Scalable in-network rate monitoring

Per Kreuger and Rebecca Steinert
Decisions, Networks and Analytics (DNA) lab.
Swedish Institute of Computer Science (SICS Swedish ICT AB),
SE-164 29 Kista, Sweden.
Email: {piak, rebste}@sics.se

***Abstract*—We propose a highly scalable statistical method for modelling the monitored traffic rate in a network node and suggest a simple method for detecting increased risk of congestion at different monitoring time scales. The approach is based on parameter estimation of a lognormal distribution using the method of moments. The proposed method is computationally efficient and requires only two counters for updating the parameter estimates between consecutive inspections. Evaluation using a naive congestion detector with a success rate of over 98% indicates that our model can be used to detect episodes of high congestion risk at 0.3 s using estimates captured at 5 m intervals.**

***Keywords*—probabilistic management; performance monitoring; statistical traffic analysis; link utilization modelling; congestion detection; in-network rate monitoring**

I. Introduction

Current development towards software-defined networking (SDN) and network function virtualization (NFV) opens the door for more dynamic management of storage, compute, and networking resources compared to traditional management where the network setup is assumed to be static. However, dynamic service provisioning and performance management rely on the ability to perform scalable and resource-efficient measurements which can be used in a proactive manner.

We propose a measurement approach based on the simple and general idea of high rate but low complexity local updates and lower and adaptive rate analysis on locally produced statistics, which can then be further distributed and collected as needed. The method is based on the use of two counters for storing the first and second statistical moments of each monitored rate. This provides richer information about the observed traffic rates when used for parameter estimation of suitable distributions, and enables robust detection of performance degradations and resource-efficient dissemination of the results. Updating these simple statistics locally and at high rates allow for accurately capturing important aspects the traffic behavior with significantly lower overhead than in a centralized setting. By adapting the rate at which the counters are queried, we can achieve flexible high quality monitoring without the cost of constant high rate sampling. Figure 1 shows a schematic rate monitoring architecture of this kind.

Experimental results based on real-world traffic rates indicate that the lognormal distribution provides the best fit for link level aggregates, and that persistent congestion symptoms at 0.3 s levels can be detected with our method at a monitoring time scale of 5 m intervals. The method may be used for resource- and fault management purposes, and has been developed as a part of the UNIFY (https://www.fp7-unify.eu) project [1], [2], which is aimed at developing an automated and flexible service creation architecture for dynamic service chaining via Cloud virtualization techniques and SDN.

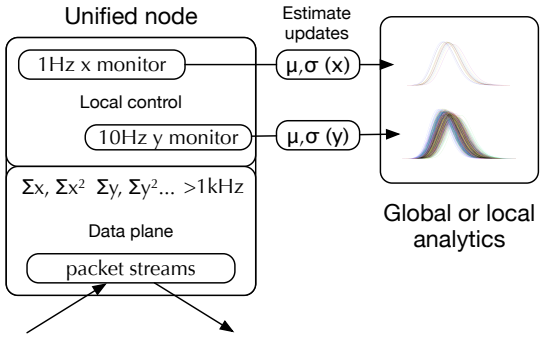

Fig. 1: Overview of rate monitoring architecture with high rate local counter updates and two lower rate estimate updates

A. Related Work

Common network monitoring tools include SNMP [3] polling for gathering statistics at regular intervals. A more advanced tool is RMON [4] which enables conditional monitoring and reporting, compared to SNMP. Other well-known tools are NetFlow [5] and sFlow [6]. NetFlow-enabled routers collect statistics on IP-traffic data, which are forwarded to a server for analysis. Such data include e.g. source and destination IP addresses, ports, type of service, packet and byte counts, timestamps, protocol flags, and routing information. sFlow is a network monitoring protocol that uses random sampling of packets (matching the headers of one or several flows) and scheduled sampling of counters. An example of a link utilization monitoring tools for SDN is FlowSense [7], based on existing information in OpenFlow control messages derived upon initiation and removal of flows. Many of the aforementioned methods are practically limited, as they often rely on external processing and analysis of the measurements, which we address by performing rate estimation locally.

In previous studies of the statistical characteristics of network traffic, it is often suggested that heavy-tailed distributions, such as lognormal or Pareto distributions can be used to model e.g. packet rates, arrival times, throughput, etc [8], [9]. Papagiannaki et. al. [10] have analysed packet delay and link utilization for the purpose of detecting congestion episodes at varying utilization levels and measurement intervals. Our method exhibits some similarities to the work in [10], but differs mainly in our focus on rate estimation at different time scales and probabilistic congestion detection for quick isolation of potential bottlenecks.

II. Approach

In current practice, flow aggregate counters are inspected at regular time intervals followed by forwarding observed statistics for analysis to a monitoring station. The use of raw packet and byte counters for each monitored flow aggregate only allows for obtaining the average rates between two consecutive inspections of the counters, and the time resolution of the estimates reflects the counter inspection rate, rather than the counter update frequency. The level of analysis possible in this setting is limited both by the restrictions to averages and the communication overhead when forwarding monitoring statistics. This situation severely limits the operator's ability to effectively manage resources and handle performance degradations in a timely manner.

Essential to the proposed approach is the computational efficiency of the high rate operations, as well as the choice of a probability distribution that is likely to provide the best fit to the observed data. For the type of rate monitoring data that we are modelling, the lognormal distribution has in general been observed to fit very well [8], [9], [11], which agrees with our own observations. For estimating the parameters of the lognormal distribution we employ a method-of-moments (MoM) approach, which produce sufficiently accurate parameter estimates with low demands on memory consumption and computational capacity, compared to a maximum likelihood estimator (MLE). The MoM-approach requires only storage of the first two statistical moments of the observed data, i.e. two counters for storing the sum and sum of squared observations. Computationally, only simple arithmetic is required for the update.

A. Model validation

To develop our model and verify our model assumptions, we have studied traffic traces recorded as part of a measurement project performed in collaboration with TeliaSonera. Bit and packet rates were captured at 1kHz on 2 links in their aggregation network in Sweden. One was a peripheral 1 Gb/s link with a moderate degree of aggregation, which was running close to overload at busy hour periods. This link was measured just before a capacity upgrade to 10 Gb/s took place, but despite the high load, no packet loss was observed. The second link was a 10 Gb/s metro level link with a high degree of aggregation. Both links where chosen to give a representative mix of home and business customers, connected through both DSL and fiber access technologies.

The observed rates vary widely on both links over the day with shifting usage patterns. Figure 2 shows a scatter plot of the raw captured data, and a histogram over the empirical bit rate distribution over 24 hours on the 10Gb/s link. This distribution is clearly multi modal, so no simple parametric distribution will fit it well. On shorter time scales however, lognormal distributions do fit the observed data very nicely. Figure 3 shows examples on two different time scales and load levels. Similar patterns appear in packet rates and on the 1Gb/s link, despite lower levels of aggregation. Our main focus is on even shorter time scales, but the simple lognormal model appears to capture a wide range of time scales very well. Figure 4 shows fits for Mb/s packet rates on two randomly selected fragments of durations of 5 m and 0.3 s and on the 1Gb/s link.

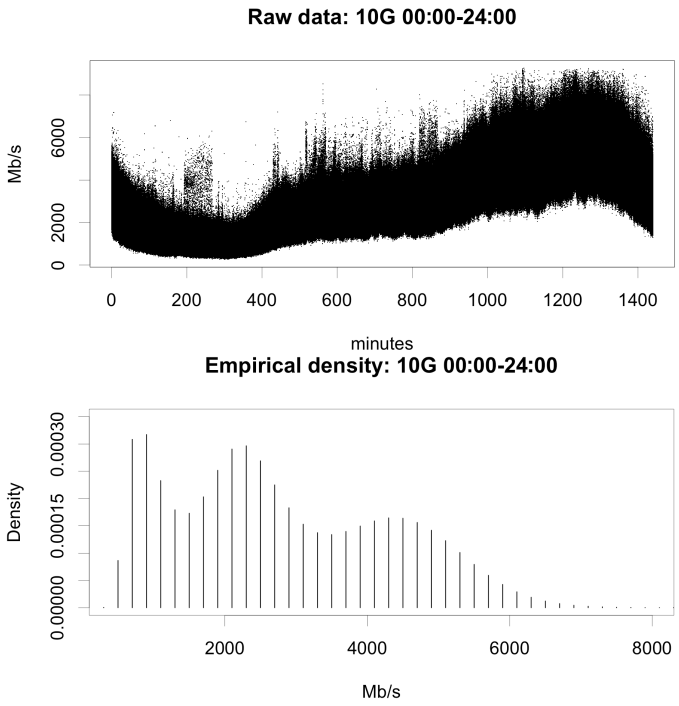

Fig. 2: Recorded rates (in Mb/s) as time series (upper) and rate distribution (lower) over 24 h on 10Gb/s link.

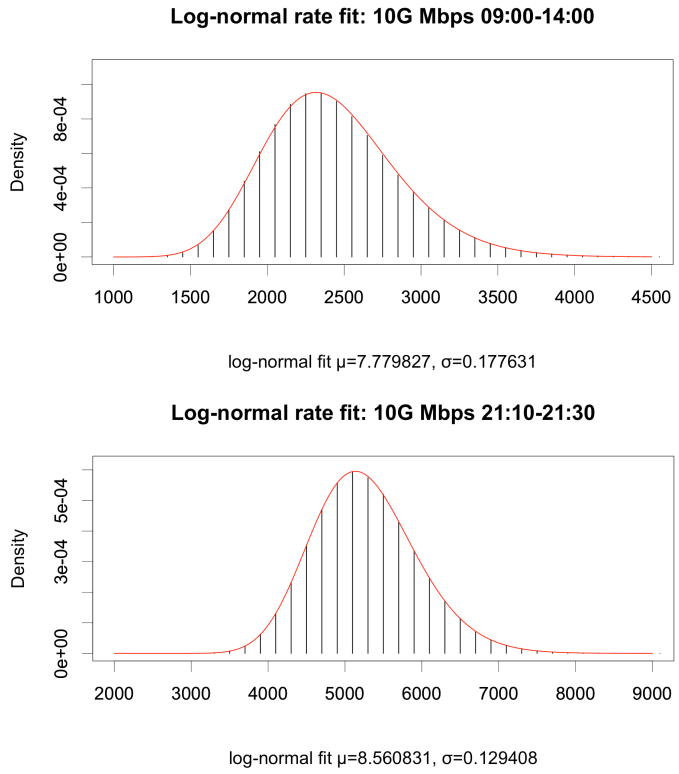

Fig. 3: Lognormal fits over 5 h and 20 m periods on a 10Gb/s link.

B. Estimation model

The most common way of estimating the parameters of lognormal distributions is maximum likelihood (ML) which has the closed form: $\hat{\mu} = \frac{\sum_i \ln x_i}{n}$, $\hat{\sigma}^2 = \frac{\sum_i (\ln x_i - \hat{\mu})^2}{n}$. Using

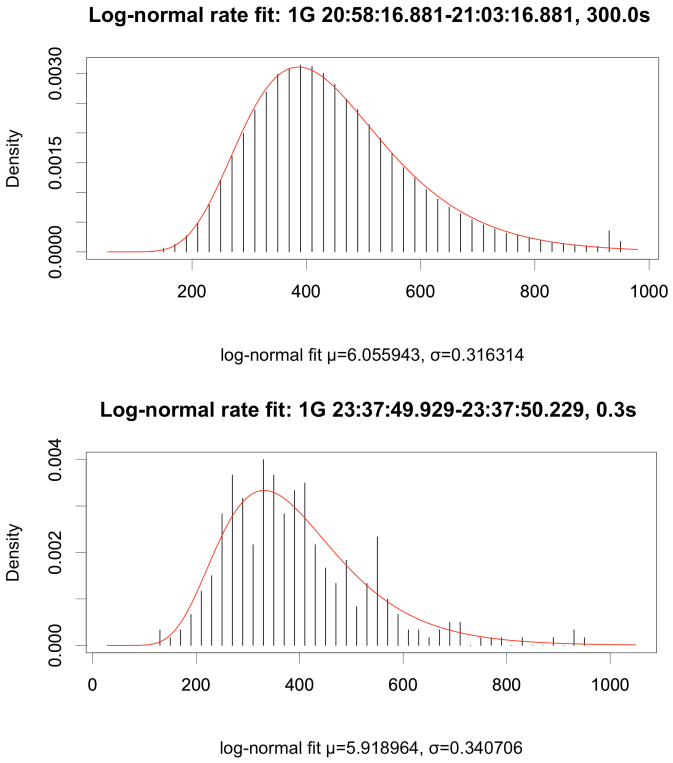

Fig. 4: Lognormal fits over intervals of 5 m and and 0.3 s durations.

this method for rate monitoring requires storing the sum of the logarithms and squared logarithms of each increment on the monitored stream.

The method of moment (MoM) however, offers an alternative parameter estimate:

$$\begin{cases} \hat{\mu} = \ln M - \frac{1}{2}\hat{\sigma}^2 \\ \hat{\sigma}^2 = \ln\left(1 + \frac{V}{M^2}\right) \end{cases} \quad (1)$$

where $M = \frac{\sum_i x_i}{n}$, and $V = \frac{\sum_i x_i^2}{n} - M^2$, the sample mean and variance, respectively, and n is the number of counter updates performed during the estimation interval.

C. Estimates at different time scales

One of the most important points of estimating the rate distributions based on simple statistics produced at high rate, is that we can query those statistics at significantly lower rates, and still obtain reasonable estimates of the underlying behavior. To illustrate how the proposed method scales with different query rates, we have investigated selected intervals of the data and produced estimates at several time scales. Table I shows the scale variance over estimates of all the subcomponents at an estimation duration covered by an estimate over a longer interval. The tables were produced by randomly selecting an interval with a length corresponding to the longer estimation duration, e.g. 5 m, and calculating the standard deviation (and coefficient of variation) of estimates produced at shorter intervals, i.e. ten 30 s intervals, and a hundred 3 s intervals. Each reported result is the mean of 1000 random selections.

TABLE I: Standard deviation (and variation coefficient) of the Mb/s 99th percentile for subdivisions of one time scale into another.

1Gb/s link	30 s	3 s	0.3 s
5 m	24.8 (0.030)	39.5 (0.048)	61.8 (0.076)
30 s		31.8 (0.039)	56.6 (0.070)
3 s			48.2 (0.060)

10Gb/s link	30 s	3 s	0.3 s
5 m	86.5 (0.026)	135.1 (0.040)	202.5 (0.060)
30 s		107.8 (0.032)	184.1 (0.054)
3 s			153.9 (0.045)

In the most extreme case, for the 10Gb/s link, we should be prepared for an error on the order of a few hundred Mb/s in the value of the 99th percentile when approximating the behaviour at 0.3 s using estimates obtained from a 5 m interval.

III. Application of the Model

There are numerous potential applications of our modelling approach once the parameter estimates have been obtained. In this section we will outline and demonstrate a probabilistic approach to dynamic monitoring at different time scales based on the risk of observing episodes of link congestion.

Standard practice for identifying increased bandwidth consumption is based on low-frequency counter inspections and reporting when the average exceeds a fixed threshold, or by manual inspection. An example of traffic rate averages from the 1Gb/s link over consecutive 5 m intervals over the last 9 h of a single day is shown as the top part of figure 5. The drawback is that these low resolution averages are far below the link capacity and any deterministic and strict threshold on such a measure is at best a rule of thumb, and will either miss real congestion events, or is likely to generate false alarms, and possibly both.

An alternative method is to exploit the cumulative density function (CDF) from the estimated parameters, and assess the risk of congestion at different percentiles. The lower two graphs of figure 5 shows the risk of exceeding the link capacity given an estimate at each consecutive interval, query rates of 1/5 m and 1/0.3 s. Compared to the top graph in the figure, we can in the lower two clearly observe episodes of increased risk of micro-congestions developing over time, and while exceeding the capacity over a 5 minute period is (unsurprisingly) much less likely than doing so over a single 0.3 s period, we can also see that even at 5 m resolution we have clear indications of the trends we see in the more highly resolved estimates. We argue that a robust and low complexity prediction of high link utilization, even for links with moderate degree of aggregation can be based on rate estimates of the proposed type. Although the complete description of such a mechanism is out of scope for the present paper, we have evaluated a simple detection mechanism based on monitoring the risk of congestion at a 5 m time scale, and triggering increased monitoring resolution to the 0.3 s scale whenever the coarser estimate indicated a risk exceeding a given threshold.

Table II shows the results for three (5 m) detection levels. The columns “true positive” and “false positive” indicate how often the 5 m detection mechanism accurately predicted at least one 0.3 s congestion in the following 5 m interval. The “true negative” column indicates how often the 5 m detector

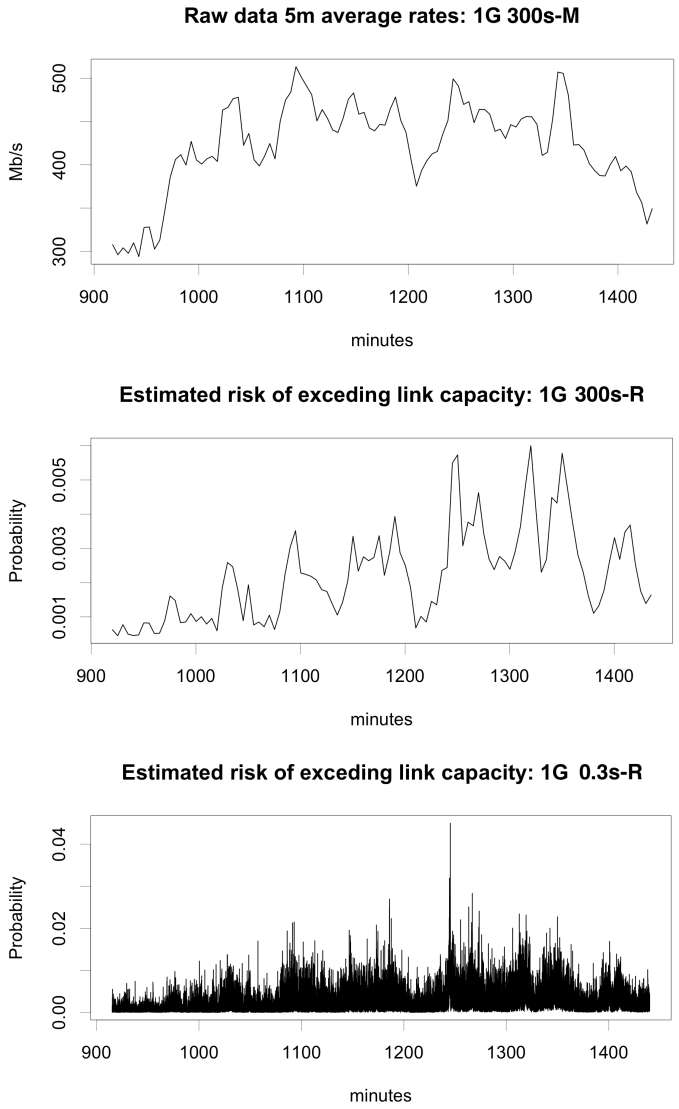

Fig. 5: Time series of 5 m averages over one day from 15-24 on the heavily loaded 1 G link (top), and estimates of the risk of exceeding link capacity over consecutive periods of 5 m, and 0.3 s respectively.

TABLE II: Detection rates for naive congestion detector.

threshold $t = 0.01$	5 m			0.3 s	
	true positive	false positive	true negative	hit	miss
t/15	72	23	9	1184	17
t/10	69	11	24	1180	21
t/5	49	6	49	1088	113

correctly predicted *absence* of congestion at 0.3 s during the following 5 m period. The "hit" and "miss" columns indicate how many of the "true" 0.3 s congestions the detector captured.

We can see that even this relatively naive mechanism has a fairly impressive hit rate. For example, a threshold of one tenth of the "true" 0.3 s threshold at the 5 m level, the detector captures 1180 out of 1201 "true" 0.3 s congestions, or 98.3%. The main problem with this simple mechanism is the amount of high rate monitoring required to find the true positives. Still, this simple assessment clearly indicates the potential for using the estimates derived at lower time resolutions as predictors for higher rate events.

IV. Conclusion and perspective

We have proposed a generic local and scalable approach to traffic rate monitoring based on high rate updates of *two* counters in the data plane for recording the first and second statistical moments of each observed rate. The moments are used to estimate the parameters (using a MoM estimator) of a lognormal distribution at predefined and/or variable intervals. Different aspects of the method have been evaluated using real-world data sets. We have verified that the data can be fitted using a lognormal distribution, compared the estimation accuracy relative to observations at different time scales and tested a naive probabilistic method for detecting increased risk of congestion on a link using probabilistic thresholds on properties of the estimated distributions.

Analysis of the percentiles of estimates obtained at low rates shows clear potential for methods for autonomously and robustly detecting high risk of congestion. Future work includes development of a more robust detector based on adaptive probabilistic thresholds, and extension of the study to less aggregated flows, and shorter time scales.

Acknowledgement

This work was supported in part by the FP7 UNIFY EU project. The authors would like to thank the staff at TeliaSonera for providing access to the traffic rate measurements: Per Tholin, Johanna Nieminen and Patrik Lindwall.

References

[1] W. John, K. Pentikousis, G. Agapiou, E. Jacob, M. Kind, A. Manzalini, F. Risso, D. Staessens, R. Steinert, and C. Meirosu, "Research directions in network service chaining," in *Future Networks and Services (SDN4FNS), 2013 IEEE SDN for*. IEEE, 2013, pp. 1–7.

[2] A. Császár, W. John, M. Kind, C. Meirosu, G. Pongrácz, D. Staessens, A. Takács, and F.-J. Westphal, "Unifying Cloud and Carrier Network: EU FP7 Project UNIFY," in *Utility and Cloud Computing (UCC), 2013 IEEE/ACM 6th International Conference on*. IEEE, 2013, pp. 452–457.

[3] R. Presuhn, "Management information base (MIB) for the simple network management protocol (SNMP)," 2002, RFC 3418, Internet Engineering Task Force.

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

[5] Cisco IOS, "NetFlow," 2008.

[6] P. Phaal, S. Panchen, and N. McKee, "Inmon corporations sflow: A method for monitoring traffic in switched and routed networks," RFC 3176, Tech. Rep., 2001.

[7] C. Yu, C. Lumezanu, Y. Zhang, V. Singh, G. Jiang, and H. V. Madhyastha, "FlowSense: monitoring network utilization with zero measurement cost," in *Proceedings of the 14th International Conference on Passive and Active Measurement*, ser. PAM'13. Springer, 2013, pp. 31–41.

[8] K. Fukuda, "Towards modeling of traffic demand of node in large scale network," in *Communications, 2008. ICC'08. IEEE International Conference on*. IEEE, 2008, pp. 214–218.

[9] A. B. Downey, "Lognormal and pareto distributions in the internet," *Computer Communications*, vol. 28, no. 7, pp. 790–801, 2005.

[10] K. Papagiannaki, R. Cruz, and C. Diot, "Network performance monitoring at small time scales," in *Proceedings of the 3rd ACM SIGCOMM Conference on Internet Measurement*. ACM, 2003, pp. 295–300.

[11] S. Miller and D. Childers, *Probability and random processes: With applications to signal processing and communications*. Academic Press, 2004.