Network Measurement

Brighten Godfrey
CS 538 April 26 2017

Science of network measurement

Measurement goes back to the inception of the Internet

By the mid-1990s: Internet and its protocols were big, wild, organic

- Complex system: hard to predict global effects of interacting components
- Distributed multi-party system: can't see everything that's happening

Network measurement moves from "just" monitoring to a science

Challenge #1: Emergent behavior

Example: Model packet arrivals over time at a link

Simplest common model: Poisson process

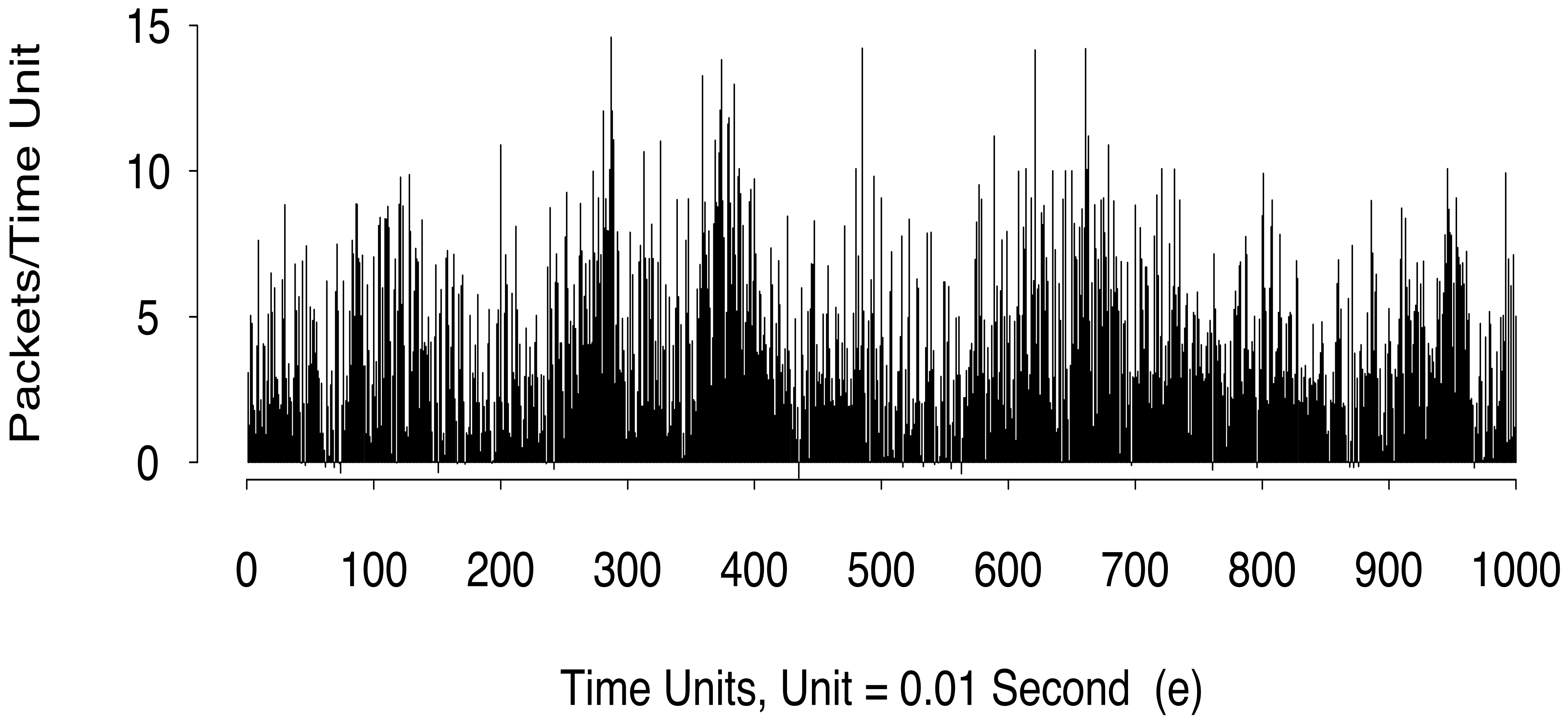
- Parameter: rate λ (mean arrivals per unit time)
- Pr[time till next arrival > t] = $e^{-\lambda t}$ (exponential dist.)

Properties

- Memoryless: Even knowing entire history gives no clue as to next arrival time
- Number of arrivals in a given time interval concentrates around expected value

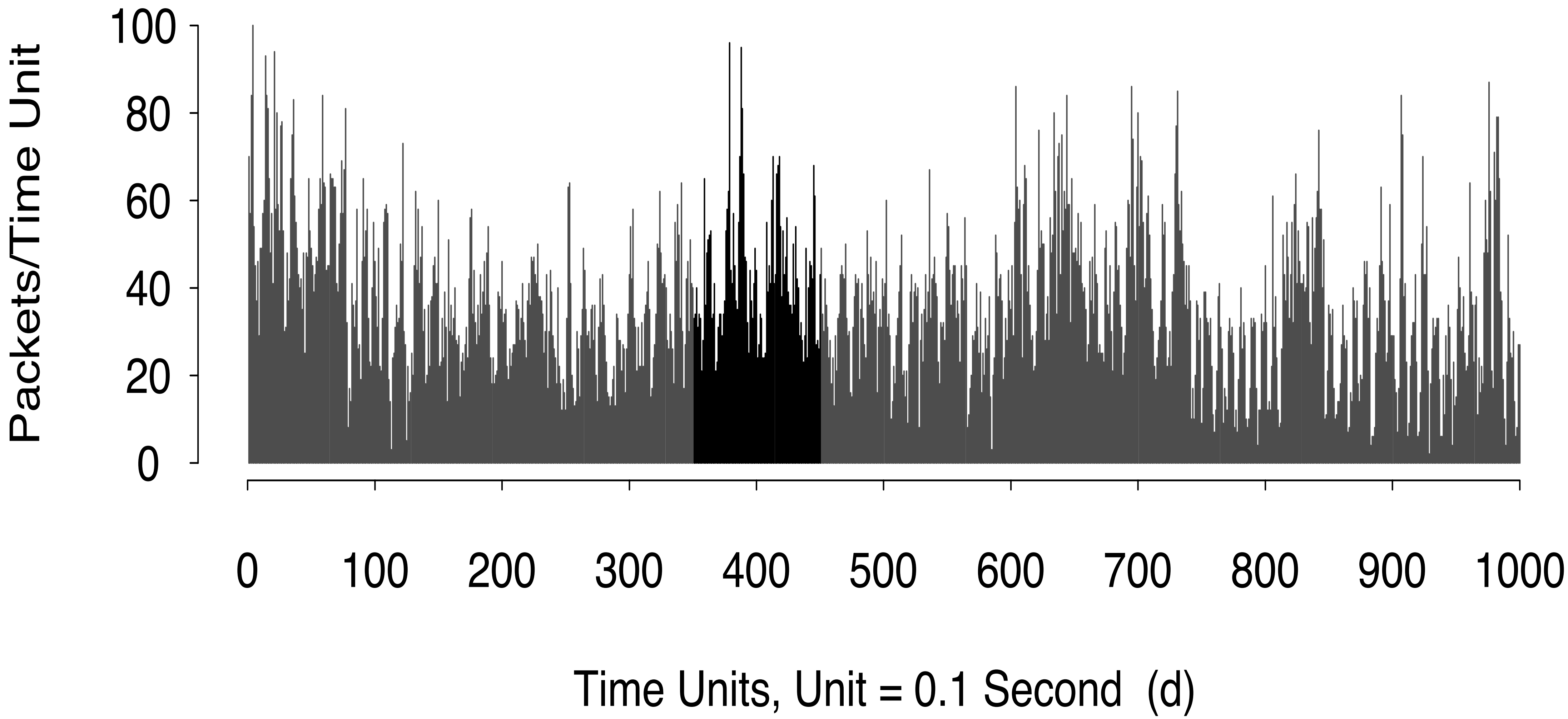
Temporal patterns of traffic

"On the Self-Similar Nature of Ethernet Traffic"
Leland, Taqqu, Willinger, Wilson, SIGCOMM 1993



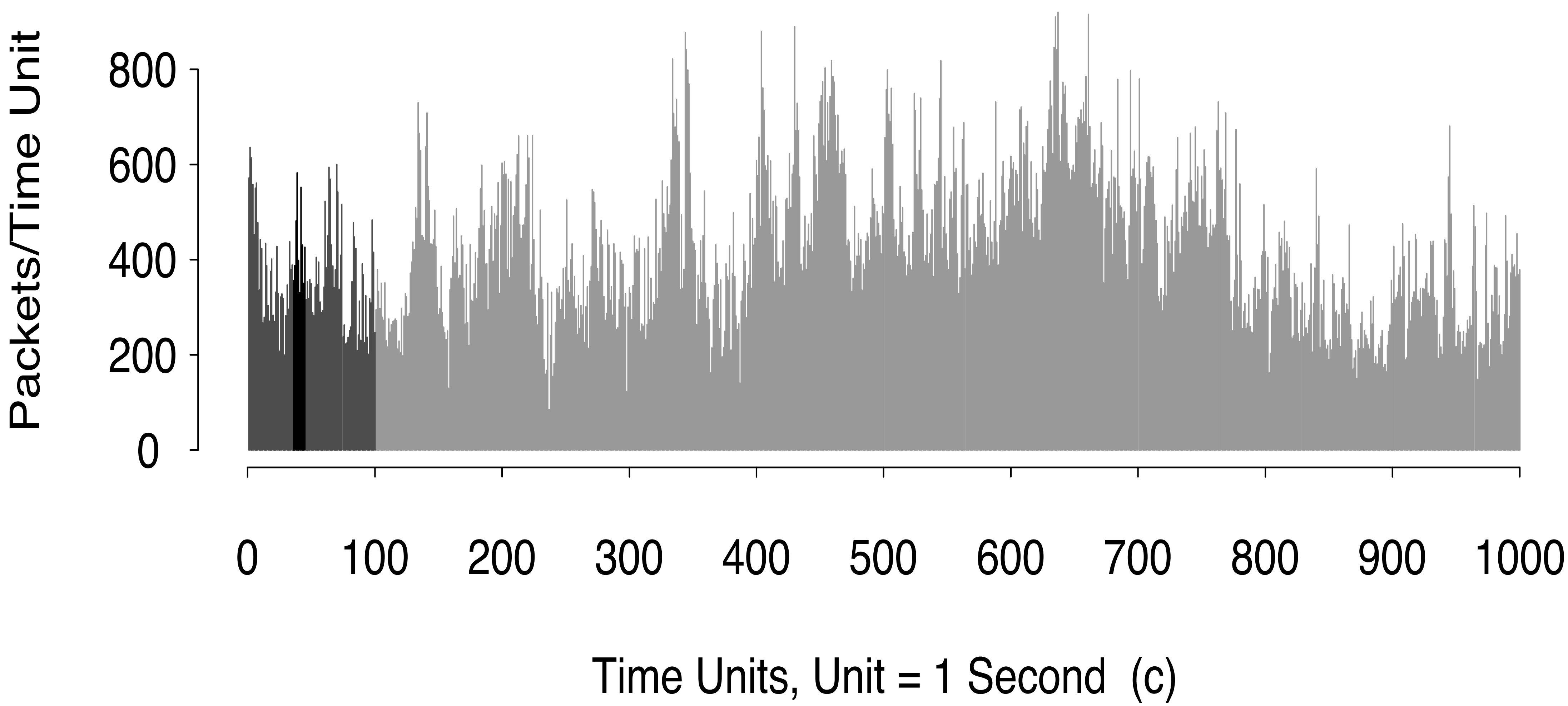
Temporal patterns of traffic

"On the Self-Similar Nature of Ethernet Traffic"
Leland, Taqqu, Willinger, Wilson, SIGCOMM 1993



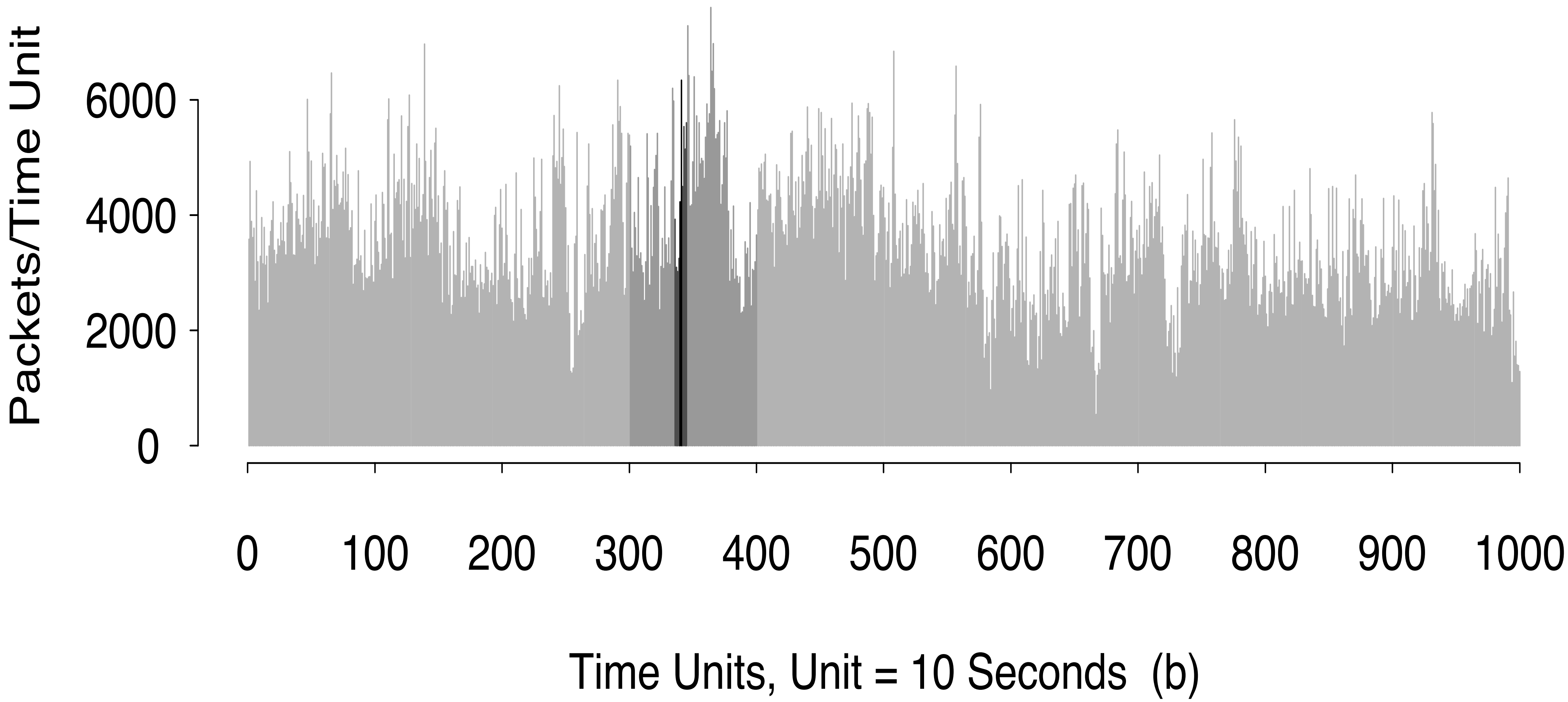
Temporal patterns of traffic

"On the Self-Similar Nature of Ethernet Traffic"
Leland, Taqqu, Willinger, Wilson, SIGCOMM 1993



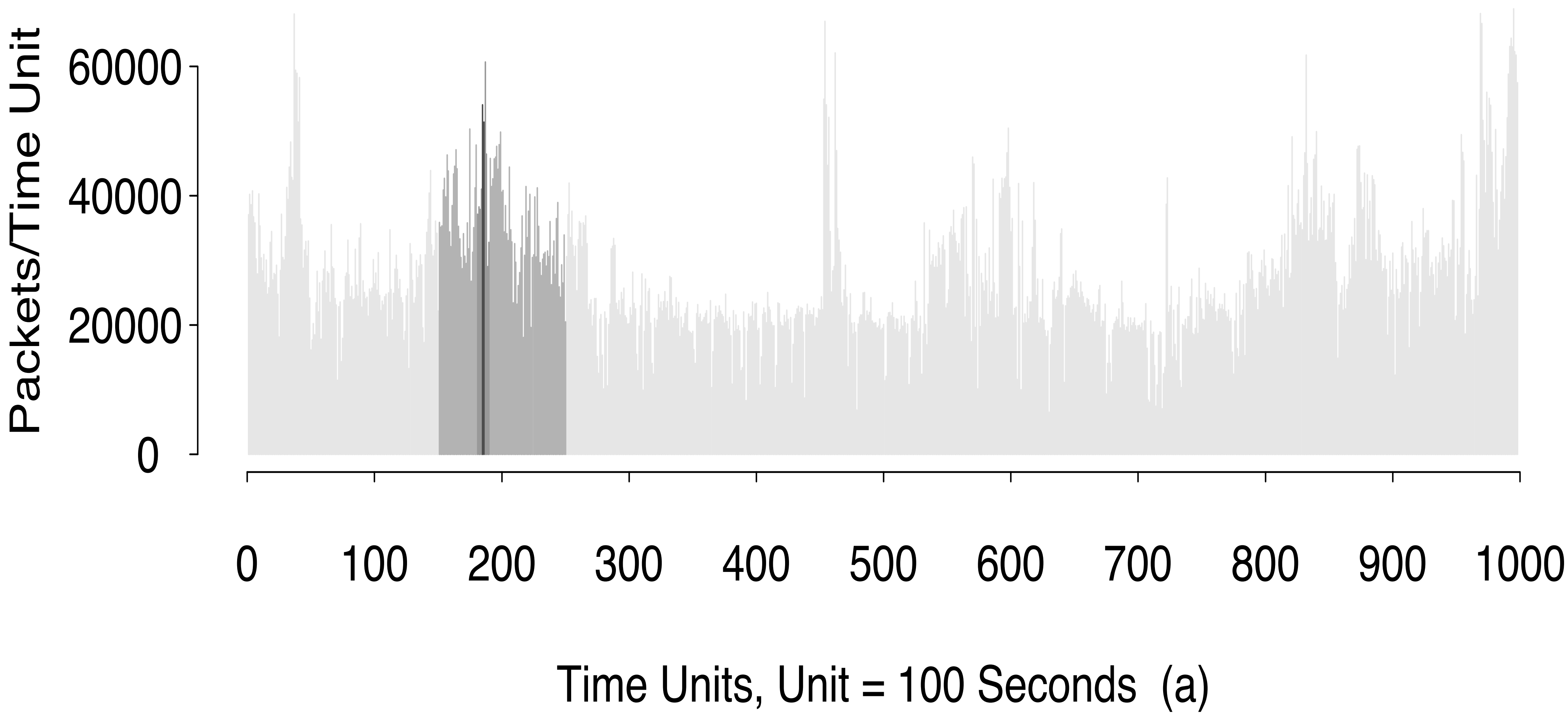
Temporal patterns of traffic

"On the Self-Similar Nature of Ethernet Traffic"
Leland, Taqqu, Willinger, Wilson, SIGCOMM 1993



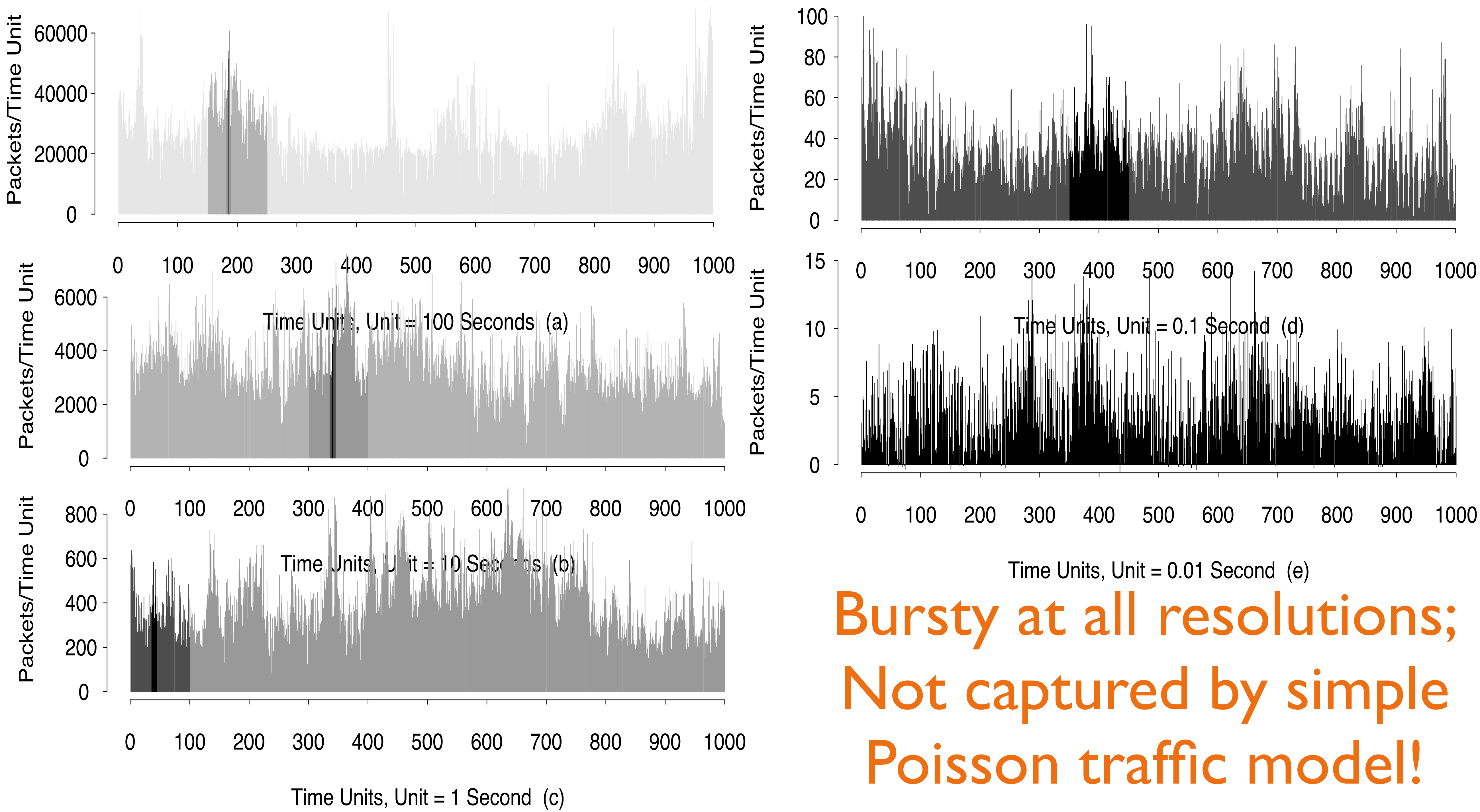
Temporal patterns of traffic

"On the Self-Similar Nature of Ethernet Traffic"
Leland, Taqqu, Willinger, Wilson, SIGCOMM 1993



Temporal patterns of traffic

"On the Self-Similar Nature of Ethernet Traffic"
Leland, Taqqu, Willinger, Wilson, SIGCOMM 1993



Bursty at all resolutions; Not captured by simple Poisson traffic model!

Challenge #2: Lack of visibility

Only a fraction of the system is visible

- For what we can observe, the cause is not obvious

Foundational work by Vern Paxson in the mid 1990s

- "End-to-End Routing Behavior in the Internet", SIGCOMM 1996
- Loops, asymmetry, instability
- Established Internet measurement methodology: "looking inside the black box" via end-to-end measurements

Name	Description
adv	Advanced Network & Services, Armonk, NY
austr	University of Melbourne, Australia
austr2	University of Newcastle, Australia
batman	National Center for Atmospheric Research, Boulder, CO
bnl	Brookhaven National Lab, NY
bsdi	Berkeley Software Design, Colorado Springs, CO
connix	Caravela Software, Middlefield, CT
harv	Harvard University, Cambridge, MA
inria	INRIA, Sophia, France
korea	Pohang Institute of Science and Technology, South Korea
lbl	Lawrence Berkeley Lab, CA
lbli	LBL computer connected via ISDN, CA
mid	MIDnet, Lincoln, NE
mit	Massachusetts Institute of Technology, Cambridge, MA
ncar	National Center for Atmospheric Research, Boulder, CO
near	NEARnet, Cambridge, Massachusetts
nrao	National Radio Astronomy Observatory, Charlottesville, VA
oce	Oce-van der Grinten, Venlo, The Netherlands
panix	Public Access Networks Corporation, New York, NY
pubnix	Pix Technologies Corp., Fairfax, VA
rain	RAINet, Portland, Oregon
sandia	Sandia National Lab, Livermore, CA
sdsc	San Diego Supercomputer Center, CA
sintef1	University of Trondheim, Norway
sintef2	University of Trondheim, Norway
sri	SRI International, Menlo Park, CA
ucl	University College, London, U.K.
ucla	University of California, Los Angeles
ucol	University of Colorado, Boulder
ukc	University of Kent, Canterbury, U.K.
umann	University of Mannheim, Germany
umont	University of Montreal, Canada
unij	University of Nijmegen, The Netherlands
usc	University of Southern California, Los Angeles
ustutt	University of Stuttgart, Germany
wustl	Washington University, St. Louis, MO
xor	XOR Network Engineering, East Boulder, CO

[Paxson's vantage points]

Collateral Damage of Censorship

"The Collateral Damage of Internet Censorship by DNS Injection" [Anonymous, CCR 2011]

Several moving parts; let's look in detail…

Discussion

We typically use many vantage points in order to "see inside the black box" of the Internet. Where were their vantage points?

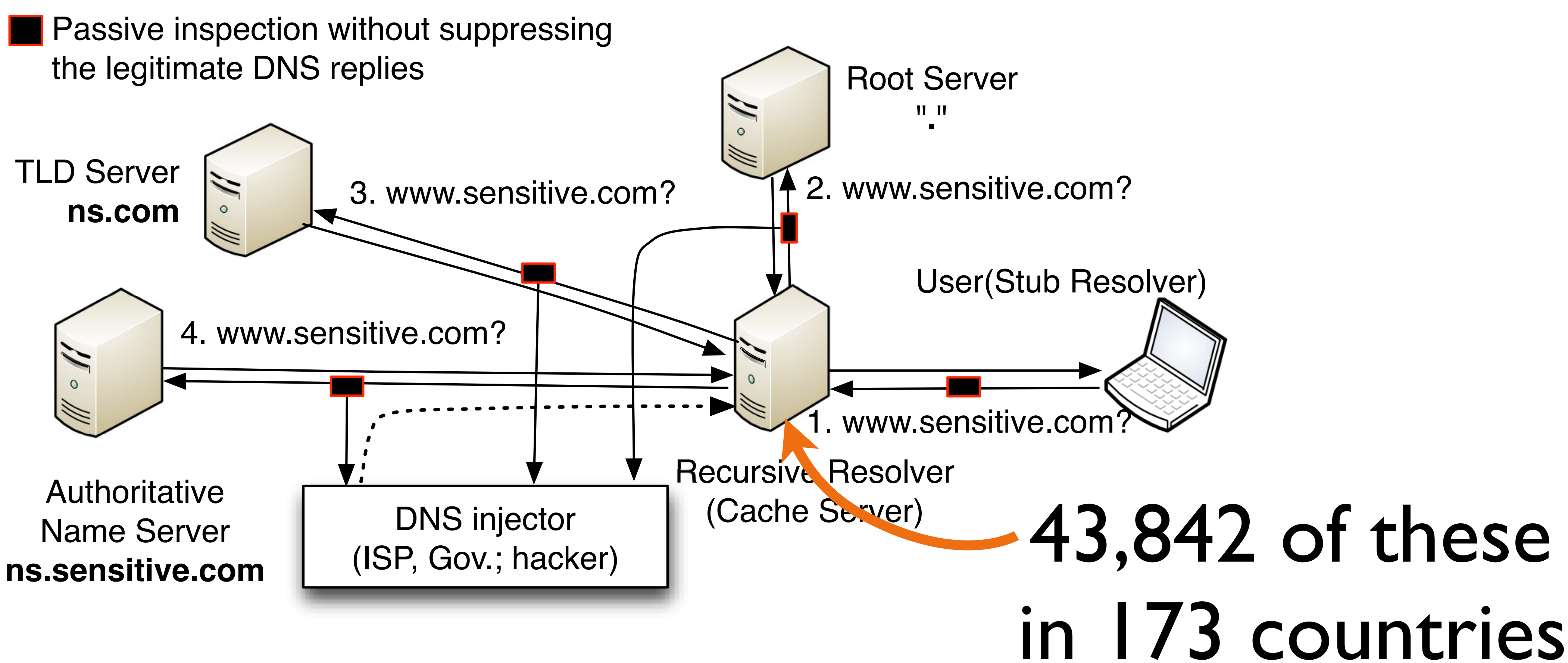


Figure 1: DNS query process and DNS injection

Discussion

How could you counteract this censorship?

How could service providers offer protection?

How could an individual client protect itself?

Discussion

How could you counteract this censorship?

How could service providers offer protection?

- Censor avoids polluting transit queries
- Threat of depeering
- DNSSEC
  - signed DNS responses
  - requires widespread deployment

How could an individual client protect itself?

- DNSSEC
- Query multiple servers, wait for all responses
- Tunnel queries through a friend in another country

Towards a Comprehensive Picture

Key points

- Centrally managed, consistent across nodes
- Pervasive (99.9% polluted)
- At one node
  - Load balancing based on (src, dst) IP across 360 processes
  - 2800 censored responses per sec

Towards a Comprehensive Picture

Discussion

- “Our results may overestimate the GFW injector locations due to the problem of false negatives”
  - If packets are dropped, wouldn’t that cause us to miss a polluted response and *underestimate* GFW locations?

A word of caution

"The most important difference between computer science and other scientific fields is that: We build what we measure. *Hence, we are never quite sure whether the behavior we observe, the bounds we encounter, the principles we teach, are truly principles from which we can build a body of theory, or merely artifacts of our creations. ... this is a difference that should, to use the vernacular, 'scare the bloody hell out of us!'"*

– John Day

Announcements

Next time: Future Internet architecture

Last class: Wed May 3