Towards Multihop Available Bandwidth Estimation

François Baccelli
INRIA and ENS
francois.baccelli@ens.fr

Bruno Kauffmann
INRIA and ENS
bruno.kauffmann@ens.fr

Darryl Veitch *
University of Melbourne
dveitch@unimelb.edu.au

ABSTRACT

We evaluate the algorithm proposed in [1], which estimates the residual bandwidth on each hop of a path using a parametric Kelly network model. The evaluation is driven by simulation based on real network traces over a short path. Correction factors are proposed and evaluated to correct for deviations from model assumptions.

1. INTRODUCTION

In [1] we describe how the techniques known as *active Internet probing* and *network tomography* can be viewed as *inverse problems in queuing theory*. Among the examples given was a FIFO Kelly network model of an Internet path, using a maximum likelihood estimator to infer available bandwidth on each hop along the path from the probe delay distribution, as exposed in section 2.

Whilst the Kelly network model is parametric and tractable, it makes the following strong assumptions that may not hold in real networks: **1** Routers are considered as FIFO queues; **2** Service times are independent across hops; **3** Packets sizes are exponentially distributed; **4** Cross traffic packets and probes have the same size distribution; **5** Cross traffic is Poisson. This article aims at verifying that in practice, the assumptions listed are either (nearly) verified, or can be dealt with. Section 3 provides numerical results and useful corrections to the base estimator, inspired by simulation based on real traces. The results of assumption 4 are similar to those of assumption 3, and are omitted.

We are grateful to Sprint ATL and Sridhar Machiraju for the access to the full router experiments, scripts and ancillary data.

2. THE MODEL AND EM ALGORITHM

We consider a 2 hop path in a Kelly network. All packets have exponential service time with mean 1, cross-traffic packets on server i (resp. probes) are sent according to a Poisson process of intensity λ_i (resp. x) and the service rate of server i is μ_i.

In such a network, the end-to-end delay d for a probe is the sum of two independent exponential random variables of parameter $\gamma_1' = \mu_1 - \lambda_1 - x$ and $\gamma_2' = \mu_2 - \lambda_2 - x$, which has density $\phi_{\gamma_1,\gamma_2}(d) = \frac{\gamma_1'\gamma_2'}{\gamma_2'-\gamma_1'}\left(e^{-\gamma_2' d} - e^{-\gamma_1' d}\right)$, $d \geq 0$.

The likelihood function for n independent probe delays $\mathbf{d} = (d_1, \ldots, d_n)$ is $f_{\gamma_1,\gamma_2}(\mathbf{d}) = \prod_{i=1}^{n} \phi_{\gamma_1,\gamma_2}(d_i)$. One way to compute the maximum likelihood estimator, which is the value $(\widehat{\gamma}_1, \widehat{\gamma}_2)$ of (γ_1, γ_2) that maximizes the likelihood function $f_{\gamma_1,\gamma_2}(\mathbf{d})$, is to use the Expectation - Maximization (E-M) algorithm, which provides a sequence of estimates converging to the maximum likelihood estimator. In our setting the algorithm reads:

E-M Algorithm: *For arbitrary initial* $(\gamma_1^{(0)}, \gamma_2^{(0)})$*, do the following for each positive integer k: compute* $\gamma_1^{(K+1)}$ *and* $\gamma_2^{(K+1)}$ *so*

$$\frac{1}{\gamma_1^{(k+1)} + x} = \frac{1}{n}\sum_{i=1}^{n} \frac{d_i e^{(\gamma_2^{(k)} - \gamma_1^{(k)})d_i}}{e^{(\gamma_2^{(k)} - \gamma_1^{(k)})d_i} - 1} - \frac{1}{\gamma_2^{(k)} - \gamma_1^{(k)}}$$

$$\frac{1}{\gamma_2^{(k+1)} + x} = \frac{1}{\gamma_2^{(k)} - \gamma_1^{(k)}} - \frac{1}{n}\sum_{i=1}^{n} \frac{d_i}{e^{(\gamma_2^{(k)} - \gamma_1^{(k)})d_i} - 1}.$$

The same approach extends easily to any number of stations. The proof, details about the model and the algorithm, and the convergence, bias and variance of estimators, can be found in [1].

3. VALIDATION

3.1 Traces and Methodology

We use a subset of the traces described in [3] and [2]. All packets traversing a Sprint gateway IP router which leave it via a particular OC-3 interface were monitored, for 9 different 5 minute time intervals.

The traces are used to feed a simulator modeling a single stage FIFO queue of desired capacity. To determine which assumptions are most crucial, we used the "semi-experimental method", where all theoretical assumptions are enforced, except the ones we specifically want to study.

3.2 Challenge: Routers as FIFO queues

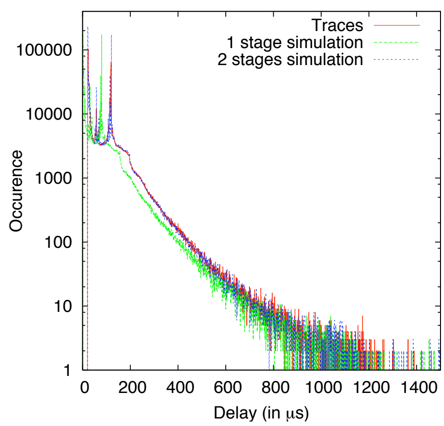

Figure 1: Comparison of true and model delays

*ARC Special Research Centre for Ultra-Broadband Information Networks (CUBIN), an affiliated program of National ICT Australia (NICTA).

Input Traces	Output	γ	$\widehat{\gamma}$	$\widehat{\gamma}$ corrected
P1-BB1 + P1-BB2	OC-3	42.5	34.6	42.0
Q1-BB1 + Q1-BB2	OC-3	89.5	55.0	77.7

Table 1: Impact of two-stage correction (1200 probes).

In [3] two models are proposed for router behavior: (i) the classical FIFO model; (ii) a 2 stage model where packets have first to wait a time $\Delta(S) = a + bS$ depending on their size S, before joining a FIFO queue (which models the output buffer).

Figure 1 shows the true versus surrogate delays for packets from the input interface 'BB1-in'. The one stage model histogram has a very similar shape to the true one, but is visibly offset from it by 50μs. The two stage model is nearly perfect.

Response : Estimation correction for two stage model
In this case the experiment corresponds to a M/M/1 queue, except that the router is replaced by a two stage model which represents a simplified but accurate ground truth. As seen in Table 1, the errors in a Kelly based E-M estimate are large, however a simple correction based on replacing $\widehat{\gamma} = \mathbb{E}[S]/\mathbb{E}[D]$ by $\widehat{\gamma} = \mathbb{E}[S]/(\mathbb{E}[D] - \mathbb{E}[\Delta(S)])$ largely succeeds in correcting it.

3.3 Challenge: Size Independence

Contrary to a Kelly network, where service times at different stations are independent, packets have a constant packet size and so correlated service times on different hops of a real network. As ground truth we set probe sizes to be independent at each hop, and compare against the real network case. Three scenarios are considered: 1 : input traces (Q1) and (P1) on stations 1 and 2 respectively, 2 : input traces (Q1, P2-BB2) and (P1, P3-BB1), 3 : input traces (Q4) and (P4). Table 2 shows that the impact can be quite large.

Response: Random Probe Split
We emulate a probe which is of exponential size and different at each hop by sending a back-to-back probe pair, the first of which will drop out after hop 1. The correction is quite effective.

3.4 Challenge: Exponential Sizes

We now investigate the impact of service times which are not exponentially distributed, by using an experiment where all aspects of the traffic are replaced by model surrogates except for the original (non-exponential) cross traffic packet sizes.

Two examples of real traffic are given in Table 3. The errors are small, however the coefficient of variation of packet sizes is close 1. To create a more challenging test, we modify (at constant mean) the packet size distribution in 2 other experiments : (i) constant packet sizes, (ii) bimodal distribution on $(40, 3000)$ bytes with probability $(0.785, 0.215)$. The E-M estimates are now significantly different.

	load	(γ_1, γ_2)	$(\widehat{\gamma}_1, \widehat{\gamma}_2)$	$(\widehat{\gamma}_1, \widehat{\gamma}_2)$ corr
1	(0.42, 0.70)	(45.5 , 87.5)	(33.9 , 166,4)	(42.7 , 88.2)
2	(0.77 , 0.96)	(5.4 , 33.7)	(5.5 , 34.7)	(6.4 , 35.8)
3	(0.36 , 0.38)	(92.7 , 96)	(50.6 , 748.0)	(94.4 , 94.4)

Table 2: Impact of Size Independence (120 000 probes).

Input Traces	κ	γ	$\widehat{\gamma}$	$\widehat{\gamma}$ corrected
P1-BB1, P1-BB2	1	42.5	42.1	38.7
Q1-BB1, Q1-BB2	1	89.5	85.7	86.5
P1 - Constant size	-0.5	42.5	70.3	44.9
P1 - Bimodal distrib	0.215	42.5	25.5	45.9

Table 3: Impact of Exponential Size assumption (1200 probes).

Input Traces	Output	γ	$\widehat{\gamma}$	$\widehat{\gamma}$ corrected
P1-BB1, P1-BB2	OC-3	42.5	33.8	44.6
Q1-BB1, Q1-BB2	OC-3	89.5	87.8	100.1
Q4-BB1 + Q4-BB2	OC-3	95.0	92.3	98.4

Table 4: Impact of Poisson assumption (1200 probes).

Scenario	(γ_1, γ_2)	$(\widehat{\gamma}_1, \widehat{\gamma}_2)$	$(\widehat{\gamma}_1, \widehat{\gamma}_2)$ corrected
Q1–P1	$(45.5, 87.5)$	$(27.8, 147.9)$	$(43.6, 97.3)$
Q4–P4	$(92.7, 96)$	$(95.0, 95.0)$	$(101.5, 101.5)$

Table 5: Poisson assumption: 2 hops (120 000 probes).

Response: Variance correction factor
We can relax the exponential packet size assumption in our M/M/1 model by considering the M/G/1 queue, where the service times are i.i.d. with general distribution. Careful derivation using the Pollaczek–Khinchin formula leads to a correction factor $F_S = 1 + \kappa\rho$, where κ is defined by $\mathbb{E}\left[S^2\right] = 2(1+\kappa)\mathbb{E}\left[S\right]^2$.

3.5 Challenge: Poisson Arrivals

In this section, the only non-ideal components left are the original non-Poisson arrival processes of the cross traffic packet streams.

Response: Poisson batch correction
The correction in this case is based on the idea that packets arrive at the input in batches of back to back packets, due to the queuing at the upstream hop. If we know the intensity β of batches and the random number N of packets that arrives in the same time in a batch, then the workload of such a process is the same as that in a M/G/1 queue with arrival rate β and service times $S = \sum_{i=1}^{N} \sigma_i$, and the correction factor is as explained in section 3.4.

The key point is that from the first and second moments $C_t^{(1)}$ and $C_t^{(2)}$ (resp. $\mathbb{E}(\tau)$ and $\mathbb{E}(\tau^2)$) of the workload arriving in an interval of length t (resp. the packet inter-arrival time τ), one can estimate β and $\mathbb{E}(S^2)$, and hence compute the correction factor $1 + \kappa\rho$. As seen in Table 4, the correction succeeds in reducing the biggest error, but increased others. The traffic here is close to Poisson so the correction factors are small.

Using the ladder epoch representation, it is also possible to "correct" end-to-end delays for a several hops path, such that they fit the delays from a corresponding Kelly path. We cannot describe this here, but some results are presented in Table 5.

4. CONCLUSION

This paper shows that it is possible to apply successfully correction techniques to cover unverified assumptions, and finally estimates the real available bandwidth on each link of the 2 hops path. These results still needs to be extended to longer paths, and we also need to explain whether the needed statistics are stable over time.

5. REFERENCES

[1] F. Baccelli, B. Kauffmann, and D. Veitch. Inverse Problems in Queueing Theory and Internet Probing. *Queuing Systems*, Special Issue for the Erlang Centennial, to appear in 2009.

[2] F. Baccelli, S. Machiraju, D. Veitch, and J. Bolot. On Optimal Probing for Delay and Loss Measurement. In *Proc. ACM SIGCOMM Internet Measurement Conf.*, pages 291–302, San Diego, California, 23-26 October 2007.

[3] N. Hohn, D. Veitch, K. Papagiannaki, and C. Diot. Bridging router performance and queuing theory. In *Proc. ACM Sigmetrics*, pages 355–366, New York, June 12-16 2004.