time services on the Internet. In: Proceedings of Measuring, Modelling and Evaluation of Computer and Communication Systems Conference, pp. 1–15 (2008)
9. Clark, A.: Modeling the effects of burst packet loss and recency on subjective voice quality. In: Proceedings of Internet Telephony Workshop, pp. 123–127 (2001)
10. Rachwalski, J., Papir, Z.: Burst ratio in concatenated Markov-based channels. J. Telecommun. Inf. Technol. **1**, 3–9 (2014)
11. Rachwalski, J., Papir, Z.: Analysis of burst ratio in concatenated channels. J. Telecommun. Inf. Technol. **4**, 65–73 (2015)
12. Chydzinski, A., Samociuk, D.: Burst ratio in a single-server queue. Telecommun. Syst. **70**, 263–276 (2019)
13. Chydzinski, A., Samociuk, D., Adamczyk, B.: Burst ratio in the finite-buffer queue with batch Poisson arrivals. Appl. Math. Comput. **330**, 225–238 (2018)
14. Samociuk, D., Chydzinski, A.: On the impact of the dropping function on the packet queueing performance. In: Proceedings of International Convention on Information and Communication Technology, Electronics and Microelectronics, pp. 473–478 (2018)
15. Smolka, B., et al.: New filtering technique for the impulsive noise reduction in color images. Math. Probl. Eng. **2004**(1), 79–91 (2004)
16. Benko, P., Veres, A.: A passive method for estimating end-to-end TCP packet loss. In: Proceedings of IEEE GLOBECOM 2002, pp. 2609–2613 (2002)
17. Bolot, J.: End-to-end packet delay and loss behavior in the Internet. In: Proceedings of ACM SIGCOMM 1993, pp. 289–298 (1993)
18. Coates, M., Nowak, R.: Network loss inference using unicast end-to-end measurement. In: Proceedings of ITC Conference on IP Traffic, Measurement and Modeling, pp. 282–289 (2000)
19. Duffield, N., Presti, F.L., Paxson, V., Towsley, D.: Inferring link loss using striped unicast probes. In: Proceedings of IEEE INFOCOM 2001, pp. 915–923 (2001)
20. Sommers, J., Barford, P., Duffield, N., Ron, A.: Improving accuracy in end-to-end packet loss measurement. ACM SIGCOMM Comput. Commun. Rev. **35**(4), 157–168 (2005)

21. Takagi, H.: Queueing Analysis - Finite Systems. North-Holland, Amsterdam (1993)
22. Chydzinski, A., Wojcicki, R., Hryn, G.: On the number of losses in an MMPP queue. In: Koucheryavy, Y., Harju, J., Sayenko, A. (eds.) NEW2AN 2007. LNCS, vol. 4712, pp. 38–48. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-74833-5_4
23. Chydzinski, A., Mrozowski, P.: Queues with dropping functions and general arrival processes. PLoS One **11**(3), e0150702 (2016)
24. Yu, X., Modestino, J.W., Tian, X.: The accuracy of Gilbert models in predicting packet-loss statistics for a single-multiplexer network model. In: Proceedings of IEEE INFOCOM 2005, pp. 2602–2612 (2005)
25. Cidon, I., Khamisy, A., Sidi, M.: Analysis of packet loss processes in high-speed networks. IEEE Trans. Inf. Theory **39**(1), 98–108 (1993)
26. Bratiychuk, M., Chydzinski, A.: On the loss process in a batch arrival queue. Appl. Math. Model. **33**(9), 3565–3577 (2009)
27. DPDK-Dump application for capturing traffic using DPDK. https://github.com/marty90/DPDK-Dump
28. TCPmkpub. http://www.icir.org/enterprise-tracing/tcpmkpub.html
29. TCPdpriv. http://ita.ee.lbl.gov/html/contrib/tcpdpriv.html
30. Jinliang, F., Jun, X., Mostafa, H.A.: Prefix-preserving IP address anonymization. Comput. Netw. **46**(2), 253–272 (2004)
31. Jinliang, F., Jun, X., Mostafa, H.A.: On the design and performance of prefix-preserving IP traffic trace anonymization. In: Proceedings of ACM SIGCOMM Internet Measurement Workshop, San Francisco (2001)
32. Jinliang, F., Jun, X., Mostafa, H.A.: Prefix-preserving IP address anonymization: measurement-based security evaluation and a new cryptography-based scheme. In: Proceedings of IEEE International Conference on Network Protocols, Paris (2002)
33. Chydzinski, A.: Duration of the buffer overflow period in a batch arrival queue. Perform. Eval. **63**(4–5), 493–508 (2006)
34. Fischer, W., Meier-Hellstern, K.: The Markov-modulated Poisson process (MMPP) cookbook. Perform. Eval. **18**(2), 149–171 (1992)
35. Lucantoni, D.M.: New results on the single server queue with a batch Markovian arrival process. Commun. Stat. Stoch. Models **7**(1), 1–46 (1991)
36. Chydzinski, A.: Queue size in a BMAP queue with finite buffer. In: Koucheryavy, Y., Harju, J., Iversen, V.B. (eds.) NEW2AN 2006. LNCS, vol. 4003, pp. 200–210. Springer, Heidelberg (2006). https://doi.org/10.1007/11759355_20