CSC2231: Availability in P2Ps

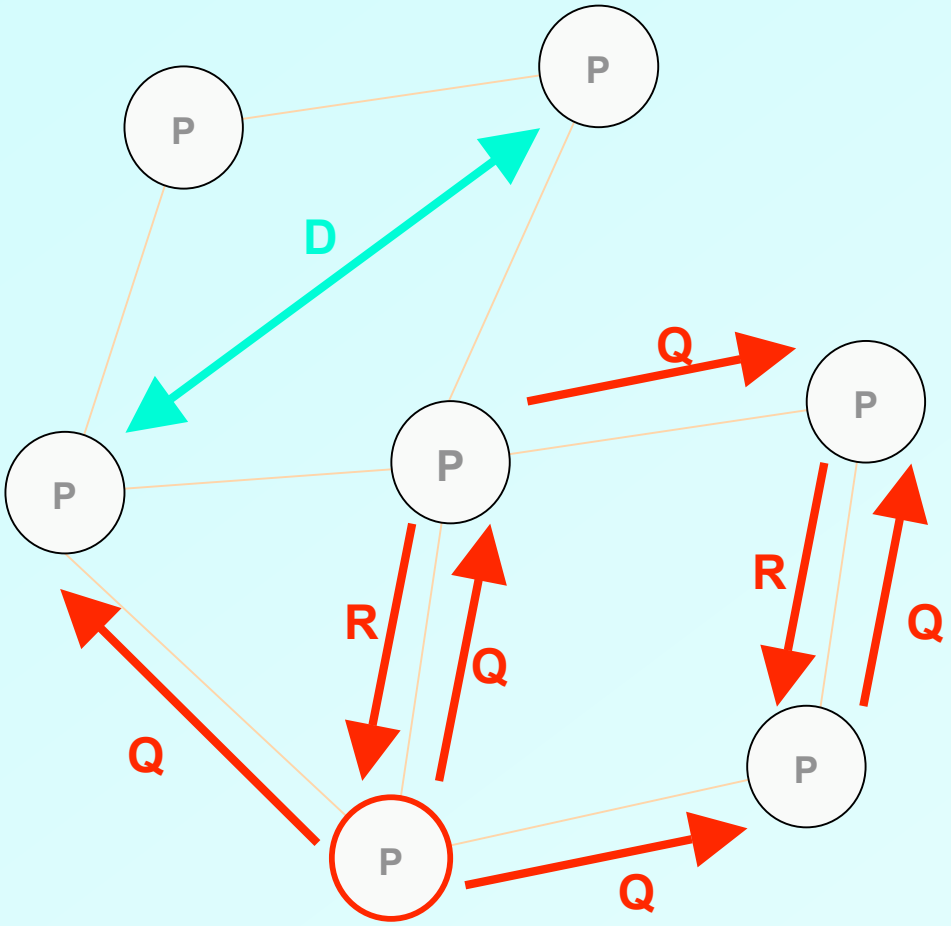
http://www.cs.toronto.edu/~stefan/courses/csc2231/05au

Stefan Saroiu

Department of Computer Science

University of Toronto

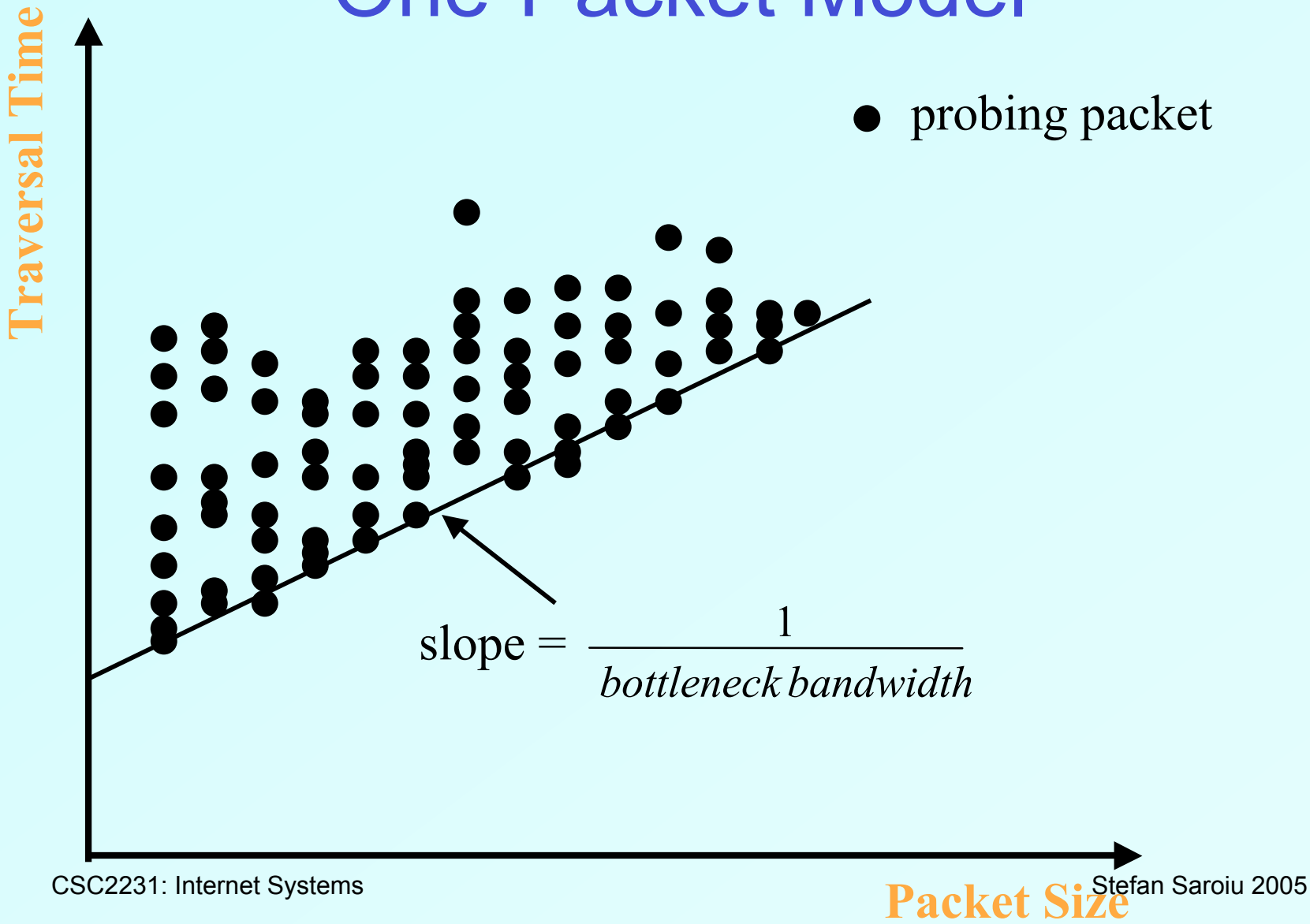
Gnutella ver. 0.4



Methodology

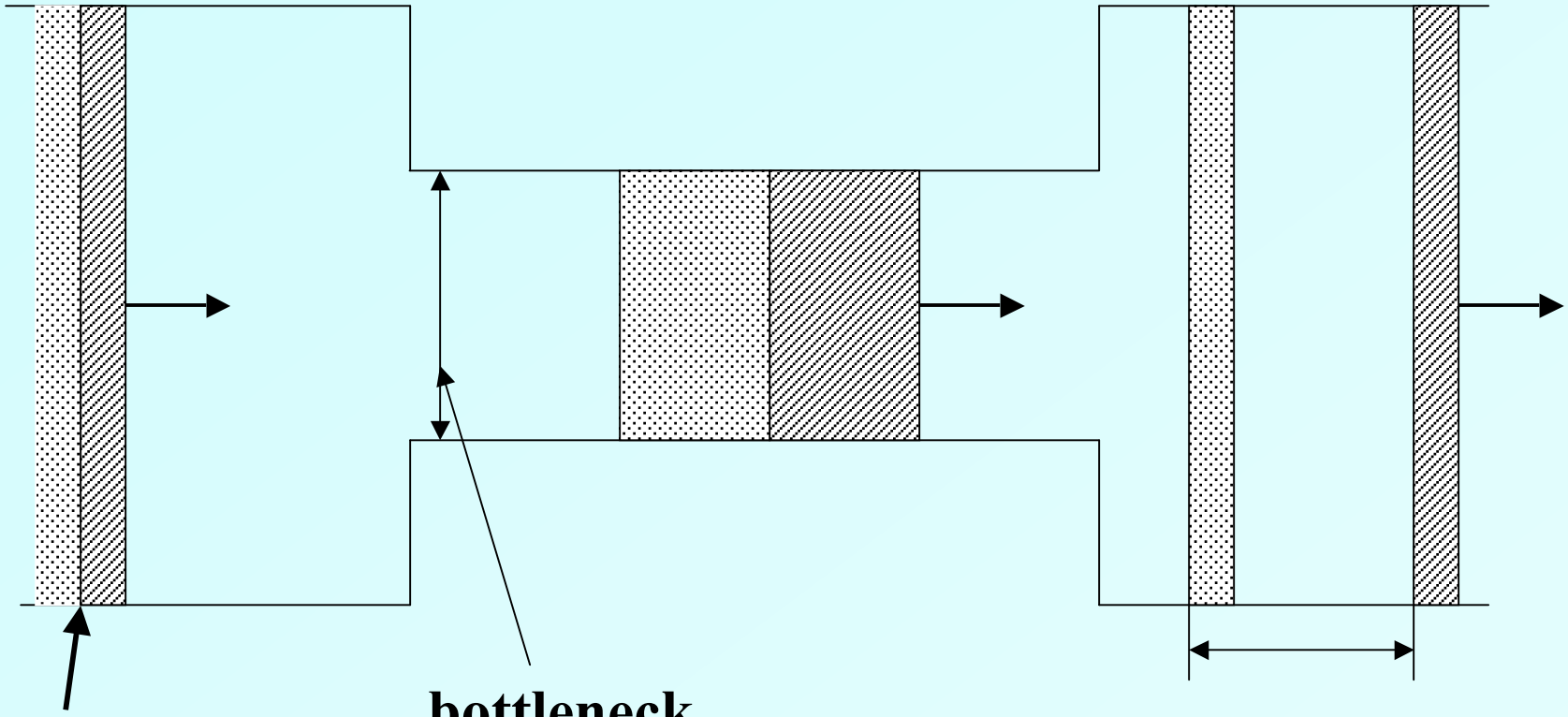
- **Two stages:**
  1. Periodically crawl Gnutella
     - Obtain periodical snapshots of the network
     - Discover peers and their metadata
  2. Probe discovered peers
     - Measure bottleneck bandwidth [SProbe]
     - Measure peer availability

One-Packet Model



Packet-Pair Model

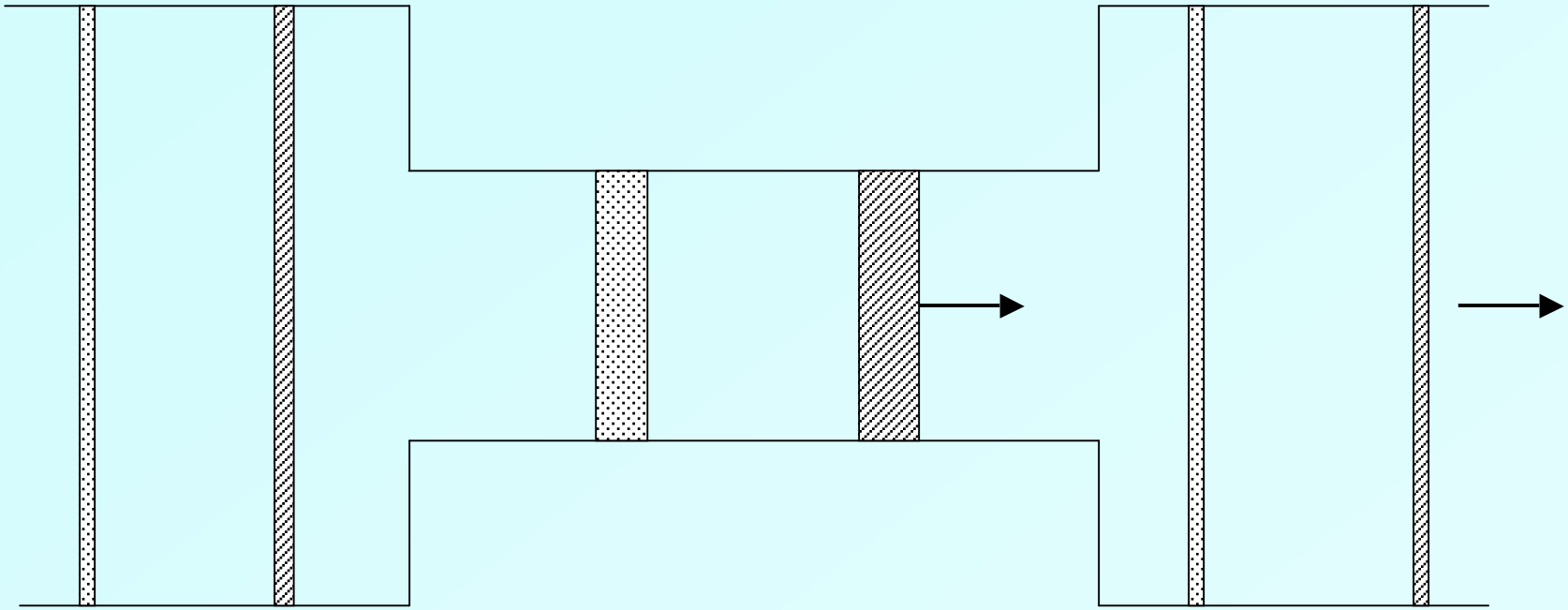
$$\text{bottleneck bandwidth} = \frac{\text{packet size}}{\Delta t}$$



bottleneck bandwidth

time dispersion proportional to bottleneck bandwidth

Packet-Pair Model



SProbe Uses TCP

- **From local host To remote host**
  - No cooperation needed

SProbe Uses TCP

- **From local host To remote host**
  - No cooperation needed

SProbe Uses TCP

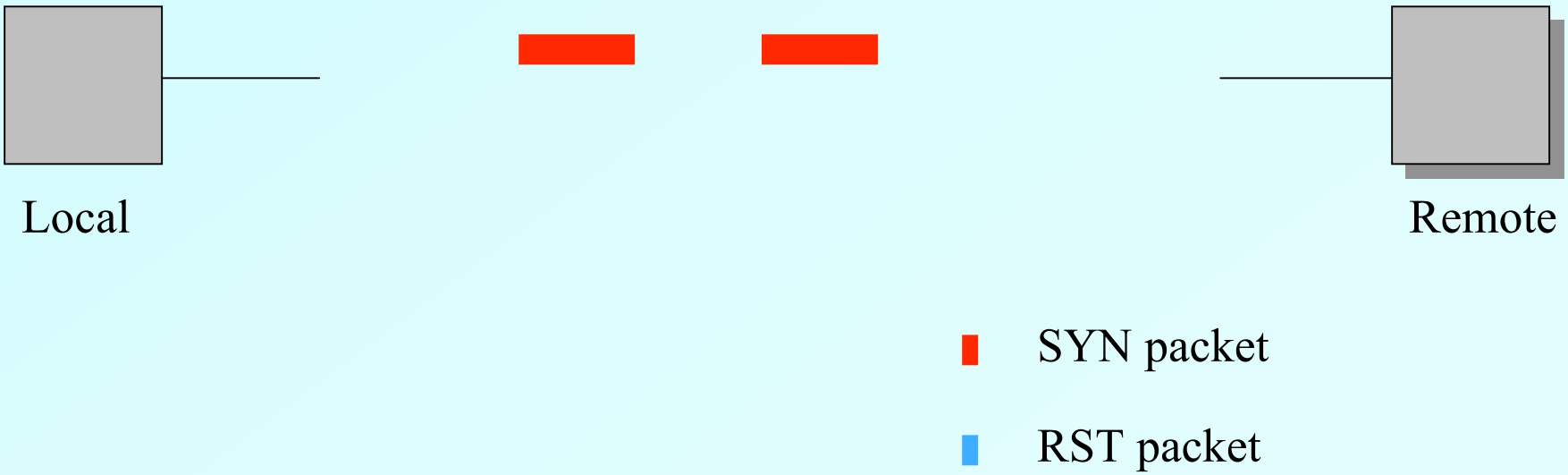
- **From local host To remote host**
  - No cooperation needed

SProbe Uses TCP

- **From local host To remote host**
  - No cooperation needed

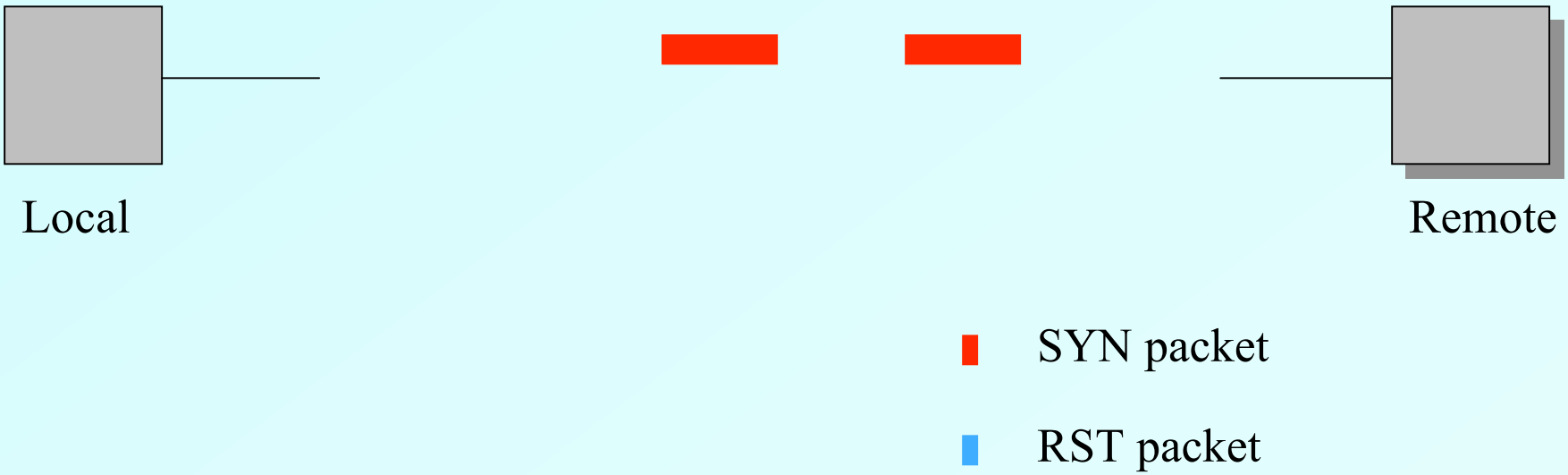
SProbe Uses TCP

- **From local host To remote host**
  - No cooperation needed



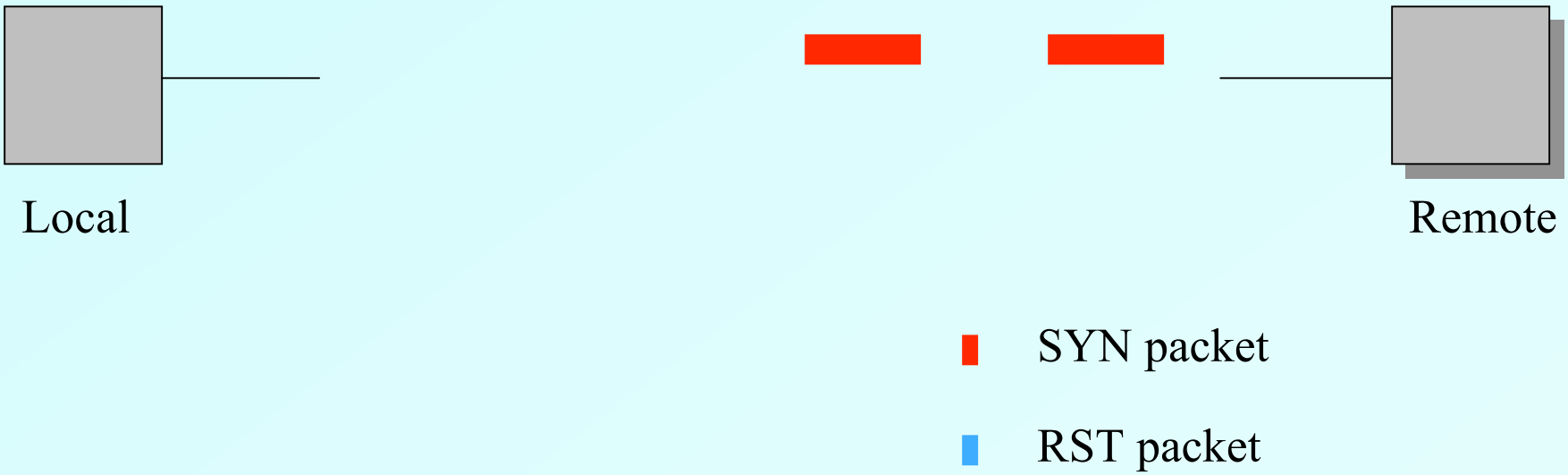
SProbe Uses TCP

- **From local host To remote host**
  - No cooperation needed



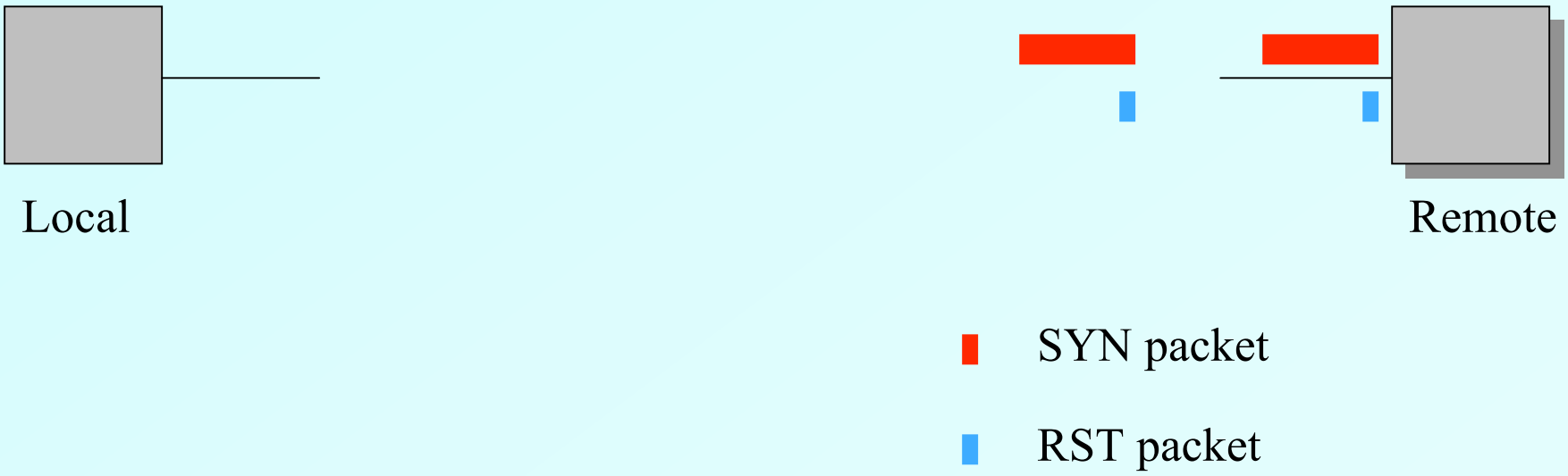
SProbe Uses TCP

- **From local host To remote host**
  - No cooperation needed



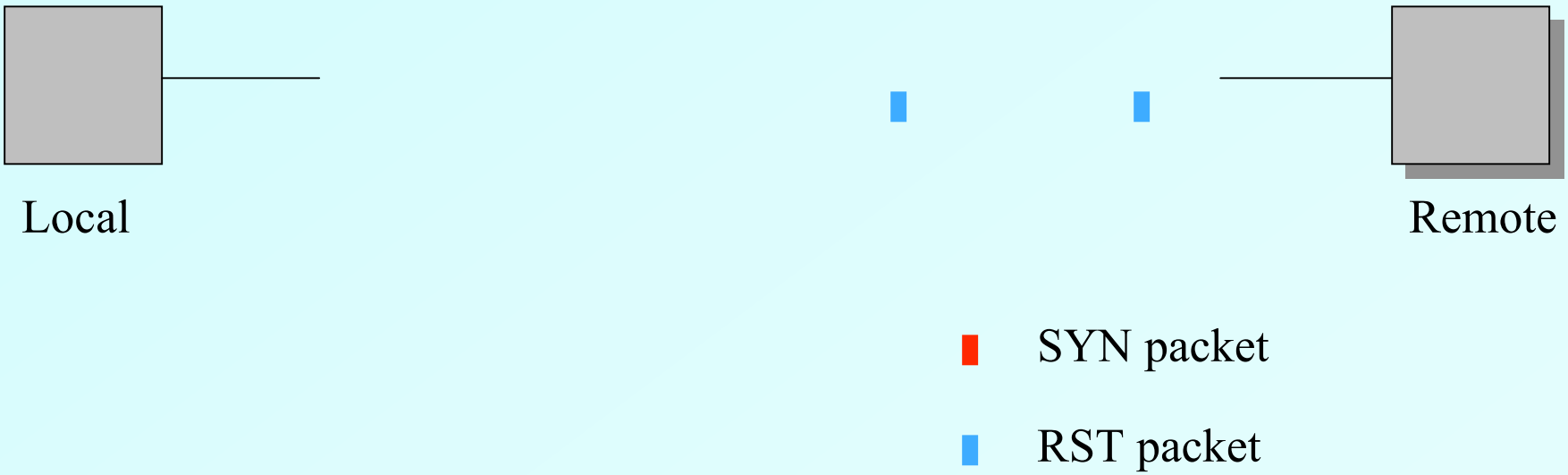
SProbe Uses TCP

- **From local host To remote host**
  - No cooperation needed



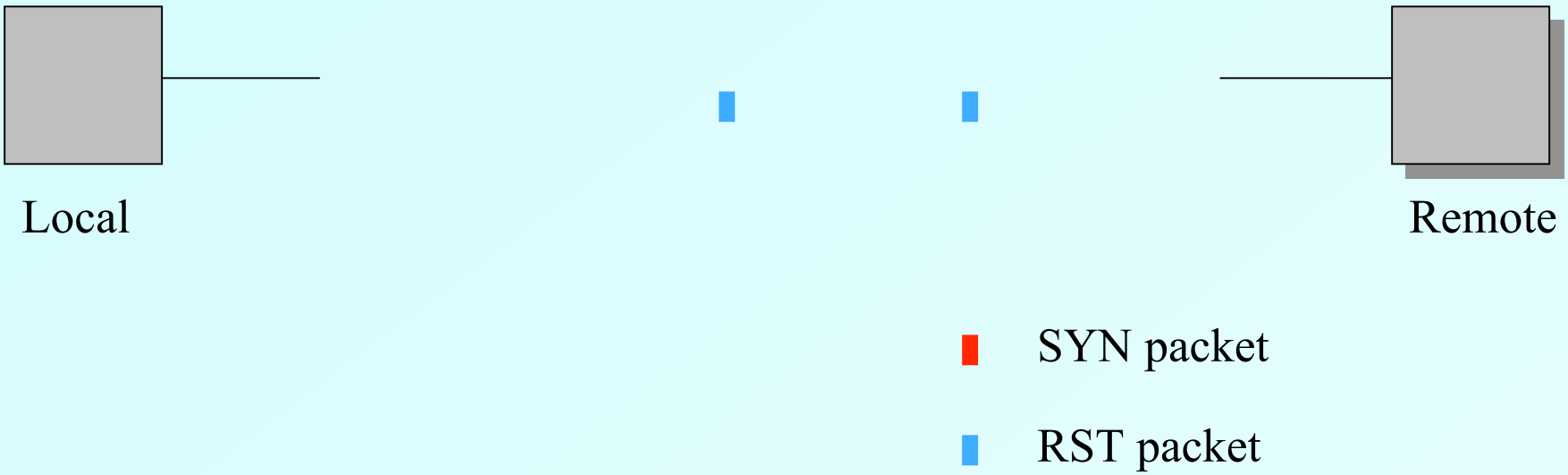
SProbe Uses TCP

- **From local host To remote host**
  - No cooperation needed



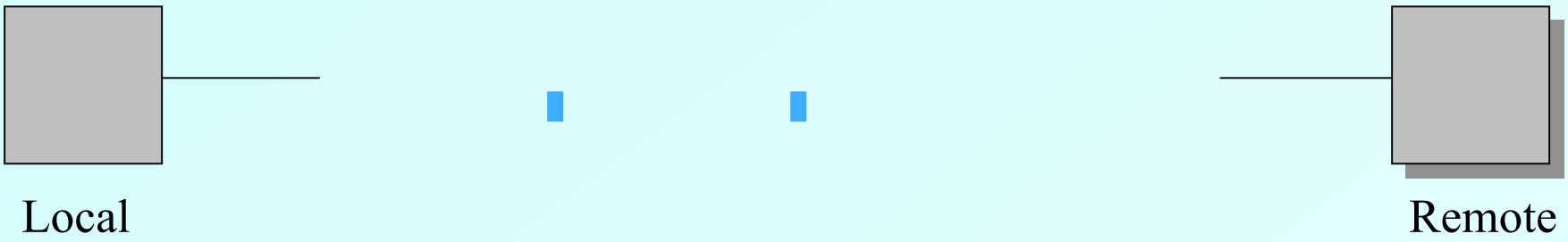
SProbe Uses TCP

- **From local host To remote host**
  - No cooperation needed



SProbe Uses TCP

- **From local host To remote host**
  - No cooperation needed

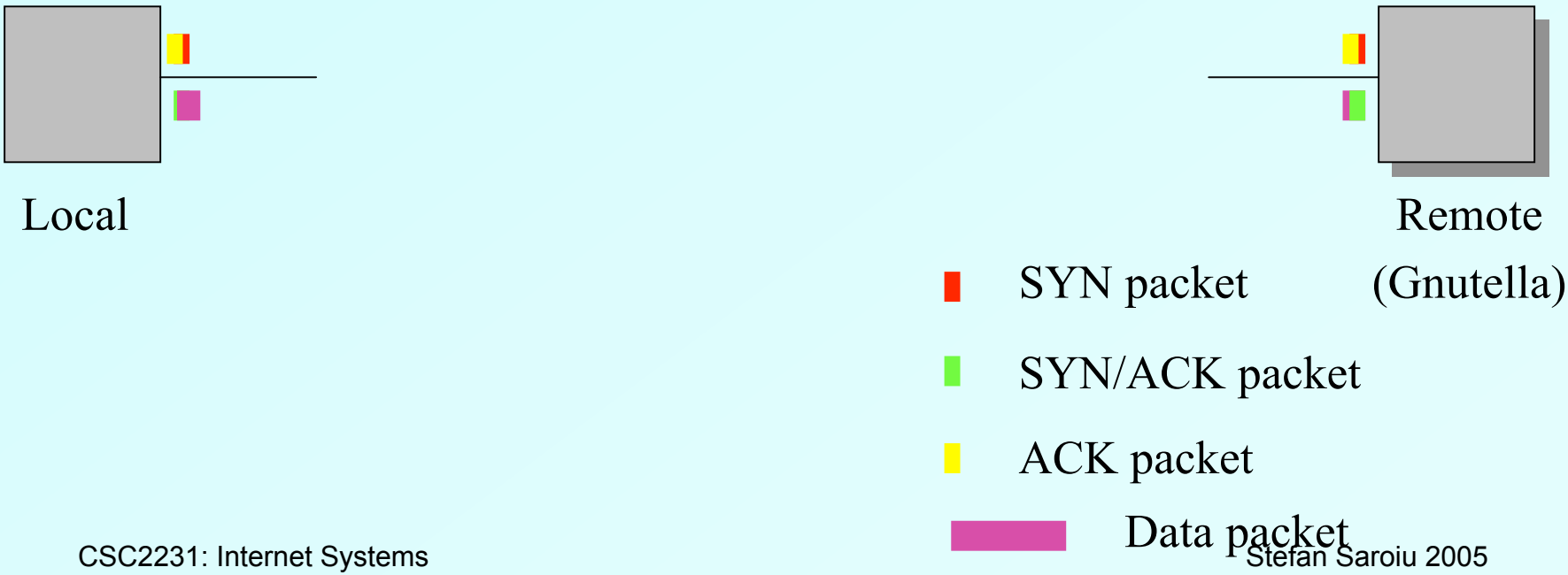


SProbe Uses TCP

- **From local host To remote host**
  - No cooperation needed

SProbe Uses TCP

- **From remote To local**
  - Involuntary cooperation of application layer



SProbe Uses TCP

- **From remote To local**
  - Involuntary cooperation of application layer

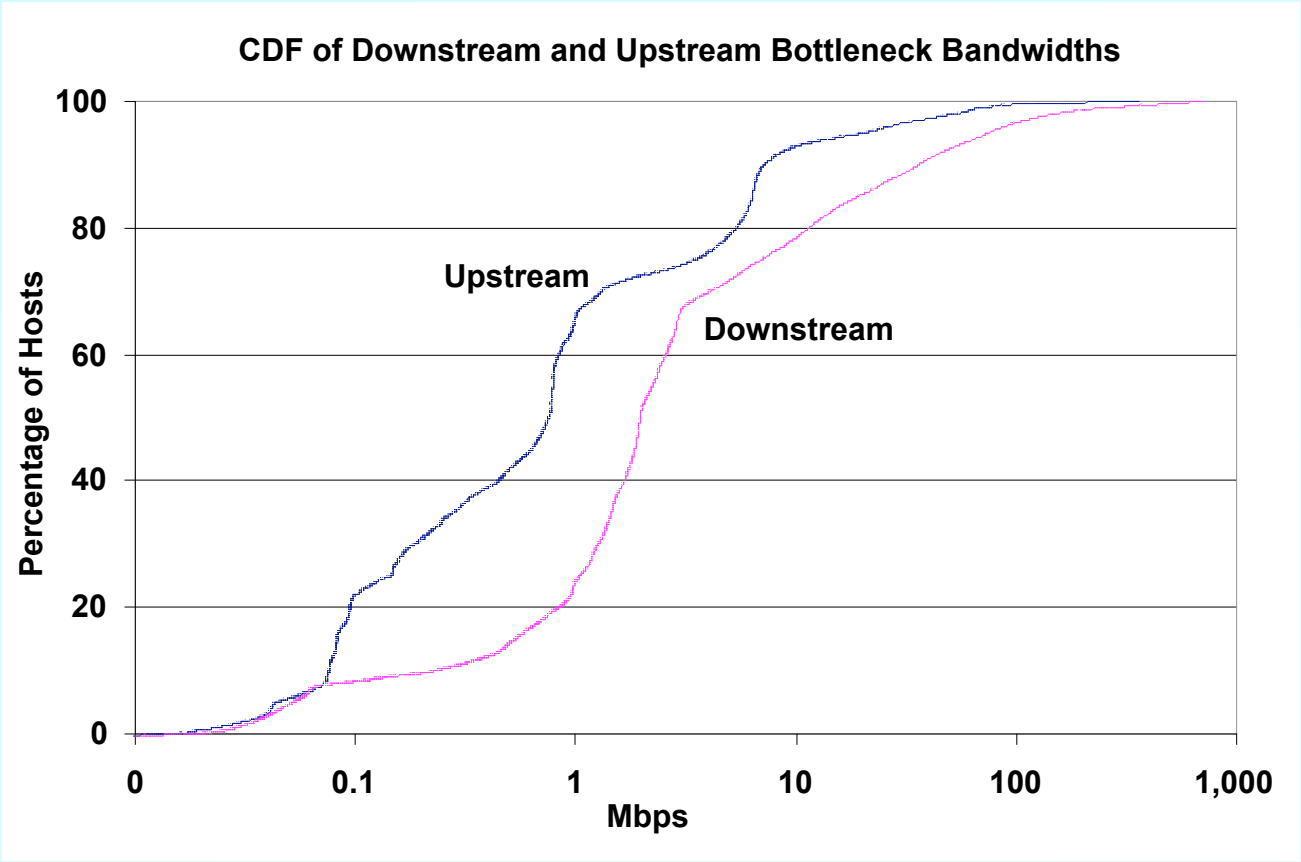
Peer Characteristics

- **P2P systems premises:**
  - Non-greedy behavior, voluntary cooperation
  - Single & uniform roles, no client/server demarcation

- **Question:**
  - Is this true in practice?

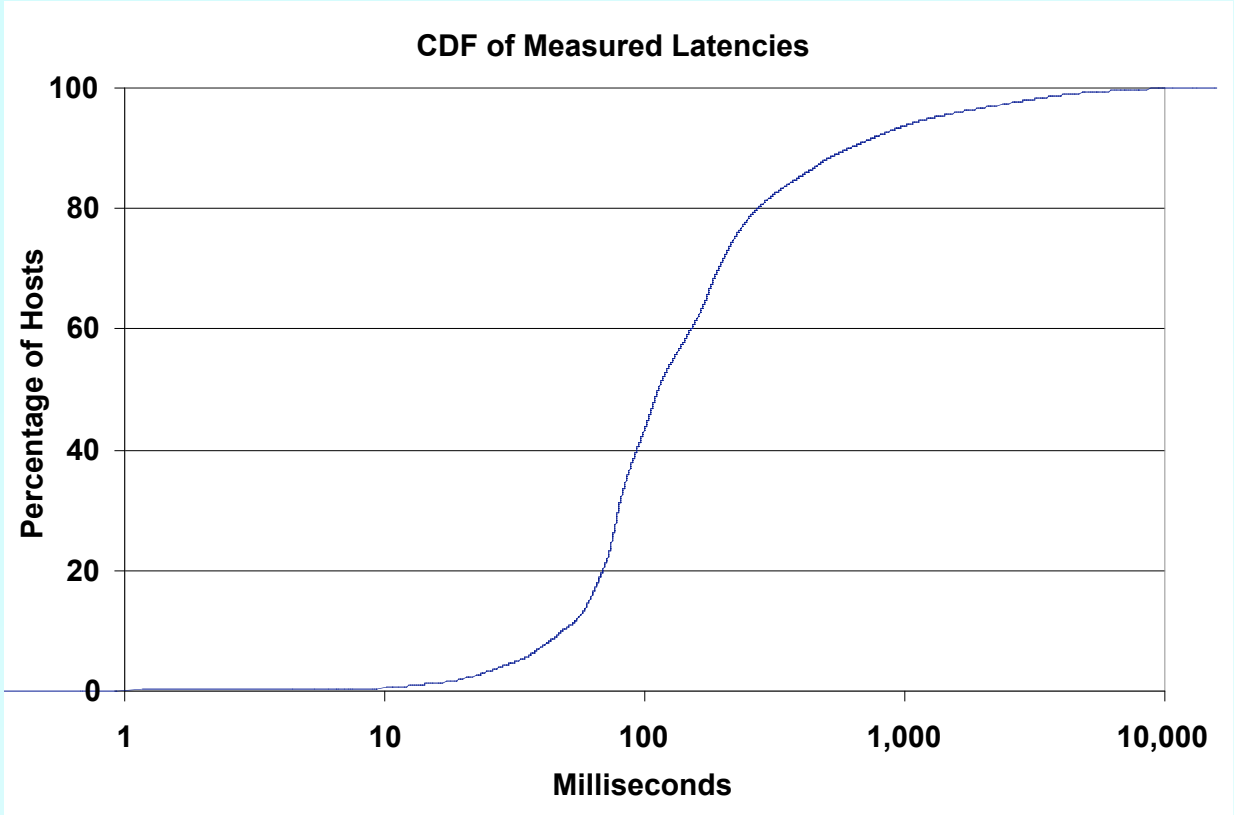
1. **In practice, how uniform are peers ?**
   Or… are some server-like, are some client-like?
2. **In practice, how well behaved are peers?**

Higher Downstream Bandwidths



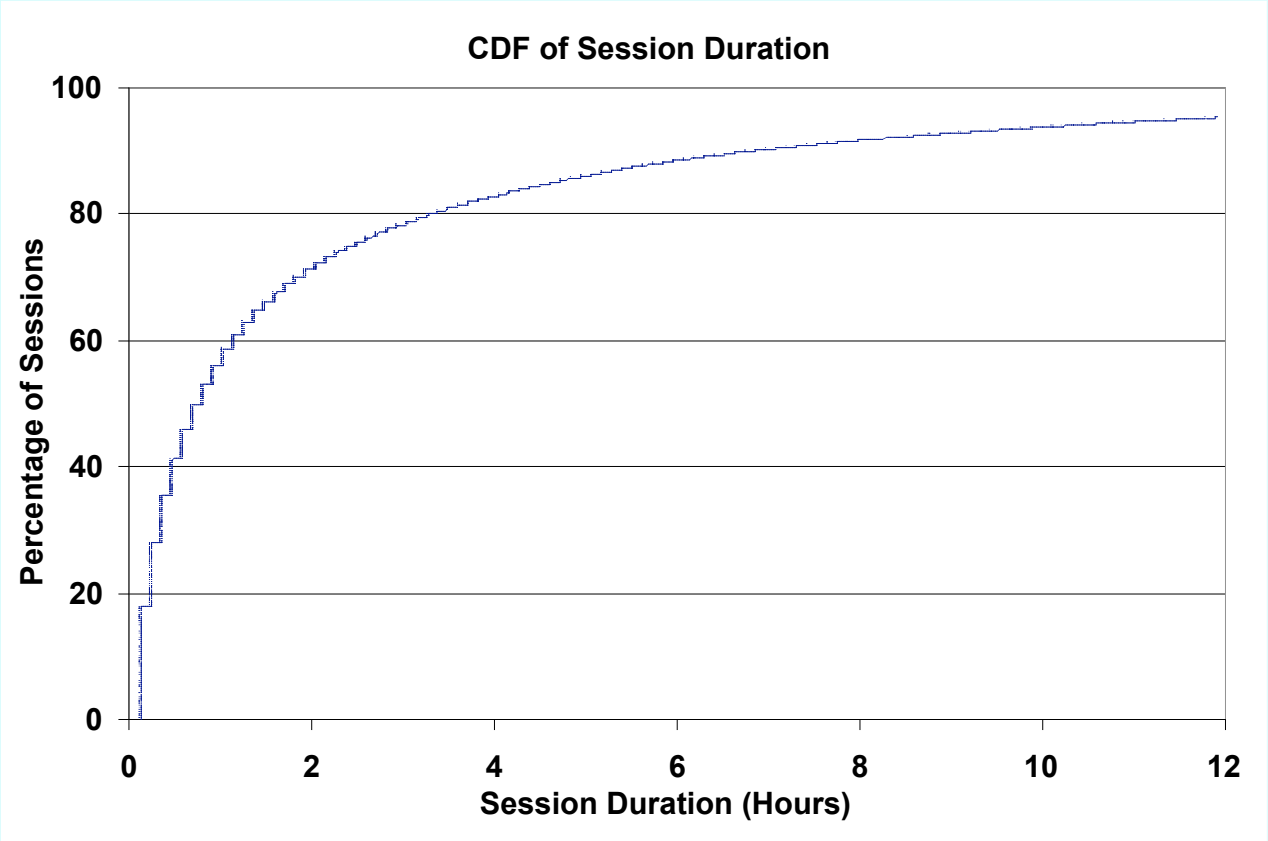
22% of peers have upstream bw <= 100Kbps
8% of peers have upstream bw > 100Mbps

Closest 20% are 50X Closer than Furthest 20%



20% of peers have latencies of at most 70ms
7% of peers have latencies of at least 1 sec

Median Session is about One Hour



50% of sessions last at most 1 hour
11% of sessions last at least 4 hours

Huge Degree of Heterogeneity

- **Heterogeneity:**
  - 3 orders of magnitude of bandwidth
    - 50Kbps-100Mbps
  - 4 orders of magnitude of latency
    - 10us-10s
  - 4+ orders of magnitude in availability
    - 1%-99.99%

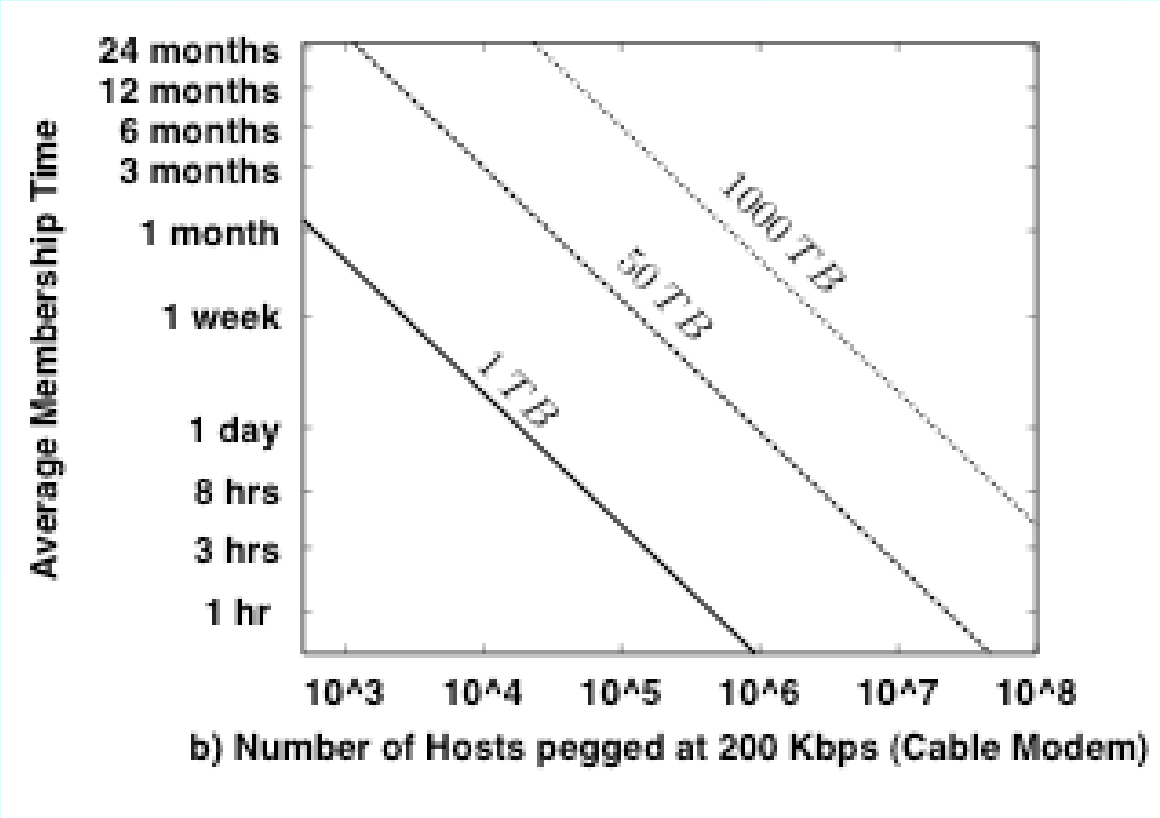
- **Lesson:**
  - Delegate responsibilities across peers in a P2P system

Implications

Data Maintenance Model

- **S = total amount of storage**
- **Storage per node = S/N**
- **rate of joins = rate of leaves = α joins per time**
  - Node lifetime = N/ α
- **Bandwidth per time:**
  - αS/N for joining
  - αS/N for leaving
  - 2 αS/N total
- **Bandwidth per node per time:**
  - 2 S/(N * Lifetime)

Model's Results



1 million peers must have 1 month lifetimes to maintain 1 PB
1 million peers only contribute 1 GB of unique data (20 GB of total)

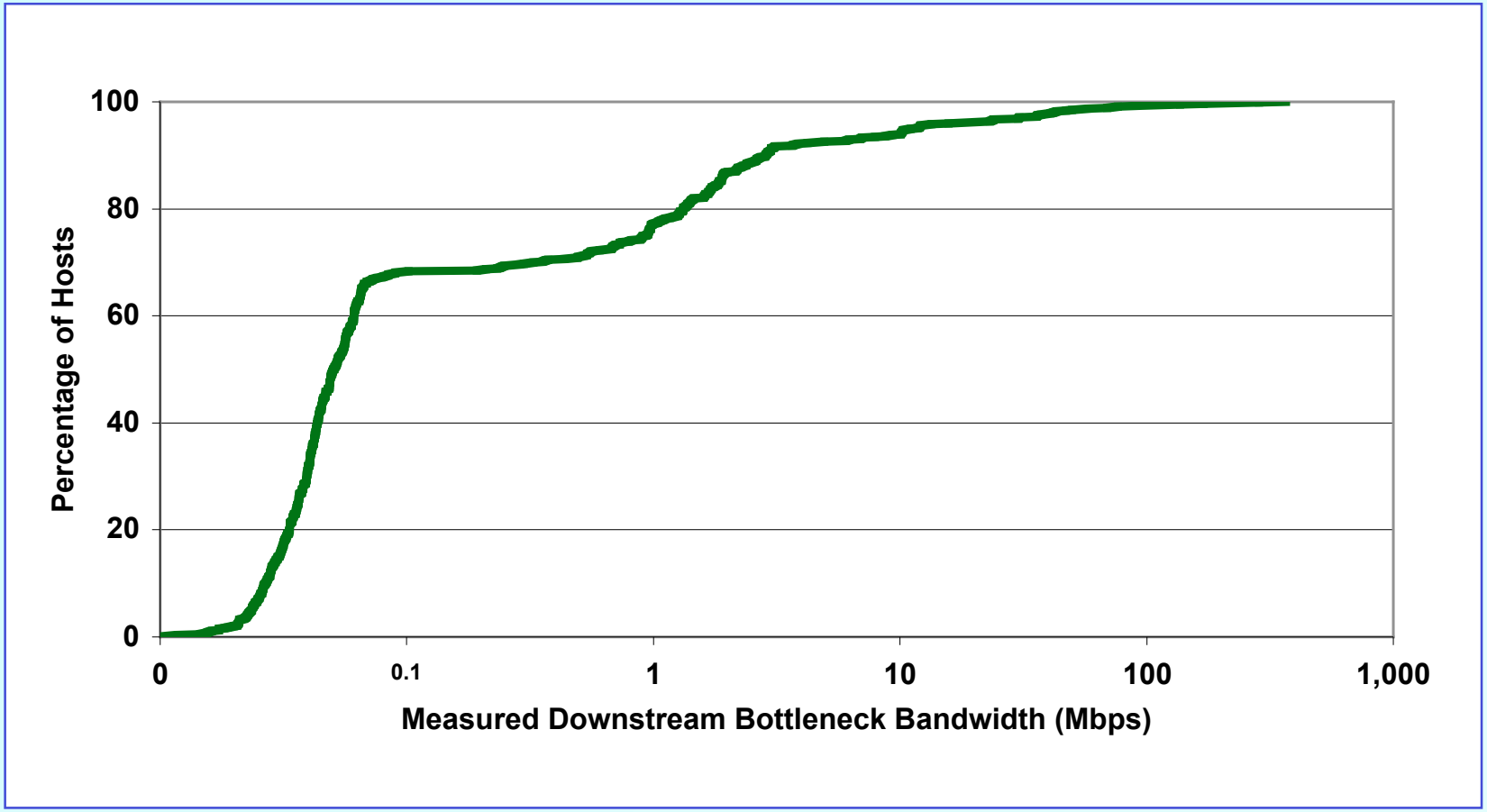
Are Peers Well-Behaved in Practice?

- **Will peers lie if there is an incentive to do so?**
  - Should we design incentive compatible systems?

Methodology

- **In Gnutella:**
  - Each peer voluntarily reports Internet connection type
    - Dialup, cable modem, DSL, T1, T3, Unknown
  - Well-connected peers have incentive to report lower bandwidths in order to shed load
- **But… we also measure each peer's bandwidth**
- **Experiment:**
  - Compare reported and measured bandwidths

Peers Reporting Dialup Bandwidths



Peers lie if there is an incentive to lie

Conclusions – Characterizing Peers

- **Delegate responsibilities across peers in a P2P system**
  - Significant amount of heterogeneity across peers

- **Build incentive in P2P designs**
- **Incorporate direct measurement techniques**
  - Peers deliberately misreport information if there is an incentive to do so

Problems

- **DHCP/Aliasing effects**
- **Lack of metrics:**
  - We have one project attempting to fix this problem!

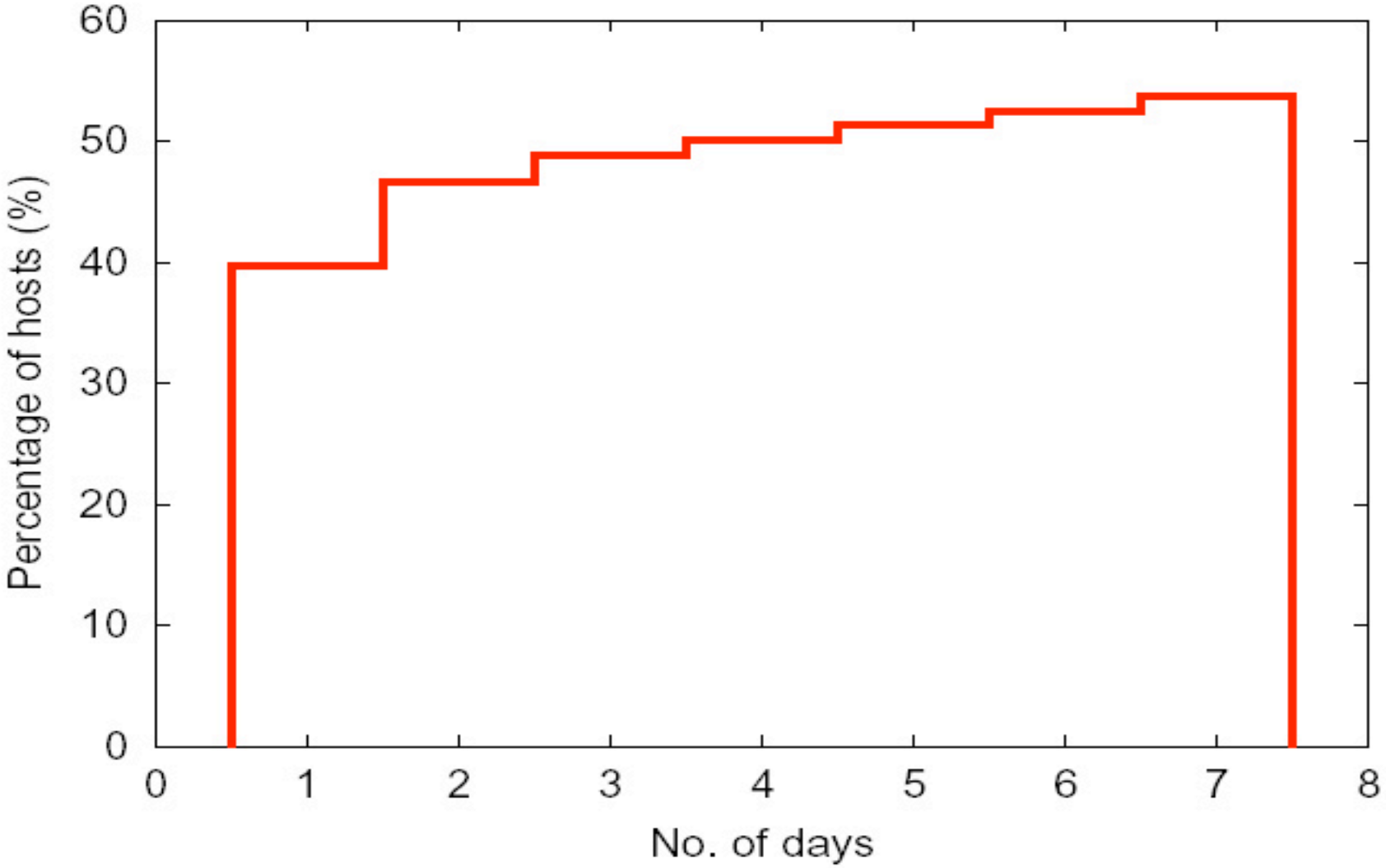


Figure 1: Percentage of hosts that have more than one IP address across different periods of time.

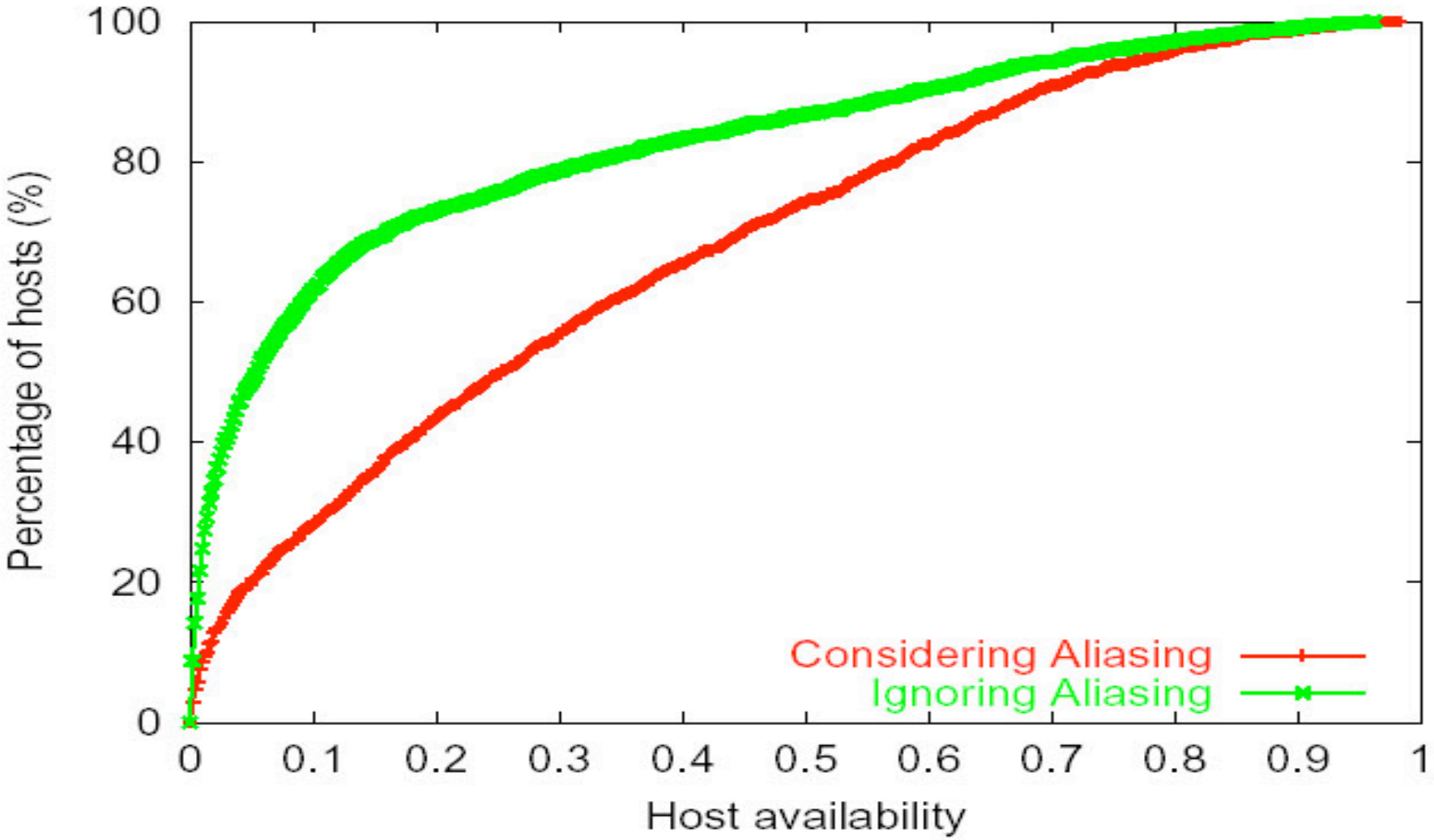


Figure 2: Host availability derived using unique host ID probes vs. IP address probes.

Implications of aliasing

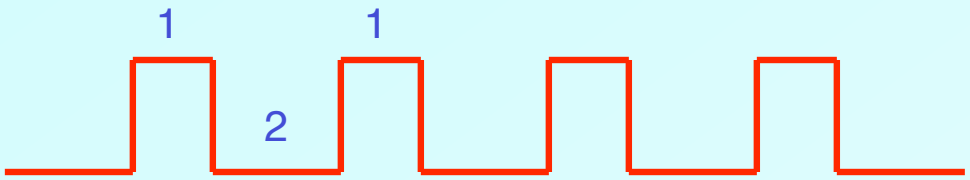
The use of IP address-based probing …

- **would thus underestimate availability by a factor of 4!**

My take on it

- **It's unclear whether "Understanding Availability" understands availability**
- **We don't have the right metrics to measure availability**
  - MTTF and MTTR do not capture a sys's availability
- **Peer's uptime is useful only when it's considered relative to other peers**

Availability vs. Number of 9s



1. **Once request made, request is pending for 3 seconds before time-out**
2. **A request takes 1 second to complete**

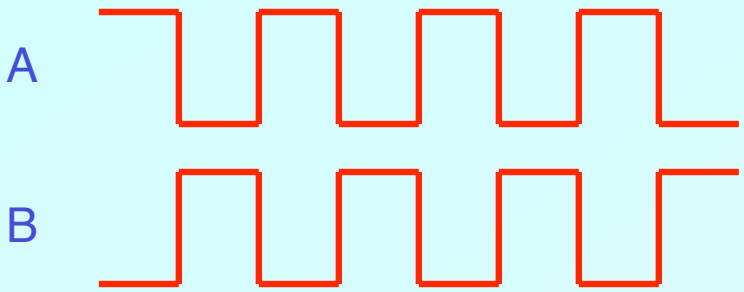
- **A system with zero 9s has perfect availability**
  - Up 1 sec, down 2 secs (zero 9s)

MTTR < time_out && MTTF > service_time

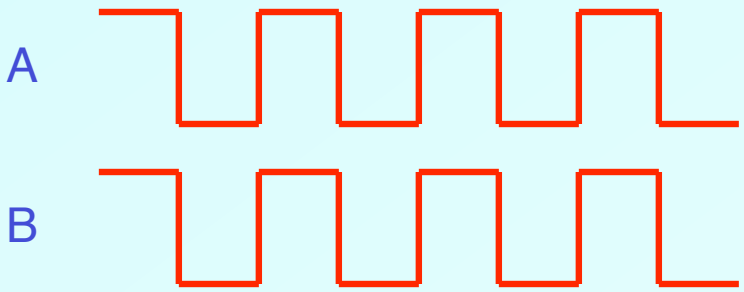
My take on it

- **We don't have the right metrics to measure availability**
  - MTTF and MTTR do not capture a sys's availability
- **Peer's uptime is useful only when it's considered relative to other peers**

Is System 1 more available than System 2?



System 1



System 2

- **Number of nines is the same in both systems 1 and 2**
- **A system is perfectly available iff every request is served**

Thesis Topics

- **Ph.D. Thesis topic:**
  - Understand what availability means in the context of distributed systems:
    - Huge open problem

- **M.Sc. Thesis topic:**
  - Almost all measurement projects are subject to aliasing effects (DHCP, NATs, multiple NICs)
    - Both active and passive measurements
  - Understand how DHCP servers allocate IP addreses
  - Propose heuristics to differentiate multiple clients behind NAT
  - Validate them and propose a model