



Telefonica

Impact of Carrier-Grade NAT on Web Browsing

- Enrico Bocchi
- Ali Safari Khatouni
- Stefano Traverso
- Alessandro Finamore
- Valeria Di Gennaro
- Marco Mellia
- Maurizio Munafò
- *Dario Rossi*



RIPE71 - Bucharest

Background

2

In February 2011, Internet Assigned Numbers Authority (IANA) ran out of IPv4 /8 unallocated networks

- After exhaustion, IANA and RIRs allowed IPv4 transfers
- Transfer policies are based on **economics** and **trading**
- IPv4 are now subject of a **growing market**
 - APNIC market +220%, RIPE market +600% in 2014^[1]
 - **10.50\$/IP** in a /24 block (7.75\$/IP in a /20 block)
 - ARIN has 12 *Registered Transfers Facilitators*
<http://www.ipv4auctions.com>, <http://www.iptrading.com/>, ...



ISPs and organizations running large IP networks are committed to huge economical investments for addresses

[1] <http://www.ipaddressnews.com/2014/04/07/343>

Background

What about IPv6 ?

- Requires a significant investment of resources (hw/sw, training, ...)
- Poses incompatibility issues with IPv4 (dual stack networks, tunnels)
- Deployment is still lagging [2]: ~6% of users access Google over IPv6

Hotelling Rule

The transition from an exhaustible resource to a new one will not occur until the price of the current resource exceeds the cost of its replacement

Workaround



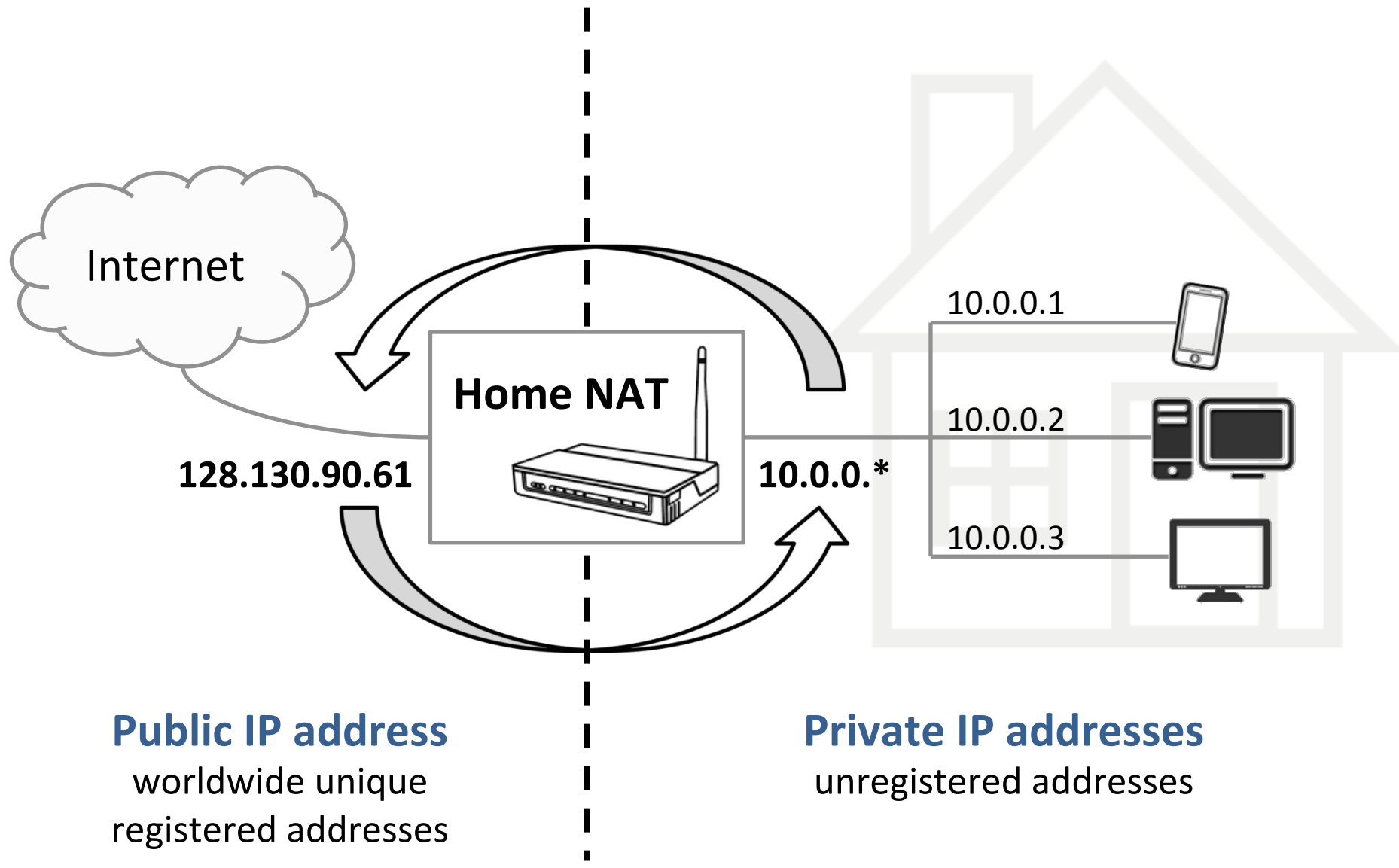
ISPs are deploying **Carrier-Grade NAT**

- Reserved pool of addresses 100.64.0.0/10 - IETF RFC 6598
- Implemented through NAT444

[2] <http://www.worldipv6launch.org/measurements/>
<https://labs.ripe.net/Members/gih/counting-ipv6-in-the-dns>

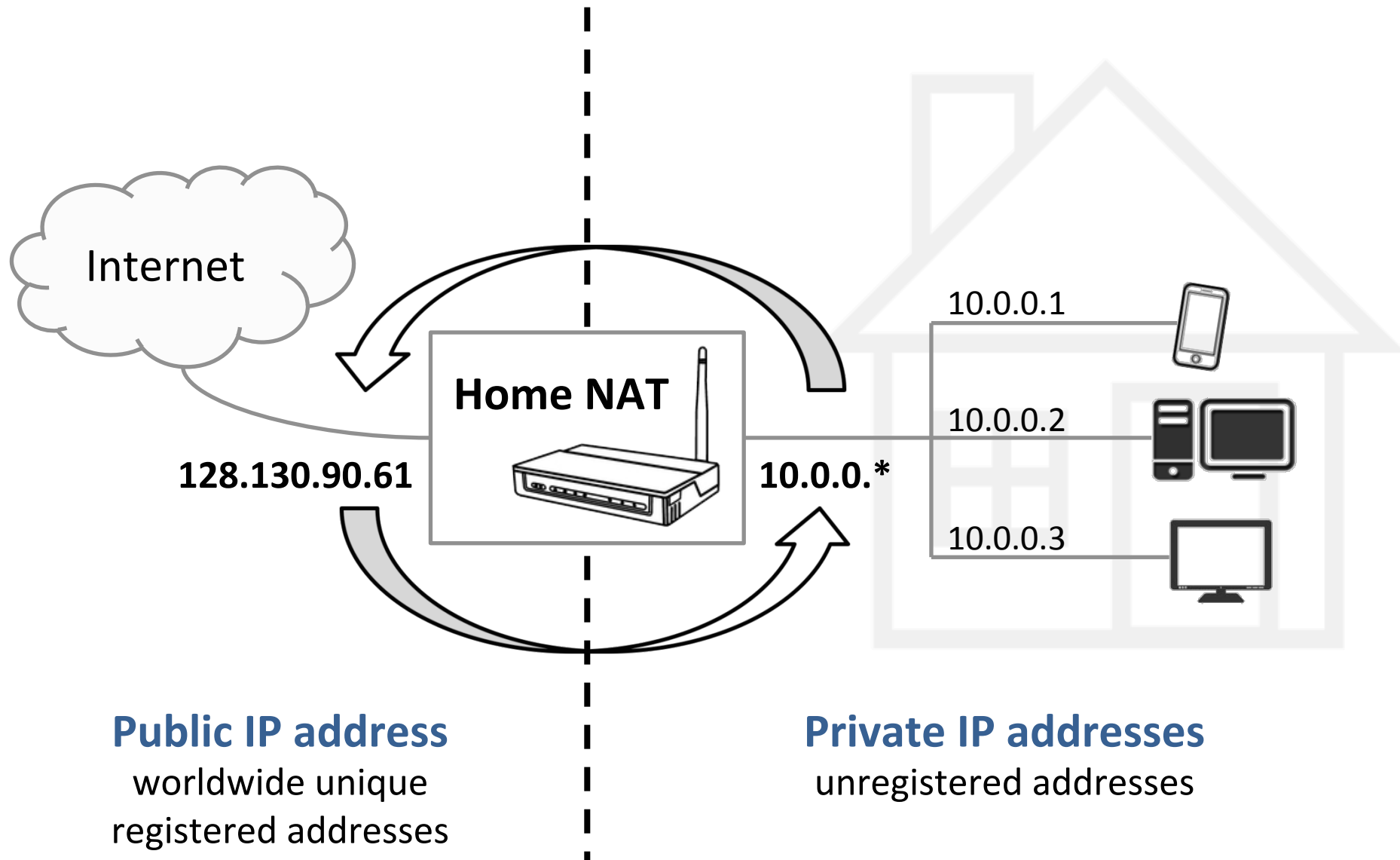
NAT at a glance - Traditional NAT44

4



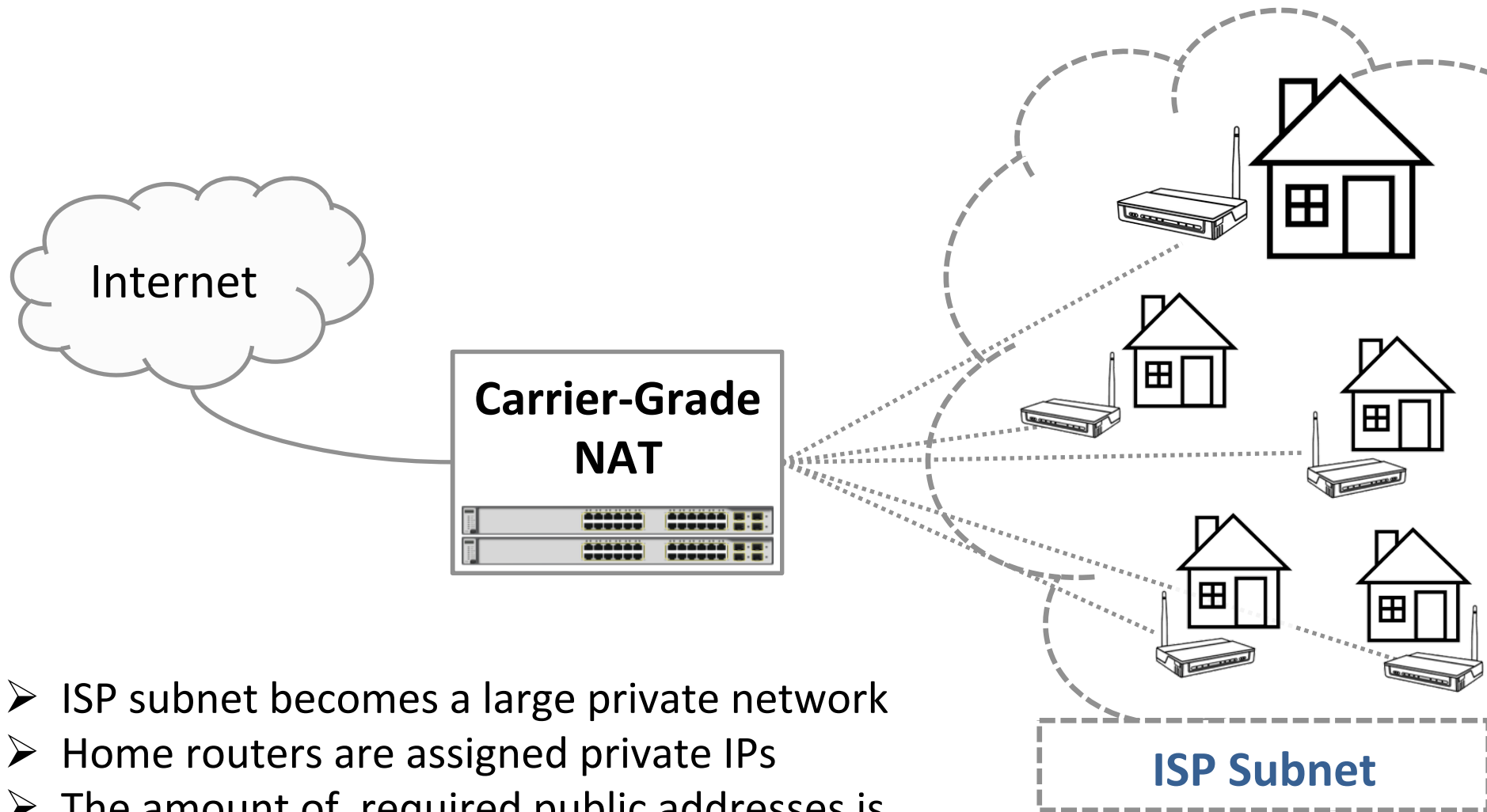
NAT at a glance - Traditional NAT44

5



...and Carrier-Grade NAT – NAT444

6



- ISP subnet becomes a large private network
- Home routers are assigned private IPs
- The amount of required public addresses is reduced as well as cost

Carrier-Grade NAT (CGN)

7

The deployment of CGN has some implications:

- Breaks the end-to-end IP connectivity
- Introduces reachability problems for NAT-ted devices
 - Need of successful NAT traversal techniques
 - Updates of non-NAT friendly applications
- Mandates the network keeps the state of the connections
- Impacts negatively lawful intercept
- **May have performance implications**

Our investigation goals

- ▶ Does CGN impact users' browsing experience?
- ▶ For users, is there any benefit in having a public IP?
- ▶ For ISPs, how many IPs would CGN let them save?

Answer with measurements

Methodology

Methodology roadmap

Large scale passive measurement

- A real ISP deployment
- Customers are offered public or private address
- Traffic monitored to extract performance metrics

Leverage statistical tools

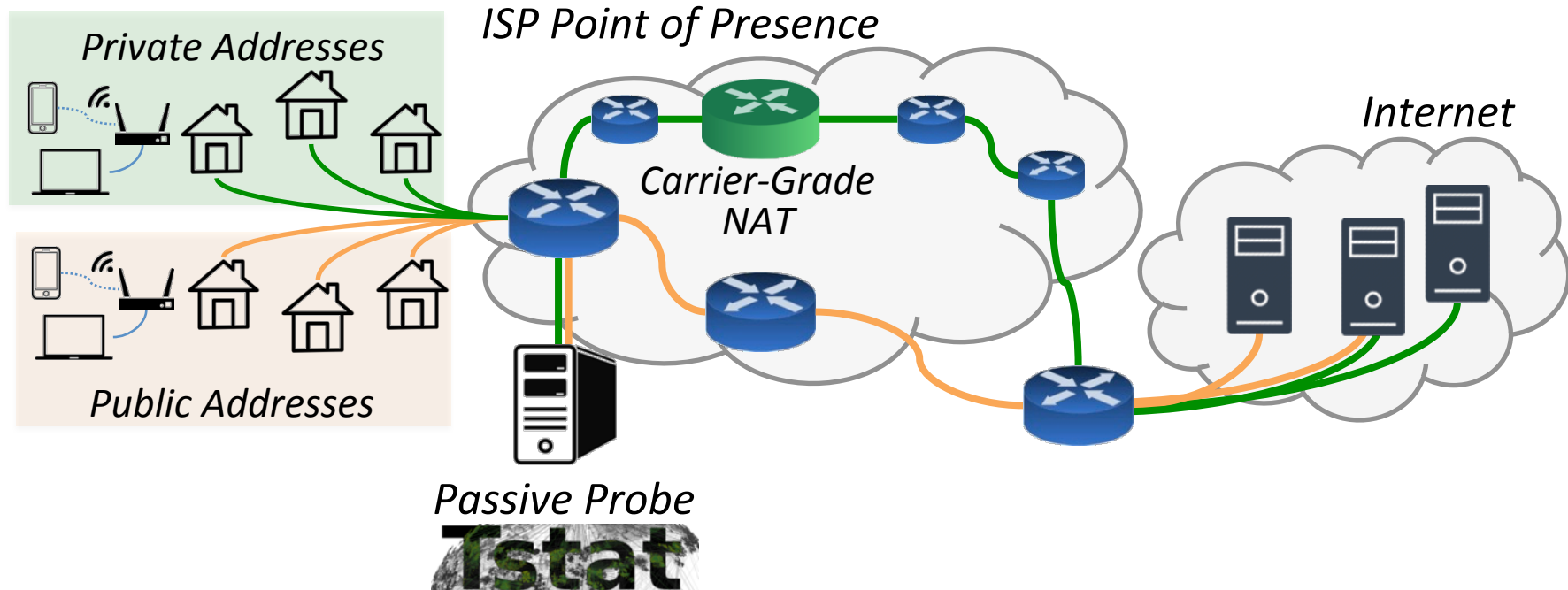
- Collect and compare empirical probability distributions
- Check and quantify eventual differences

Focus on

- Web traffic and performance

Monitoring Scenario

10



Dataset

- **1 month** of real traffic recorded, October 2014
- **17,000 household** monitored, **residential customers**
- 1.7Billion TCP flows, 0.7Billion HTTP requests

<http://tstat.polito.it>

Monitoring Scenario

11

Client



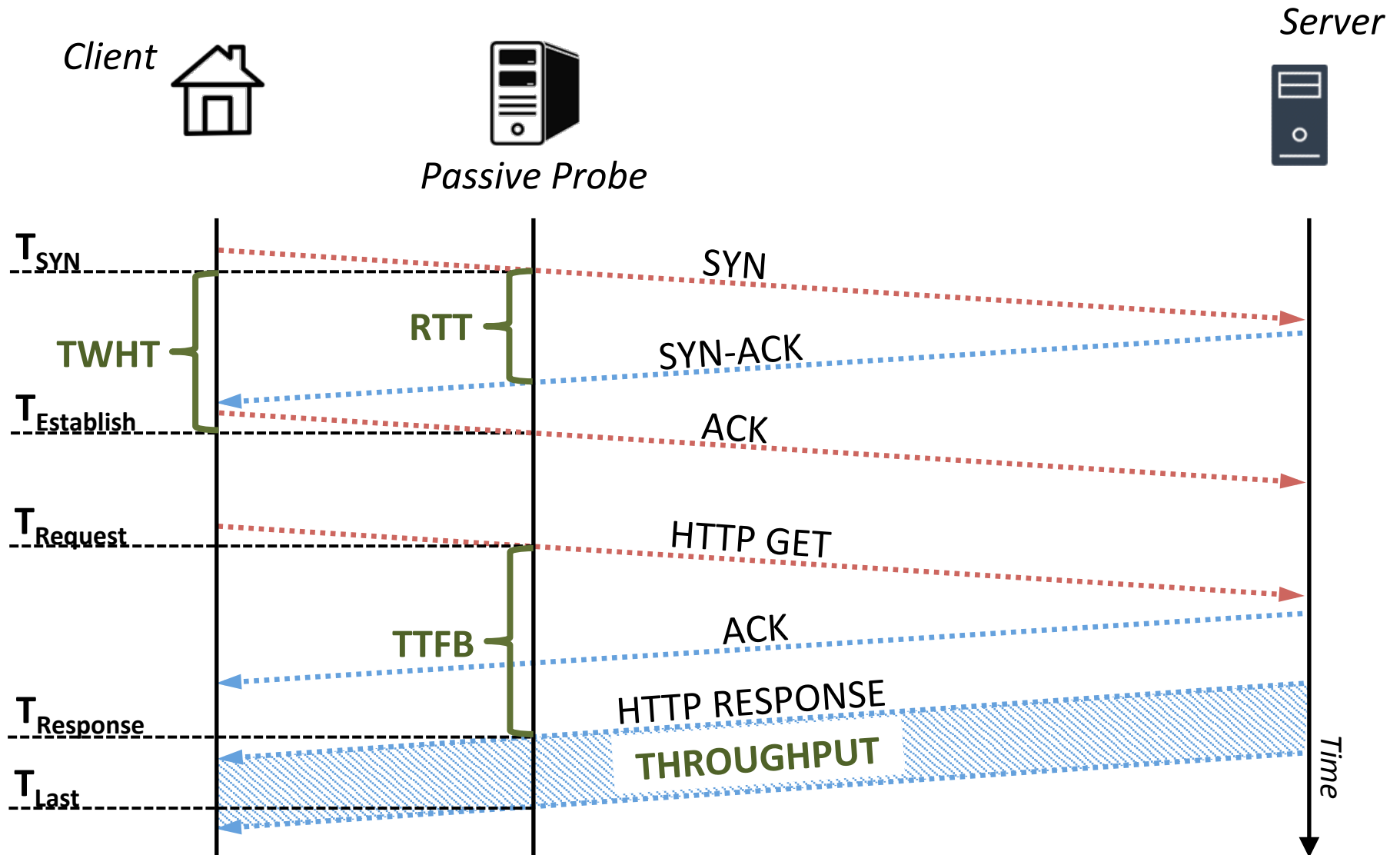
Passive Probe

Server



Monitoring Scenario

12



Assessing the Impact of CGN

13

- Consider **9 performance metrics**
- **Measure** distinct probability distributions for each metric
 - Coupled distributions for clients with private or public address

Jensen-Shannon Divergence

- **Quantify the difference** between a pair of probability distributions
- Based on the Kullback-Leibler divergence
 - + Symmetric
 - + Bounded to finite value $[0, \ln(2)]$

$$JS\text{-}div = \sum_i \left\{ \frac{1}{2} p_i \ln \left(\frac{p_i}{\frac{1}{2} p_i + \frac{1}{2} q_i} \right) + \frac{1}{2} q_i \ln \left(\frac{q_i}{\frac{1}{2} p_i + \frac{1}{2} q_i} \right) \right\}$$

with p_i and q_i being the samples of the two distributions to compare

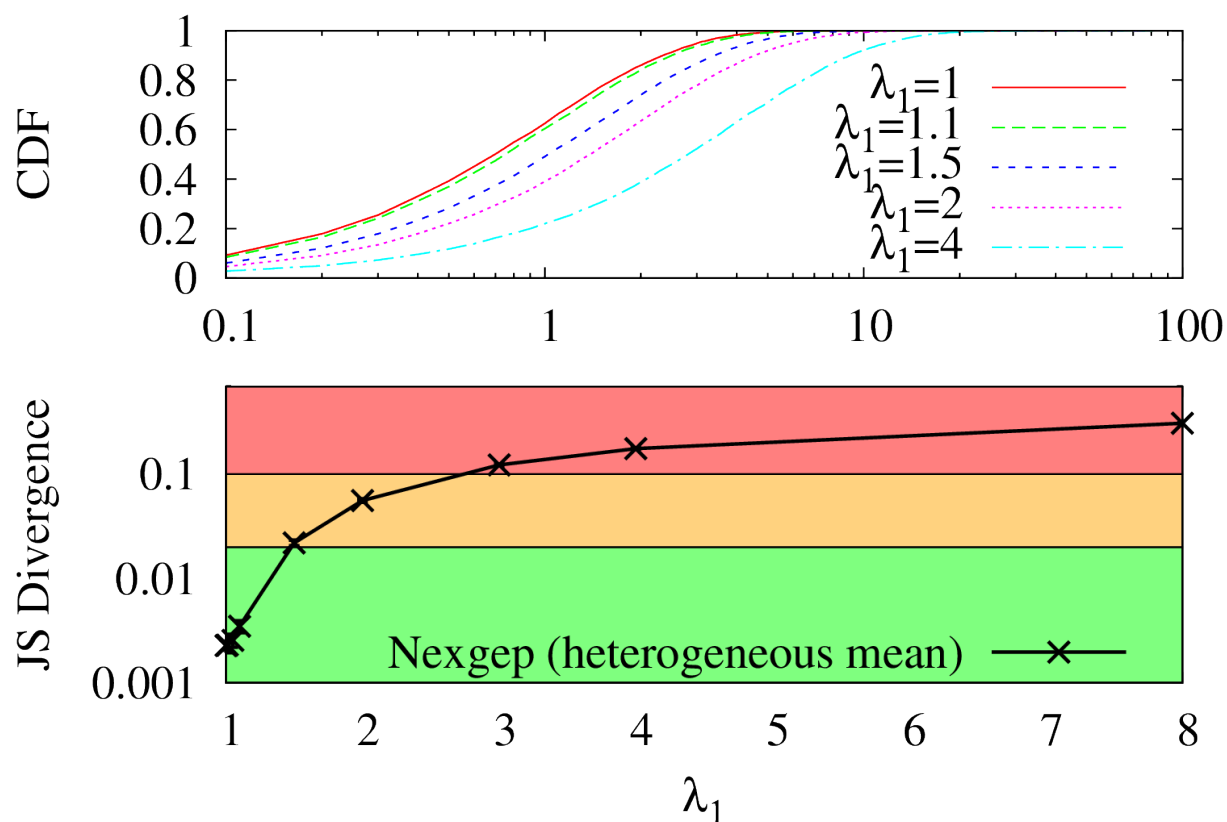
Jensen-Shannon Calibration

14

Need of a **threshold** to discriminate between **significant** and **negligible** differences

Example calibration:

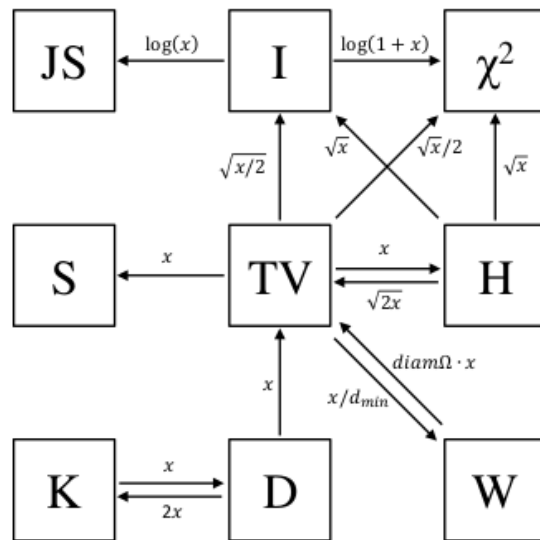
- Negexp CDF
- $\lambda_0 = 1$, fixed
- λ_1 varies $[1 \div 8]$



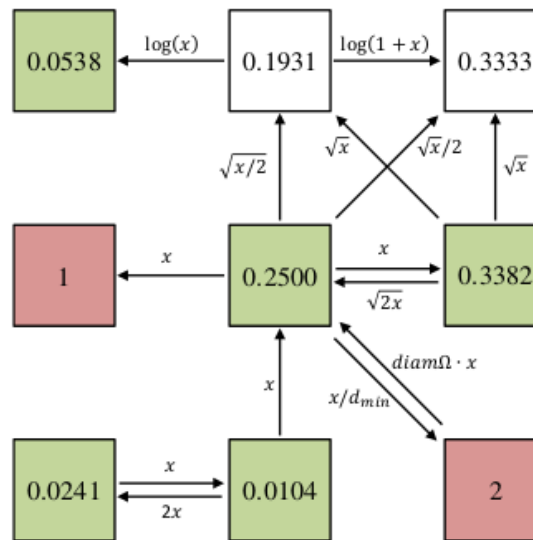
Divergence metrics, reloaded

15

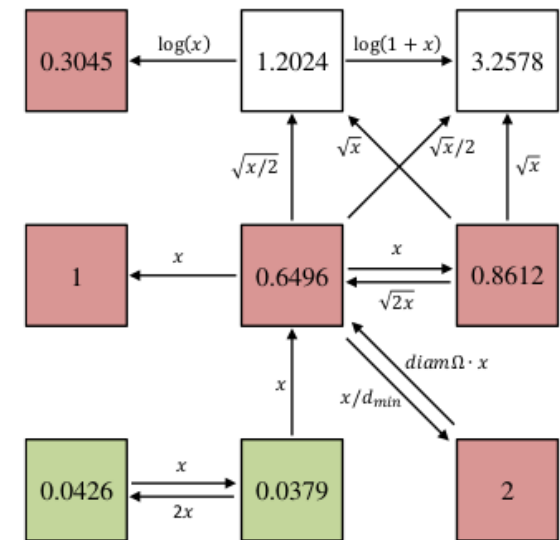
- JS just one of several possibilities
- Total Variation (TV) or Hellinger (H) also equivalent



(a) Statistical distance measures and dependency relationships.



(b) Computed distance values with $\lambda_1=2$. Notice Separation and Wasserstein reaching the upper bound.



(c) Computed distance values with $\lambda_1=8$. Notice Kolmogorov and Discrepancy being non-responsive.

Figure .16: Distance measures overview and computed values for negative exponential distributions.

Performance Analysis

Does CGN impact users' browsing experience?

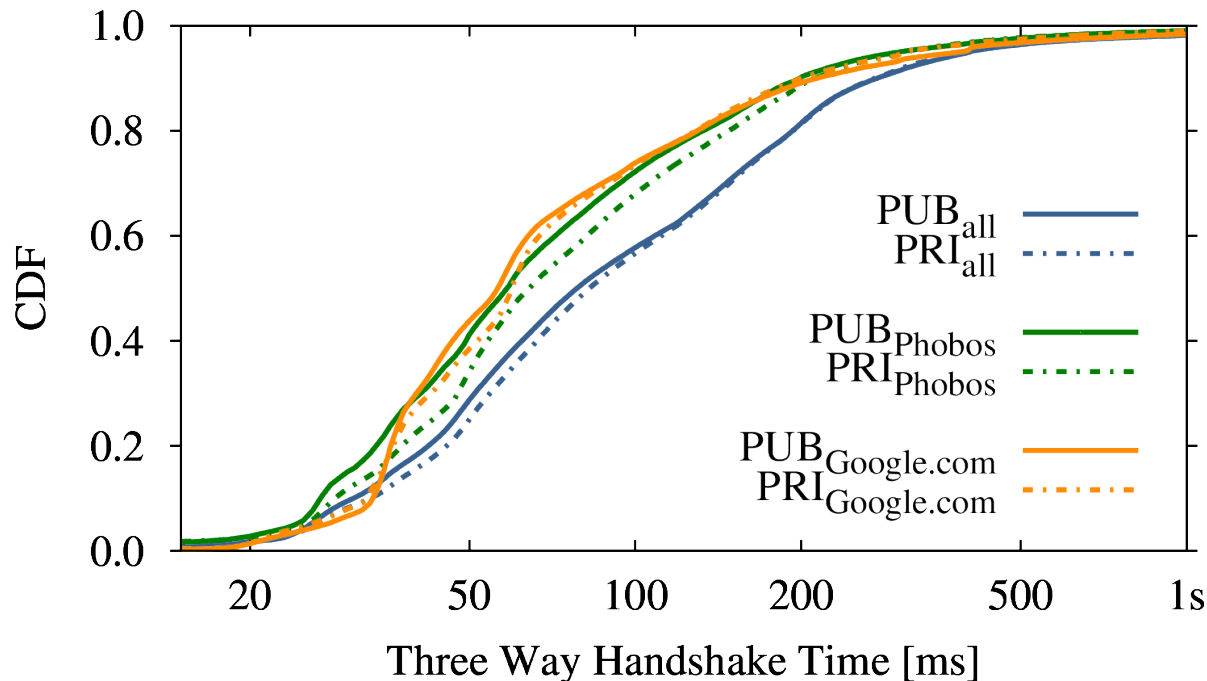
Performance Metrics – TWHT

17

1. Three Way Handshake Time (TWHT)

- Any remote server (*all*)
- iTunes contents (*phobos.apple.com - Akamai*)
- Google Search (*Google.com*)

Service	JS Div
<i>All</i>	0.002
<i>Phobos</i>	0.016
<i>Google.com</i>	0.010



Performance Metrics – Throughput

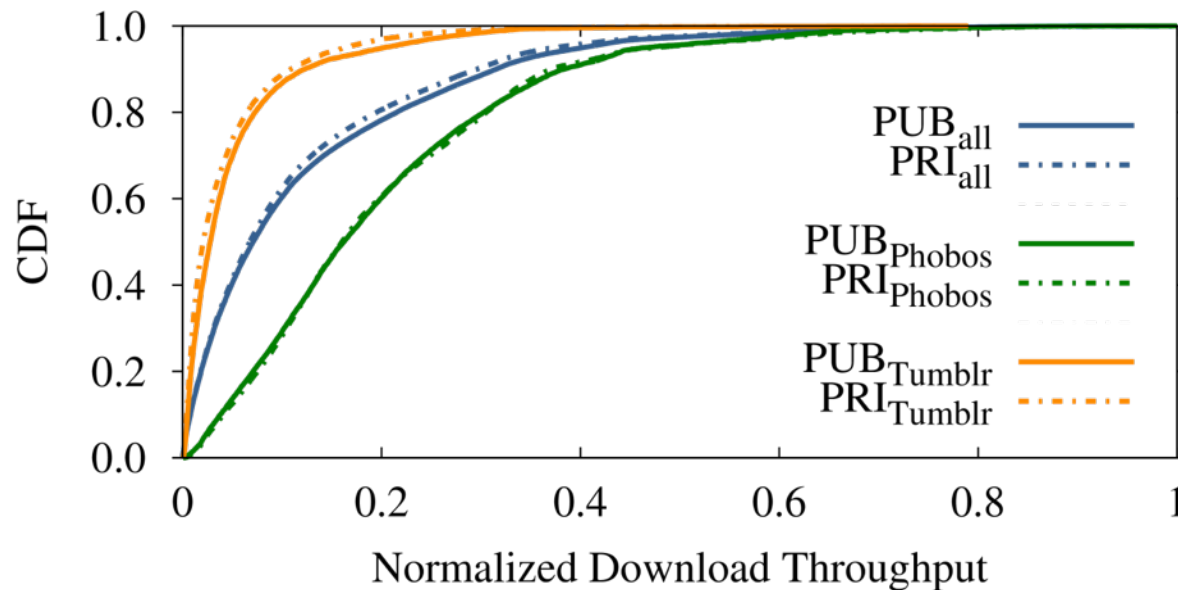
18

1. Three Way Handshake Time (TWHT)

2. Download Throughput

- Any remote server (*all*)
- iTunes contents (*phobos.apple.com - Akamai*)
- Tumblr Blogging Platform (*Tumblr.com*)

Service	JS Div
<i>All</i>	0.001
<i>Phobos</i>	0.022
<i>Tumblr</i>	0.021



Performance Metrics – Number of Hops

19

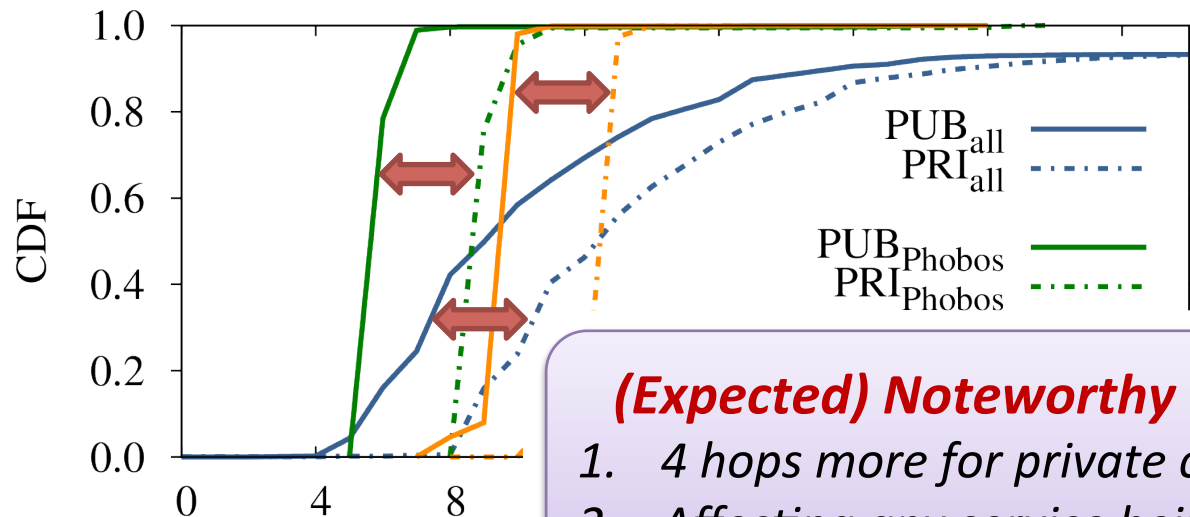
1. Three Way Handshake Time (TWHT)

2. Download Throughput

3. Number of Hops

- Any remote server (*all*)
- iTunes contents (*phobos.apple.com - Akamai*)
- Google Search (*Google.com*)

Service	JS Div
<i>All</i>	0.223
<i>Phobos</i>	0.689
<i>Google.com</i>	0.666



(Expected) Noteworthy differences

1. 4 hops more for private customers
2. Affecting any service being contacted

Jensen-Shannon Results

20

Three intervals identified

- **Significant** differences $JS_{div} \geq 0.1$
- **Noticeable** differences $0.02 \leq JS_{div} < 0.1$
- **Negligible** differences $JS_{div} < 0.02$

Metric	Any Server	Google.com	Phobos.com
Number of Hops	0.223	0.666	0.689
Latency (RTT)	0.001	0.006	0.007
Establish			
HTT			
Thr			
Nu			
Out			
Dup			

Our investigation goal

- ▶ Does CGN impact users' browsing experience?

We observe no significant impact (for these KPI)

- ▶ Is there any benefit in having a private IP?

Benefits of having a Public / Private IP address

For users, is there any benefit in having a public IP?

Active Servers

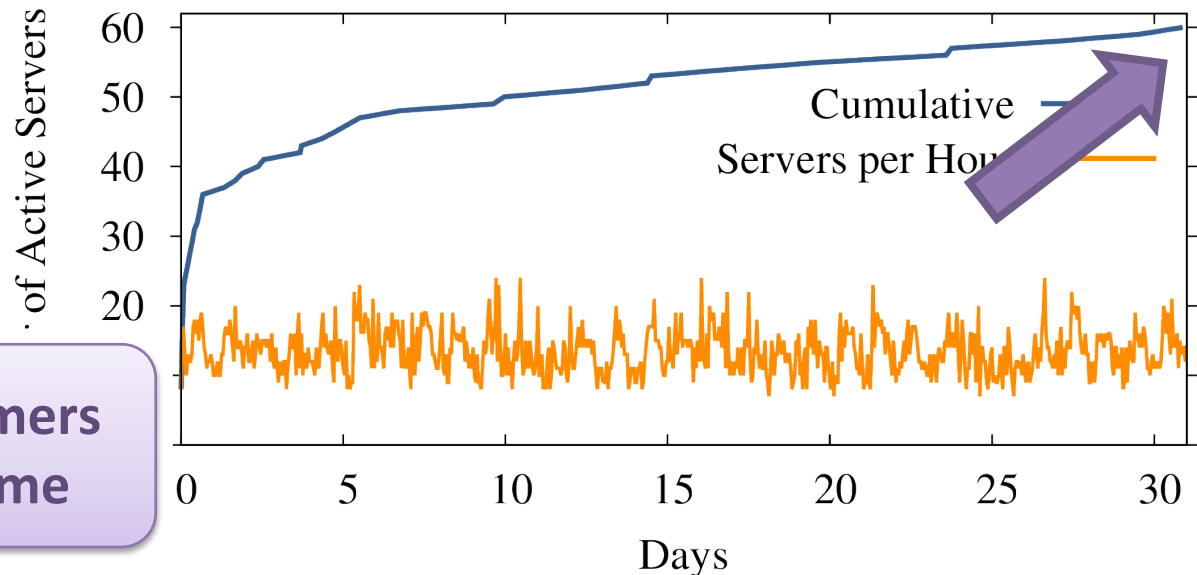
22

Does the customer **need IPv4 reachability**?

- Is there any ISP customer running a server at home?

Detection technique

- Look for customers answering at least one incoming connection
 - Protocols: HTTP(S), IMAP(S), POP(S), SMTP(S)
 - No P2P



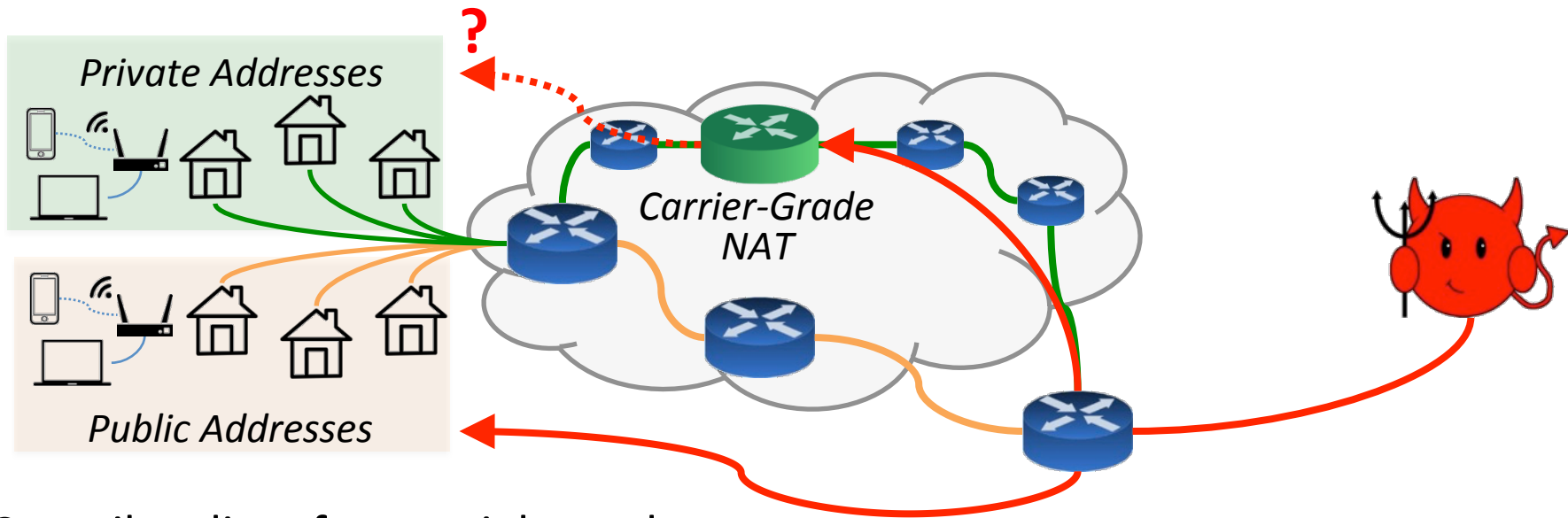
**Only 0.6% of customers
runs servers at home**

Unsolicited Traffic

23

What about **unsolicited traffic**?

- How many home routers are victims of port-/net- scans?



- Compile a list of potential attackers
 - Remote hosts making unsuccessful TCP connection attempts to more than 50 IPs in the PoP,
 - Private addresses only reachable by attackers within the ISP network
- Focus on destination ports with well-known services or vulnerabilities

Unsolicited Traffic

24

Destination Port	Description	Percentage of victims in PoP	
		PRI	PUB
80	HTTP	1.8	78.5
443	HTTP Secure (HTTPS)	0.1	78.9
143	Internet Message Access Protocol (IMAP)	<0.1	79.3
995	Post Office Protocol (POP3 over SSL)	<0.1	79.2
25	Simple Mail Transfer Protocol (SMTP)	0.1	79.0
22			

Our investigation goals

- ▶ Does CGN impact users' browsing experience?
- ▶ Is there any benefit in having a private IP?

0.6% of customers needs IPv4 reachability

Public IPs are up to 800x more likely to be victim of attacks

CG-NAT Dimensioning and Saving Estimation

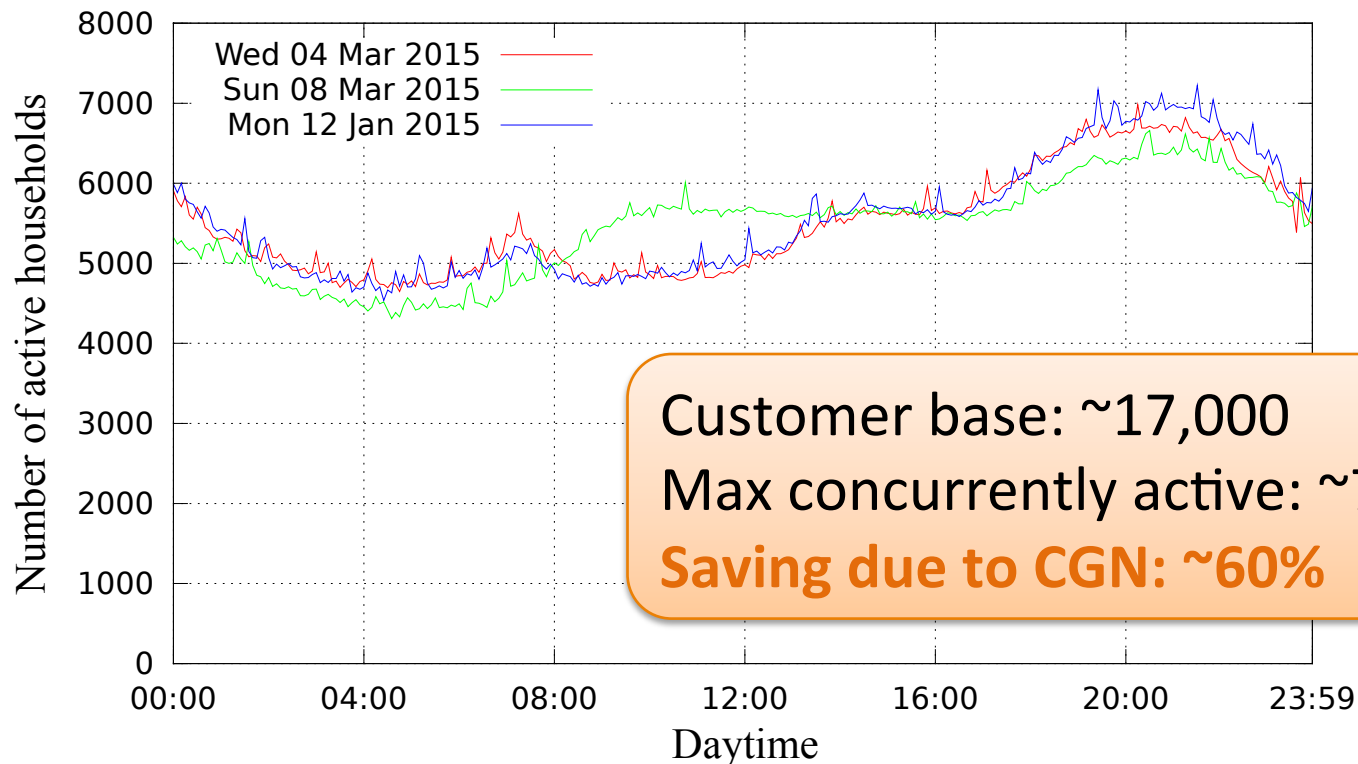
For ISPs, how many IPs would CGN let them save?

How to properly dimension CGN?

26

How many households are concurrently active?

- Assume an idle timer of 5min at the CGN
- **Active** – those who generate one connection in the last 5min

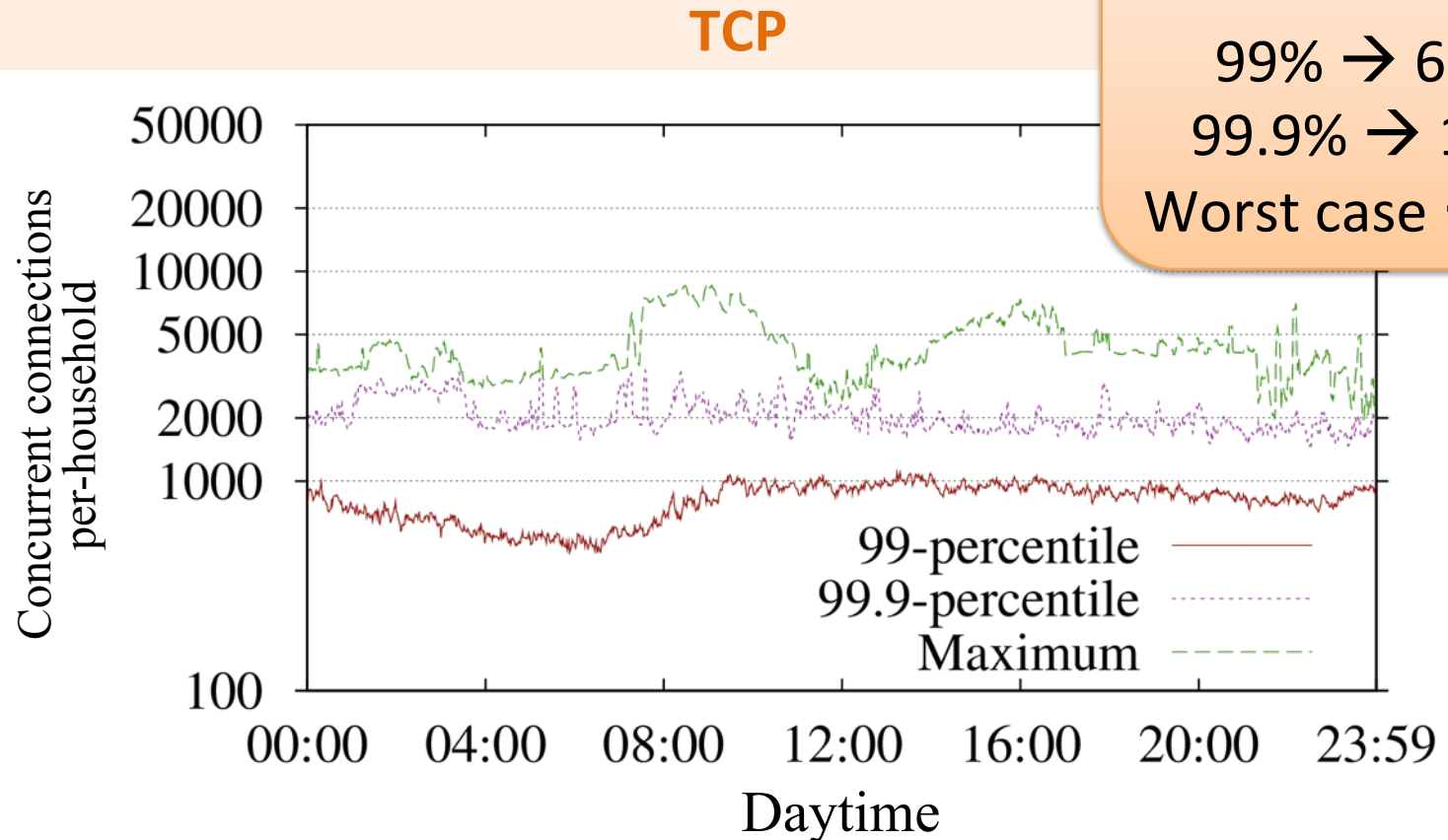


What if Port Address Translation (PAT)?

27

How many concurrent connections?

- **Concurrent** – active in the past 5 min



Potential Saving
(with 65k ports)

99% → 65x

99.9% → 18x

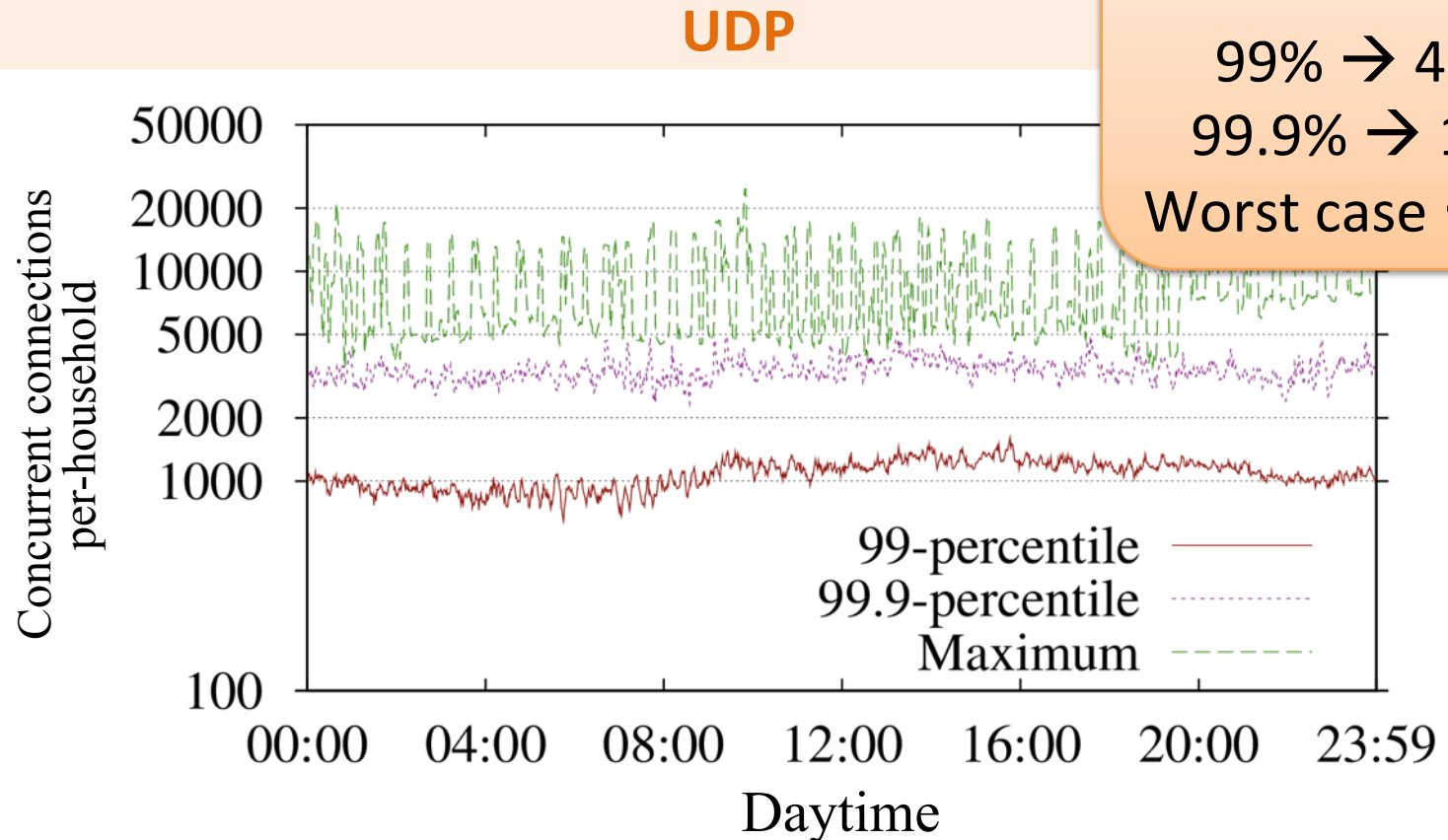
Worst case → 1x

What if Port Address Translation (PAT)?

28

How many concurrent connections?

- **Concurrent** – active in the past 5 min



Potential Saving
(with 65k ports)

99% → 43x

99.9% → 13x

Worst case → 1x

Conclusions

29

- Goal: assess the impact of CGN on users' web browsing
 - Large scale **passive measurements**
 - Multiple performance metrics considered
 - Jensen-Shannon to pinpoint relevant statistical differences
- CGN does **not harm** users' web browsing
 - Results show **negligible impact**
 - Customers with private address get same performance as customers with public addresses
 - **Positive side-effects** against unsolicited traffic
- Saving in terms of \$\$\$ could be significant (15x)
 - Still, temporary patch (IPv6 anyone?)

?? || //