Tracking down Skype Traffic

Dario Bonfiglio
Marco Mellia
Michela Meo
Nicolo' Ritacca
Dario Rossi

TELECOM PARIS
école nationale supérieure des télécommunications

POLITECNICO DI TORINO
1859 1906

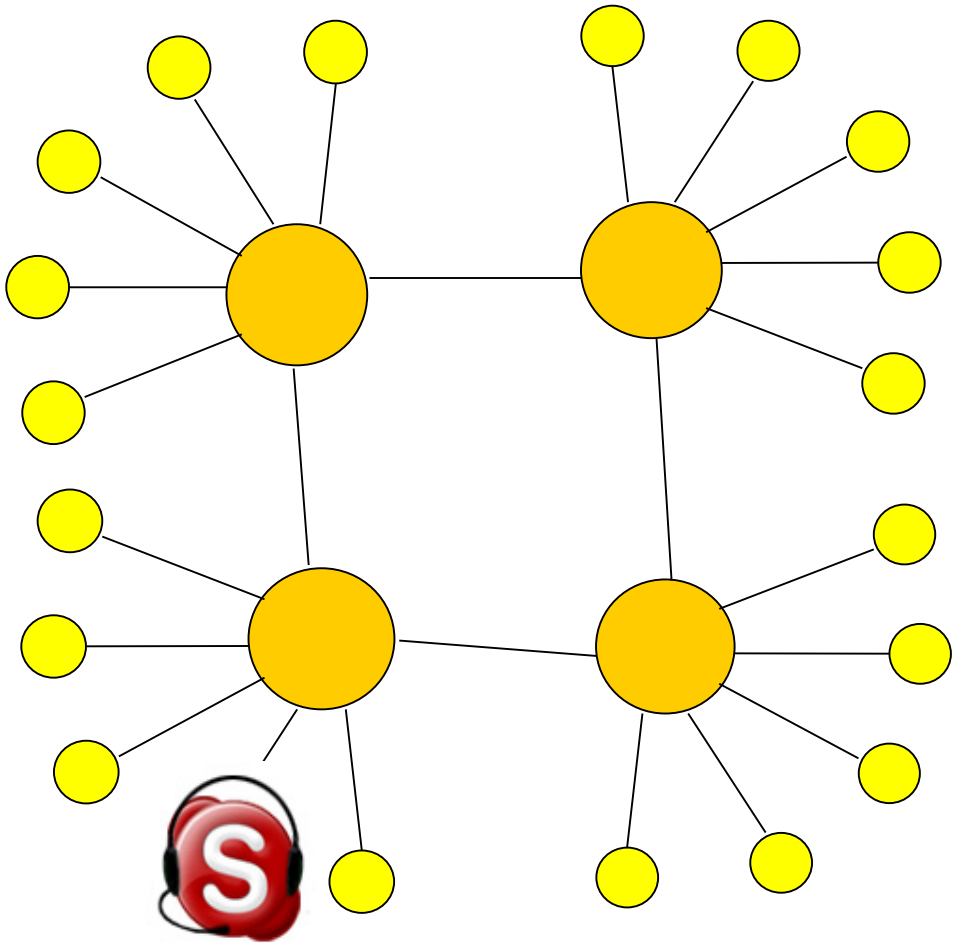
Agenda

- A few words about Skype
  - Known facts
  - Preliminary definitions
- Investigate Skype "Traffic"
  - Voice traffic
    - Reaction to network performance degradation
  - Signaling traffic
    - Signaling patterns & peer selection
  - Users' behavior
    - Please, see the paper

Why Skype ?

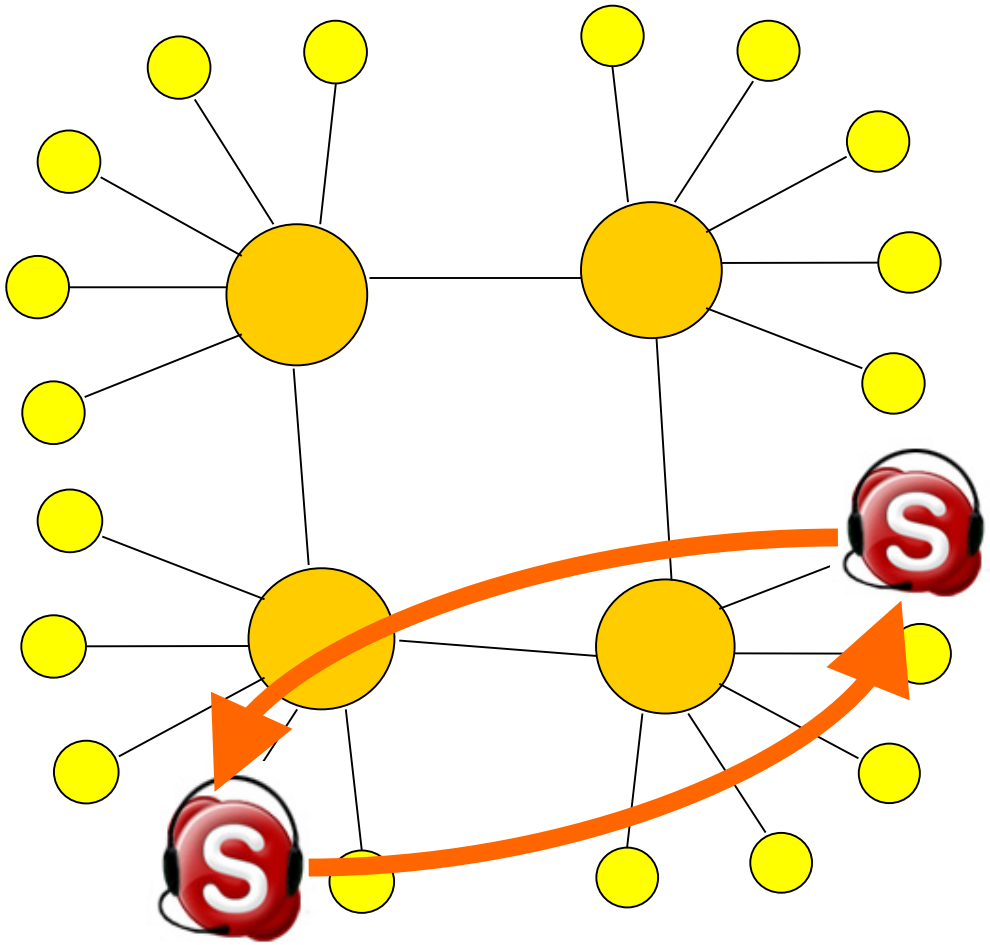
- Skype is very popular
  - More than 100M users, 5% of all VoIP traffic
  - Easy to use, many free services
    - voice / video / chat / data transfer over IP
- Understanding Skype is a challenging task
  - Closed design, proprietary solutions
  - Almost everything is encrypted
  - Uses a P2P architecture
  - Lot of different flavors

Skype for Dummies



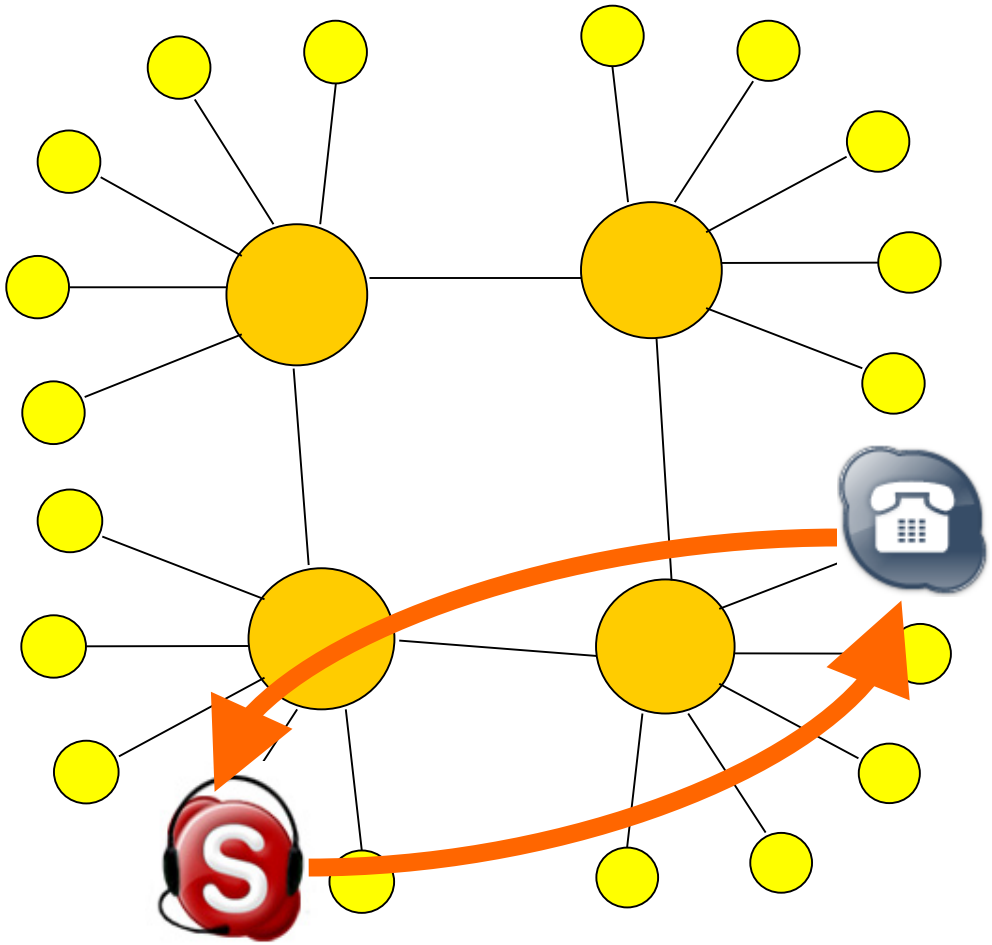
- Architecture
  - P2P design

Skype for Dummies



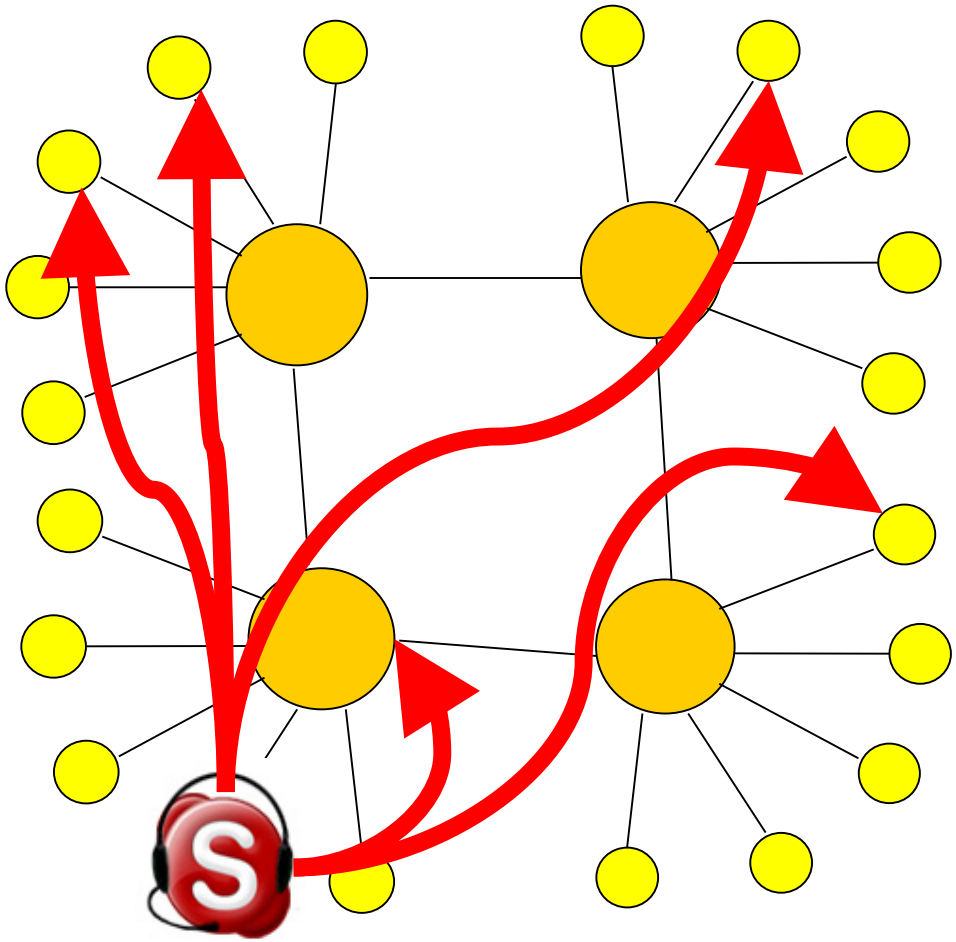
- Architecture
  - P2P design
- **Service traffic**
  - Voice calls
  - Video calls
  - Chat
  - Data transmission

Skype for Dummies



- Architecture
  - P2P design
- **Service traffic**
  - Voice calls
  - Video calls
  - Chat
  - Data transmission
  - Skypeout/Skypein

Skype for Dummies



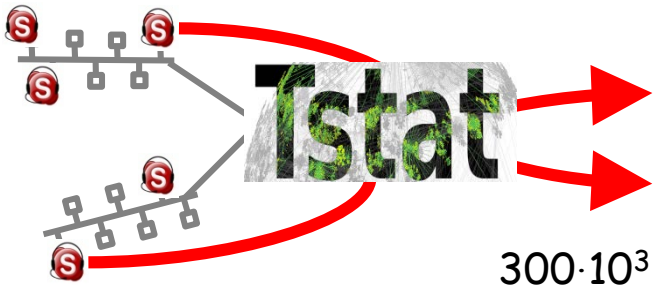
- Architecture
  - P2P design
- **Service traffic**
  - Voice calls
  - Video calls
  - Chat
  - Data transmission
  - Skypeout/Skypein
- **Signaling traffic**
  - Login & auth.
  - Look for buddies
  - ....

Methodolody

- **Service traffic**
  - Small scale active testbed
  - Controlled bandwidth, packet loss
  - voice service,
    many Codecs, TCP/UDP traffic

- **Signaling traffic**
  - Passive measurement technique
  - Adopt a black-box approach
  - Inspect and quantify UDP signaling
  - Classification framework:

    D.Bonfiglio, M.Mellia, M.Meo, D.Rossi, P.Tofanelli, *Revealing Skype Traffic: When Randomness Plays with You*, SIGCOMM'07



7000 hosts
1700 peers

$300 \cdot 10^3$
external
peers

Preliminary Definition

- **Useful information**
  - At installation, Skype chooses a port at random
  - The port is never changed (unless forced by the user)
  - All traffic multiplexed over the same socket (UDP preferably)

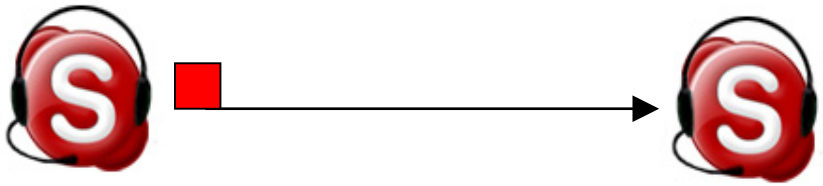
Skype peer

- A Skype peer can be identified by its endpoint
- Consider only peers that were ever observed making a call

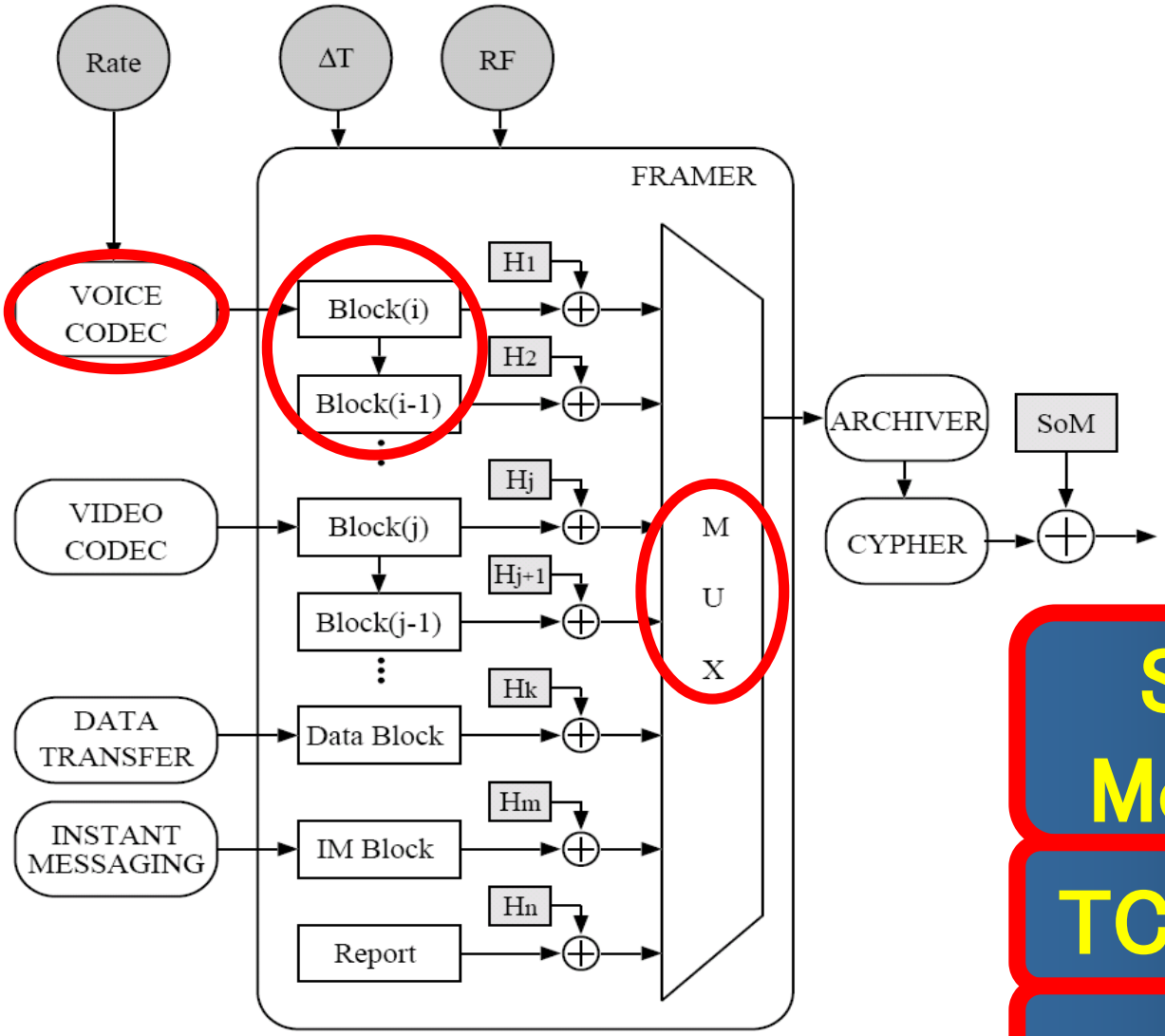
(IP addr, UDP port)

Skype flow

- A sequence of packets originated from a Skype peer (and destined to another skype peer)
- Flow starts when the first packet is observed
- Flow ends when no packet is observed for a given inactivity timeout (200s)



Skype Source Model



Skype Message

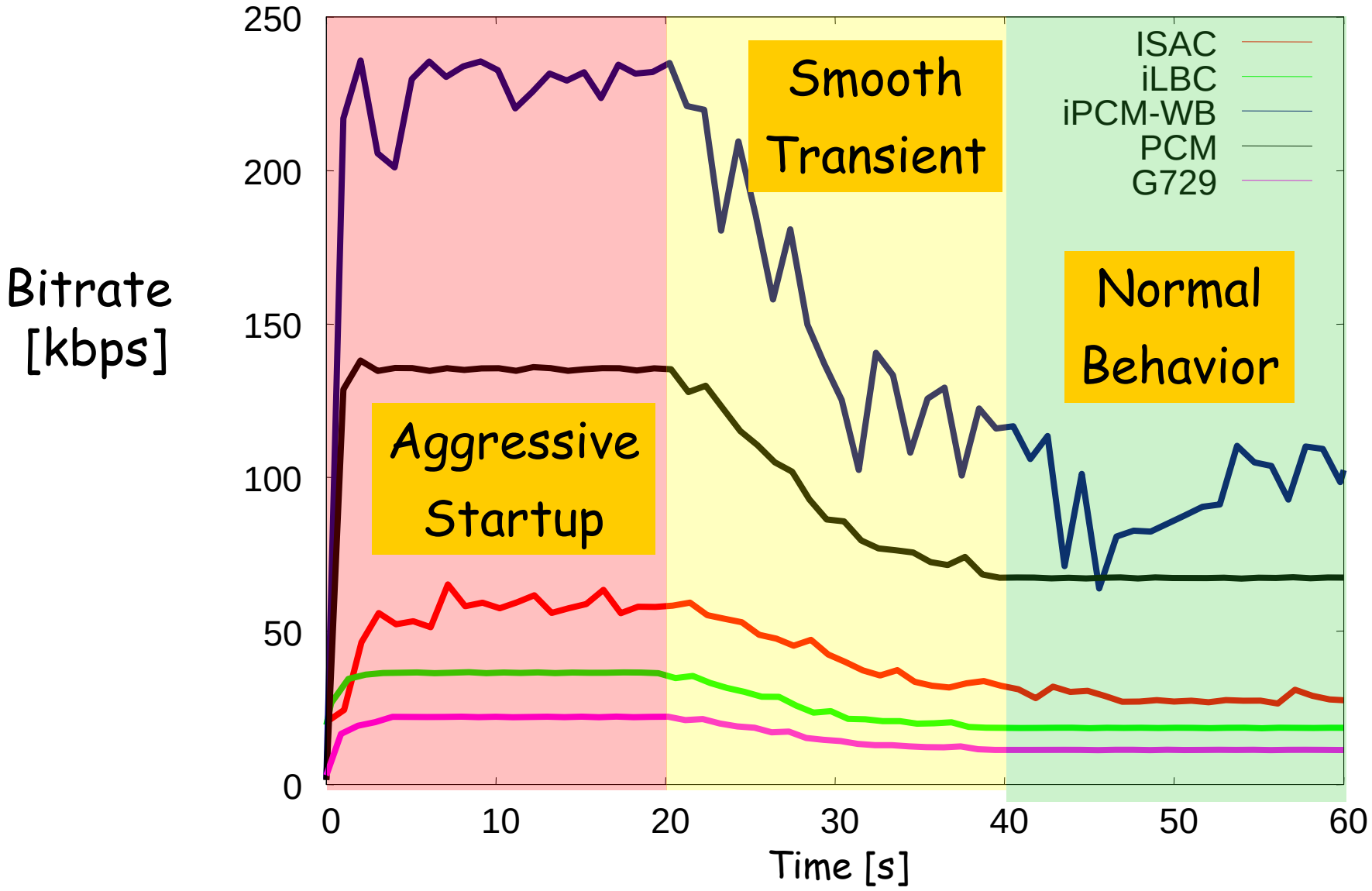
TCP/UDP

IP

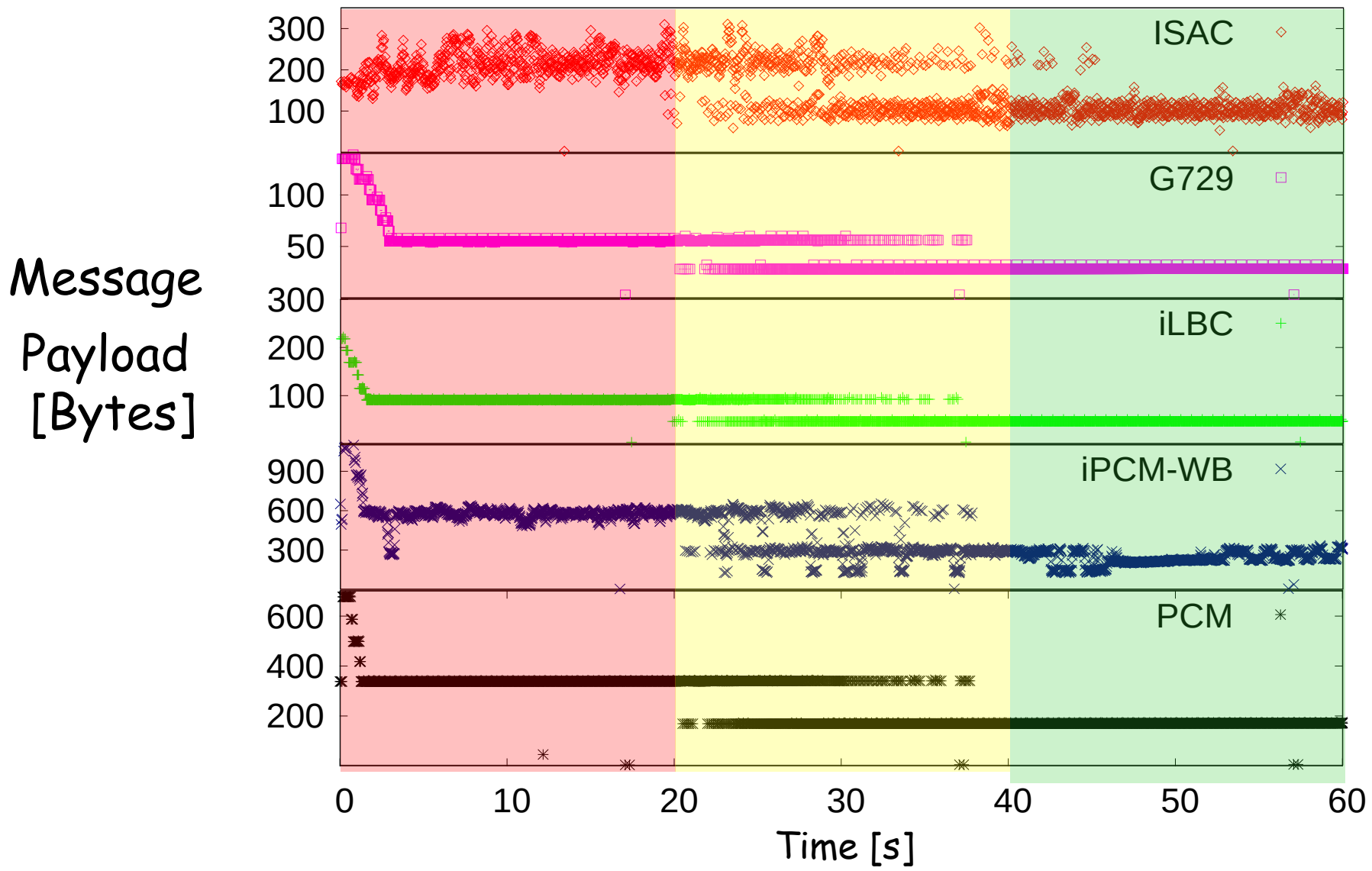
Service traffic

Codec Impact

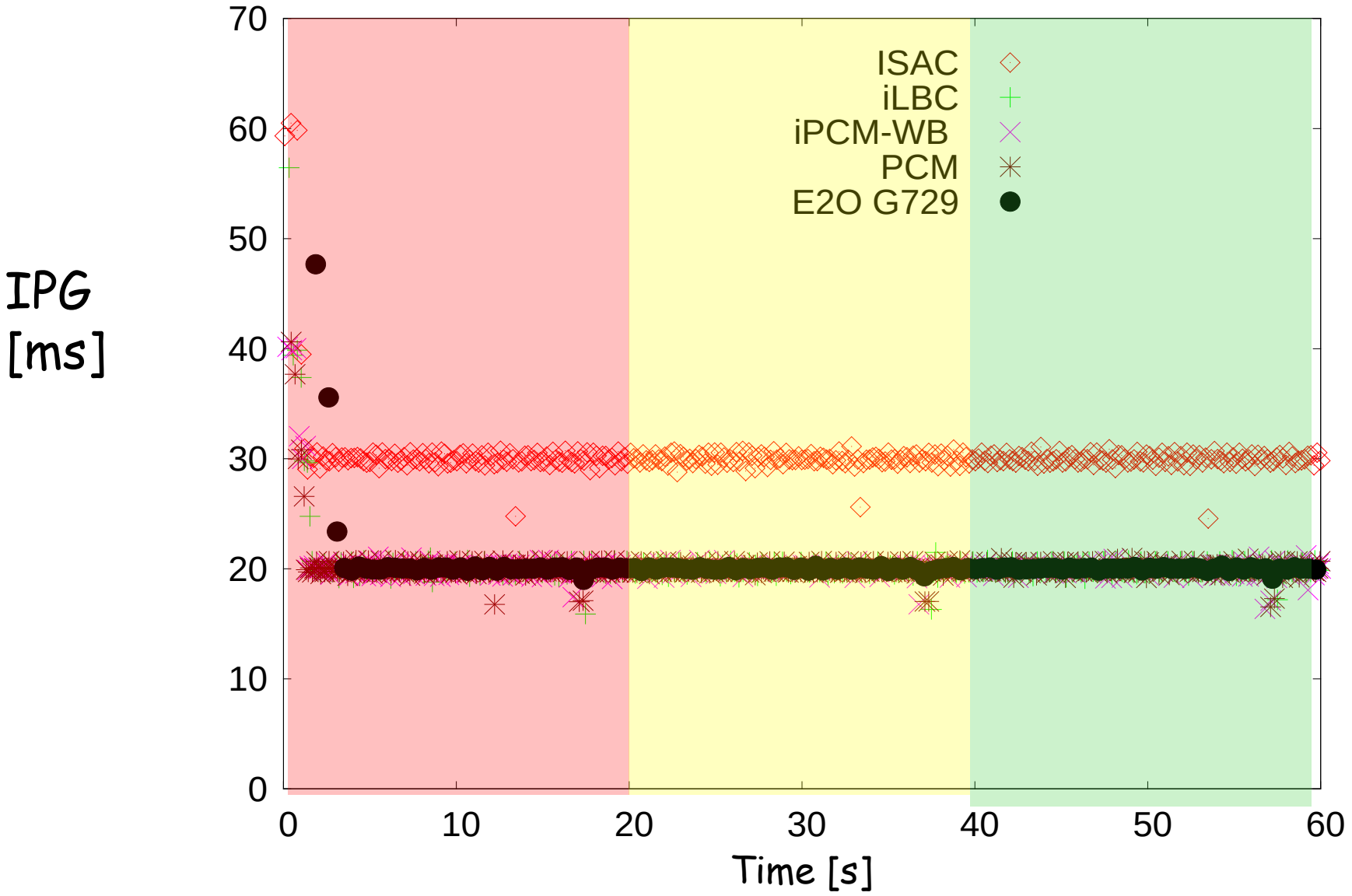
Service Traffic: Normal Condition



Service Traffic: Normal Condition



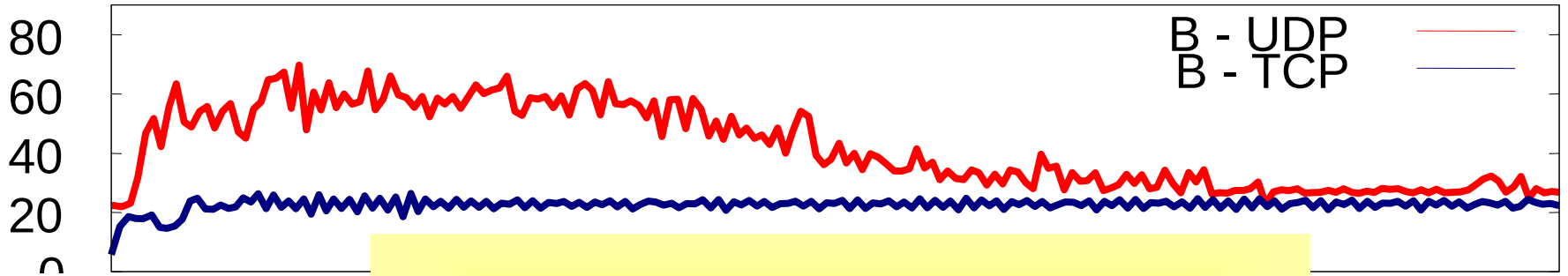
Service Traffic: Normal Condition



Service traffic

Transport Layer Impact

Service Traffic: TCP vs UDP

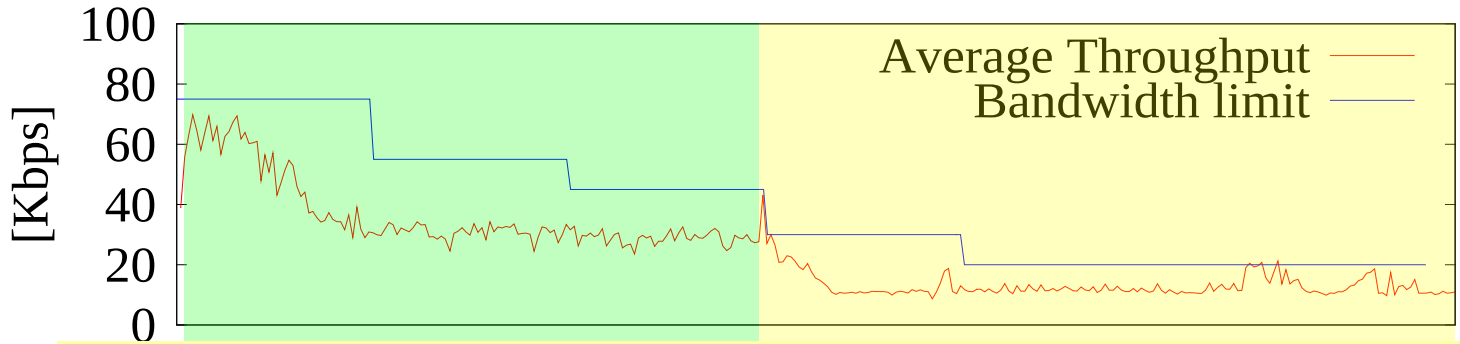


TCP/UDP have no impact

Service traffic

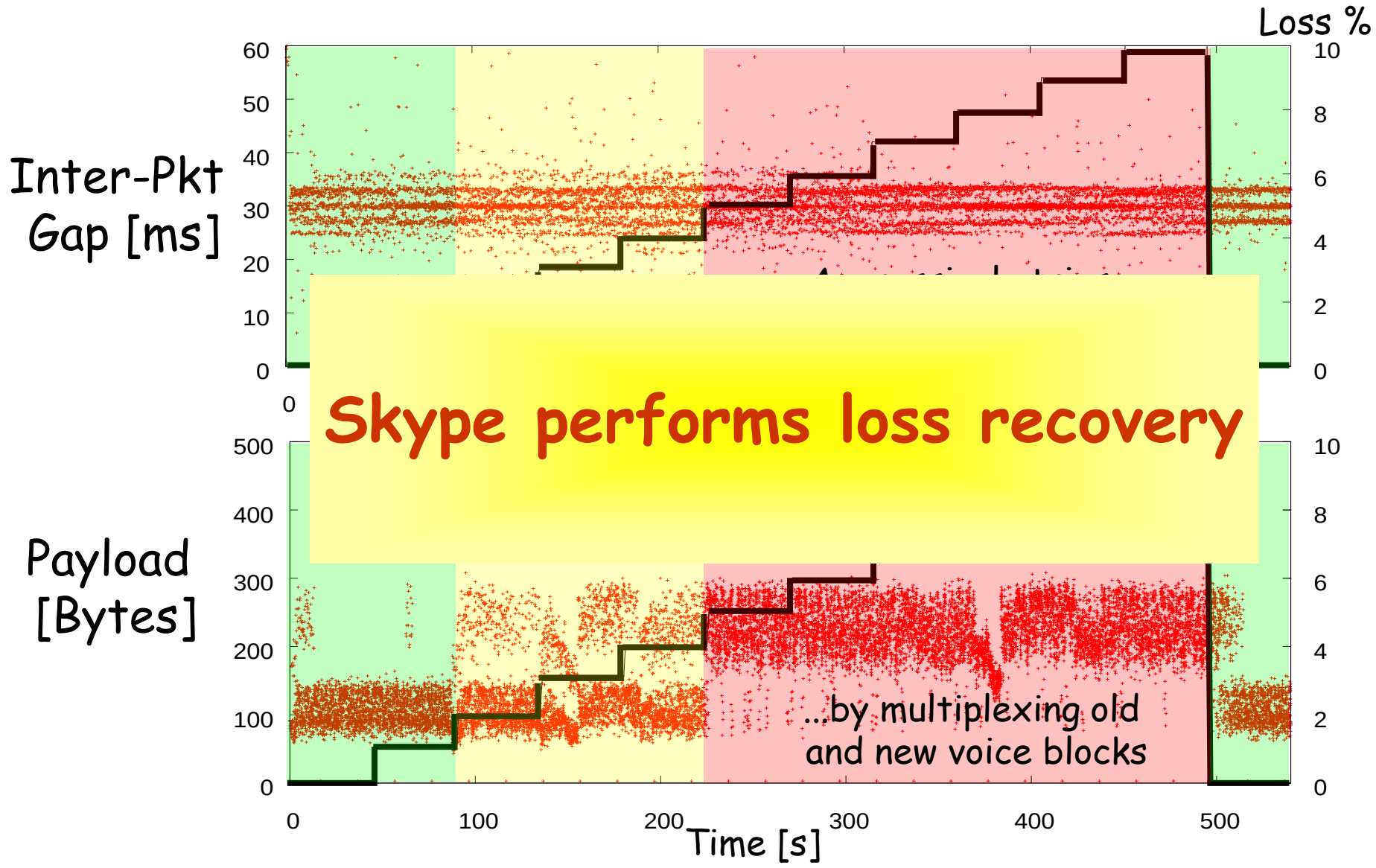
Network Impact

Service Traffic: Bandwidth Limit



Skype performs congestion control

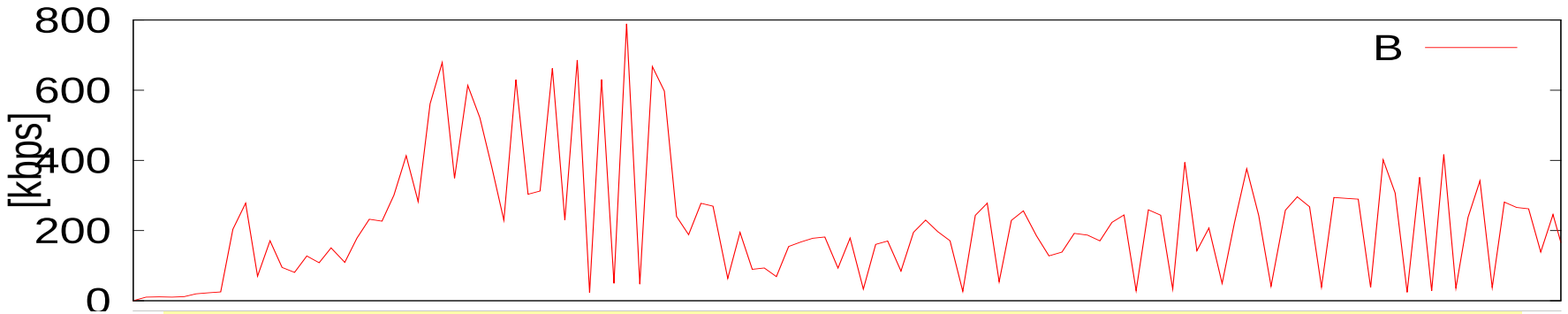
Service Traffic: Packet Loss



Service traffic

Video Traffic

Service Traffic: Video Source

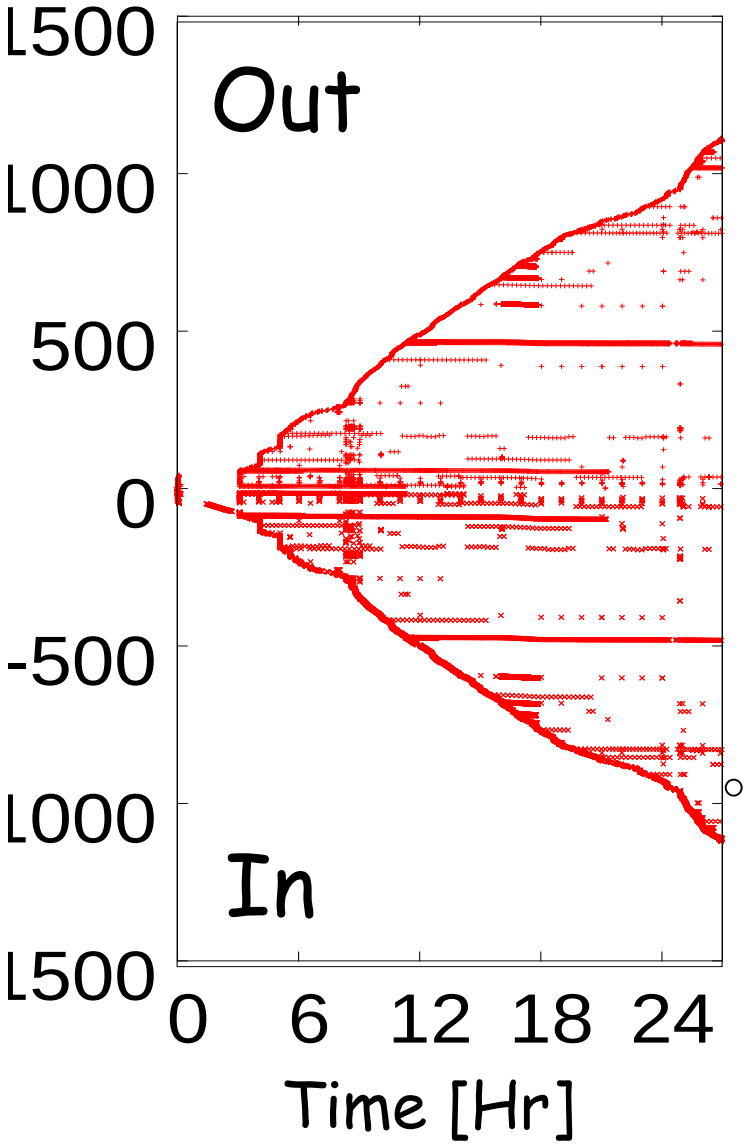


Skype multiplexes different sources

Usual IPG for
Back-to-back video
Pure audio messages
Massages => frame

Signaling traffic

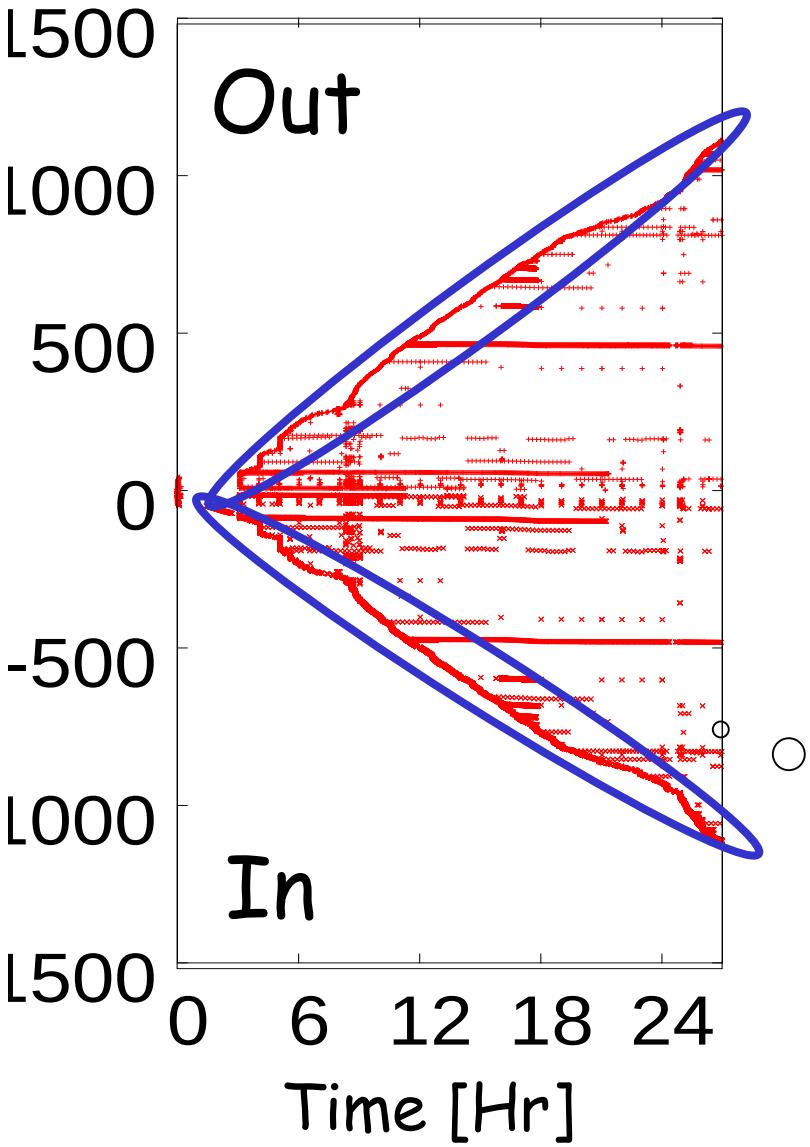
Signaling Traffic: Activity Pattern



- **Legend**
  - Consider a single client
  - Each dot is a packet
  - Top: outgoing, Bottom: incoming
  - For every new peer, increment the ID
  - For every old peer, use the previous ID

Rather different patterns emerge from the plot

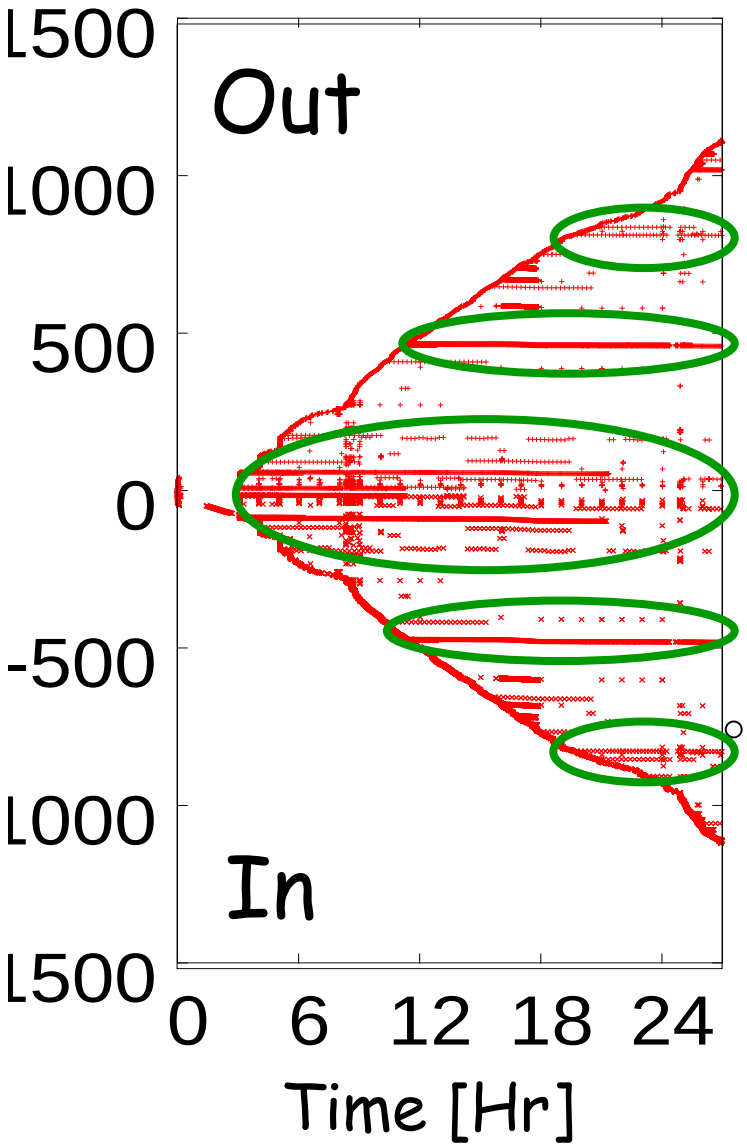
Signaling Traffic: Activity Pattern



- **Probes**
  - Single packet
  - Sent toward unknown peers
  - Reply possibly follows
  - No further traffic between the same peers pair
  - Majority of the flows

Peer discovery is a continuous task

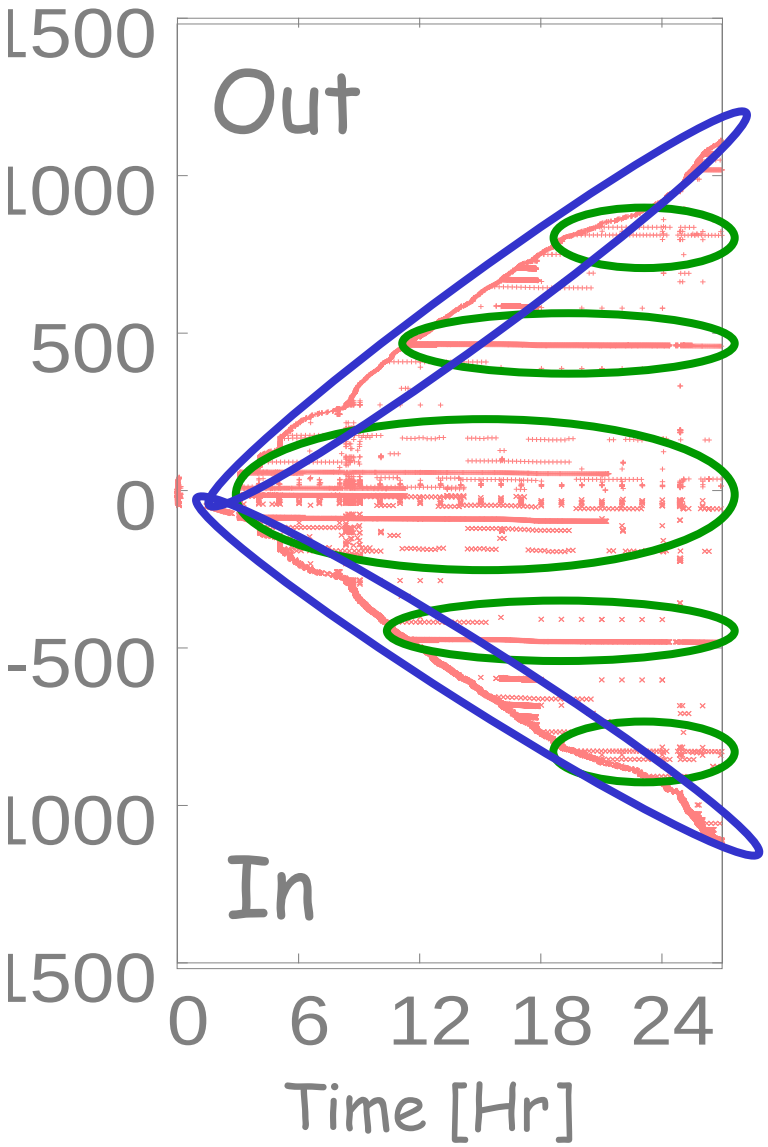
Signaling Traffic: Activity Pattern



- **Non-Probes**
  - Flows longer than one packet
  - Series of single-packet flows
  - Sent toward the same peer
  - Carry most signaling bytes

Talk to super peers, notify buddies of status change, …

Signaling Traffic: All Peers



- **Probes**
  - Majority of the flows
- **Non-probes**
  - Carry most signaling bytes
- **Signaling bandwidth**
  - 95% generate <100 bps
  - Only 1% exceeds 1 Kbps
- **Signaling spread**
  - 95% of peers contact <40 peers (in 5 min)
  - 1% exceeds >75 (in 5 min)

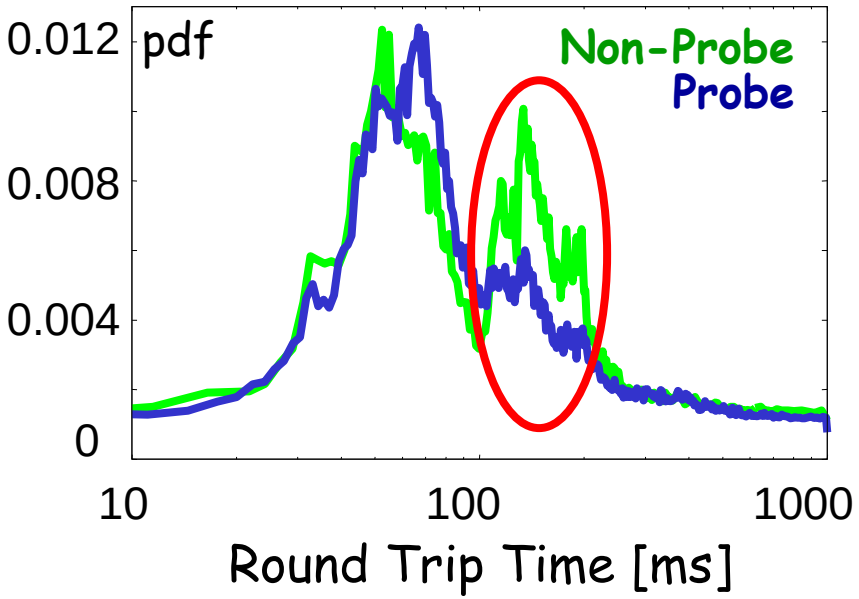
Conclusions

- Service traffic
  - Active testbed
  - Skype implements a congestion control
    - Aggressive with losses
    - Conservative with bottlenecks
- Signaling traffic
  - Passive measurement
  - Two different threads shapes the overlay
    - Probes
    - Non-Probes
  - Signaling rate and spread
    - Very limited bitrate
    - Large number contacted peers
- User Characterization
  - Number of calls per unit of time
  - Call duration for different services
  - Peer Lifetime

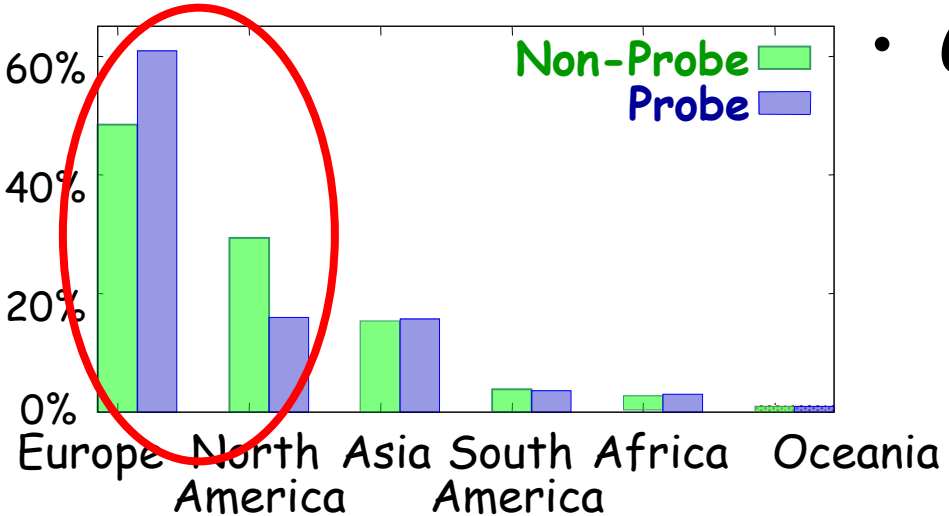
Details are in the paper, not in this talk ☺

- Future Work
  - Extensive measurement in different networks
    - Campus LAN
    - ADSL installation
    - Cellular Network

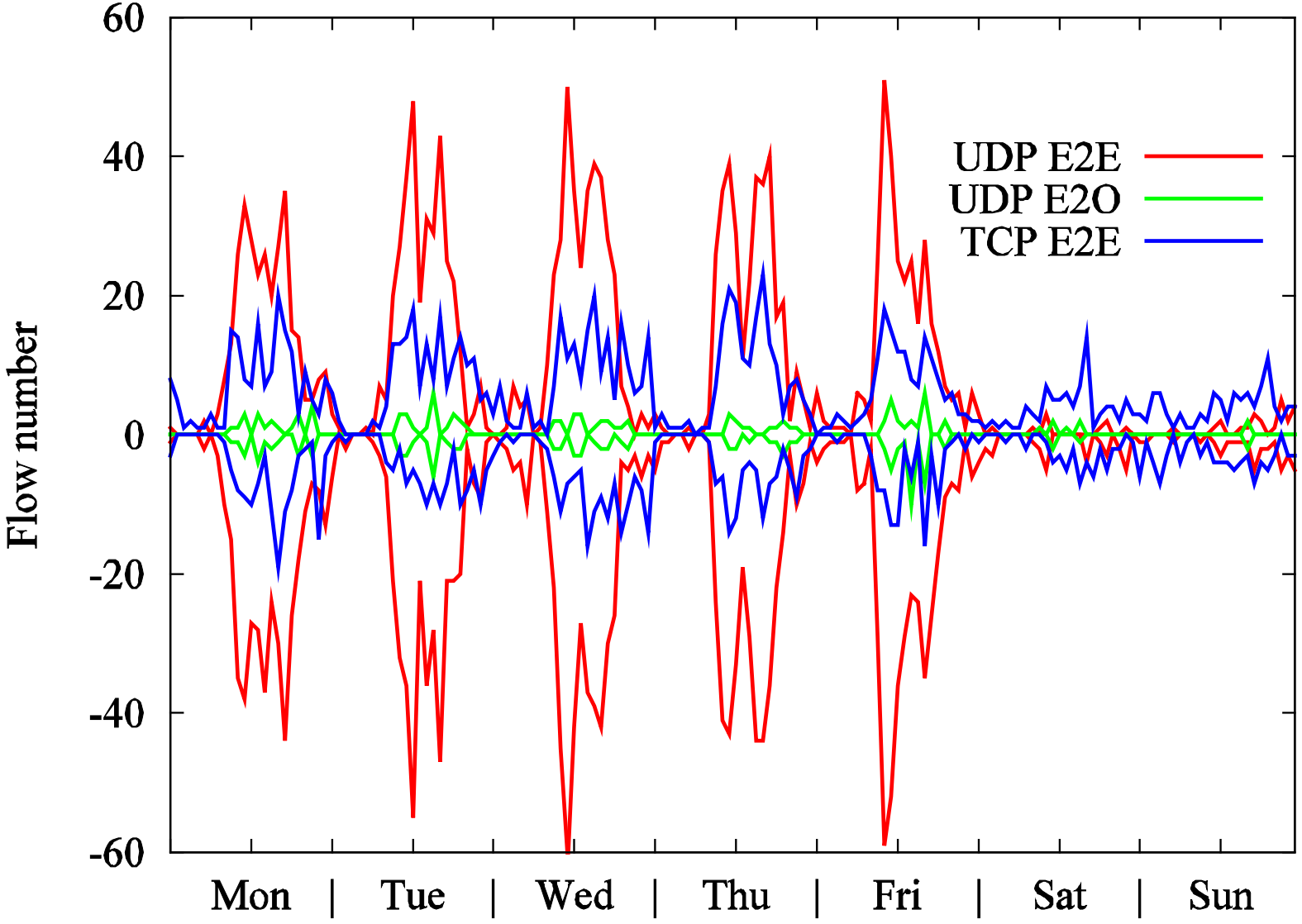
Signaling Traffic: Peer Selection

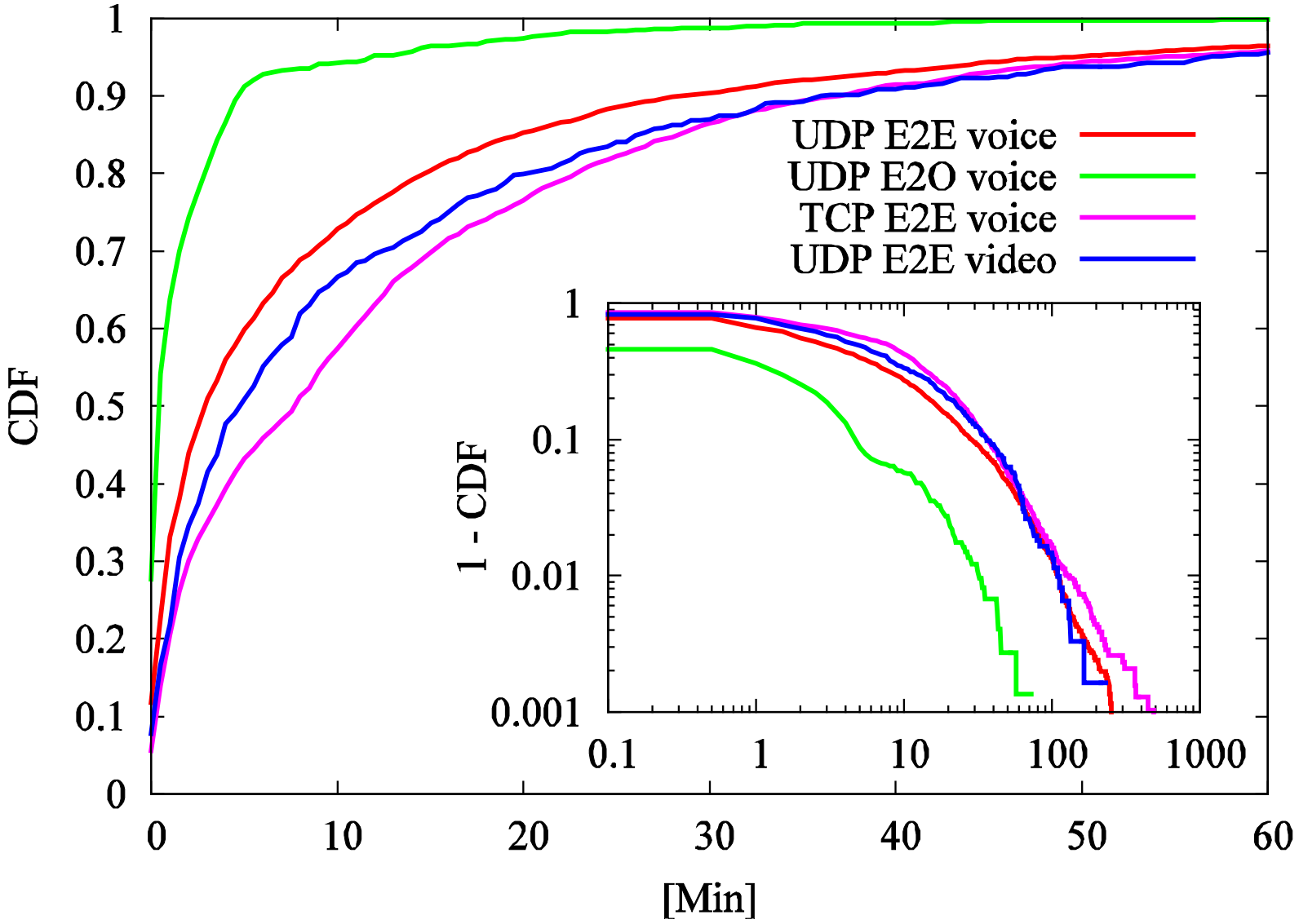


- **RTT distance**
  - RTT between first request-reply packets
  - Probe RTT smaller w.r.t. non-probe traffic

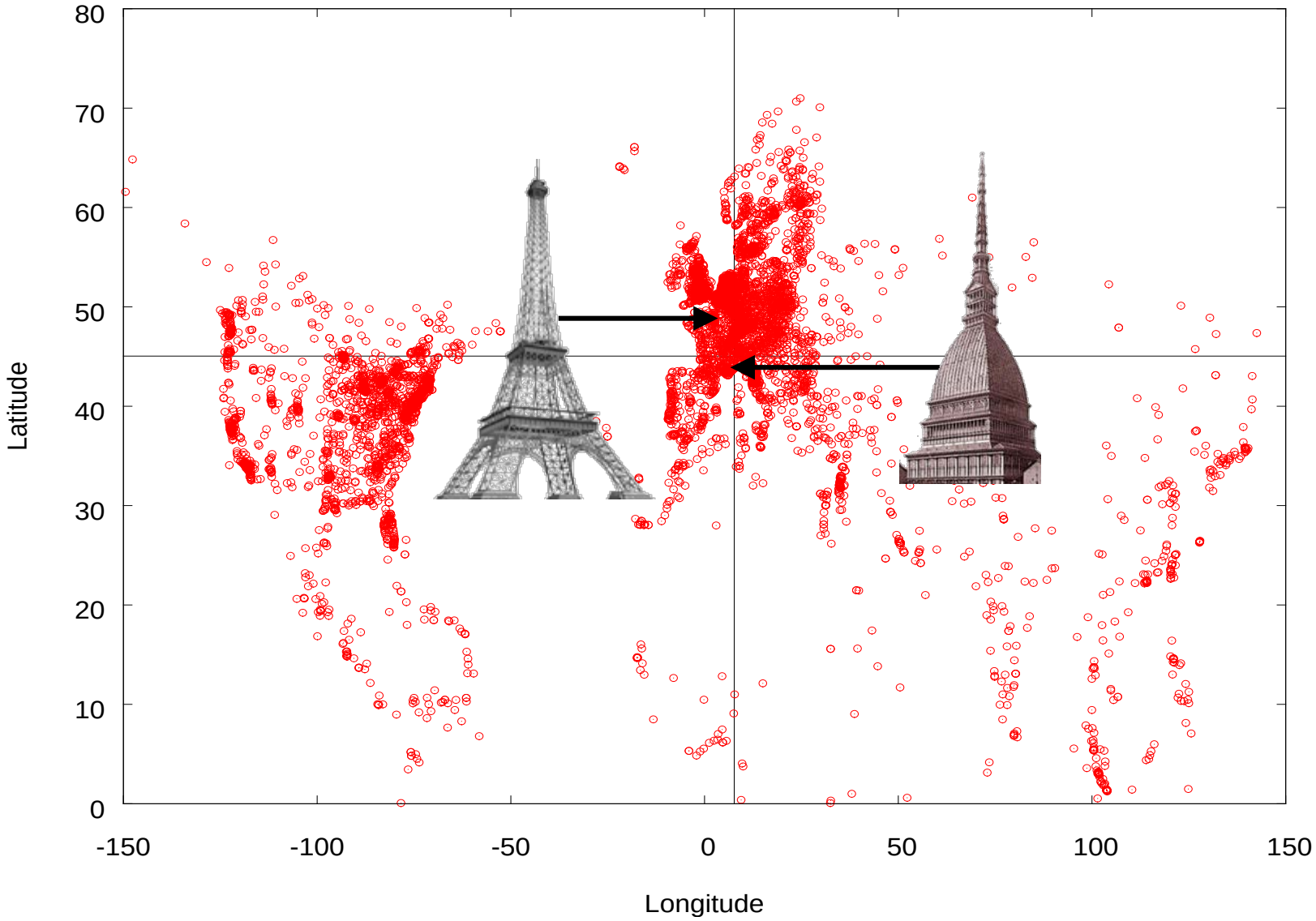


- **Geolocation breakdown**
  - Probes favor discovery of nearby hosts
  - Non-probes driven by social network


60
40
20
0
-20
-40
-60
Flow number
Mon
Tue
Wed
Thu
Fri
Sat
Sun
UDP E2E
UDP E2O
TCP E2E


CDF
[Min]
UDP E2E voice
UDP E2O voice
TCP E2E voice
UDP E2E video
1 - CDF
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
10
20
30
40
50
60
0.001
0.01
0.1
1
10
100
1000

Signaling Traffic: Peer Selection



Signaling Traffic: Inferring Churn