Per-hop and end-to-end capacity estimation

Fall 2003
CS 8803: Network Measurement Seminar

Ravi S. Prasad

Outline

- Introduction
- Definitions
- Capacity estimation methodologies
  - Hop capacity
    - Variable packet size (VPS)
      - Effect of layer 2 switches
    - Delay variation methods
  - Path capacity
    - Packet pair dispersion
      - Effect of cross-traffic
    - Packet train dispersion

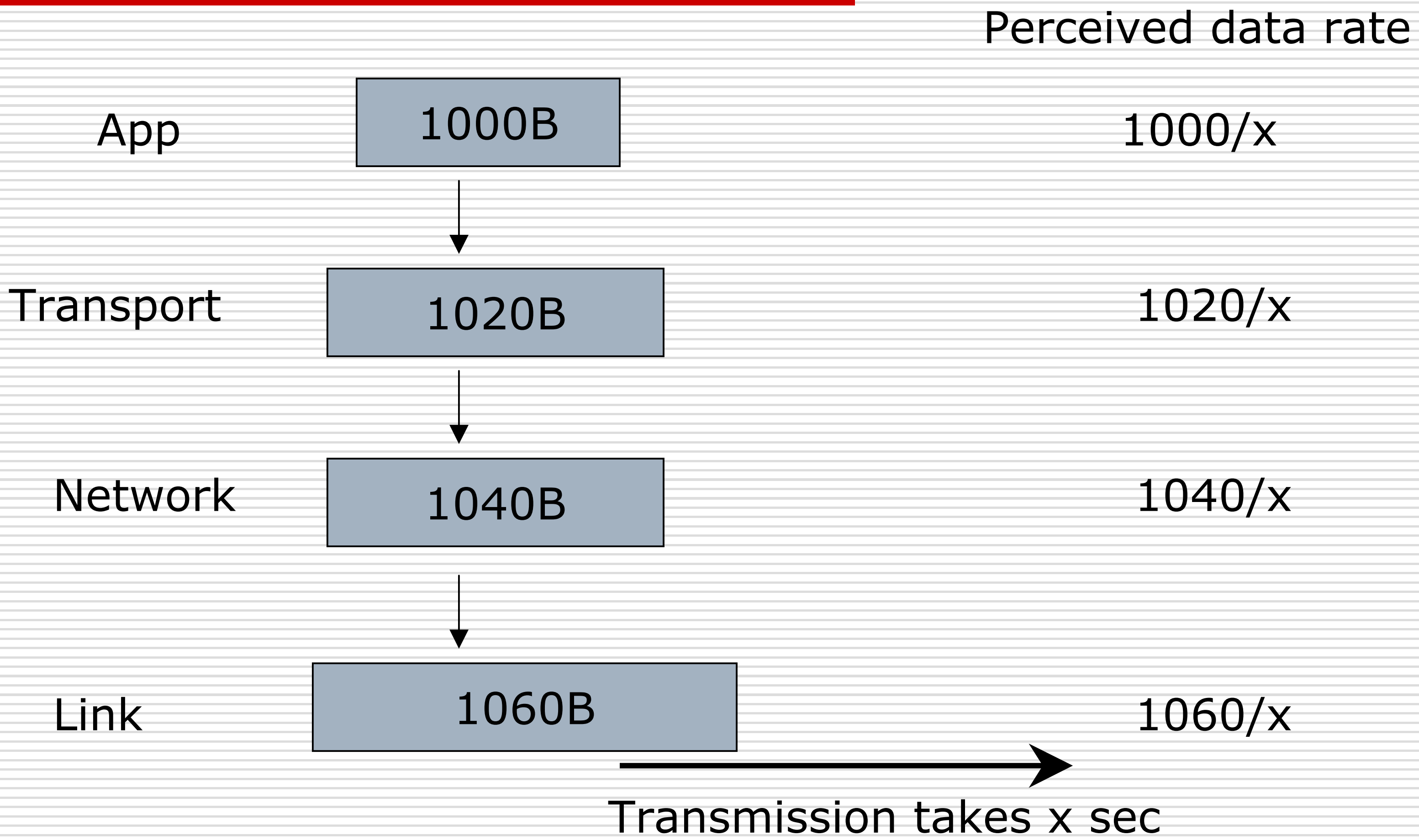
Network Bandwidth

- Bandwidth: Metric for data transfer rate
- Important for many applications
  - File transfer
  - Multimedia streaming
  - Overlay network routing
  - Service level agreement (SLA) and its verification

Bandwidth & layering

- Link layer transmits at fixed data rate
- All higher layers add overhead (header and/or trailer)
- Different layers observe different data rate
- For IP network
  - Bandwidth = IP layer data rate

Example


Perceived data rate
App
1000B
1000/x
Transport
1020B
1020/x
Network
1040B
1040/x
Link
1060B
1060/x
Transmission takes x sec

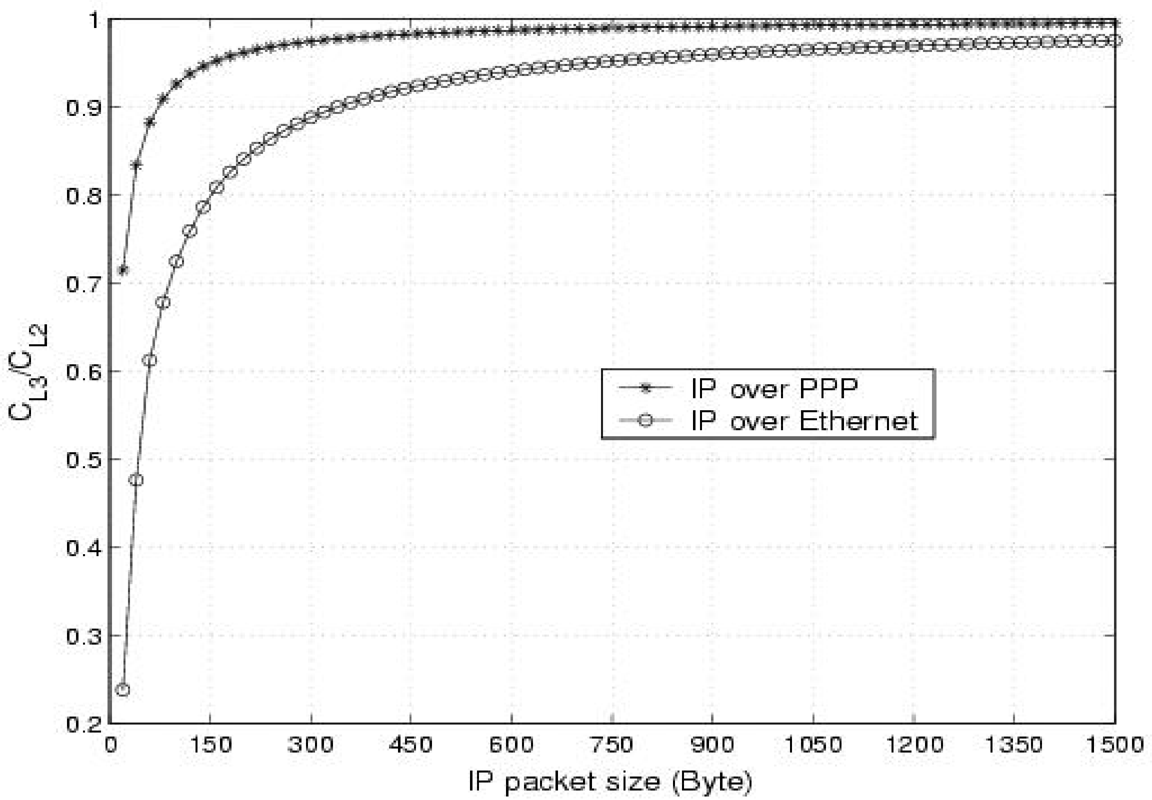
Capacity

- Maximum IP layer data rate
- Packet size dependence due to Layer2 encapsulation

$$C_{L3} = C_{L2}\left(\frac{L_{L3}}{L_{L3} + H_{L2}}\right)$$

  - H_{L2} = Layer 2 header
  - L_{L3} = Layer 3 (IP) data size
- Capacity = data rate for MTU packets

Example



- Ethernet overhead 38B
- PPP overhead 8B(default)

Hop and path capacity

- Hop capacity
  - Maximum possible bandwidth
- Path capacity
  - Minimum of hop capacities
  - Narrow link: hop with minimum capacity

$$C^{path} = \min_{i=1,\ldots,H} \left(C_i^{hop} \right)$$

Estimation Methodologies

Capacity estimation

- Per-hop capacity
  - Variable Packet Size (VPS)
    - Pathchar, clink, pchar, tailgater
  - Delay Variations
    - Packet quartet methods
- End-to-end capacity
  - Packet Pair Dispersion
    - Pathrate, Sprobe, Bprobe
  - Packet Train Dispersion

Per Hop Capacity Estimation

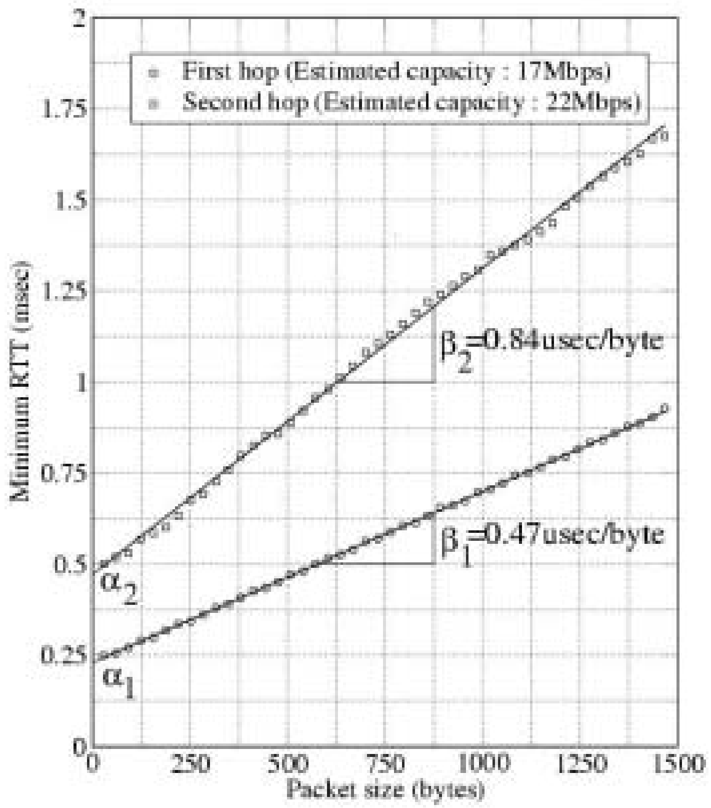
1. VPS methodology

- Obtain RTT up to I^{th} hop for different size (L) packets (using ICMP response or tailgating)
- Components of RTT
  - Propagation delay (constant)
  - Queuing delay
  - *Serialization delay* (L/C)
- Assume:
  - Minimum RTT (T_I) for a packet size didn't see queuing.

$$T_I(L) = \alpha + \beta_I L$$

$$\beta_I = \sum_{i=1}^{I} \frac{1}{C_i}$$

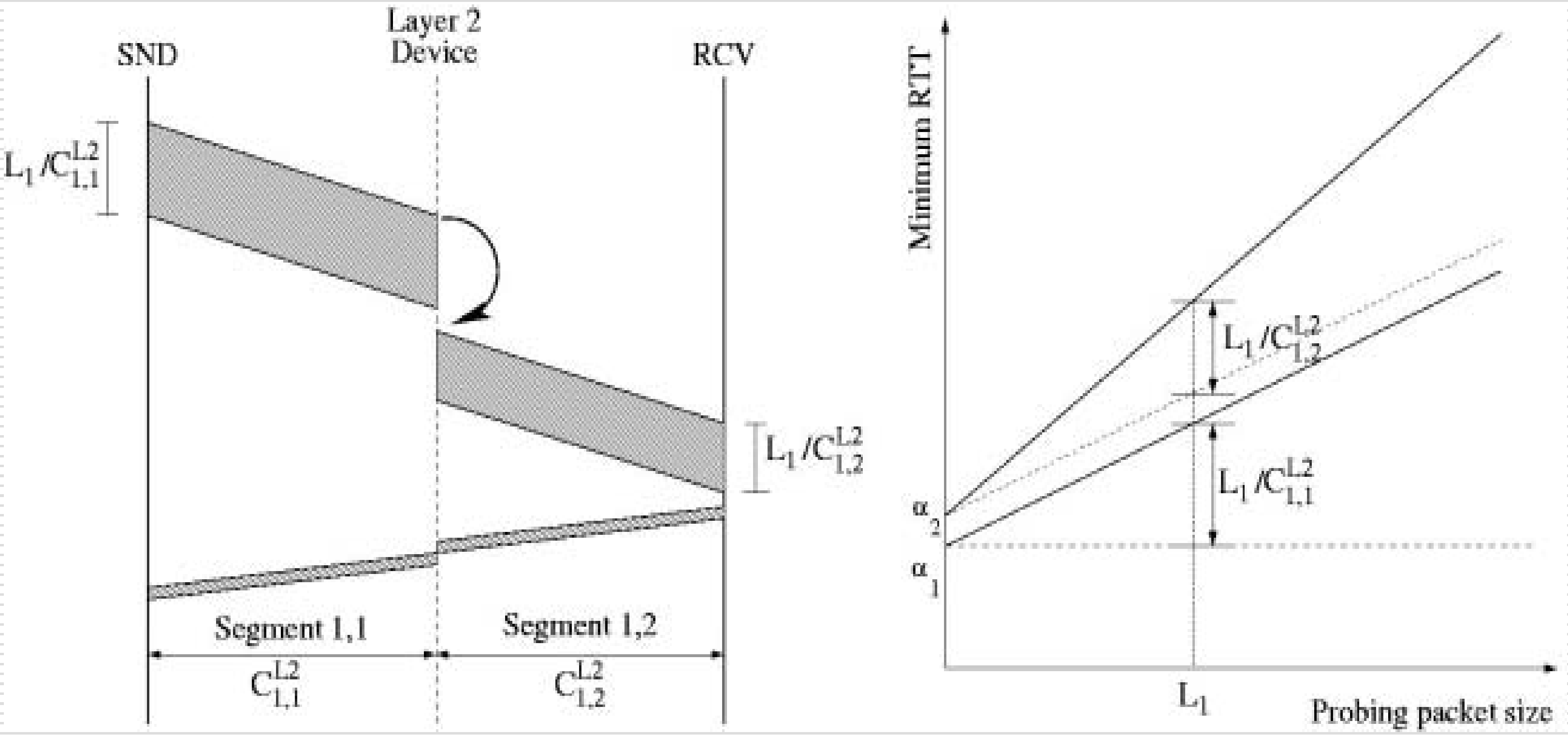
Example



$$C_1 = \frac{1}{\beta_1}$$

$$C_2 = \frac{1}{\beta_2 - \beta_1}$$

Effect of layer2 switches



- Increase serialization delay

Effect of layer2 switches (cond.)

- Increase serialization delay

$$\beta = \frac{1}{C_1^{L2}} + \frac{1}{C_2^{L2}}$$

- Capacity underestimation

$$\frac{1}{\beta} < \frac{1}{C = \min(C_1^{L2}, C_2^{L2})}$$

- Can't be detected with TTL expiration

Does this error propagate?

- If no layer2 switches in I^{th} hop

$$\beta_I = \frac{1}{C_I^{L3}} + \beta_{I-1}$$

  - Path up to $(I-1)^{th}$ hop may have switches

- Estimated capacity for I^{th} hop

$$\frac{1}{\beta_I - \beta_{I-1}} = C_I^{L3}$$

  - Localized error

Other sources of error for VPS

- Non-zero queuing delays
- Limited clock resolution
- Error propagation/amplification
- ICMP generation latencies

2. Delay variation methods

- Pasztor and Veitch (IWQoS `02)
- Use difference of one-way delay of consecutive probes
  - Instead of VPS like regression
  - More than one useful probe per size
- Delay variation after *h* hops

$$\delta_i^h = \sum_{j=1}^{h} \left(\frac{L_i}{C_j} - \frac{L_{i-1}}{C_j} \right) + \sum_{j=1}^{h} (q_i^j - q_{i-1}^j)$$

• q_i^j Queuing delay for i[th] packet at j[th] hop

Example

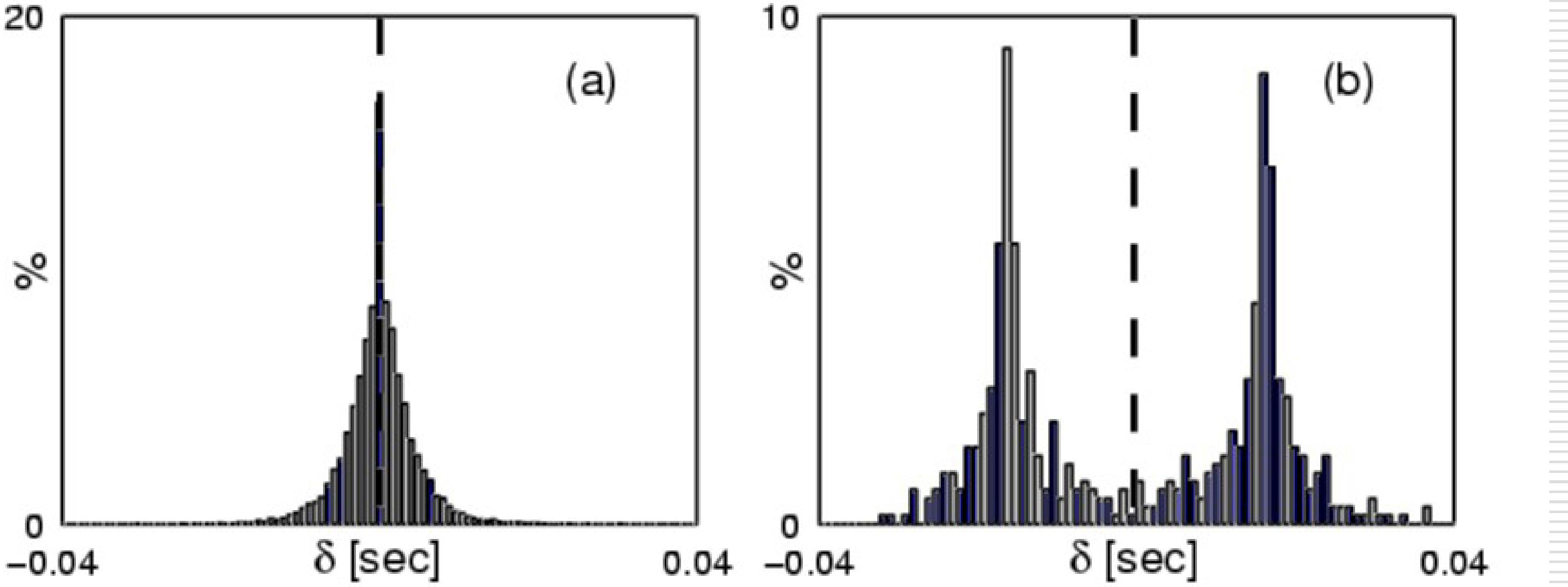
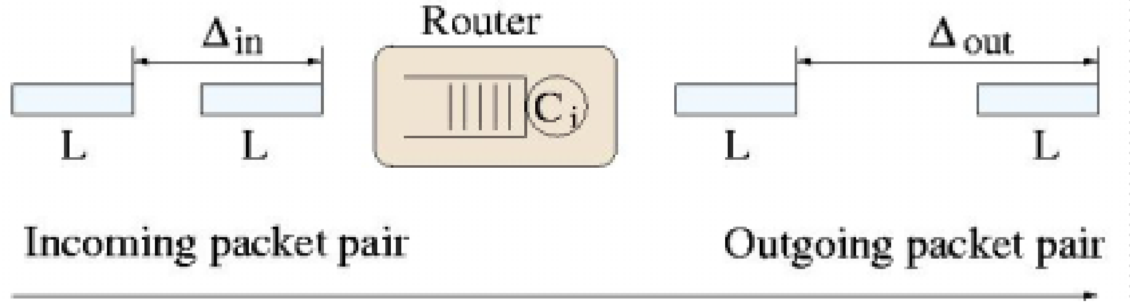


Fig. 2. Accumulation signature in delay variation histograms. (a) 1 hop simulation. Constant probe size: the service time cancels, the noise is symmetric, (b) 12 hop measurement. Alternating probe size: the peak splits and the signature emerges: a symmetric bimodal histogram.

End-to-End Capacity Estimation

Packet pair dispersion

- Relate dispersion of two back-to-back packets to path capacity



- Empty link $$\Delta_{out} = \max\left(\Delta_{in}, \frac{L}{C_i}\right)$$

- Empty path $$\Delta_R = \max_{i=1,\ldots,H}\left(\frac{L}{C_i}\right) = \frac{L}{C}$$

Effects of probe size

- Small probes
  - Incorrect estimate
    - Higher layer-2 header
  - Small dispersion value
    - Require high resolution clock
- Large probes
  - Higher probability of cross-traffic interference

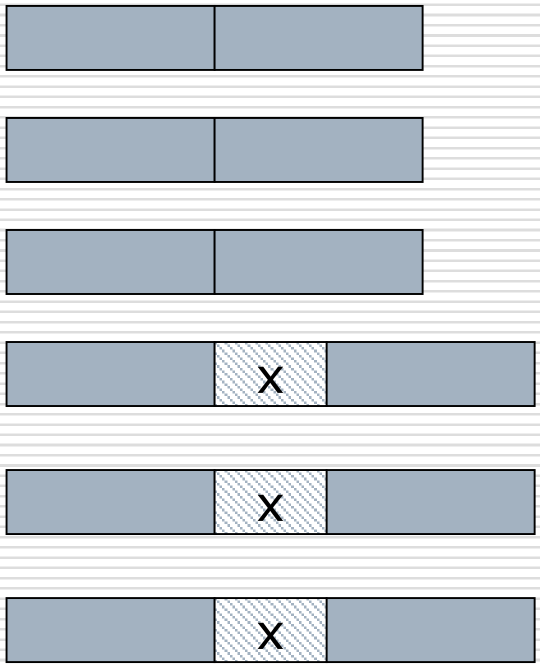
Cross-traffic effects

- Before the narrow link
  - Can increase dispersion
- After the narrow link
  - Can increase or decrease dispersion

- Large error possible in any estimate
- Correct estimation requires statistical analysis

Example of statistical analysis

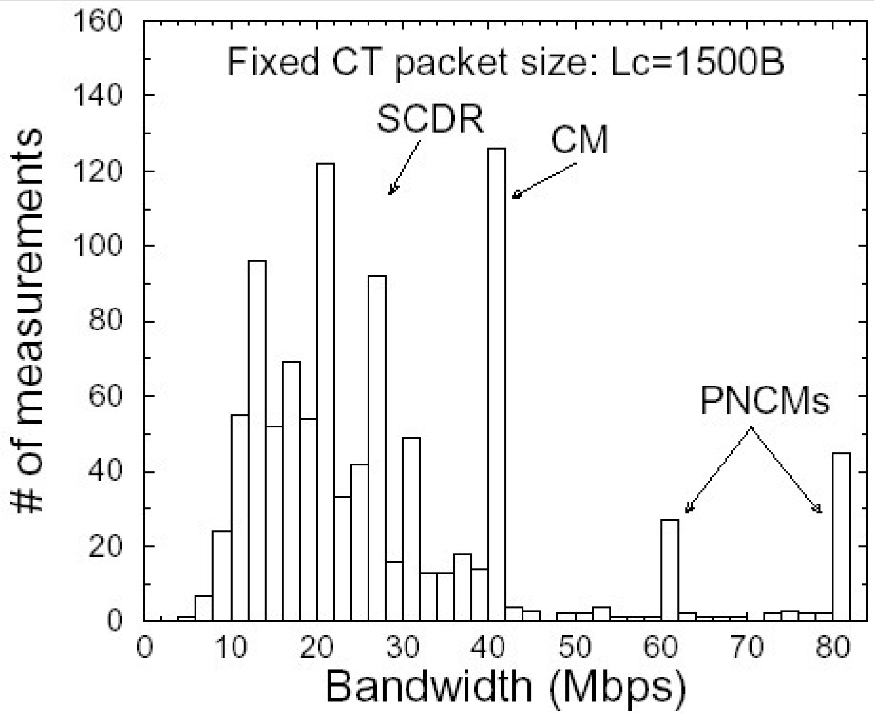
- Dispersion with highest frequency gives capacity
  - Fails when probe size and cross-traffic size constant
  - Internet traffic is mostly 40, 512,1500B packets
  - Have different size packet pairs

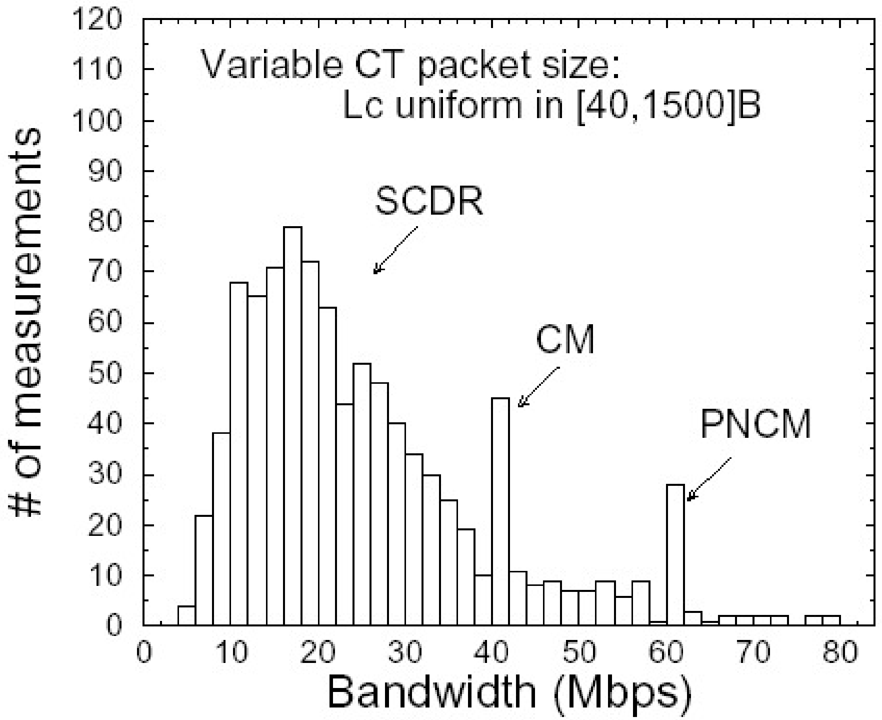


Characteristics of packet pair estimates

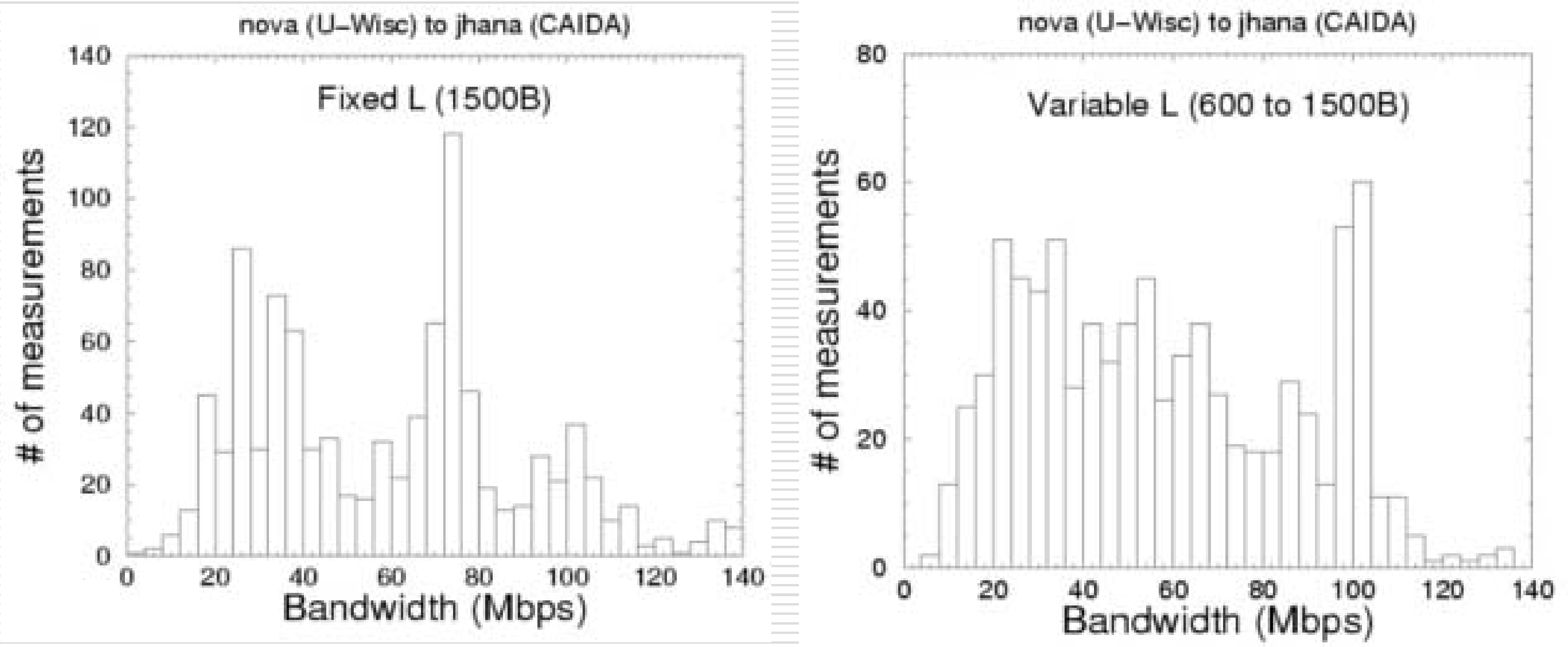
- Multi-modal estimates
- Local modes
  - Peaks in pdf of bandwidth estimates
  - Formed by repetitive measurement
  - Potential candidates for capacity
- Three zones of local modes
  - Capacity mode(CM)
    - Captures path capacity
  - Post Narrow Capacity Mode (PNCM)
    - Captures capacity after narrow link
  - Sub-Capacity Dispersion Range(SCDR)
    - Increased dispersion due to cross-traffic

Effect of cross-traffic size variation





Effect of probe size variation



Packet train dispersion

- Extension of packet pair
- Send N packets back-to-back
- Dispersion rate D(N)

$$D(N) = \frac{L(N-1)}{\Delta_N}$$

- In absence of cross-traffic
  - D(N) = C (Capacity)

Packet train dispersion

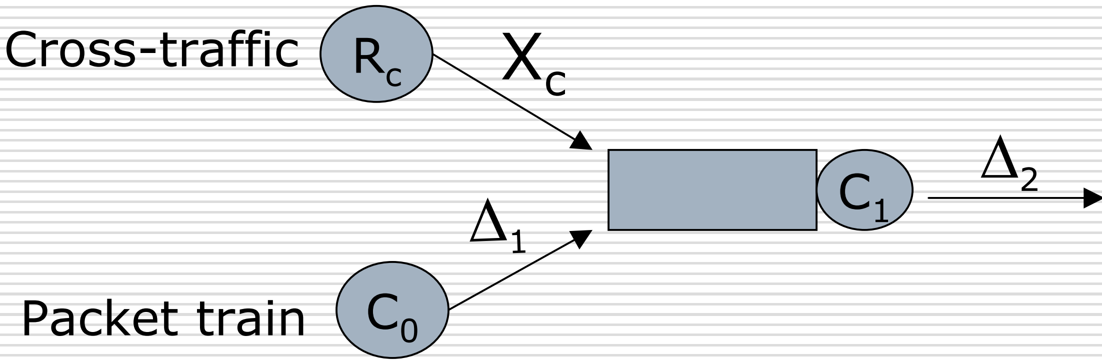
- cprobe assumes
  - D(N)=Available bandwidth (A)
- Dovrolis et. al. (Infocom 2001)
  - $D(N) \neq A$
  - D(N)=Average Dispersion Rate (ADR)
    - If N is large enough
  - ADR = f(hop capacities, hop utilizations)
  - $C \geq ADR \geq A$

Example

$$\Delta_1 = \frac{(N-1)L}{C_0}$$

$$X_c = R_c \Delta_1$$

$$\Delta_2 = \frac{X_c + (N-1)L}{C_1}$$



$$E(D) = \frac{(N-1)L}{E(\Delta_2)} = \frac{C_1}{\left(1 + \frac{R_c}{C_0}\right)}$$

- For large N, Δ_1 is large
  - Variation in X_c is small
  - D(N) = E(D)

$$E(D) \neq A = C_1 - R_c$$

End-to-end capacity estimation tool

- Pathrate
  - 1000 packet pair estimation with different size probe (600B-1500B)
    - Obtain multi-modal estimate distribution
  - 500 packet train estimation with maximum size probes
    - Obtain ADR
  - Use ADR as lower bound on possible capacity modes
    - Capacity mode: strongest local mode higher than ADR

Thank You