Some background on Lab 1

Lab 1

- Lab 1 is about comparing a simple model for network traffic (Poisson traffic) with actual network traffic (LAN traffic, video traffic)

- Lab 1 retraces one fo the most fundamental insights of networking research ever:

> "Typical network traffic is not well described by Poisson model"

Poisson

- In a **Poisson process** with rate λ, the number of events in a time interval *(t, t+τ]*, denoted by *N(t+τ) – N(t)*, is given by

$$P[N(t+\tau)-N(t) = k] = \frac{(\lambda\tau)^k}{k!}e^{-\lambda\tau} \qquad ,k = 0, 1, \ldots,$$

- In a Poisson process with rate λ, the time between events follows an **exponential distribution:**

$$P[\text{Time between two events} \leq X] = 1 - e^{-\lambda X}$$

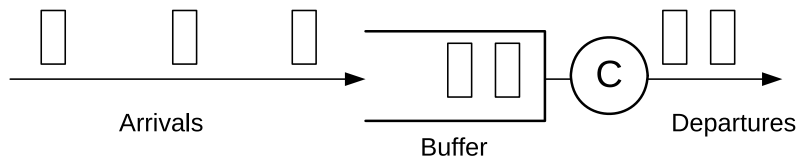
In the Past…

- Before there were packet networks there was the circuit-switched telephone network
- Traffic modeling of **telephone networks** was the basis for initial network models
  - – Assumed Poisson arrival process of new calls
  - – Assumed Poisson call duration

Source: Prof. P. Barford (edited)

… until early 1990's

- Traffic modeling of packet networks also used Poisson
  - Assumed Poisson arrival process for packets
  - Assumed Exponential distribution for traffic



Source: Prof. P. Barford (edited)

The measurement study that changed everything

- **Bellcore Traces**: In 1989, researchers at (Leland and Wilson) begin taking high resolution traffic traces at Bellcore
  - Ethernet traffic from a large research lab
  - 100 μ sec time stamps
  - Packet length, status, 60 bytes of data
  - Mostly IP traffic (a little NFS)
  - Four data sets over three year period
  - Over 100 million packets in traces
  - Traces considered representative of normal use

The data in part 3 of Lab 1 is a subset of the actual measurements.

Source: Prof. C. Williamson

Extract from abstract

Results were published in 1993

- **"On the Self-Similar Nature of Ethernet Traffic"**
  Will E. Leland, Walter Willinger, Daniel V. Wilson, Murad S. Taqqu

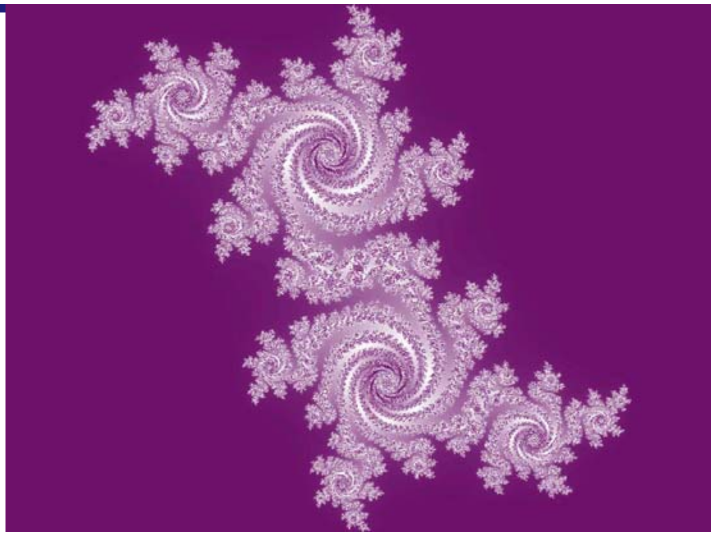
"*We demonstrate that Ethernet local area network (LAN) traffic is statistically self-similar, that none of the commonly used traffic models is able to capture this fractal behavior, that such behavior has serious implications for the design, control, and analysis of high-speed…"*

That Changed Everything…..

Source: Prof. V. Mishra, Columbia U. (edited)

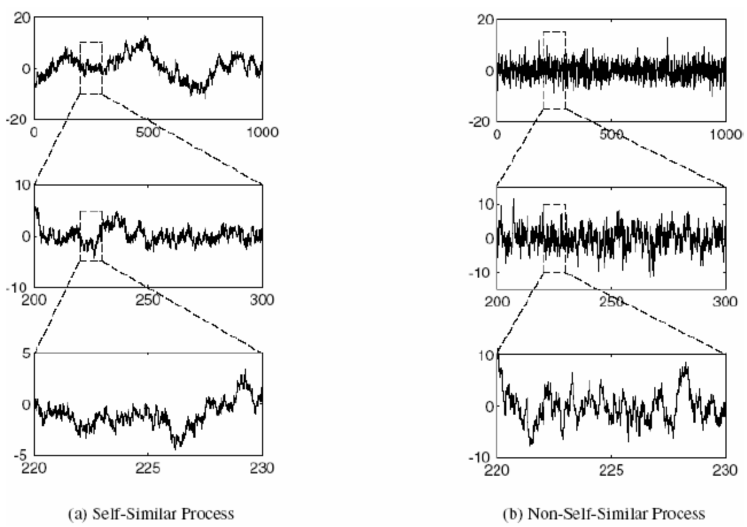
ECE 466

Fractals

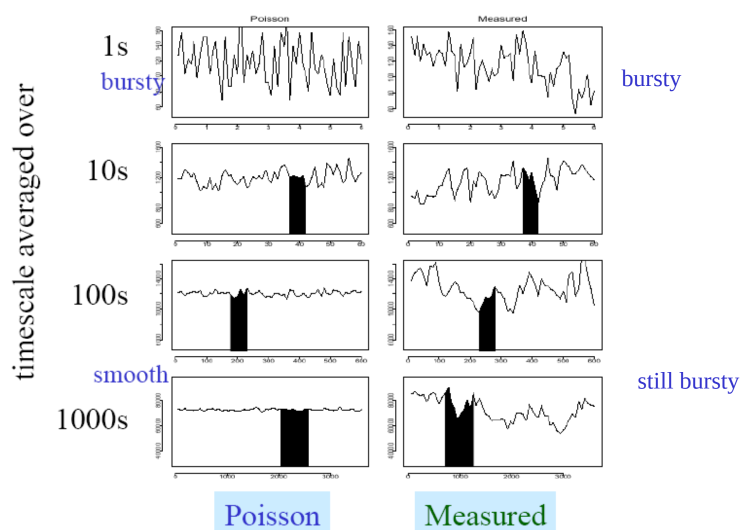


Source: Prof. P. Barford, U. Wisconsin

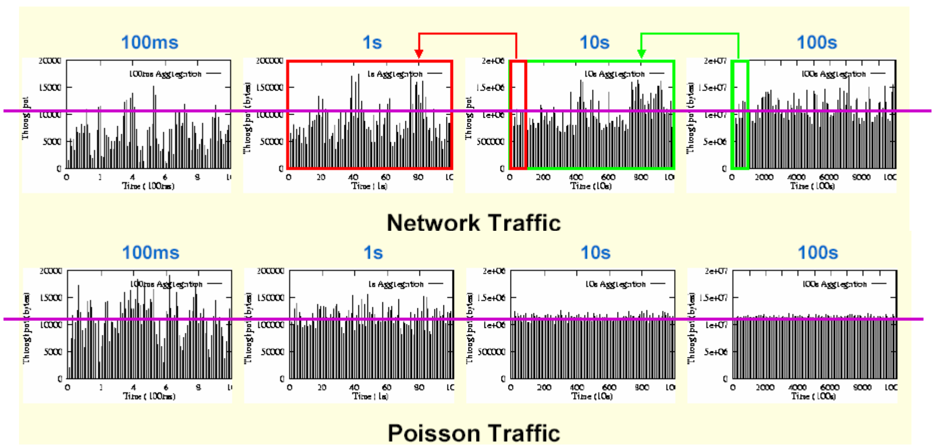
ECE 466



Traffic at different time scales (Bellcore traces)



Source: Prof. P. Barford (edited)



Source: Prof. V. Mishra, Columbia U.

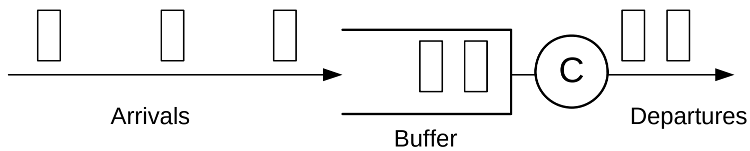
What is the observation?

- A Poisson process
  - When observed on a fine time scale will appear bursty
  - When aggregated on a coarse time scale will flatten (smooth) to white noise
- A Self-Similar (fractal) process
  - When aggregated over wide range of time scales will maintain its bursty characteristic

Source: Prof. C. Williamson

Why do we care?

- For traffic with the same average, the probability of a buffer overflow of self-similar traffic is much higher than with Poisson traffic
  - Costs of buffers (memory) are 1/3 the cost of a high-speed router !
- When aggregating traffic from multiple sources, self-similar traffic becomes burstier, while Poisson traffic becomes smoother
  –



Self-similarity

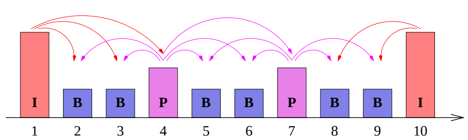
- The objective in Lab 1 is to <u>observe</u> self-similarity and obtain a sense.
- The challenge of Lab 1:
  - The Bellcore trace for Part 4 contains 1,000,000 packets
  - The computers in the lab are not happy with that many packets
  - Reducing the number of packets in plots, may reduce opportunities to discover self-similarity effect

Part 2: Video Traffic

- Compressed video traffic is very bursty, as a result of the compression algorithm
- Uncompressed digital video consists of a sequence of images. The frame rate is 24-30 frames/seconds
- Video compression algorithms (MPEG-1,2,4, H.263) exploit that subsequent frames are very similar
  - Only transmit the difference between subsequent frames

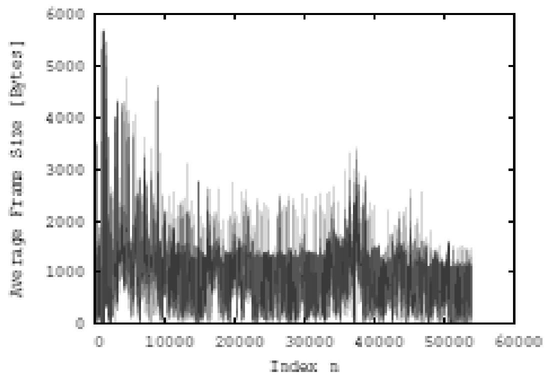
VBR Compression

- Most compression algorithms generate three types of frames:
  - **I Frames** (intra-coded frames):
    - Still images, just like JPEG
  - **P Frames:** (predictive frames)
    - encode the differences from the most recent I or P frame
  - **B Frames** (bi-directional frames):
    - interpolations of the next and previous I or P frames
- The frames are generated in a repeated pattern, called the Group-of-Picture (GOP) pattern, e.g., IBBPBBPBB
- The result is that compressed video is sending at different times at different rates. This is called Variable Bit Rate or VBR.

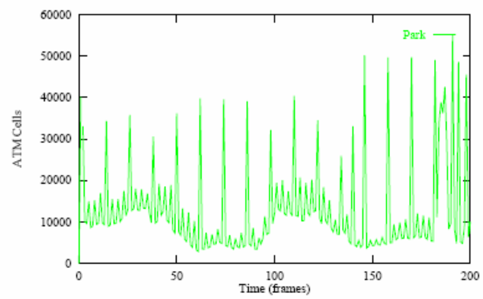


Video Traffic

- This plot shows the size of frames from a compressed version of "Silence of the Lambs" (MPEG-4 part 2, Q=28)
- Encoding done at Arizona State University

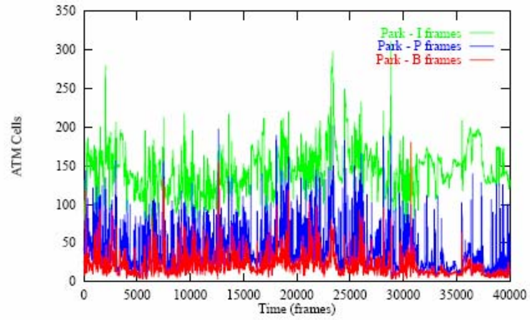


MPEG-1 Video Trace



- **30-minute trace from *Jurassic Park*, 24 fps**
- **Resolution: 384x288; YUV (4:1:1), 8-bit color**
- **Q-factors: I=10, P=14, B=18**
- **GOP Size: 12; GOP Pattern: IBBPBBPBBPBB**
- **Encoding done at University of Wurzburg (1995).**

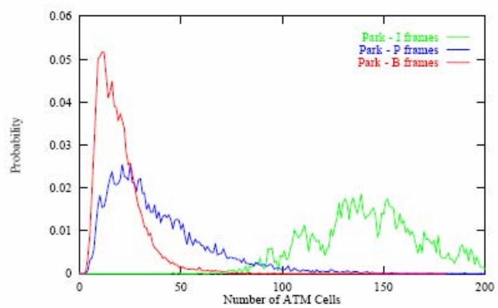
MPEG-1 Video Trace



- **Frame Types:**

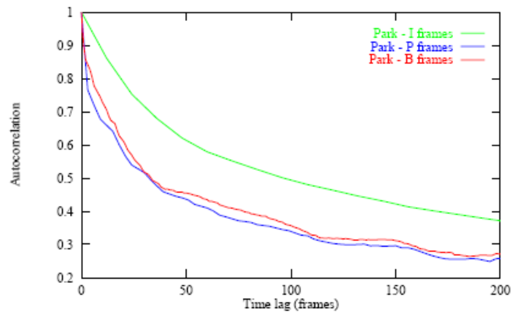
top - I-frames   middle - P-frames   bottom - B-frames

Marginal Distribution



- **Marginal distribution per frame type**
- **Has been matched with Log-Normal distribution and Weibull distribution.**

Autocorrelation Function



- Autocorrelation function broken down by frame type.
- Autocorrelation function

$$R(k) = \frac{Cov[X_n, X_{n+k}]}{Var[X_n])} = \frac{E[(X_n - \overline{X})(X_{n+k} - \overline{X})]}{Var[X_n]}$$

Making the concept of "Self-similarity" precise

- Self-similarity is a mathematically well-defined property
  - We do not cover this
- Self-similar traffic has certain statistical properties
  - Slow decay of autocorrelation function
    - → long-range dependent
  - "Hurst parameter" and "index of dispersion" has values in characteristic ranges

This is done in the optional Part 4 of the lab

Long Range Dependence

- Autocorrelation coefficient can range between:
  - +1 (very high positive correlation)
  - -1 (very high negative correlation)
- Zero means no correlation
- Autocorrelation function shows the value of the autocorrelation coefficient for different time lags k

o **A long-range dependent process has an autocorrelation function** *R(k)* that decreases hyperbolically (not exponentially)

$\Sigma_k R(k) = \infty$ *(long range dependence)*

Autocorrelation Function