Bell Labs DARPA Project

Part I

1. Simulation: Statistical Models for Open-Loop Generation of Best-Effort Packet Traffic
2. QoS: Statistical Models for Bandwidth Allocation

Jin Cao, Bill Cleveland, and Yuan Gao
{cao, wsc, yuangao}@bell-labs.com

Part II

QoS: Algorithms for MPLS Traffic Engineering

Anwar Walid, Karun Nithi, and Debasis Mitra

Presentation yesterday by Anwar Walid

Fractional Sum-difference (FSD) Statistical Models

Open-loop generation of best-effort packet streams

- arrivals and sizes arriving at a link
- first parsimonious models for generation
- our trick is an initial transformation to approximate gaussian

Software and integration

- generation software in C completed
- with Jennifer Hou and Guanhui He, implemented in JavaSim

Specification of generated stream is easy

- packet size distribution (default)
- number of packets/sec/connection (default)
- traffic load: either number of active connections or bit rate (no default)

Fitting model (training) to live trace is easy

- small number of parameters to estimate
- Fourier methods and pseudo-likelihood

Fast and scalable

- specify a number of active connections
- statistical properties depend on number
- a single thread: generation time independent of number

Accuracy

- extensive validation
- mimics statistical properties of real-world traces for number of simultaneous connections of 64 and above

Bandwidth Allocation for Internet Links

For a specified traffic load τ (bits/sec), what link bandwidth β (bits/sec), or equivalently utilization $u = \tau/\beta$, is needed to maintain quality-of-service criteria?

A Problem of Statistical Modeling

statistical properties of packet arrivals and sizes

↓

queue-height process

↓

quality-of-service bandwidth

A widely studied, classical problem, much of it under the name of “effective bandwidth”

The most fundamental problem of traffic engineering

Critical for those who run networks, like the major ISPs

Basic input to path optimization routines

Our Approach

Treat the specific types of traffic on the Internet, because different types have different statistical properties

- best-effort (today's presentation)
- VoIP (just beginning with Global Crossing)

To solve the bandwidth allocation problem for a specific traffic type

- build a statistical model specifically for that type
- study the problem for that statistical model
- the best chance for a parsimonious, exact solution for the type

Formulation of the Problem for Best-Effort Traffic

Negligible packet loss in getting on the link

- choose buffer size sufficiently large
- relevant for sufficiently highly aggregated traffic
- consistent with the quality goals of service providers on their core links
- bottleneck is elsewhere for the connections using the link
- can use open-loop simulation

QoS delay delay criteria

- target maximum delay, δ (sec)
- delay frequency, ω (fraction of packets delayed by more than δ)

Simulation Study

Given a traffic stream with bit rate τ

- QoS criteria: fix delay δ and delay frequency ω
- find QoS bandwidth β and QoS utilization $u = \tau/\beta$ that satisfy criteria

Synthetic packet streams

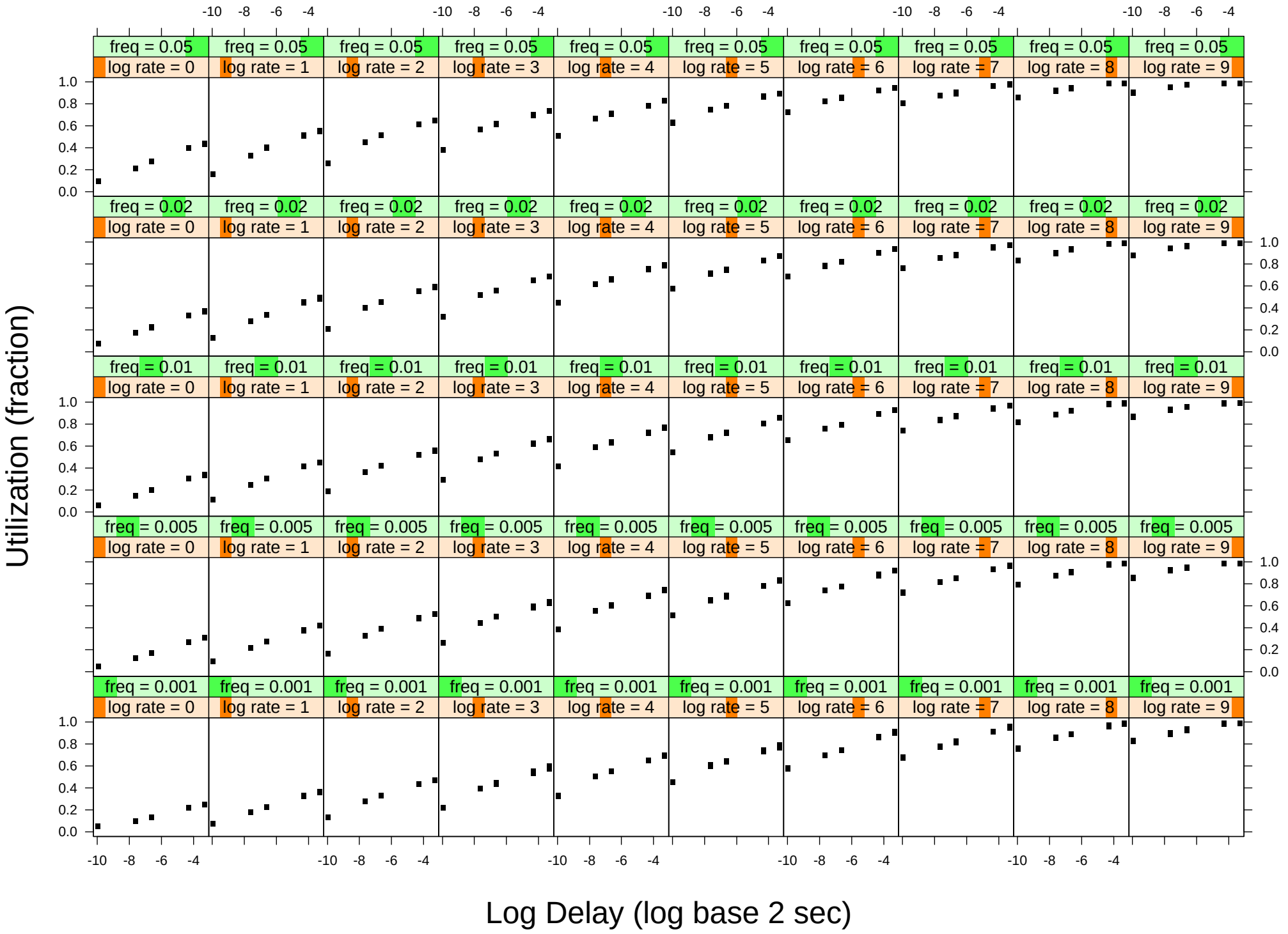
- generated 20 synthetic streams with FSD models
- 10 values of τ, 2 streams for each, increasing by powers of 2 from about 1 mbps to about 512 mbps

Live packet streams

- 532 traces, 90 sec to 5 min to ensure stationarity
- τ from about 1 mbps to 348 mbps
- little delay on links

For each live or synthetic stream

- feed into fifo, infinite-buffer queue
- find QoS utilization u for 25 combinations of δ and ω
- δ: 1, 5, 10, 50, 100 ms
- ω: 0.001, 0.005, 0.01, 0.02, 0.05


freq = 0.05 log rate = 0
freq = 0.05 log rate = 1
freq = 0.05 log rate = 2
freq = 0.05 log rate = 3
freq = 0.05 log rate = 4
freq = 0.05 log rate = 5
freq = 0.05 log rate = 6
freq = 0.05 log rate = 7
freq = 0.05 log rate = 8
freq = 0.05 log rate = 9
freq = 0.02 log rate = 0
freq = 0.02 log rate = 1
freq = 0.02 log rate = 2
freq = 0.02 log rate = 3
freq = 0.02 log rate = 4
freq = 0.02 log rate = 5
freq = 0.02 log rate = 6
freq = 0.02 log rate = 7
freq = 0.02 log rate = 8
freq = 0.02 log rate = 9
freq = 0.01 log rate = 0
freq = 0.01 log rate = 1
freq = 0.01 log rate = 2
freq = 0.01 log rate = 3
freq = 0.01 log rate = 4
freq = 0.01 log rate = 5
freq = 0.01 log rate = 6
freq = 0.01 log rate = 7
freq = 0.01 log rate = 8
freq = 0.01 log rate = 9
freq = 0.005 log rate = 0
freq = 0.005 log rate = 1
freq = 0.005 log rate = 2
freq = 0.005 log rate = 3
freq = 0.005 log rate = 4
freq = 0.005 log rate = 5
freq = 0.005 log rate = 6
freq = 0.005 log rate = 7
freq = 0.005 log rate = 8
freq = 0.005 log rate = 9
freq = 0.001 log rate = 0
freq = 0.001 log rate = 1
freq = 0.001 log rate = 2
freq = 0.001 log rate = 3
freq = 0.001 log rate = 4
freq = 0.001 log rate = 5
freq = 0.001 log rate = 6
freq = 0.001 log rate = 7
freq = 0.001 log rate = 8
freq = 0.001 log rate = 9
-10 -8 -6 -4
0.0 0.2 0.4 0.6 0.8 1.0
Utilization (fraction)
Log Delay (log base 2 sec)

The FSD Delay Formula (from empirical study and theory)

A very simple model fits the simulation data very closely

$$\log_2\left(\frac{u}{1-u}\right) = \mu + o_\delta \log_2(\delta) + o_\omega(-\log_2(-\log_2(\omega))) + \epsilon$$

ϵ

- expresses random error variability of the formula for predicting utilization
- approximately a normal random variable
- mean 0 and variance $\sigma^2(\epsilon)$

o_δ and o_ω

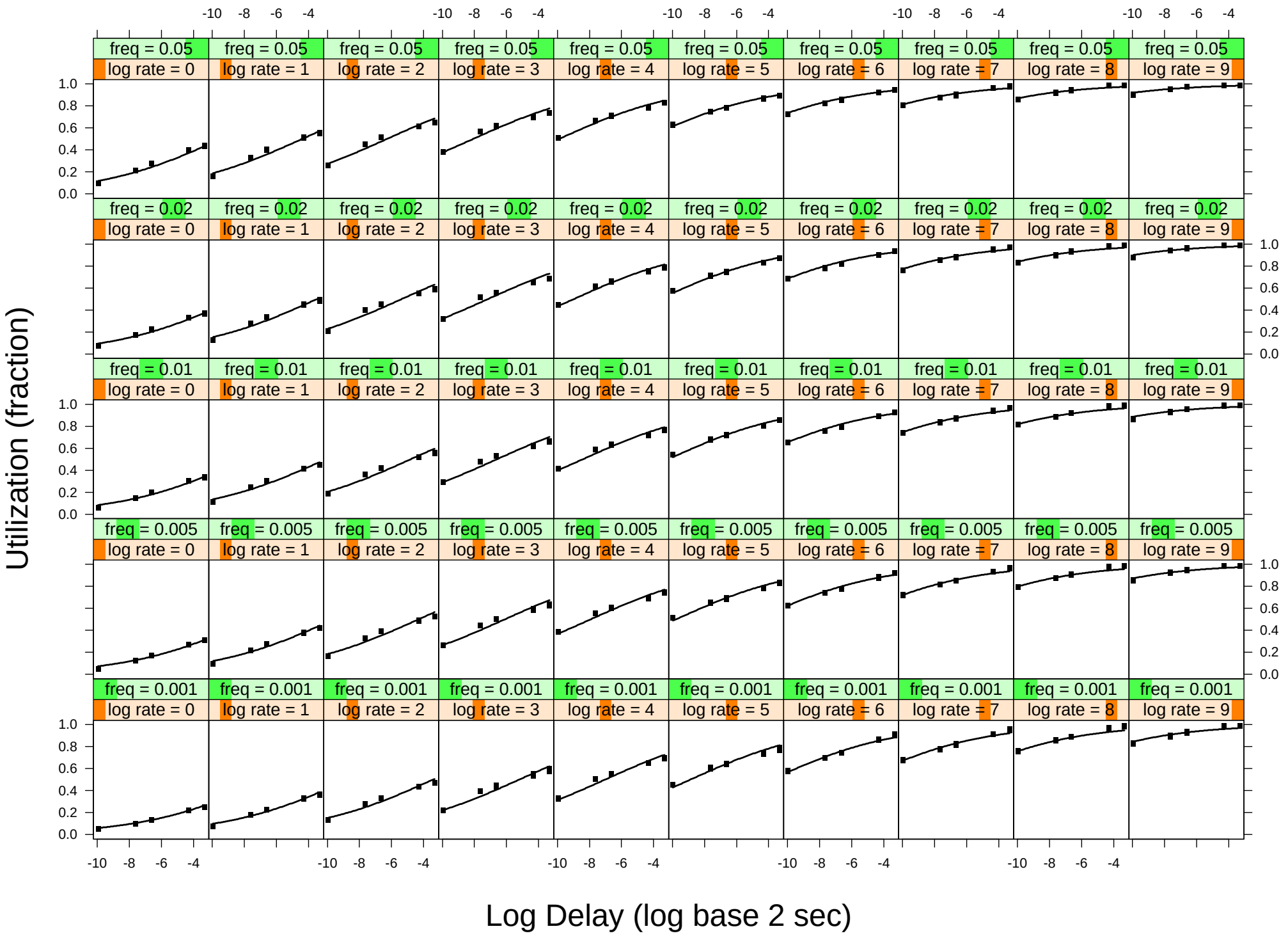
- constant across streams
- estimates for synthetic streams: $\hat{o}_\delta = 0.378, \hat{o}_\omega = 0.903$
- estimates for live streams: $\hat{o}_\delta = 0.432, \hat{o}_\omega = 0.823$
- live and synthetic very close

μ

- changes with the packet stream
- a single parameter characterizes the statistical properties
- more on this later

$\sigma(\epsilon)$

- estimates are very small
- estimate for synthetic streams: $\hat{\sigma}(\epsilon) = 0.120$
- estimate for live streams: $\hat{\sigma}(\epsilon) = 0.163$.


freq = 0.05 log rate = 0
freq = 0.05 log rate = 1
freq = 0.05 log rate = 2
freq = 0.05 log rate = 3
freq = 0.05 log rate = 4
freq = 0.05 log rate = 5
freq = 0.05 log rate = 6
freq = 0.05 log rate = 7
freq = 0.05 log rate = 8
freq = 0.05 log rate = 9
freq = 0.02 log rate = 0
freq = 0.02 log rate = 1
freq = 0.02 log rate = 2
freq = 0.02 log rate = 3
freq = 0.02 log rate = 4
freq = 0.02 log rate = 5
freq = 0.02 log rate = 6
freq = 0.02 log rate = 7
freq = 0.02 log rate = 8
freq = 0.02 log rate = 9
freq = 0.01 log rate = 0
freq = 0.01 log rate = 1
freq = 0.01 log rate = 2
freq = 0.01 log rate = 3
freq = 0.01 log rate = 4
freq = 0.01 log rate = 5
freq = 0.01 log rate = 6
freq = 0.01 log rate = 7
freq = 0.01 log rate = 8
freq = 0.01 log rate = 9
freq = 0.005 log rate = 0
freq = 0.005 log rate = 1
freq = 0.005 log rate = 2
freq = 0.005 log rate = 3
freq = 0.005 log rate = 4
freq = 0.005 log rate = 5
freq = 0.005 log rate = 6
freq = 0.005 log rate = 7
freq = 0.005 log rate = 8
freq = 0.005 log rate = 9
freq = 0.001 log rate = 0
freq = 0.001 log rate = 1
freq = 0.001 log rate = 2
freq = 0.001 log rate = 3
freq = 0.001 log rate = 4
freq = 0.001 log rate = 5
freq = 0.001 log rate = 6
freq = 0.001 log rate = 7
freq = 0.001 log rate = 8
freq = 0.001 log rate = 9
-10 -8 -6 -4
0.0 0.2 0.4 0.6 0.8 1.0
Utilization (fraction)
Log Delay (log base 2 sec)

Modeling the Stream Coefficient μ

For traffic engineering must predict μ by packet stream statistical variables that are available in practice

Start with the most important statistical variable: bit rate τ

A very simple formula explains μ as a function of τ

$$\mu = o + o_\tau(\log_2(\tau) - 24) + \zeta$$

ζ

- expresses random error variability of the formula for predicting utilization
- approximately a normal random variable
- mean 0 and variance $\sigma^2(\zeta)$

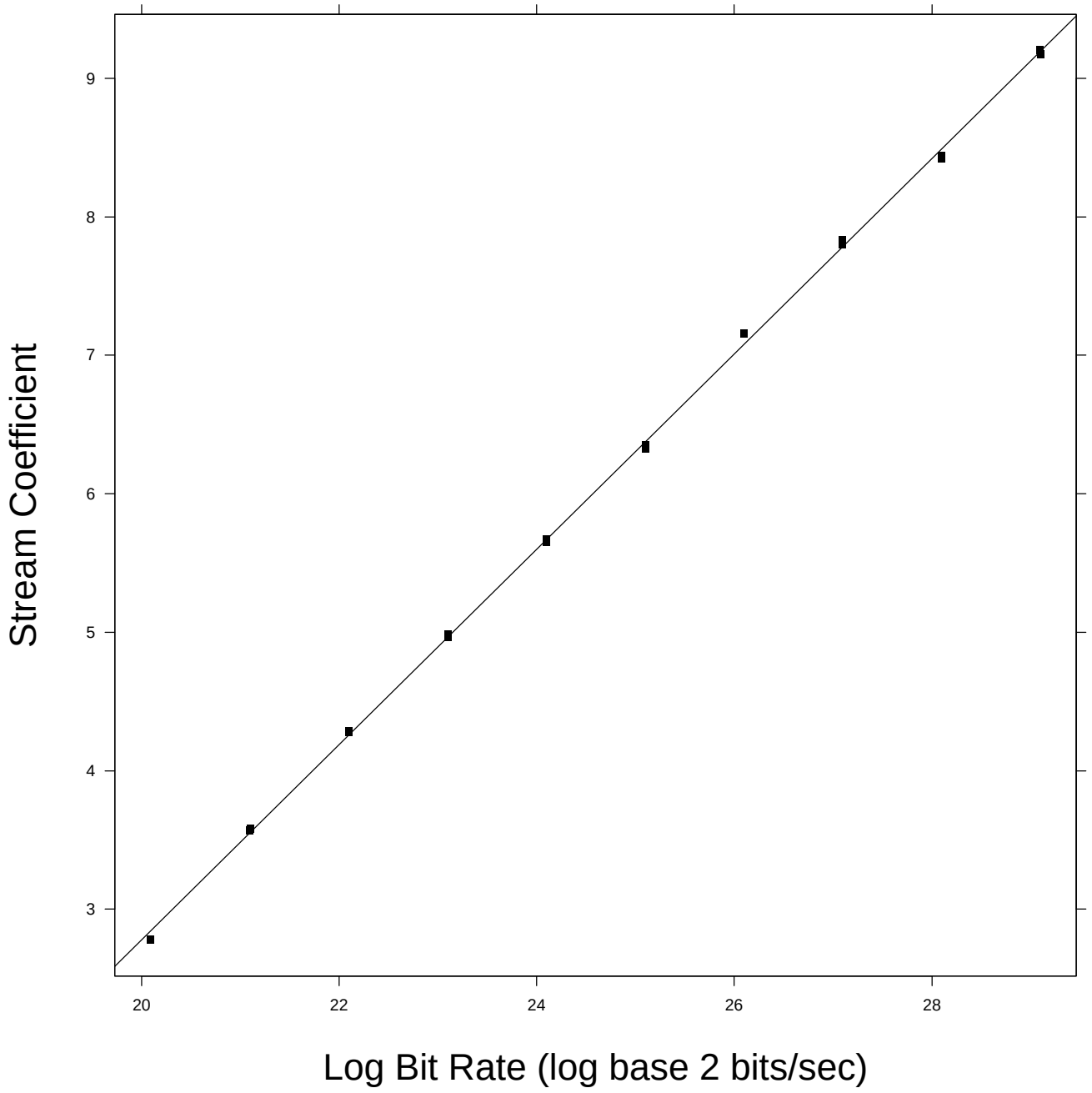
Hierarchical statistical model

- fsd delay formula
- stream coefficient formula

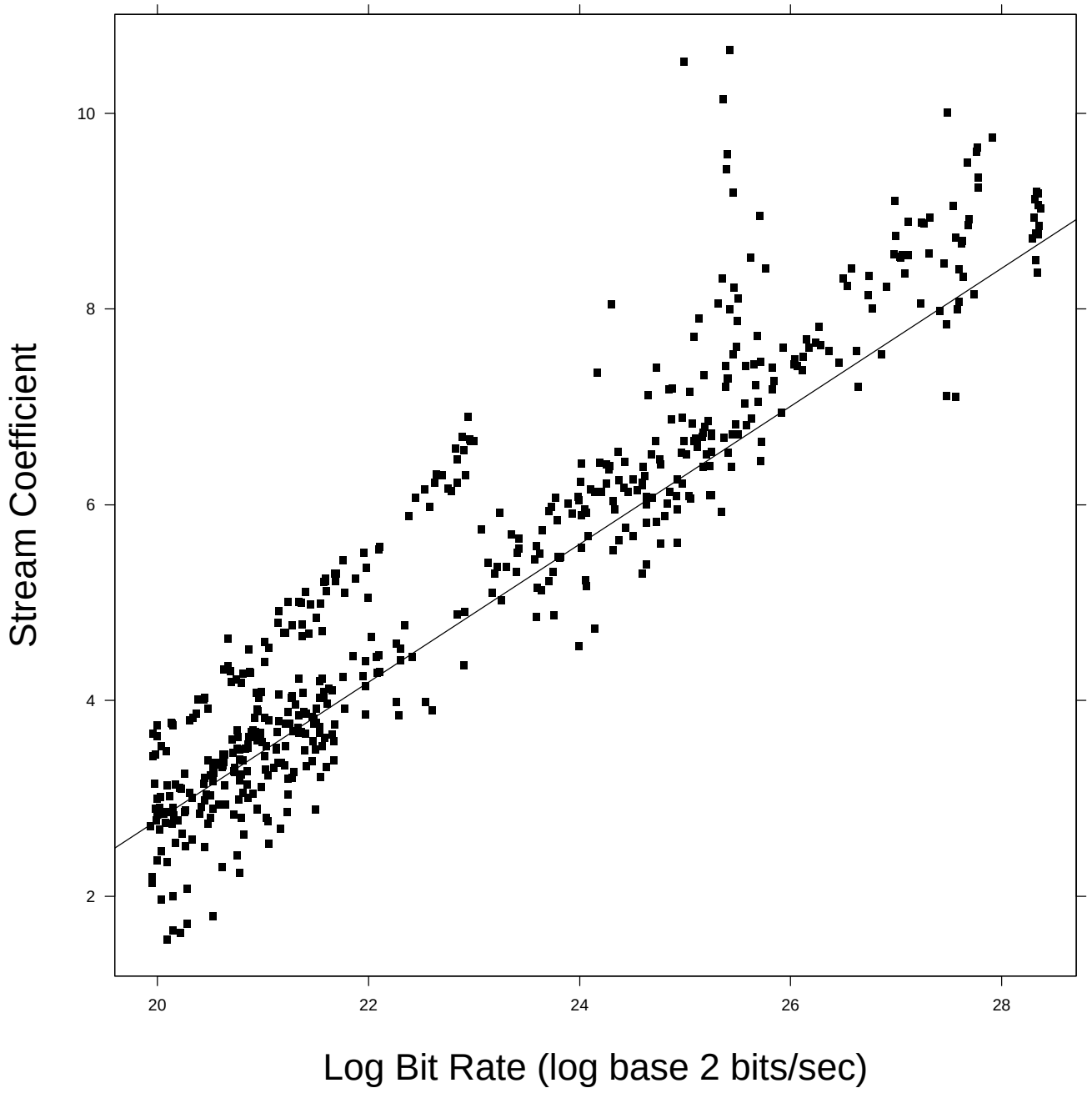
o and o_τ

- estimates for synthetic streams: $\hat{o} = 5.598, \hat{o}_\tau = 0.705$
- the estimates for the live streams: $\hat{o} = 5.908, \hat{o}_\tau = 0.758$
- the two fits are very close

QoS Utilization vs. Log δ Given τ and ω



QoS Utilization vs. Log δ Given τ and ω



Modeling the Stream Coefficient μ

$\sigma(\zeta)$

Estimate for synthetic streams: $\hat{\sigma}(\zeta) = 0.0191$

Estimate for the live streams: $\hat{\sigma}(\zeta) = 0.265$

More variability for live streams

Possible causes: live streams have

- varying packet size marginals and packets/sec/connection
- much smaller sample sizes, so there is more statistical variability

Future work: study this issue