

## INTERNET TRAFFIC AND QoS

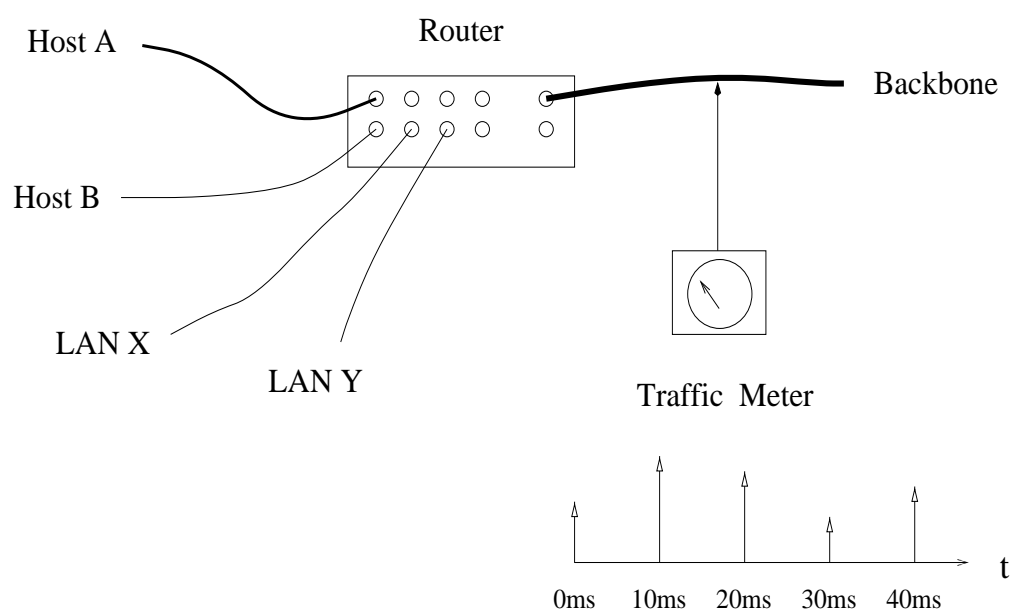
Simplest of all: constant bit rate (CBR)

- flat is good
- because predictable
- e.g., telephone call, real-time MP3 audio

Internet data traffic: variable bit rate (VBR)

- session arrivals: random
- exponential interarrival time
- swimming with the fishes: Poisson process
- e.g., telephone calls, TCP connections, fast food

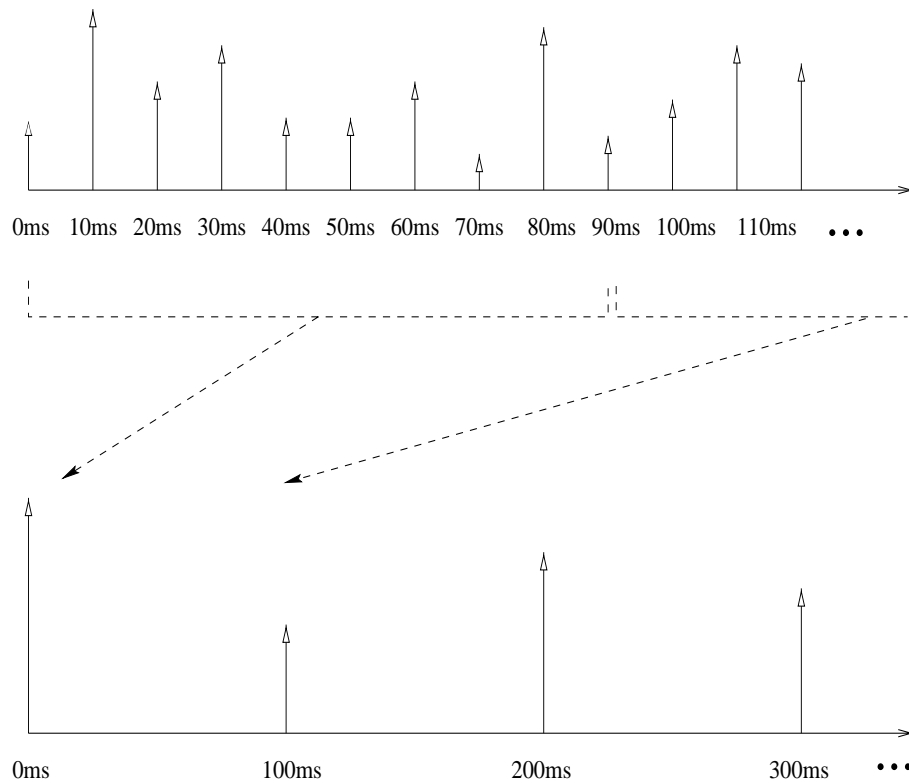
What does Internet traffic look like?



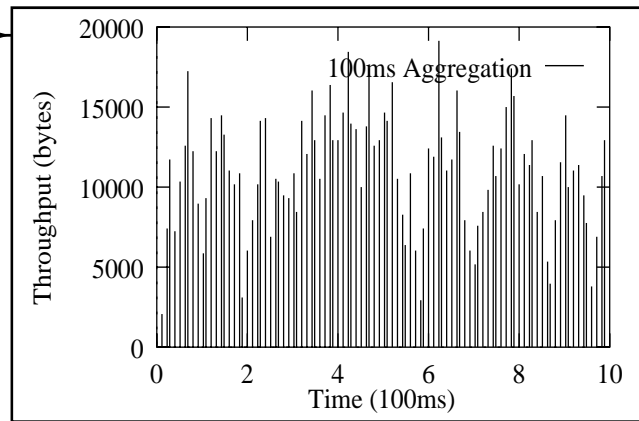
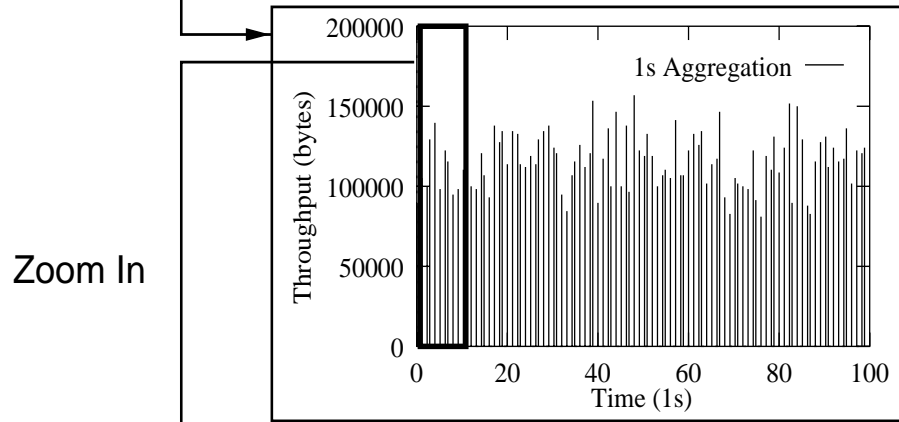
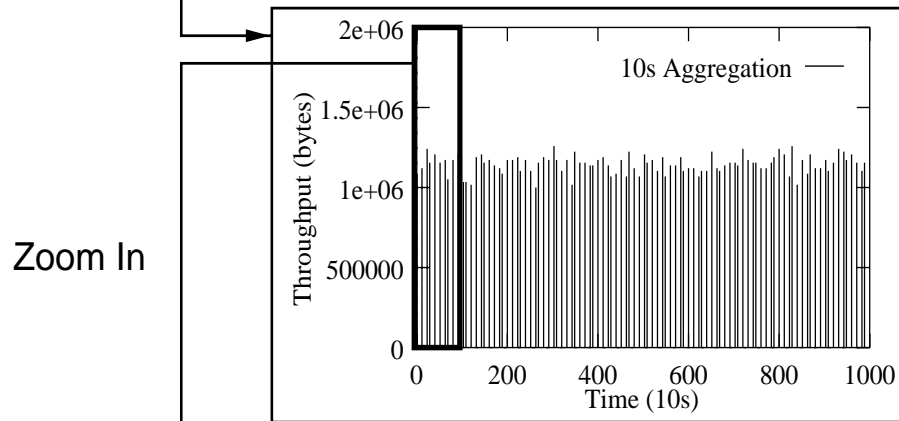
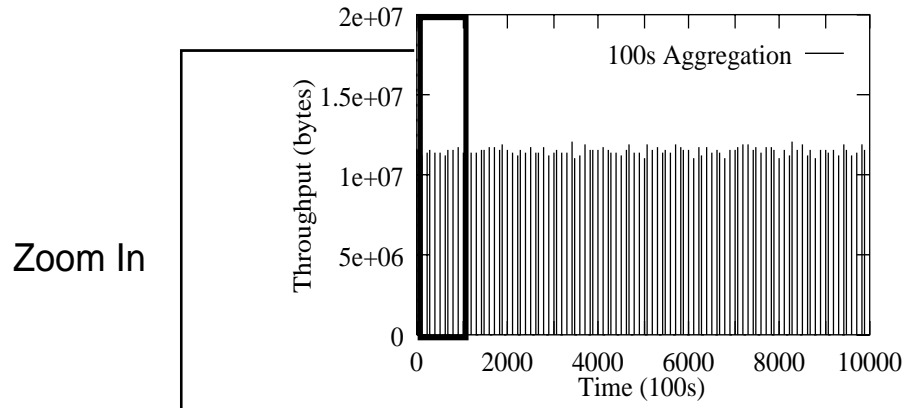
→ logging

→ traffic time series (at 10ms granularity)

Aggregation (time):



- analogous to computing sample mean
- aggregation over multiple time scales
- what to expect?
- first, telephone traffic



Deaggregation

Aggregation

Aggregated traffic becomes flat

- “flat is good” rule for QoS provisioning
- bandwidth dimensioning
- technically: law of large numbers in action
- not correlated in time
- efficient and happy customers

Also aggregation over multiple users

- called statistical multiplexing
- assuming independence between different users
- nice normal distribution shape
- allows non-peak reservation
- consider peak-to-mean ratio: less costly

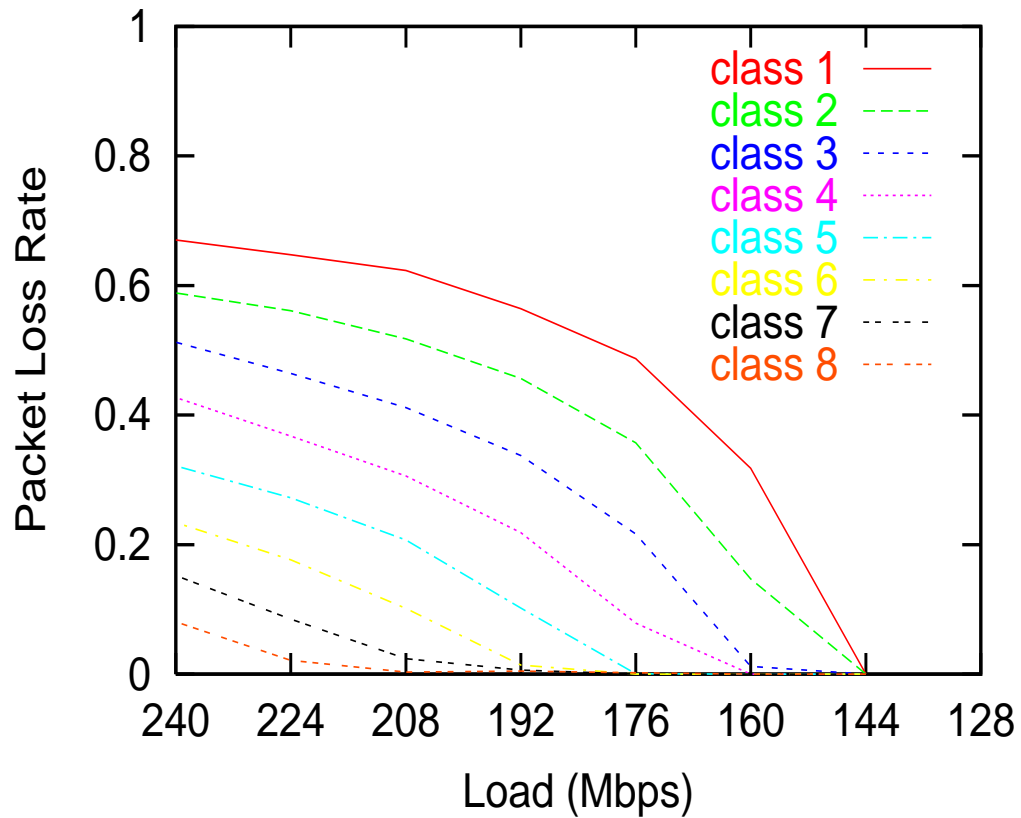
QoS policy:

- Use per-user (or flow) reservation for super-quality service
  - guaranteed service
- Use shared service classes (platinum, gold, silver, bronze) for good service
  - differentiated service

Internet standards:

- IETF IntServ
  - RSVP protocol
  - analogous to leasing a line
- IETF DiffServ
  - different types of router behavior
  - AF, EF, Cisco's LLQ for VoIP

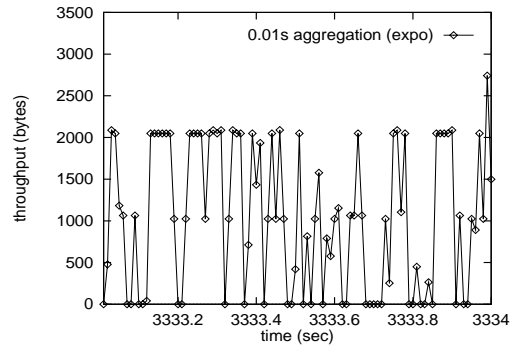
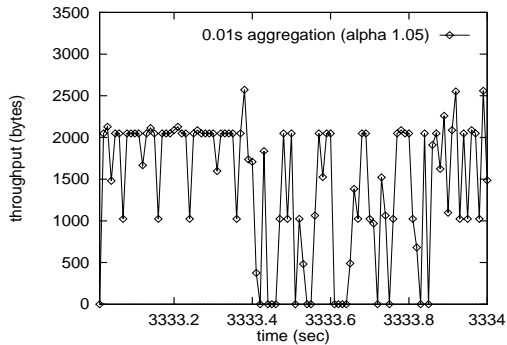
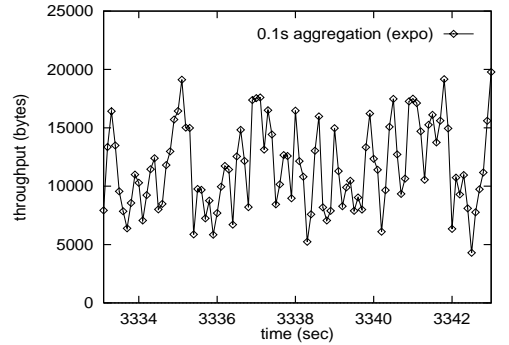
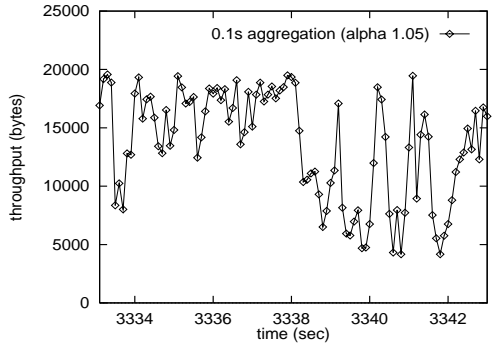
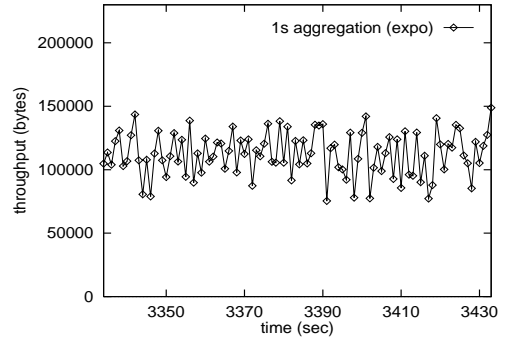
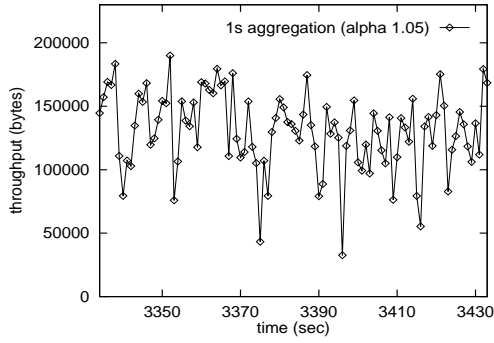
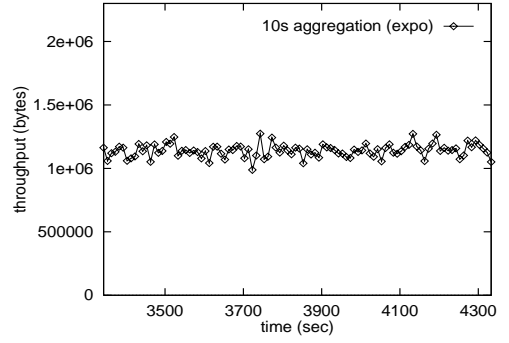
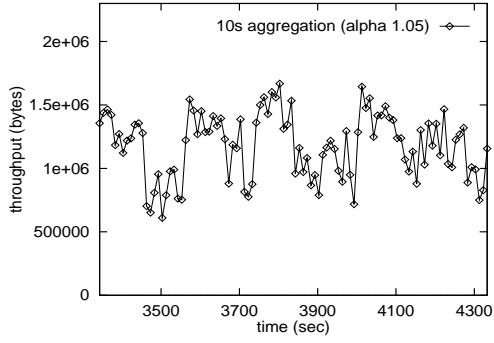
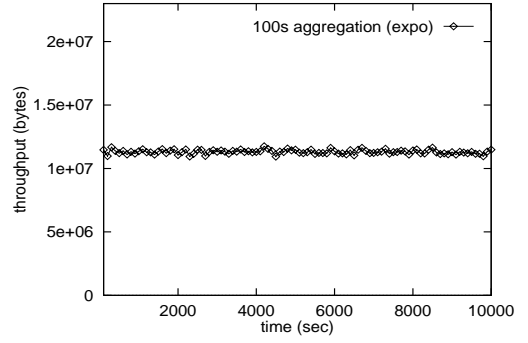
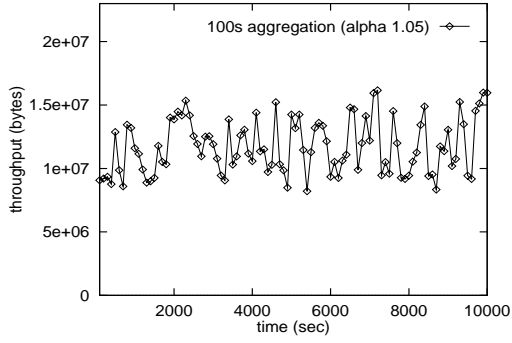
Cisco 7206VXR router: packet loss rate



- 8 classes
- OC-3 link
- varying offered load

Back to Internet traffic:

- what does it look like?
- doesn't become flat with time aggregation
- stays bursty!
- in a peculiar fashion: self-similar or fractal

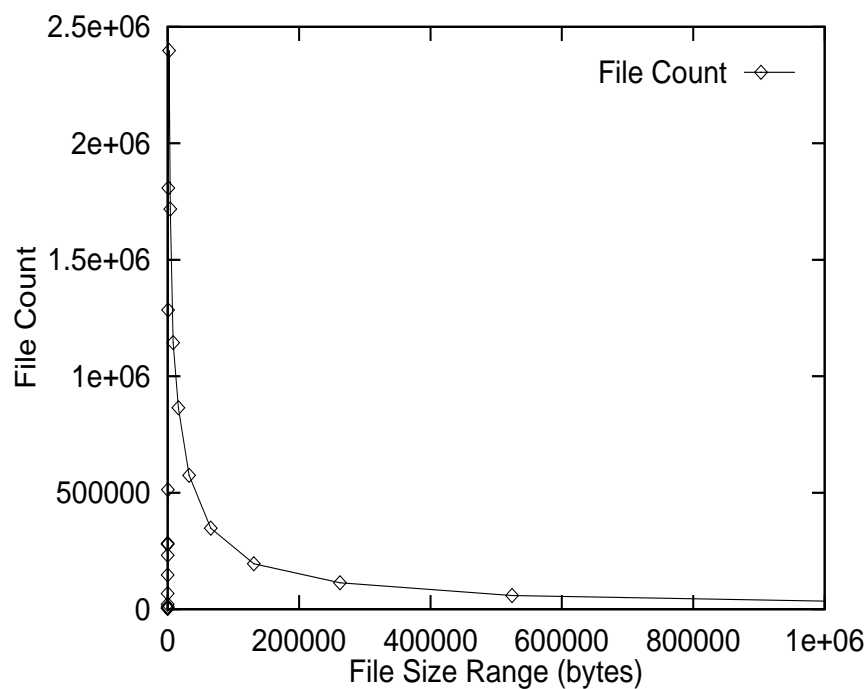


Consequences:

- cannot use “flat is good” method anymore
- intrinsic trade-off between QoS and efficiency
- bad news for QoS provisioning
- traffic must be correlated in time (why?)

Tale of elephants and mice:

→ UNIX and WWW file systems



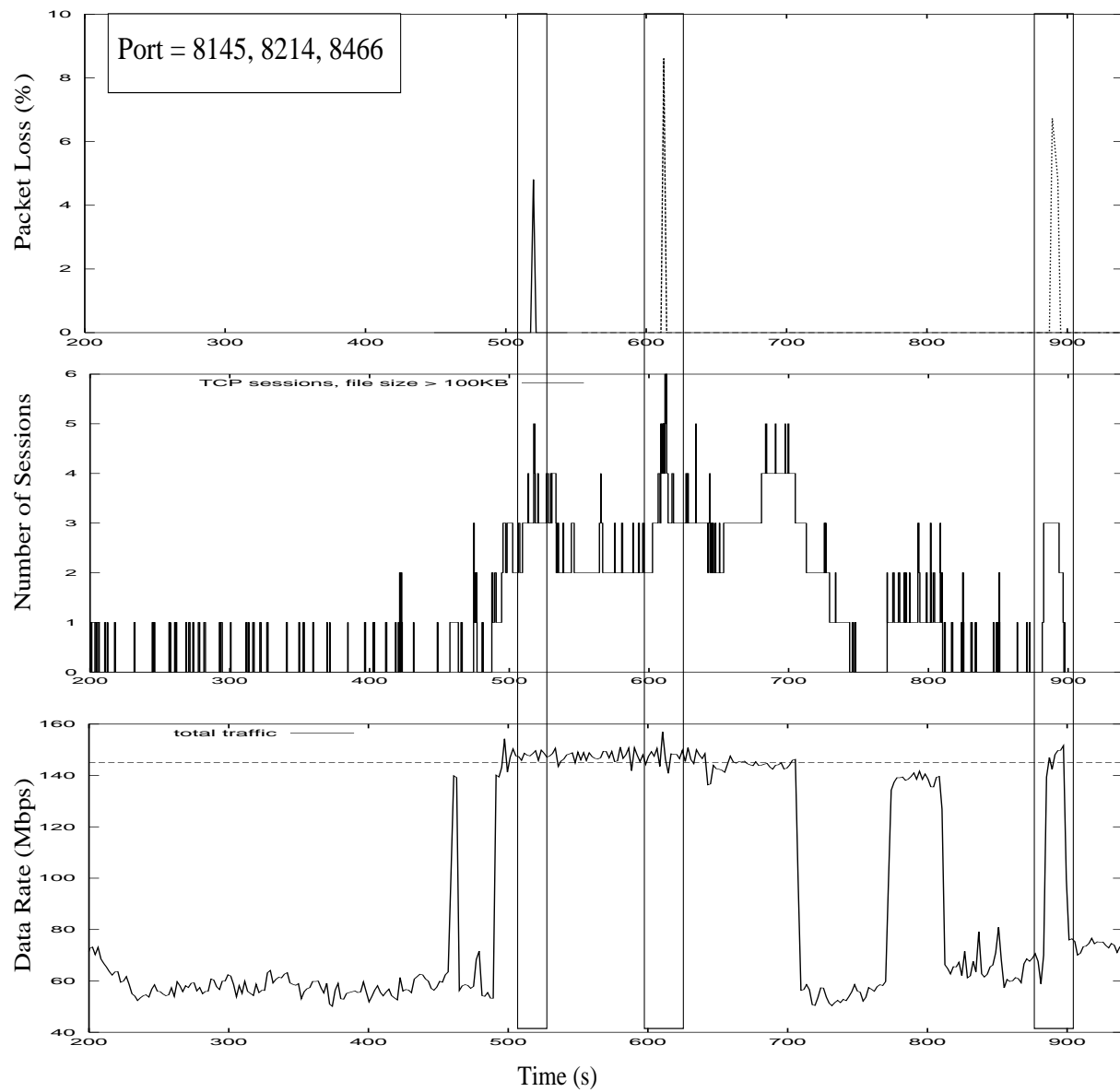
→ many small ones (mice)

→ a few very large ones (elephants)

→ 10% consumes 90% of bandwidth

Elephants in action:

→ at backbone router



Can the problem be solved?

→ no: as long as elephants and mice holds

Turns out to be a wide-spread phenomenon is sociology, networks, and elsewhere

→ size (population) of cities

→ popularity

→ frequency of words in books

→ etc.

In the real world:

→ norm: skewed distribution of sizes

→ power-law