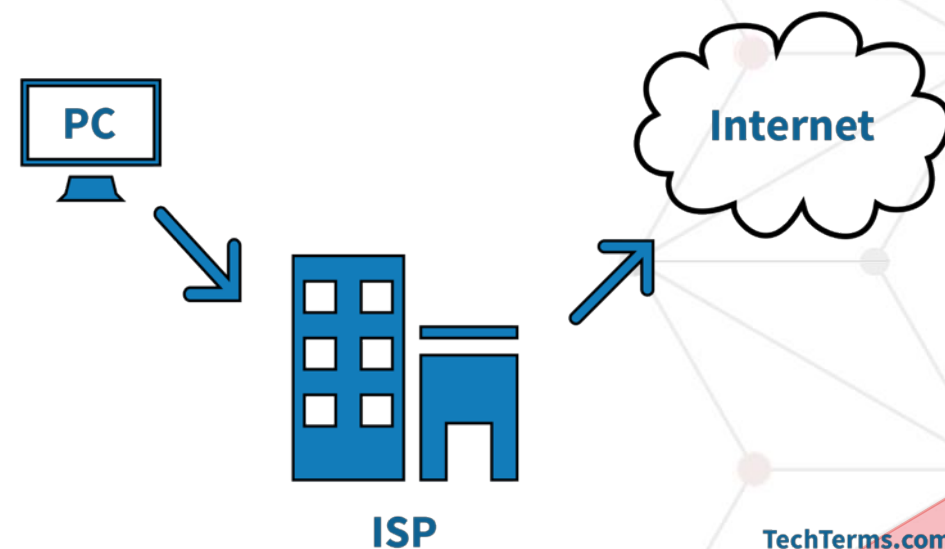




Five Years at the Edge: Watching Internet from the ISP Network

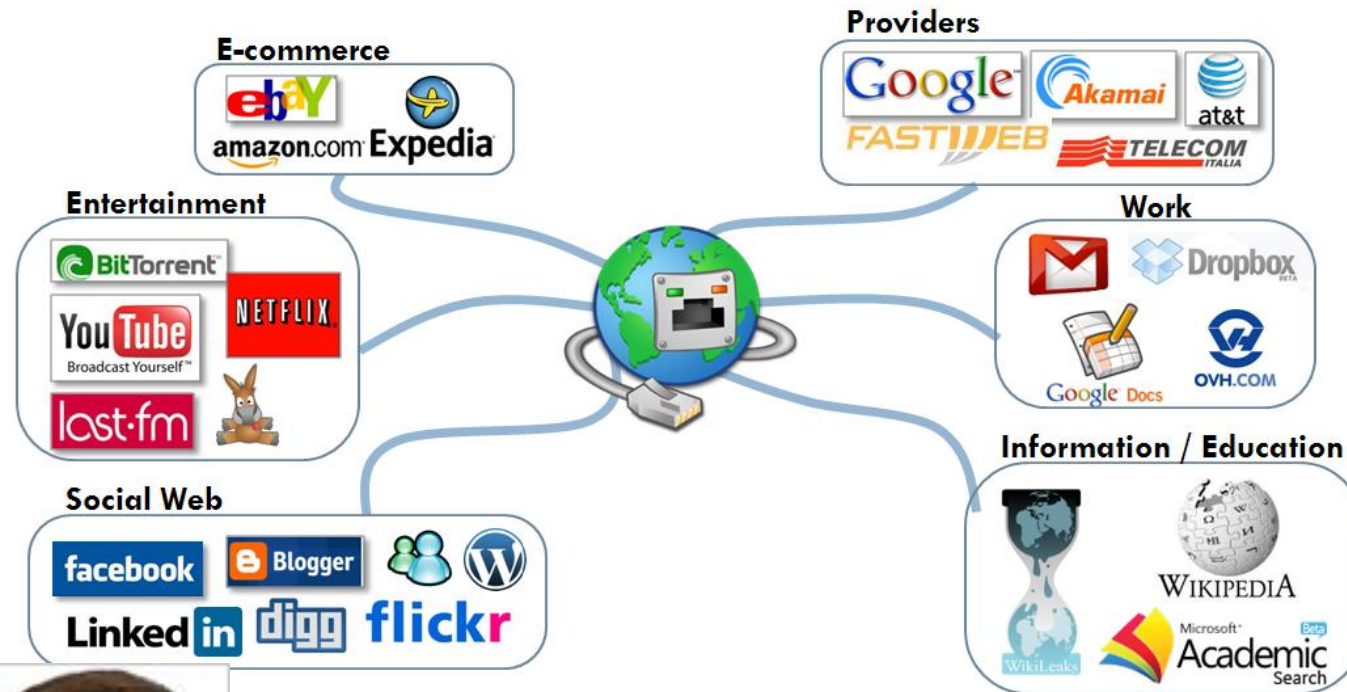
Danilo Giordano, Martino Trevisan, Idilio
Drago, Marco Mellia, Maurizio Munafò



The Goals



POLITECNICO
DI TORINO



*"The Internet is the first thing that humanity has built that humanity doesn't **understand**, the largest experiment in **anarchy** that we have ever had."*

Eric Schmidt – ex Google Exec. Chairman

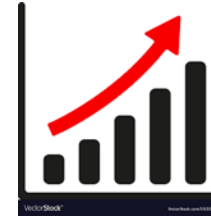
The Questions



POLITECNICO
DI TORINO



1. How much does the user cost?



2. How users' habits changed?

3. How the protocols changed?



4. How the infrastructure changed?



Measurement Setup

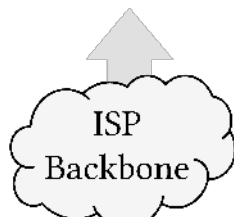


POLITECNICO
DI TORINO

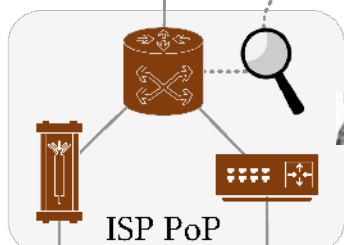


Internet

Data Center

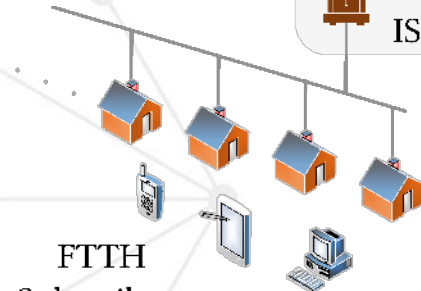


ISP
Backbone

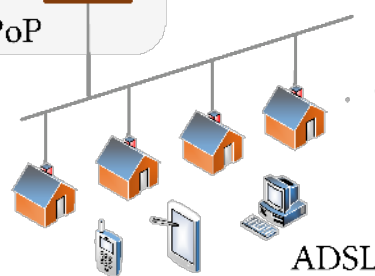


ISP PoP

Passive Probe
Tstat



FTTH
Subscribers



ADSL
Subscribers

5 k households

Download 100 Mbps
Upload 10 Mbps

10 k households

Download 4-20 Mbps
Upload 1 Mbps

5 Years 04/2013 – 10/2017

247 Billions flows

31.9 TB of compressed data

*Data at <https://smartdata.polito.it/five-years-at-the-edge-watching-internet-from-the-isp-network/>

Measurement Setup



POLITECNICO
DI TORINO



Internet

Data Center

5 Years 04/2013 – 10/2017

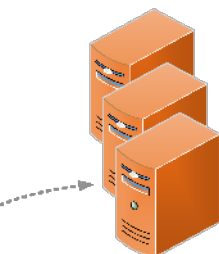
247 Billions flows

31.9 TB of compressed data

Tstat

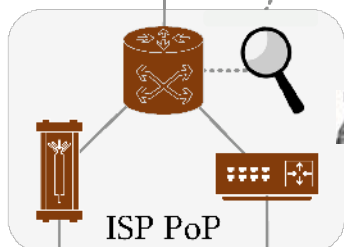


ISP
Backbone

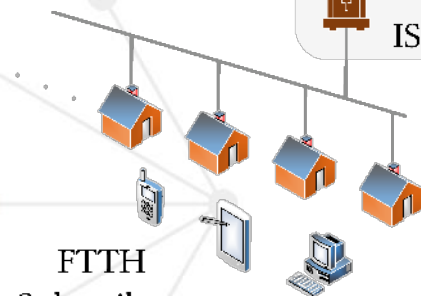


Passive Probe

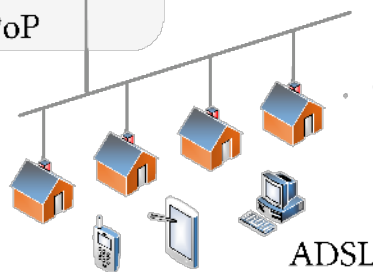
Tstat



ISP PoP



FTTH
Subscribers



ADSL
Subscribers

- Process traffic in **real-time**
- Statistics computed at the end of the flow
- Implements DN-Hunter
 - To Track DNS conversations
 - To tag the network flows with the FQDN

c_ip	c_port	c_pkts	c_bytes	s_ip	s_port	s_pkts	s_bytes	Protocol	FQDN	TLS/SNI	HTTP Host:
12.132.54.94*	1197	1	18938	87.250.137.92	443	1	992221	HTTP	Facebook.com	Facebook.com	--

*Client IP address is anonymized

Tstat

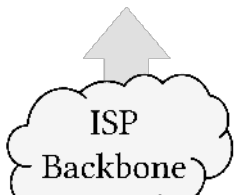
DN-Hunter

Measurement Setup



Internet

Data Center



ISP
Backbone

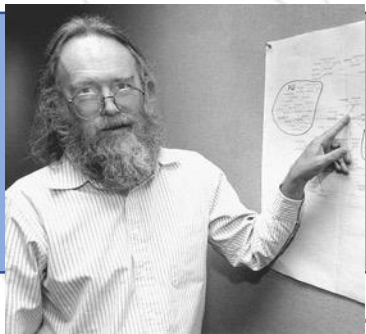


5 Years 04/2013 – 10/2017

247 Billions flows

31.9 TB of compressed data

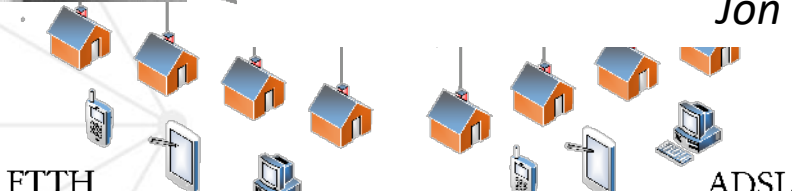
Tstat



*"A name indicates what we seek.
An address indicates where it is.
A route indicates how we get there."*

ISP PoP

Jon Postel – RFC 791



FTTH

ADSL

*Full list at <https://smartdata.polito.it/five-years-at-the-edge-watching-internet-from-the-isp-network/>

- Process traffic in real-time

When	FQDN
Current	facebook.com, fbcdn.com
Before	fbstatic-a.akamaihd.net

- To tag the network flows with the FQDN



Regex*	Service
facebook.com, fbcdn.com, ^fbstatic-[a-z].akamaihd.net\$	Facebook
netflix.com, nflxvideo.net	Netflix

c_ip	c_port	c_pkts	c_bytes	s_ip	s_port	s_pkts	s_bytes	Protocol	FQDN	TLS/SNI	HTTP Host:
12.132.54.94*	1197	1	18938	87.250.137.92	443	1	992221	HTTP	Facebook.com	Facebook.com	--

*Client IP address is anonymized

Tstat

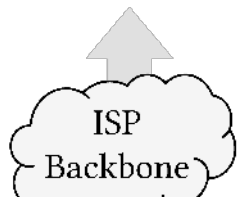
DN-Hunter

Measurement Setup



Internet

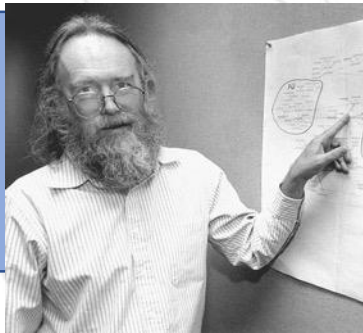
Data Center



5 Years 04/2013 – 10/2017

247 Billions flows

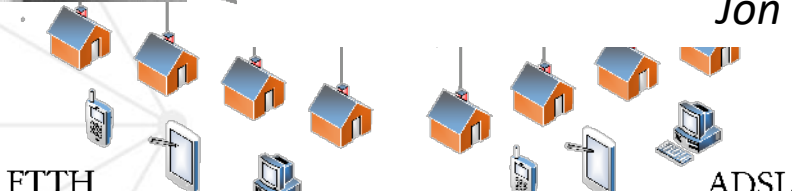
31.9 TB of compressed data



*“A name indicates what we seek.
An address indicates where it is.
A route indicates how we get there.”*

ISP PoP

Jon Postel – RFC 791



*Full list at <https://smartdata.polito.it/five-years-at-the-edge-watching-internet-from-the-isp-network/>

- Process traffic in real-time

When	FQDN
Current	facebook.com, fbcdn.com
Before	fbstatic-a.akamaihd.net

- To tag the network flows with the FQDN



Regex*	Service
facebook.com, fbcdn.com, ^fbstatic-[a-z].akamaihd.net\$	Facebook
netflix.com, nflvideo.net	Netflix

c_ip	c_port	c_pkts	c_bytes	s_ip	s_port	s_pkts	s_bytes	Protocol	Service	HTTP Host:
12.132.54.94*	1197	1	18938	87.250.137.92	443	1	992221	HTTP	Facebook	--

*Client IP address is anonymized

Tstat

DN-Hunter

Measurement Setup

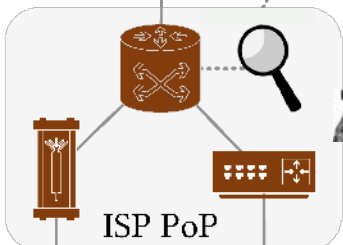


Internet

Data Center



ISP
Backbone



ISP PoP



Passive Probe

5 Years 04/2013 – 10/2017

247 Billions flows

31.9 TB of compressed data



Aggregation

Analytics

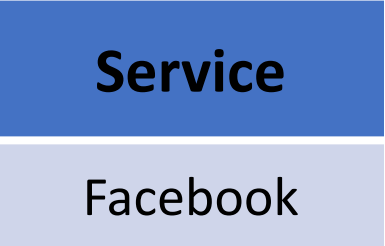
- Process traffic in real-time

When	FQDN
Current	facebook.com, fbcdn.com
Before	fbstatic-a.akamaihd.net

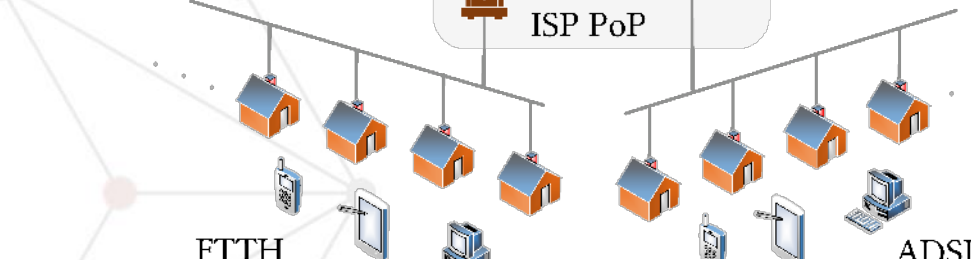
- To tag the network flows with the FQDN



Regex*	Service
facebook.com, fbcdn.com, ^fbstatic-[a-z].akamaihd.net\$	Facebook
netflix.com, nflvideo.net	Netflix



DN-Hunter



*Full list at <https://smartdata.polito.it/five-years-at-the-edge-watching-internet-from-the-isp-network/>

c_ip	c_port	c_pkts	c_bytes	s_ip	s_port	s_pkts	s_bytes	Protocol	Service	HTTP Host:
12.132.54.94*	1197	1	18938	87.250.137.92	443	1	992221	HTTP	Facebook	--

*Client IP address is anonymized

Tstat



1. How much does the user cost?

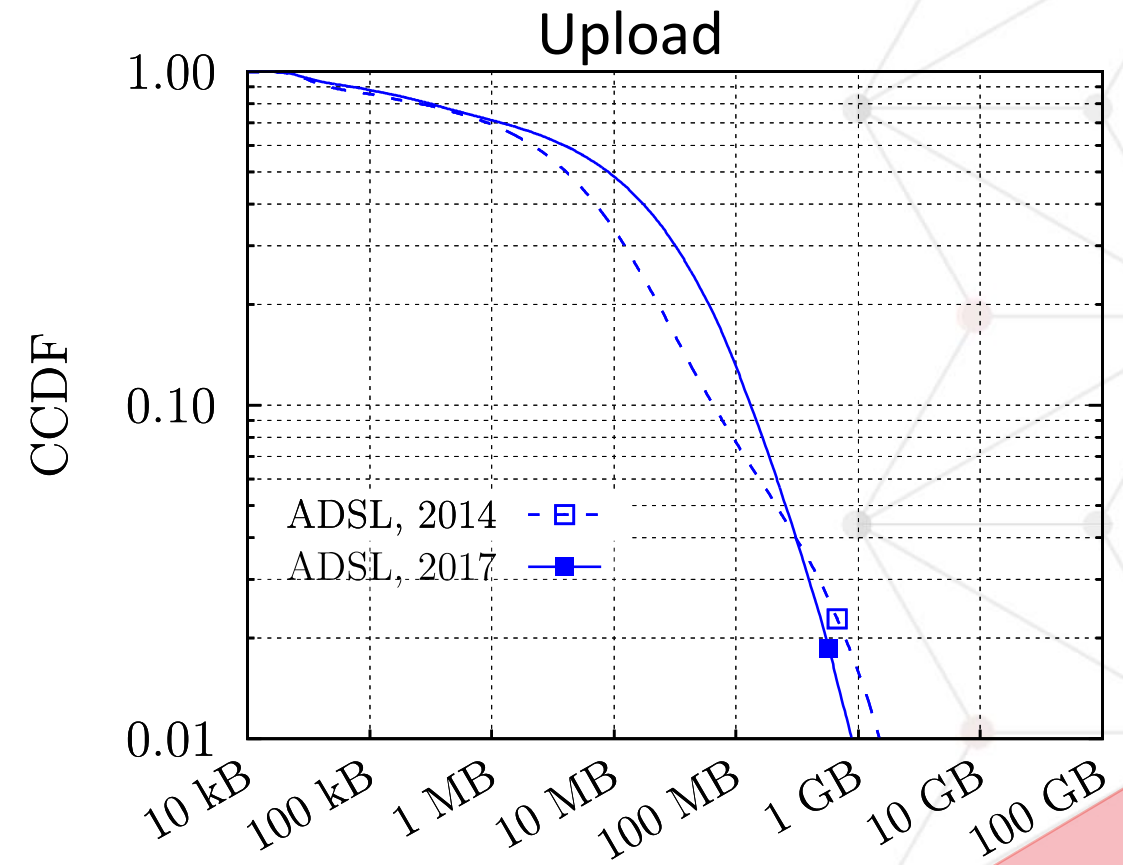
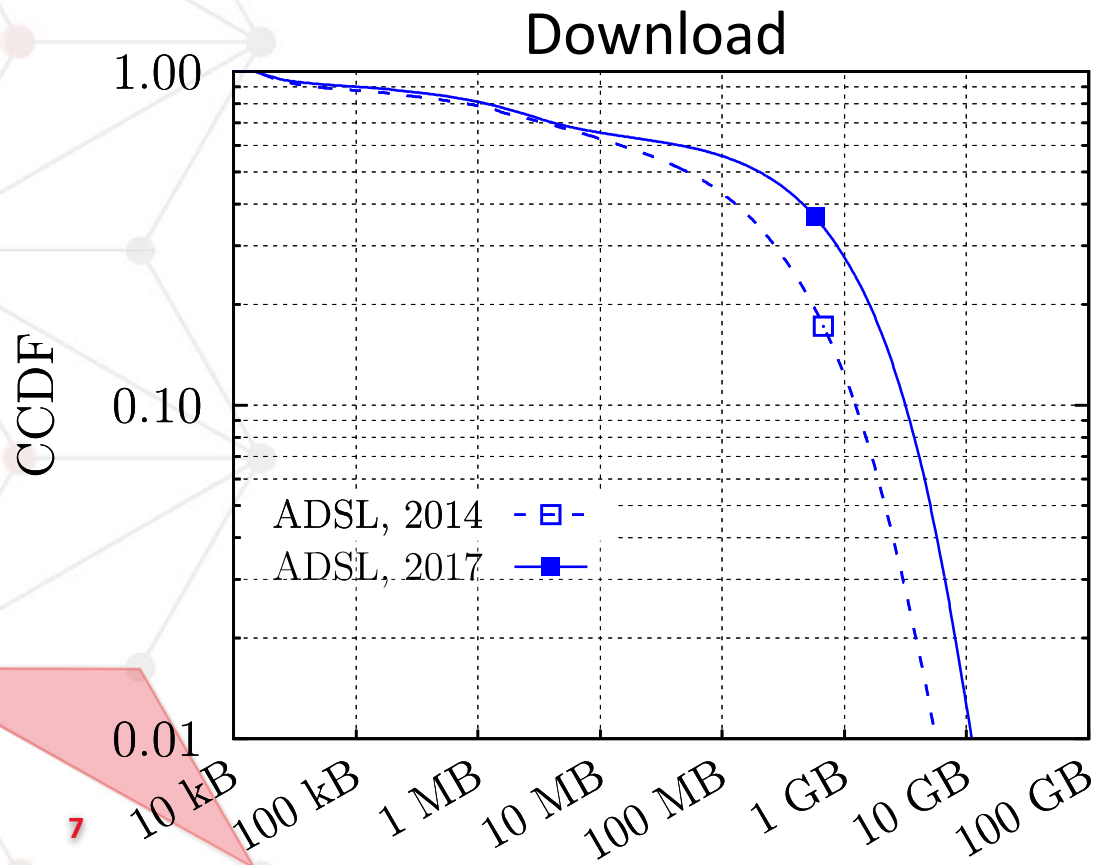
Daily Traffic Consumption



POLITECNICO
DI TORINO



- 2014 vs 2017
 - **ADSL** 2x Download and Upload



Daily Traffic Consumption

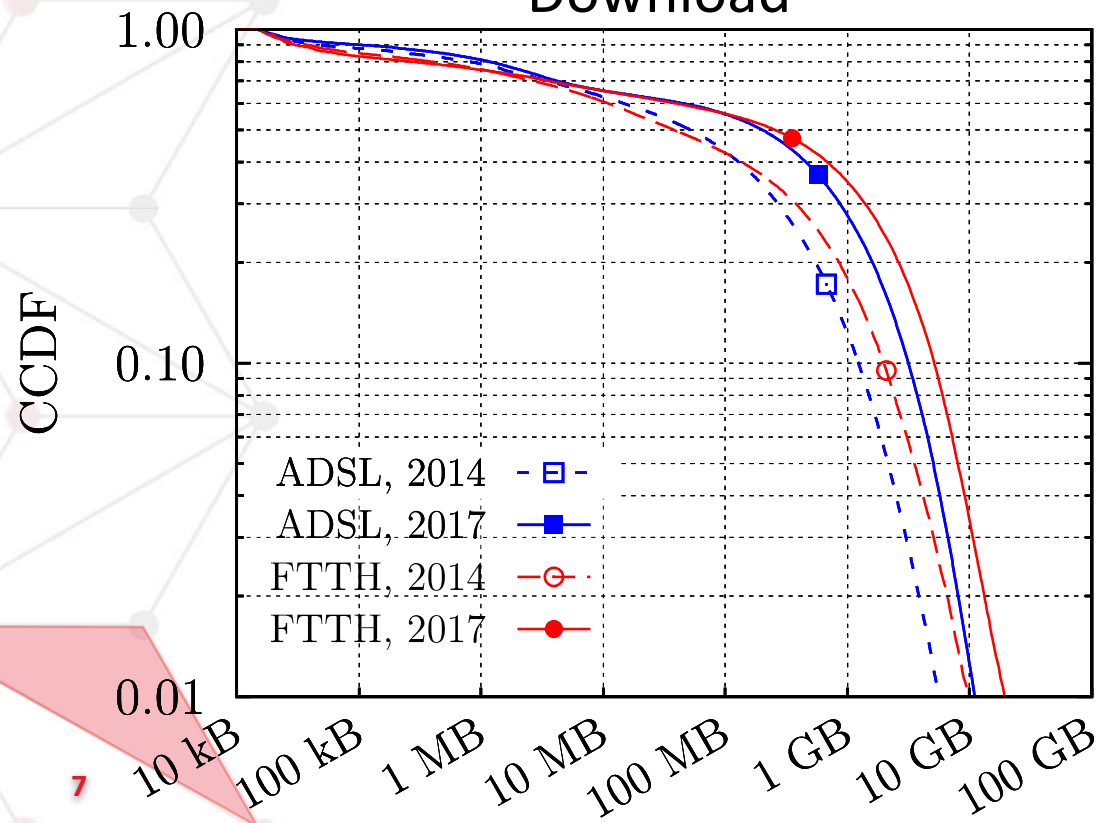


POLITECNICO
DI TORINO

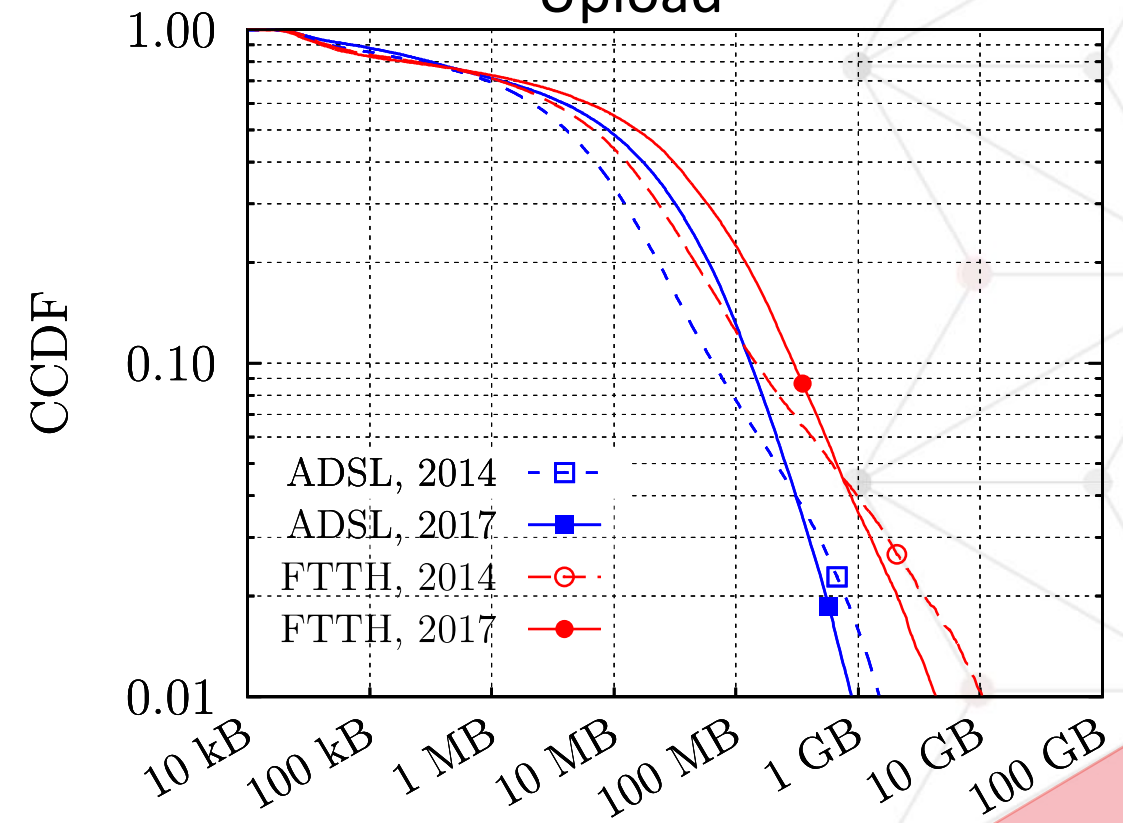


- **CCDF of Daily** traffic consumption per household
- **FTTH** vs **ADSL**
 - 25% + Download, 2x Upload

Download



Upload



Daily Traffic Consumption

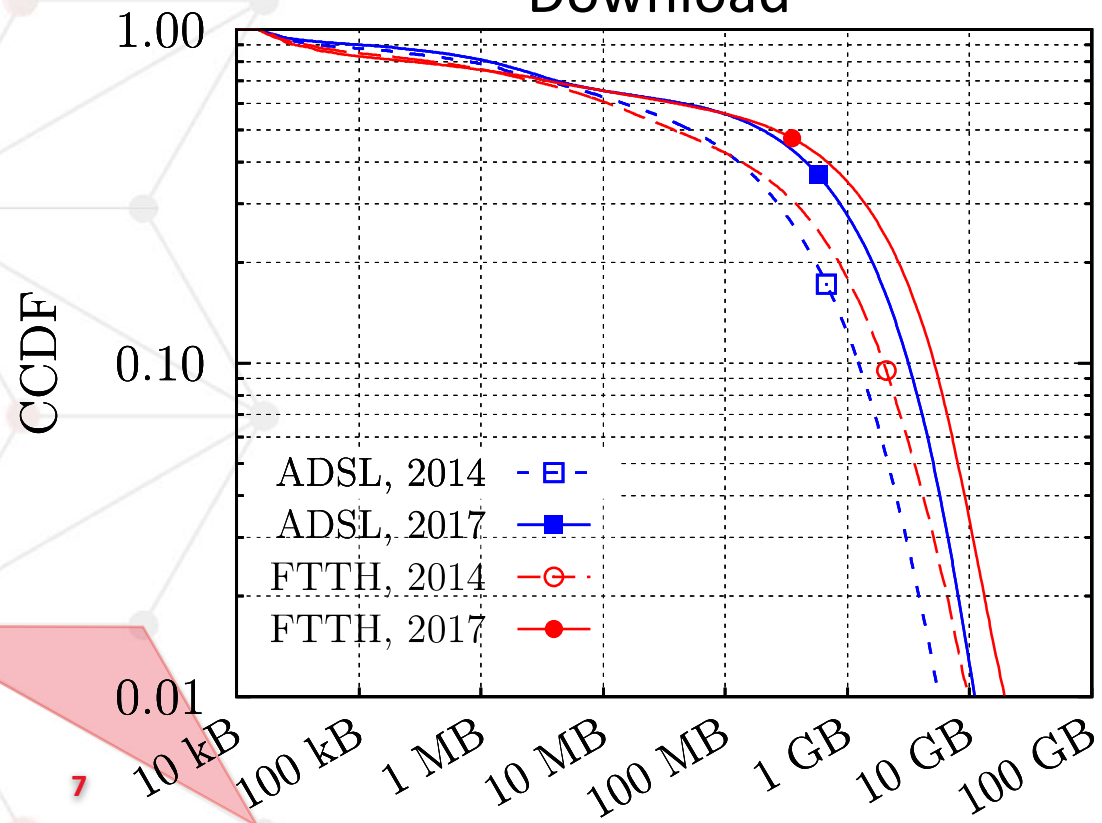


POLITECNICO
DI TORINO

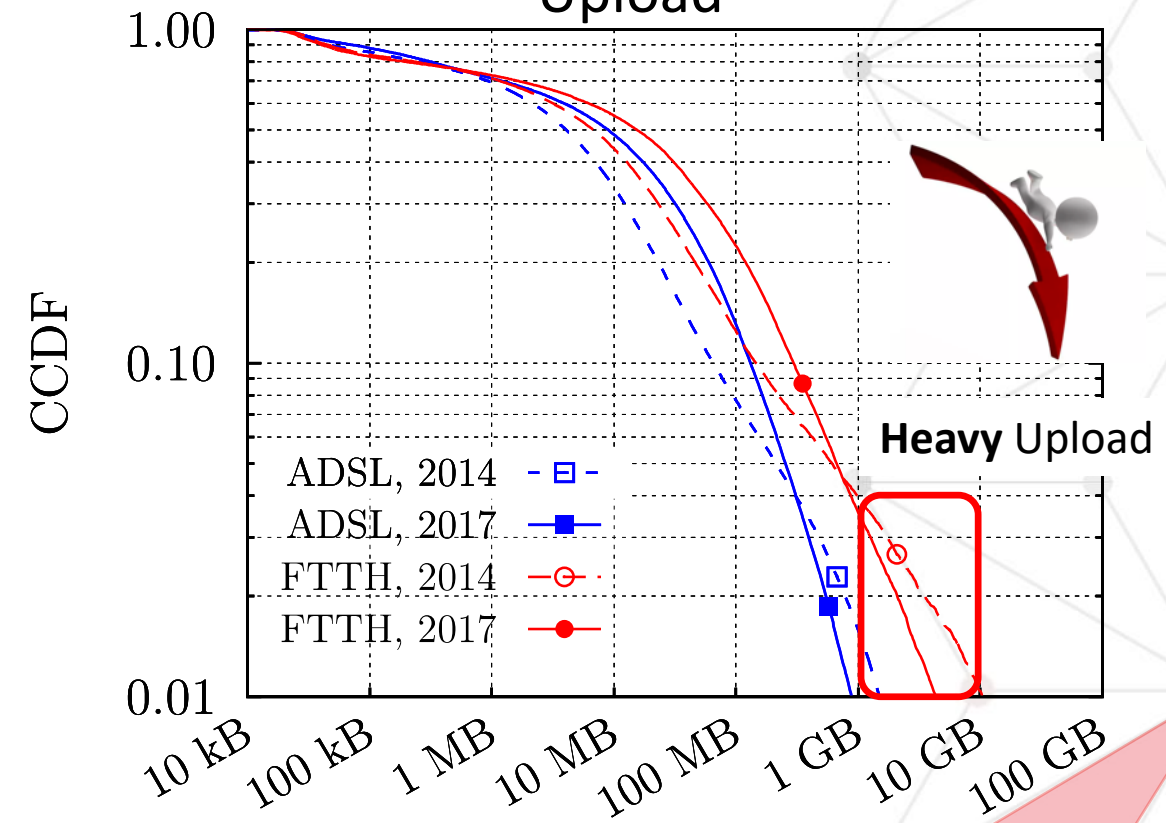


- **CCDF of Daily** traffic consumption per household
- **FTTH** vs **ADSL**
 - 25% + Download, 2x Upload

Download



Upload





2. How users' habits changed?

Service Popularity



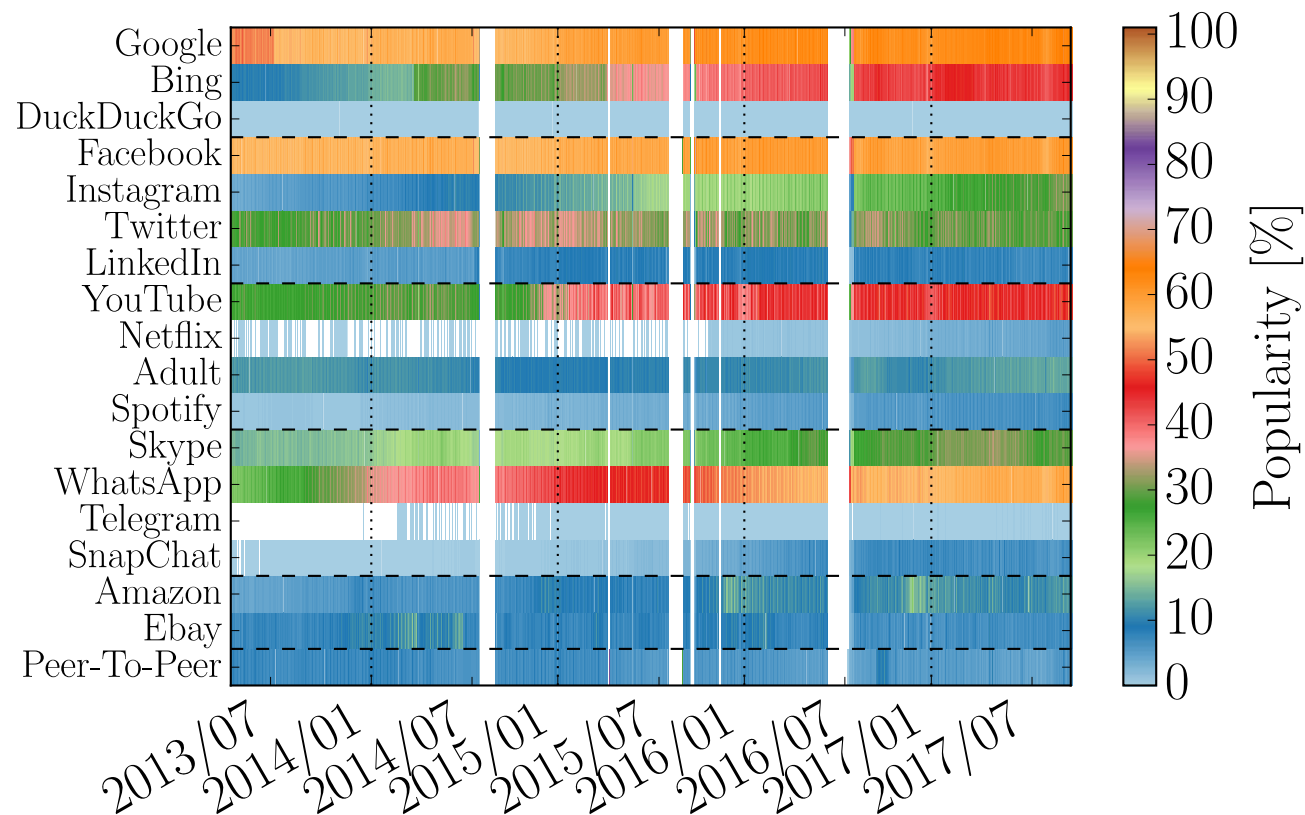
POLITECNICO
DI TORINO



- % of active households accessing a given service



Only Real Visits



Service Popularity



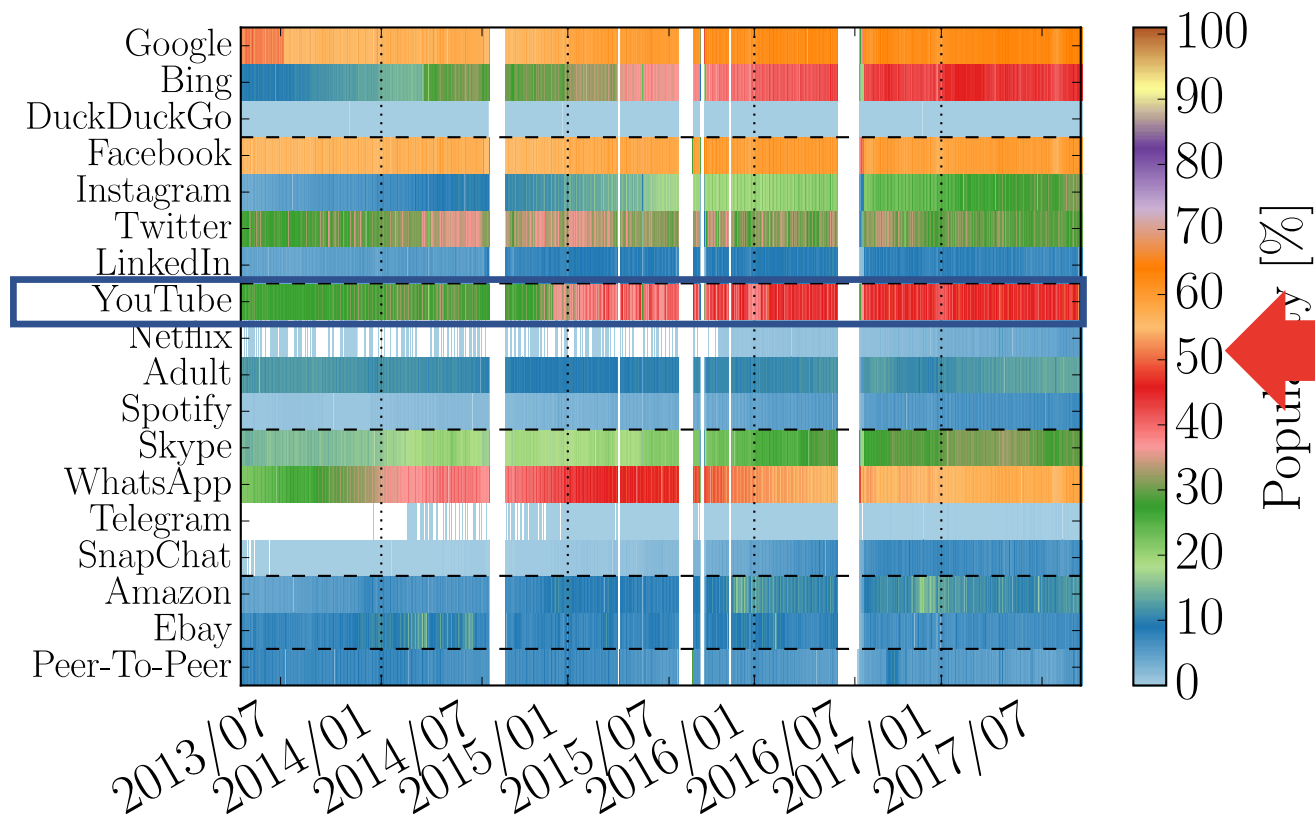
POLITECNICO
DI TORINO



- % of active households accessing a given service



Only Real Visits



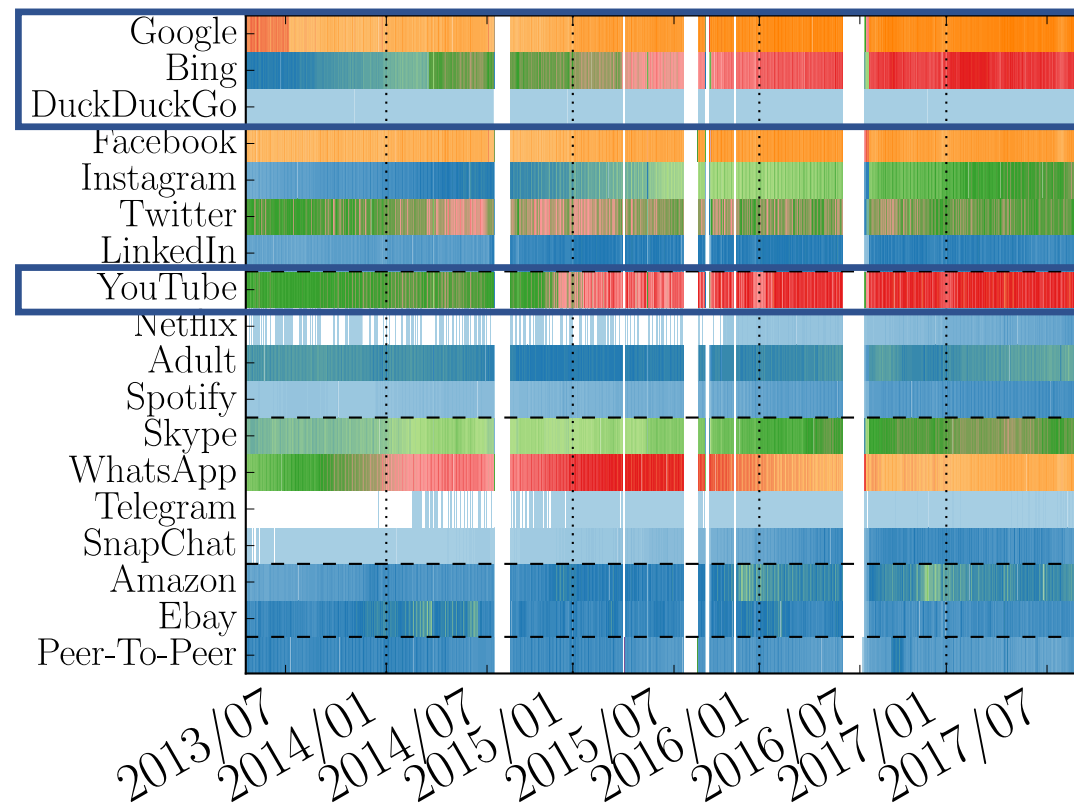
Service Popularity



POLITECNICO
DI TORINO



- % of active households accessing a given service
- Only Real Visits



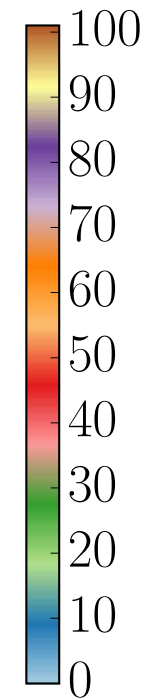
15 % → 45 %



DuckDuckGo

0.3 % → 0.3 %

Popularity [%]



Service Popularity



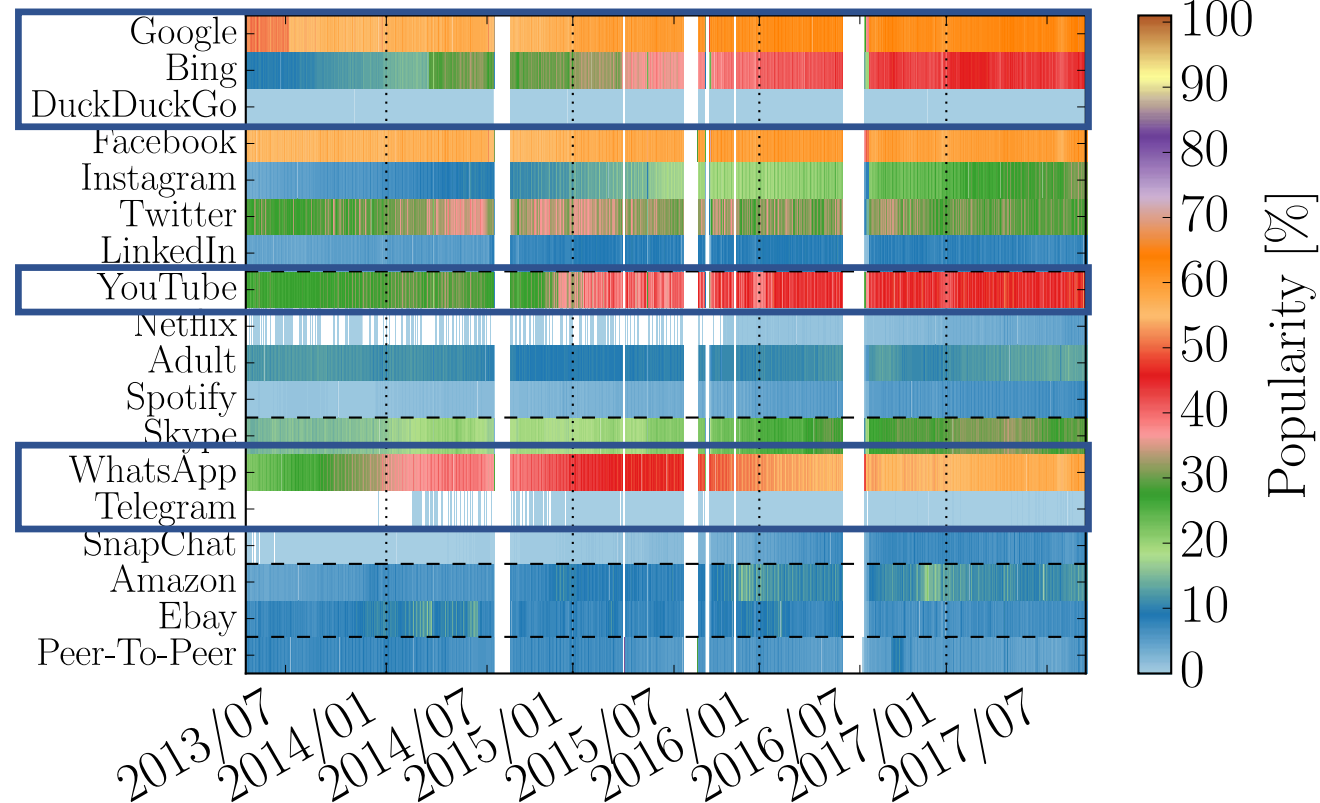
POLITECNICO
DI TORINO



- % of active households accessing a given service



Only Real Visits



15 % → 45 %



DuckDuckGo

0.3 % → 0.3 %



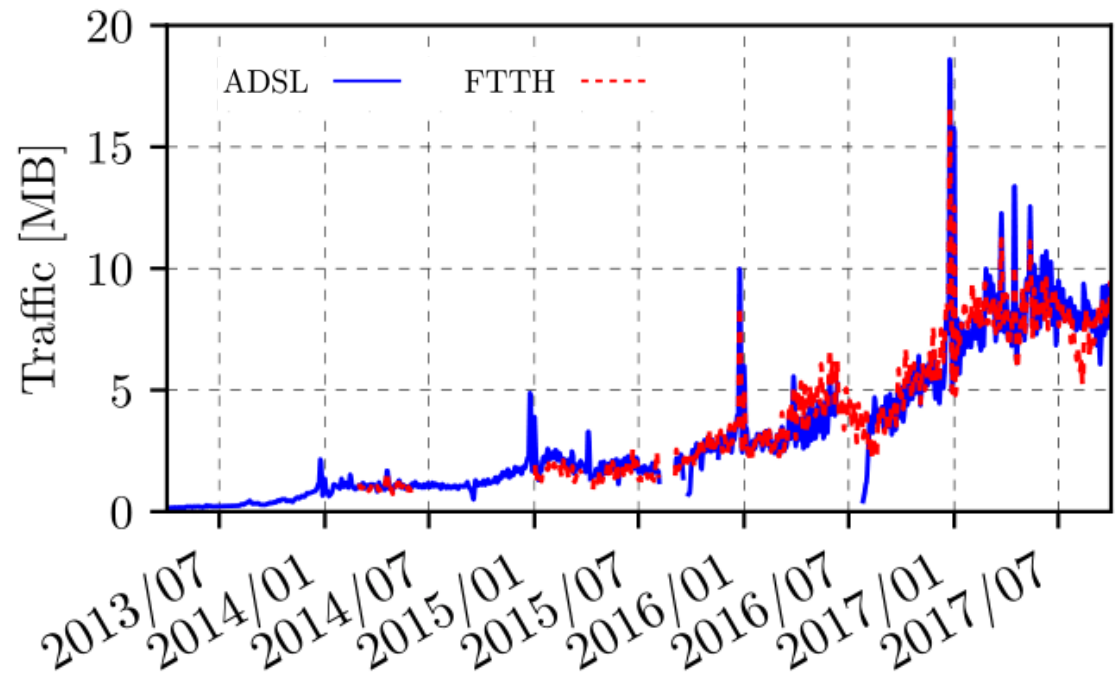
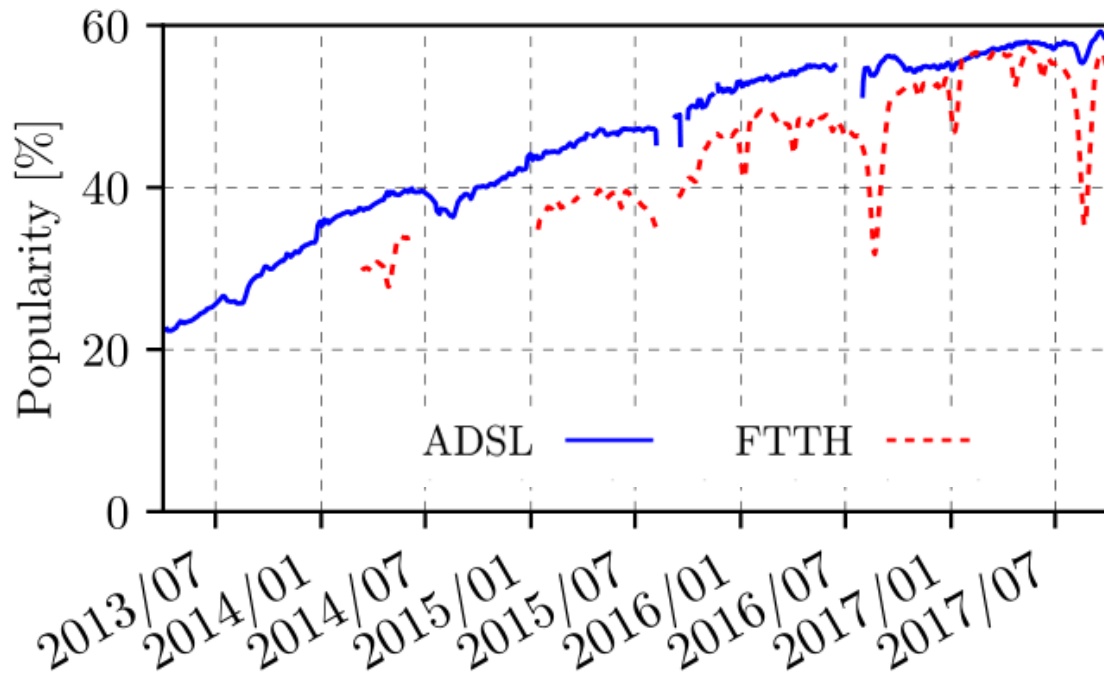
What is the traffic load imposed by web services?



POLITECNICO
DI TORINO

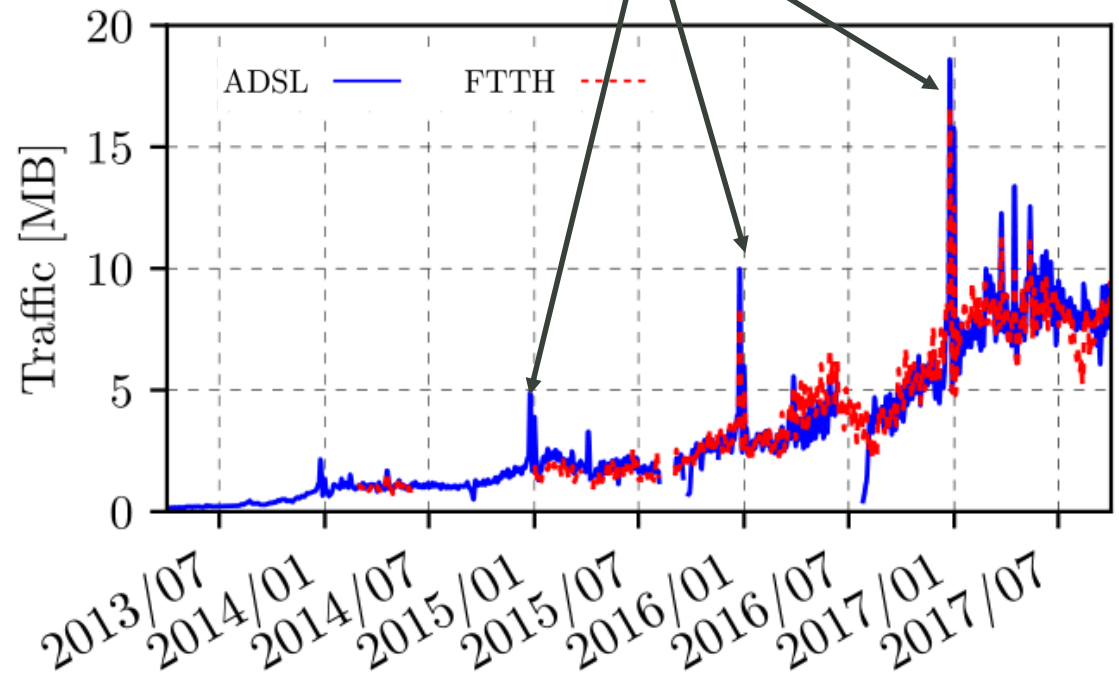
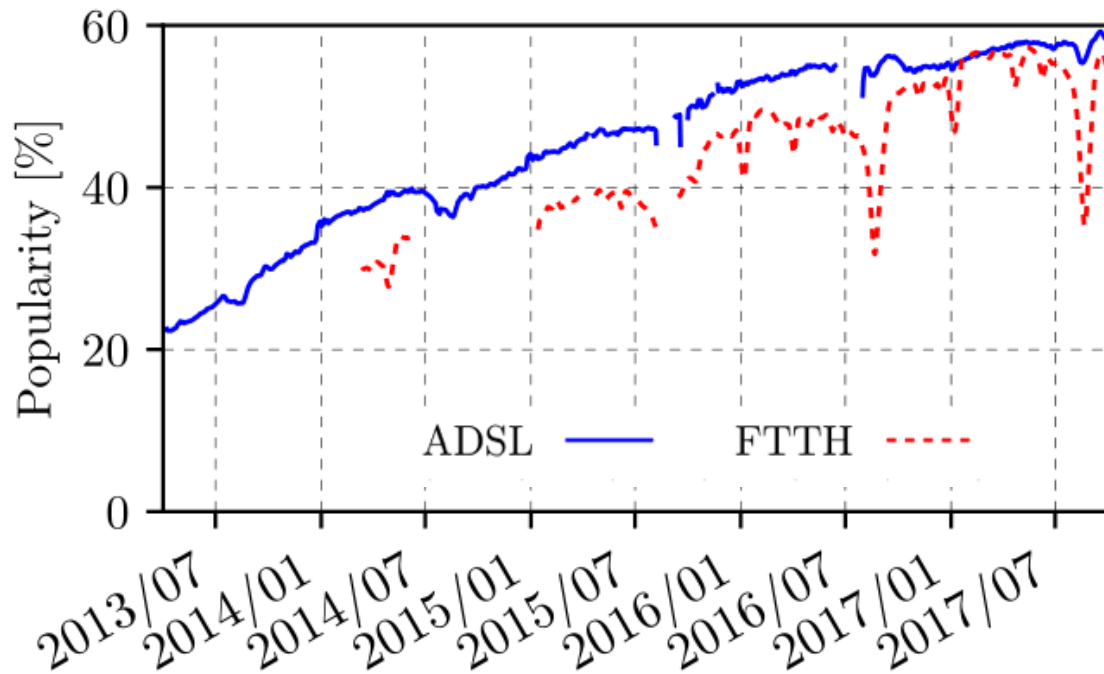


- Popularity almost reached **saturation**
- Per-user traffic now **10 MB** per day



What is the traffic load imposed by web services?

- Popularity almost reached **saturation**
- Per-user traffic now **10 MB** per day
 - Christmas and New Year's Eve peaks



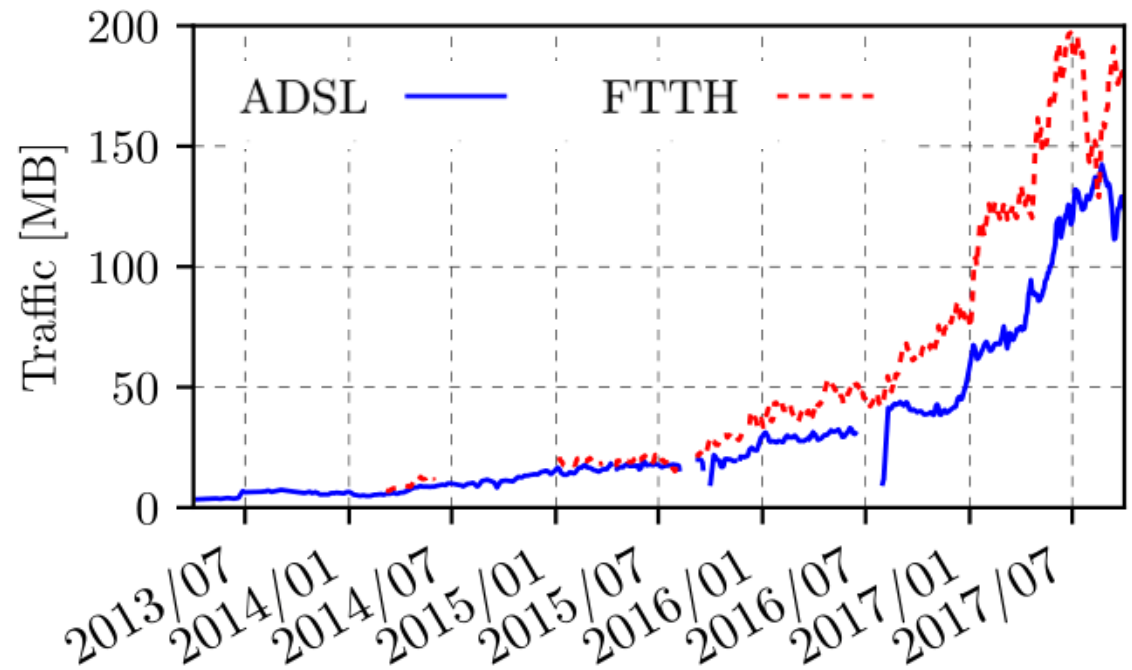
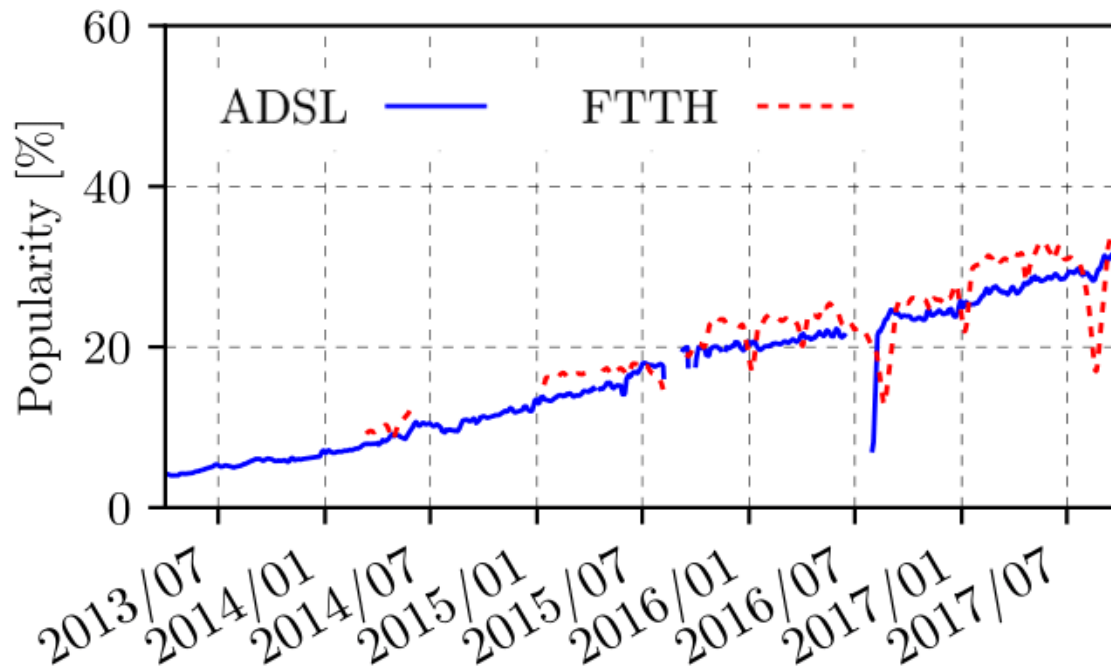
What is the traffic load imposed by web services?



POLITECNICO
DI TORINO



- Popularity still **growing**
- Per-user traffic still **exploding**



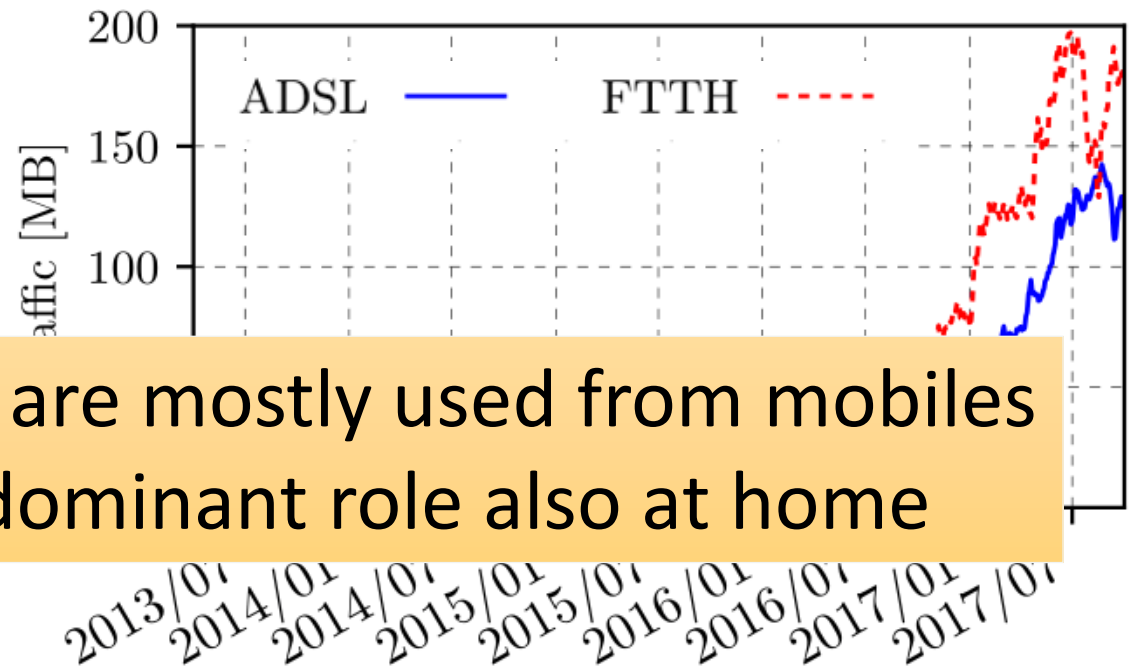
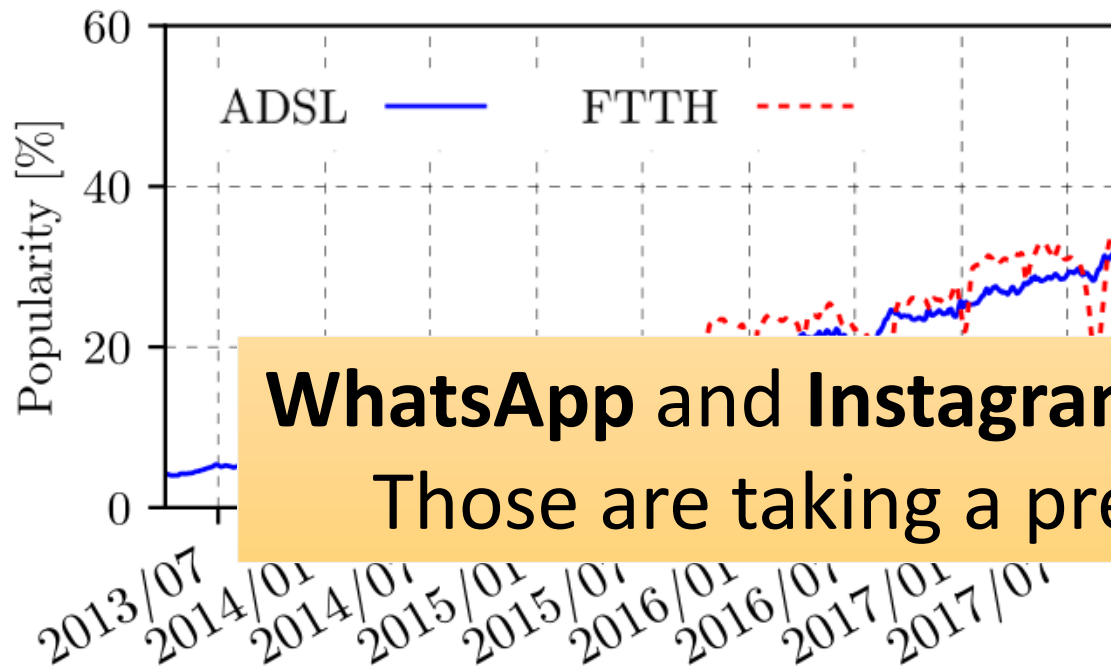
What is the traffic load imposed by web services?



POLITECNICO
DI TORINO



- Popularity still **growing**
- Per-user traffic still **exploding**



WhatsApp and Instagram are mostly used from mobiles
Those are taking a predominant role also at home

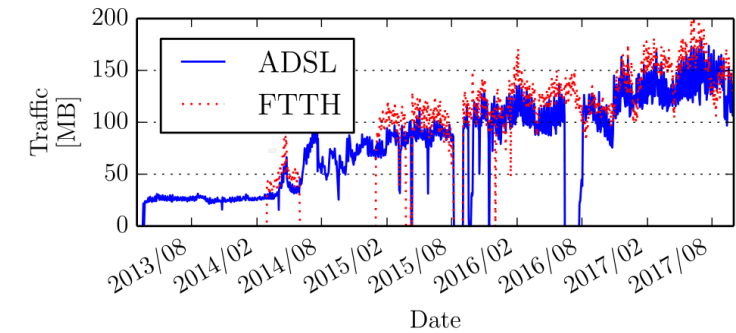
What is the traffic load imposed by web services?



POLITECNICO
DI TORINO



- Average traffic per-client per day



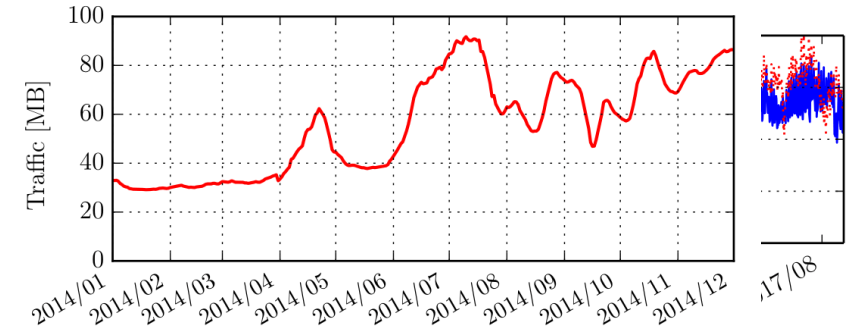
What is the traffic load imposed by web services?



POLITECNICO
DI TORINO



- Average traffic per-client per day



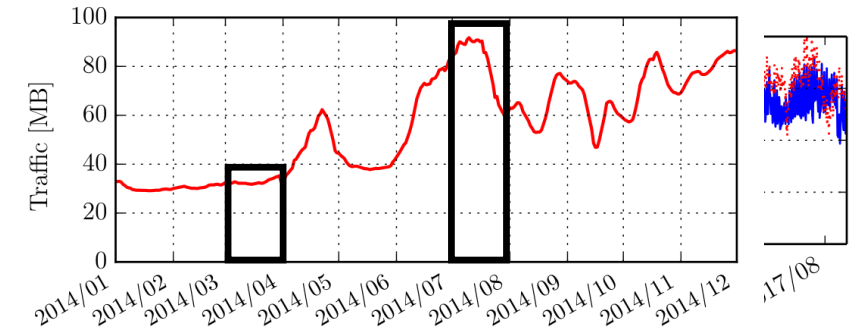
What is the traffic load imposed by web services?



POLITECNICO
DI TORINO



- Average traffic per-client per day
- Video auto-play enabled in 2014
 - April traffic from 35 MB to 60+ MB
 - After video auto-play stopped
 - July daily traffic topped at 90 MB
 - 2.5 times March 2014!



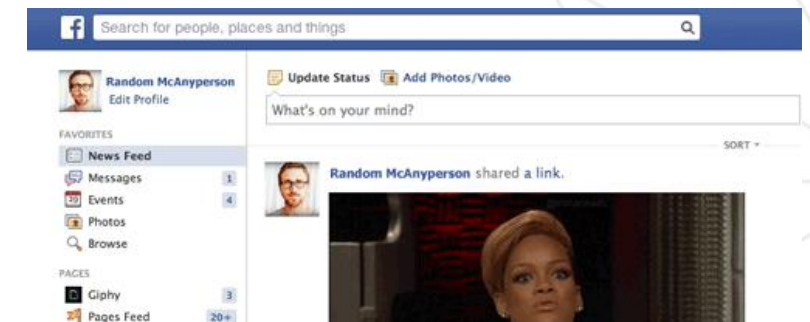
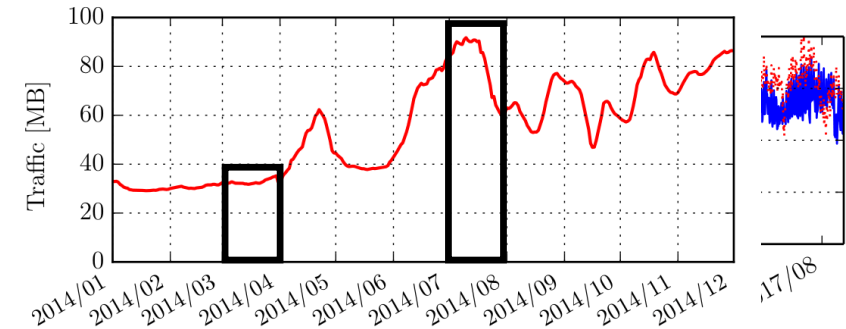
What is the traffic load imposed by web services?



POLITECNICO
DI TORINO



- Average traffic per-client per day
- Video auto-play enabled in 2014
 - April traffic from 35 MB to 60+ MB
 - After video auto-play stopped
 - July daily traffic topped at 90 MB
 - 2.5 times March 2014!



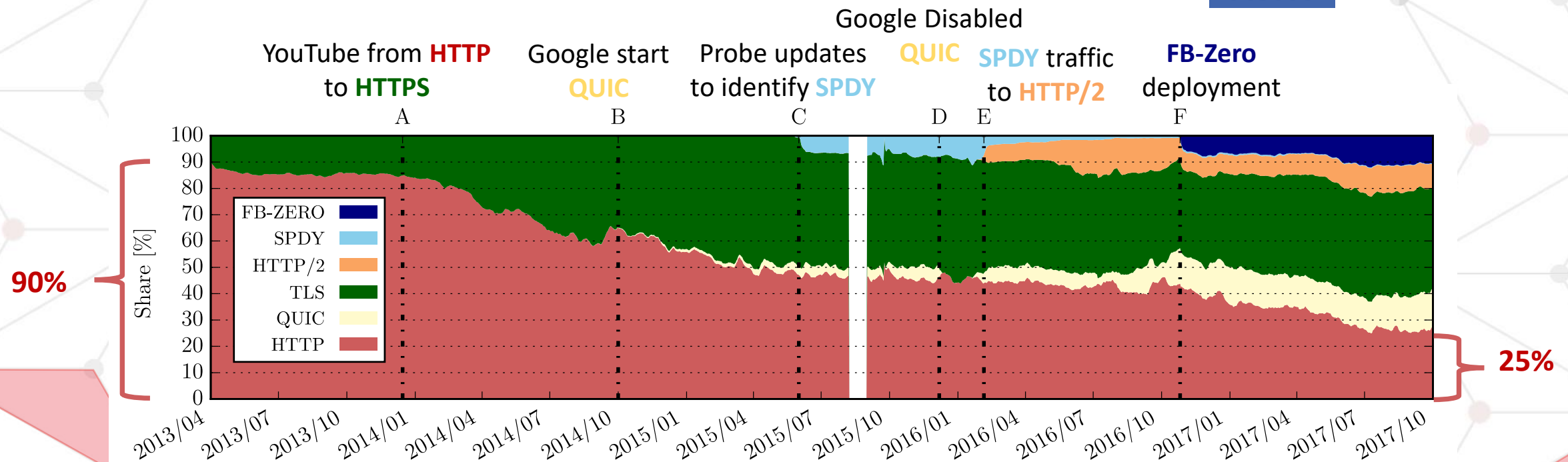
Big players owning servers and clients
can dramatically impact traffic consumption
complicating the management of ISP networks

3. How the protocols changed?

Web Protocols Over Five Years



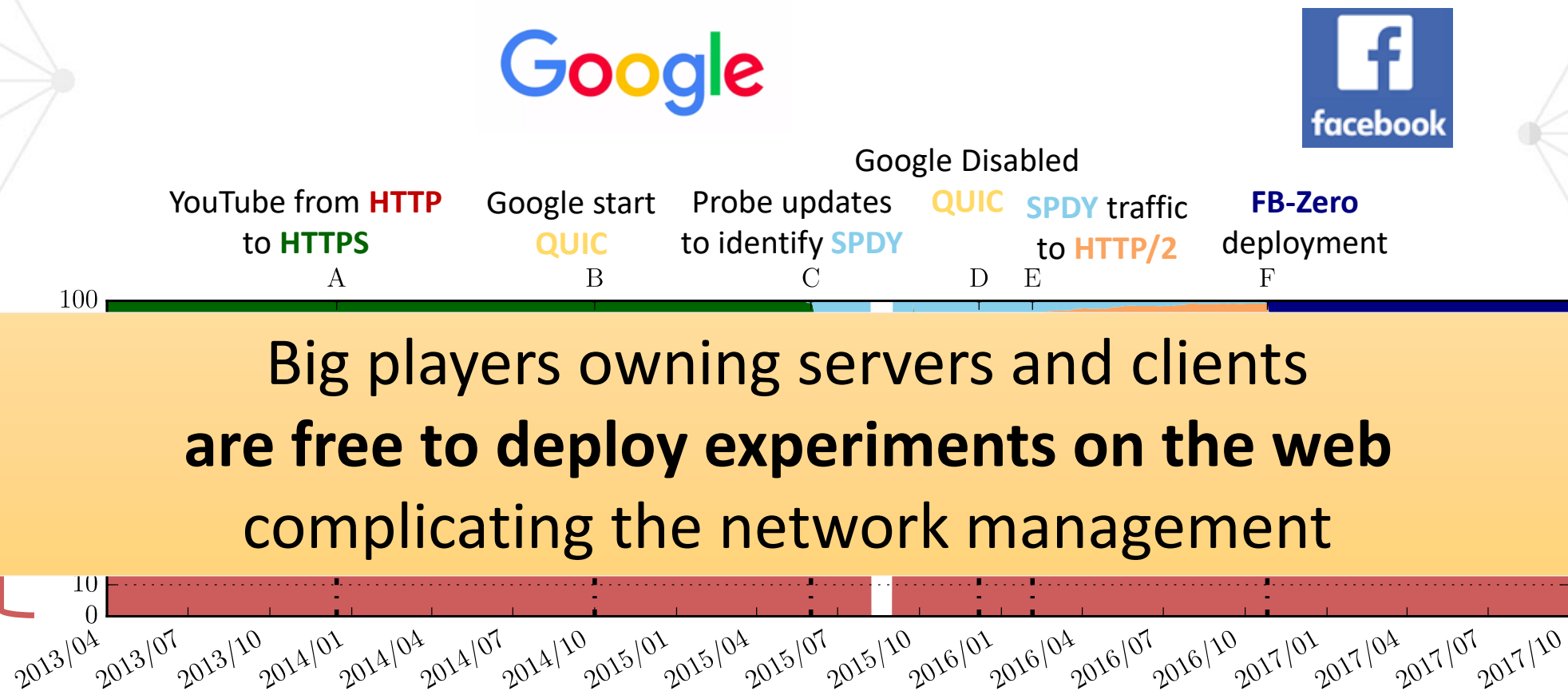
- Traffic share for several Web protocols



Web Protocols Over Five Years



- Traffic share for several Web protocols



4. How the infrastructure changed?

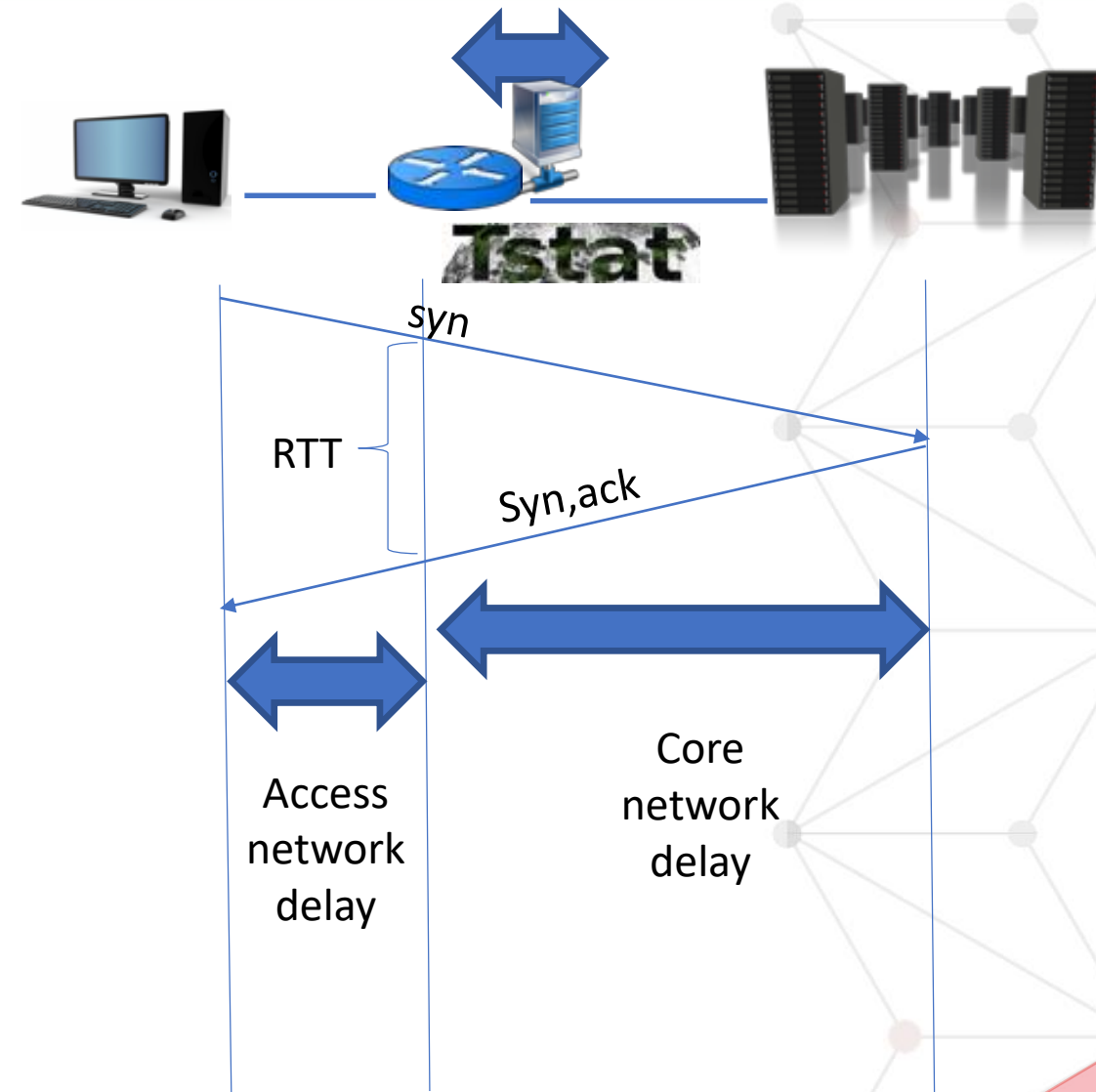
CDN Monitoring Methodology



POLITECNICO
DI TORINO



- Study CDNs impact
 - Content access latency as Users' Quality of Experience metric
 - Focus on the Round Trip Time (RTT)
- RTT from the probe to servers only
 - Core Network Delay
- Excluding RTT from clients to the probe
 - Access network delay



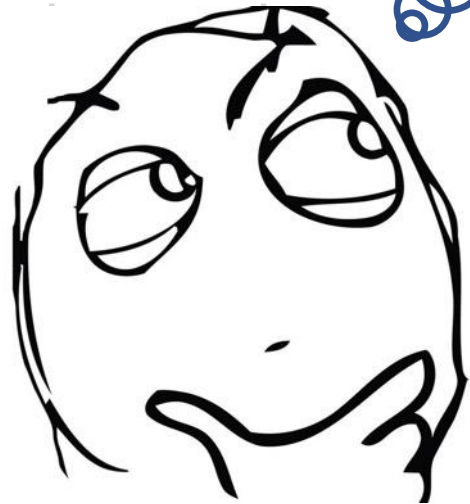
CDN Monitoring Methodology



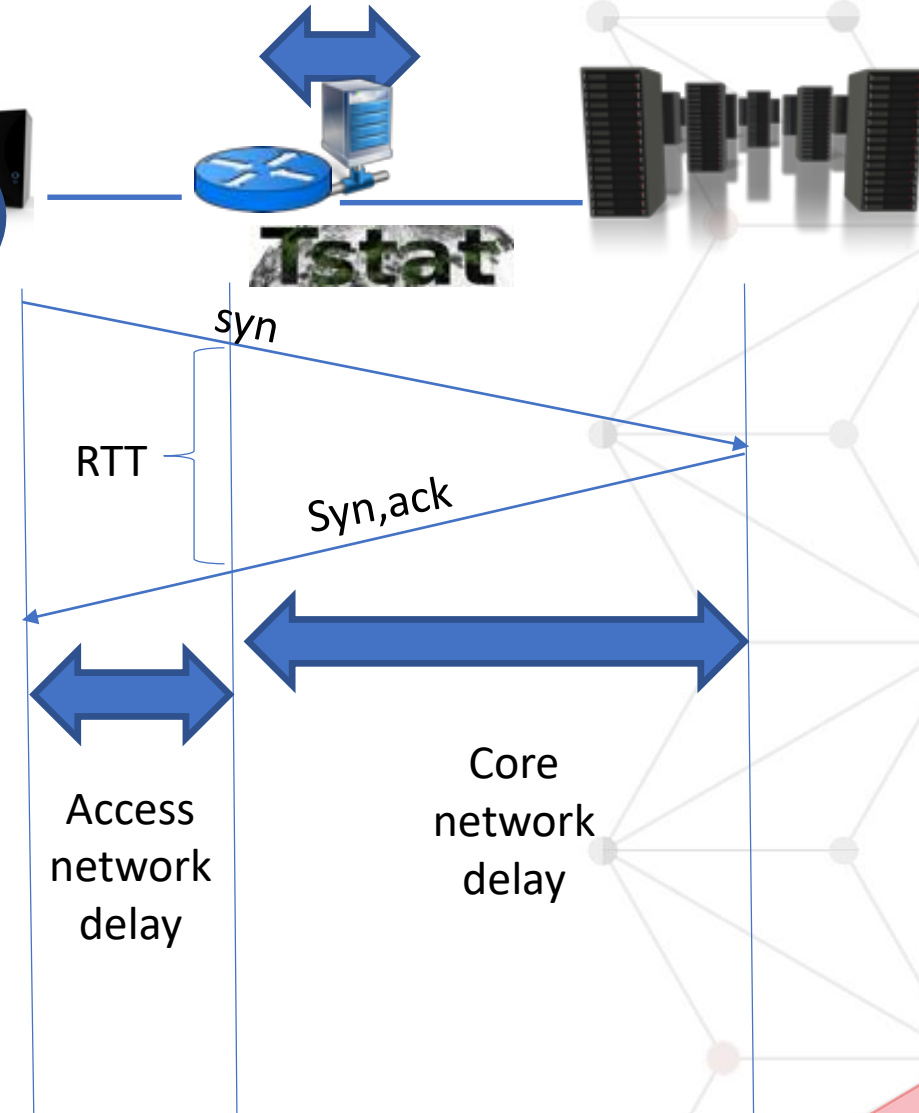
POLITECNICO
DI TORINO



- Study CDNs in
 - Content access
 - Quality of Experience
 - Focus on the Round Trip Time (RTT) to the servers only
- RTT from the client to the server
 - Core network delay
 - Excluded probe delay
 - Access network delay



...Because access
delay is negligible...
... right?

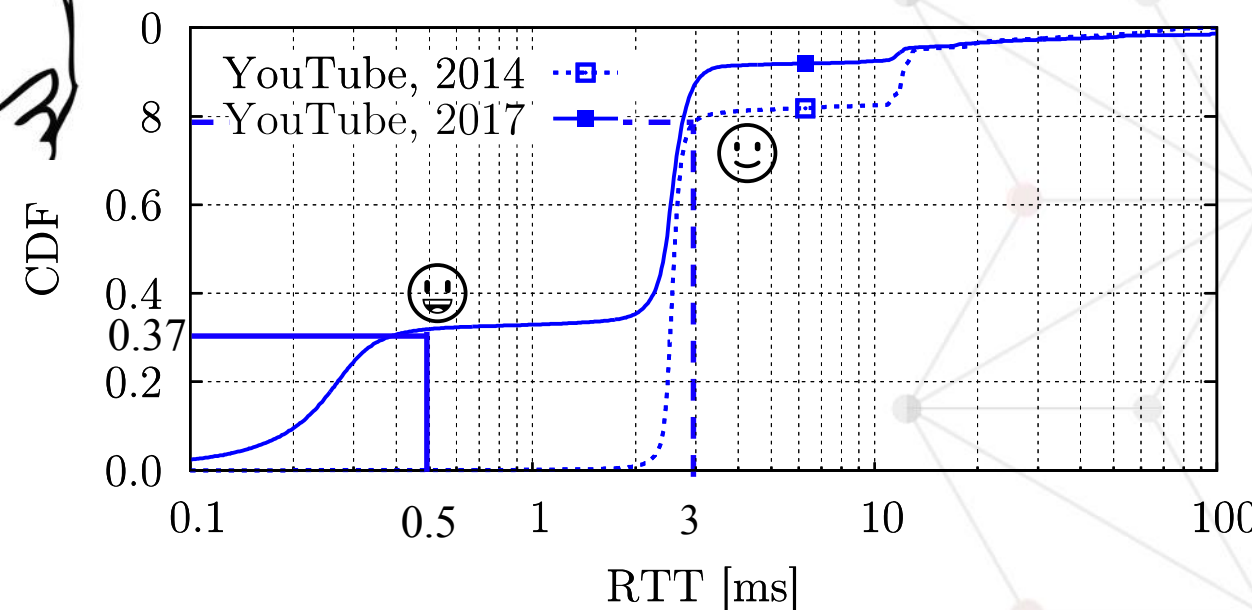
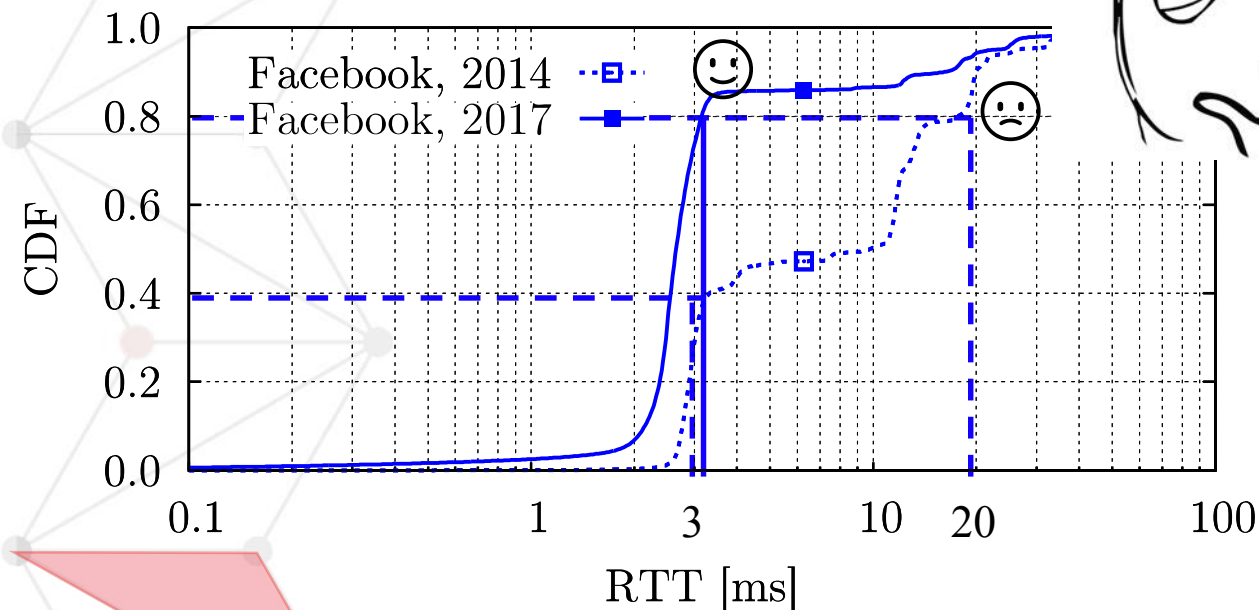


The birth of the sub-millisecond Caching



- Compare RTT Distributions
 - April 2014 vs April 2017

The CDN is inside
the PoP!

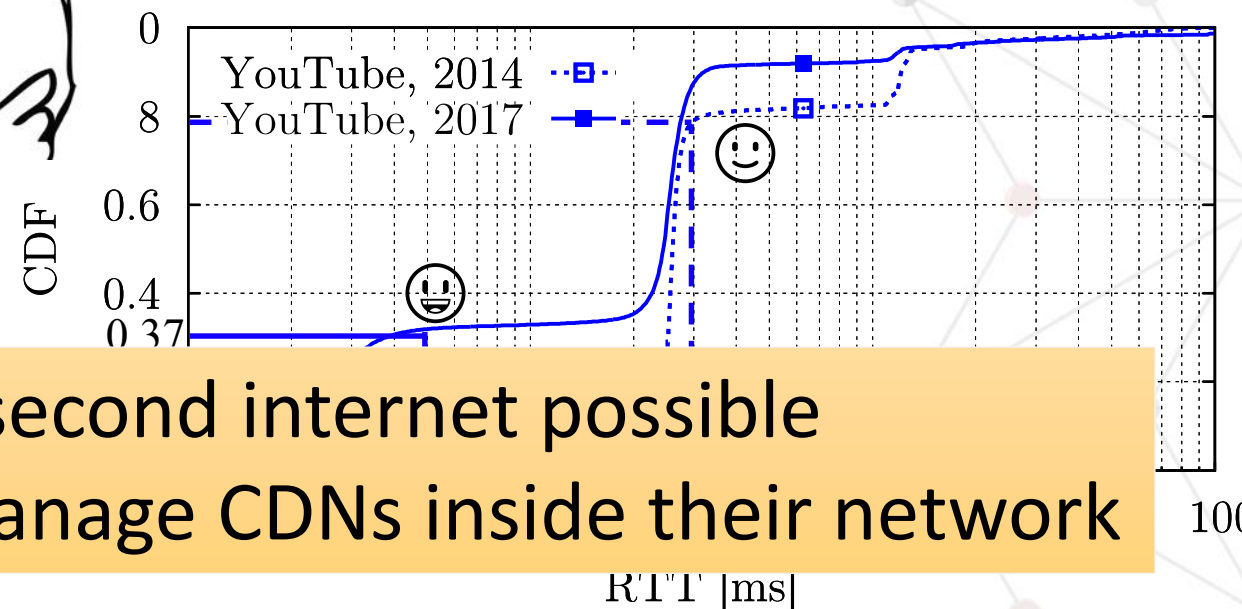
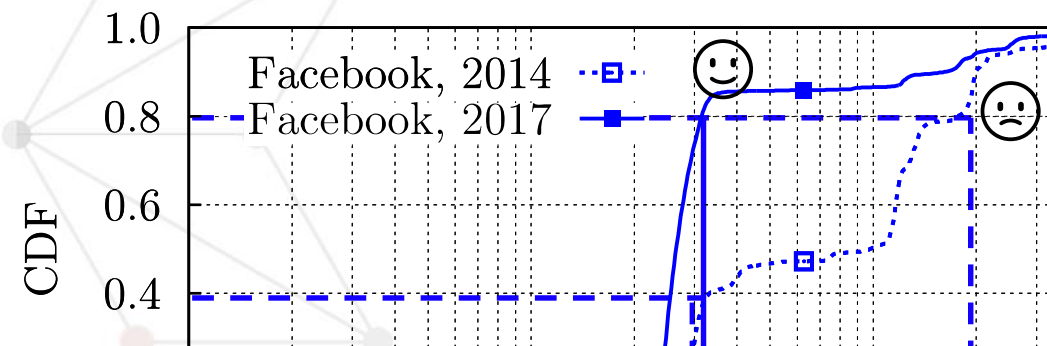


The birth of the sub-millisecond Caching



- Compare RTT Distributions
 - April 2014 vs April 2017

The CDN is inside
the PoP!



CDNs are making sub-millisecond internet possible
New headache for the ISPs: host/manage CDNs inside their network

Conclusion & Future Work



- Q1: Households daily traffic **more than doubled**
- Q2: Quick **rise** of social applications
 - Able to generate massive amount of data
- Q3: **Big players** experiment new protocols (& imposing standards)
- Q4: The infrastructure keeps **concentrating** and **specializing**
- Current and Future Work
 - **Web portal** to make this data playable
 - **New analytics** for...

Perguntas
Fragen Domande Galdera
Otázky
Questions
Spørsmål Pertanyaan kysymykset
Frågor Spørsmål Cwestiynau
вопросы Preguntas Sorular
Въпроси
Vragen
Pytania

What is the traffic load imposed by web services?



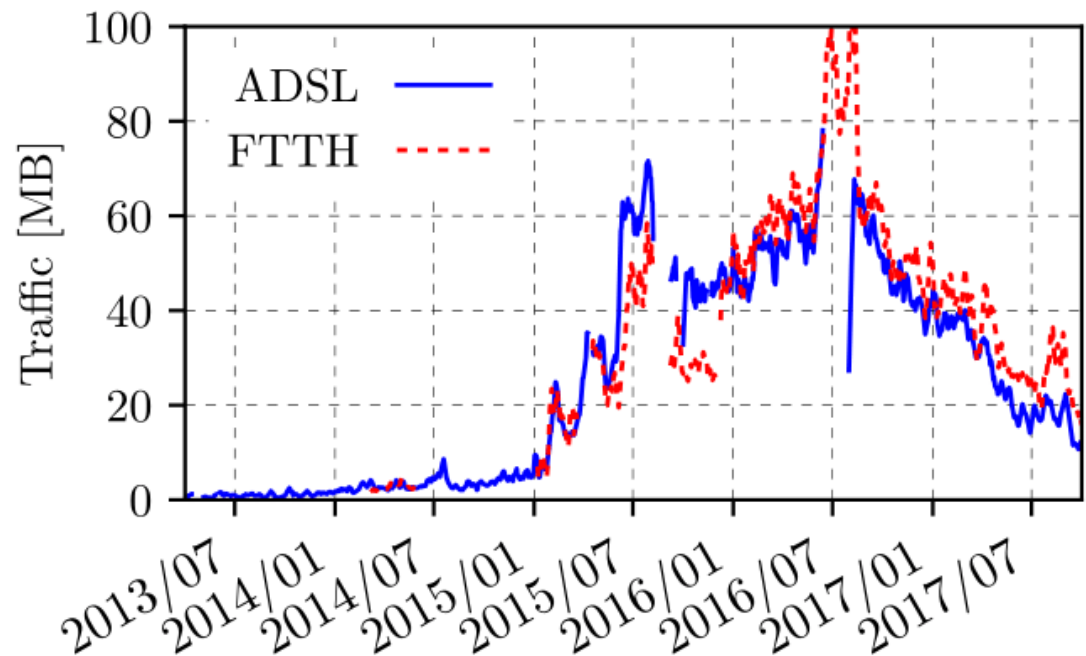
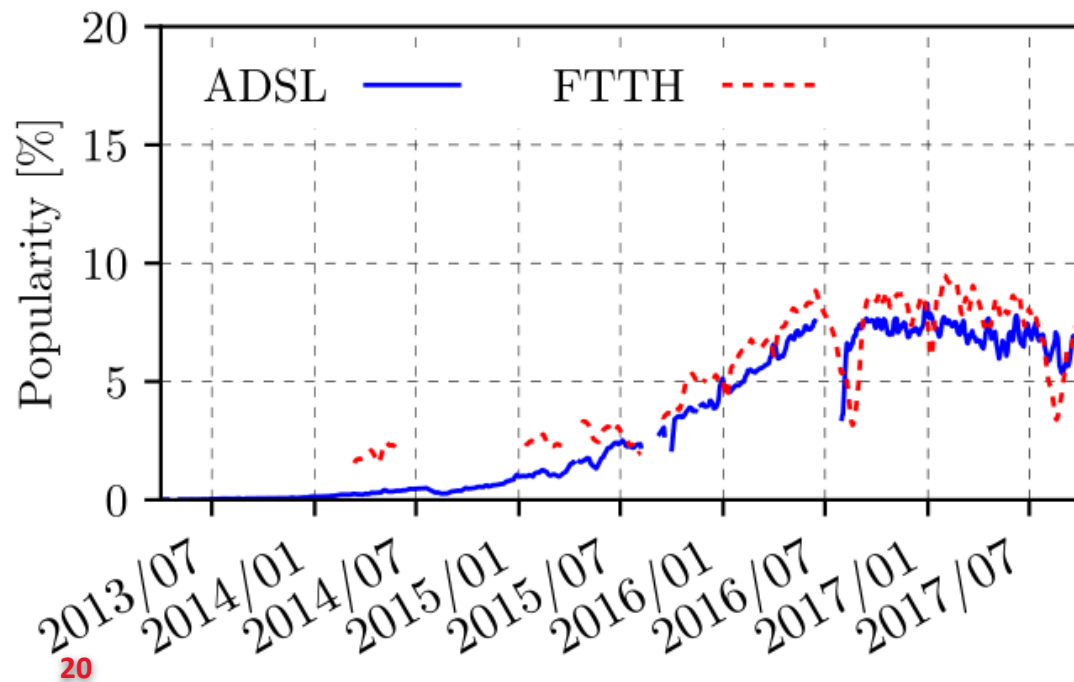
POLITECNICO
DI TORINO



- Popularity **tops** in 2016
- Per-user traffic **drops** mid 2016



snapchat



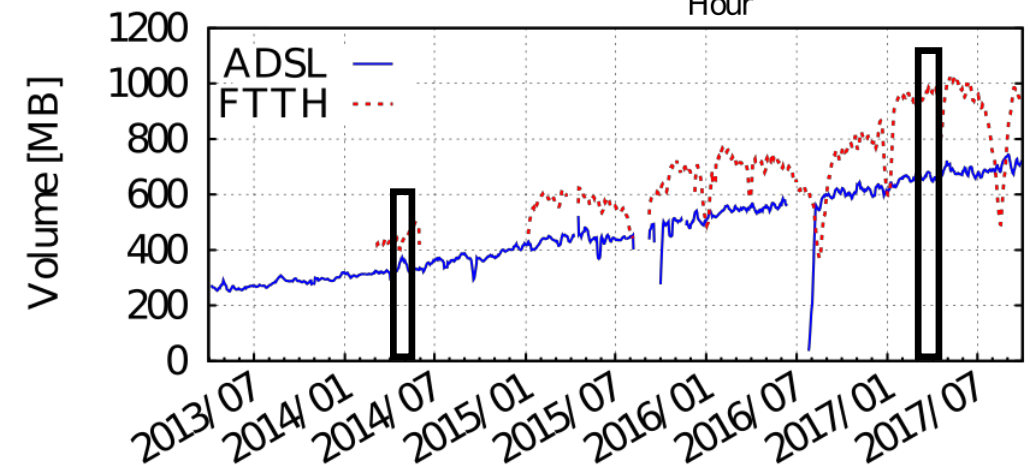
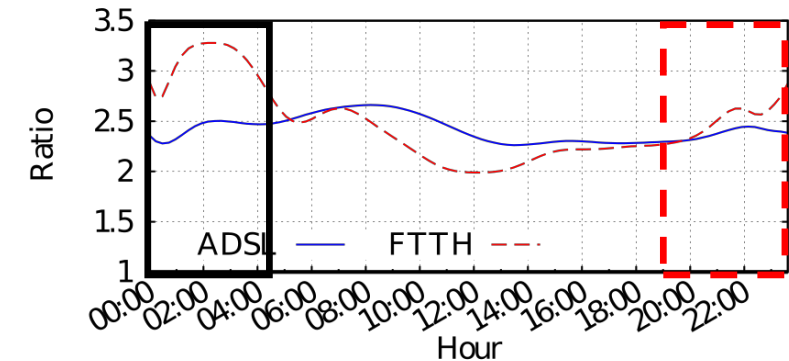
Traffic Consumption Increase



- Average download per user per day
 - Increases at constant rate
- Does it hold at every hour of the day?
- Yes!
 - Late night traffic increases
 - automatic download of updates, machine-generated traffic like IoT devices
 - Larger for **FTTH** due to different profiles
 - During prime time **FTTH** has a higher increase
 - Video streaming content in higher definition



WINDOWS UPDATE



NETFLIX

You Tube

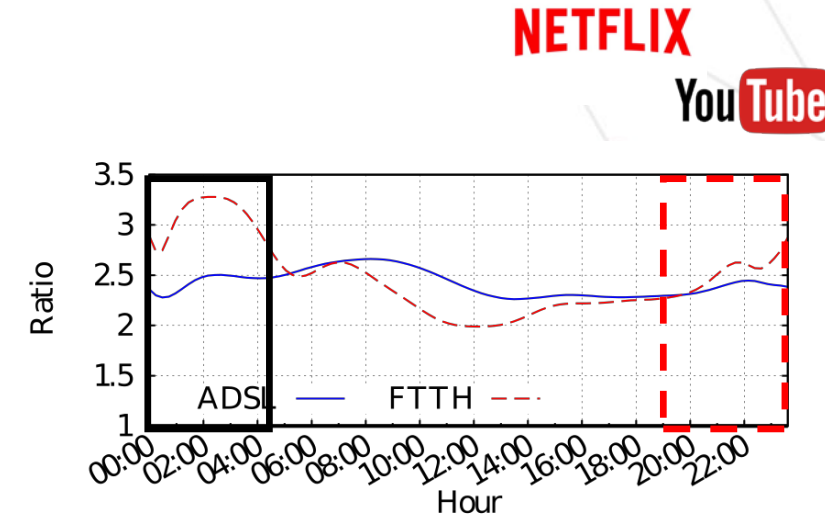
Traffic Consumption Increase



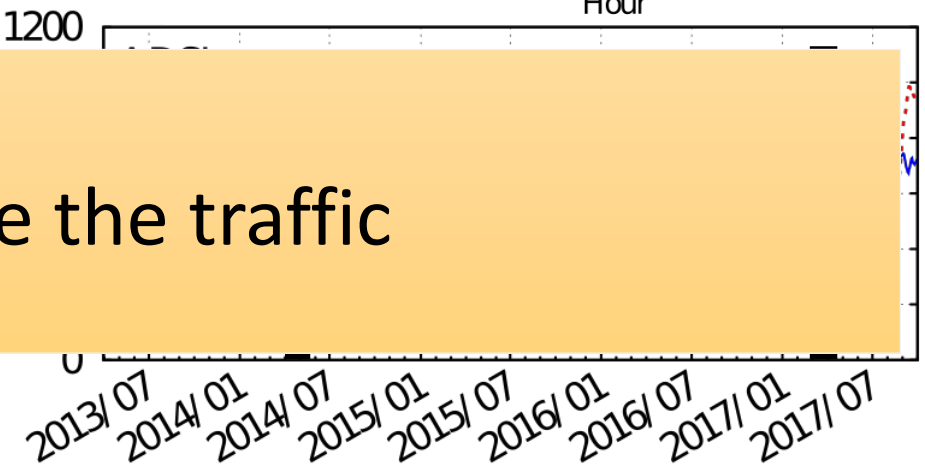
POLITECNICO
DI TORINO



- Average download per user per day
 - Increases at constant rate
- Does it hold at every hour of the day?
- Yes!
 - Late night traffic increases
 - automatic download of updates, machine-generated



The same user consume twice the traffic



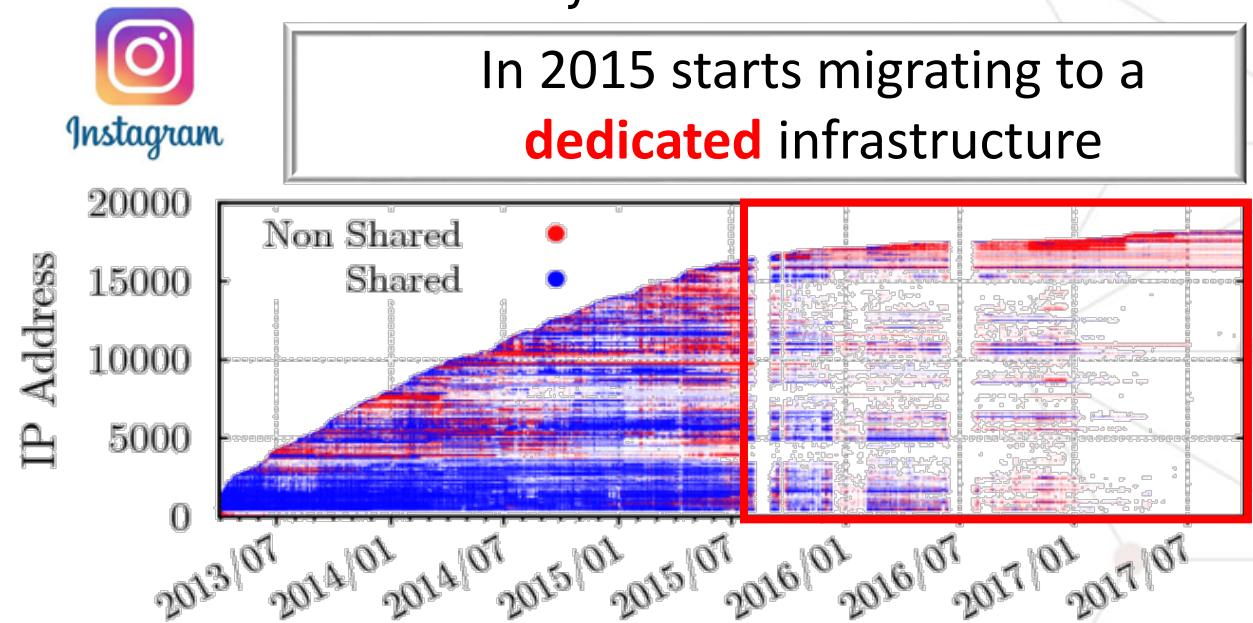
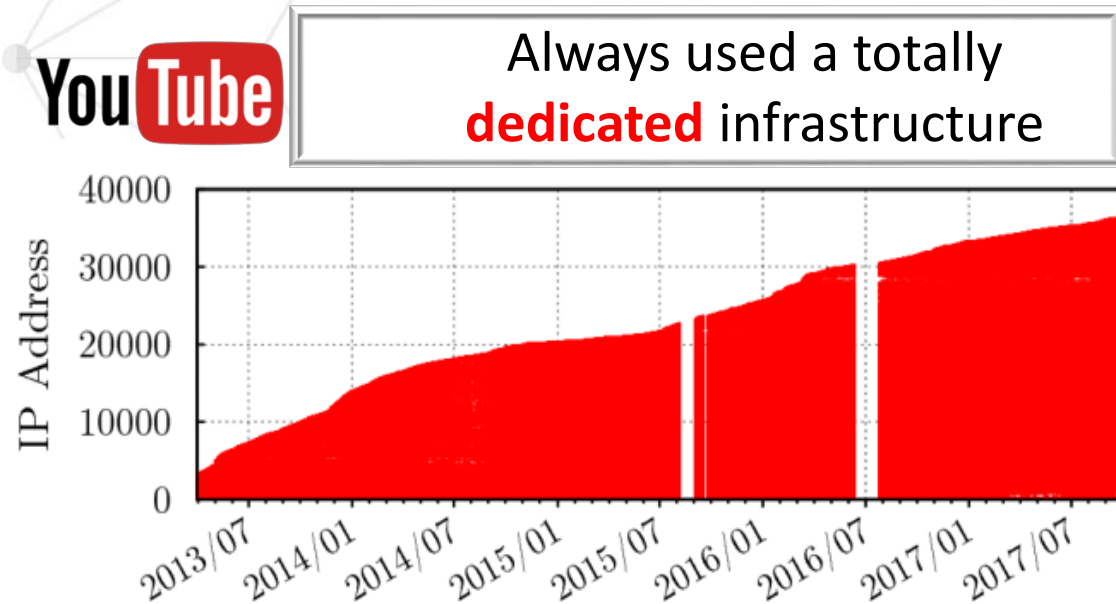
The Server Infrastructure



POLITECNICO
DI TORINO



- Server IP evolution over time
 - **Red:** the IP address was dedicated
 - **Blue:** the IP address
 - no dot: the IP address was not contacted in that day



The Server Convergence



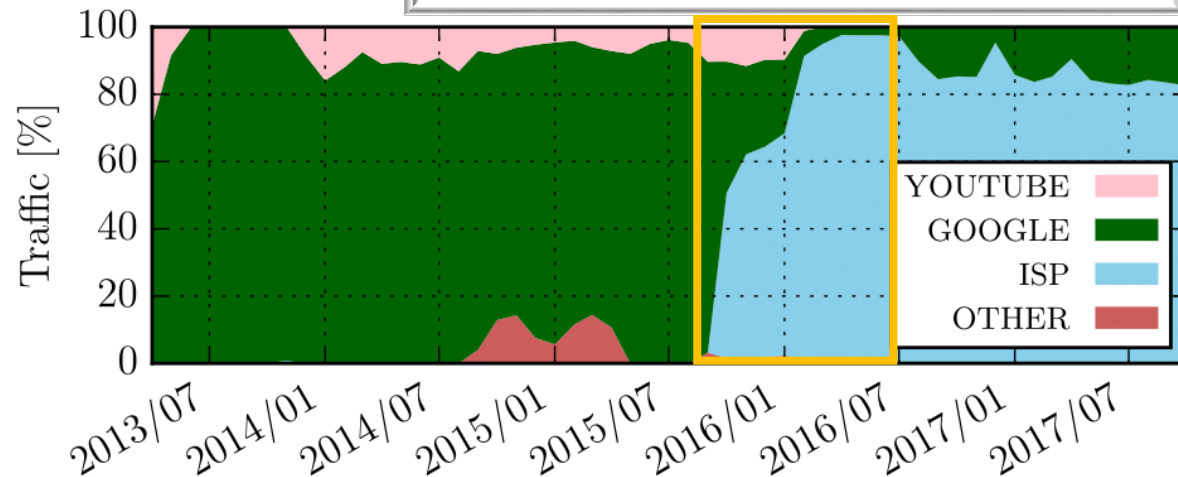
POLITECNICO
DI TORINO



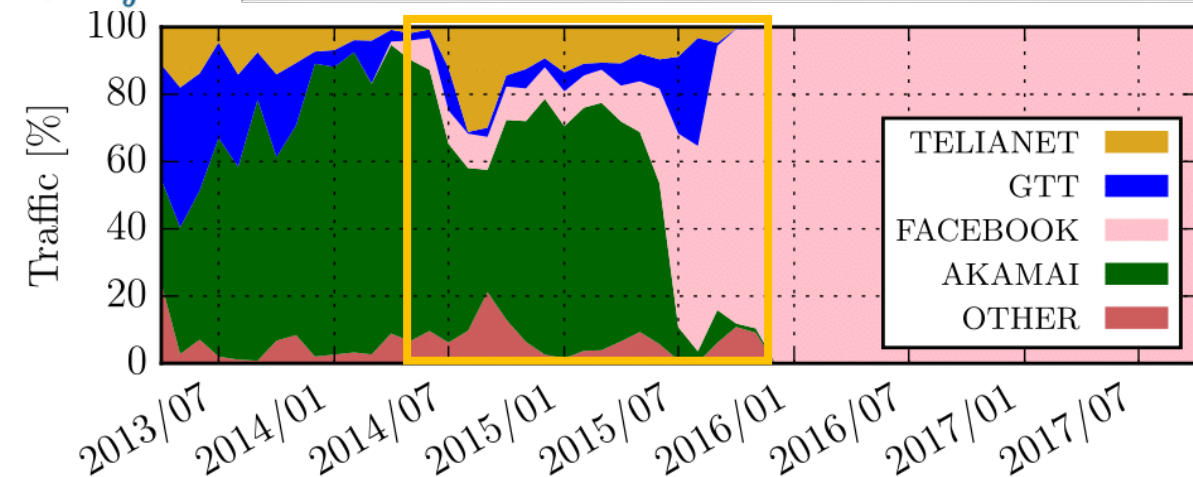
- Which Autonomous Systems are involved?

YouTube

YouTube PoP integration
during 2015



Migration to Facebook infrastructure
started in 2014 ended in 2015



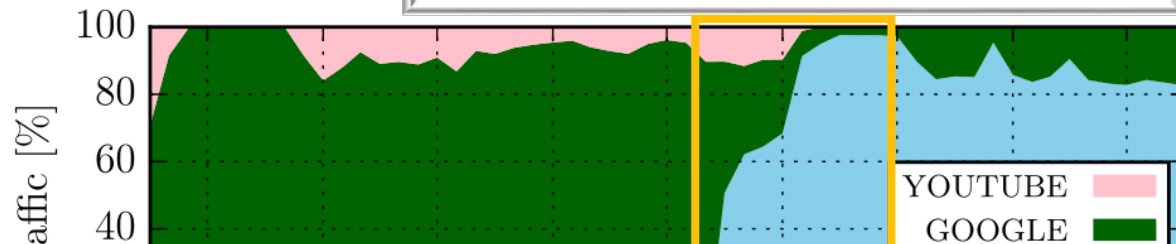
The Server Convergence



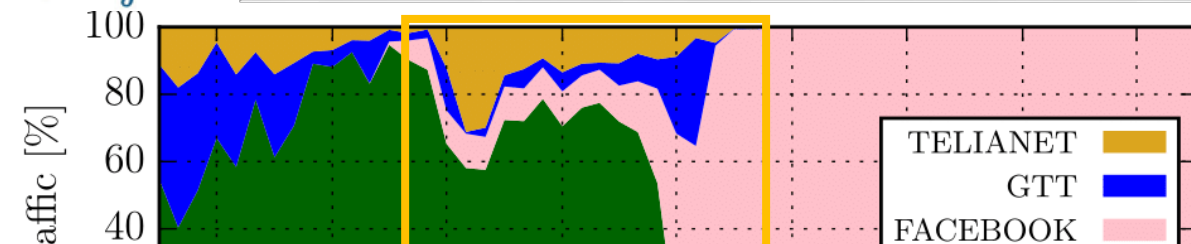
- Which Autonomous Systems are involved?

YouTube

YouTube PoP integration
during 2015



Migration to Facebook infrastructure
started in 2014 ended in 2015



Content providers deploy their own infrastructures
with thousands of IP addresses