Inferring Interconnection congestion using Mlab NDT and Paris Traceroute data
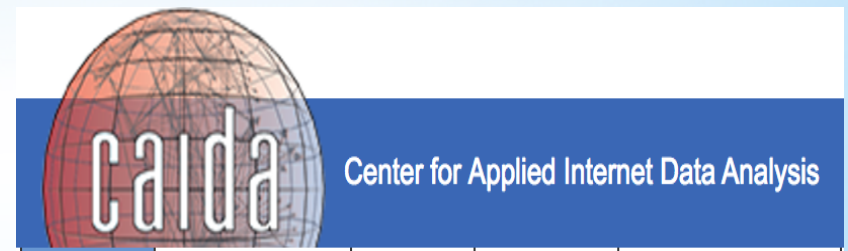
Xiaohong DENG

Yun FENG

Vijay Sivaraman

Amogh DHAMDHERE

Kc CLAFFY

UNSW SYDNEY

caida Center for Applied Internet Data Analysis

* Aims: Inferring inter-connection point congestions

* Challenges of using crowd sourcing measurement platforms

* Addressing Challenge 1

* Build Fine-grained network tomography techniques

* IMC'17 accepted

* Overview: This Talk

*Infer and localize congestion on end-to-end paths

*With particular focus on points of interconnections between ISPs

*Aims

* As of February 2017 the M-Lab infrastructure was able to measure between 0.3% and 9% of AS-level interconnections of access ISPs in US.

* Between 79% to 90% of AS-level interconnections traversed on paths from US ISPs to popular content providers were not testable using M-Lab's server infrastructure.

* Challenges:
Crowd source testing platform
AS link coverage

* 1.Fine-grained network tomography techniques not supported by existing throughput measurement platforms.

* 2. Existing measurement platforms do not provide sufficient visibility of paths to popular content sources, and only capture a small fraction of interconnections between ISPs.

* 3. Crowdsourcing measurements inherently risks sample bias: using measurements from volunteers across the Internet leads to uneven distribution of samples across time of day, access link speeds, and home network conditions.

* Challenges:
To be addressed -1

* Path information is not always available from large-scale throughput measurement platforms.

* M-Lab collects *paris-traceroutes* from M-Lab servers toward clients that run measurements against their infrastructure.

* Using it as input to a tomography algorithm is challenging due to issues with measurement synchronization and traceroute artifacts.

* One alternative is to use a simplified form of tomography at the AS-level; M-Lab's studies of interconnection congestion used this simplified approach.

* State of Art

* Assumption 1: There is no congestion internal to ASes.

* Assumption 2: The server and client ASes directly interconnected.

* Assumption 3: All router-level interconnections over which an inference is made for the AS interconnection behave similarly.

* **Assumptions**

of **Simplified AS-level tomography approach**
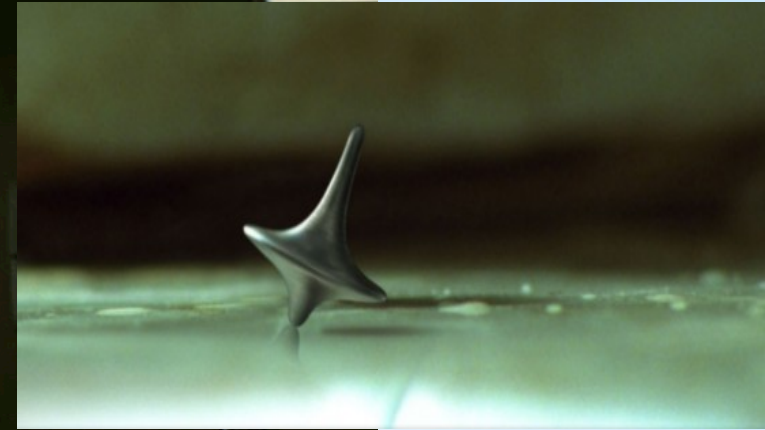
Your password's strength
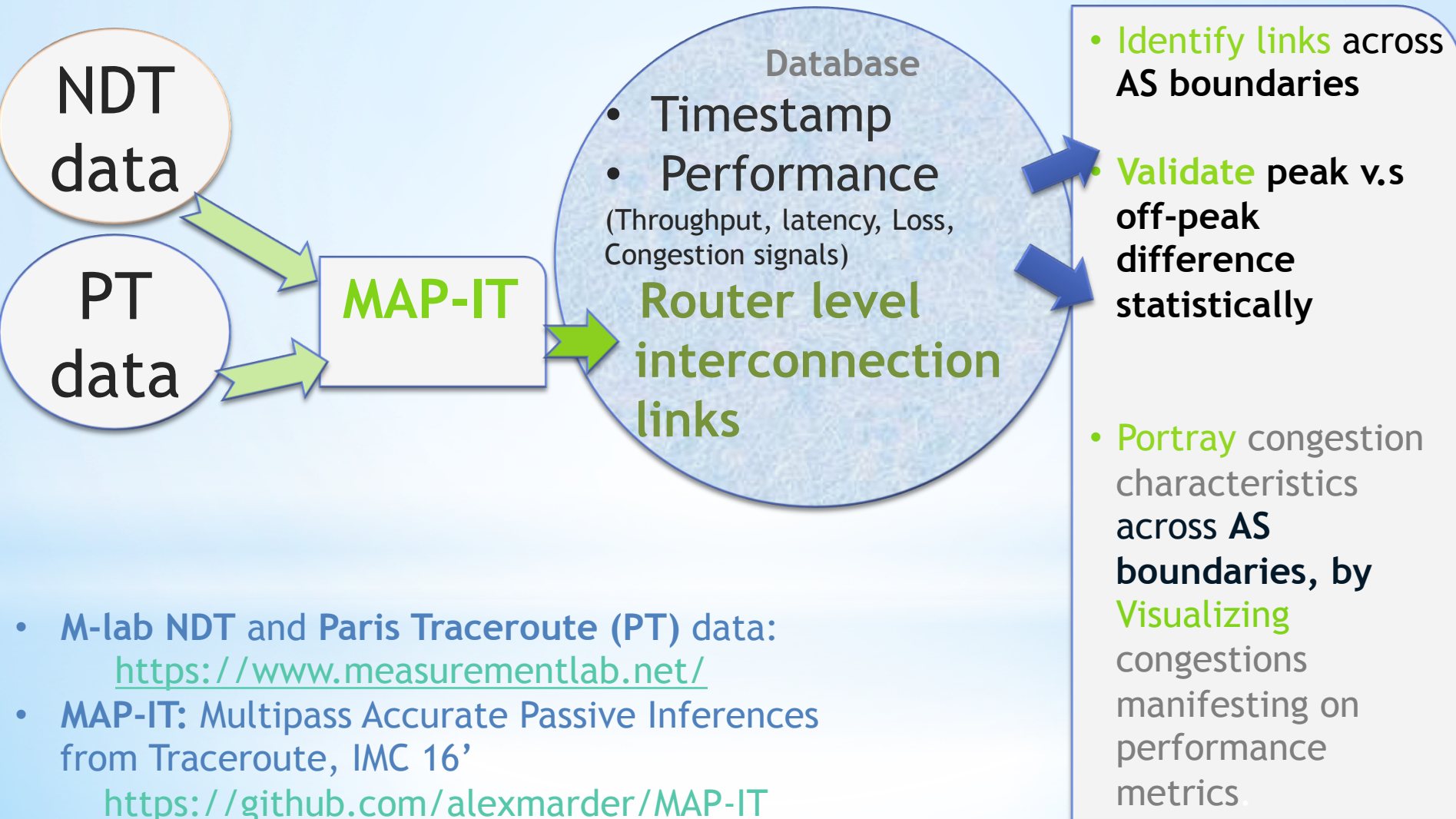
in your mind

in reality

*How valid are these assumptions?

* Lucid dream gives you direct access to your unconscious mind and its assumptions.

The dream explorers in Inception use special totems for their reality checks.

NDT data

PT data

MAP-IT

Database
- Timestamp
- Performance

(Throughput, latency, Loss, Congestion signals)

Router level interconnection links

*Totems

- **Identify links** across **AS boundaries**

- **Validate peak v.s off-peak difference statistically**

- **Portray** congestion characteristics across **AS boundaries, by Visualizing** congestions manifesting on performance metrics

- **M-lab NDT** and **Paris Traceroute (PT)** data:
  https://www.measurementlab.net/
- **MAP-IT:** Multipass Accurate Passive Inferences from Traceroute, IMC 16'
  https://github.com/alexmarder/MAP-IT

* Traces from all U.S. M-Lab servers to clients in 12 major U.S. ISPs listed in the Measuring Broadband America report.

* Match NDT tests traceroute from the server to that same client within a 10-minute window. Even with this wide window
  * May 2015 allowed us to match only 77% (572,564 out of 743,780 NDT tests) from clients to M- Lab servers (with both endpoints in the U.S.).

  * We found that in March 2017, we were able to match about the same fraction, 76% of NDT tests (4,689,239 out of 6,185,394) from U.S. M-Lab servers to U.S. clients.

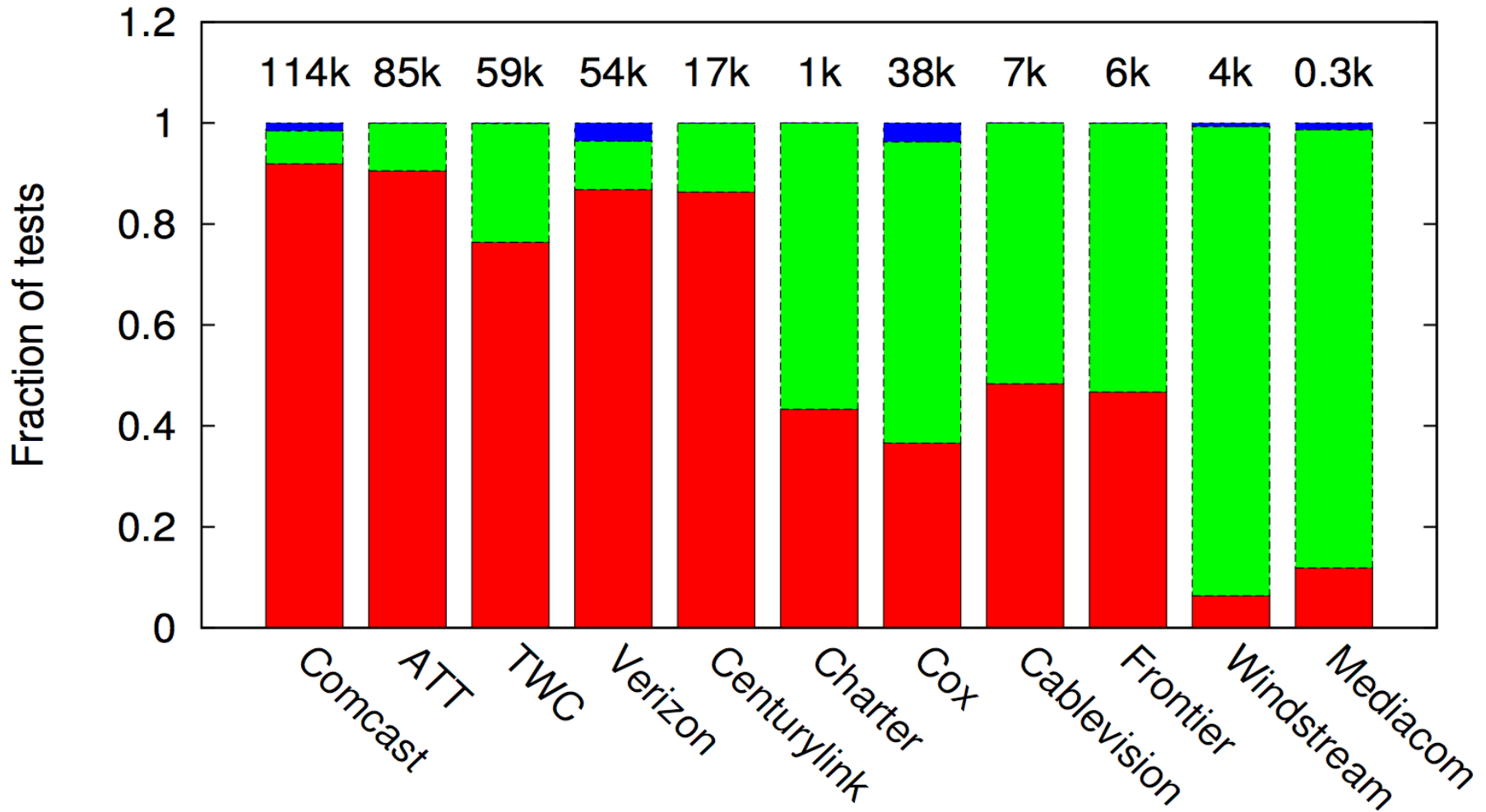* Step 1: Associate NDT with Traceroute

* In a transition between ASes A and B, the inter-domain link interfaces could be numbered out of either A or B's address space.

* Third party addresses that appear in traceroute that may confuse the identification of AS boundaries.

* Mistakes due to low visibility of certain interfaces in traceroute paths.

* 90% accuracy on the datasets tested.

* Address link inference challenges: MAP-IT algorithm

* May 2015, 894,408 interface adjacencies

* CAIDA's prefix-AS mapping derived from BGP routing tables from May 1-5, 2015, CAIDA's AS-Organization mapping from July 2015

* IXP prefixes obtained from peeringDB and PCH as input to MAP-IT.

* Sibling ASes as the same AS hop using information from CAIDA's AS-to- Organization dataset

* Step 2: Apply MAPIT to infer interconnection links

Legend: 1 hop (red), 2 hops (green), 2+ hops (blue)

ISP values (number of tests): Comcast 114k, ATT 85k, TWC 59k, Verizon 54k, Centurylink 17k, Charter 1k, Cox 38k, Cablevision 7k, Frontier 6k, Windstream 4k, Mediacom 0.3k

Y-axis: Fraction of tests
X-axis: ISP

*Reality check 1: Direct connection between server AS and client AS ?

*Care must be taken to ensure that the server and client AS are directly connected, using traceroutes and a technique to identify AS boundaries in traceroutes

*What we learned from 1st reality check?

* Does aggregating tests across multiple links ok?

* Aggregating tests across links is intuitively problematic if those links are in different geographical regions, as they could vary widely in terms of diurnal throughput patterns.

* The M-Lab service uses proximity-based server selection to try to ensure a client performs its measurement to the geographically closest M- Lab server.   -how WELL it is doing?
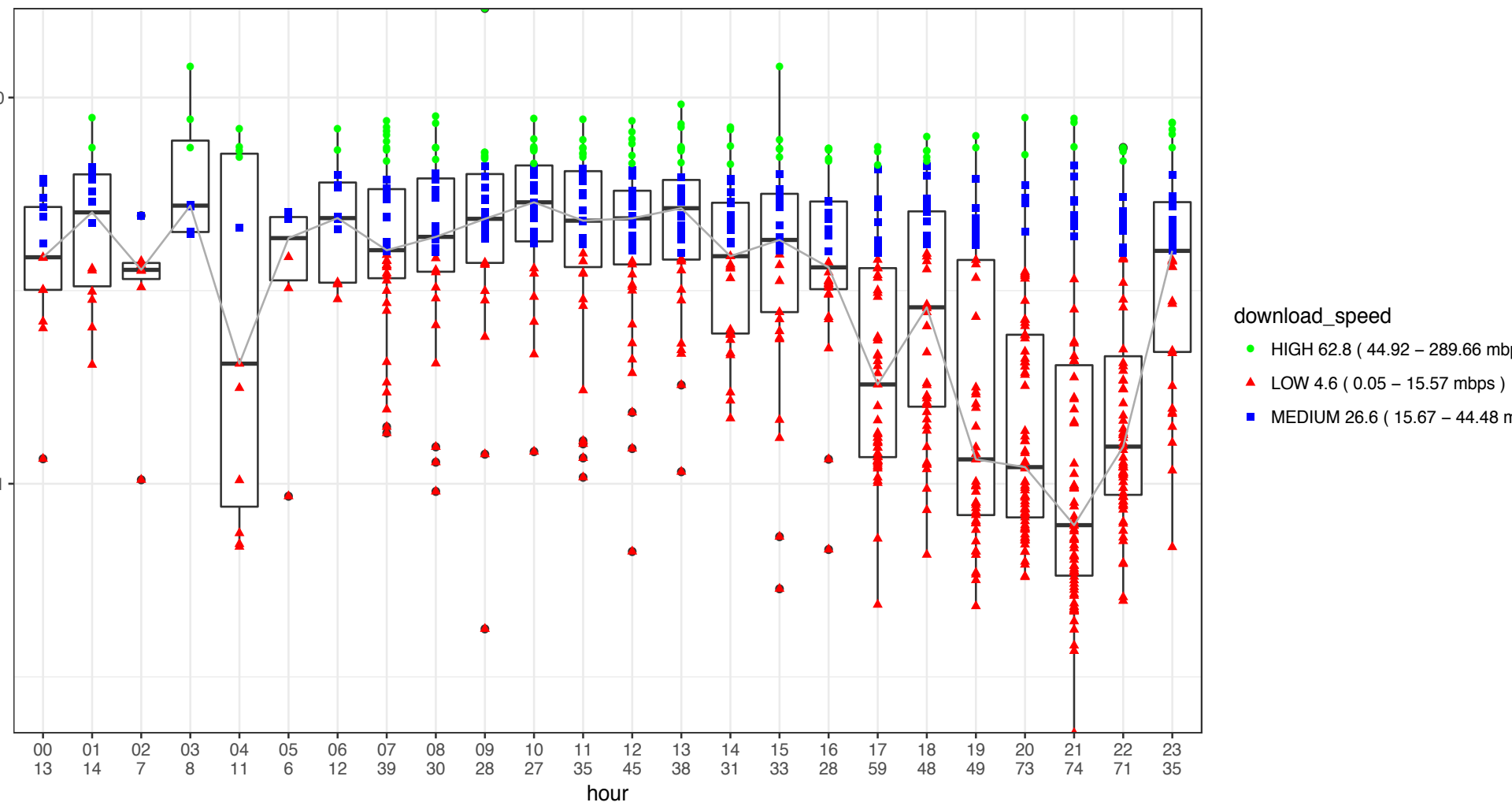
* Diversity of interconnection to access providers

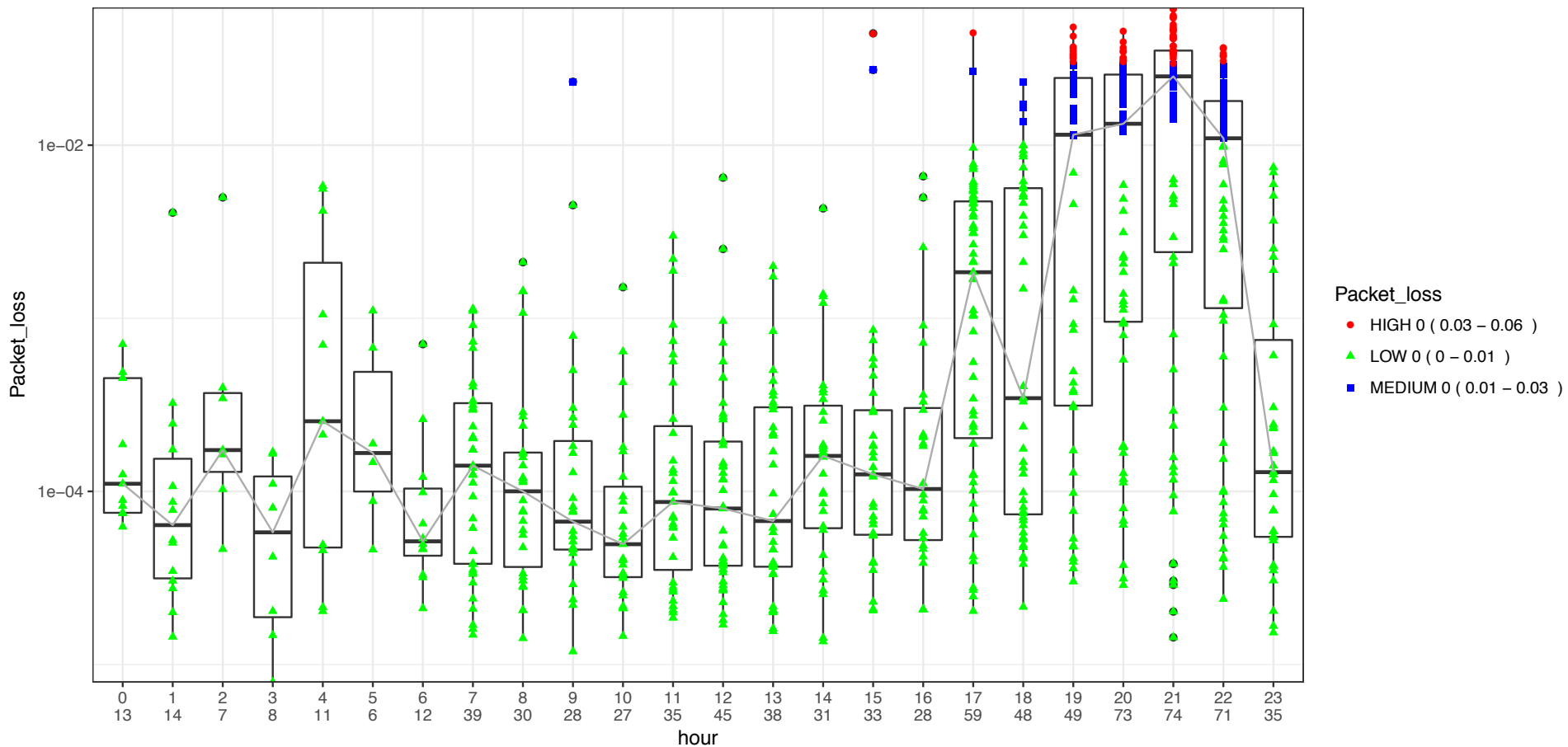| Client ISP (ASN)	# Links	# NDT tests per link
Comcast (AS7922)	2	1763,8
Comcast (AS7725)	1	1640
Comcast (AS22909)	1	1135
AT&T (AS7018)	14	2402,817,683,216,136 88,25,21,19,19,18,8,2,1
Verizon (AS701)	7	536,61,51,41,20,1,1
Verizon (AS6167)	2	3,3
Cox (AS22773)	40	total 816, max 376
Frontier (AS5650)	1	108
Centurylink (AS209)	4	382,38,17,1

Inter-domain links to top U.S. ISPs seen by M-Lab server atl01 (Level 3) in Atlanta (May 2015)

* Aggregating NDT throughput measurement results at an AS granularity

  * Masks the fact that different measurements could cross different IP-level links

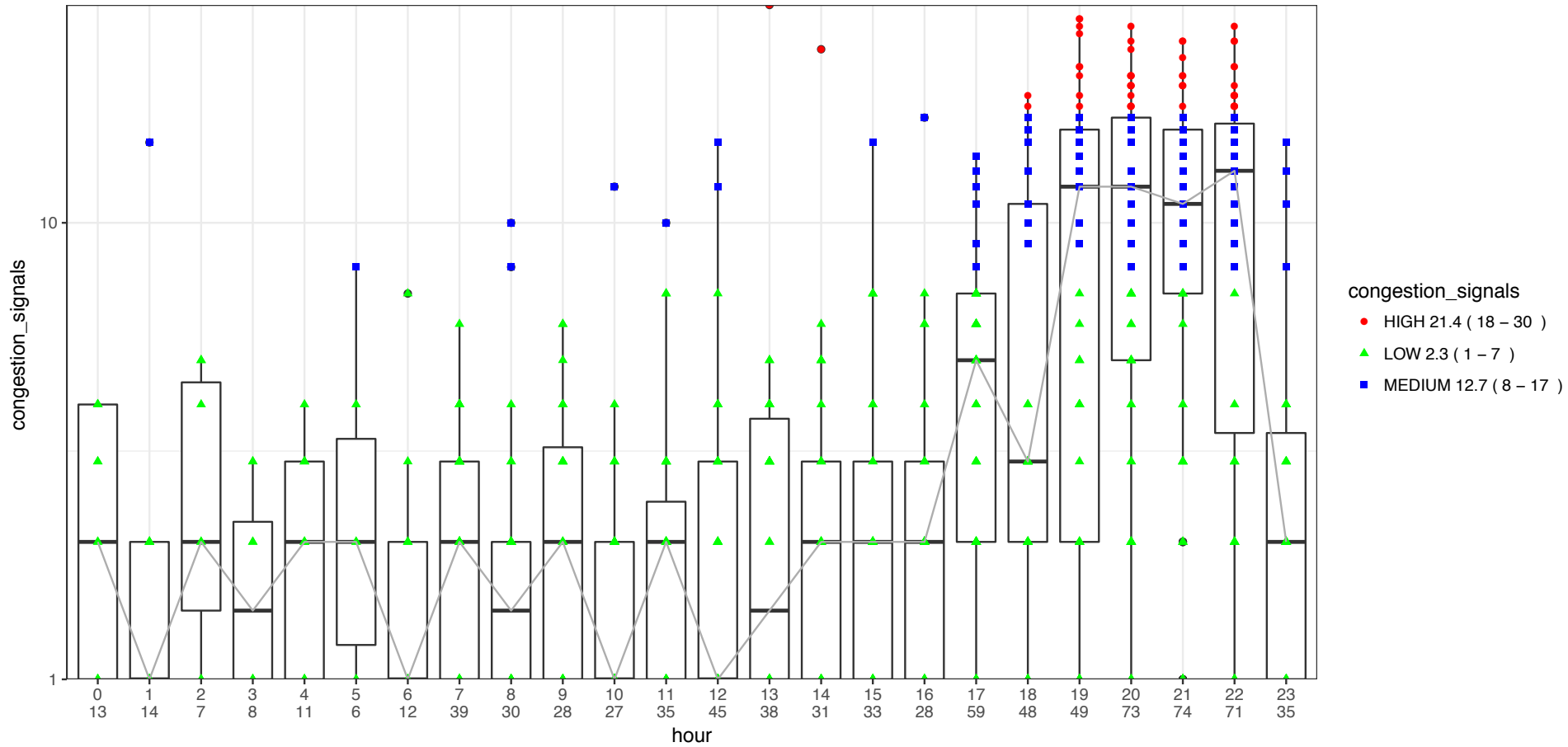 * Sometimes in different geographical regions that may have vastly different performance characteristics.
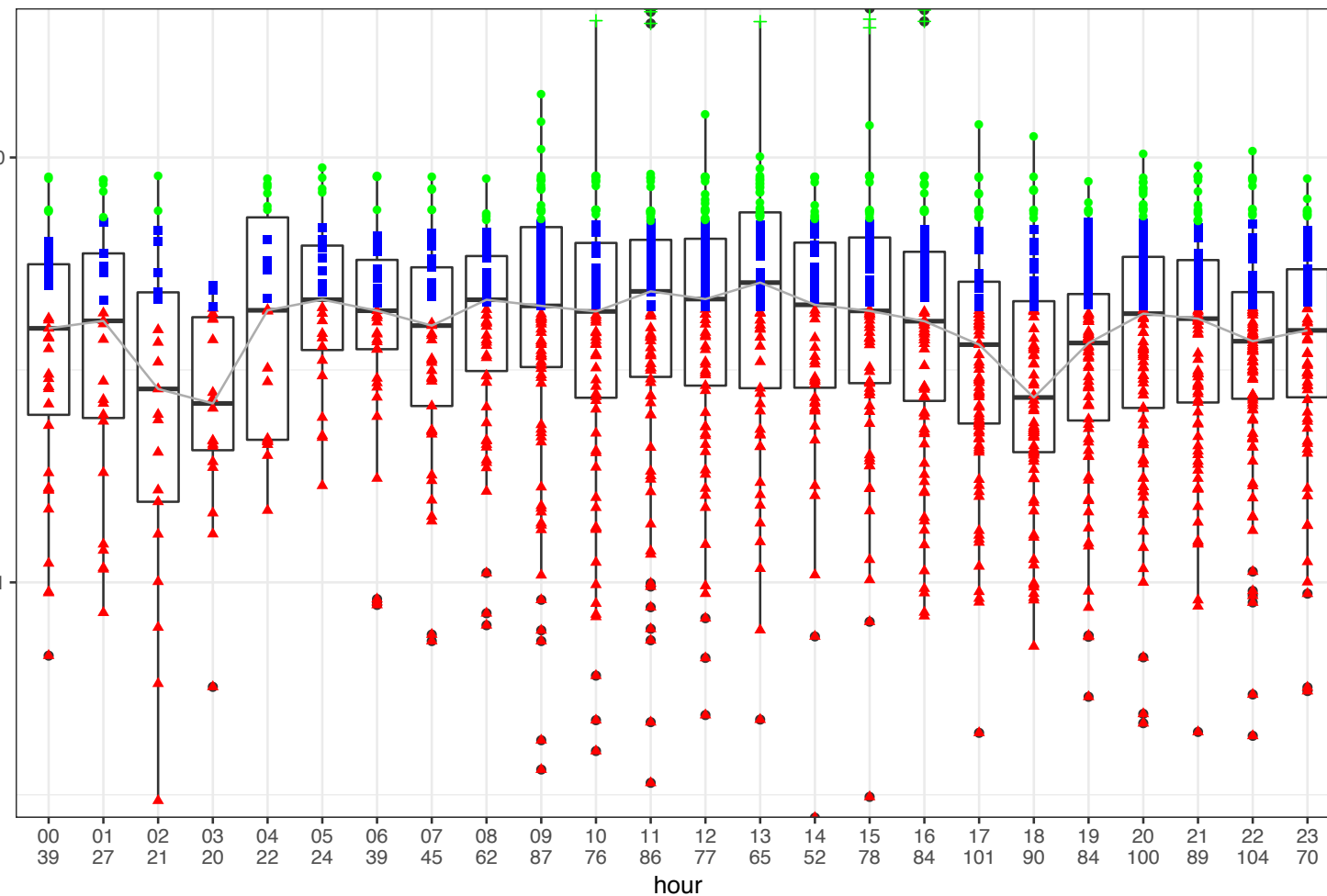
* **What we learned from 2nd Reality check?**

Speed in hour bin between verizon 701 206.82.141.134 and Tata (6453) .pdf

download_speed
- HIGH 62.8 (44.92 – 289.66 mbp
- LOW 4.6 (0.05 – 15.57 mbps)
- MEDIUM 26.6 (15.67 – 44.48 m

hour		
00	13	
01	14	
02	7	
03	8	
04	11	
05	6	
06	12	
07	39	
08	30	
09	28	
10	27	
11	35	
12	45	
13	38	
14	31	
15	33	
16	28	
17	59	
18	48	
19	49	
20	73	
21	74	
22	71	
23	35	

*Portray Link Characteristics:
Throughput - Link A of Verizon 701- Tata (6453)

Packet_loss in hourly bin between verizon 701 206.82.141.134 and Tata (6453) .pdf

*Portray Link Characteristics:
Loss- Link A of Verizon 701- Tata (6453)

Congestions in hourly bin between verizon   701   206.82.141.134 and Tata   (6453) .pdf

*Portray Link Characteristics:
Congestions- Link A of Verizon 701- Tata (6453)

Speed in hour bin between verizon 701 64.86.252.98 and Tata (6453) .pdf

*Portray Link Characteristics:
Throughput - Link B of Verizon 701- Tata (6453)

Packet_loss in hourly bin between verizon 701 64.86.252.98 and Tata (6453) .pdf
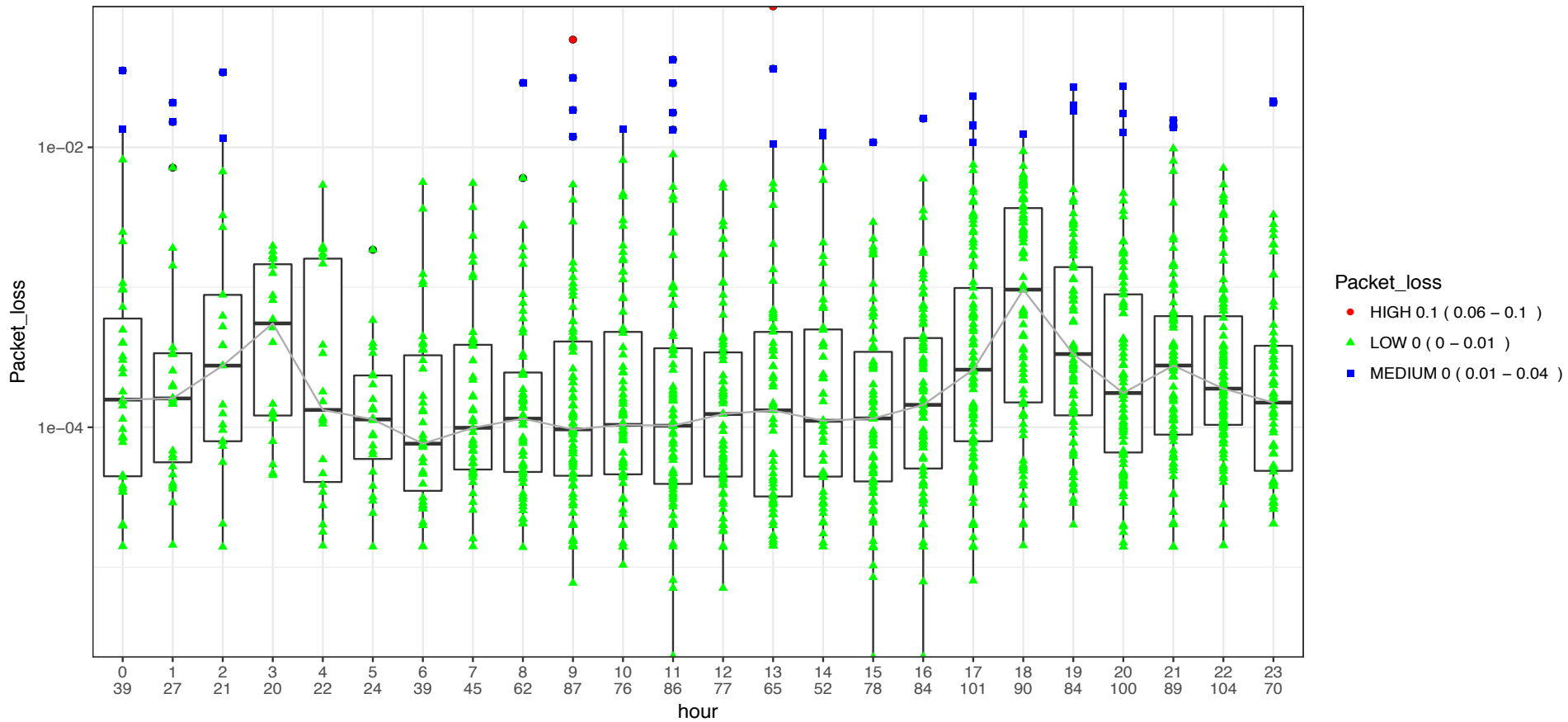
*Portray Link Characteristics:
Loss - Link B of Verizon 701- Tata (6453)

Congestions in hourly bin between verizon 701 64.86.252.98 and Tata (6453) .pdf
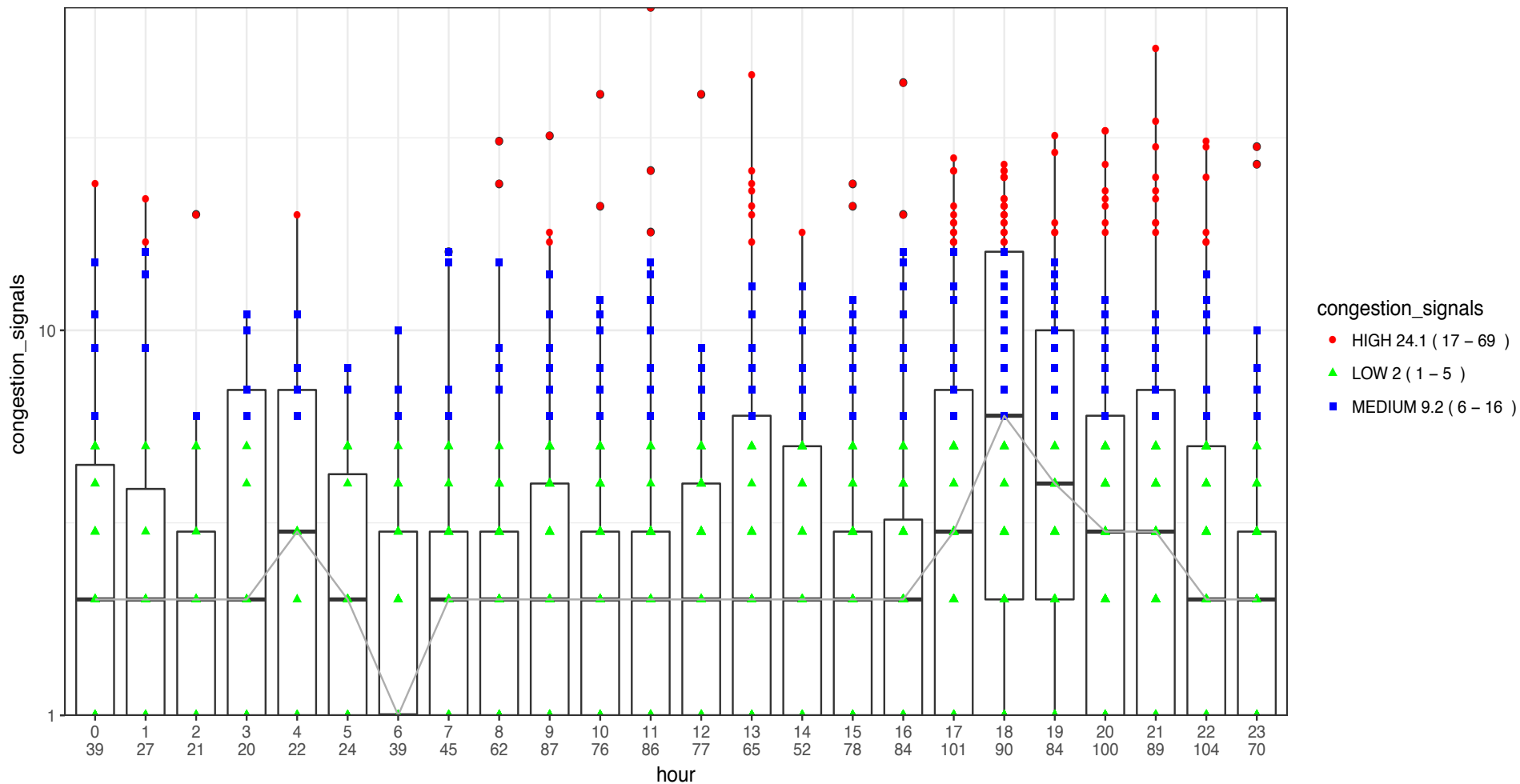
*Portray Link Characteristics:
Congestions- Link B of Verizon 701- Tata (6453)

*Portrayals of Two links belong to same owners do show **different** Characteristics

* Aims: Inferring inter-connection point congestions

* Addressing Challenge 1

* Build Fine-grained network tomography techniques

   * Reality Check 1: Tests samples are not always represent the case when server and client AS are directly connected
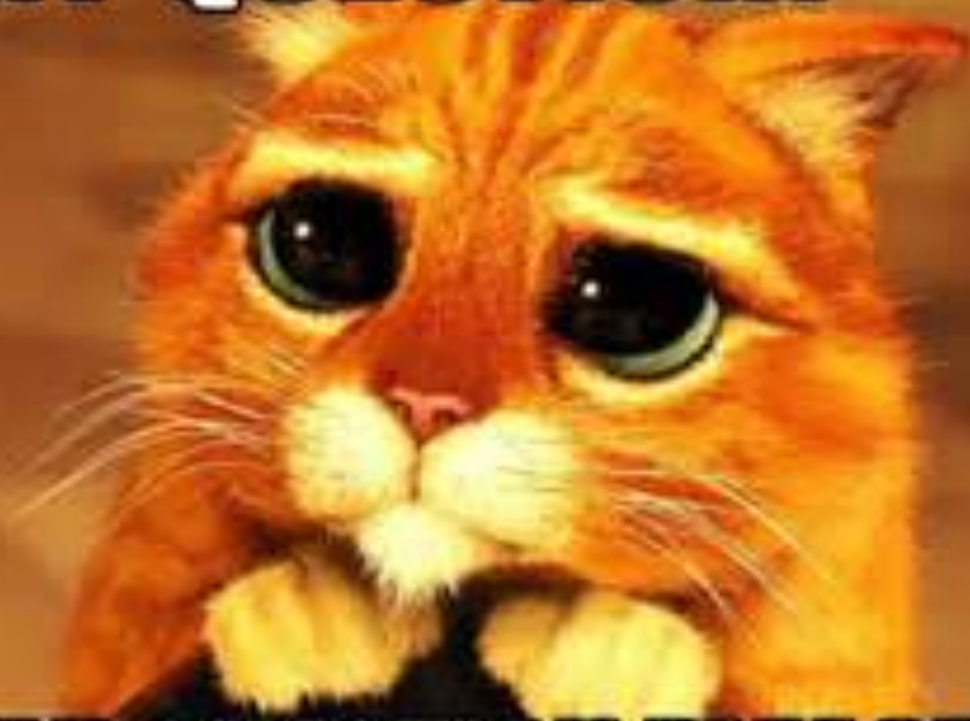   * Reality Check 2: Two links belong to same owners do show different congestion Characteristics

*Recap: This Talk

* Propose:
   * De-aggregating AS level to inter-domain router-level link granularity
   * Portray router-level link congestion events and signatures