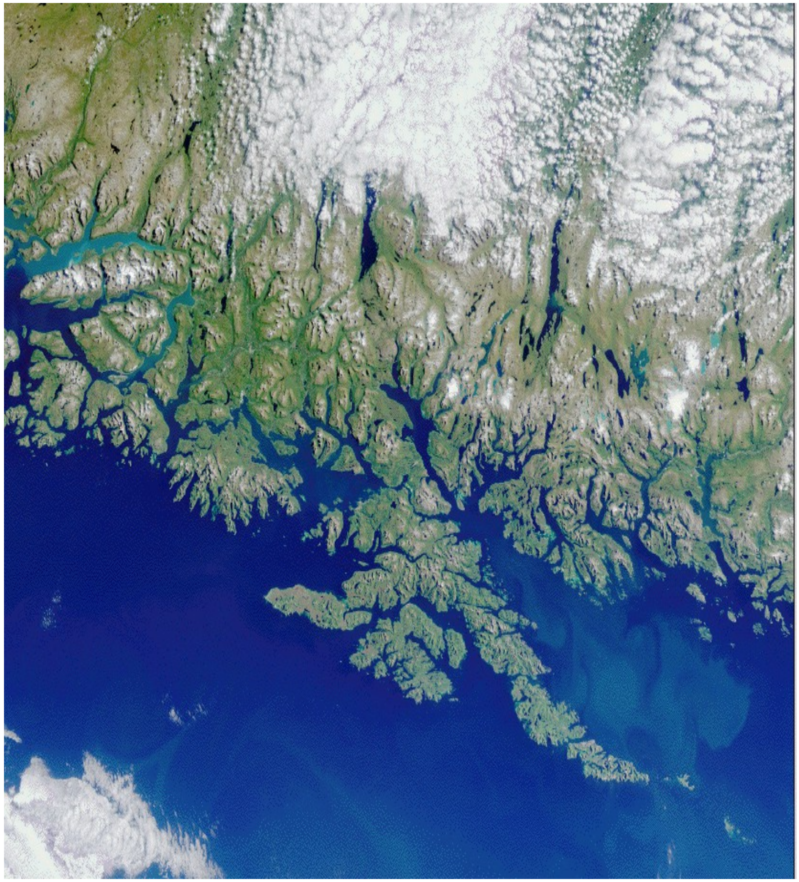
Self-similarity

Self-similarity

a self-similar object is exactly or approximately similar to a part of itself

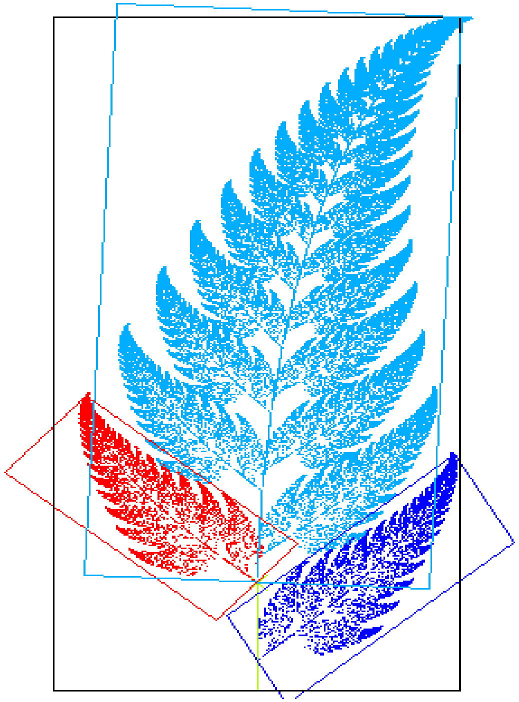


Examples: Nature (rivers, coastline)

Self-similarity

a self-similar object is exactly or approximately similar to a part of itself

Romanesco broccoli



Fern

Snowflake

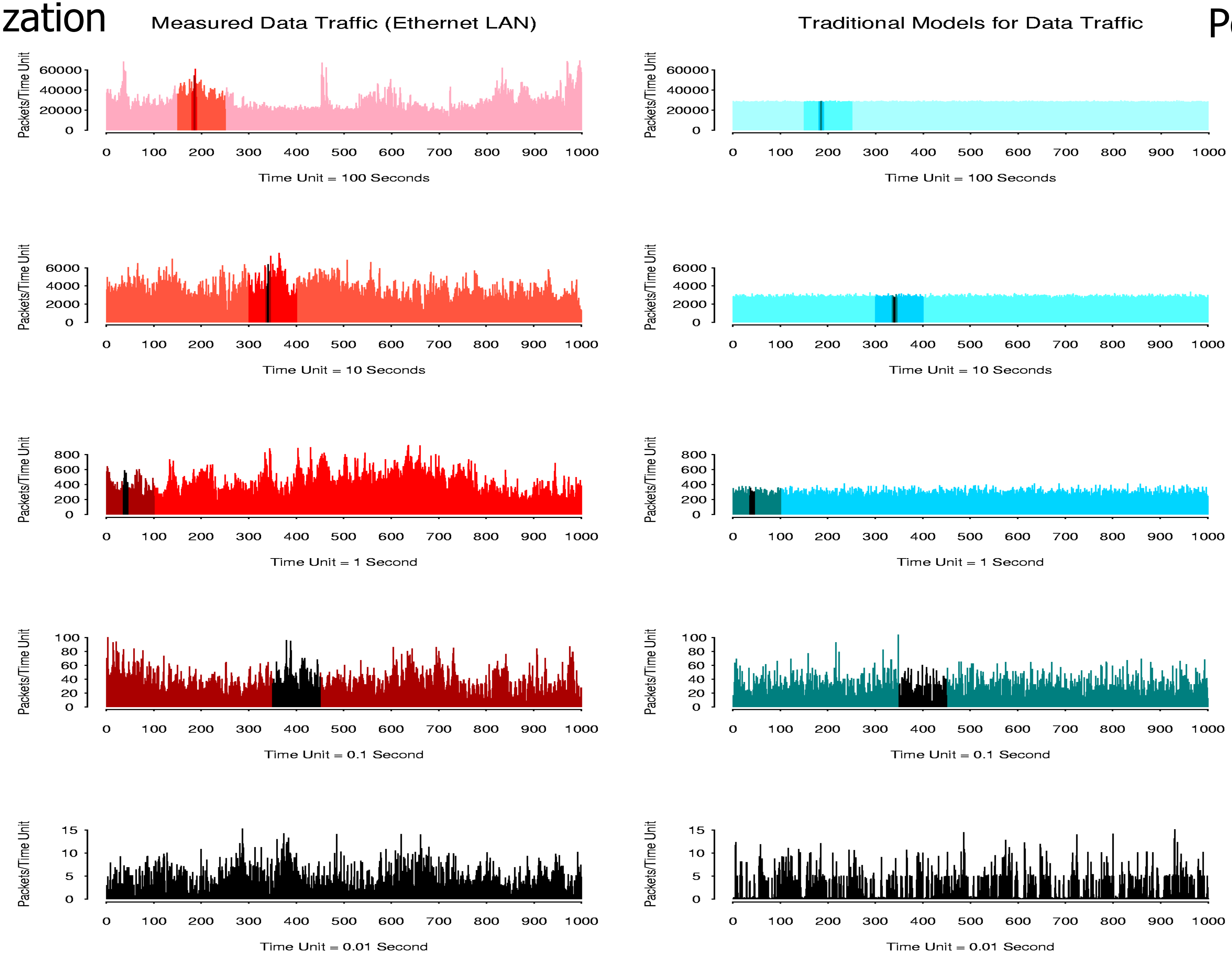
Self-similarity

- Just a mathematical concept?
- What does it mean?

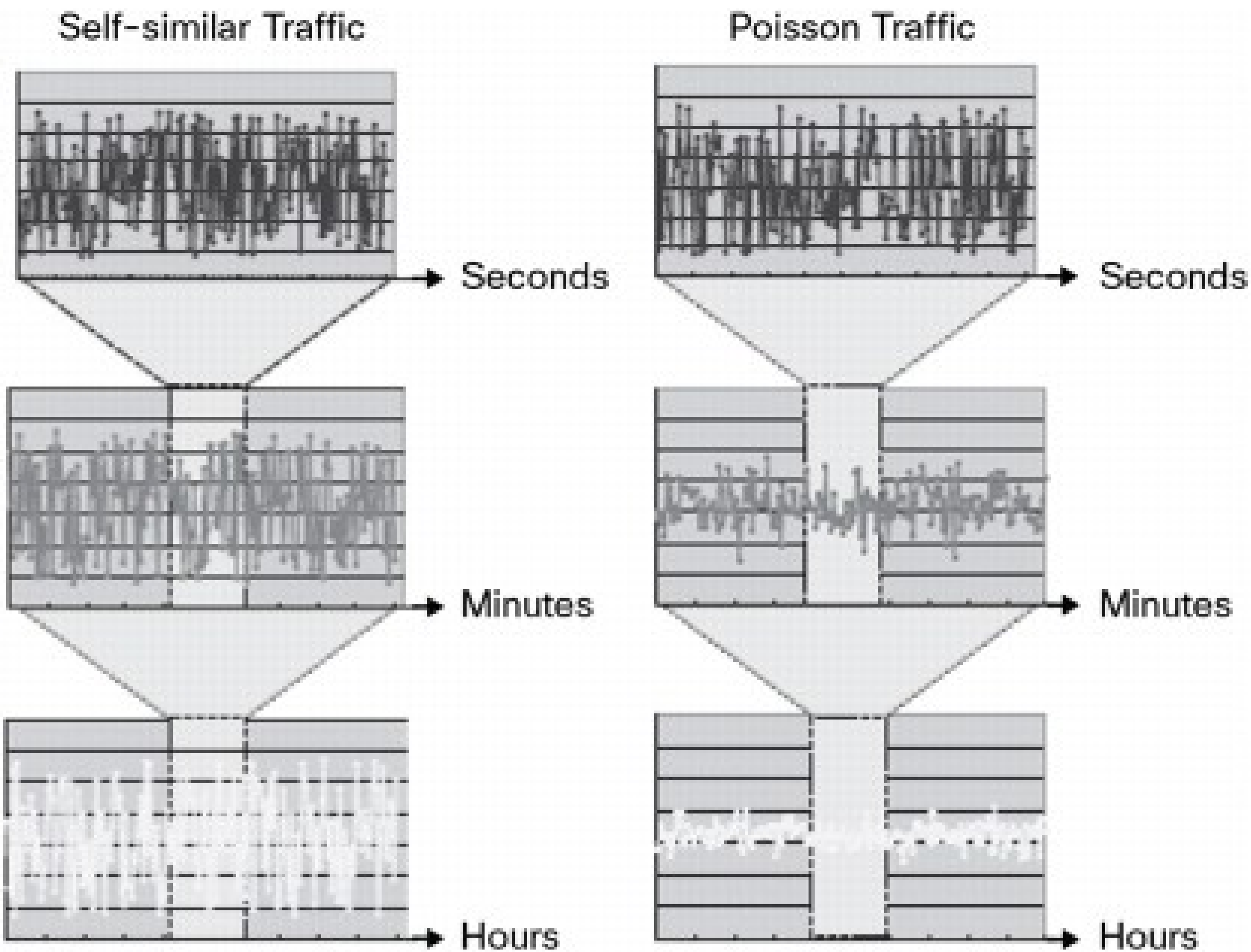
Self-similarity and Network Traffic (I)

Realization

Poisson



Self-similarity and Network Traffic (II)



Overview of Traffic Analysis

Question:
Capacity planning
Traffic Engineerng
Performance Evaluation

Typical Modeling:
Non-stationary Models
Stationary Models
Signal+noise
Self-Similarity
Multi-fractal

Timescale:
Months
Days
Hours
Seconds
mSec
μSec

Factors:
“trends”: Long-term changes
In demand, diurnal/ weekly patterns
High variability in packet trains/ connections
TCP’s closed loop
Queue Fluctuation

Aggregate traffic - self-similarity (I)

Intuition: self-similar processes "look the same" at all (i.e., over a wide range of) time scales

Def.: aggregate (or m-scaled) process $X^{(m)}$

Let $X=(X_t : t= 0,1,2,\ldots)$ be a stationary stochastic process, then

$$X^{(m)} = (X_k^{(m)} : k=1,2,\ldots)$$

where

$$X_k^{(m)} = 1/m(X_{km-m+1} + \ldots + X_{km}), k>0.$$

*stationary process: stochastic process whose probability distribution does not change when shifted in time or space

Aggregate traffic - self-similarity (II)

Intuition: self-similar processes “look the same” at all (i.e., over a wide range of) time scales

Def.: A zero mean stationary process* X is called self-similar (with self-similarity parameter H), if for all m $\geq$ 1,

$m^H X$ has the same Distribution as $X^{(m)}$

H is called the Hurst parameter. To be self-similar, 0.5 < H < 1

In networks: [LTWW94] LAN traffic is consistent with self-similarity as well as WAN traffic [PF95] and web traffic [CB96]

Aggregate traffic - self-similarity (III)

Intuition: self-similar processes "look the same" at all (i.e., over a wide range of) time scales

Def.: A zero mean stationary process* $X = (X_k : k \geq 0)$ is called self-similar (with self-similarity parameter H),
if for all $m \geq 1$,

$m^H X$ has the same Distribution as $X^{(m)}$

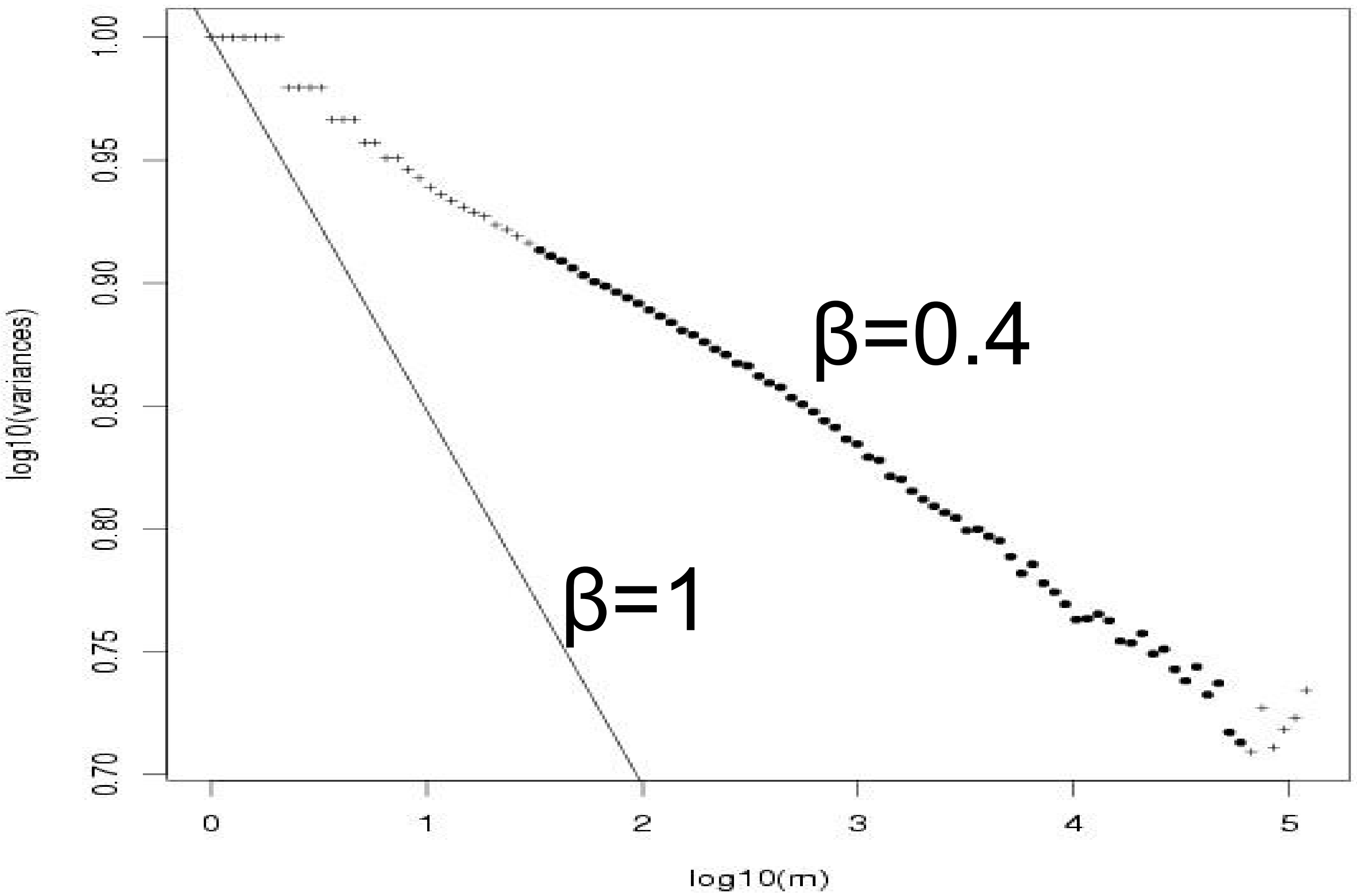
H is called the Hurst parameter. To be self-similar, $0.5 < H < 1$

Test of *variance*, i.e., the measure of how far a set of samples is spread out, $var(X)=E((X - E(X))^2)$.

$$\mathrm{var}\left(X^{(m)}\right) \sim c m^{-2H-2} \; as \; m \rightarrow \infty$$

Variance time plot

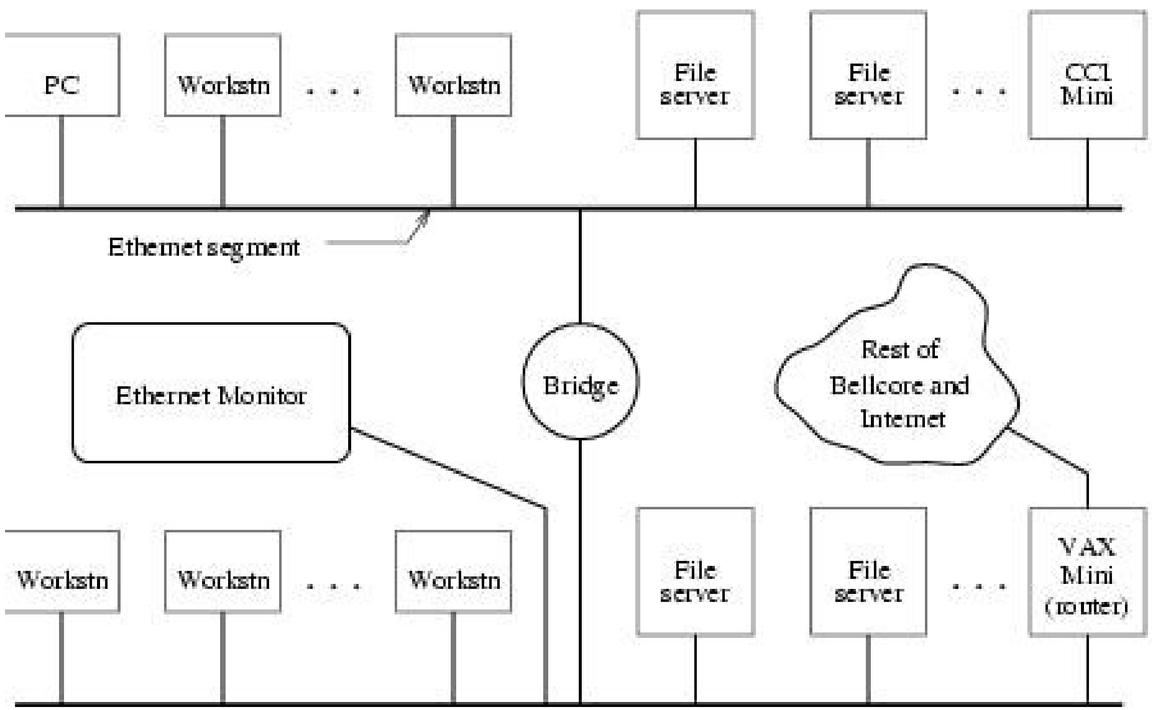
H=1-β/2 =0.8 in (0.5,1)



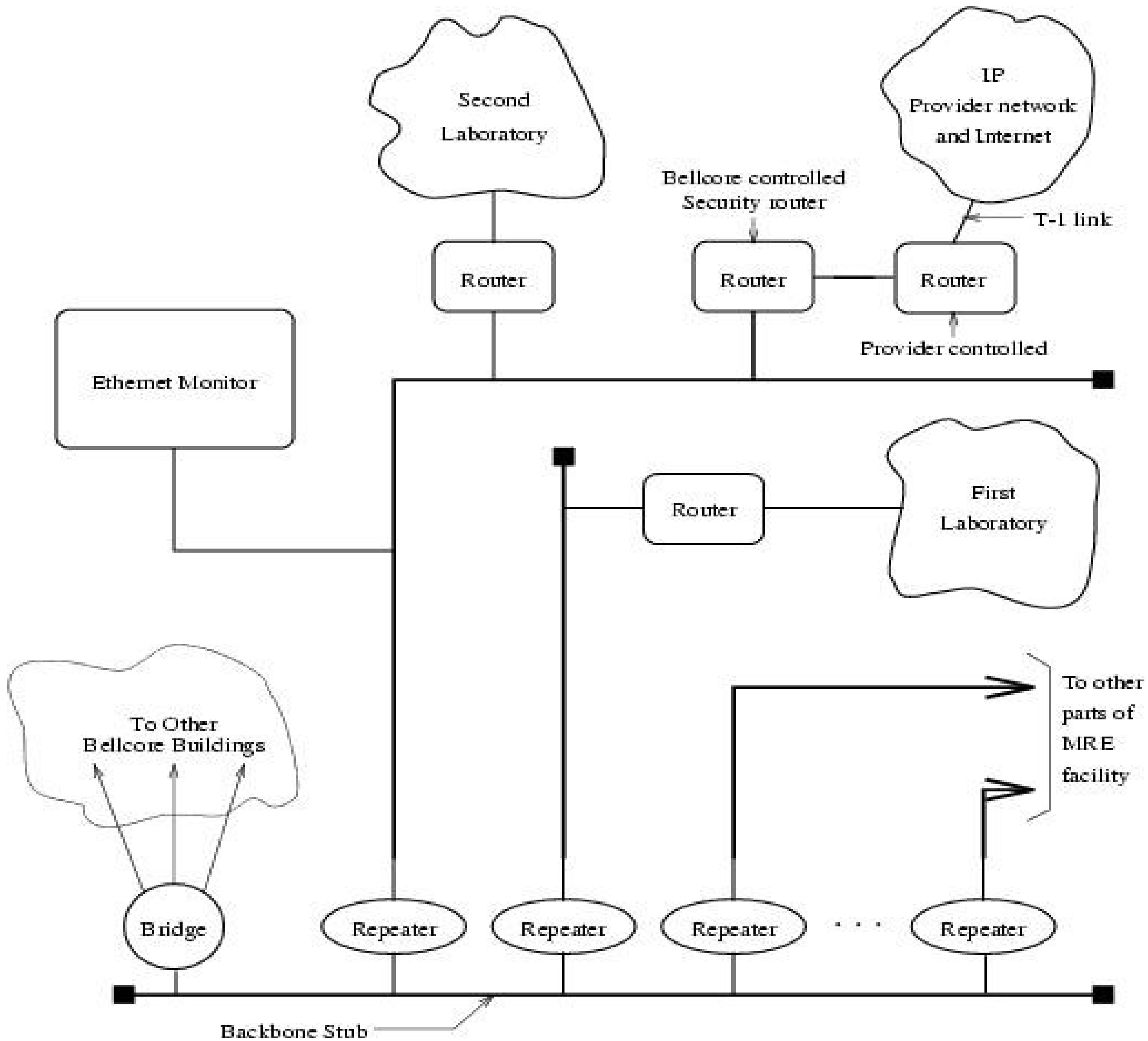
Self-similarity

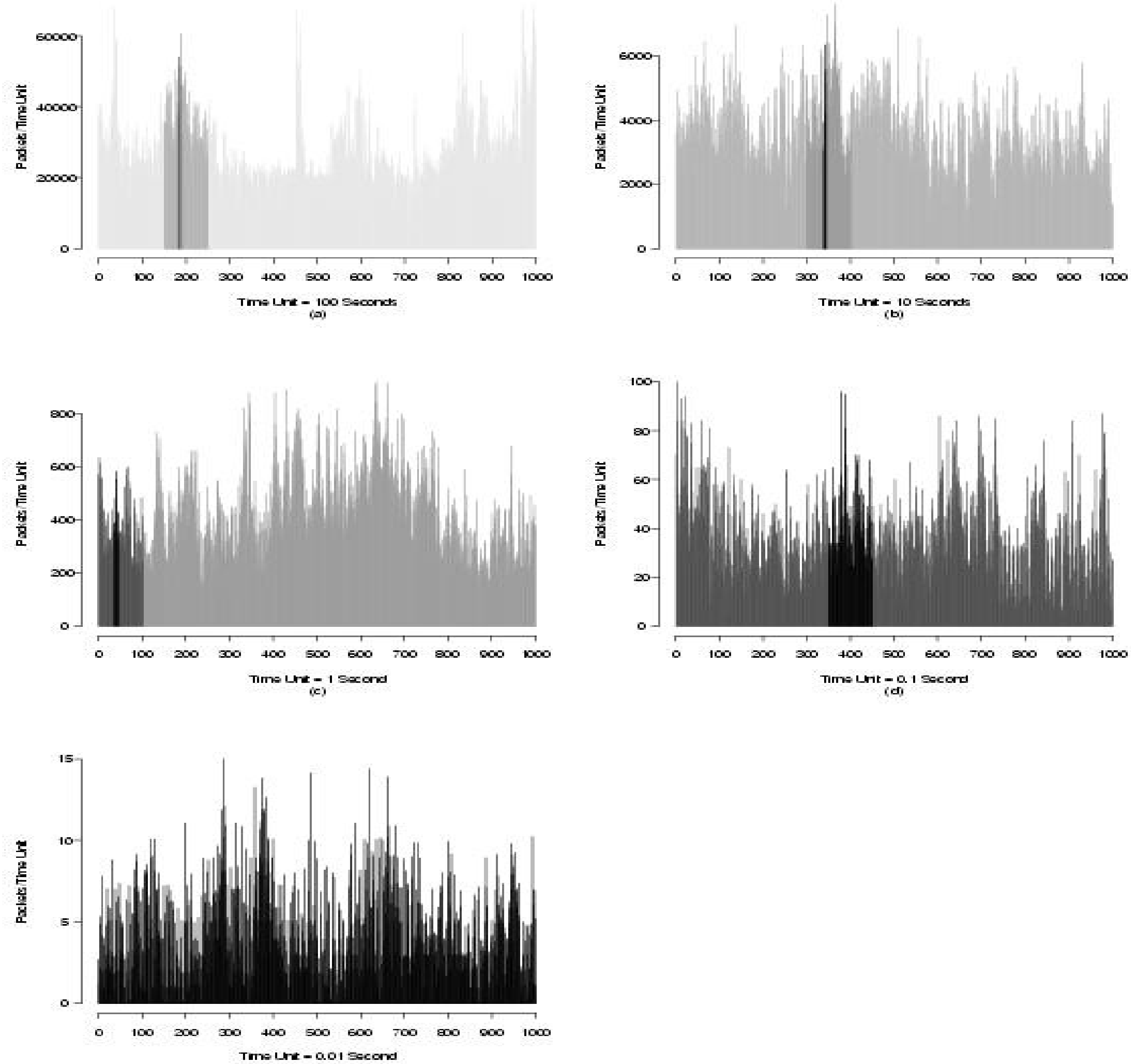
- Just a mathematical concept?
- What does it mean?

Network topology 1989

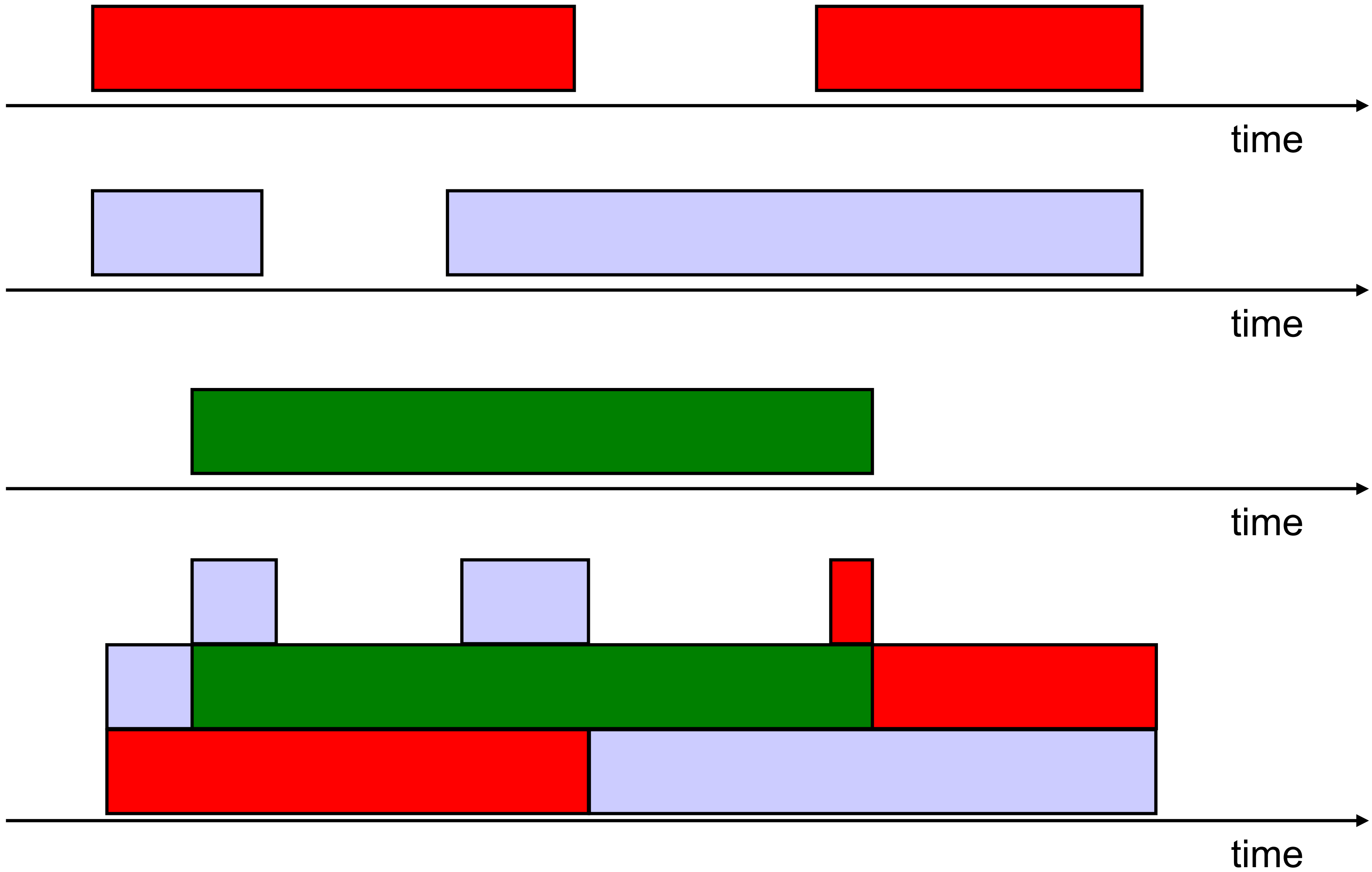


Network topology 1992




Packets/Time Unit
Time Unit = 100 Seconds
(a)
Time Unit = 10 Seconds
(b)
Time Unit = 1 Second
(c)
Time Unit = 0.1 Second
(d)
Time Unit = 0.01 Second
(e)

Route Causes? Superposition of sources (e.g., computer processes, human activity)



Self-similarity via heavy tails

Self-similar:

Superposition of independent ON/OFF sources is self-similar, if durations of periods are heavy-tailed with infinite variance

Short-range Dependent:

Superposition of independent ON/OFF sources is short-range dependent, if durations of periods are light-tailed

Self-similarity

- Light-tailed and heavy-tailed distribution
- Short-range dependence and long-range dependence

Covariance

- Is the measure of how much two random variables change together
- Given two random variables x, y with means μ_x and μ_y, their covariance is:

$$Cov(x,y)=\sigma_{xy}^2=E\left[(x-\mu_x)(y-\mu_y)\right]$$

- Their correlation coefficient is the normalized covariance is a measure of linear dependence

$$Cor(x,y)=\rho_{xy}=\frac{\sigma_{xy}^2}{\sigma_x\sigma_y}$$

Autocorrelation Function (ACF)

- Is a Cross-correlation of a process with itself
- It describes the correlation between values of the process at different times, as a function of the two times or time difference

$$r(k) = \frac{E\left(X_t - \mu\right)\left(X_{t+k} - \mu\right)}{\sigma^2}$$

- If ACF ~ 0 then traffic observations that are far apart are independent.

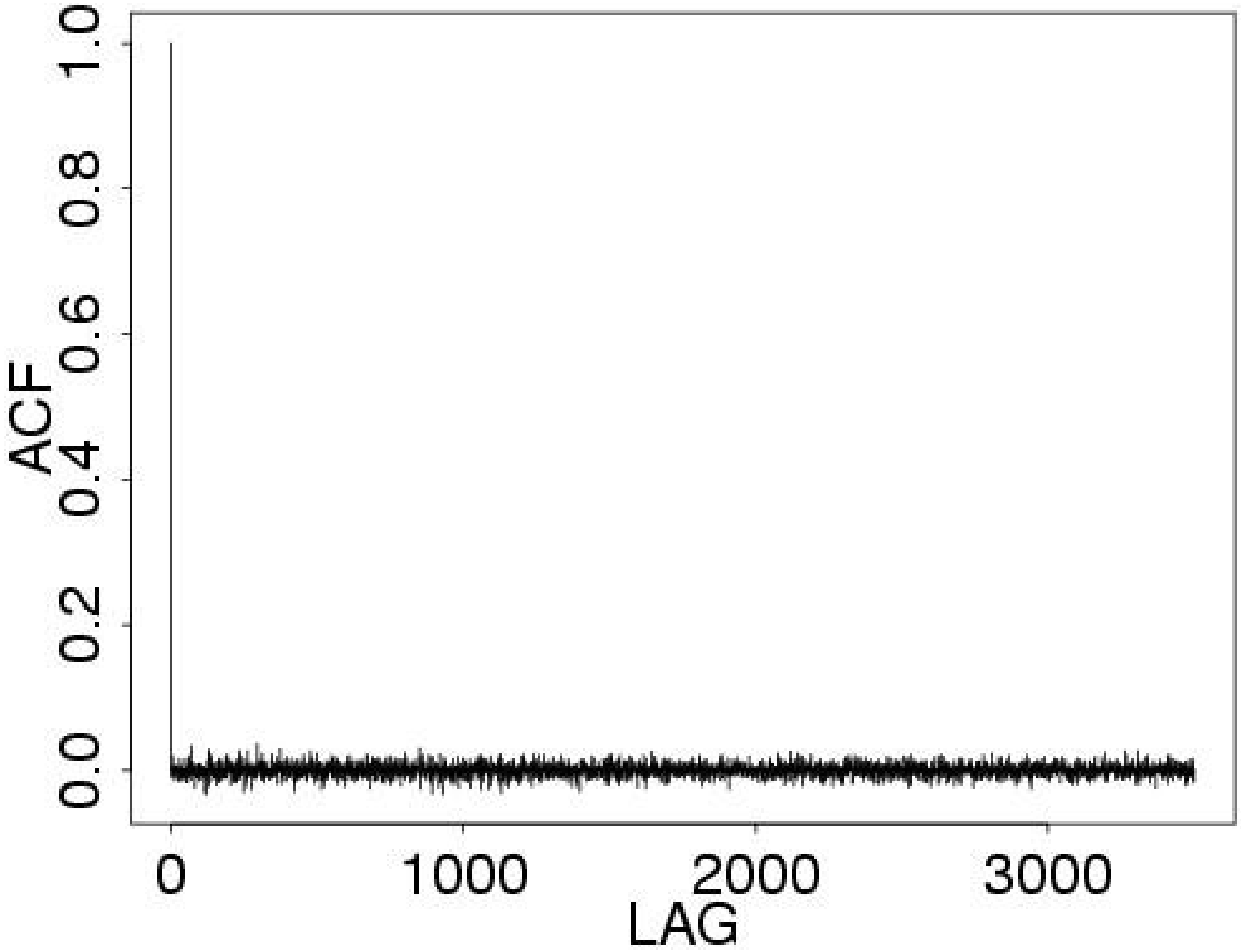
Short-Range Dependence (SRD)

- A stationary process X with autocorrelation function r(k), $k \geq 1$, exhibit short-range dependence (SRD) if there exists $0 < \rho < 1$ and $\tau > 0$ with

$$r(k)\tau\rho^{-k} \rightarrow 0 \quad as \quad k \rightarrow \infty$$

- There are no observations "in the past" that are correlated with current observations
- Important feature: Autocorrelations decay (at least) exponentially fast for large lags k

Poisson process: an SRD processes



Short-range dependence (SRD)

- The aggregated process $X^{(m)}$ tends to second-order white noise, as

$$r^{(m)}(k) \rightarrow 0 \ as \ k \rightarrow \infty$$

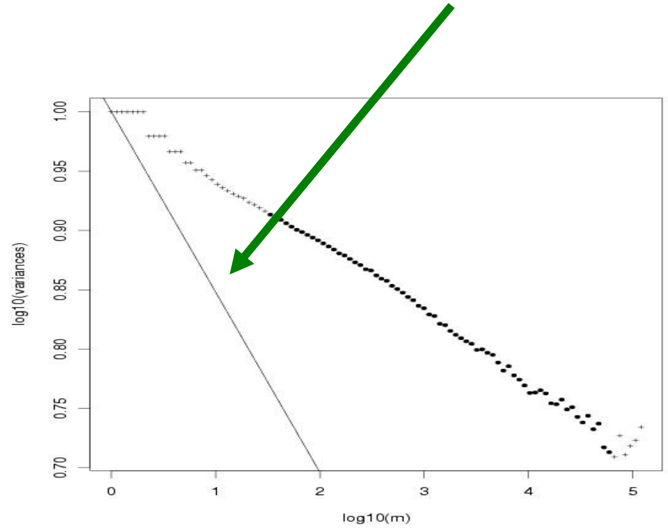
where $r^{(m)}$ denotes autocorrelation function of $X^{(m)}$

- The variance-time function, i.e., the variance of the sample mean, as a function of m, satisfies:

$$\operatorname{var}\left(X^{(m)}\right) \sim c m^{-1} \ as \ m \rightarrow \infty$$

Short-range dependence (SRD)

- Key features
  - Short range dependence = finite correlation length
  - Fluctuations over narrow range of time scales
  - Plotting var($X^{(m)}$) vs. m on log-log scale shows linear relationship for large m, with slope –1 (β=1)



Light-tailed distributions

- X random variable with distribution function F.
- F is said to be light-tailed if there exists c > 0

$$\left(1-F\left(X\right)\right)e^{cx} \rightarrow 0 \ as \ x \rightarrow \infty$$

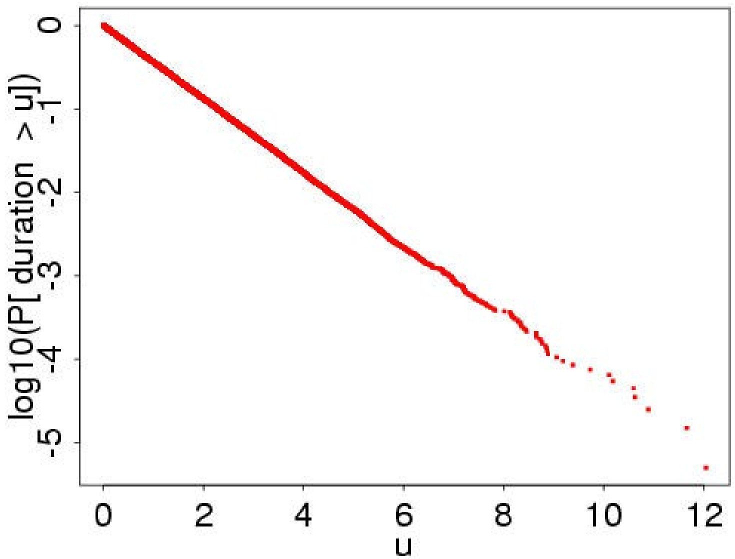
- Important feature: tails decay exponentially fast for large x; i.e.,

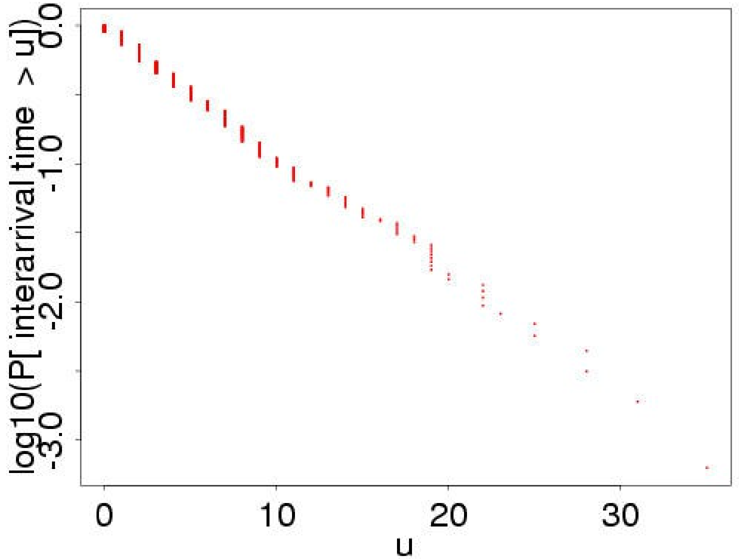
$$P\left[X>x\right]=1-F\left(X\right)\sim e^{-x} \ as \ x \rightarrow \infty$$

Where P[X>x] is the complementary cumulative distribution function (CCDF)

Light-tailed distributions

- Examples: Exponential, Normal, Poisson, Binomial
- Key features:
  - F has limited variability
  - F is tightly concentrated around its mean
  - F has finite moments
  - P[X > x] (CCDF) vs. x on log-linear scale is linear for large x





Summary of light-tails and SRD

- Distributional assumptions
  - Light-tails imply limited variability in space
- Assumptions about temporal dynamics
  - SRD implies limited variability over time
- Common characteristics of traditional traffic processes
  - Limited burstiness (in time and space)

Long-range dependence (LRD)

- A stationary process X with autocorrelation function r(k), exhibits long-range dependence (LRD) if for some 1/2 < H < 1

$$r(k) \sim ck^{2H-2} \quad as \ k \rightarrow \infty$$

- H is called the Hurst parameter
- Important features of LRD
  - Infinite correlation length
  - Fluctuations over all time scales
  - No characteristic time scale

Heavy-tailed distributions

- X random variable with distribution function F
- F is said to be heavy-tailed if there exists c > 0

$$P[X > x] \sim cx^{-\alpha} \text{ as } x \rightarrow \infty$$

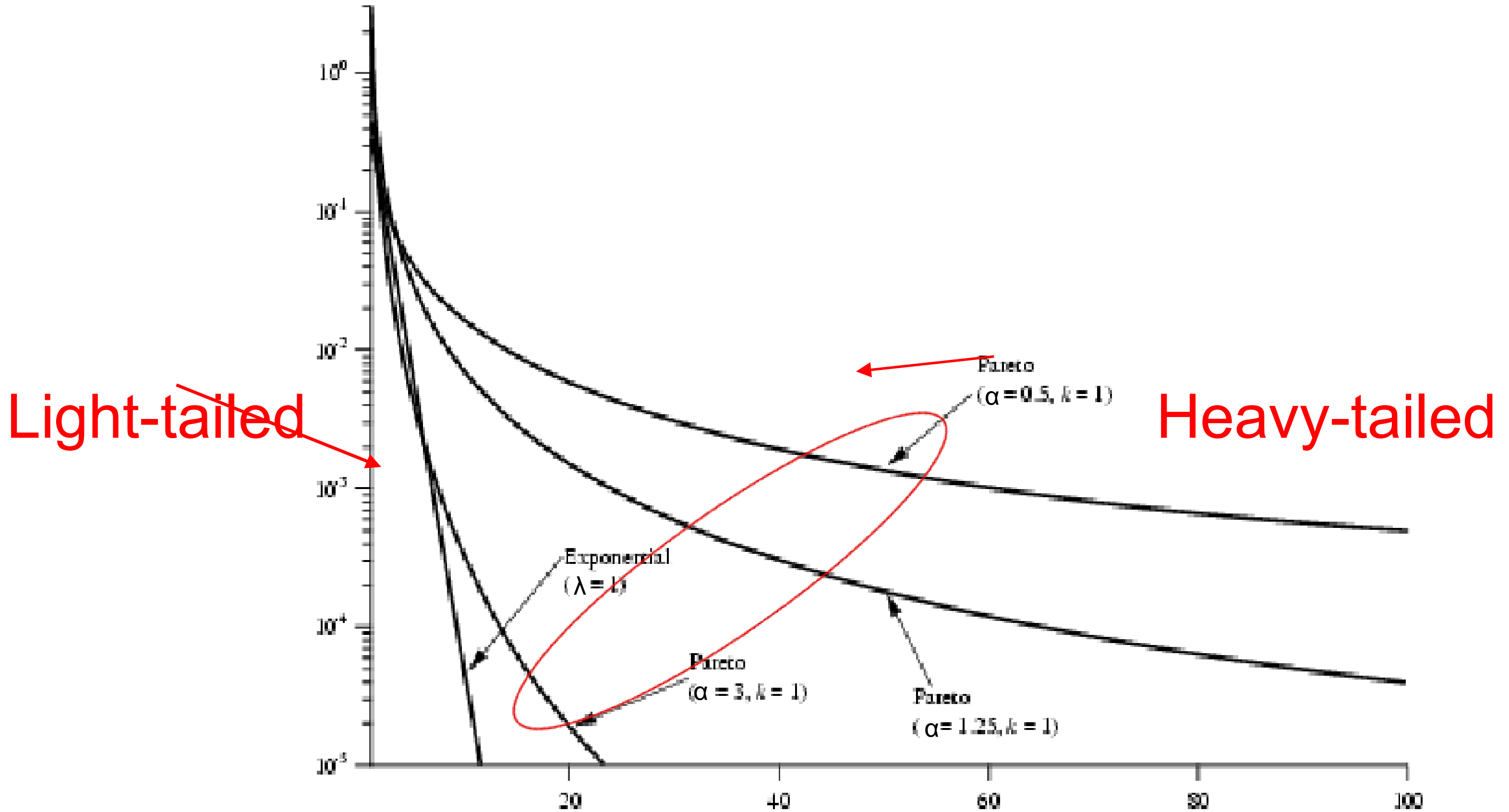
$1 < \alpha < 2$, X has finite mean but infinite variance.

→parsimonious model (small number of parameters)

- **Note:** Stationary models such as self-similarity/Heavy-tailed distributions and LRD are appropriate models for time-scale of seconds to minutes to hours. With them we model in (time) "invariants"

Heavy-tailed distribution

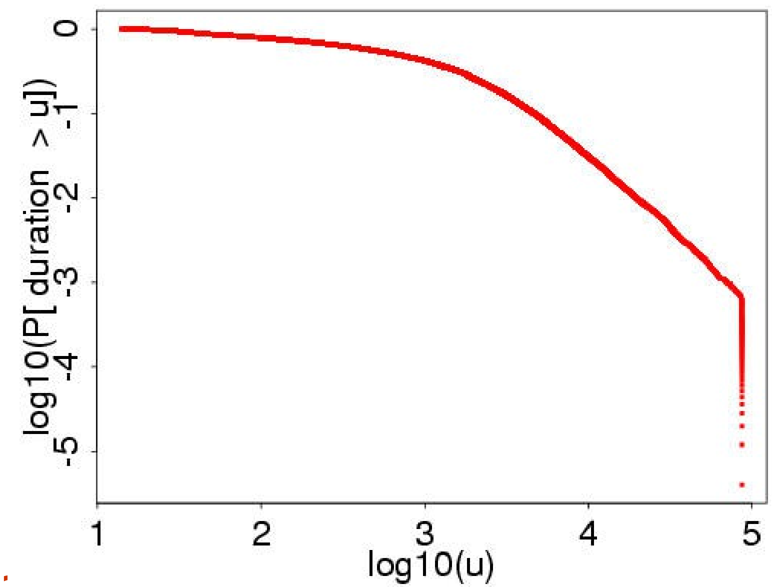
Ex. Pareto distribution: $p(x) = \alpha k^{\alpha} x^{-\alpha-1}$



if α≤2, the variance is infinit

if α≤1, the mean is also infinit

Heavy-tailed distributions



- Important features:
  - Finite mean but infinite variance
  - Heavy-tailed implies high variability
  - Tail decays like a power, hence power-law distribution
  - Plotting P[X > x] (CCDF) vs. x on log-log scale is linear for large x with slope α
- LRD is not a characteristic of only Heavy-tailed distributions

Long-range dependence (LRD) & Heavy-tail Distributions

- There are observations in the past that are correlated with current observations
- Parsimonious models available
- It changes the way we design systems

(e.g., how to deal with bursts, effect of queuing, protocol design)