Delay and Loss Modeling in Internet

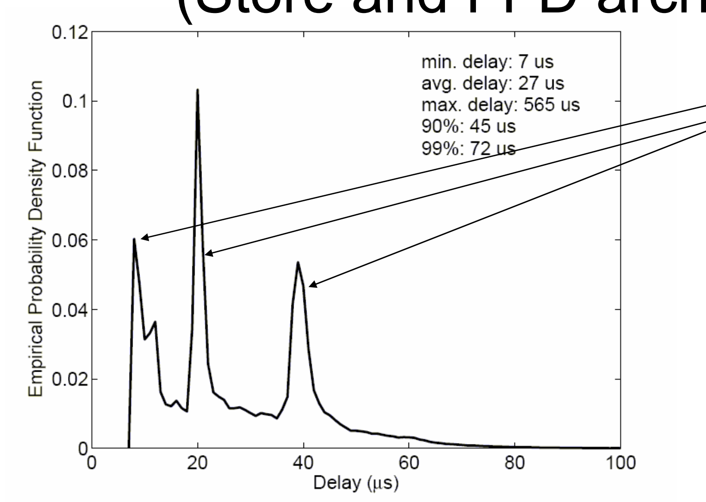
By

Sanjeev Dwivedi
CS8803-Network Measurement

What causes Delays & Losses? [MoonJSAC03]

- Bugs in router implementations.
- Speed of EM waves in media.
- CPU Power (e.g. routing updates).
- Packets on the slow path.
- Congestion (Queuing).
- Packet sizes.
- Noisy channels.
- Route flapping.

Delays (w/in a router) proportional to packet sizes (Store and FFD arch)

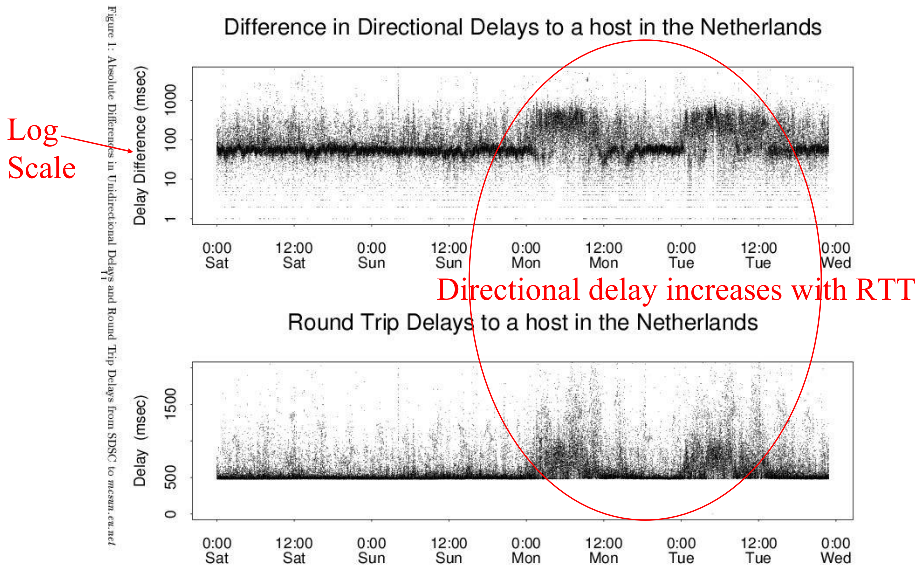


Trimodal Graph (Peaks for most freq. packet sizes)

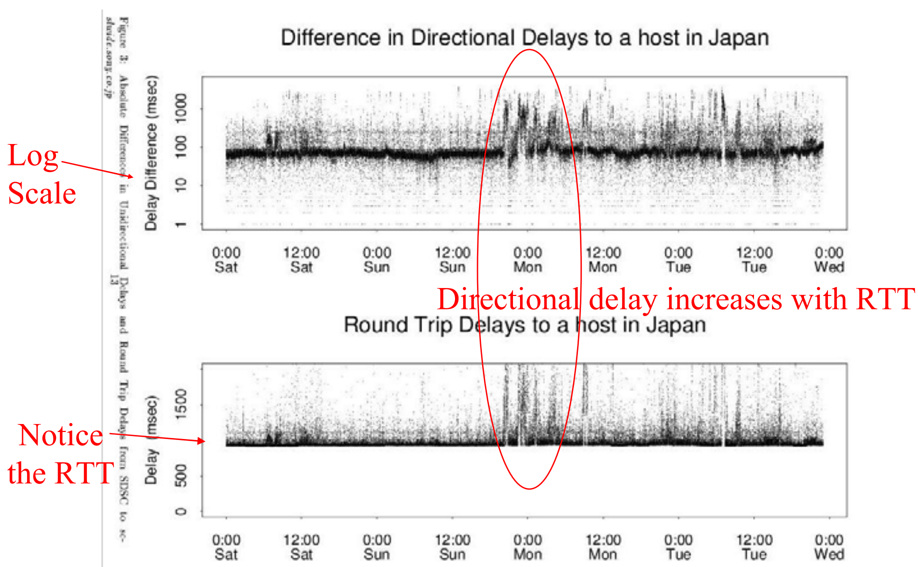
Are delays symmetric on Internet Paths? [Claffy]

All delays in following sections assume end-to-end measurements.

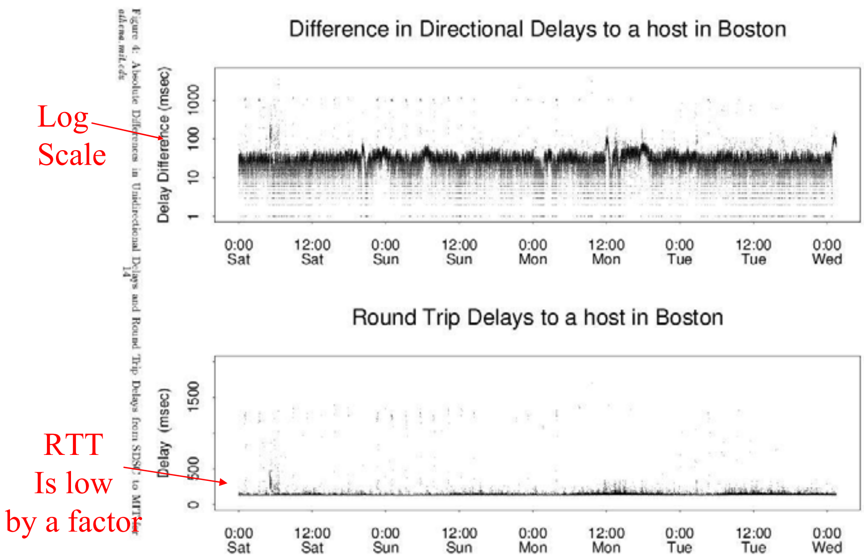
Transatlantic Link – I (1992)



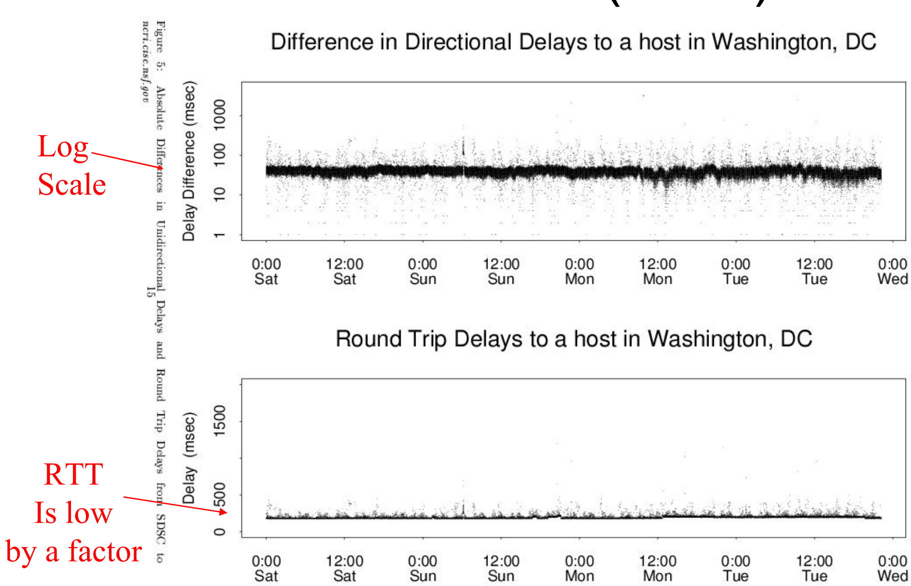
Transatlantic Link – II (1992)



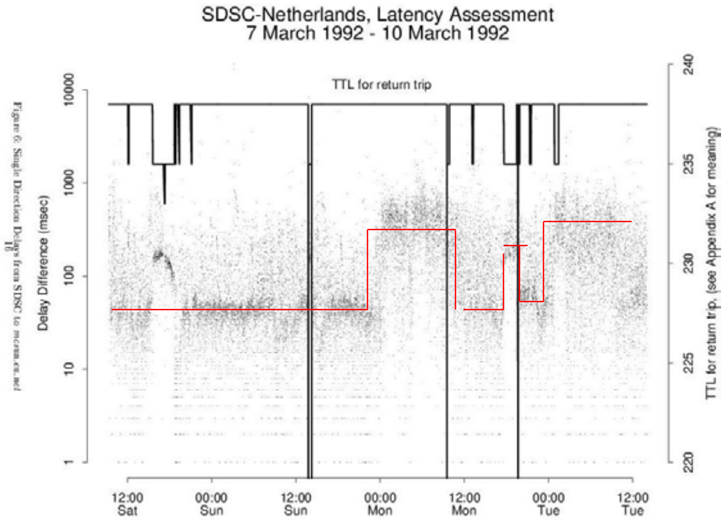
USA Route – I (1992)



USA Route – II (1992)



TTL Variation (Route Flapping) [Cottrell]



RTTs

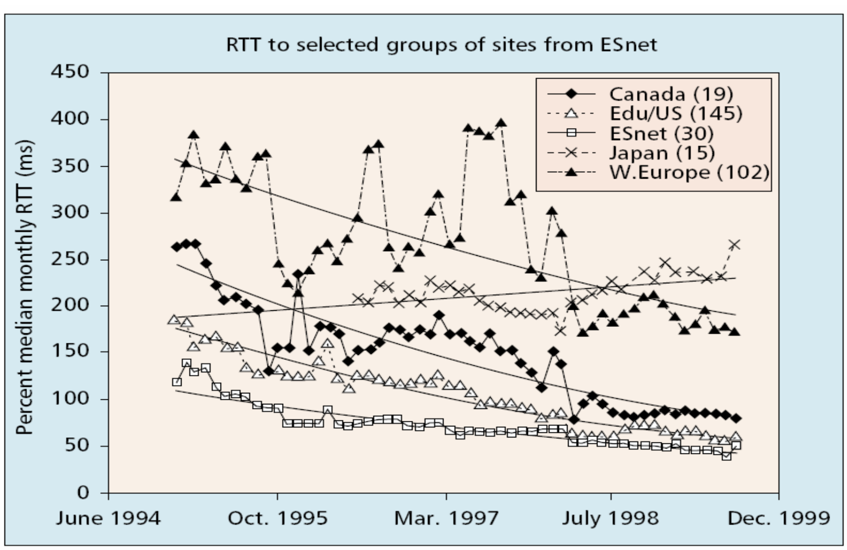
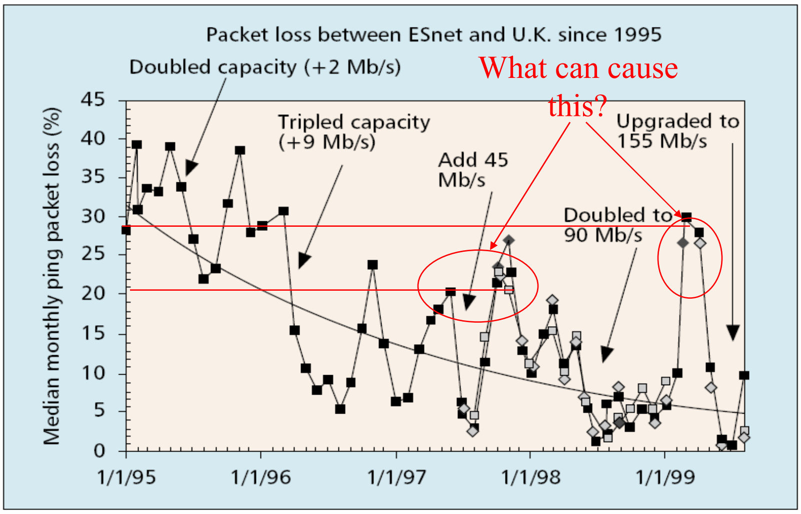


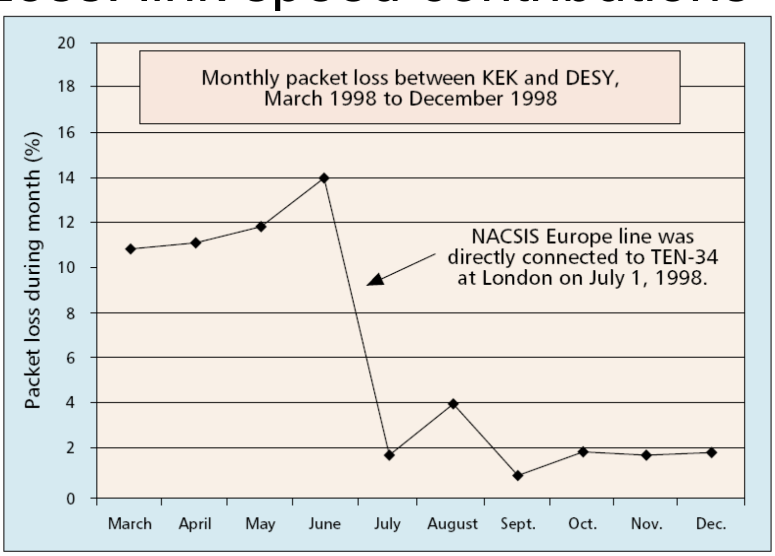
Figure 1. *Round-trip time between ESnet and sites in several groups. The numbers in parentheses in the legend are the number of pairs aggregated.*

Losses



■ **Figure 2.** *Packet loss between U.S. sites connected to ESnet and sites at universities in the U.K. connected to the JAnet NRN. Packet loss decreases significantly as the capacity of the link is upgraded, and during university holidays.*

Loss: link speed contributions



■ **Figure 3.** *Packet loss between KEK in Japan and DESY in Germany. Packet loss decreased from unusable to acceptable after the NACSIS line from Japan to Europe was relocated to London and connected directly to the TEN-34 network.*

Route changes

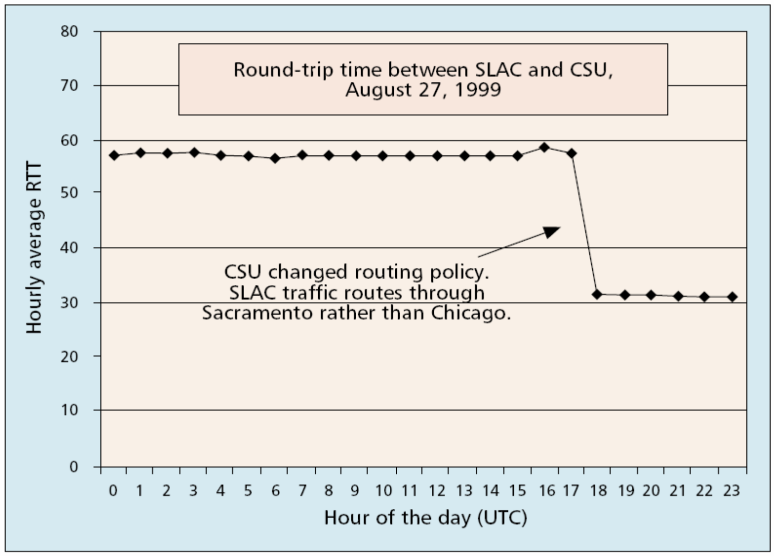


Figure 4. *Average round-trip time between SLAC and Colorado State University dropped after CSU changed routing policy so that packets destined for SLAC were sent to Sacramento rather than the ESnet hub in Chicago.*

ICMP ping Measurements: Effectiveness

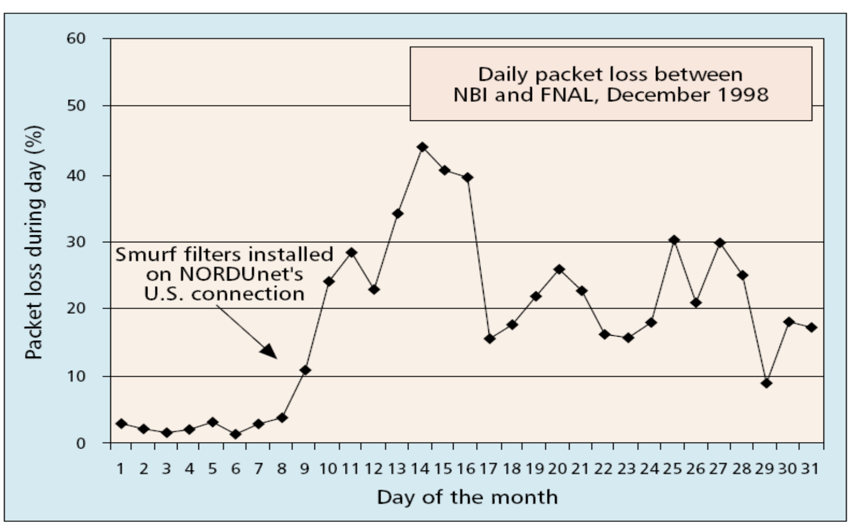


Figure 5. *Packet loss between NBI and FNAL appeared to increase wildly in December 1998. However, this is an effect of a filter used to prevent security attacks and is not a true indication of network performance.*

Link speed effects

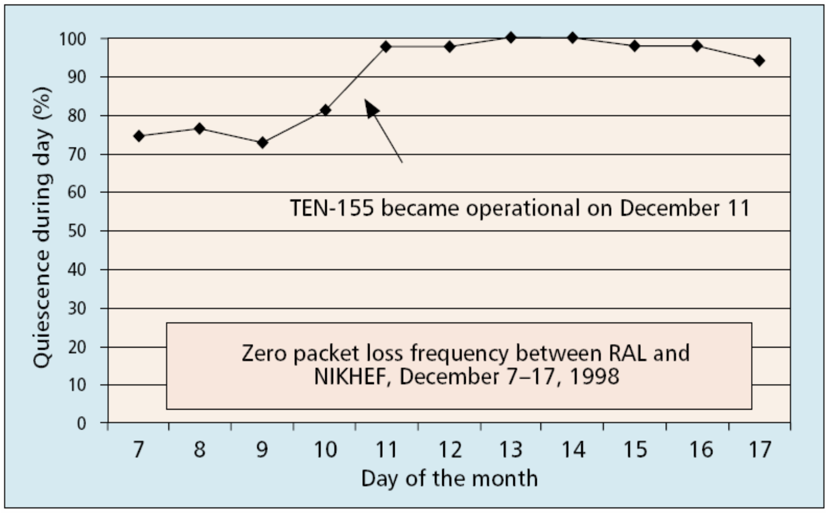


Figure 6. *Frequency of zero packet loss (quiescence) between RAL and NIKHEF improved when the TEN-155 network became operational in December 1998.*

How to measure and sanitize? [Pax97]

- Measuring between the same pairs
- Measuring at poisson intervals because it gives unbiased measurement even if sampling rate varies [pax96]
- Network pathologies: Reordering (violates FIFO model), packet replication (how and why can this even happen? Link layer replication, stack implementation fault…), packet corruption.
- TTL shift removal (Routing changes)
- Compressed timing
- Suspect clocks

Some stats

	N1 (1994)	N2 (1995)
Reorder (%)	36	12
Reorder (data pkt) %	2	0.3
Reorder (Ack pkt) %	0.6	0.1
Packet replication (count)	1	65
Packet corruption %	0.02	0.02

Do large windows cause higher loss rates?

- Hypothesis: it can be done by looking at loss rates of *data* packets vs. those of *ack* packets.
- Reason: Because Data packets stress the ffd path more than smaller ack packets, also ack rate is usually less than data rate by a factor.
- But: TCP data packets adapt to the current conditions, whereas the acks do not unless a whole flight of packets is lost, causing a sender timeout. Hence ack losses give a clearer picture of overall internet loss patterns, while data losses tell specifically about the conditions as perceived by TCP connections.
- Proof (look?):

	N1	N2
Data %	2.65	5.28
Ack %	2.88	5.14

Geographical Effects

Region	$Quies_1$	$Quies_2$	$Busy_1$	$Busy_2$	Δ
Europe	48%	58%	5.3%	5.9%	+11%
U.S.	66%	69%	3.6%	4.4%	+21%
Into Europe	40%	31%	9.8%	16.9%	+73%
Into U.S.	35%	52%	4.9%	6.0%	+22%
All regions	53%	52%	5.6%	8.7%	+54%

Except for the links going into US (check what this means: Probably the Into Europe thing because the delta is 73%) the proportion of quiescent connections is fairly stable. Hence loss rates increases are primarily due to higher loss rates during the already-loaded "busy" periods.

Loss Evolution & behavior over time

- Observing a zero loss connection at a given point in time is quite a good predictor for longer time scales (hrs, even days & weeks). Similar holds for lossy durations.
- Observation of the predictive behavior of lossy/non-lossy periods gives the notion that network lives in an "on-off" state or "quiescent-busy" periods, and that both periods might be long-lived.
- Predictive power of observing a *specific loss rate* is much lower than that of observing the presence of zero/non-zero loss

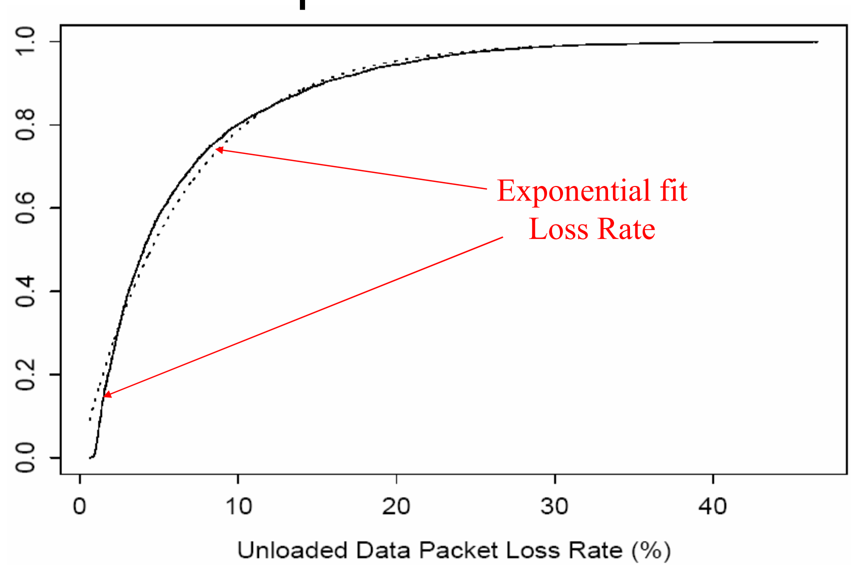
Losing data packets and acks

Loaded	Unloaded	Acks
47%	65%	68%

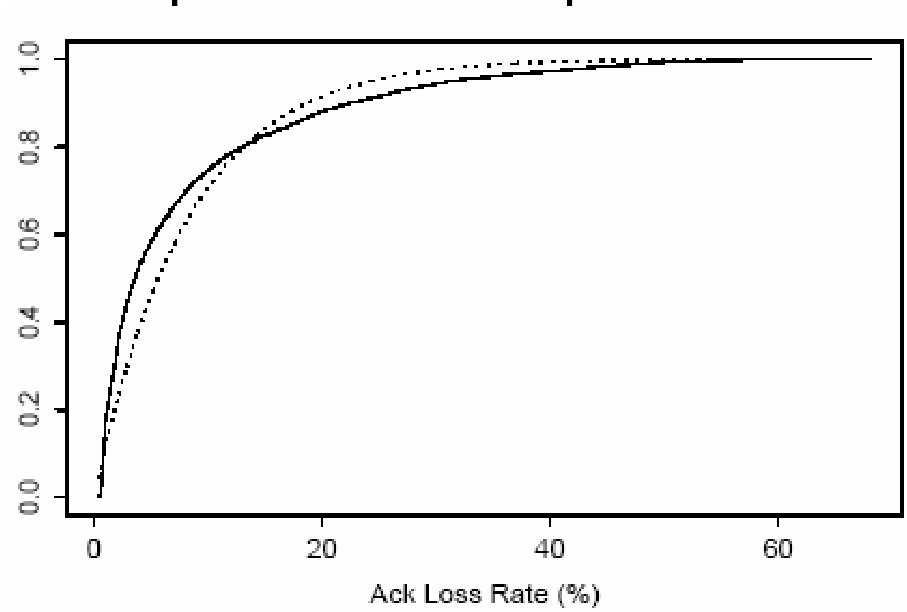
For all of these extremes (the above stats are for different connections, worst case scenario in each category), *no* packet was lost in the reverse direction (TCP Good)! Clearly, packet loss on the ffd and reverse directions is sometimes completely independent

Non-zero portions of both the unloaded and the loaded data packet loss rates agree closely with exponential distributions, while that for acks is not so persuasive match.

Data packet loss and exponentials



Ack packets and exponentials

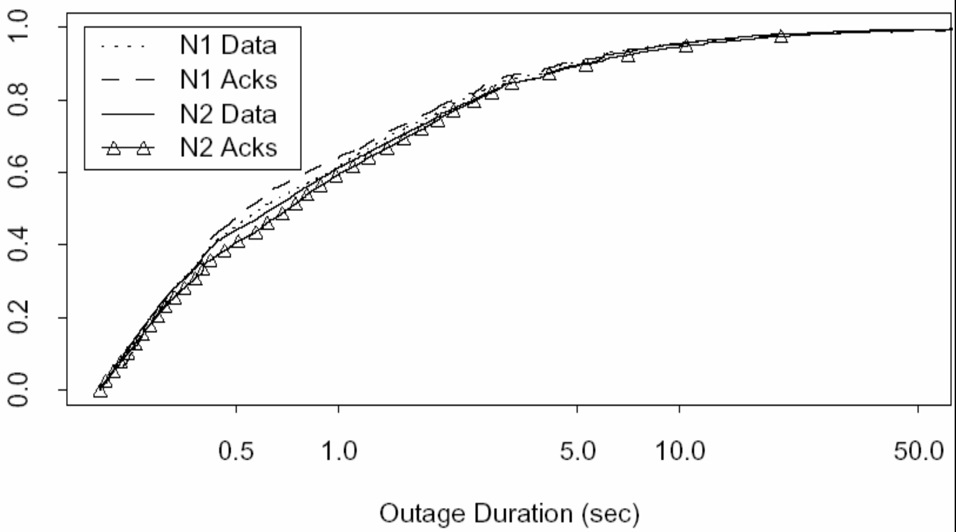


Loss Bursts and Independence

Type of loss	P_l^u		P_l^c	
	$\mathcal{N}_1$	$\mathcal{N}_2$	$\mathcal{N}_1$	$\mathcal{N}_2$
Loaded data pkt	2.8%	4.5%	49%	50%
Unloaded data pkt	3.3%	5.3%	20%	25%
Ack	3.2%	4.3%	25%	31%

- Loss events are well modeled as independent does not hold.
- If successive packet losses (in bursts) are grouped as *outages,* then the outage duration span several orders of magnitude

Outage duration spans large orders



Delay: Why timing compression does not matter?

$$\xi = \frac{\Delta T_r + C_r}{\Delta T_s - C_s},$$

Compression is present if: $\xi < 0.75$.

Ack compression is rare and small in magnitude. Also, upper extremes can be dealt by removing extremes from sender based measurements.

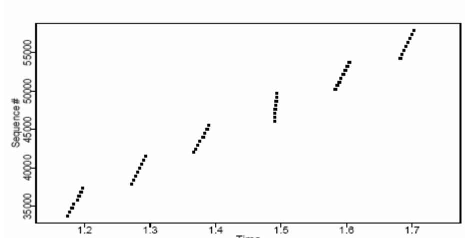


Figure 9: Data packet timing compression

Data packet compression is rarer than ack compression, but it recurs between specific pairs of hosts and indicates that it happens repeatedly, and is sometimes due to specific routers. It is rare enough not to present a problem, but outlier filtering should be applied in measurements.

Delays: Queuing time scales

Are there particular timescales at which most queuing variations occur?

Methodology (For OTTs):

1) **Sanitize the trace.**

2) **any time interval t, divide it into two halves. If $N_l < (1/4)N_r$ *or* either chunk is zero, discard the interval. Else let M_l and M_r be medians of left and right halves, and Q_t be median of $|M_l - M_r|$, the interval's queuing variation.**

For each connection find value of *t* such that Q_t is greatest.

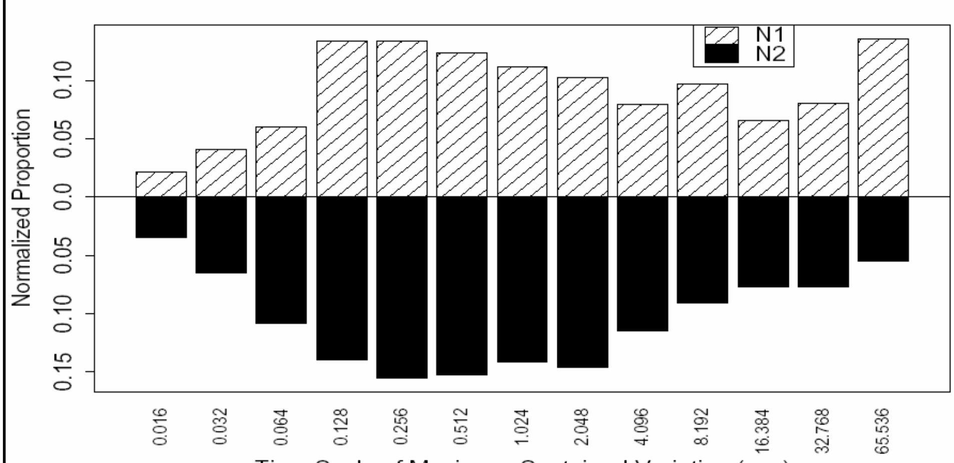
Find the frequency of each *t*.

Normalize by dividing the frequency of *t* by all *t*'s at least as large as this *t*.

Find the proportion of each *t* from the numbers that we got in previous process

Queuing time scales - II

Summary: Internet delay variations occur primarily on time scales of 0.1-1sec, but extend out quite frequently to much larger scales.



Getting Modern

Data collection Methodology [Moon99]

•End-to-End data collection

•Loss measurement by sending probes in network periodically (periods = 20ms, 40ms, 80ms, 160ms).

•Both Unicast and Multicast probes.

Trace Sanitization

•Making sure that the trace segments are stationary, i.e. splitting the traces in 2 hr intervals and checking for stationarity. This is done by using a moving average filter (of window size 2000 packets) to judge the extent of variation.

•Other non-stationary effects (sudden increase in loss rate, slow decays or increase in loss percentages) are observed and those traces are not included in the analysis.

Analyzing the data

•A packet whose sequence number is not recorded is assumed to be lost.

•Under the above assumption, the loss data can be represented as a binary time series

$$\{x_i\}_{i=1}^{n}$$

$$x_i = \begin{cases} 1 \text{ if probe i is lost} \\ 0 \text{ otherwise} \end{cases}$$

How to check whether the losses are independent?

- We should check if the binary time series of the loss data has any dependence.
- To do that we need to learn a little math.
- We know that for an independent process the autocorrelations at all lags are zero. We could try this with our loss data series.
- But since we have only an observed sample sequence, the autocorrelations may not be zero exactly.
- But maybe the autocorrelations of our time series are so small that they are insignificant. Excellent! But we need to check if it really is so. We need to have confidence in its insignificance.

Finding Confidence!

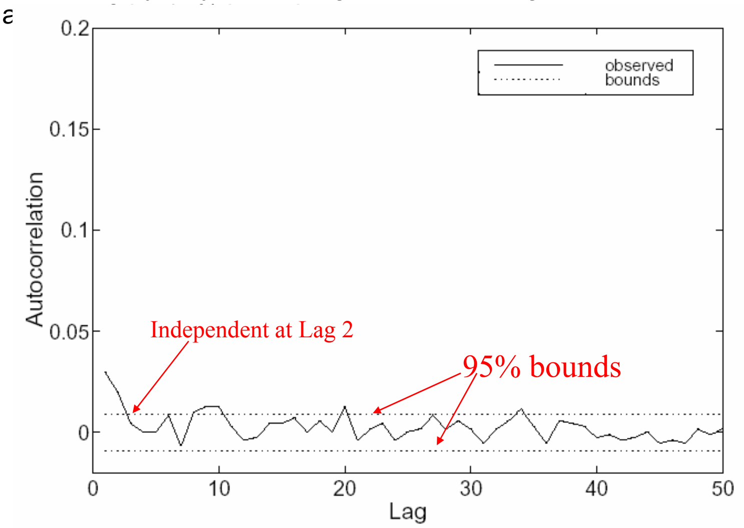
- For large *n,* the sample autocorrelations of an IID sequence with finite variance are approximately IID with a normal distribution *N(0,1/n).*
- Thus if we have a sample of an IID sequence with finite variance, then 95% of the autocorrelation values for this sample should lie between: $\mu \pm 1.96\sigma$

$$= \pm 1.96 / \sqrt{n}$$

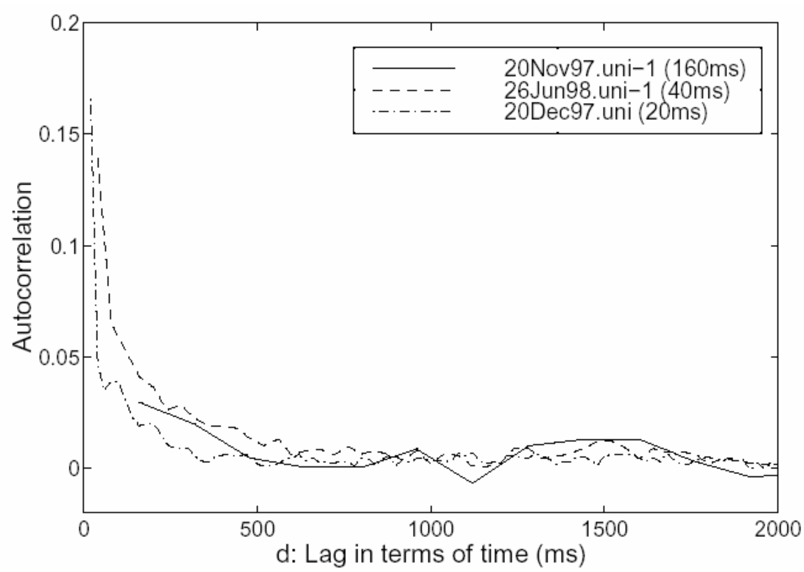
- If we divide the autocorrelations above by the standard deviation the resulting distribution will become a unit normal distribution.
- Thus if $|S = \hat{\rho}(h)/\sigma = \hat{\rho}(h)\sqrt{n}|$ < 1.96 at all h, then our sample sequence is independent. Else if x is the smallest lag at which |S| > 1.96, then we say that the sequence is independent at lag x-1.

Making the hypothesis

- Thus if we make the hypothesis that "the autocorrelations from the sample loss sequence are independent at lag h", we can verify it by calculating the smallest lag at which we a



Autocorrelations with lag as timescales



Frequency dist. Of correlation timescales

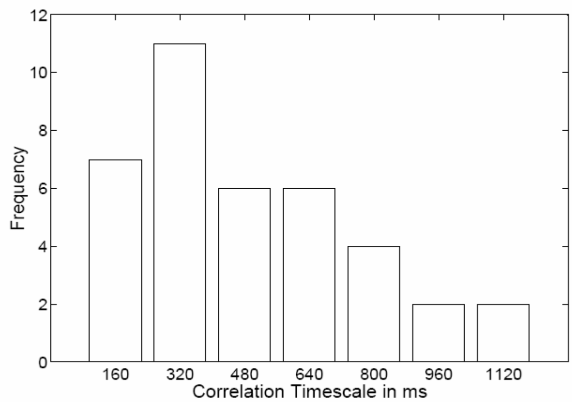


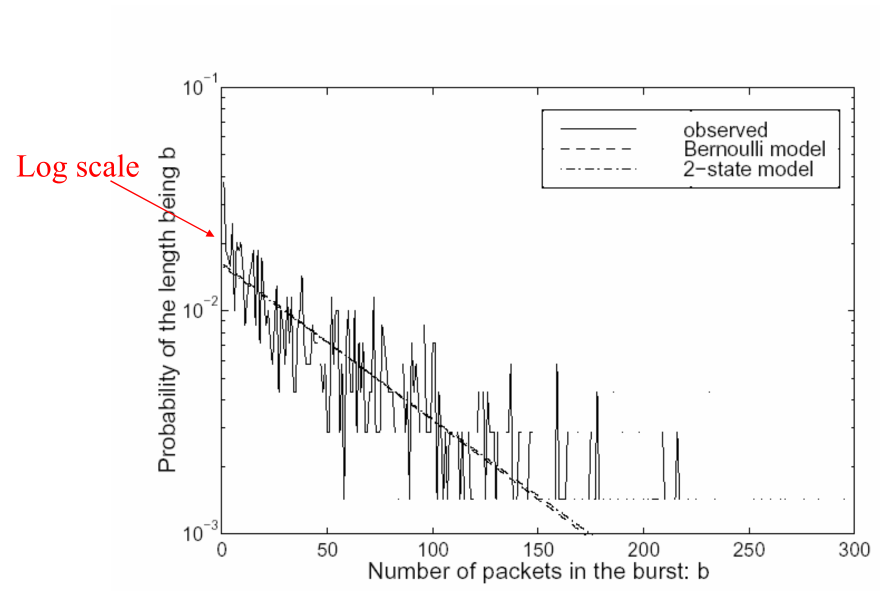
Fig. 3. Frequency Distribution of Correlation Timescale

We find that the losses are not independent because they are correlated at various lags. Thus the hypothesis is rejected.

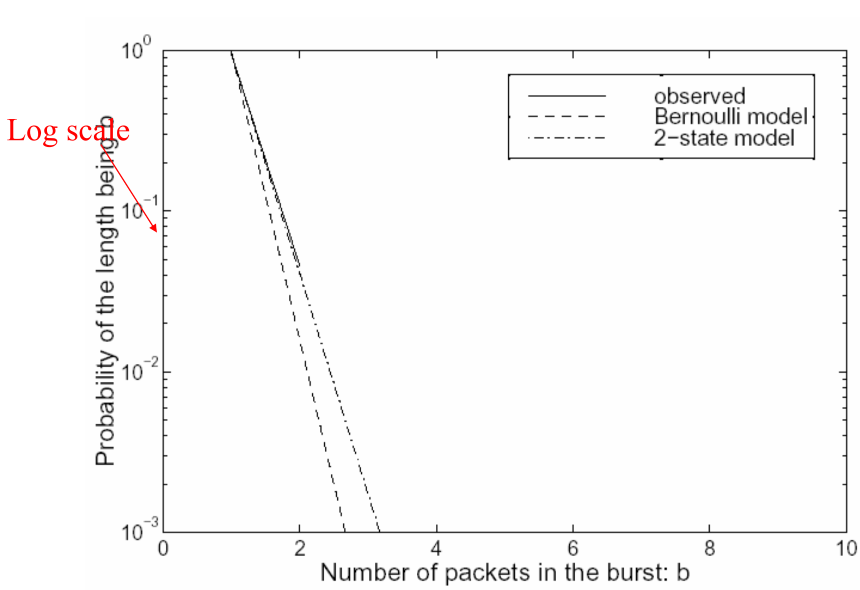
Modeling good and loss run lengths

- The data can also be represented as sequences of good run lengths (no loss) and loss run lengths (losses). The dependence between the good runs and loss runs can be found by plotting the autocorrelation functions of the good run lengths, loss run lengths and crosscorrelation function between the good and loss run lengths.
- From the graphs, it seems that the distributions of good and loss run lengths follow the exponential and Gamma distributions.
- Hypothesis: The good run lengths are IID random variables with the exponential and Gamma distributions.
- Hypothesis Testing: using the 95% confidence intervals.

Distribution of Good run lengths–trace 1



Distribution of loss run lengths-trace1



Hypothesis rejected!

But the hypothesis is rejected:

- 11 out of 33 for exponential distribution
- 2 out of 33 for gamma distribution

Modeling the data - I

- Since none of the previous hypothesis held successfully, we try to model the binary time series of losses as Markov chains of increasingly complex orders.
- Bernoulli Loss Model: It is described by only one parameter, $r = p(x(i))$ being $1 = n(1)/n$ where $n(1)$ is the number of times the value 1 occurs in the observed time series.
- Two state Markov Chain Model: The current state of the stochastic process depends only on the previous value. The maximum likelihood estimators:

$p=n(01)/n(0), q=n(10)/n(1)$

Here $n(01)$ is the number of times 1 follows 0 and $n(10)$ is the number of times 0 follows 1, $n(1)$ is the number of ones and $n(0)$ is the number of 0s

Modeling the Data - II

- Kth order Markov chain:

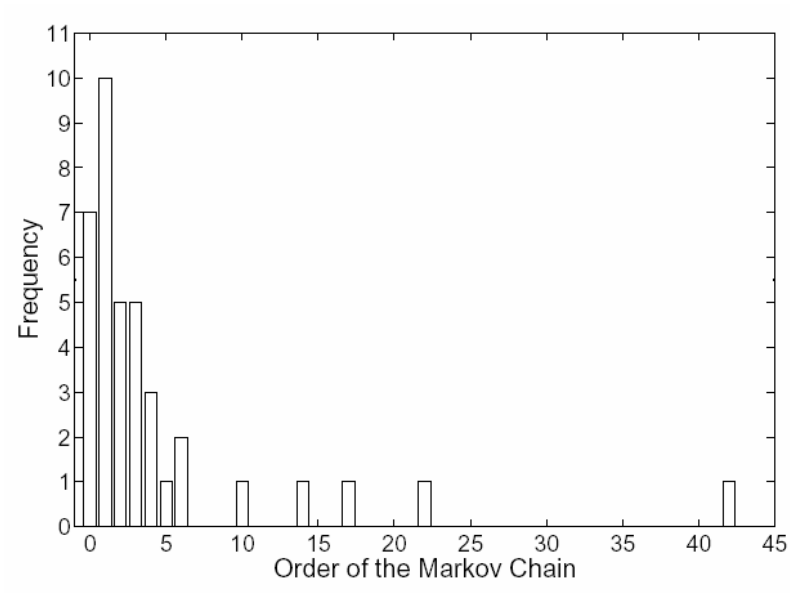
Let $b = (b_1, b_2, \ldots b_k)$ be a given state of the chain and n_{ba} be the number of times state *b* is follows by value *a* in the sample sequence and n_b is the number of times state *b* is seen. Then if $p^k_{b;a}$ be an estimate of the probability that system moves from state *b* to state $(b_2, \ldots b_k, a)$ then the maximum likelihood estimators of the *k*-th order Markov chain are:

$p^k_{b;a}$={ n_{ba} / n_b *if* n_b *>0* and 0 otherwise}

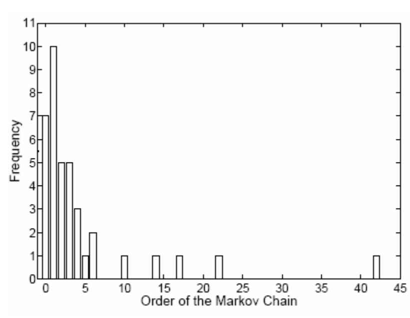
Finding the order of the Markov process

- It can be estimated by estimating the minimum lag beyond which the process is independent.
- i.e. Test Hypothesis 1 for lags 1 to 50. Let *l* be the first lag at which the hypothesis is not rejected. Then the estimated order is *l*-1.

Order of the Markov Chains



Finding the order - II



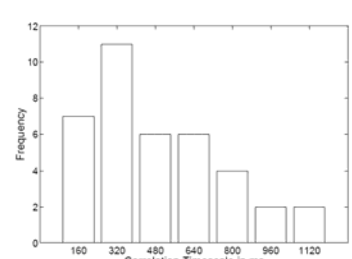


Fig. 3. Frequency Distribution of Correlation Timescale

Open issues

- How to actually use this modeling (Markov Chains) in simulations?
- Internet packet loss rates (measured using non-adaptive sampling) vs. loss rates experienced by a transport connection's packets. There might be a link [Pax97]

References

- Measurement and Modelling of the Temporal Dependence in Packet loss Maya Yajnik, Sue Moon, Jim Kurose, Don Towsley IEEE INFOCOM 99 [Moon99]
- End-to-End Internet Packet Dynamics (Sigcomm 1997) Vern Paxson [Pax97]
- The PingER Project: Active Internet Performance Monitoring for the HENP Community Warren Matthews and Les Cottrell [Cottrell]
- Measurement Considerations for assessing unidirectional latencies

  Kimberly C. Claffy, George C. Polyzos and Hans-Werner Braun [Claffy]
- Characterizing End-to-End Packet Delay and Loss in the Internet (1993)

  Jean-Chrysostome Bolot Sigcomm 93 [Bolot93]
- "Packet-Level Traffic Measurements from the Sprint IP Backbone," C. Fraleigh, S. Moon, B. Lyles, C. Cotton, M. Khan, D. Moll, R. Rockell, T. Seely, C. Diot. *IEEE Network, 2003.* [MoonNetwork03]
- "Measurement and Analysis of Single-Hop Delay on an IP Backbone Network," K. Papagiannaki, S. Moon, C. Fraleigh, P. Thiran, C. Diot. *IEEE JSAC Special Issue on Internet and WWW Measurement, Mapping, and Modeling*, 3rd quarter., 2003. [MoonJSAC03]