

Measuring the Internet using BitTorrent

Tomas Isdal

Royal Institute of Technology, Stockholm, Sweden

University of Washington, Seattle, USA

Joint work with Harsha, Michael, Colin, Tom, Arvind, Arun

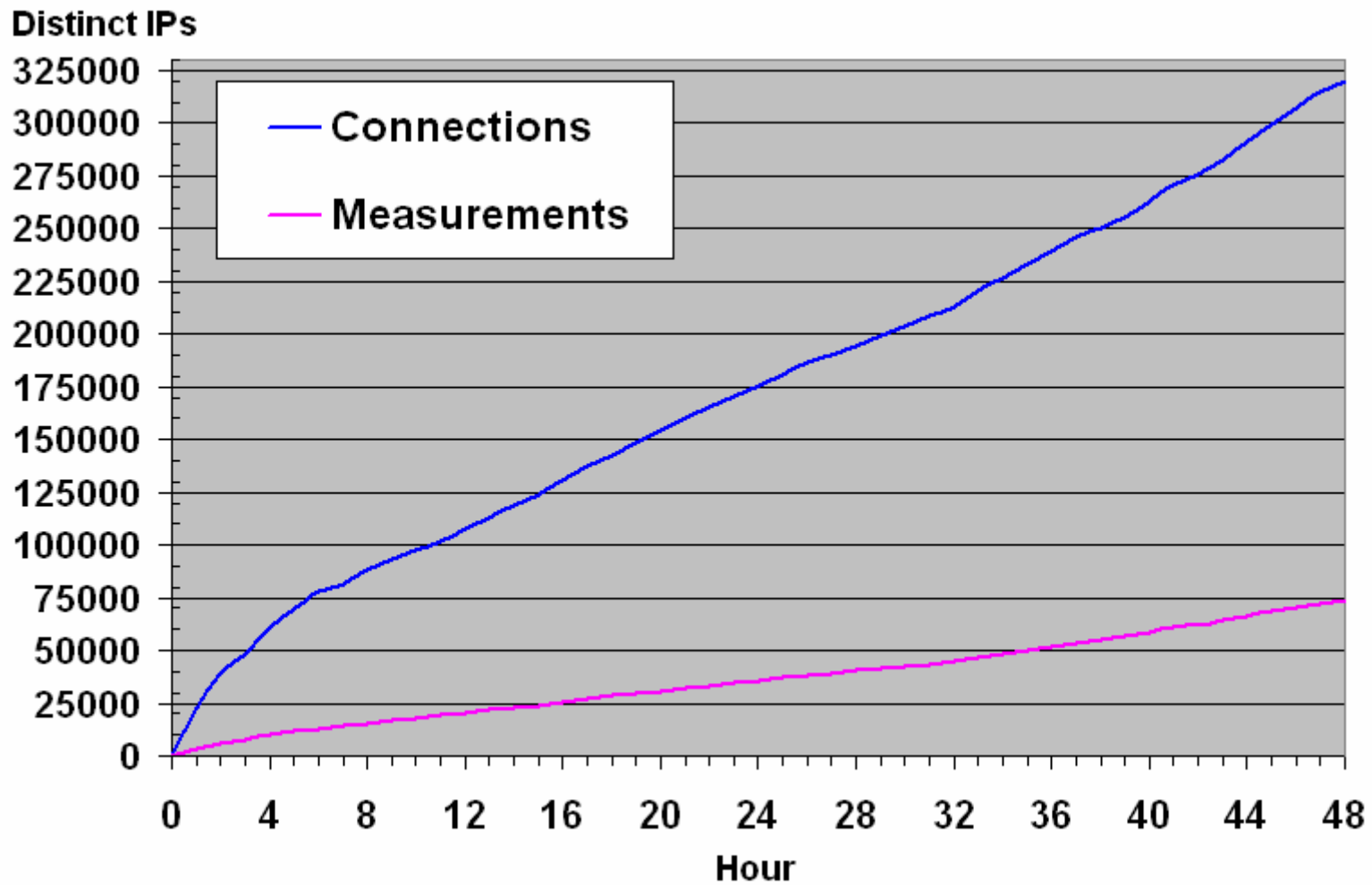
Measurements to end hosts

- End host measurements are tricky
 - Uncooperative hosts
 - Triggering of firewalls / IDSs
 - Finding targets can be difficult
- BitTorrent has useful properties
 - Millions of users / targets
 - Optimistic unchoke
 - No unexpected packets are sent

Measurements using BitTorrent

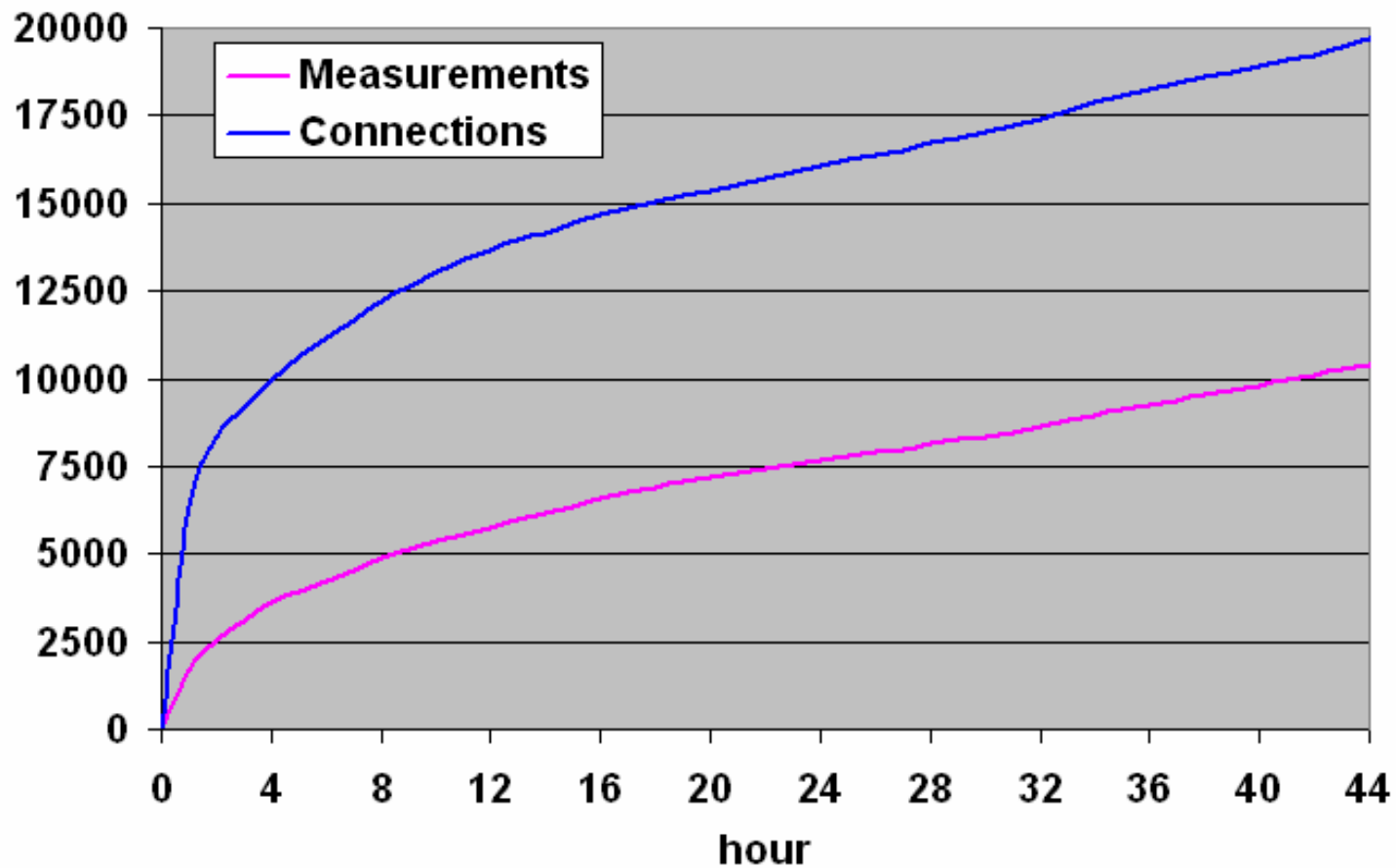
- Parse BitTorrent sites for popular torrents
- Join torrents from PlanetLab machines
- Wait for optimistic unchoke
- Mangle TCP packet stream to determine:
 - Bandwidth capacity
 - (Latency)
 - (Topology)

Coverage

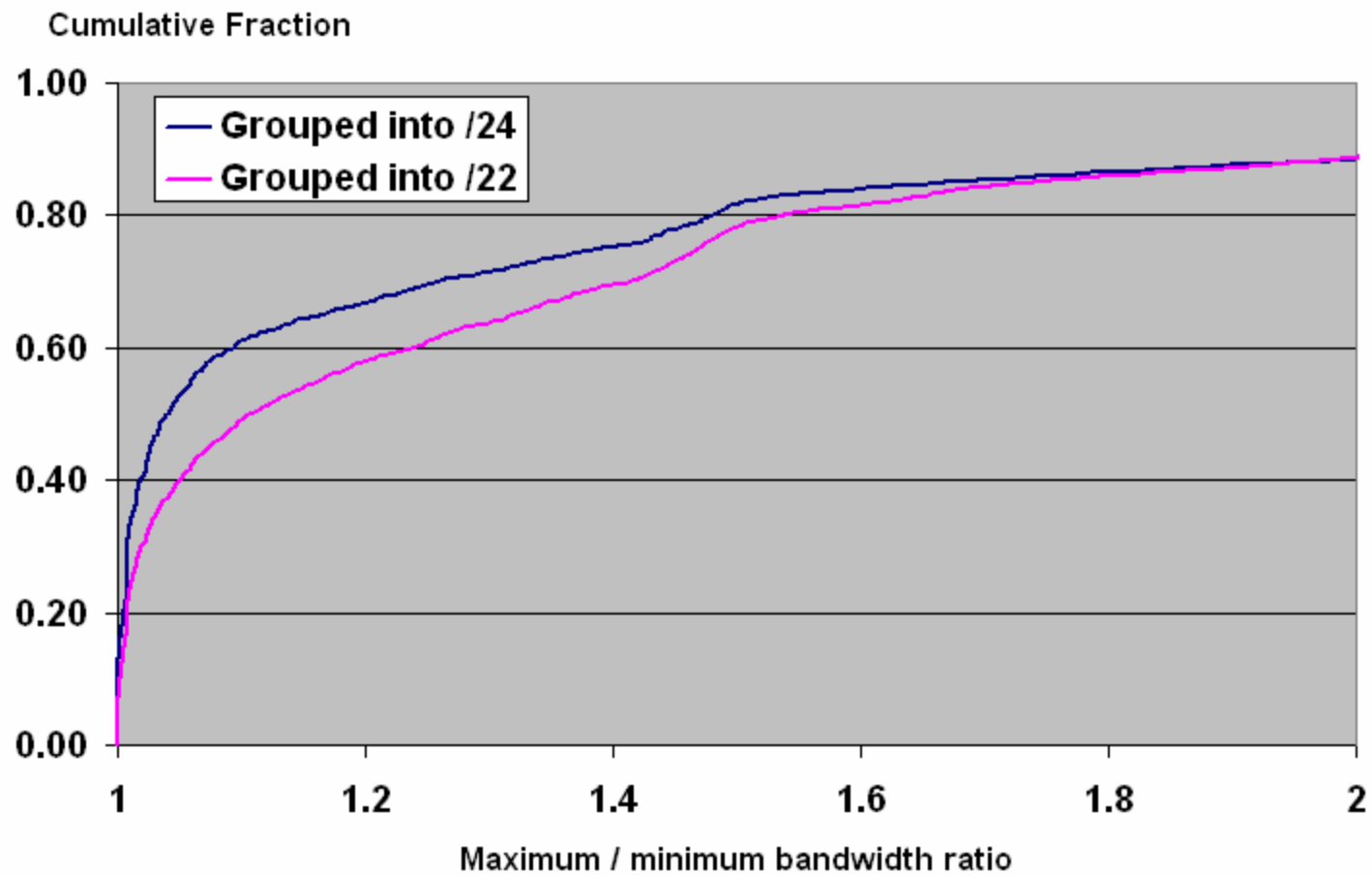


Coverage cont.

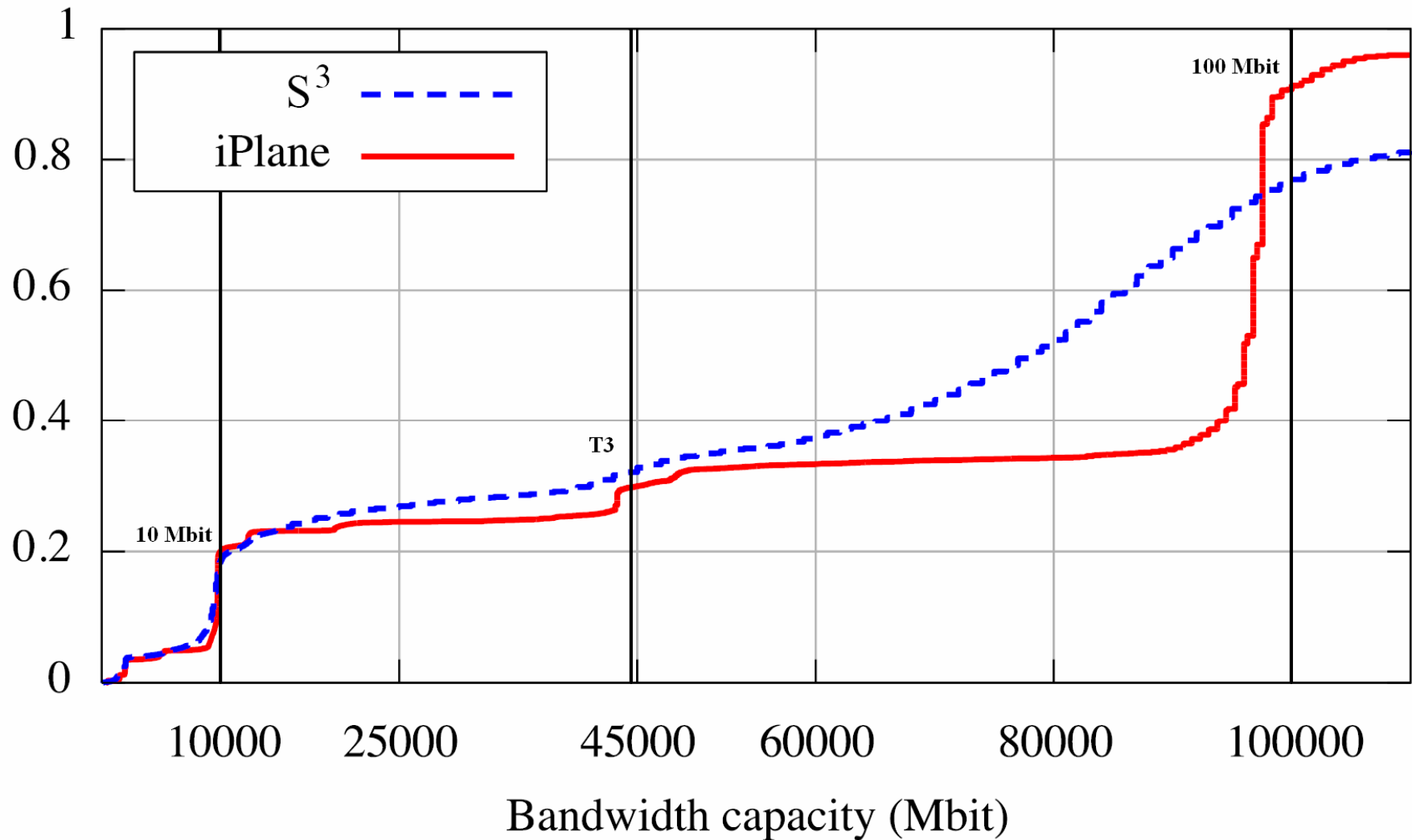
BGP prefixes



Similarity of hosts



Validation on PlanetLab



Preliminary results

- Good coverage
- Hosts “close” to each other have similar characteristics
- Measurement of uplink capacity accurate

Problems

- Difficult to measure downlink capacity
 - Downlink measurements requires uploading
 - Have to rely on ACK packet streams
 - Limited number of “legal” torrents
- Fairly bandwidth expensive
 - ~500 packets required for accurate results
- Opportunistic selection of measured hosts

Future tasks

- Increase accuracy of downlink capacity measurements
 - Prevent TCP delayed ACKs → twice as many datapoints
- Legal torrents would be better...
 - Makes it possible (not illegal) to upload / store data
 - Increased unchoking → more measurements
 - Enables measurements of downlink capacity
 - World of Warcraft patches
 - 5 million subscribers
 - Worldwide distribution
 - Legal distribution of movies
 - Warner Bros

Future tasks cont.

- Intelligent selection of torrents
 - During the 48 h run only 60% of torrents provided measurement data
 - Some torrents are much more useful than others
- Improvement of MultiQ algorithm
 - Only 75,000 of 130,000 possible measurements