Pedro Casas

Telecommunications Research Center Vienna – FTW

Network Traffic Monitoring, Characterization and Analysis in the Internet of Contents

IIE – FING – ARTES

1–5 September 2014

The Internet in the Content Age

- **Today's Internet** = **Internet-scale** (Cloud) Web **Apps**, Content Delivery Networks (**CDNs**) and **mobile devices**
- Internet **contents** and popular apps (Facebook, YouTube, Netflix, WhatsApp) largely **delivered by major CDNs** like Akamai, Google CDN, OpenConnect, SoftLayer, etc.
- **Access to content in mobile networks** has drastically **increased**, and **Quality** has the potential to become a **key differentiator** in a fully covered market
- **Understanding Internet traffic** and how this reach the end customer is highly **valuable for ISPs** (content caching, troubleshooting support, traffic engineering, trend analysis, quality of experience, etc.)

The Internet in the Content Age

- **Today's Internet** = **Internet-scale** (Cloud) Web **Apps**, Content Delivery Networks (**CDNs**) and **mobile devices**
- Interne … **red by major** …
- **Acces** … otential to becc …

This course presents basic concepts of **network traffic monitoring and analysis** to **tackle** different **problems associated to the Internet of todays**

- **Understanding Internet traffic** and how this reach the end customer is highly **valuable for ISPs** (content caching, troubleshooting support, traffic engineering, trend analysis, quality of experience, etc.)

Outline of the Course

- Module 1 – **Network Traffic Monitoring and Analysis**
- Module 2 – **Machine Learning for Network Traffic Analysis**
- Module 3 – **Network Traffic Classification**
- Module 4 – **Quality of Experience in Mobile Networks**
- Module 5 – **Network Traffic Anomaly Detection**

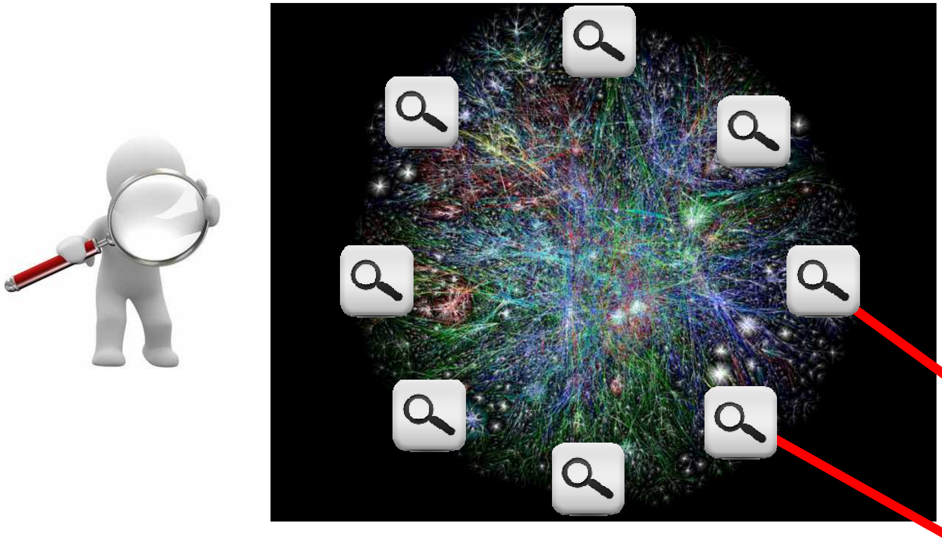
Evaluation of the Course

- **Short-paper** (IEEE 2-columns, 4/6-pages) tackling one or more of the **topics of the course.**
- Traffic traces/measurements publicly avaible @Internet, e.g.,
  - **CAIDA** data (http://www.caida.org/data/overview/)
  - **WIDE** backbone network data (http://mawi.wide.ad.jp/mawi/)
  - **WITS** data (http://wand.net.nz/wits/)
  - **CRAWDAD** data (http://crawdad.cs.dartmouth.edu/)
  - **SPEED.net** data (http://www.netindex.com/)
  - **UMass** Trace Repository (http://traces.cs.umass.edu/)
  - **Simple Web** Traces (http://www.simpleweb.org)
  - and more…or **even your own traffic measurements**

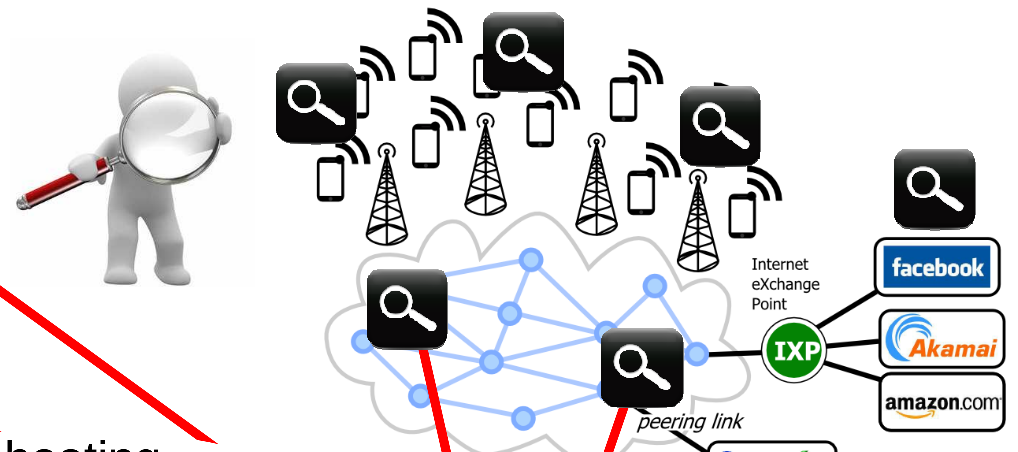
Network Traffic Monitoring and Analysis

ftw Creating Communication Technologies

- The Internet is a complex tangle → **understand how it works** (services, infrastructure, users, performance, etc.)

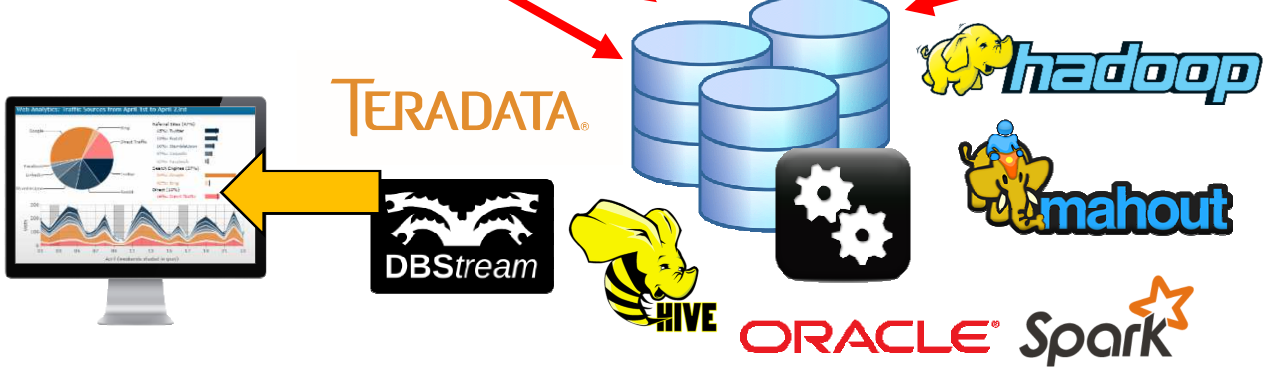


- Internet access is mobile, applications are mobile → **understand mobile traffic**



- Applications span multiple players, troubleshooting requires large number of vantage points, Internet is global → **large-scale, distributed traffic measurements**

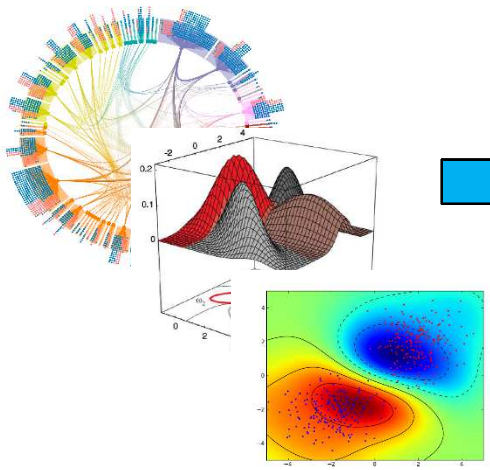
- Heterogeneous data from large number of vantage points (end devices, access network, core network, etc.)→ **platforms for big monitoring data analysis**



Machine Learning for Network Traffic Analysis

- The **value of the traffic measurements** is not on the data itself, but on the **extracted knowledge**

- Large amounts of data, difficult to make sense out of it → **machine learning approaches for data exploration, automation of processes, and knowledge discovery**

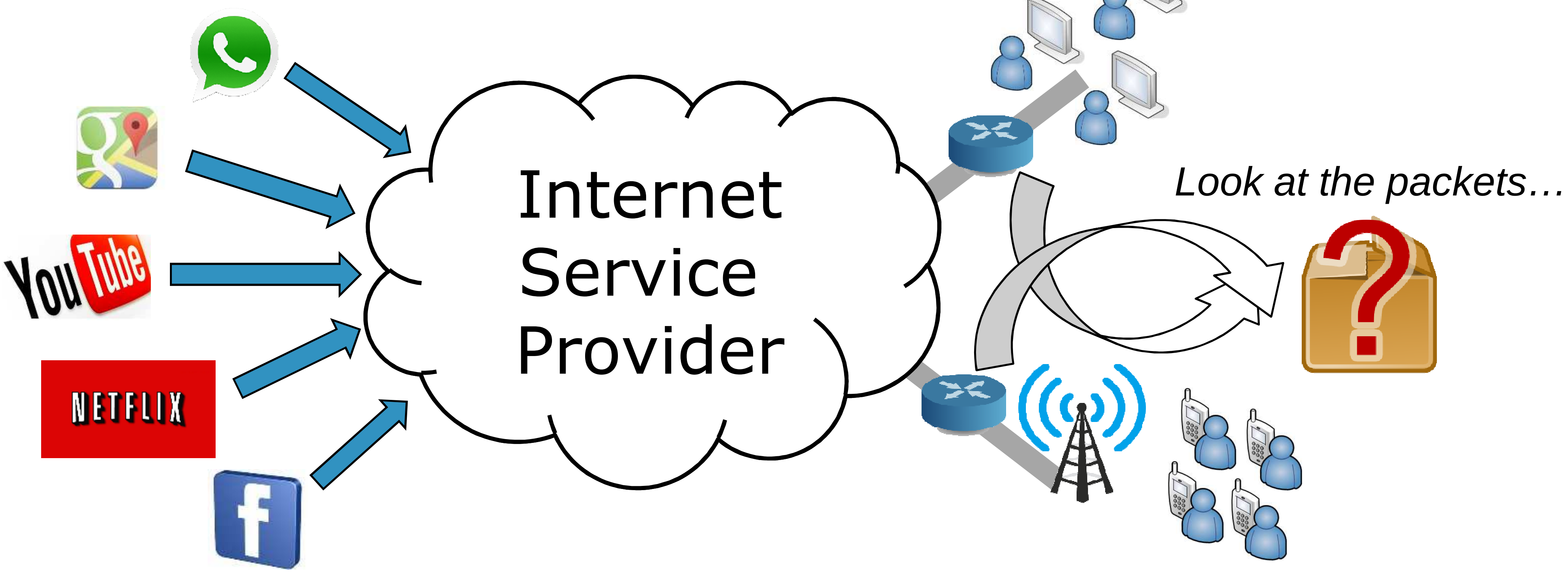


- **Supervised learning**

- **Unsupervised learning**

- **Feature selection/extraction**

Network Traffic Classification



- How to get visibility on the traffic transported through my network? → **automatic traffic classification**

- **many challenges associated** → encryption, obfuscation, OTT providers, proprietary closed implementations, P2P-based apps, HTTP apps through darknets – anonymous networks (e.g., Tor browsing), etc.

Tell me which **protocol** and/or **application** generated them

Quality of Experience in Mobile Networks

- How to measure QoE?
  → **QoE modeling**
- How good is performing my network? → **QoE based monitoring**

- Where to monitor QoE ?
  → **QoE measurements in mobile devices**

- Which is the impact of the network in Web & Cloud services?
  → **Cloud QoE**

QoS

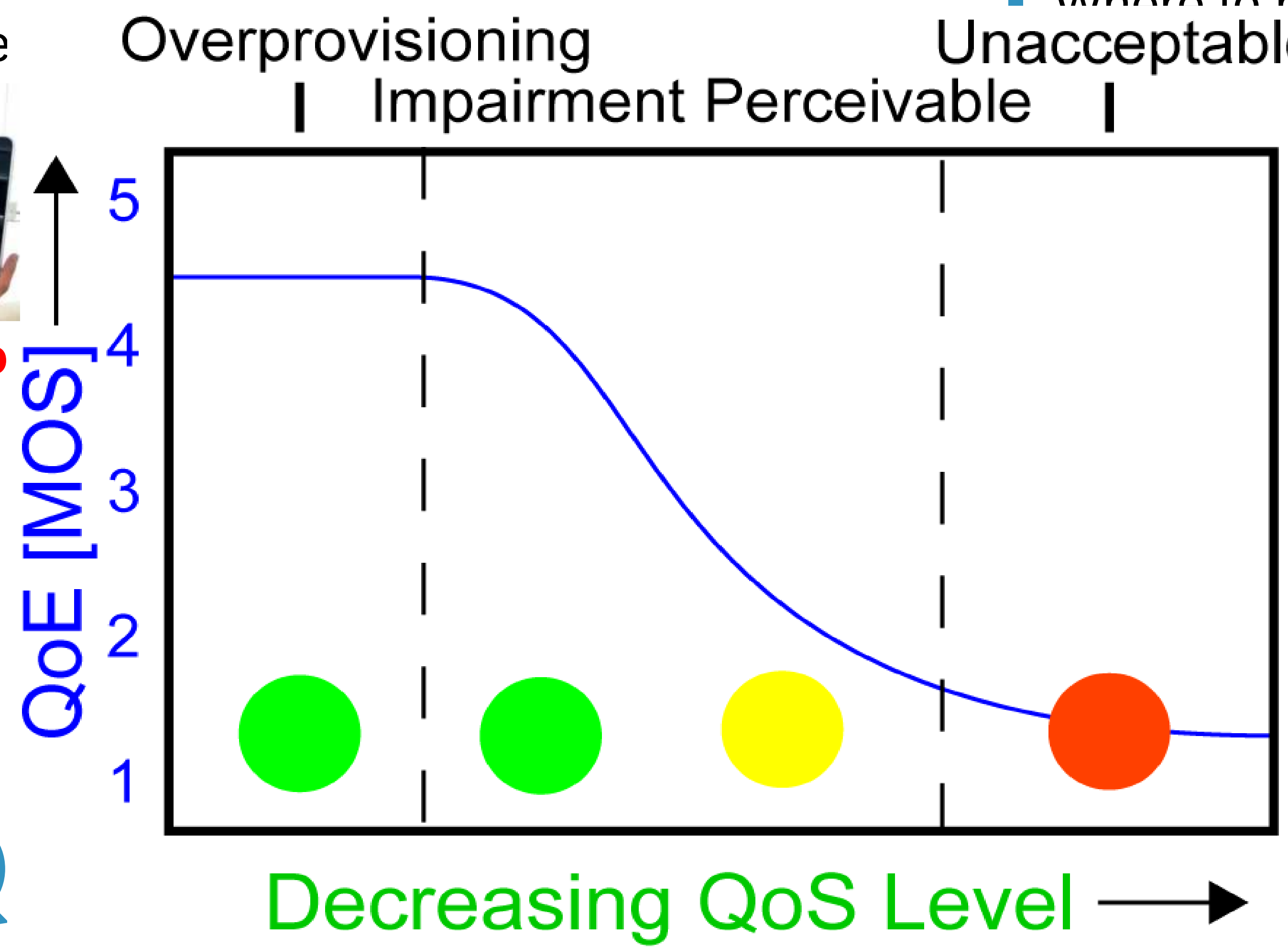
QoE

- **Technical KPIs:**
  - throughput, delay, packet loss
- **User centric KPIs:**
  - what really matters to the end-user
  - responsiveness, interactivity, availability, acceptability, satisfaction

Quality of Experience in Mobile Networks

- How to me

based mo



- Where to monitor QoE ?
asurements
evices

s the impact of
vork in Web &
ervices?
→ **Cloud QoE**

entric KPIs:
- what really matters to the end-user
- responsiveness, interactivity, availability, acceptability, satisfaction

- **Technical KPIs:**
  - throughput, delay, packet loss

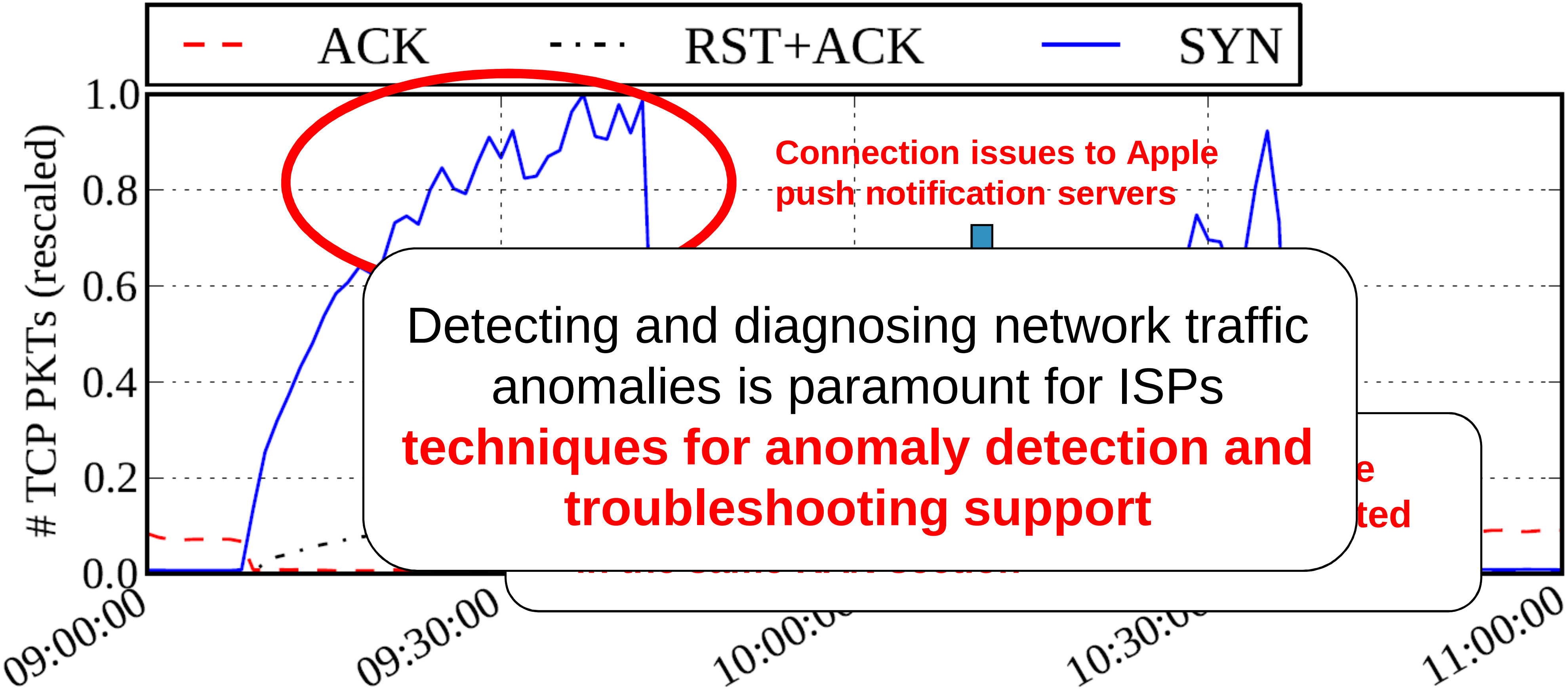
Quality of Experience in Mobile Networks

- **Marketing driver:** intensifying competition in telecom markets

Customer perception and judgement
becoming increasingly relevant

- Avoid **customer churn** for quality dissatisfaction
- Attract new customers with **better service provisioning**
- Understand **what matters the most to customers** for product recommendation

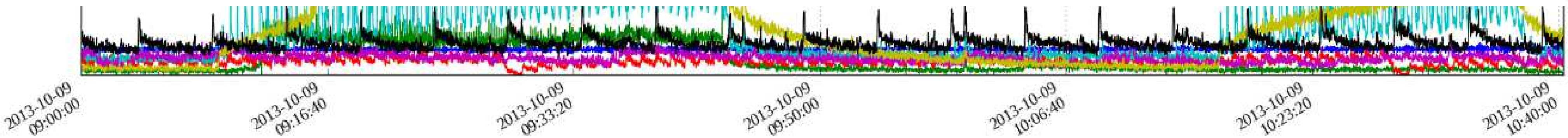
Network Traffic Anomaly Detection



Detecting and diagnosing network traffic anomalies is paramount for ISPs

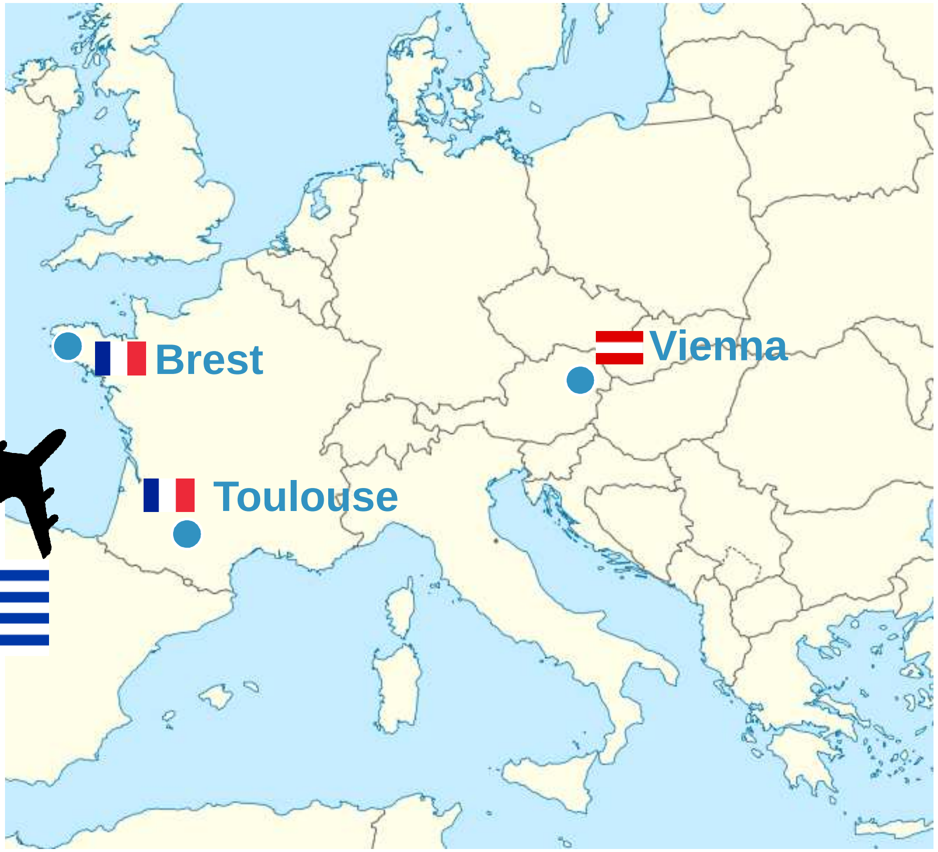
techniques for anomaly detection and troubleshooting support

uplink/downlink TCP packets



iOS7, iOS<7

Short bio



ring

rch Assitant
RTES

2003 – 2010

QoE Assessment in Multimedia Networks Performance Evaluation Traffic Measurements

Network Anomaly Detection and Traffic Estimation

ience

Machine Learning Approaches for Network Security

- CNRS

...

Network Traffic Monitoring and Analysis

Brest (France)

Toulouse (France)

Vienna (Austria)

Austria → Vienna → FTW

- Forschungszentrum Telekommunikation Wien (FTW)

7 Research Topics

- Channel Characterization
- Cross-layer Transceiver Design
- Cooperative Communication
- **Network Monitoring**
- **Quality in Communication Ecosystems**
- Information Exploitation
- Context-Aware Interfaces and Systems

3 Application Fields

- **Telecommunications**
- Transport
- Energy

23 Partners

- 15 Industrial partners
- 8 Academic Partners

Technical Employees

- **65 Researchers**
- 10 Engineers

- International research team with expertise in the management of R&D projects

http://www.ftw.at

Projects I'm currently working on

DARWIN – Data Analysis and Reporting for Wireless Networks

- Started in 2004 → traffic monitoring in mobile networks
- **Partners:** Telekom Austria, A1, Nokia, Technical Univeristy of Vienna
- Implementation of a monitoring system in the mobile network of A1 (8+ M users)
- **Topics:** traffic characterization, troubleshooting support, performance analysis, etc.

ACE – Advancing the Customer Experience

- Started in 2006 → understanding, measuring and managing quality in comnets
- **Partners:** Vodafone, Telekom Austria, A1
- Guidelines for dimensioning and operating mobile networks with improved QoE
- **Topics:** QoE modeling, subjetive lab tests and field trials, QoE–based monitoring

mPlane – an Intelligent Measurement Plane for the Internet

- EU FP7 IP project started in 2012 → Internet scale traffic measurements and analysis
- **Partners:** Telefonica, Telecom Italia, Fastweb, NEC, Alcatel, +8 research insitutions
- Implementation of an Internet-scale traffic measurement and analysis platform
- **Topics:** traffic measurements, big data analysis, machine learning

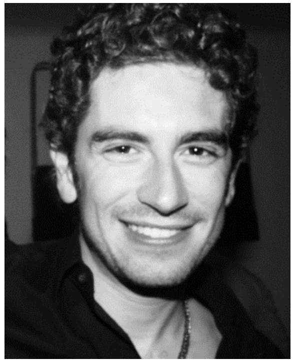
Thanks giving to many colleagues

- *The material presented in these and following slides is also the result of the work of other colleagues in the Traffic Monitoring and Analysis domain:*

Marco Mellia
Politecnico di Torino

Raimund Schatz
FTW

Arian Bär
FTW



Pierdomenico Fiadino
FTW

Ernst Biersack
EURECOM

Alessandro D'Alconzo
FTW

Tobias Hossfeld
Würzburg Universoty

Mirko Schiavone
FTW

Philippe Owezarski
CNRS

Alessandro Finamore
Politecnico di Torino

And what about you?

Outline of Module 1

- *Why Traffic Measurements → **the art of Measurement***
- ***Traffic Monitoring and Analysis**: two types of vantage points to understand and characterize the traffic and the network*
- *Several **Case Studies of Traffic Analysis***
- *mPlane – a platform for **Internet-scale measurements** and traffic analysis*
- ***Big monitoring data** → how to process and anlayze it?*

The Art of Network Measurement

Why Traffic Measurements?

- As **input** for a system design:
  - whenever you build an artifact such as a caching system, VOD service, DNS/name look up service, you need to have a workload model that informs the design
- To **evaluate the performance** of a system:
  - understand performance
  - behavior validation by measurements
  - find security vulnerabilities
- To **identify** normal and **anomalous behaviors**
- To **characterize** the **network** and its **users**
- For **filtering** unwanted traffic
- **To understand Internet traffic**

The Art of Network Measurement
Why Traffic Measurements?

- As **input** for a system design:
  - whenever you build an artifact such as a caching system, VOD service, D... ...d to have a workload model tha...
- To **evaluat**...
  - understa...
  - behavior
  - find secu...
- To **identify** ... **rs**
- To **characteriz**...
- For **filtering** unwanted traffic
- **To understand Internet traffic**

- *Traffic matrix estimation*
- *Topology discovery*
- *Bandwidth estimation*
- *Anomaly detection*
- *Trouble shooting*
- *Traffic classification*
- *etc…*

The Art of Network Measurement

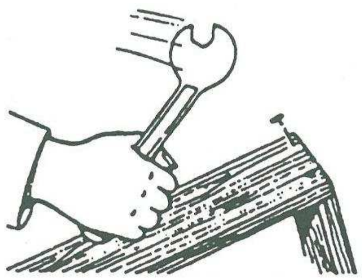
- **Measuring is actually pretty hard**
- **Imperfect** measurement **devices**
  - Data collected is often not complete (data loss, duplication)
- Dealing with a **large volumes of data**
  - Need to **capture** the data, **store** it, perform the **analysis**, etc.
- Misconception: equating **what we are** actually **measuring** with **what we wish** to measure
- **Problem of vantage point**
  - The location of exactly where a measurement is performed can significantly skew the interpretation of the measurement
  - Degree to which individual collections of Internet measurements are often **not representative**

Data Reuse/Misuse

- After nearly three decades of Internet measurement, measurement-based networking research is still a “hot topic” area in science…
- …but **many times**, drawn **conclusions are WRONG!**
- specially when you are a **consumer of measurements done by others**

→ **you may suffer from this in the work you’ll do!**

- If the **original data** gathering was **not “clean”**, the problem is compounded if the consumers were either unaware of it or did not take it into account
- **Even with properly gathered data it is possible for it to be misused by the consumers**



Data Manipulation

- **Manipulating measurements require:**
  - understand the set-up, placement of measurement device, topology
  - must not just collect data but also keep detailed **meta data**.
  - Meta data should encompass **all relevant information about the data**
  - **allows subsequent** assessment of the data fidelity and **usability**

- Meta data typically contain:
  - **what measurement techniques** were used,
  - **conditions of the network** at the time of data gathering, and
  - information about the **location of the data gathering**

Rules for Data Manipulation (1/2)

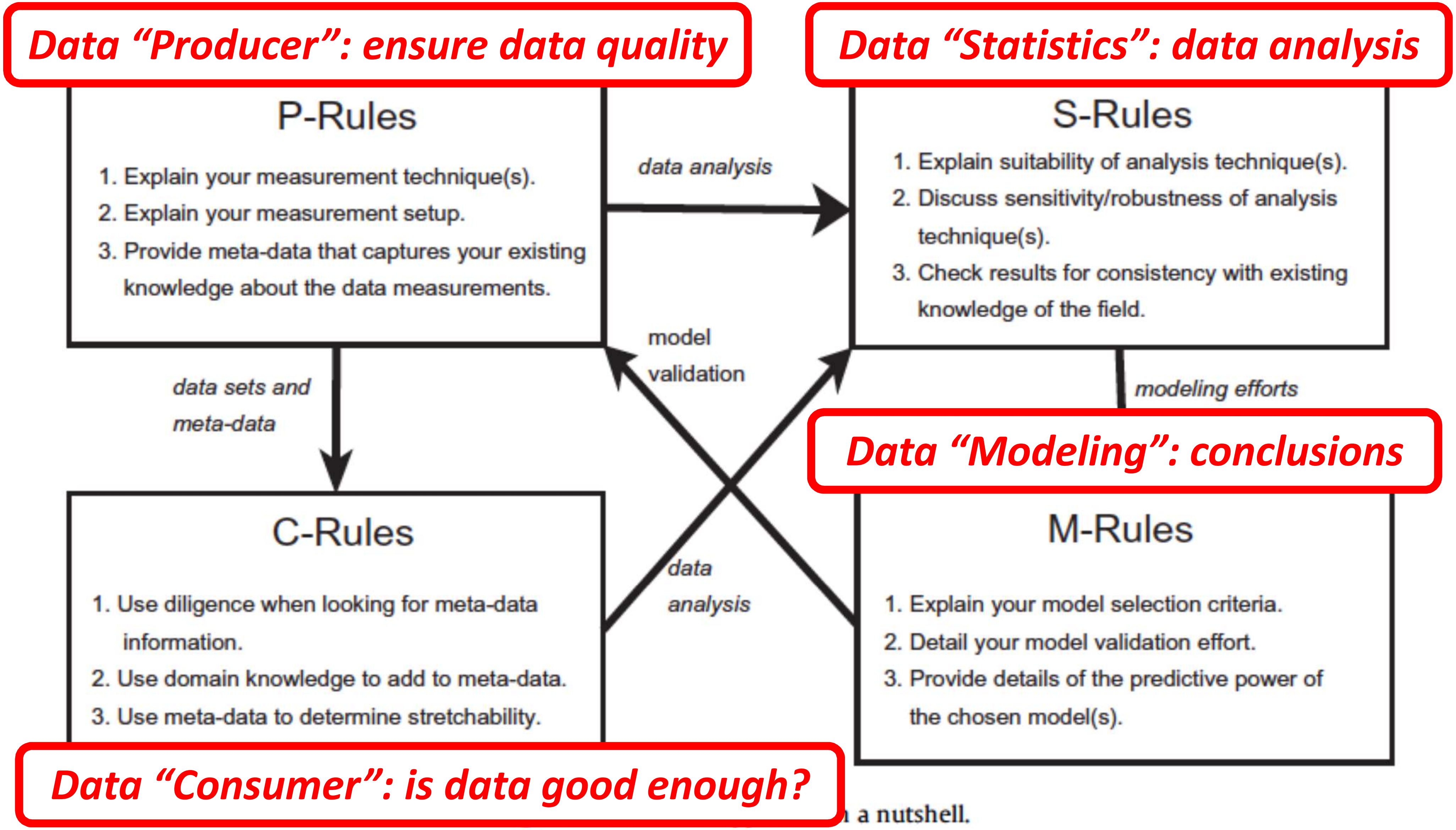
- Despite the maturity of the field, there is a **lack of clearly articulated standards** that **reduce** the probability of **common mistakes** involving measurements, their analysis and modeling.

- A **community-wide effort is likely to foster fidelity in datasets** obtained from measurements and reused in subsequent studies.

- **Rules for how to "manipulate" data**

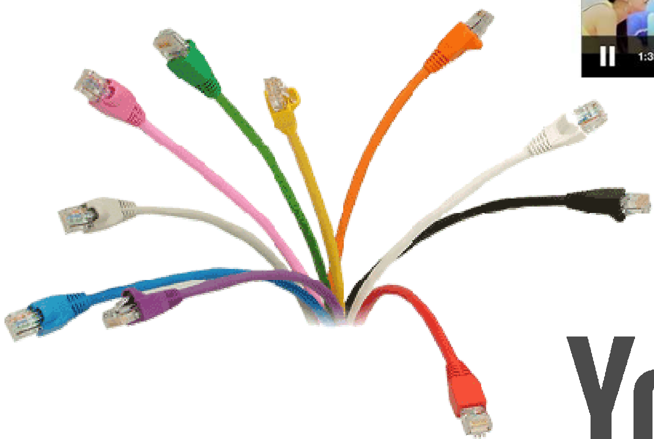
- check the paper "**A Socratic method for validation of measurement-based networking research**", from Bala Krishnamurthy, Walter Willinger et al., @Computer Communications 2011.

Rules for Data Manipulation (2/2)



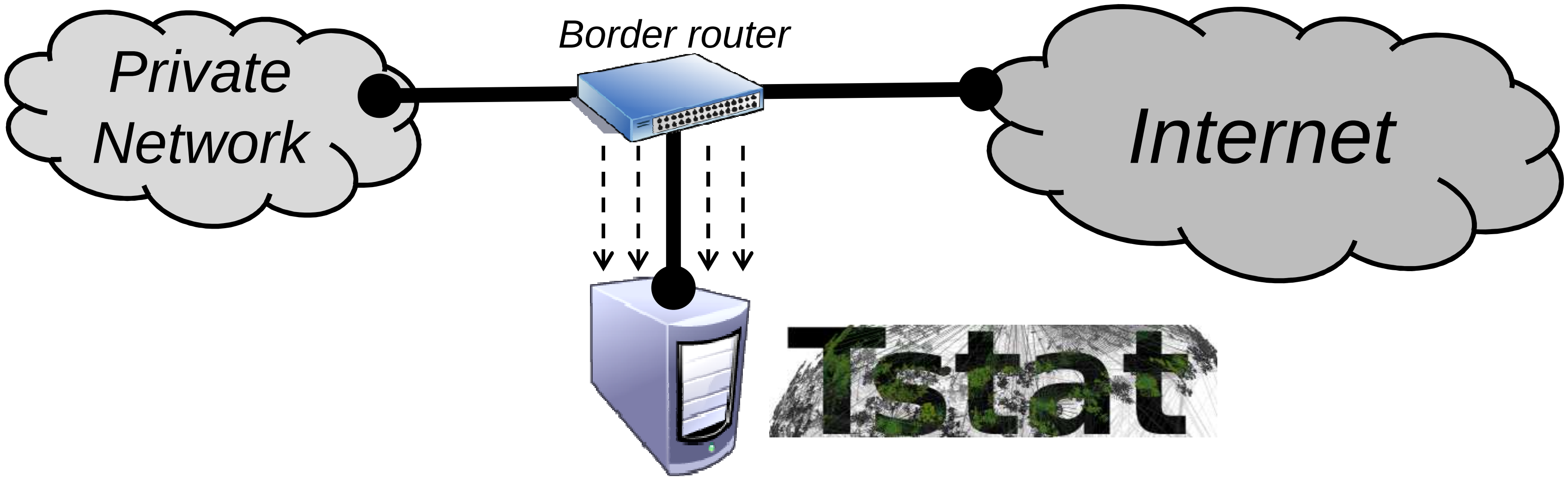
Traffic Monitoring and Analysis

Understand and Characterize the Traffic in Mobile & Fixed-line Networks



Traffic Monitoring and Analysis (TMA)

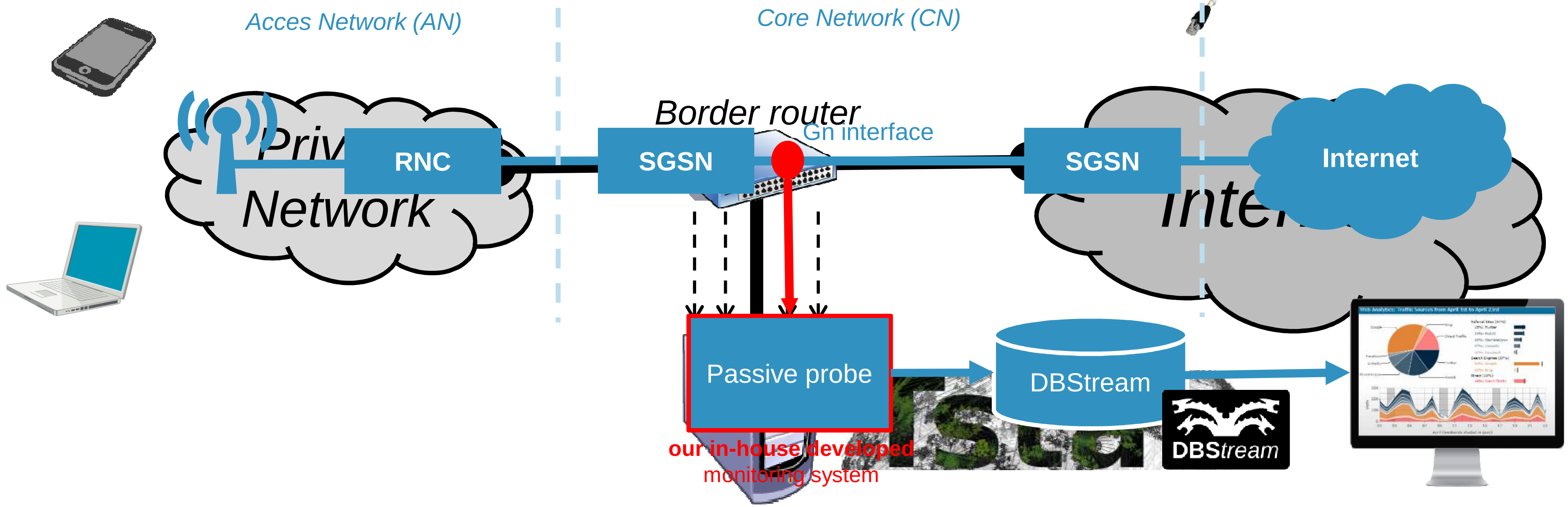
- One of the **biggest challenges for advancing research in TMA** is **accessing real traffic** from a wide variety of large–scale (representative) vantage points
- Two main projects developped in the past 10 years for **monitoring fixed-line** and **mobile networks**



Traffic Monitoring and Analysis (TMA)

- One of the **biggest challenges for advancing research in TMA** is **accessing real traffic** from a wide variety of large–scale (representative) vantage points

- Two main projects developped in the past 10 years for **monitoring fixed-line** and **mobile networks** Tstat and DBStream



Tstat – TCP Statistic and Analysis Tool

- **Open source** tool for network links passive TMA
- Developped by the **TNG group of Politecnico di Torino**
- **Online traffic classification** (DPI, statistical methods)
- Captures and analyzes **traffic flows**, outputs log-dumps and RRD
- Runs using either **common PC hardware** or more sophisticated ad-hoc cards such as **DAG cards**
- **Fixed–line network monitoring** (no 3GPPP stack support)
- **Running in a large number of fixed-line vantage points in EU**

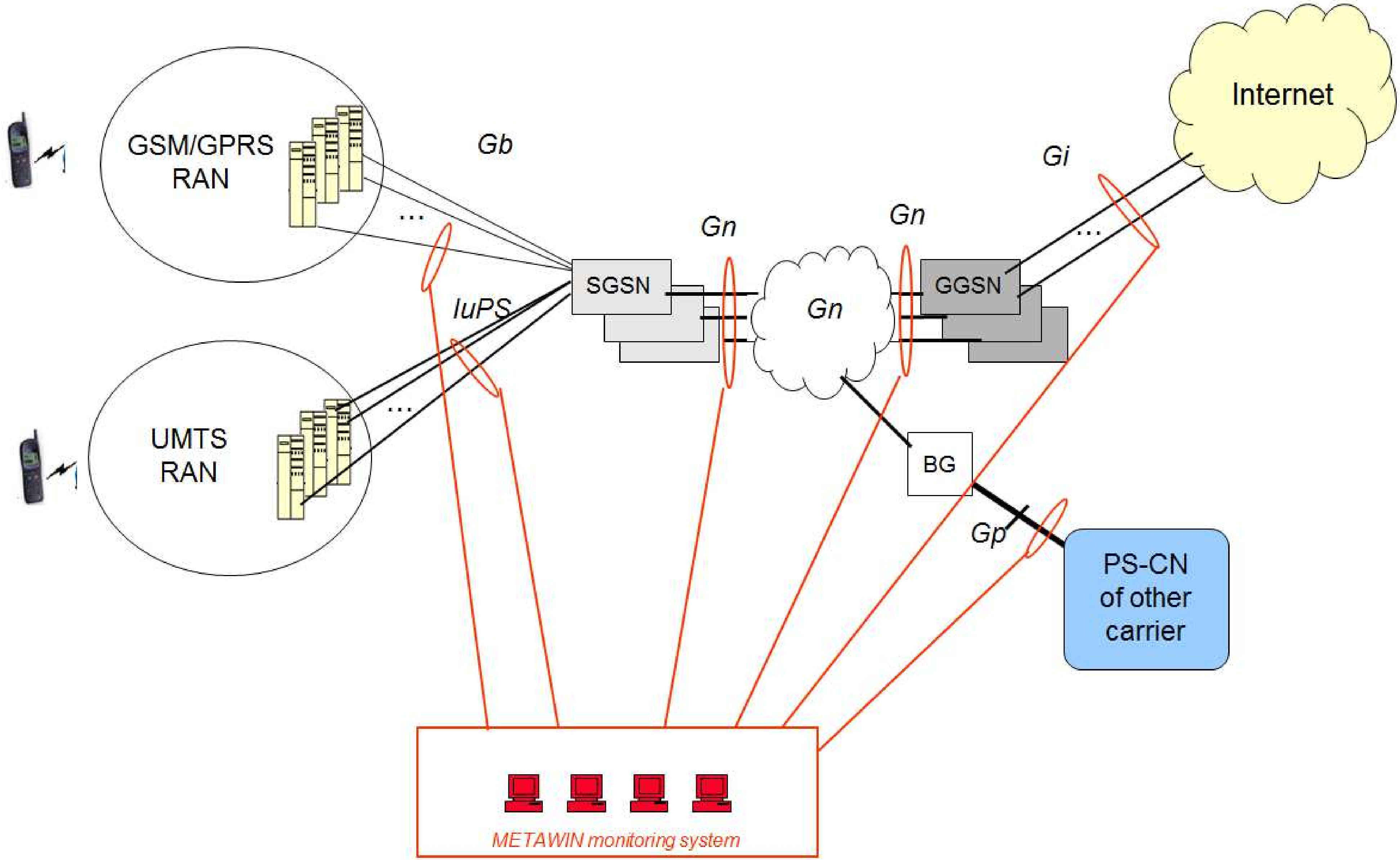
Tstat – TCP Statistic and Analysis Tool

- **Ope**
- Dev
- **Onli**
- Cap
- Run d-hoc
  carc
- **Fixe**
- **Run**

Metawin Probe + DBStream

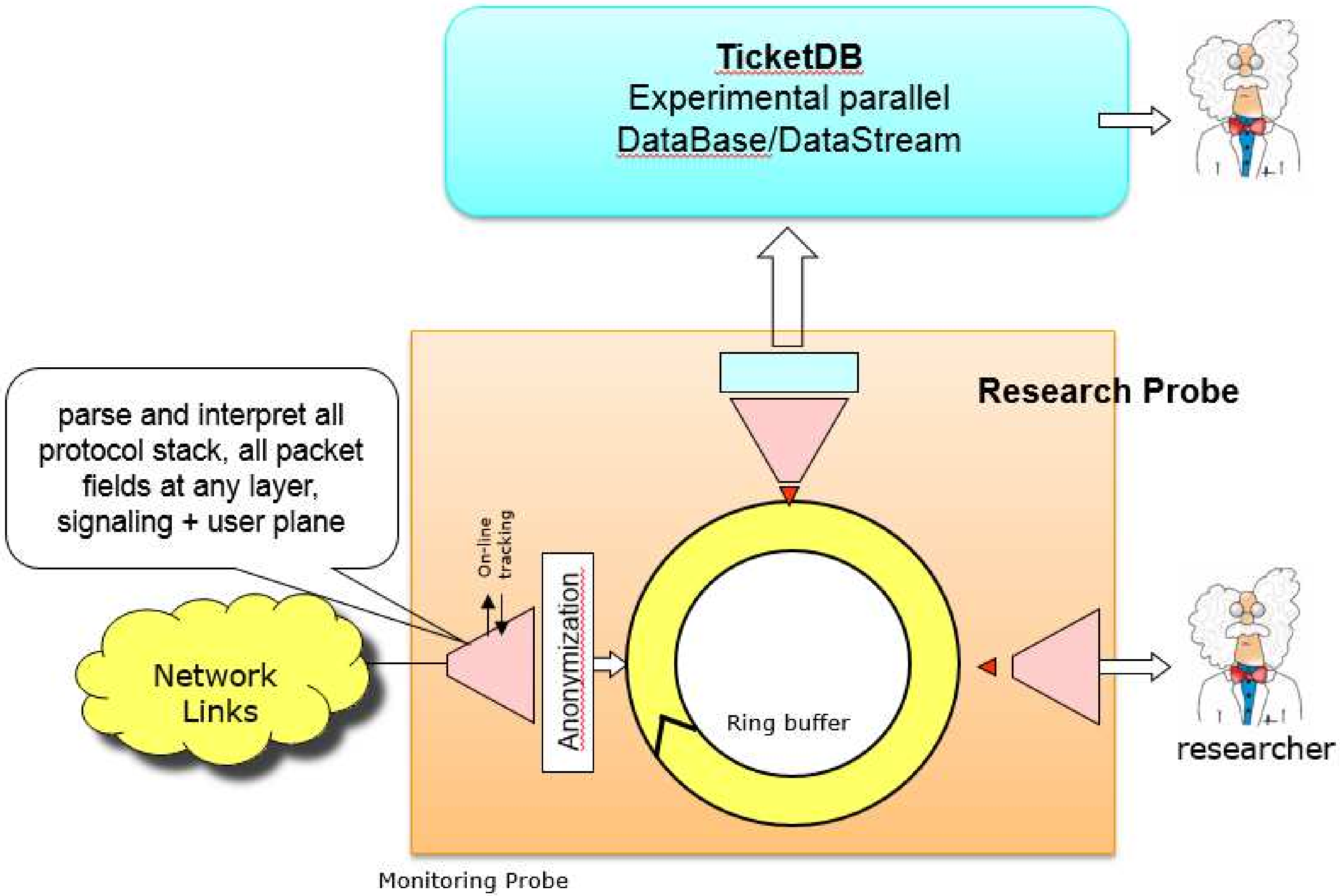
- Tool for network links passive TMA in **mobile networks**
- Developped from scratch by **FTW**
- Includes a **passive probe** and a **Data Stream Warehouse** (DBStream)
- Control-plane and user-plane monitoring
- Full 3GPPP stack support (all 2G/3G/4G core-network interfaces + Iub)
- Captures and analyzes **packets**, local storage of **micro-data** for several days (**full packet copy** plus meta-information)
- Centralized **storage of reduced data (tickets)** in a **DBStream** for several months
- **Real-time tracking of user/terminal data** (IMSI, IMEI, cell location …) and **correlation with user-plane data and payload** (including DPI)
- **Research probe running in operational mobile network from A1**
- **Core component of a commercial monitoring system installed in A1**

Metawin Probe + DBStream



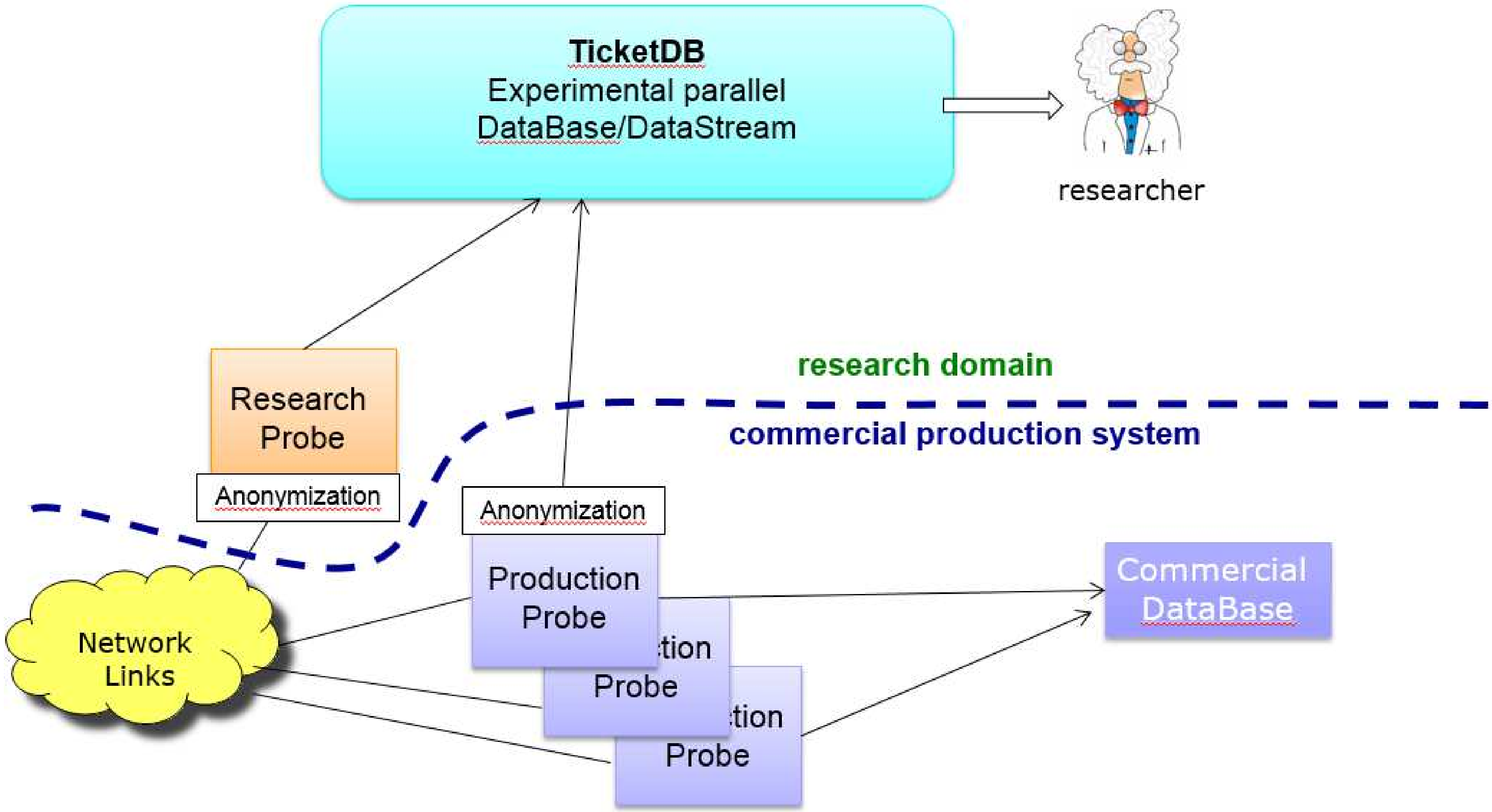
3G TMA @FTW – a bit of history (1/2)

- Original concept → pure research perspective
- **Research monitoring probe + research database**



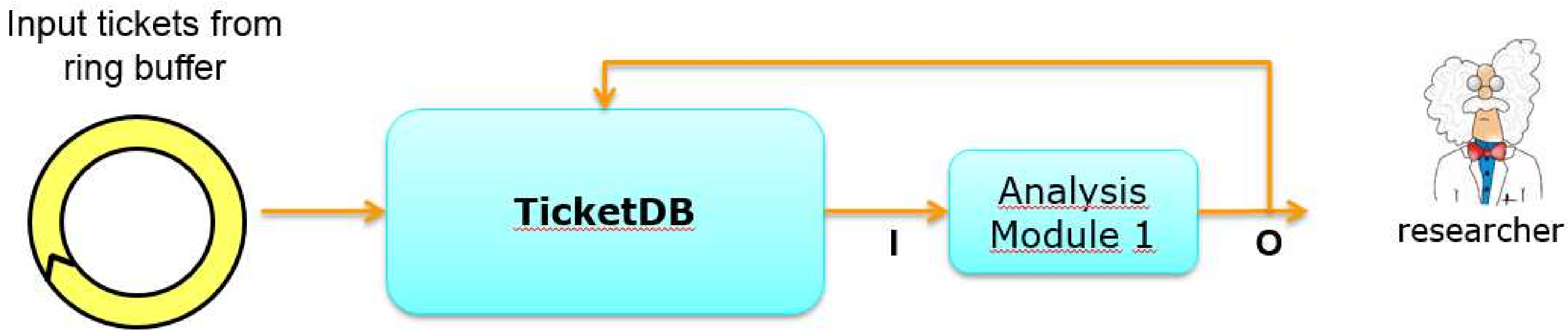
3G TMA @FTW – a bit of history (2/2)

- Evolved into an **hybrid research/commercial system**



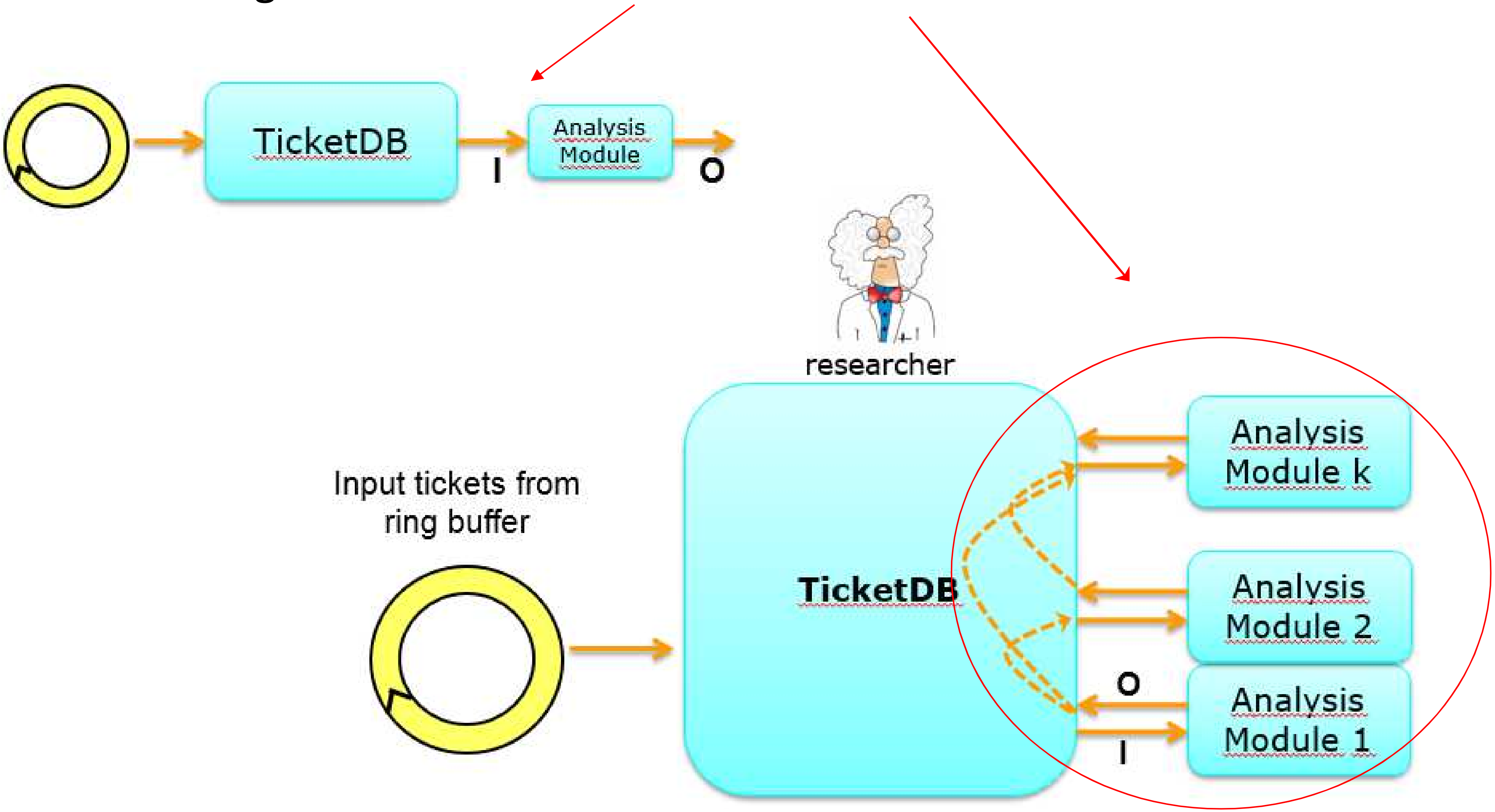
Evolution of the TMA process (1/3)

- Shift from **trace analysis to DB query processing**
- Evolving **off-line to on-line** analysis (quasi-real time)
  - **quasi-real time**: findings are relevant NOW!
  - **possibility to drill-down to packet traces for recent data**
  - allows **historical long-term analysis**
  - easier automation of **recurrent analysis processes**



Evolution of the TMA process (2/3)

- Processing workflow: from ***linear*** to ***network***

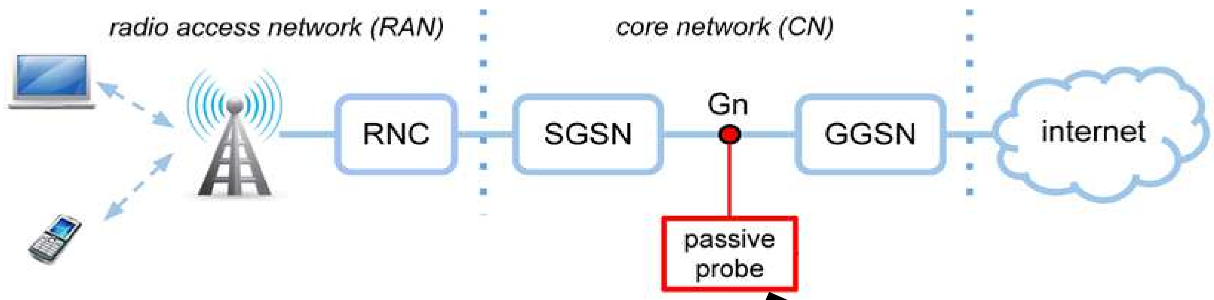


Evolution of the TMA process (3/3)

- The evolution of the processing workflow has evidenced the **need for a novel data management platform**…
- …that combines the two traditional paradigms:
- **datawarehouse** + **datastream** = **DBStream**

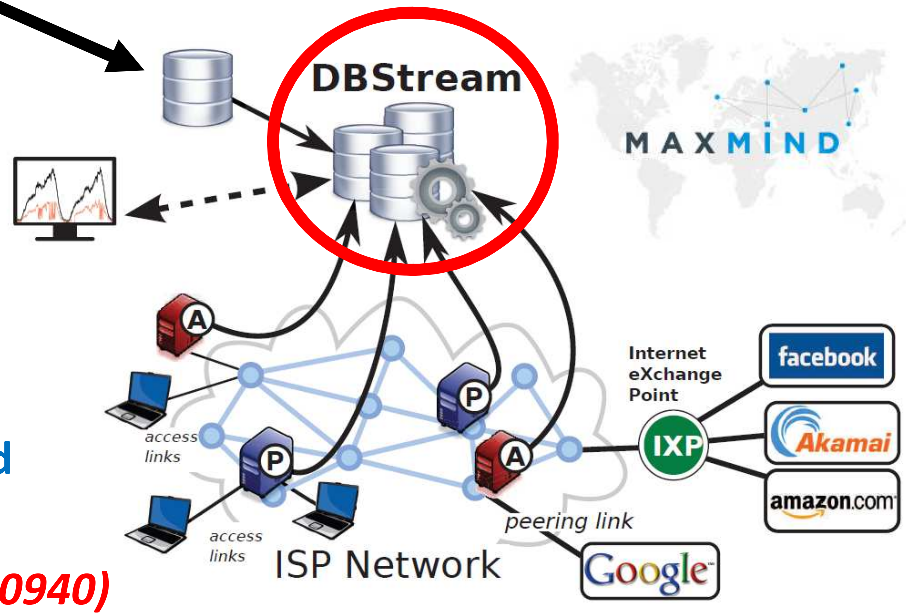
What to do with the Data?

- Nowadays Internet traffic volume is mainly HTTP + P2P



- National-wide **Mobile Network**
- Traffic captured at the **Gn interface**, using METAWIN
- **HTTP flows filtered with HTTPTag** system (module 3 on Wednesday)

- **ADSL/FTTH** links aggregating **40k+ residential customers** in Italy
- Traffic captured and filtered with **Tstat**



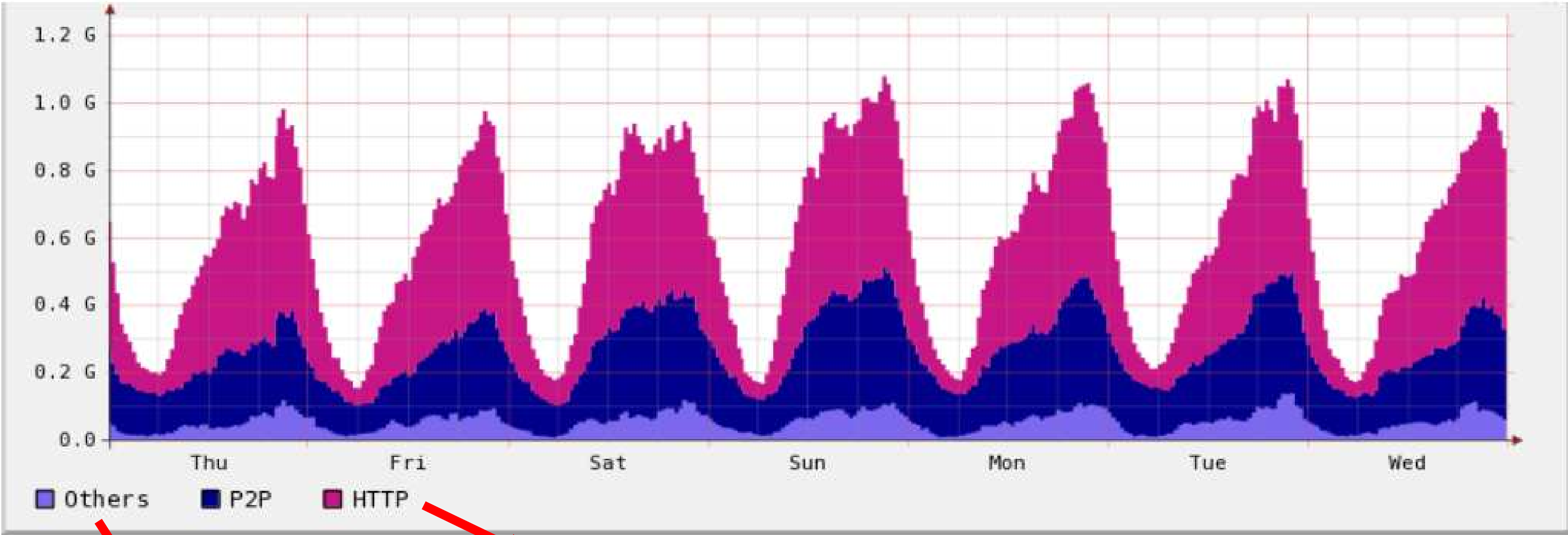
- All **analysis done in the DBStream** system
- Measurements complemented with **MaxMind for IP → AS mappings, e.g.:**

92.122.208.73 → Akamai (AS 20940)

What to do with the Data?

- Nowadays Internet traffic volume is mainly HTTP + P2P

Breakdown of downstream traffic of residential customers

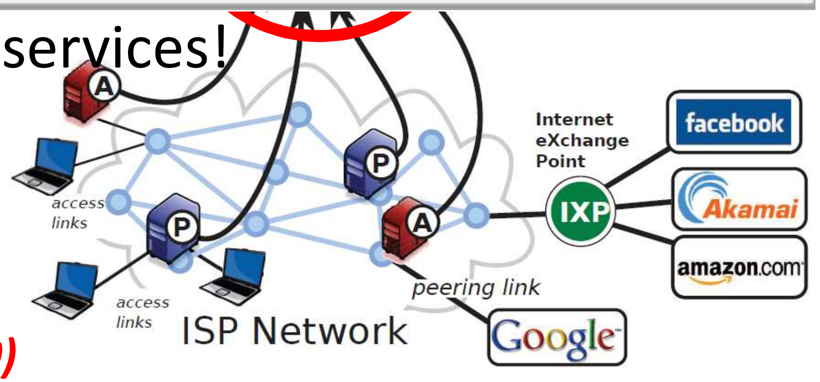


a plethora of services!

SSH, VoIP, DNS, etc.

- All **analysis done in the DBStream system**
- Measurements complemented with **MaxMind for IP → AS mappings, e.g.:**

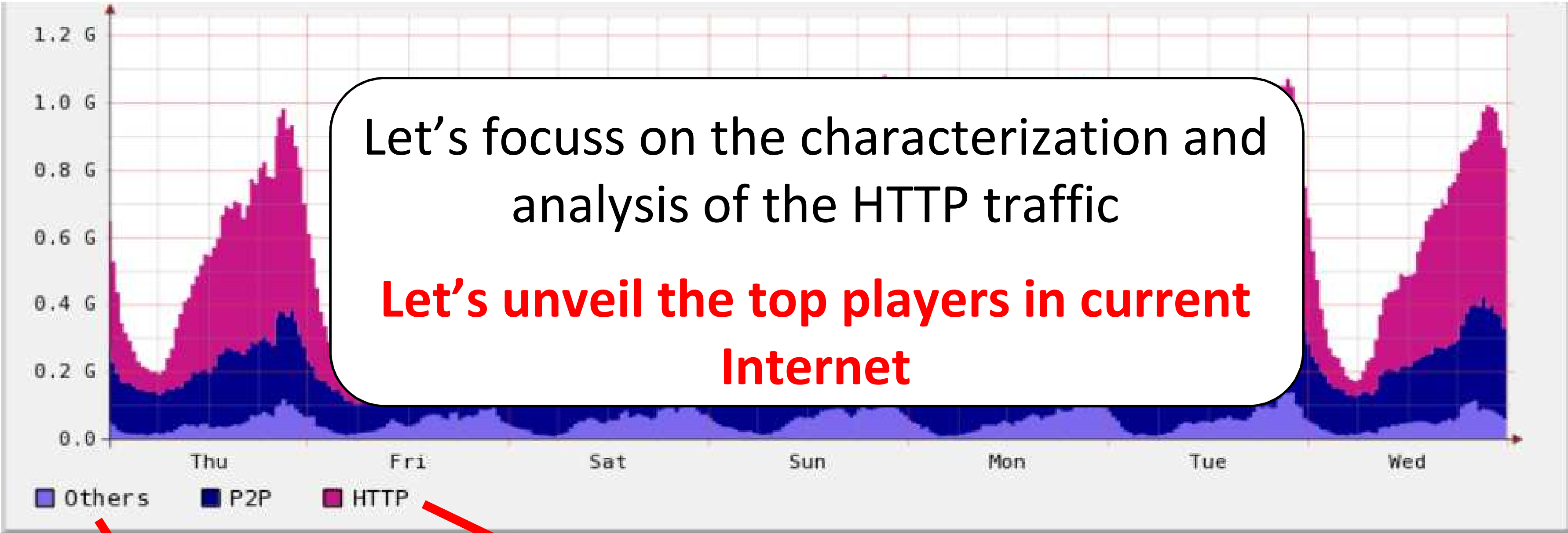
92.122.208.73 → Akamai (AS 20940)



What to do with the Data?

- Nowadays Internet traffic volume is mainly HTTP + P2P

Breakdown of downstream traffic of residential customers



Let's focuss on the characterization and analysis of the HTTP traffic

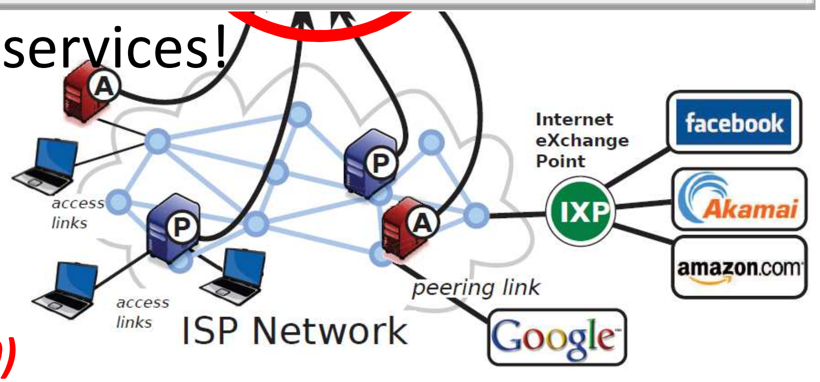
Let's unveil the top players in current Internet

a plethora of services!

SSH, VoIP, DNS, etc.

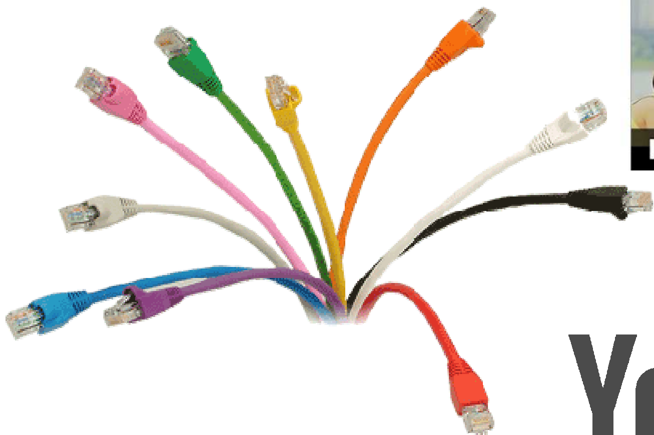
- All **analysis done in the DBStream** system
- Measurements complemented with **MaxMind for IP → AS mappings, e.g.:**

92.122.208.73 → Akamai (AS 20940)



Traffic Monitoring and Analysis

The big players in the Internet
A view from mobile and fixed-line networks



A view from a fixed-line network

- We shall use an off-line dataset collected at 3 vantage points of an ISP in Italy, using Tstat
- Residential customers, 2 weeks of data
  - FTTH (VP1)
  - ADSL access (VP2, VP3)
  - 20-24 June 2011 and 1-7 April 2012

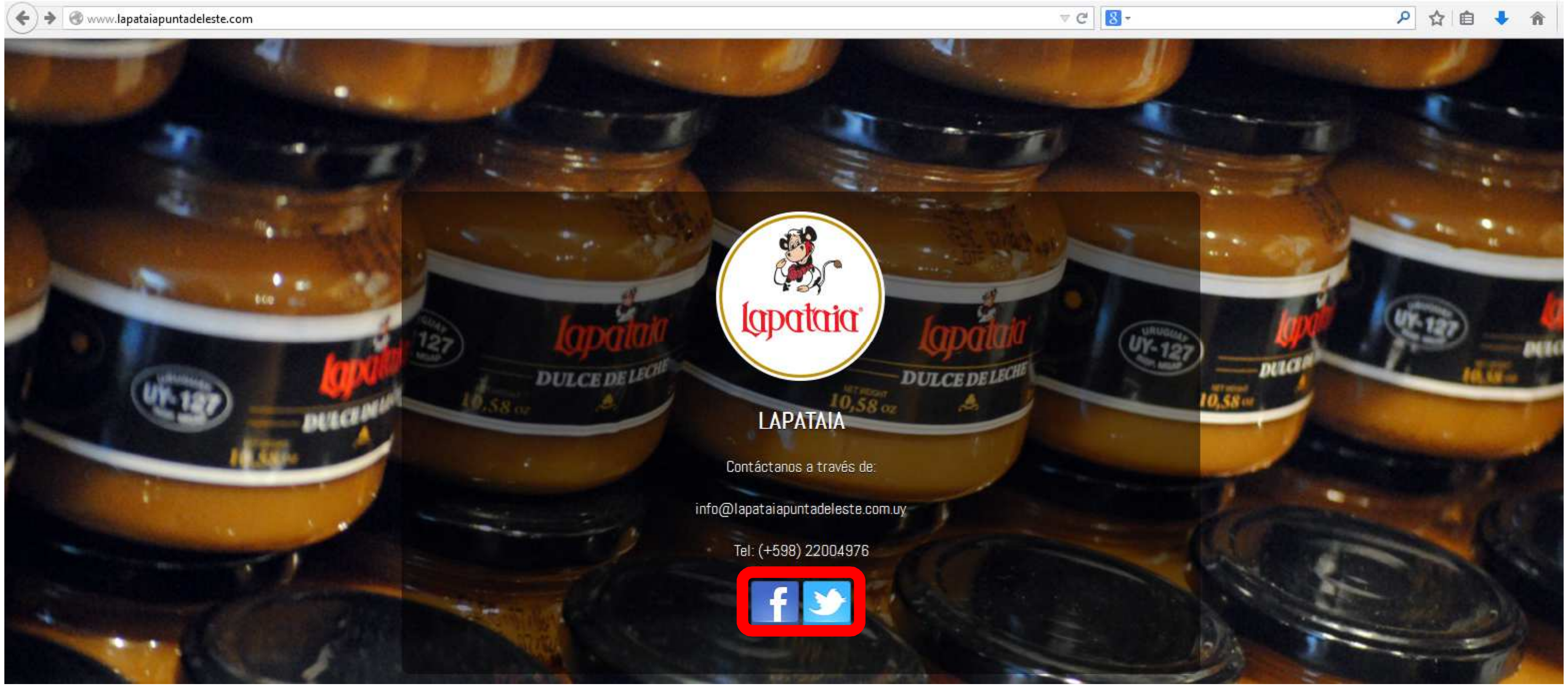
Name	Volume [GB]	Flow [M]	# Servers	# Clients
VP1	1745 (35%)	16 (63%)	77,000 (0.14%)	1534 (99%)
VP2	10802 (44%)	84 (53%)	171,000 (0.6%)	11742 (97%)
VP3	13761 (35%)	125 (52%)	215,000 (0.5%)	17168 (98%)

Top players hosting HTTP contents

- 65% of the HTTP volume is hosted by 11 organizations

	Volumes			Most known services		
Organization	%B	%F	%Clients	Video Content	SW Update	Adv. & Others
Google	22.7	12.7	97.1	YouTube	-	Google services
Akamai	12.3	16.7	97.2	Vimeo	Microsoft, Apple	Facebook static content, eBay
Leaseweb	6.3	1.1	64.3	Megavideo	Mozilla	publicbt.com
Megaupload	5.5	0.2	15.6	Megavideo	-	File hosting
Level3	4.7	1.9	79.7	YouPorn	-	quantserve, tinypic, photobucket
Limelight	3.9	1.6	72.5	Pornhub, Veoh	Avast	betclick, wdig, trafficjunky
PSINet	3.2	0.2	44.6	Megavideo	Kaspersky	Imageshack
Webzilla	2.9	0.3	13.2	Adult Video	-	filesonic, depositfiles
Choopa	1.5	0.01	5.7	-	-	zShare
OVH	1.0	0.7	63.1	Auditude	-	Telaxo, m2cai
Facebook	0.9	4.2	90.6	Facebook	-	Facebook dynamic content
total	64.9	39.6	-			

Why Facebook sees 91% of customers?



- I want to eat something sweet → visit ***http://www.lapataiapuntadeleste.com/***
- There is an **embedded object pointing to FB page** of Lapataia
- So there is a connection to FB → **FB knows that I like "dulce de leche"**
  → privacy???

Privacy Issues?

- facebook -> HTTPS
- twitter -> HTTPS
- Google France -> HTTPS
- YourFavouriteSite -> HTTPS
- …
- This is to protect your **privacy**…

… but then why the facebook app on iOS uses HTTP?!?!?

Content hosting Evolution

- The scenario is in ***constant evolution***

June-11

Rank	Organization	Bytes
1	Google	22.7%
2	Akamai	12.3%
3	Leaseweb	6.3%
4	Megaupload	5.5%
5	Level3	4.7%
6	Limelight	3.9%
7	PSINet	3.2%
8	Webzilla	2.9%
9	Choopa	1.5%
10	OVH	1.0%
11	Facebook	0.9%
12	Zynga	0.01%
	Total	***64.9%***

April-12

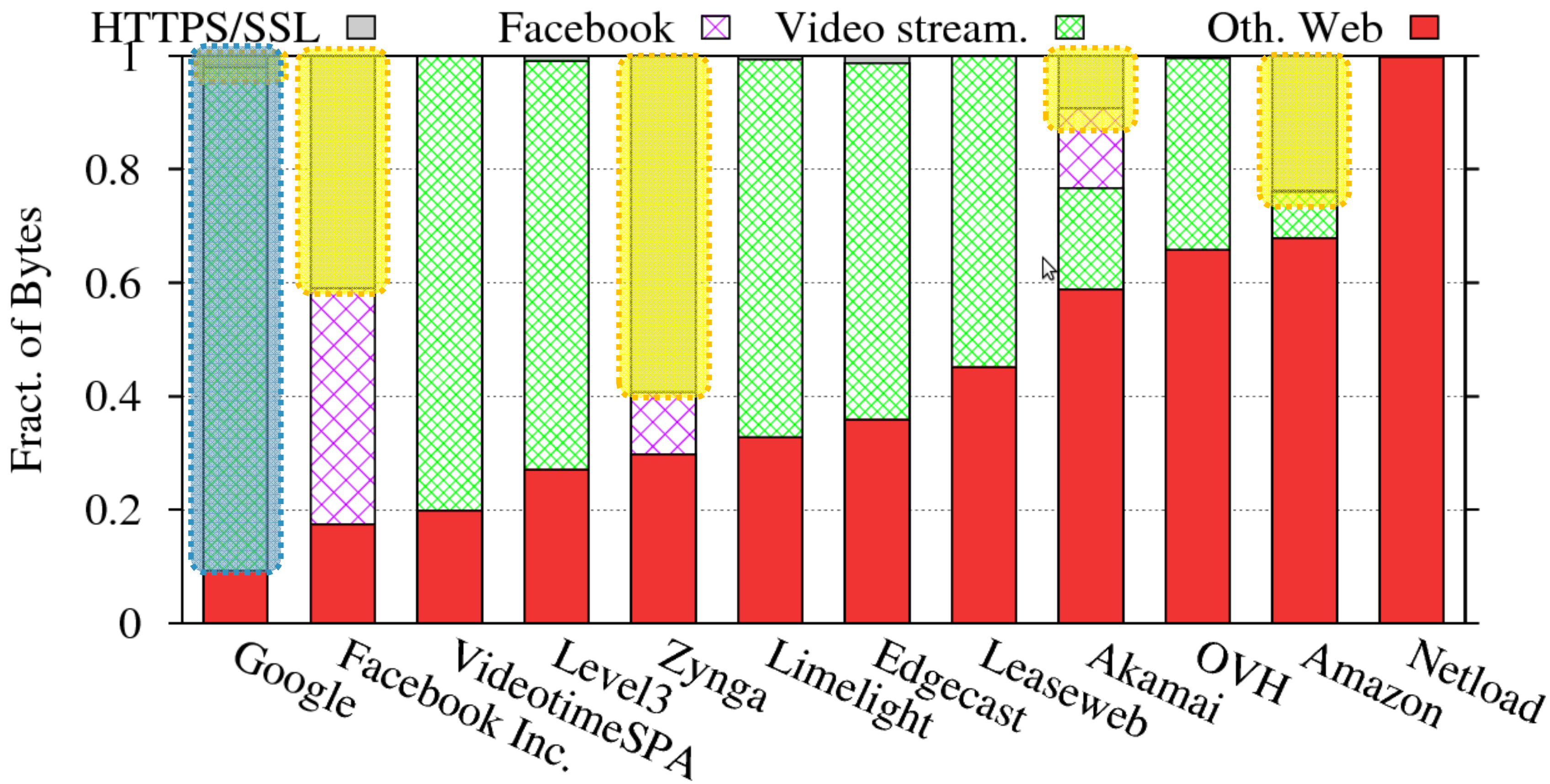
Rank	Organization	Bytes	
1	Google	29.8%	+++
2	Akamai	19.2%	+++
3	Level3	5.2%	++
4	Limelight	4.5%	++
5	Netload	3.1%	NEW
6	Leaseweb	2.0%	---
7	Edgecast	1.8%	NEW
8	VideotimeSpa	1.6%	NEW
9	OVH	1.2%	+
10	Facebook	1.1%	+
11	Amazon	1.1%	NEW
12	Zynga	0.14%	+
	Total	***70.6%***	

49% (Google + Akamai)

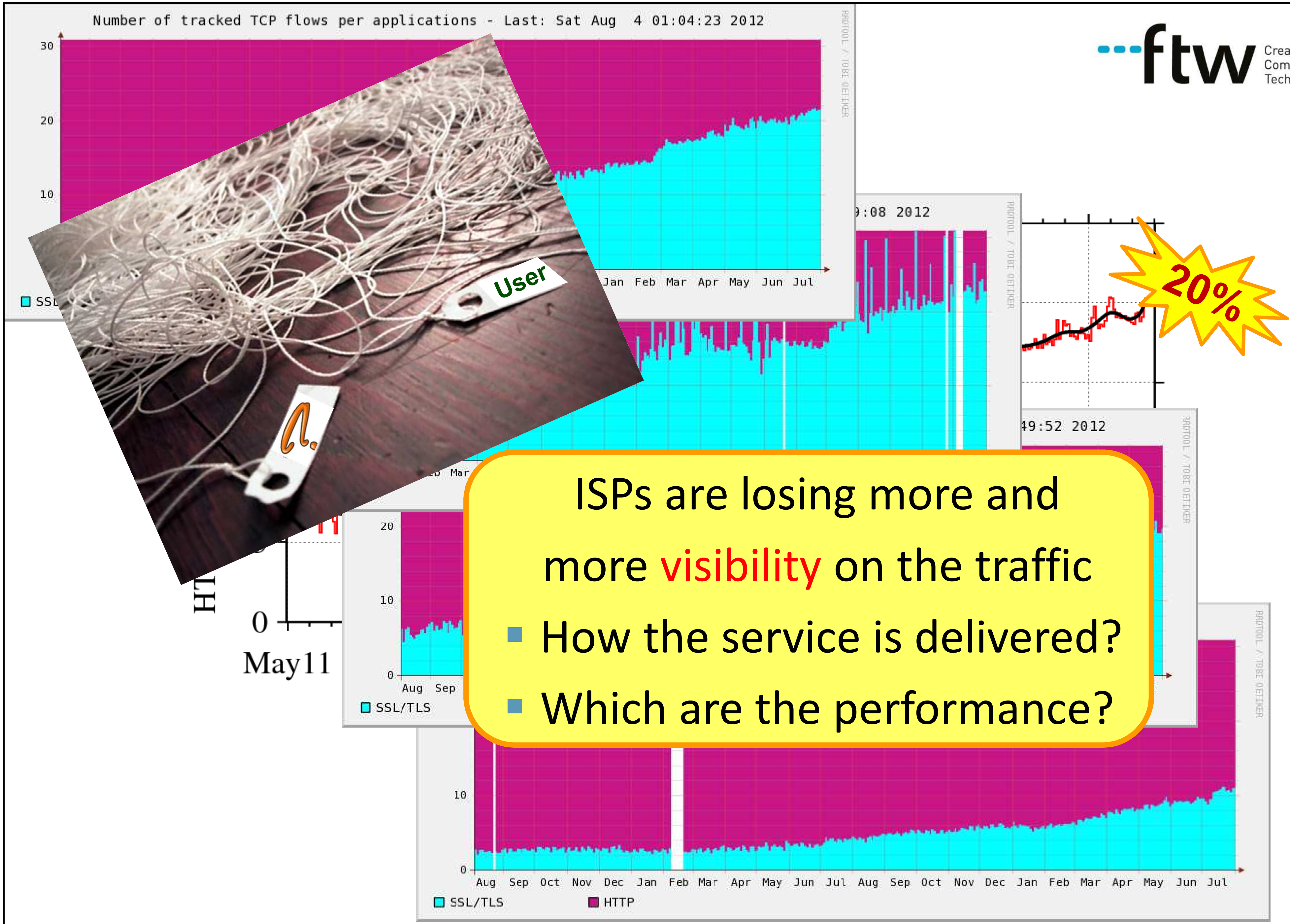
CLOSED (Megaupload)

+5.7%

Volume breakdown



- 90% of Google traffic is ***YouTube --> the biggest service in todays Internet***
- **HTTPS/SSL** is not used by all the top organizations
- ...but **can represent a large share of the volume**

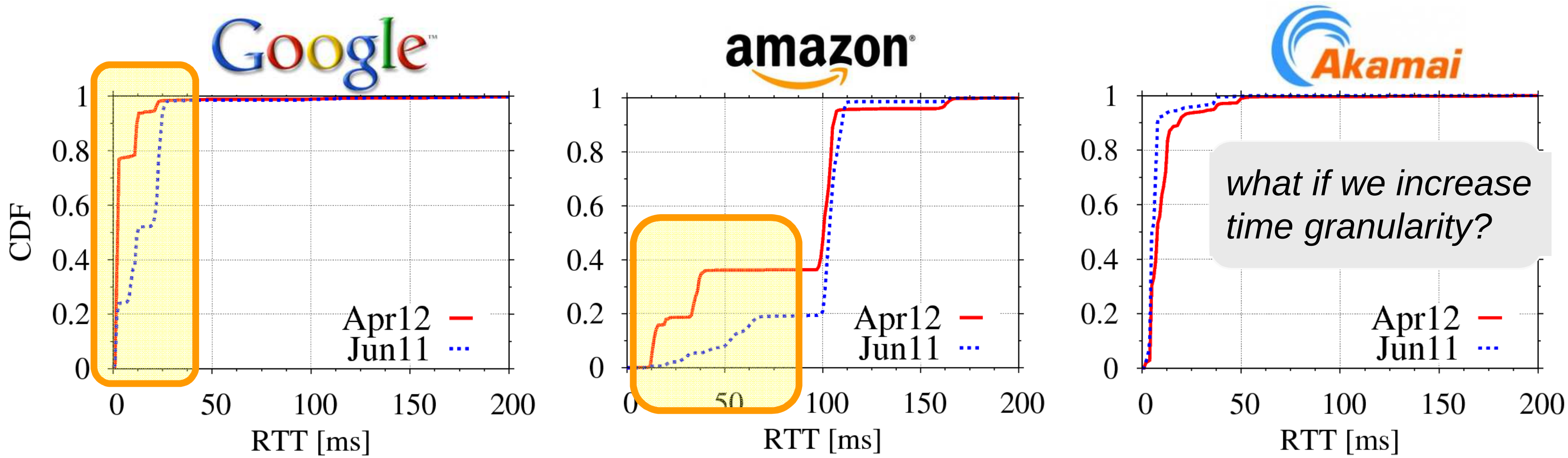


ISPs are losing more and more visibility on the traffic

- How the service is delivered?
- Which are the performance?

Data Centers distance

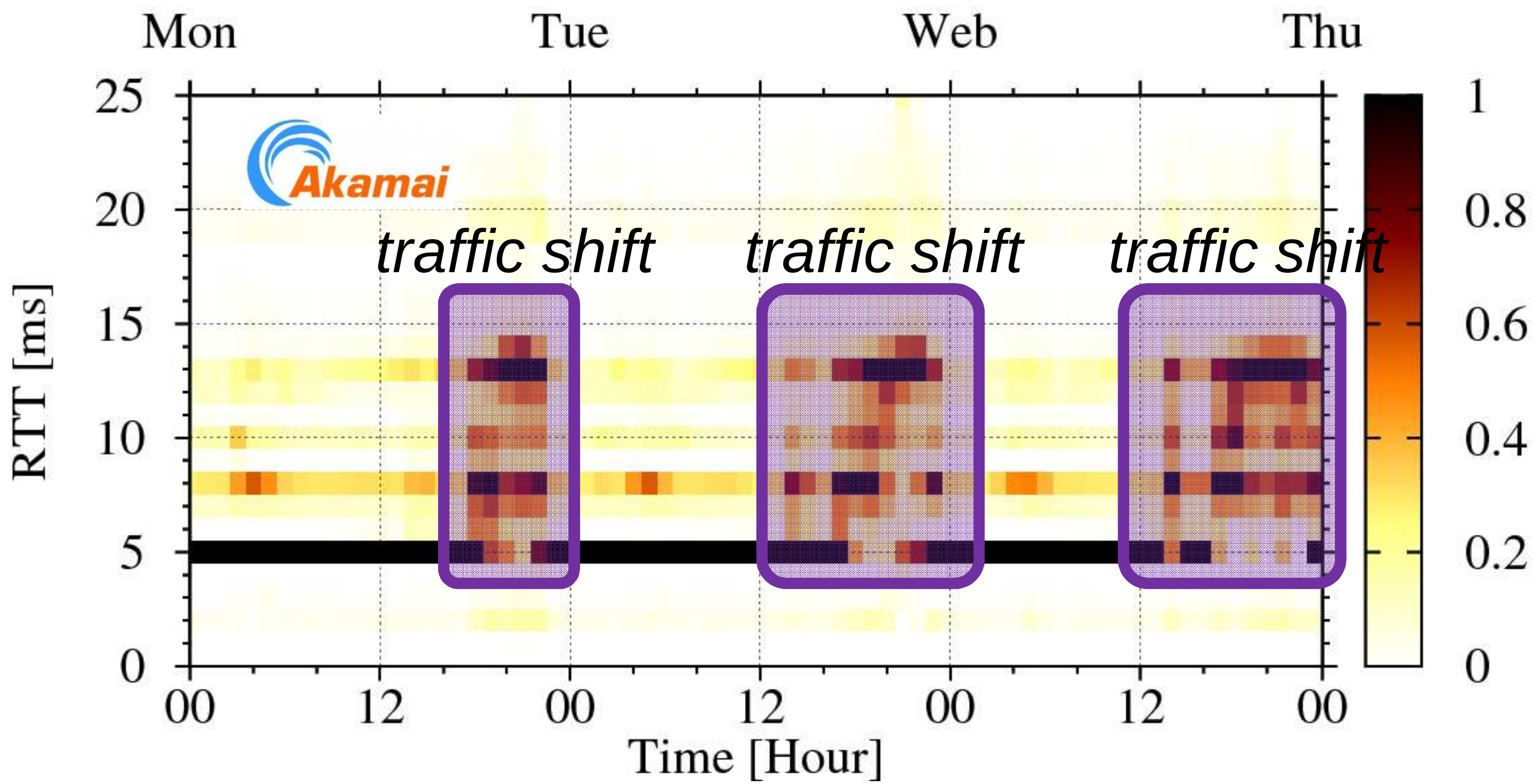
min RTT can be used as metric of distance



Organizations' networks evolve

- Different load balancing policies (Google)
- New Data Centers can be added/removed (Amazon)

Load balancing policies (April-12)



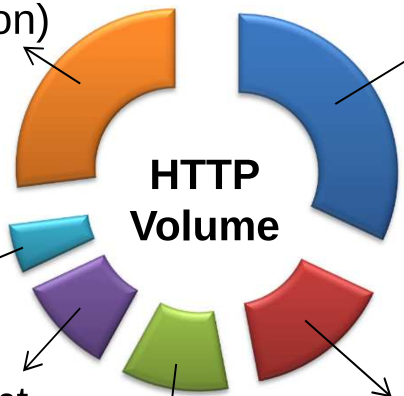
Time variant policies are a strong component of the services

- Long-term scales are important as well as short-term scales

A view from a mobile network

- We use now an off-line dataset collected at a **mobile network**
- 1 week of data in April 2012, more than **1/2 billion of HTTP flows**
- The **top-10 services** account for almost **60% of HTTP traffic volume**, and are accessed by 80% of the customers
- Top services: ***YouTube***, ***Facebook***, Google Services, Apple (iTunes and Store), ***Adult Video Services***, Windows Update Services, etc.

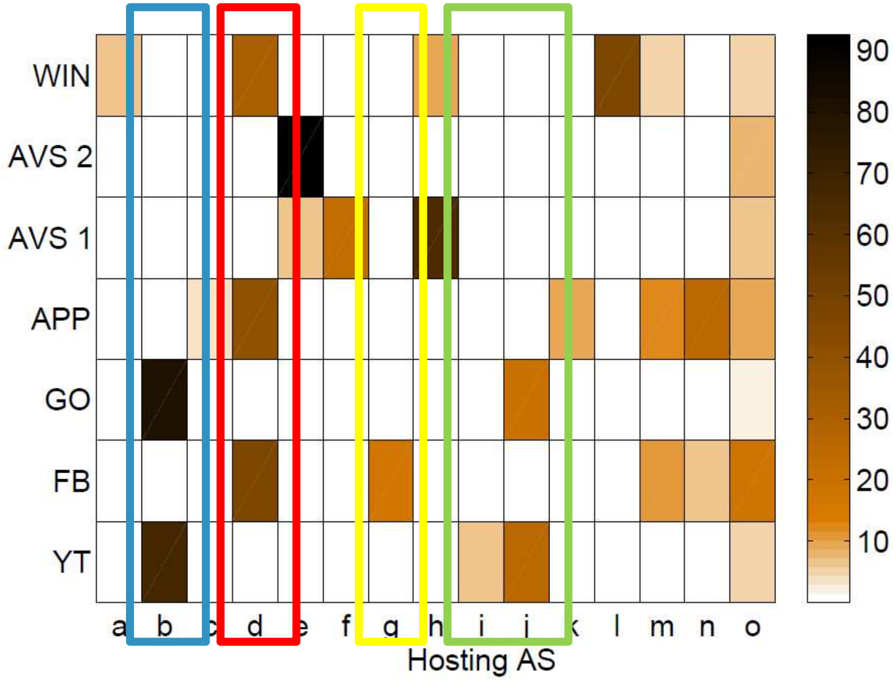
Who is Hosting the HTTP Content of the Internet?



CDN	#IPv4	#ASes	#/24 (% full)
Google	~660K	8	~2.5K (99.9)
Akamai	~3.5M	22	~14K (99.9)
Limelight	~110K	8	~400(99.8)
LeaseWeb	~480K	3	~1.8K (100)

- **A small number of CDNs is dominating the landscape of Internet content hosting**
- **Google CDN, Akamai, Limelight, and LeaseWeb** host together more than **40% of the HTTP content** observed at our vantage point
- HTTP **transparent caching in ISPs** is very spread

Who is Hosting the Top HTTP Services?

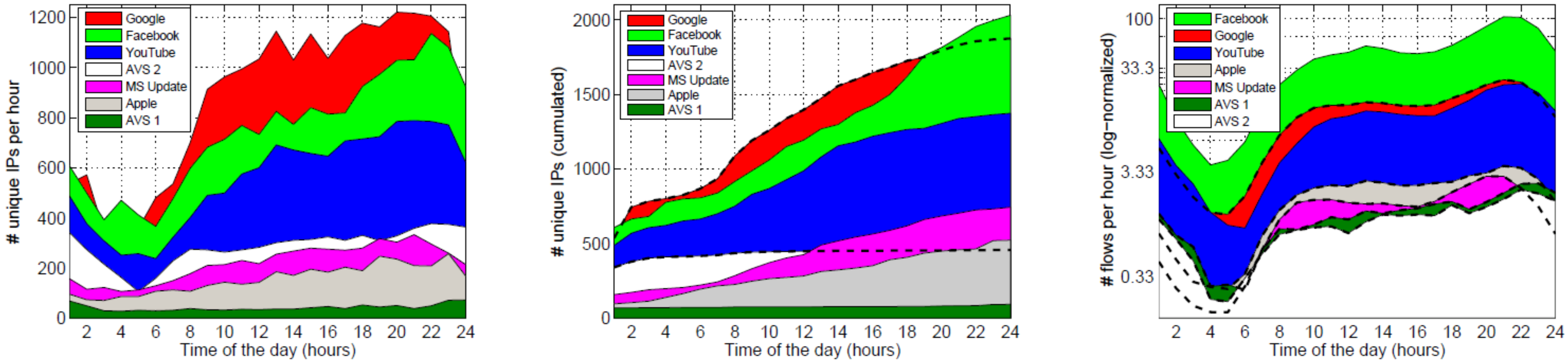


(a) IPs associated to top services

Organization	id	Organization	id
Hotmail	a	Level 3	h
Google	b	YouTube AS1	i
Akamai EU	d	YouTube AS2	j
LimeLight	e	Apple	k
Web Hoster	f	Microsoft	l
Facebook	g	ISPs	m,n,o

- Different **services** are **hosted by multiple organizations**
- **Akamai EU hosts** a large share of the servers hosting **Facebook, Apple Services, and Windows Services**
- **Non-cached** content from **Google Search** and **YouTube** is exclusively hosted by **Google CDN** (YouTube ASes included)
- **Most of the IPs serving the top HTTP services are hosted by Google and Akamai.**

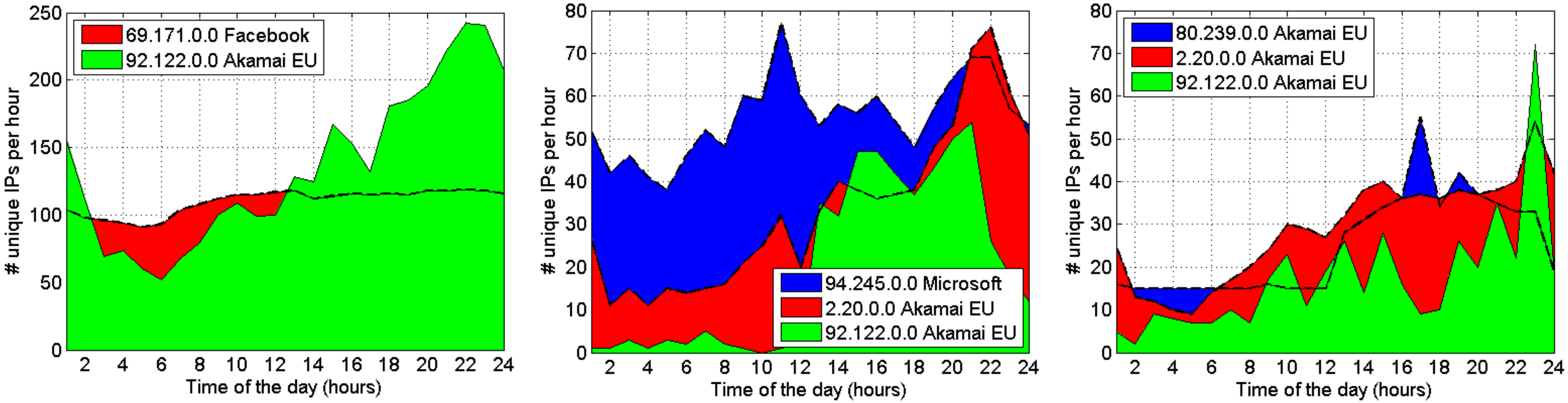
How many IPs are used to provision each Service?



(a) Unique IPs per hour. (b) Cumulative number of unique IPs. (c) Number of flows per hour.

- The number of **single IPs** per hour providing the top HTTP services vary during the day (e.g., 250 IPs per service at 5 am to up to 1200 in the case of Google Search)
- **Google Search, Facebook and YouTube dominate the IP space**
- Thanks to Akamai, **Facebook is the most IP-distributed service**, using more than **2000 different IPs on a single day** (Akamai hosts the static content)
- Some services (e.g., AVS 1) are provisioned by very stable delivery infrastructures

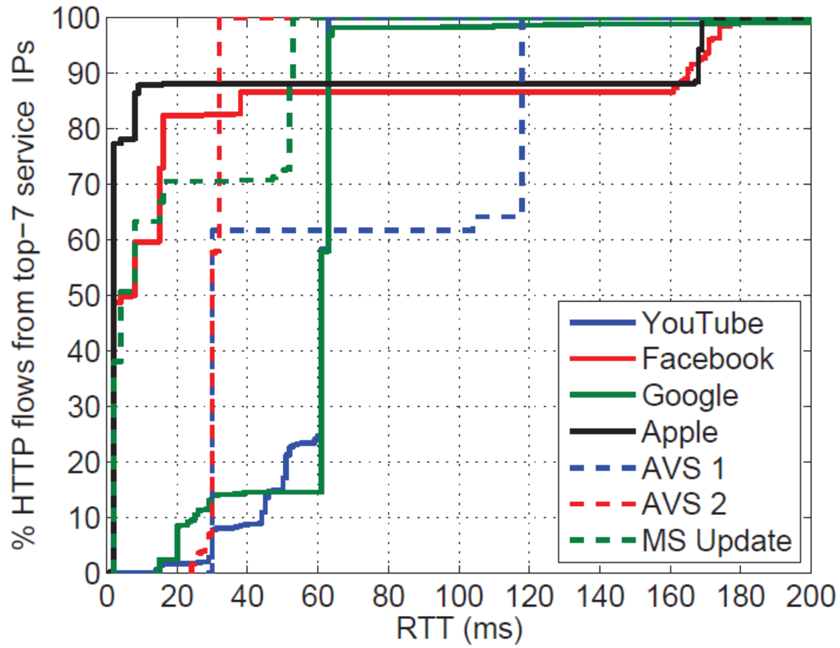
Different Subnets Utilization – not only time of the day



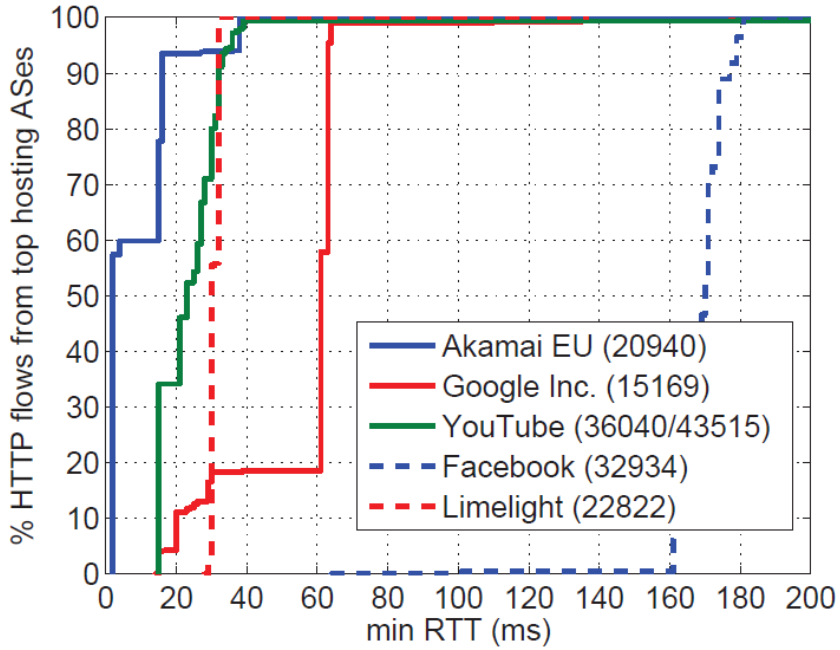
Facebook IPs | Apple Store and iTunes IPs | Windows Update IPs

- **Different subnets** of the CDNs **are more dynamic** than others
- **Akamai flows are served** from very dynamically **changing locations**
- Provisioning **servers “on-demand”** is extensibly used by **Akamai**
- **Akamai flows are, a-priori, more difficult to track using server IPs**

Where are the Caches?



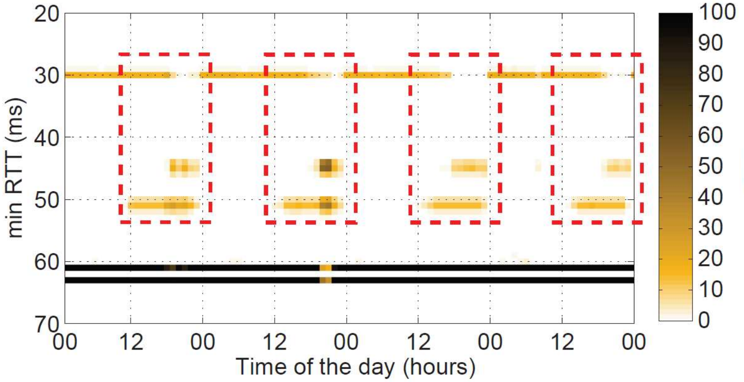
(a) min RTT per services.



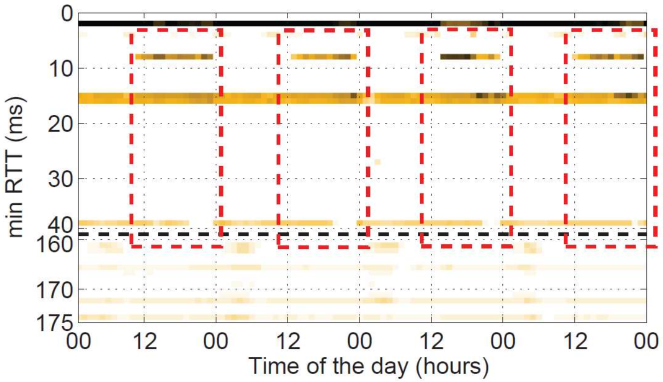
(b) min RTT per hosting organization.

- Distribution of min RTT per service and per hosting organization (num flows weighted).
- **Steps in the CDFs potentially reveal differently located caches/data-centers.**
- A big share of **Facebook, Apple, and Windows Update** flows come from servers **located in the same city of the vantage point** (min RTT < 5ms)
- **Dynamic Facebook content** (Facebook AS) is **located in the US** (min RTT > 150ms)
- More than **60% of the Akamai HTTP flows come from servers “inside the ISP”**, with min RTT values smaller than 5ms.

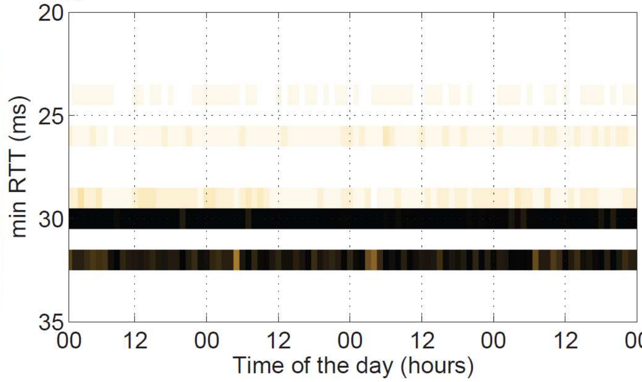
Distribution Policies – Load Balancing



(a) min RTT of YouTube flows.



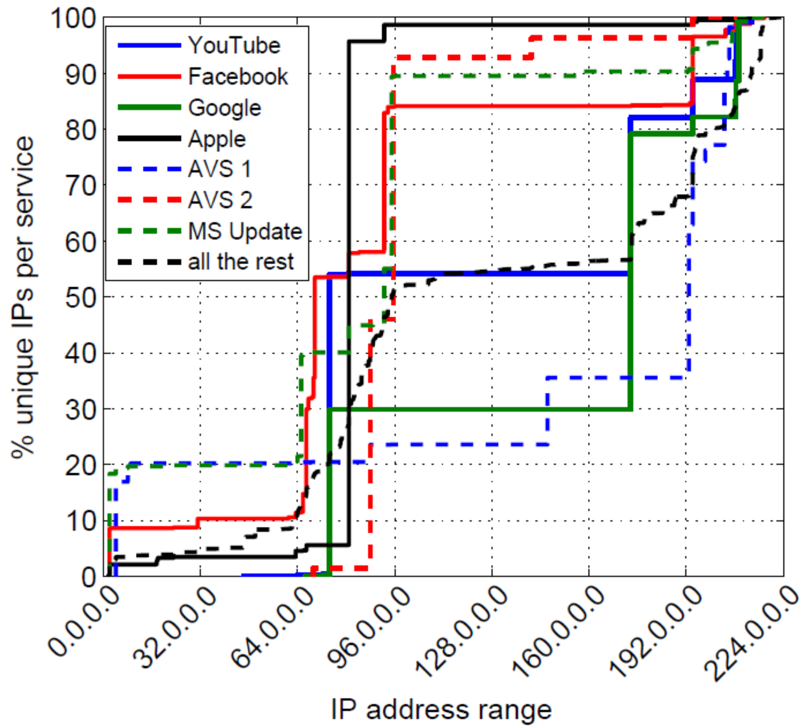
(b) min RTT of Facebook flows.



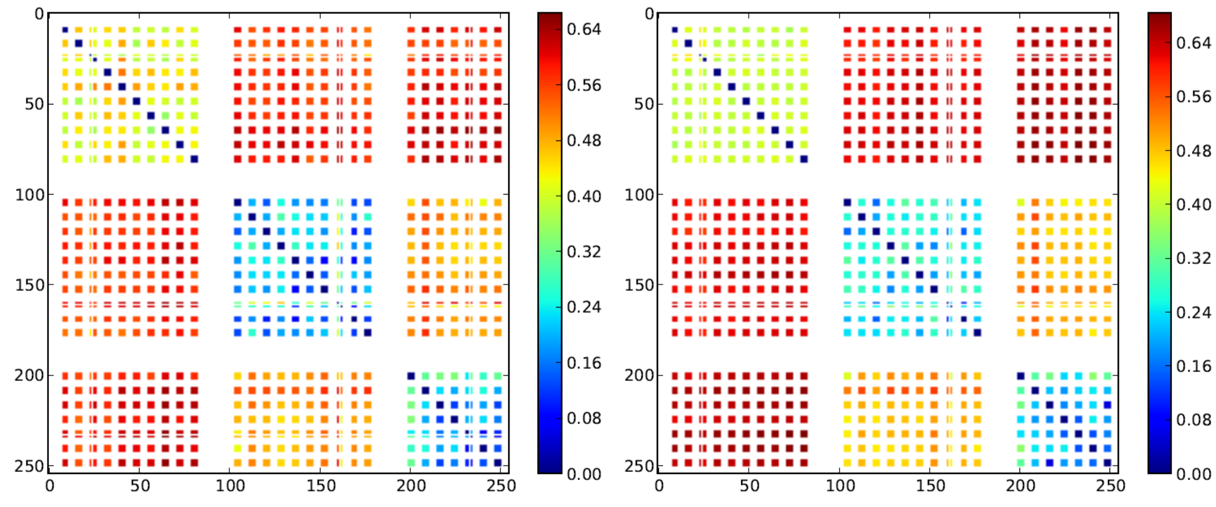
(c) min RTT of AVS 2 flows.

- **4 days min RTT evolution** for YouTube (mainly Google CDN), Facebook (mainly Google Akamai), and AVS 2 (mainly Limelight)
- Google CDN and Akamai make **use of load balancing policies** to serve **content from different caching locations**
- **YouTube and Facebook**: **markedly min RTT shifts** occur every day at exactly the **same time slots**, showing a min RTT periodic pattern
- No observable temporal patterns for AVS 2, suggesting that **Limelight is not applying load balancing techniques**, at least from our vantage point perspective

IP Collisions and Mappings‘ Stability



(a) IP range distribution on a single day

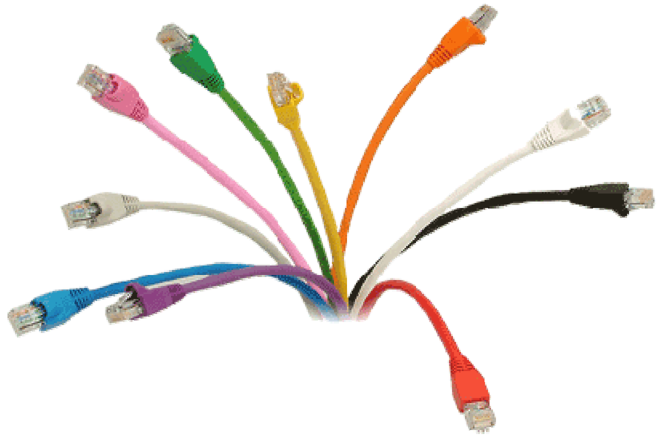


(a) 2 days (b) 10 days

Temporal stability of IPs–services association in Akamai, for Facebook

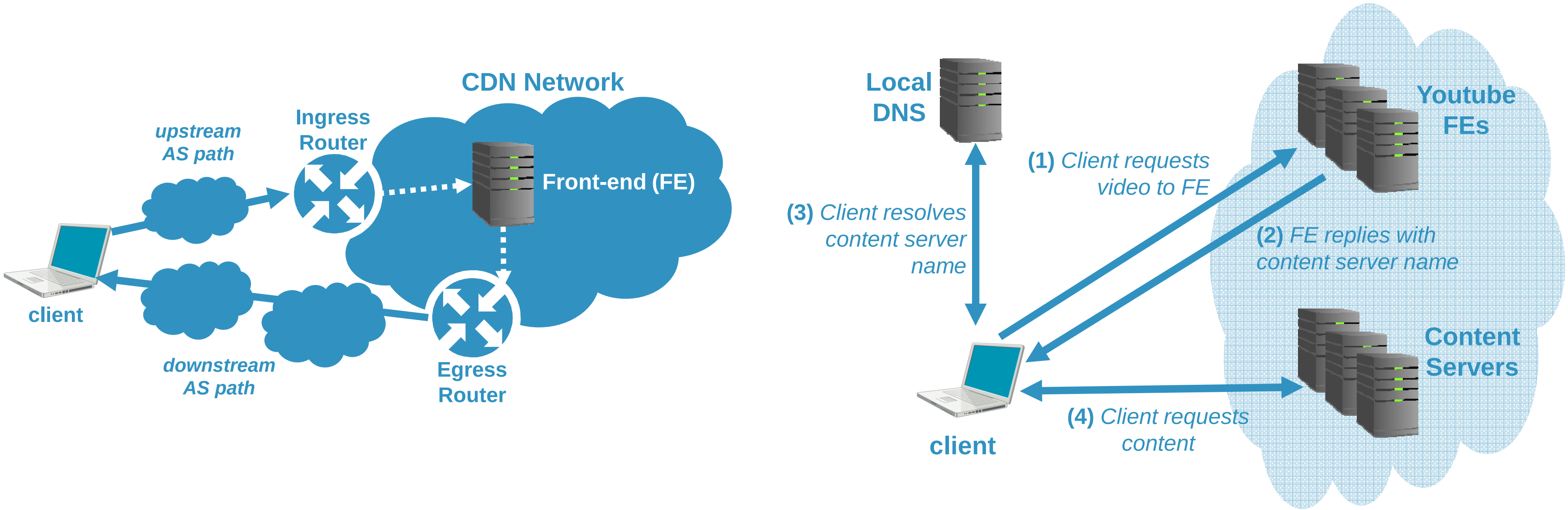
- **IP collisions → ** different services are provisioned by the same IP address at different times of the day (same CDN, IP anycast, ISP’s caching, etc.)
- For example, **Facebook collides with Apple Services and Windows Services (same CDN – Akamai)**, **Google Search** and **Facebook** collide (**ISP caching**), etc.
- Yet, **some regions of the Akamai IP space are very stable** and **used exclusively for some services** (check the Facebook example)

How does YouTube look like in Mobile & Fixed-line Networks?



A typical CDN architecture
Google CDN for Youtube

- Google CDN employes a complex server selection strategy for:
  - load balancing
  - optimize client-server latency
  - increase QoE in general
- DNS used for re-direction based on content popularity and location.



Youtube load-balancing

- DNS-driven users redirection
- Goals:
  - Load balancing
  - Optimize choice of content servers aimed at reduce latency for clusters of users (cluster: <AS,country>)
- Is it always optimal? Look at the next example...

Load balancing events impacting QoE

Requestes served by different /24 subnets

..which correspond to different data centers

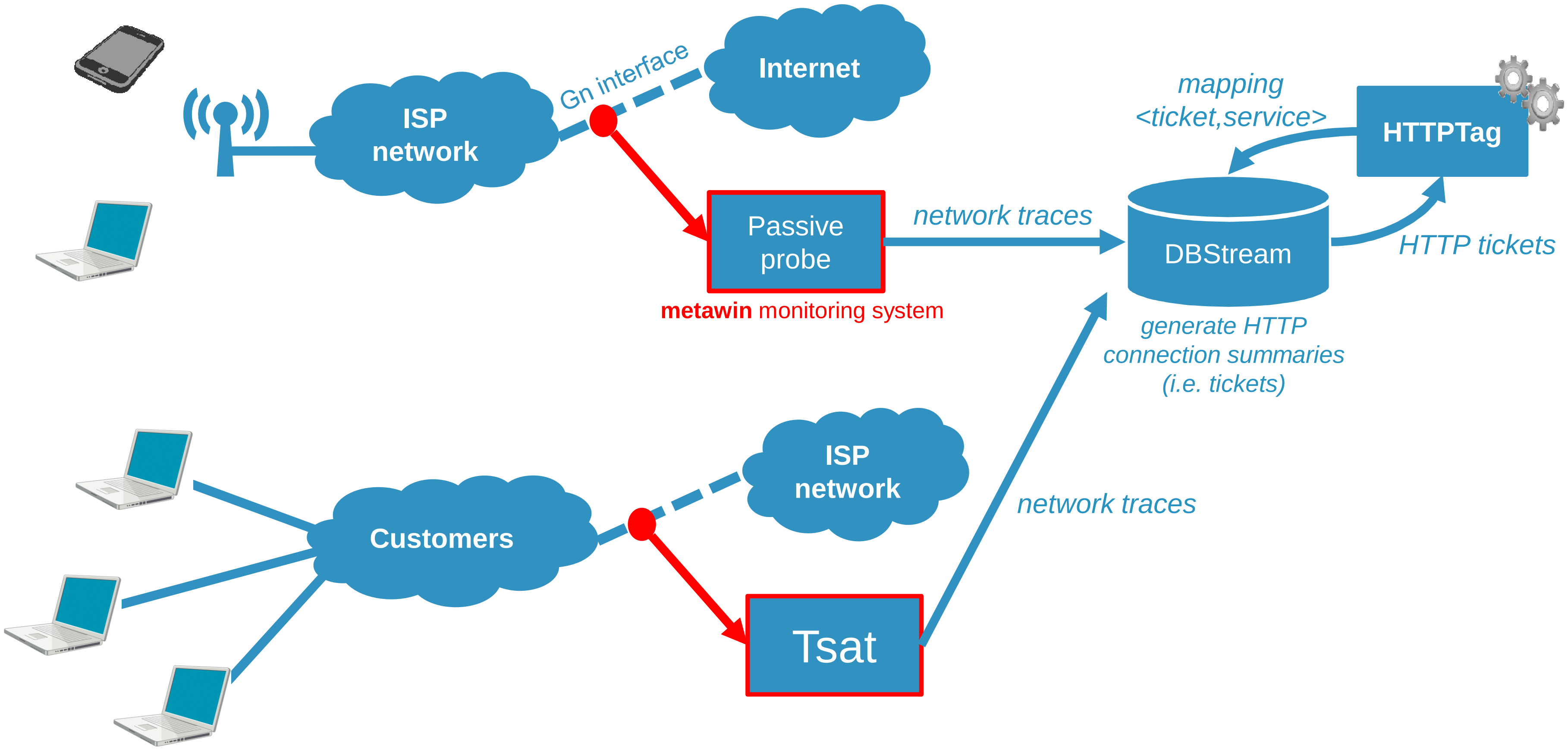
- Measurement from a single vantage point (European fixed-line ISP)

SUBNET	NAME with AIRPORT code	5-May #flow	5-May Tru avg	6-May #flow	6-May Tru avg	7-May #flow	7-May Tru avg
173.194.18	fra02s08.c.youtube.com	-1	-1	-1	-1	-1	-1
173.194.19	fra02s15.c.youtube.com	-1	-1	-1	-1	-1	[illegible]
173.194.2	mil01s12.c.youtube.com	17054	1333.46	15470	1276.31	13655	**125[illegible].63**
173.194.20	par08s06.c.youtube.com	-1	-1	-1	-1	-1	-1
173.194.208	par08s06.c.youtube.com	-1	-1	-1	-1	-1	-1
173.194.5	lhr14s08.c.youtube.com	449	1819.57	283	1658.45	-1	-1
173.194.6	fra07s13.c.youtube.com	-1	-1	-1	-1	-1	-1
173.194.62	fra07s19.c.youtube.com	-1	-1	-1	-1	-1	-1
173.194.9	par03s06.c.youtube.com	-1	-1	-1	-1	-1	-1
208.117.236	par03x04.c.youtube.com	179	164.18	4250	540.16	957	496.91
208.117.248	mia02s11.c.youtube.com	-1	-1	77	552	-1	-1
208.117.250	ams09x06.c.youtube.com	41430	679	49437	656.39	57675	**653.81**
208.117.252	dfw06x02.c.youtube.com	-1	-1	51	285.63	-1	-1
208.117.254	fra07x03.c.youtube.com	838	667.29	2130	852.53	-1	-1
74.125.105	lhr22s16.c.youtube.com	1829	1551.78	1655	1185.94	3957	[illegible]42.47
74.125.13	zrh04s03.c.youtube.com	719	1074.15	499	2264.09	82	1302.03
74.125.14	mil02s01.c.youtube.com	48366	1234.82	37968	1253.01	37182	**1162.[illegible]**
74.125.216	bru02t11.c.youtube.com	-1	-1	-1	-1	-1	-1
74.125.218	fra07t13.c.youtube.com	8697	1355.33	12579	1338.71	8560	1239
74.125.4	lhr22s11.c.youtube.com	1496	1846.25	2488	1034.78	4146	**1363.63**
74.125.99	fra07s03.c.youtube.com	-1	-1	-1	-1	-1	-1

YouTube

Datasets for YouTube Characterization

- YouTube data from **two different vantage points** (3 days of YouTube flows in mid 2013, 2 EU countries):



ftw Creating Communication Technologies

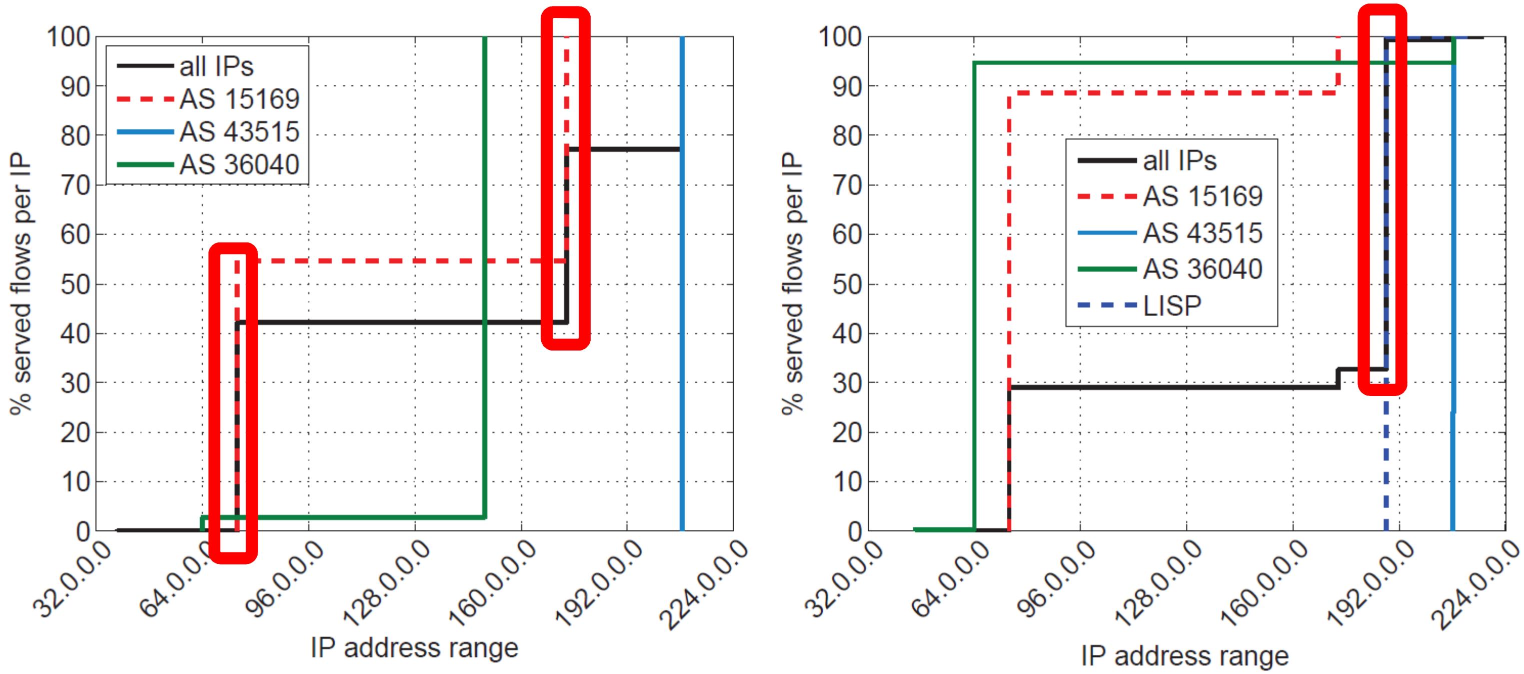
Hosting Infrastructure (1/3)

Autonomous System	**# IPs**	**#/24**	**#/16**
All server IPs fixed-line	3646	97	22
15169 (Google)	2272		
43515 (YouTube)	1222		
36040 (YouTube)	43		
All server IPs mobile	2030		
15169 (Google)	1121		
43515 (YouTube)	844		
LISP	35	4	3
36040 (Google)	26	5	3

(Network) Autonomous System	**% bytes**	**% flows**
(FL) 15169 (Google)	80.8	77.3
(FL) 43515 (YouTube)	19.1	22.5
(M) LISP	69.3	66.7
(M) 15169 (Google)	30	32.7

- Almost the **double of IPs in fixed-line access**, even if the population is much lower.
- **Servers** are highly distributed among **2 Google ASes** (**15169** and **43515**).
- The **Local ISP** (LISP) plays a key role in the distribution of YouTube videos in mobile, **serving about 70% of the video flows** (Google Global Cache – CDN inside the ISP approach, following Akamai).

Hosting Infrastructure (2/3)

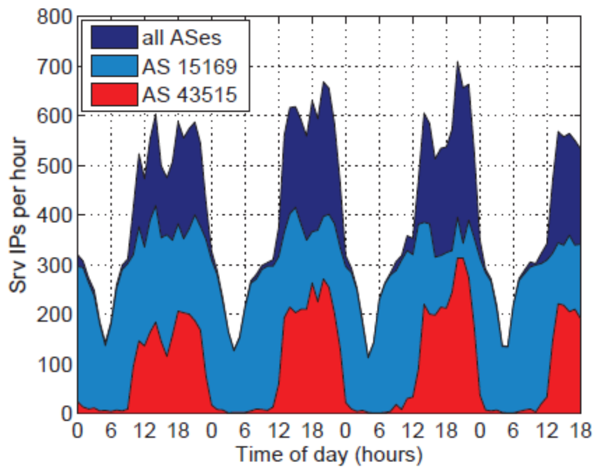


Flows per server IP - fixed-line.   Flows per server IP - mobile.

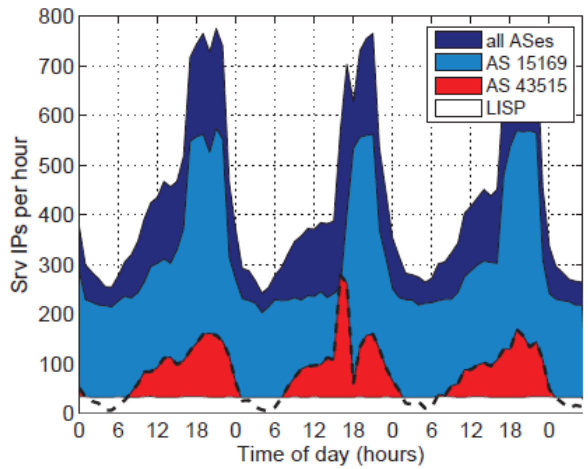
- **Flows are mainly served from AS 15169 in fixed line**, from 2 /16 prefixes, and complemented from 1 /16 prefix in AS 43515.
- The **LISP uses mainly a single /16 prefix** for servers hosting YouTube, and the **same 2 /16 prefixes from AS 15169.**

Hosting Infrastructure (3/3)

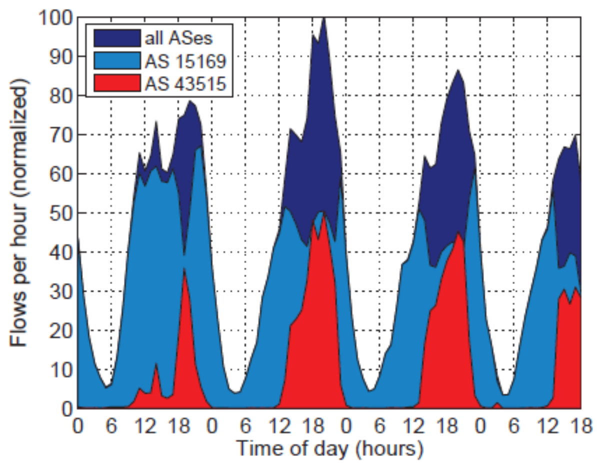
- Up to 700 YouTube server IPs active per hour at peak times
- Load balancing based on time of the day is much more evident in fixed-line (abrupt increase in #IPs from AS 43515)
- LISP IPs are constantly used during the complete period
- As a consequence, the dynamics on the # of served flows are much easier to predict in mobile
- This results in a potentially much easier traffic management at the core of the mobile network.



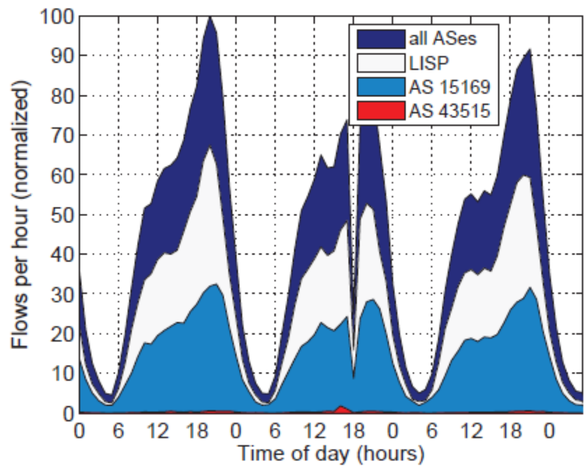
(a) IPs per hour hosting YouTube - fixed-line.



(b) IPs per hour hosting YouTube - mobile.

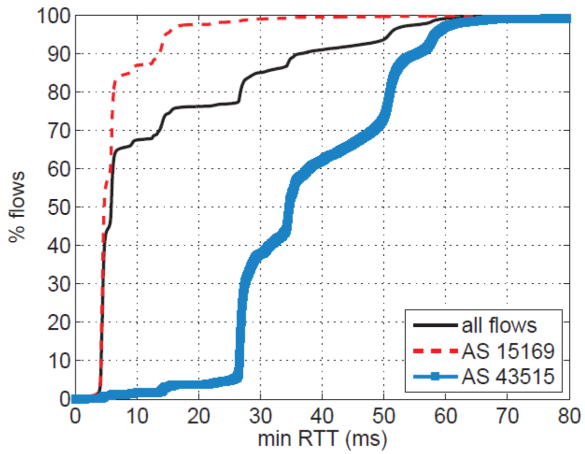


(c) Flow counts per hour - fixed-line.

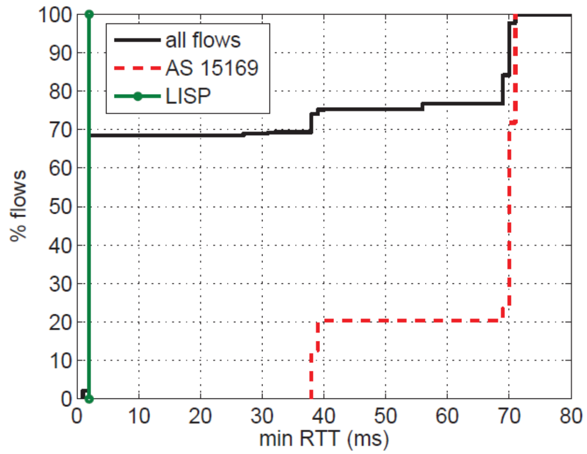


(d) Flow counts per hour - mobile.

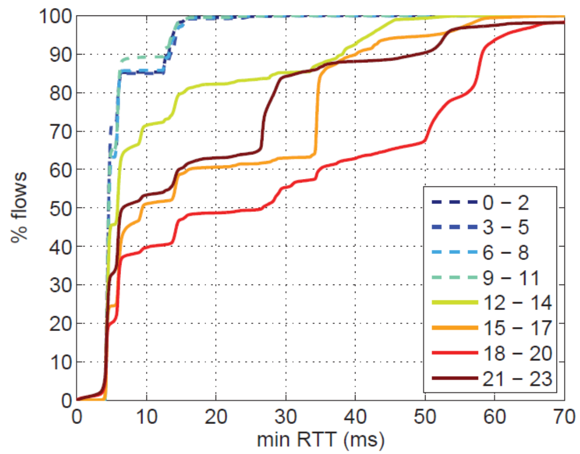
How far are the YouTube servers?



(a) min RTT (passive) in fixed-line.



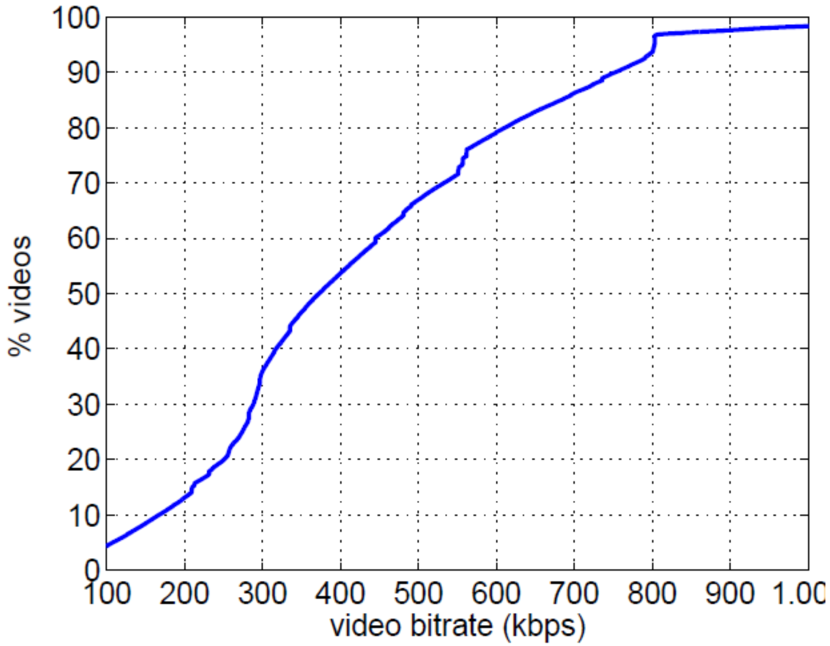
(b) min RTT (active) in mobile.



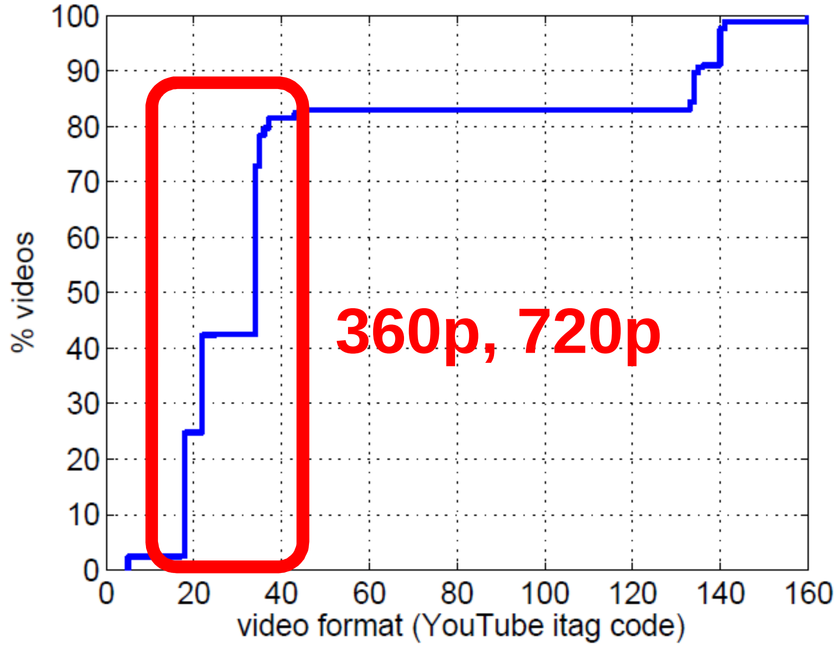
min RTT - daily variation.

- min RTT from vantage point as a measure of server location → passive in fixed-line, active in mobile (avoid acceleration middle-boxes)
- AS 15169 servers are very close to fixed-line customers → direct peering to Google at the IXP. AS 43515 servers at further locations, still in EU
- AS 15169 servers in other EU country(ies), LISP servers directly connected to the core mobile network.
- Temporal load-balancing (fixed–line) → servers at further distances from AS 43515 are selected at peak hours → **latency–map based decisions over-ruled by YouTube balancing policies**

Flow Characteristics



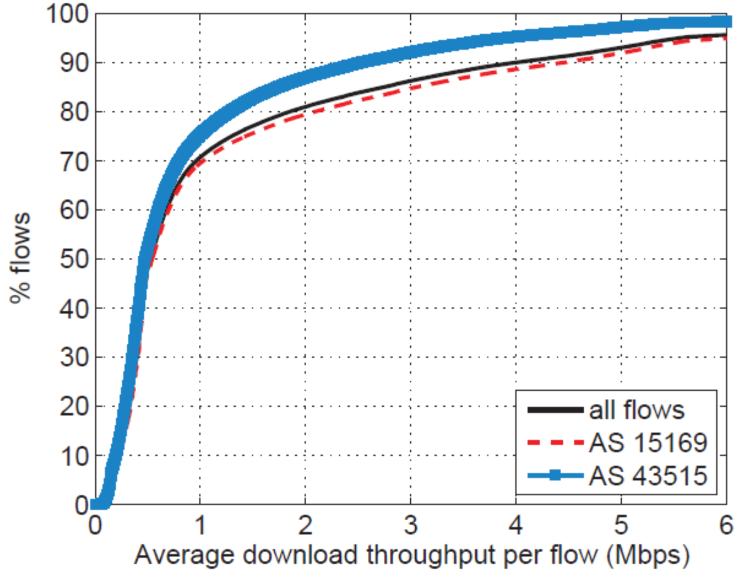
YouTube video bitrate.



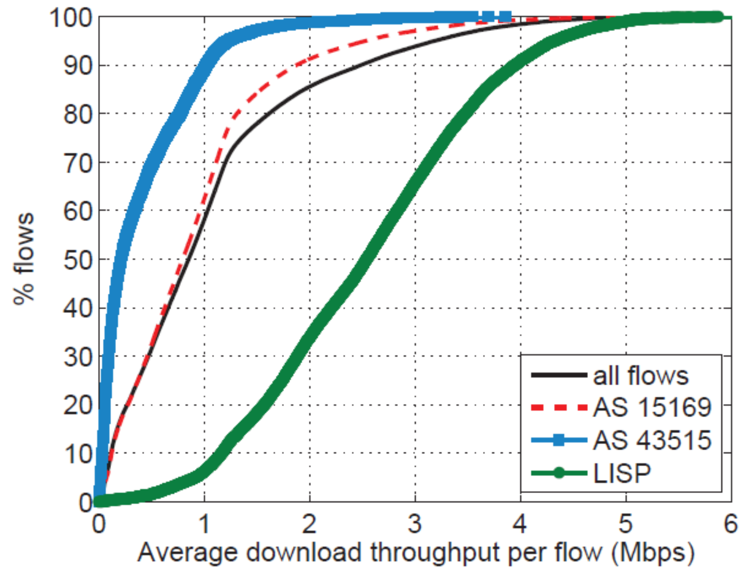
YouTube video format.

- in mobile, only flows bigger than 1MB (for accurate throughput computations)
- steps in the CDF correspond to YouTube chunking (1.8MB, 2.5MB, etc.)
- LISP flows size varies slightly between 2MB and 4MB
- AS 43515 serves larger–size YouTube flows

Network Performance – flow throughput



YouTube throughput- fixed-line.



YouTube throughput - mobile.

- we take flows bigger than 1MB only (for accurate throughput computations).
- more than 15% of the flows achieve a throughput above 2 Mbps in both networks.
- throughput is partially governed by the specific video bitrates and the YouTube flow control and not exclusively by the specific access technology (mobile or fixed-line).
- flows served by the LISP are the ones achieving the highest performance, with an average flow downlink throughput of 2.7 Mbps.
- benefits of local caching and low-latency servers for provisioning YouTube flows.

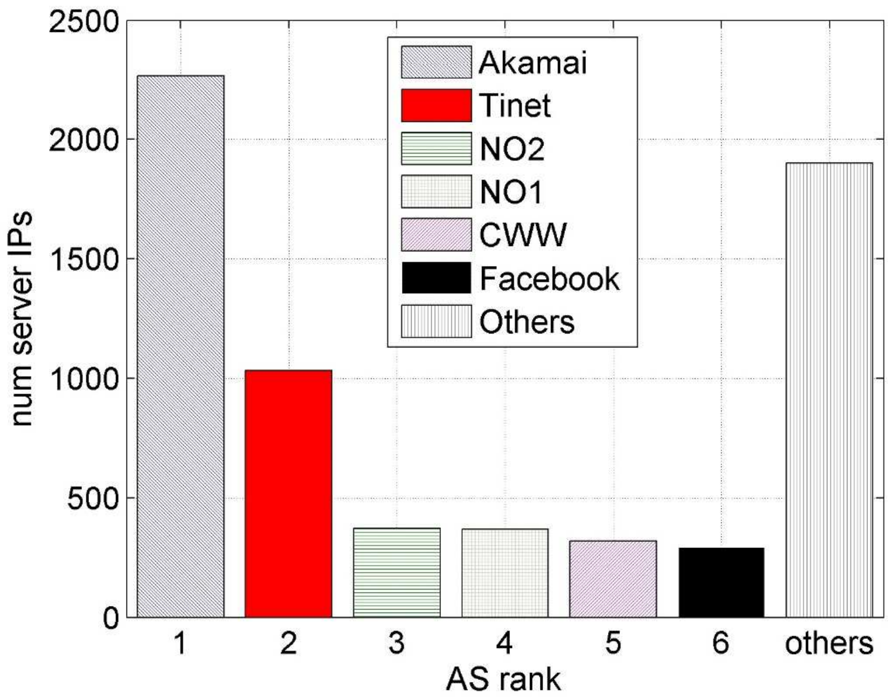
How does Facebook look like in Fixed-line Networks?

Why Facebook?

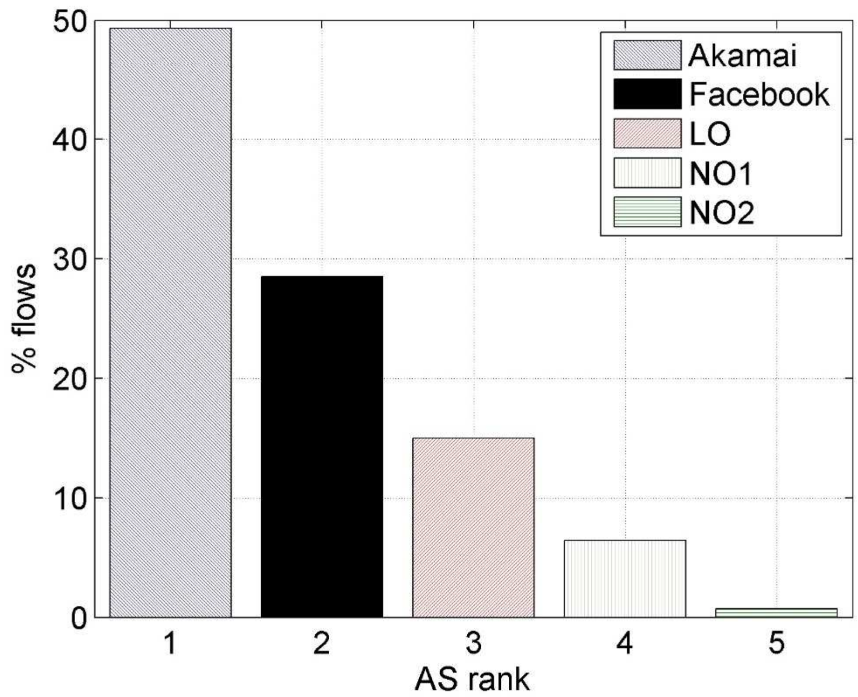
- Most popular and wide-spread Online Social Network (OSN)
- Hosted by Akamai
- Some numbers:
  - 1.28 billion of active users (as of March 2014)
  - 137.000 servers in 85 countries / 1200 networks
- From our dataset:
  - 70% of users in our dataset
  - 10% of total traffic volume
  - ~6000 different IP addresses
  - in ~250 Autonomous Systems
  - In ~20 countries across the globe

Facebook is the perfect study case to understand large services' provisioning systems

Hosting Infrastructure Overview



Number of IP addresses per A.S.



Share of flows per A.S.

- Top hosting companies: Akamai, two neighbor ISPs, Tinet, Cable&Wireless
- However: **Akamai plays key role (50% of traffic, 2600 IP addresses)**
- The others: mostly caches and spurious contents

Geographical Diversity [1/2]

Localizing Facebook IPs through MaxMind

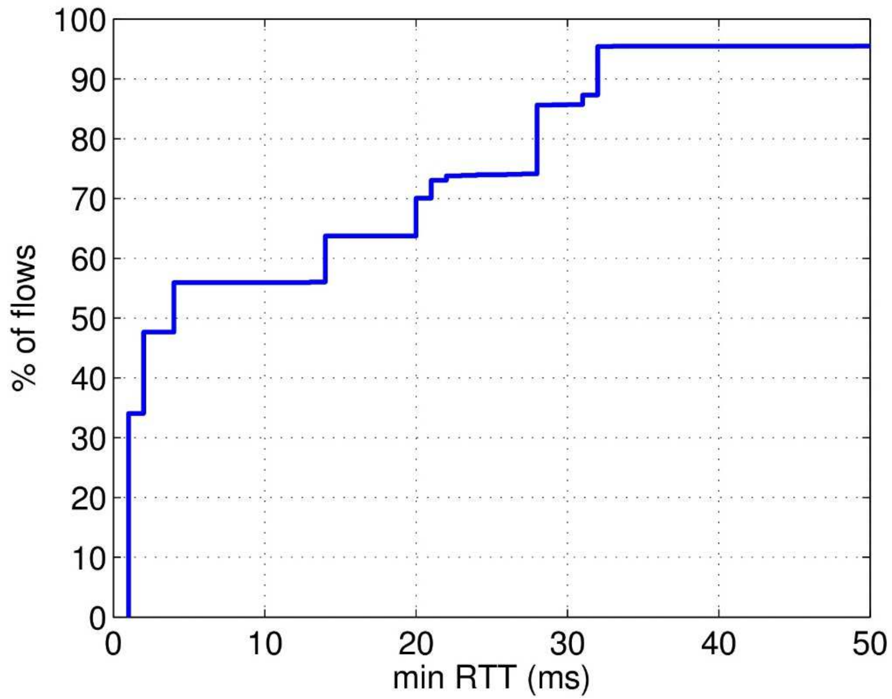
- **Austria**: *37.2%*
- **Ireland**: *12.7%*
- **Germany** *2.1%*
- **USA**: *1.1%*
- **Europe** (uncl.): *46.8%*

99% traffic from within Europe

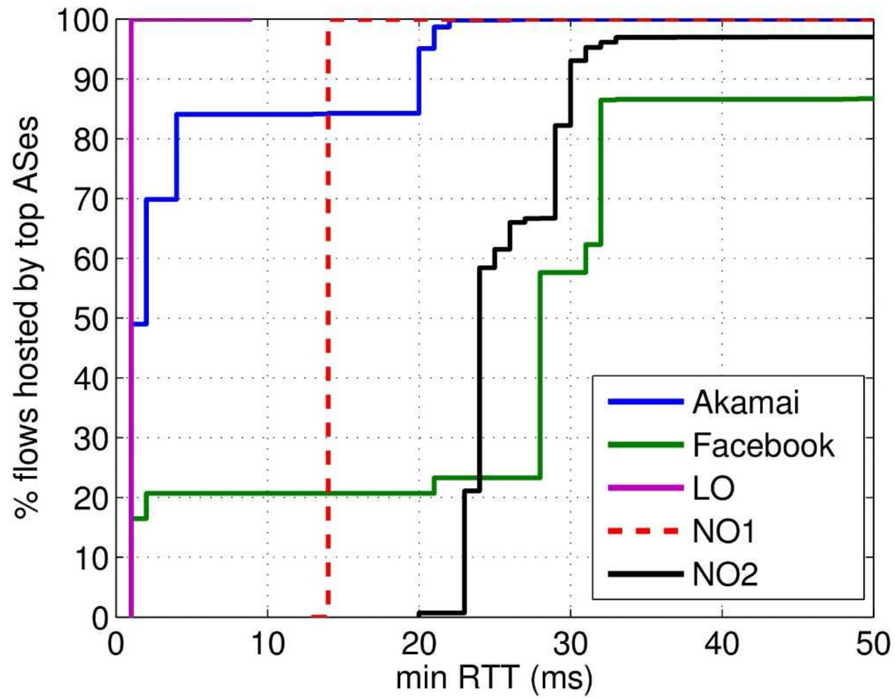
- Strong content localization
- Akamai Datacenters in Europe play biggest role
- ISPs caching: local and neighboring countries

Geographical Diversity [2/2]

Estimating servers distance through min RTT



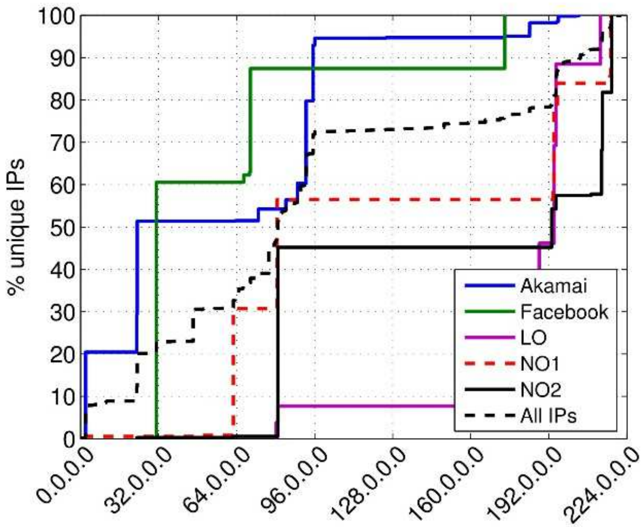
minumum RTT for all server IPs



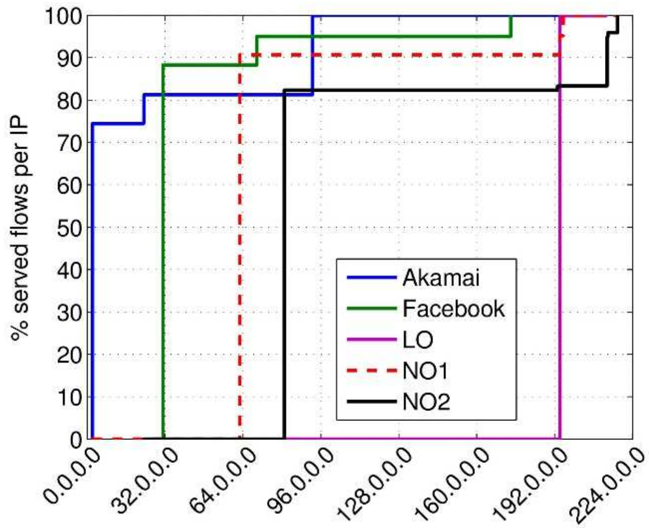
minimum RTT for top hosting A.S.

- Akamai AS: very short RTT (highly distributed and close to final users)
- Facebook AS: three knees (Ireland + locations in USA through local IXP)

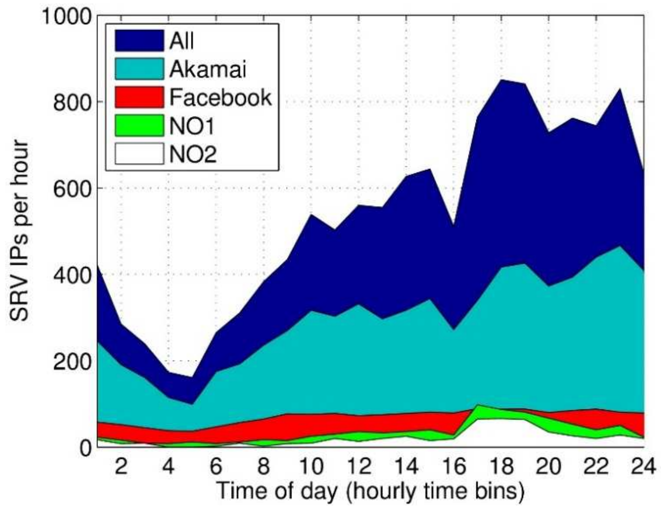
Addressing Space



Distrib. of IP addr. across ranges



Weighted distrib.

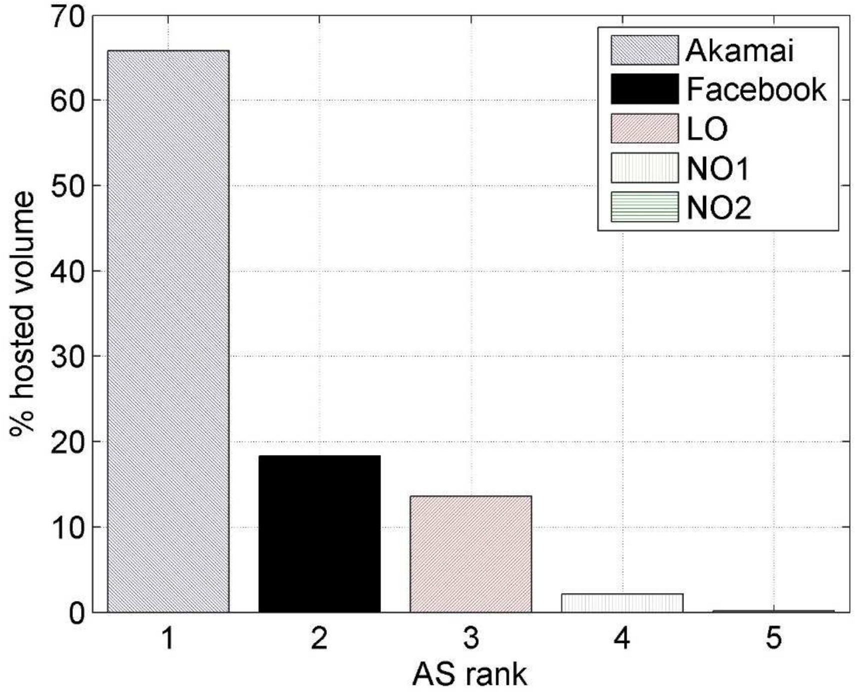


Deployment of IPs over a day

- Akamai's 75% of flows from a single subnet
- Facebook AS' 89% from a single range
- Neighbor Operators' 91% and 82% from a single range
- Local Operator only deploys a small range of IPs

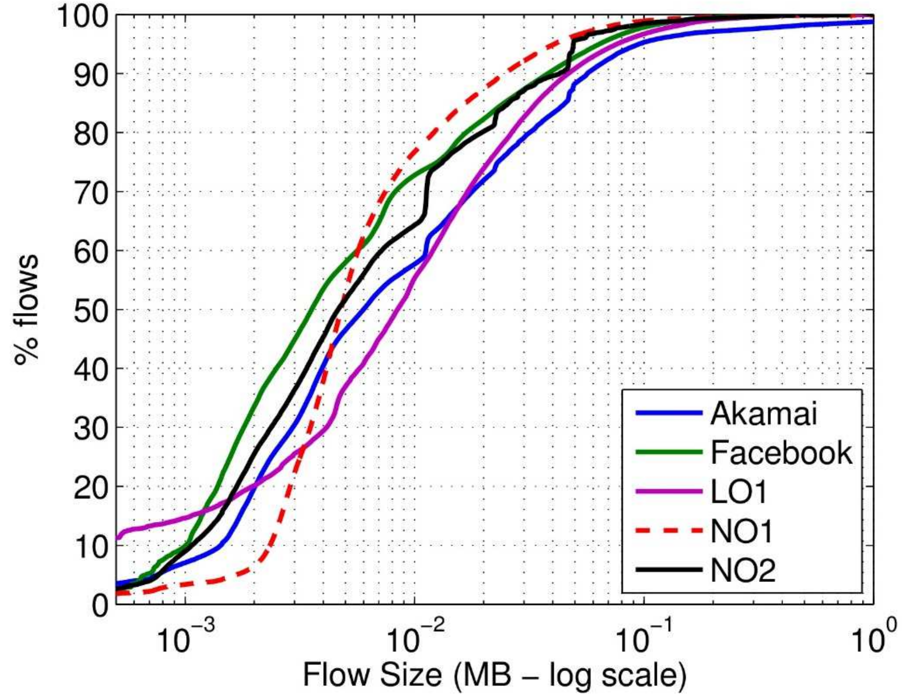
- Facebook AS IPs are always active (dynamic contents)
- Akamai strictly follows daily usage patterns (static contents)

Hosting players and roles



Shares of hosted volume per org/A.S.

- Akamai hosts more than 65% of traffic volume
- Facebook AS responsible for 20% of volume
- Local Operator (15%) is responsible for caching

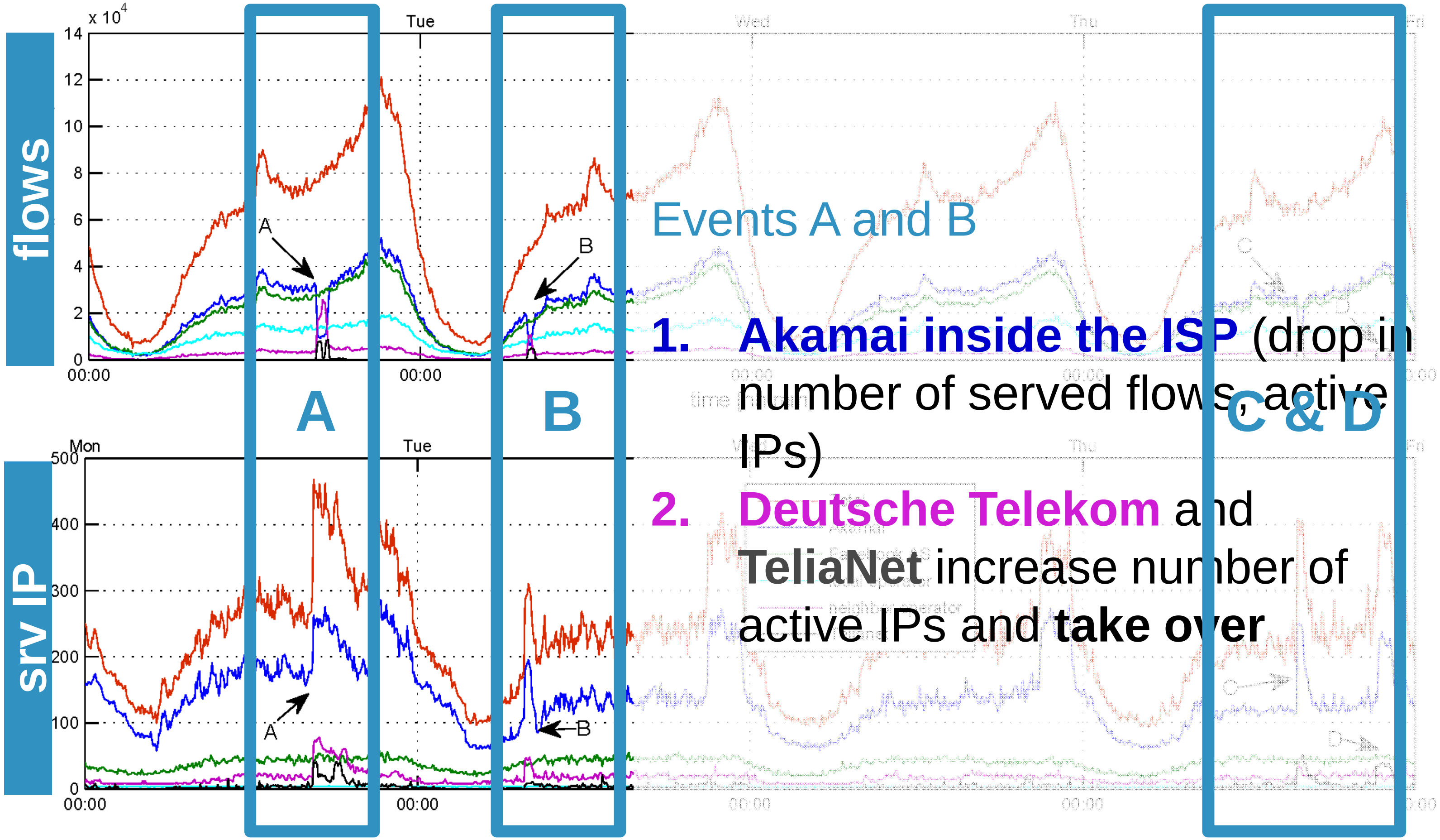


Distribution of flow sizes

- Distribution of flow sizes gives hint on the role of each hosting AS
- Akamai serves big flows (media/static contents)
- Facebook AS dedicated to dynamic contents

CDN Inter-play [1/3]

Time Series (4 days)

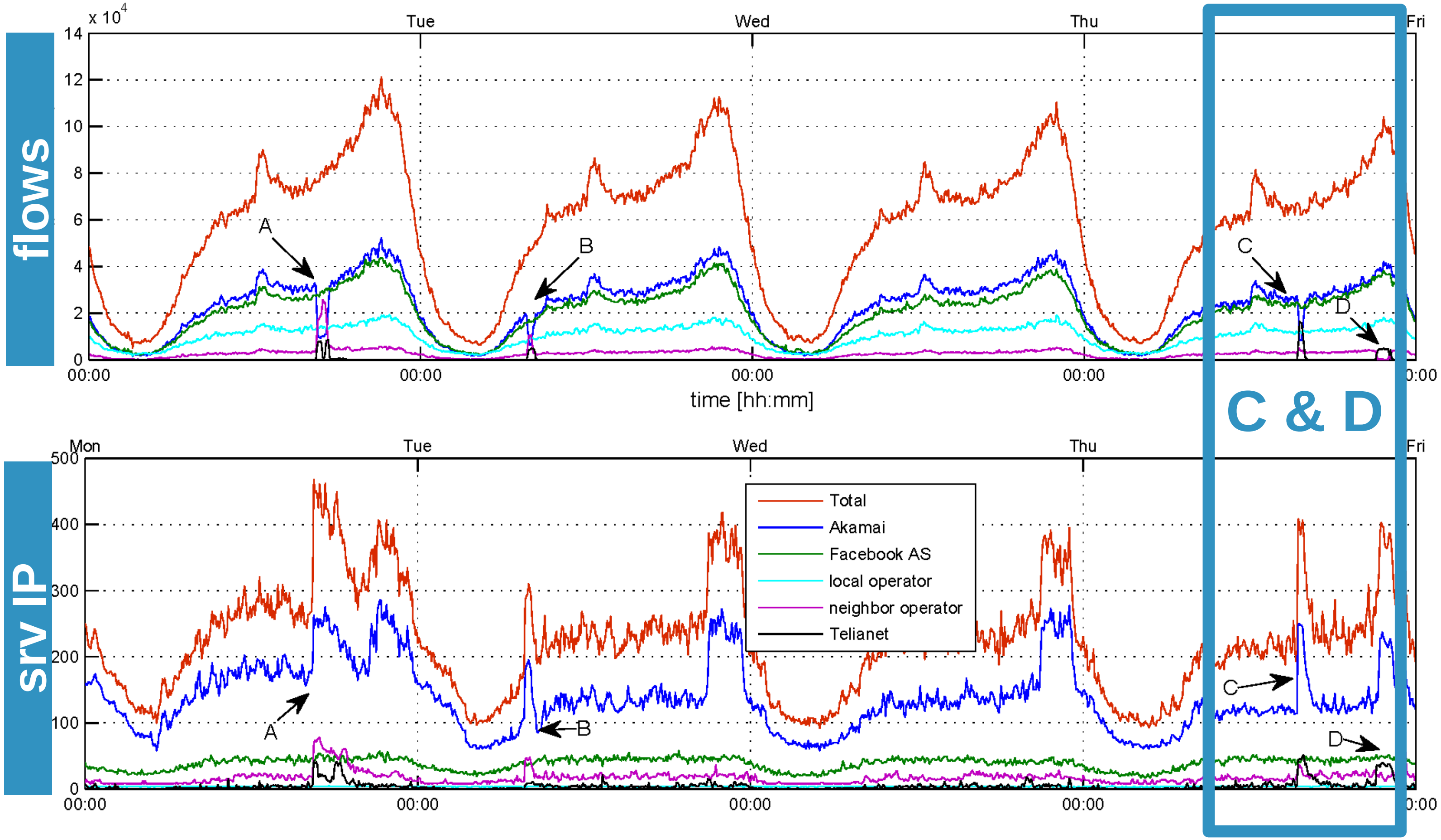


Events A and B

1. **Akamai inside the ISP** (drop in number of served flows, active IPs)
2. **Deutsche Telekom** and **TeliaNet** increase number of active IPs and **take over**

- CDNs have ~*constant* share of deployed IPs and number of flows
- Facebook AS and Akamai lead the number of served flows
- Akamai employs largest share of active IPs per time-bin

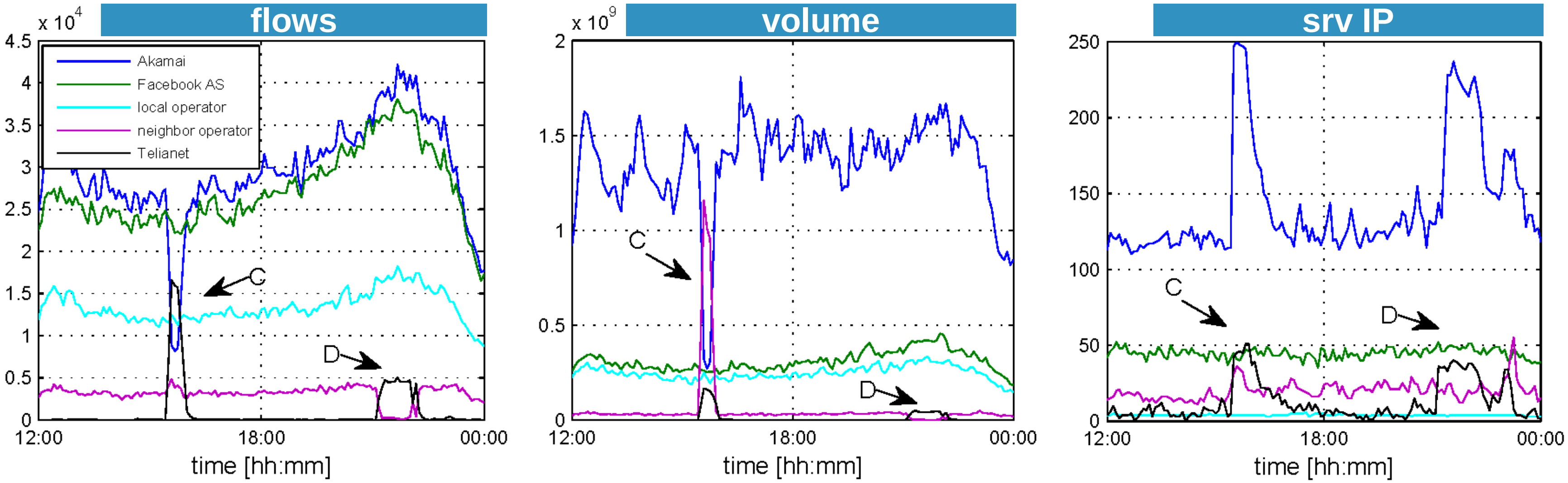
CDN Inter-play [1/3]
Time Series (4 days)



- CDNs have *~constant* share of deployed IPs and number of flows
- Facebook AS and Akamai lead the number of served flows
- Akamai employs largest share of active IPs per time-bin

CDN Inter-play [2/3]
Time Series (12 hours zoom-in)

- Zoom on last 12 hours:

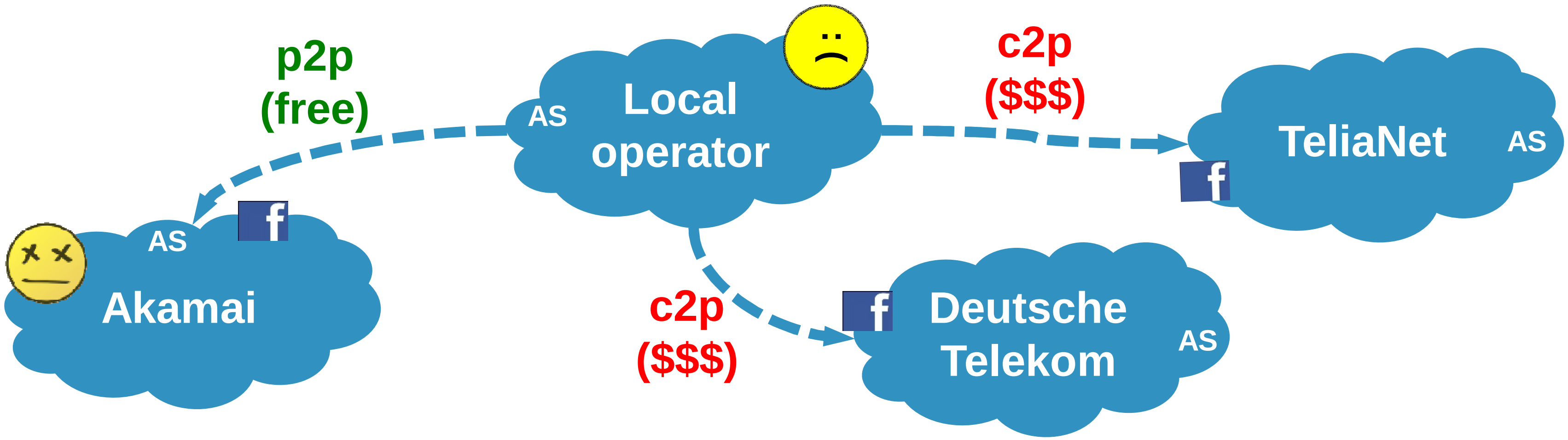


- Event C
  1. Akamai drops in number of flows, served volume but NOT active IPs
  2. TeliaNet increases number of active IPs, served number of flows and volume
  3. Deutsche Telekom keeps same number of active IPs, but increase served volume (takes over Akamai's larger flows)
- Event D
  - Akamai not involved
  - Swap between Deutsche Telekom and TeliaNet w.r.t. number of flows

CDN Inter-play [3/3]

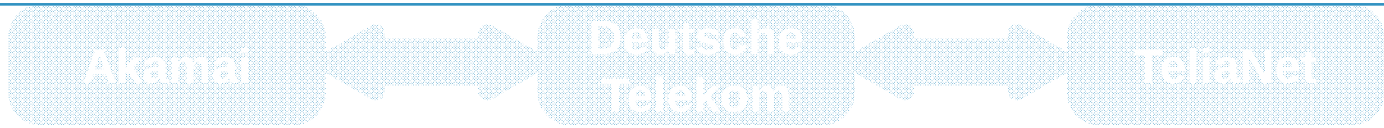
Potential Impacts on Transit Costs

- Events A-D reveal **chain of agreements in serving contents**
- According to Akamai policies, it is possible that **Akamai servers are installed in D.T. and TeliaNet networks** (Akamai directly manages the shift)
- **No performance impact from user pers**pective (normal RTT, throughput, number of erroneous HTTP response codes)
- But **different commercial agreements for peering**:



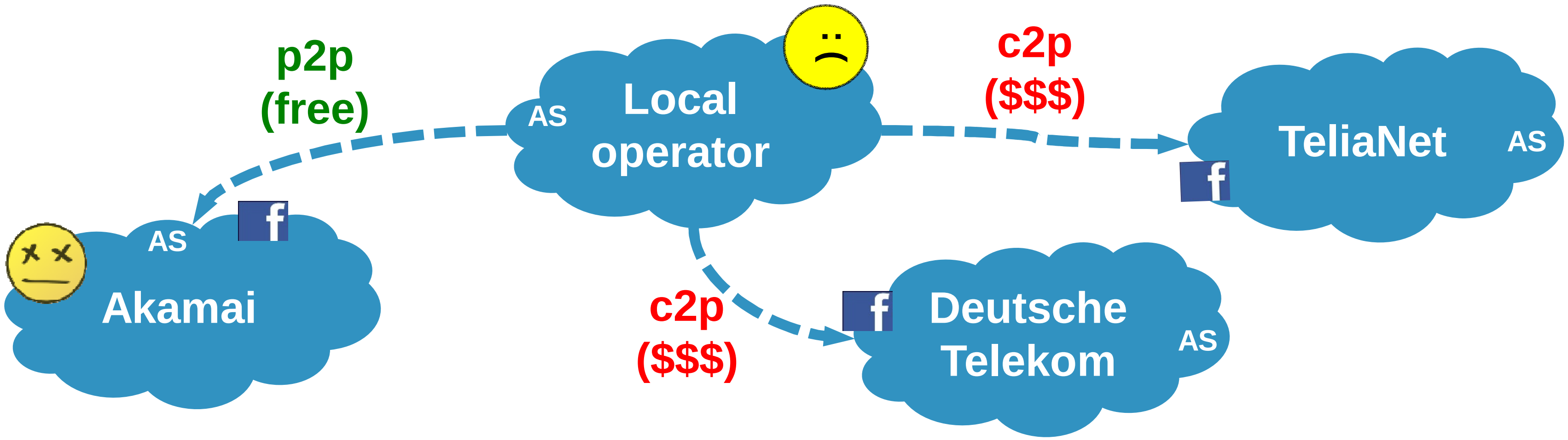
Topology from: http://irl.cs.ucla.edu/topology/

CDN Inter-play [3/3]
Potential Impacts on Transit Costs



- Events A-D reveal chain of agreements in serving contents
- No performance impact from user perspective (normal RTT, throughput, number of erroneous HTTP response codes)
- But different commercial agreements for peering:

Depending on the nature of **commercial agreements** for peering, it is possible that huge shifts of traffic volumes from one AS to another imply an **economical loss** for the ISP

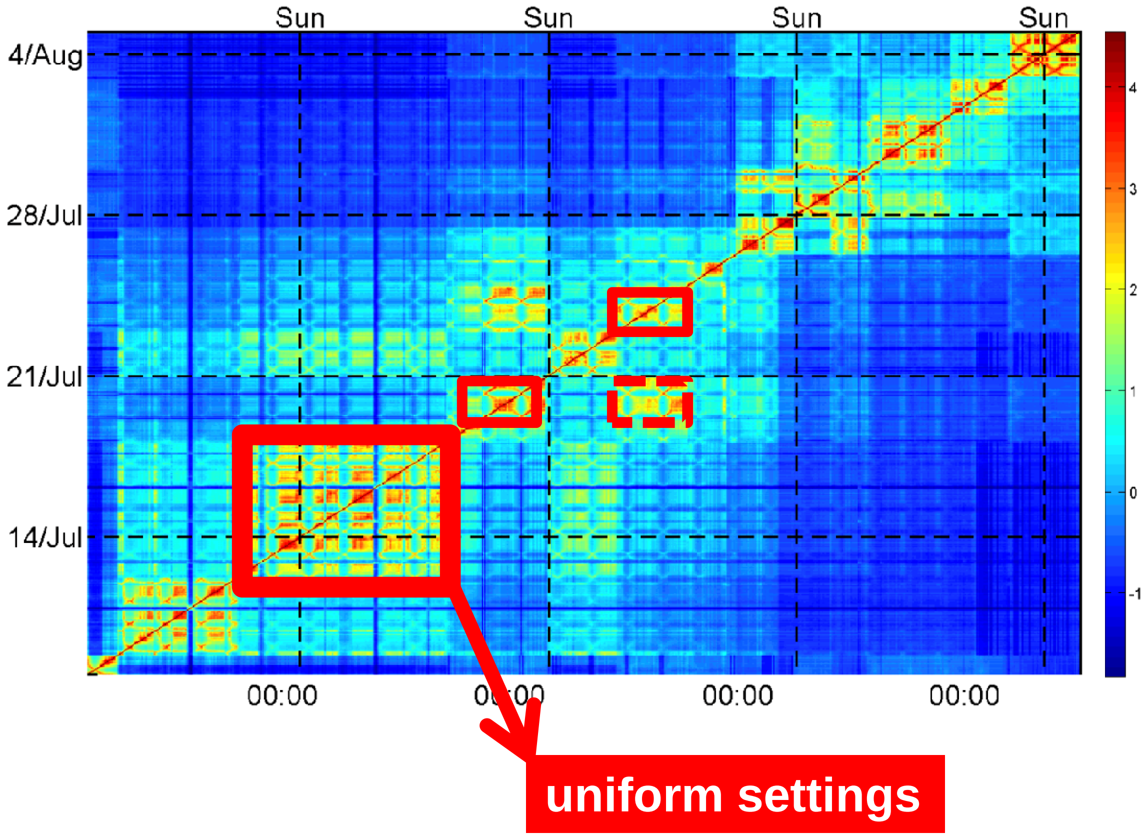


Topology from: http://irl.cs.ucla.edu/topology/

Temporal Similarity Plots (TSP)

A powerfull tool to visualize temporal patterns

- Discover **temporal patterns** and ***(ir)*regularities** in distribution timeseries



1. *For every IP: flow counts*
2. *Counters cumulated over different time scale (eg. 1hour)*
3. *For every time-bin: distribution of counters across IPs*
4. *Distribution compared with Kullback-Leibler metric*
5. *Comparisons plotted on heatmap (logscale)*

Characterization of Popular Services

The case of Whatsapp

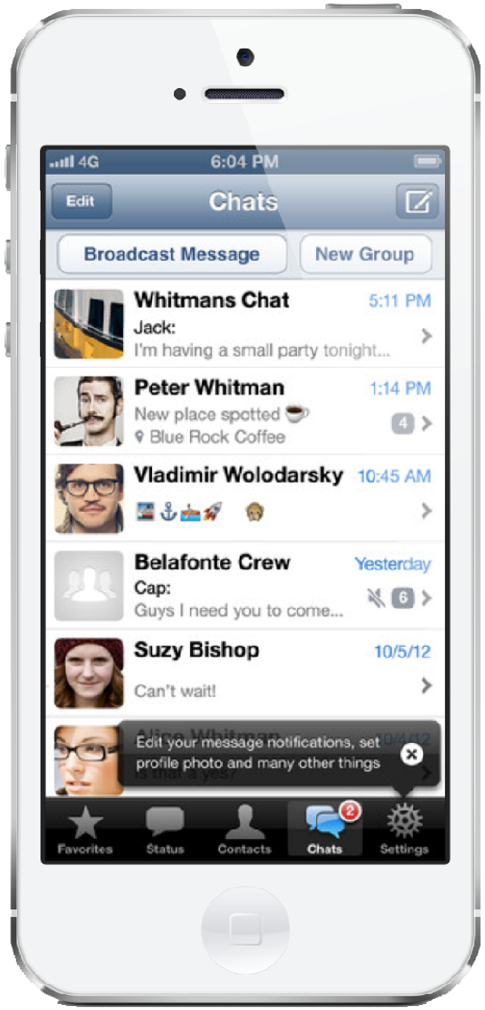
Whatsapp overview

Hard facts:

- 64 billion messages per day
  - 700 million photos
  - 100 million videos
- 500 million of daily active users
- Company with the quickest growing user base in history
- Acquired by Facebook for 19 billion $
  - Each user is worth 40$

Operators need to investigate it because:

- It is taking over (or already has...) the SMS/MMS market
- They need to learn how to track its usage
- They need to understand its impact on their networks



Reverse engineering Whatsapp naming scheme

Hybrid measurements

Testbed:

- Traffic (chat and medie exchange) actively generated at end devices (Android and iOS)
- Passively captured at a gateway (**Wireshark**)
- Focus on DNS requests

Findings:

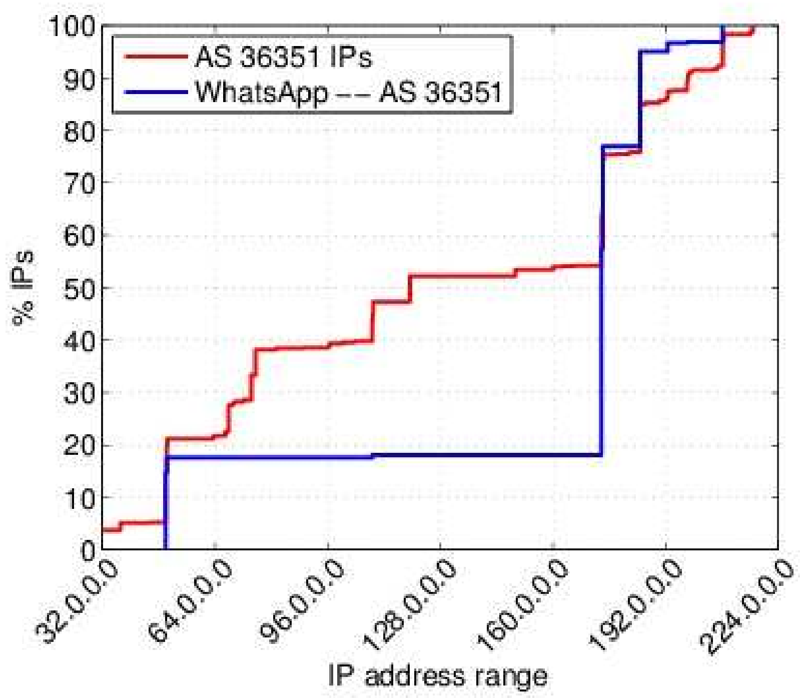
- Whatsapp used custom XMPP protocol
- Media exchange via HTTPS servers
- One persistent SSL connection to XMPP servers while the app is running
- Dedicated TLS connections to HTTPS servers for each media transfer

Servers naming scheme:

domain	prot. (port)	type
`cX, eX, dX`	XMPP(5222,443)	chat & control
`mmiXYZ,mmsXYZ`	HTTPS (443)	media (photo,audio)
`mmvXYZ`	HTTPS (443)	media (video)

Revealing Hosting Infrastructure

Through large-scale passive measurements



- 386 IP adresses used by Whatsapp (chat and media)
- All in AS36351 (Softlayer)

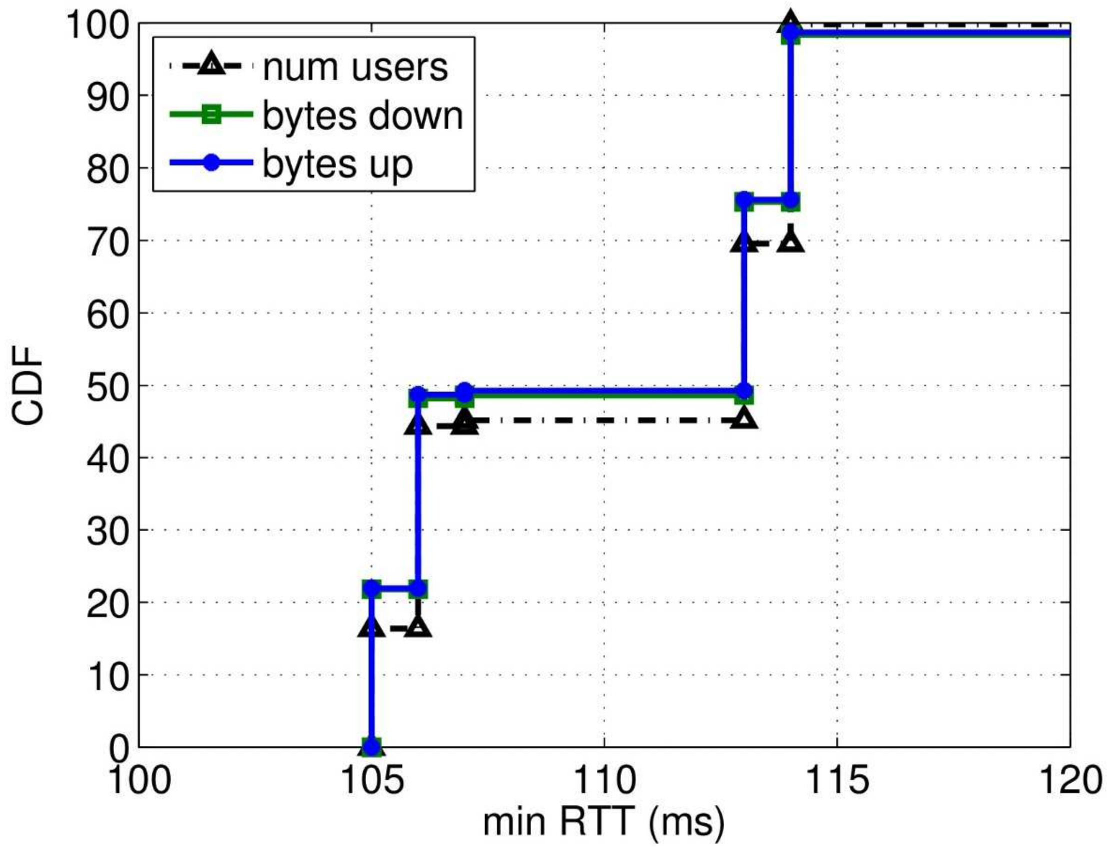
SOFTLAYER®
an IBM Company

Service/AS	#IPs	# /24	# /16	# /8
WhatsApp	386	51	30	24
SoftLayer (AS36351)	1364480	5330	106	42

Revealing Hosting Infrastructure
Through large-scale passive measurements

Creating Communication Technologies

Localization of servers through RTT measurements

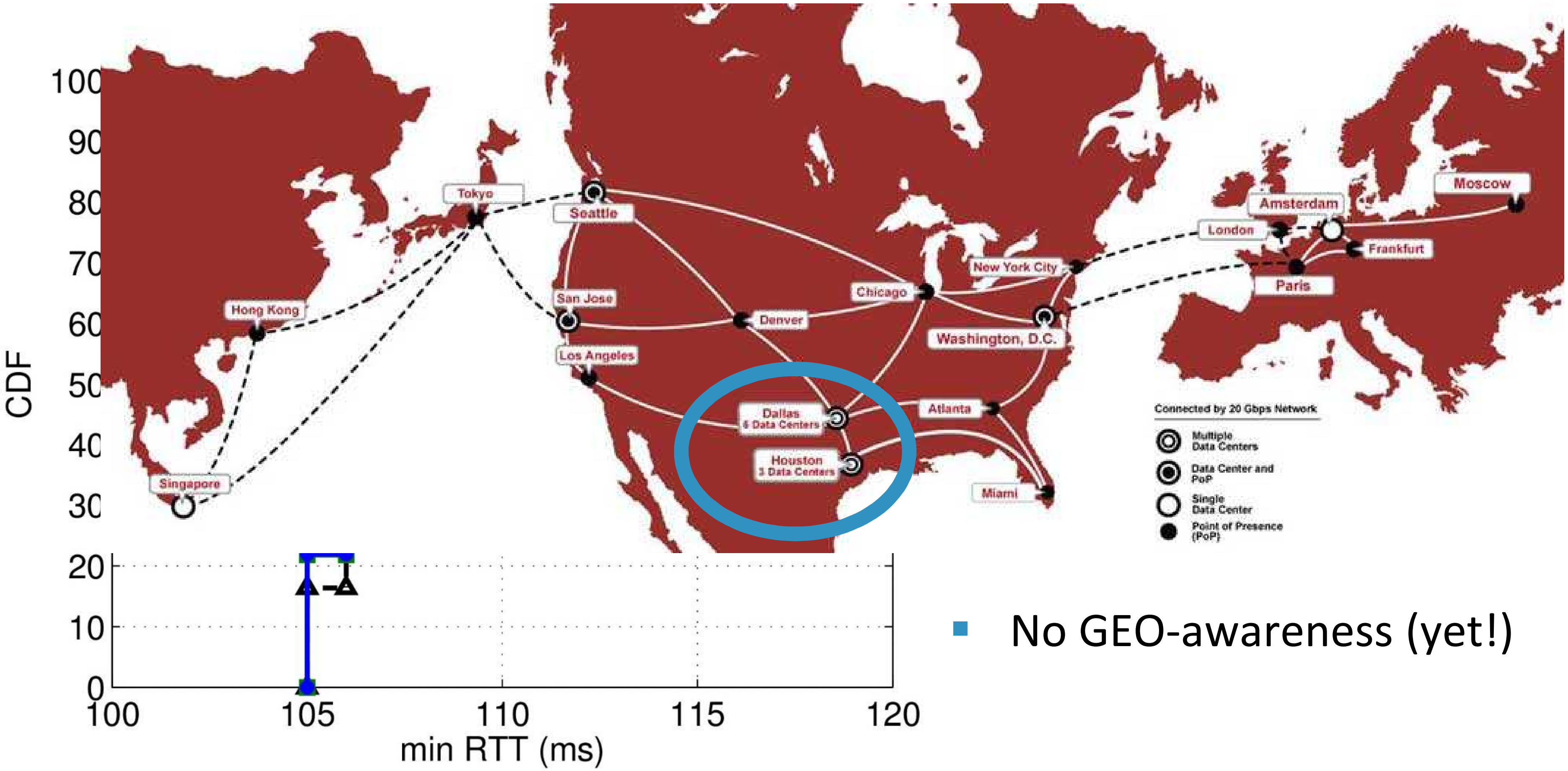


- ~400 IP addresses in Softlayer AS
- Two big steps in RTT distribution at 106ms and 114ms
- Localized by MaxMind in **Houston** and **Dallas** (Texas)
- No GEO-awareness (yet!)

Revealing Hosting Infrastructure

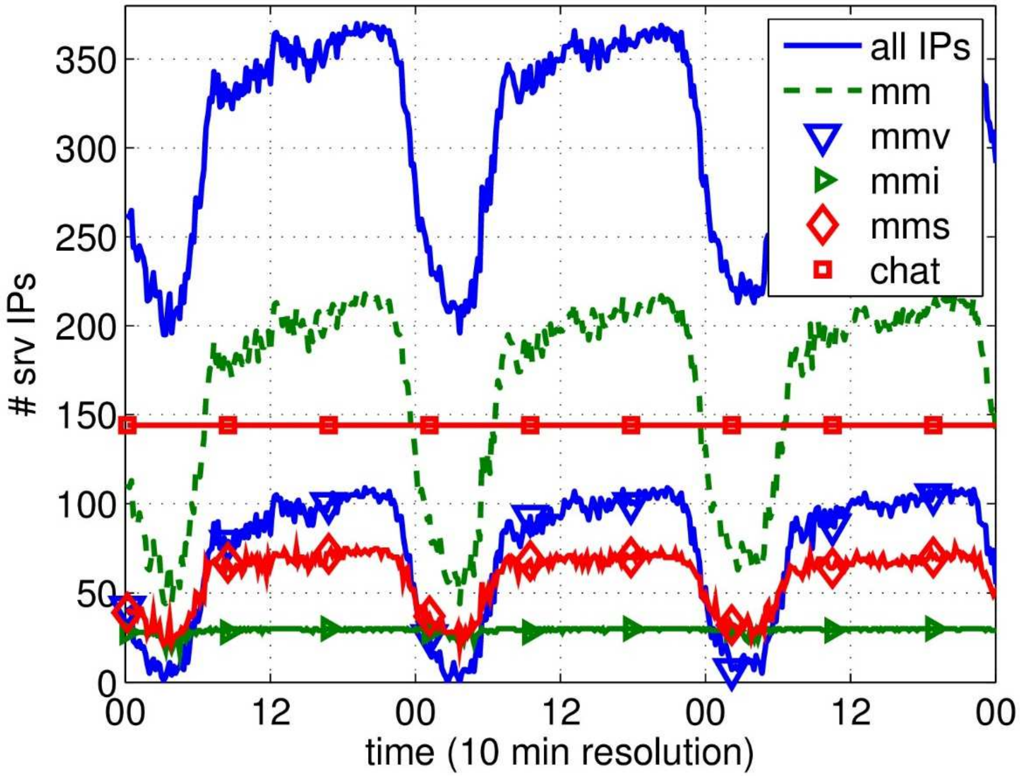
Through large-scale passive measurements

Localization of servers through RTT measurements



- No GEO-awareness (yet!)

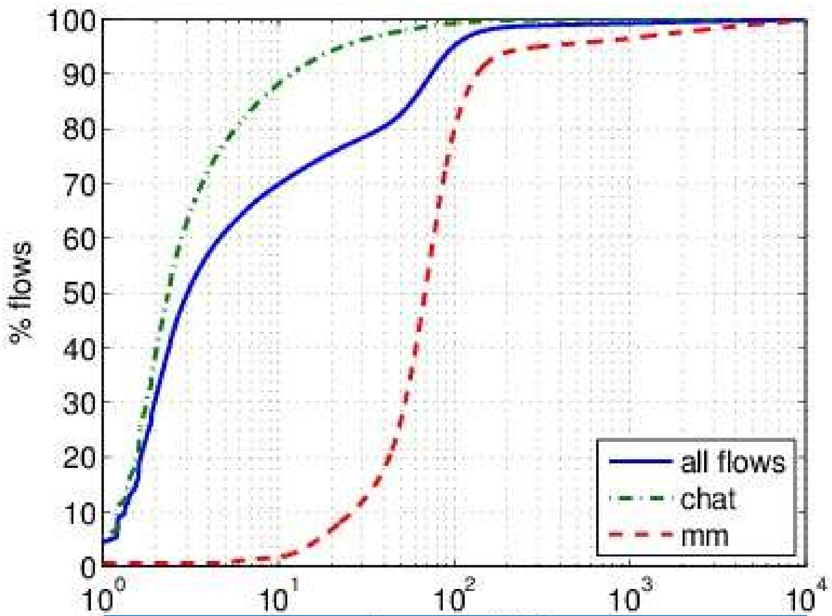
Active IPs



- More than 350 IPs during peak hours
- At least 200 IPs always active (chat servers)
- ~25 IPs always active (`mmi` servers)

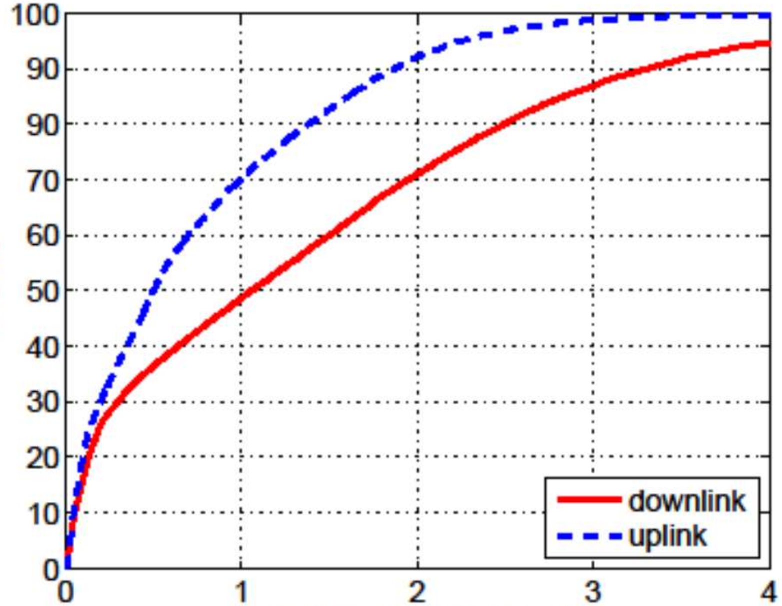
Whatsapp traffic characteristics

flow size and throughput



flow size

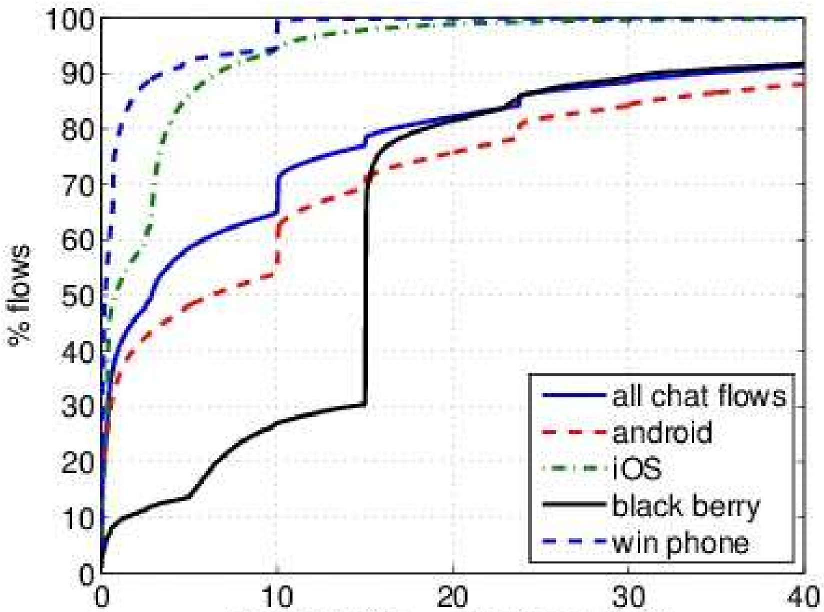
- Smaller chat/control flows and heavier mm flows
- 90% of chat flows < 10KB
- 50% of mm flows > 70KB



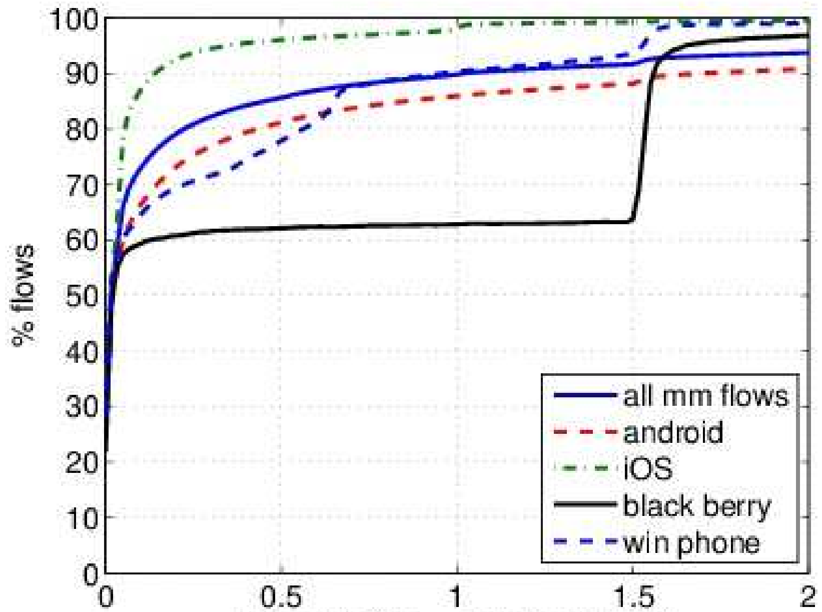
flow throughput

- Only bigger flows (<1MB) considered
- Up to 1.5Mbps in downlink
- Up to 800Kbps in uplink

Whatsapp traffic characteristics
flow duration with OS breakdown



flow duration (chat flows) [m]



flow duration (mm flows) [m]

Timeouts:

- Android: 10/15/25 min
- iOS: 3 min
- Blackberry: 15 min
- Windows Phone: 10 min

Timeouts:

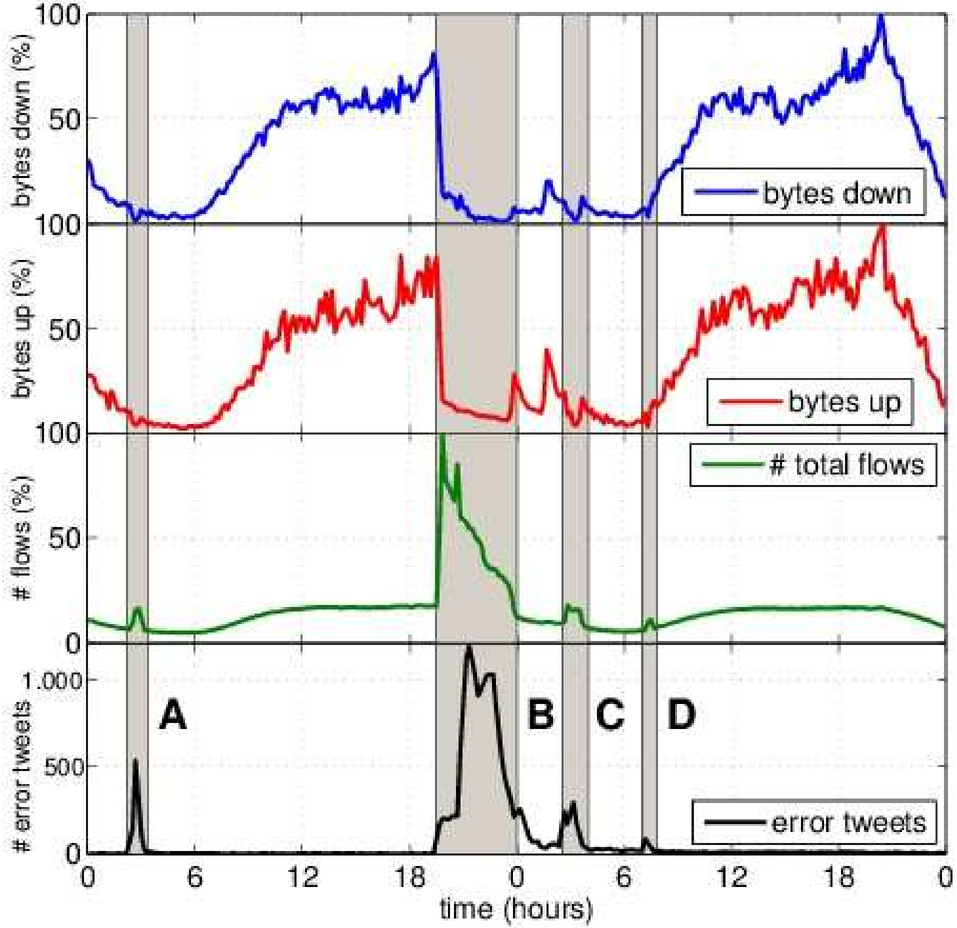
- Blackberry: 90 sec

The big outage (Feb. 22nd, 2014)

press reaction

ftw Creating Communication Technologies

The big outage (Feb. 22nd, 2014)
as seen from passive measurements and social feeds



drop in volume down

drop in volume up

ramp-up on flow counts

#whatsappdown

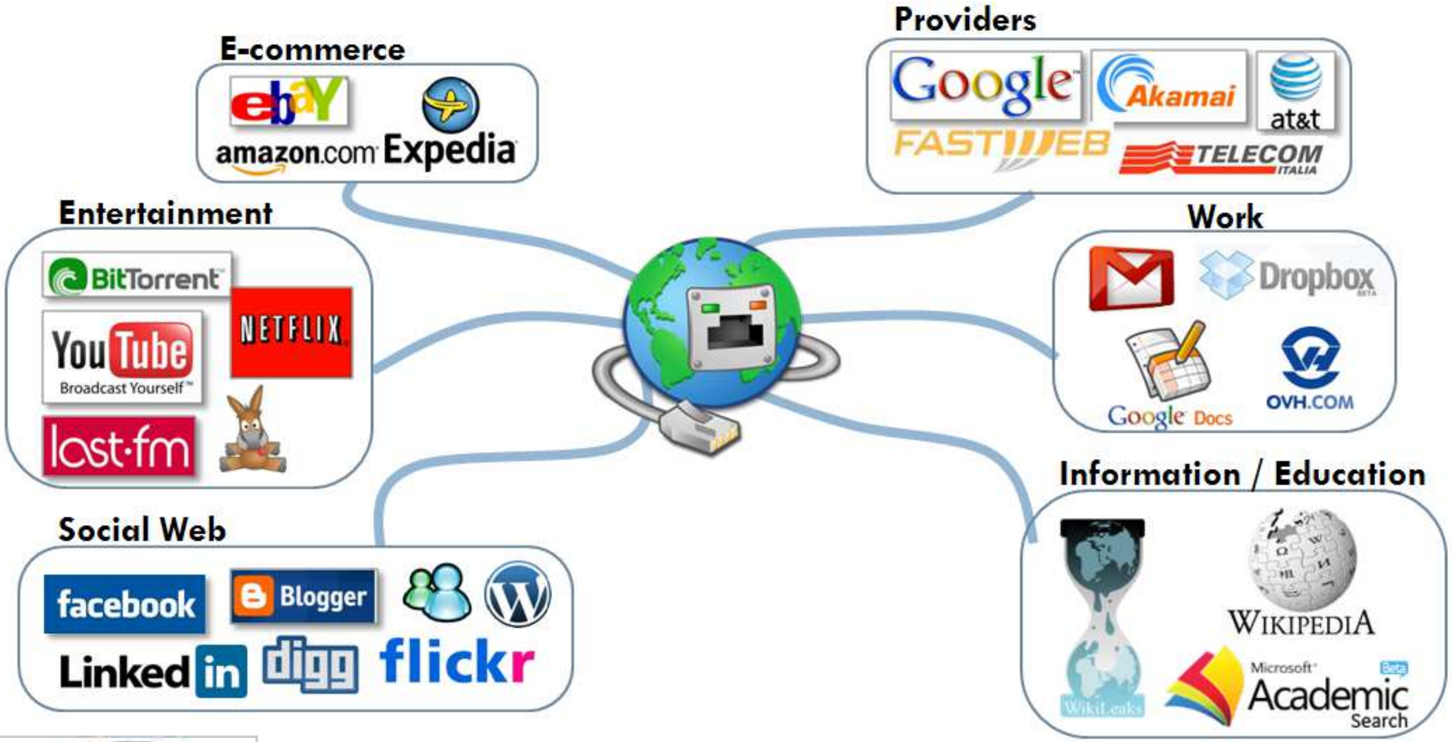
ftw Creating Communication Technologies

Large Scale
Traffic Monitoring and Analysis

mPlane – Building an Intelligent Measurement Plane for the Internet

mPlane

The nowadays Internet



"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 – President of Google

A complicated technology…

…that no one controls and understands

- Why Skype is not working?
- Which is the best ISP in my area?
- Where is YouTube traffic coming from?
- How to optimize my LTE network for Netflix

We need an intelligent system that **collects, analyzes**, **provides visibility to support** better management: **an oracle that provides answers!**

Understanding the Internet

- **How?**
  - Measuring and classifying network traffic – passive measurements
  - Testing network performance – active measurements
- **Where?**
  - Software/plugins installed by users @end devices
  - Network active probes @the edge
  - Measurements on network devices (e.g., routers)
- **What for?**
  - Troubleshooting
  - Traffic control
  - Anomaly detection
  - Performance evaluation
  - And more….

Understanding the Internet

What has been done so far?

Project	Objective			Approach		
Name	Network Mapping	Performance	Troubleshooting	SW plugin	Active probe	Passive at network devices
•Atlas •Archipelago •Merlin	✓				✓	
•Bismark •Dasu •M-Lab •Netalyzr		✓		✓		
•NetViews •RouteViews •TopHat •ASP	✓					✓
perfSONAR	✓	✓	✓			✓
CCAMP		✓		✓		✓
DIMES	✓			✓		
MOMENT		✓				✓

RIPE Atlas infrastructure
for geo-distributed active measurements

- **RIPE NCC**: Regional Internet Registry for Europe (equivalent of LACNIC)
- **RIPE Atlas**: a large measurement network composed of geographically distributed active probe used to measure connectability and reachabiltiy

http://atlas.ripe.net

RIPE Atlas probe v3
TP-Link MR3020 router with custom firmware

Understanding the Internet

EU projects

perfSONAR GÉANT — **Perf**ormance focused **S**ervice **O**riented **N**etwork monitoring **AR**chitecture – 2007-still running

MOMENT (Seventh Framework Programme) — **Mo**nitoring and **Me**asurement in the **N**ext generation **T**echnologies – **STREP**, 2007-2013

leone (Seventh Framework Programme) — From global measurements to local management – **STREP**, 2012 – still running

Project	Objective			Approach		
Name	Network Mapping	Performance	Troubleshooting	SW plugin	Active probe	Passive at network devices
mPlane	✓	✓	✓	✓	✓	✓

The mPlane project

- mPlane is an FP7 Integrated Project
  - 3 years project, started late 2012
  - 16 partners (8 industrial, 8 research)

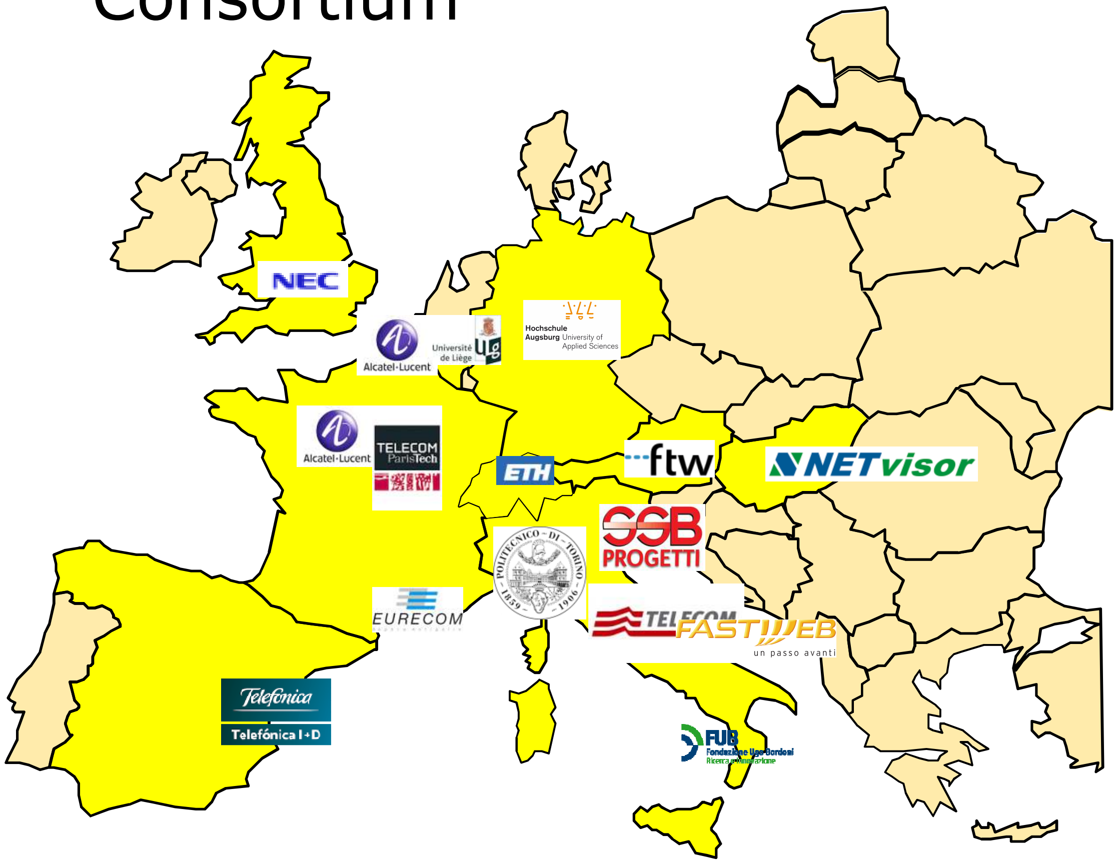
- Goal: design and demonstration of an "intelligent measurement plane for the Internet"
  - mPlane is about **large scale network measurements**,
  - and **intelligent big-data analysis** for troubleshooting support
  - **embedding measurement into the Internet as an additional capability**

ftw Creating Communication Technologies

Who we are?

General
Coordinator
Prof. Marco Mellia
Politecnico di Torino - IT

Consortium



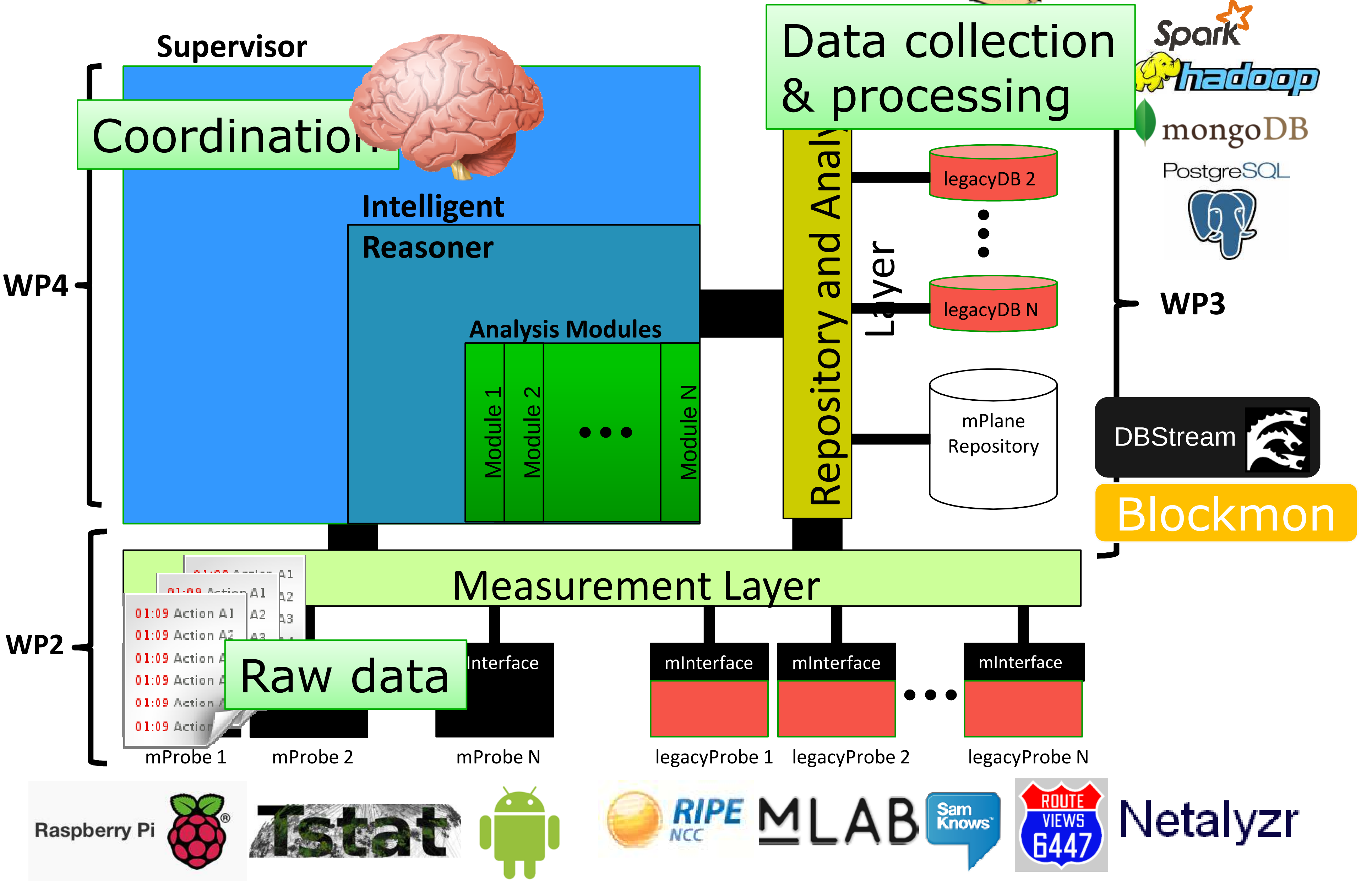
- 3 Constructors
- 3 Operators
- 2 SMEs
- 2 Research Centers
- 6 Research Groups

https://www.ict-mplane.eu

mPlane in a slide

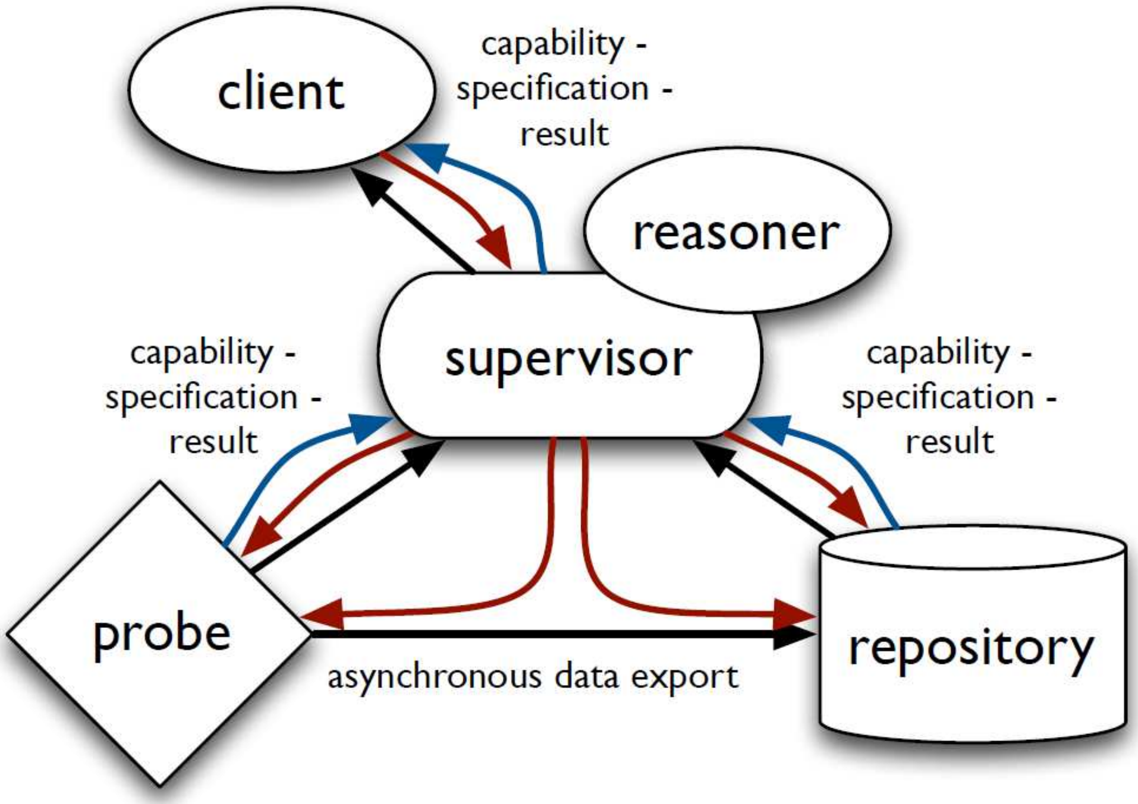
- Build a distributed, open, standard measurement infrastructure for the Internet
  - **Probes (WP2) – get the data**
    - **Build on existing tools/methodologies**
    - Offer a **flexible, programmable,** open platform to run and collect **passive, active, hybrid measurement**
  - **Repositories (WP3) – store and preprocess the data**
    - **Collect measurements** in a standard way
    - **Pre-process large amounts of data** in efficient ways
    - **Grant access to interested parties** (ISP, content providers, end-users, regulation agencies, etc.) subject to authorization rules
  - **Intelligent reasoner (WP4) – dig into the data**
    - Mine automatically the data and **extract useful information**
    - **Drill down to the root cause of a problem**
    - Allows **structured, iterative, and automated analysis**

mPlane in a picture



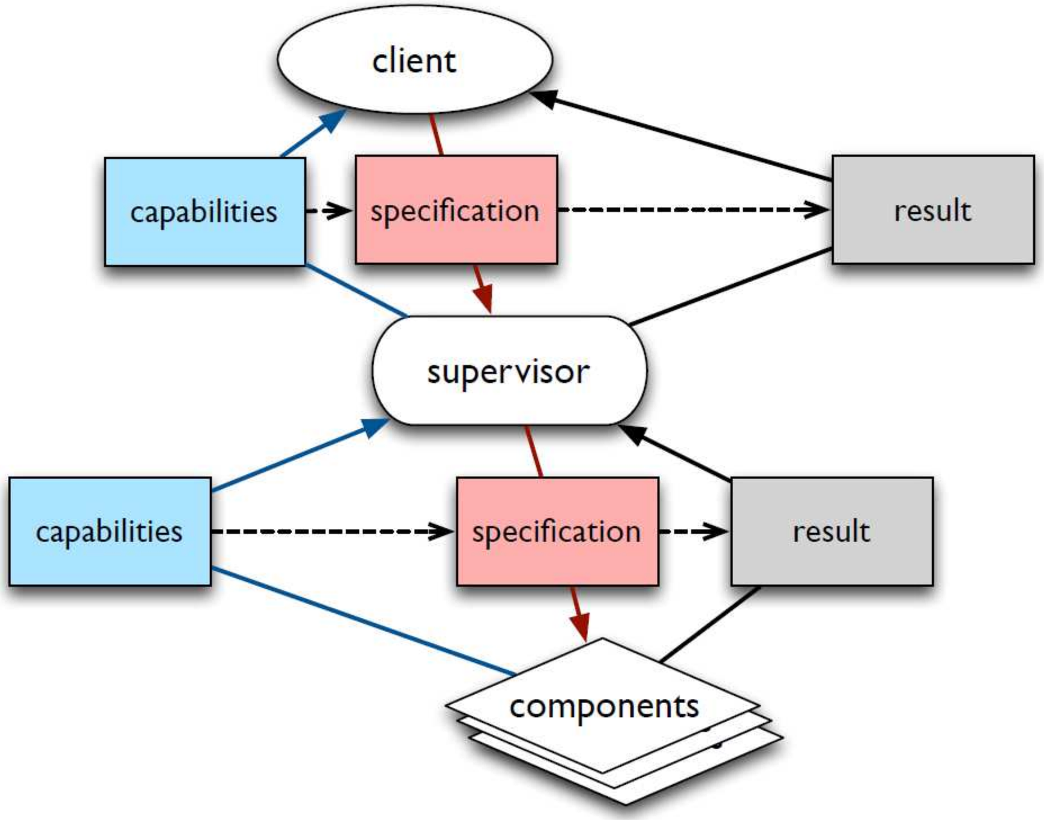
Some mPlane Architectural Details

An Overview on mPlane's Architecture



- **Components** and **interactions** in mPlane:
  - **blue lines** are **capabilities** announcements,
    red lines indicate **control messages (measurement sepecification)**,
    black lines correspond to **data**.

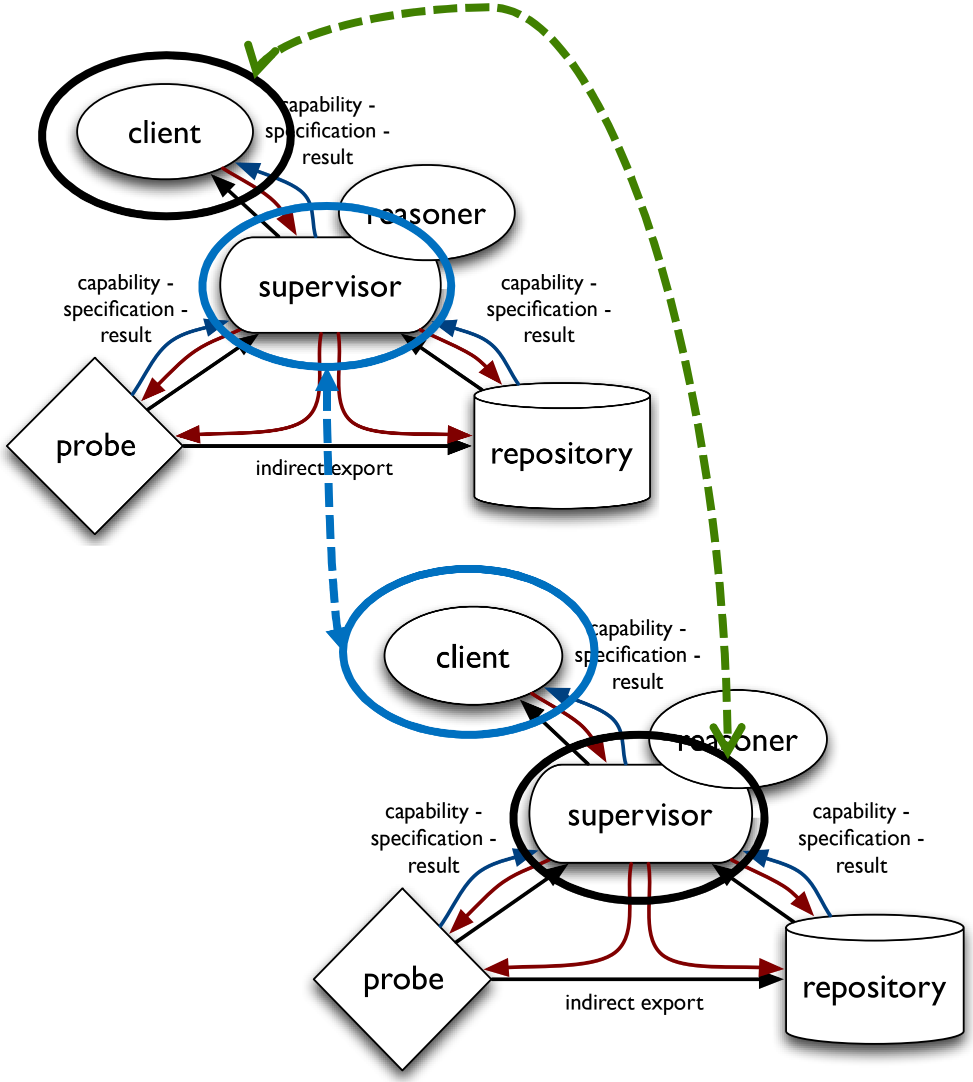
mPlane Workflow: how it works?



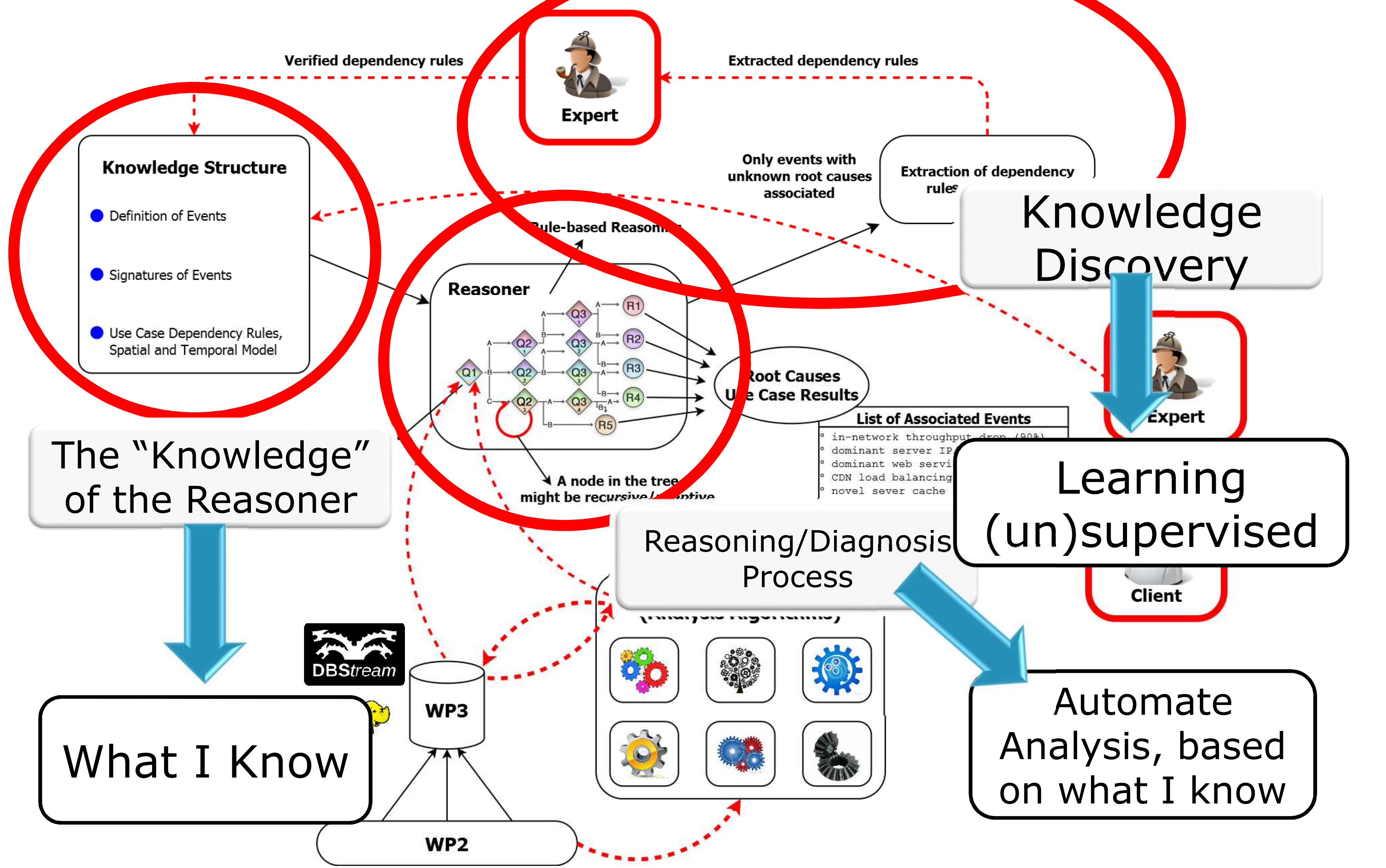
- **Capabilities** define the tasks a component can perform.
- **Specifications** consist of a description of which measurement have to be performed, how, and when.
- **Components announce their capabilities** when registering to the supervisor

mPlane inter-domain measurements

- **Each domain** collects and **owns its measurements**
- **Different mPlanes** under the control of **different players** (ISP, CDN, etc.)
- **Multi-domain measurements** handled as **communications among supervisors**



The Reasoner – The Overall Picture



Some of the mPlane Use Cases

- **Cloud Services** Troubleshooting
- **Mobile** Network **Performance** Troubleshooting
- **Web Browsing QoE** Troubleshooting
- Traffic **Anomaly Detection and Diagnosis**
- Multimedia **Content Delivery** Troubleshooting
- Content **Popularity Estimation & Caching**
- **SLA** Verification and Certification

Who benefits from mPlane?

- mPlane benefits everyone:
  - **ISPs** get a fine-grained picture of the network status, empowering effective management, operation, and troubleshooting.
  - **Content and Application providers** gain powerful tools for handling performance issues of their delivery systems and applications.
  - **Regulators** and end-users can verify adherence to SLAs, even when these involve many parties.
  - **Customers** of all kinds can objectively compare network performance, improving competition in the market.
  - **The Research Community** gets a system to accelerate the pace of research driven by Internet measurements

mPlane Case Study

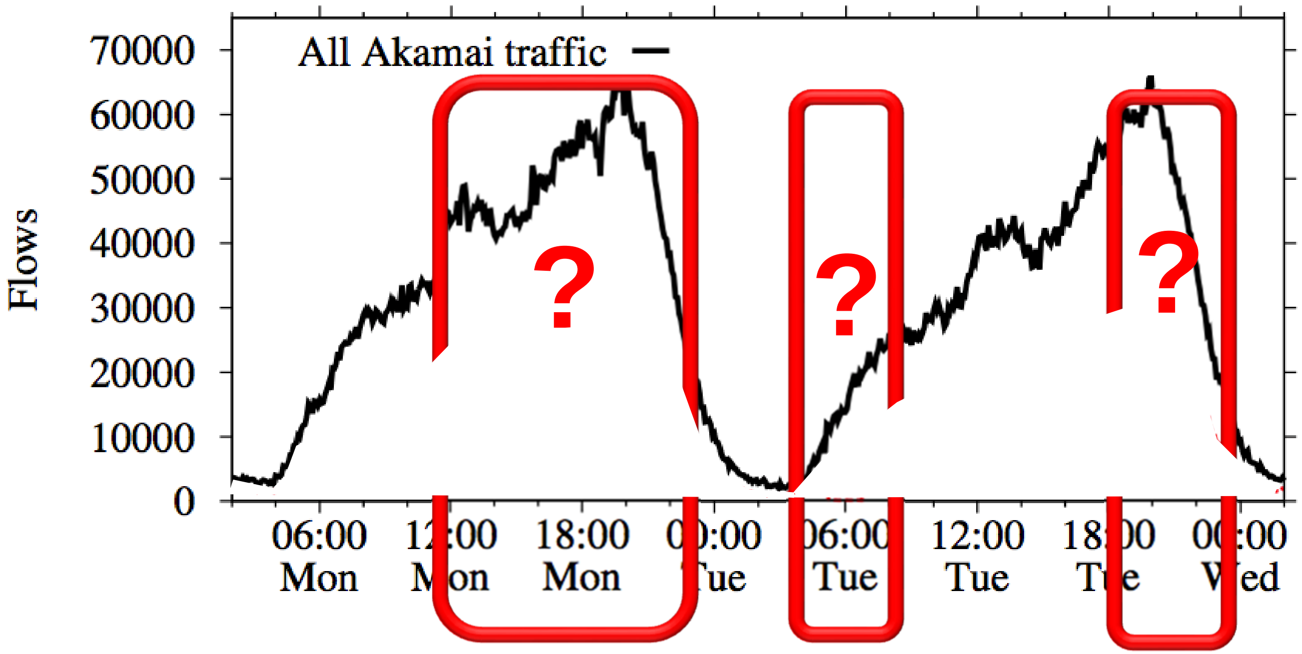
Understanding Akamai Cache Selection

Case-study: tracking CDN behaviour

- Internet: **large-scale web apps** and **Content Delivery Networks (CDNs)**
- Internet content (**YouTube, Facebook, Apple Store**) is largely **delivered by major CDNs** like **Akamai** and **Google CDN**
- **CDN's dynamics pose a challenge for ISPs** as they impact traffic engineering and possibly end-user QoE → it's **worth tracking and diagnosing shifts in the CDN traffic**

CDN makes complicated things

- Focusing on vantage point of ~20k ADSL customers
- 1 week of HTTP logs (May 2012), captured through Tstat
  - Content served by Akamai CDN
  - The ISP hosts an Akamai “preferred cache” (a specific /25 subnet)



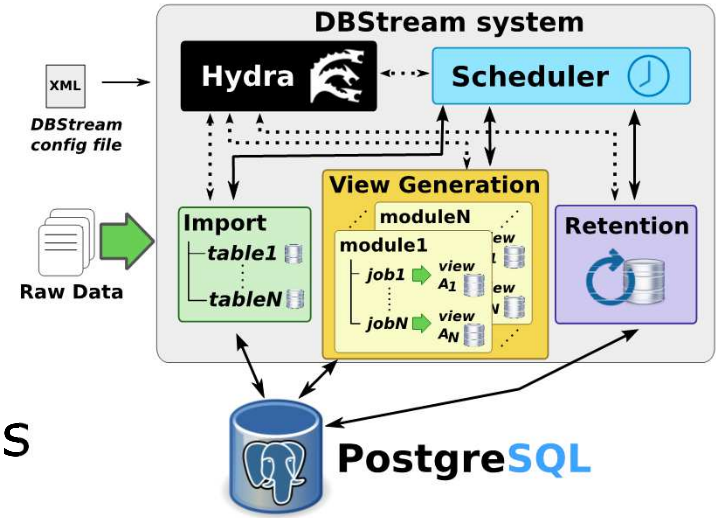
Reasoning about the problem

- Q1: Are the variations due to "faulty" servers?
- Q2: Is this affecting specific services?
- Q3: Was this triggered by CDN performance issues?
- Etc…

How to automate/simplify this reasoning?

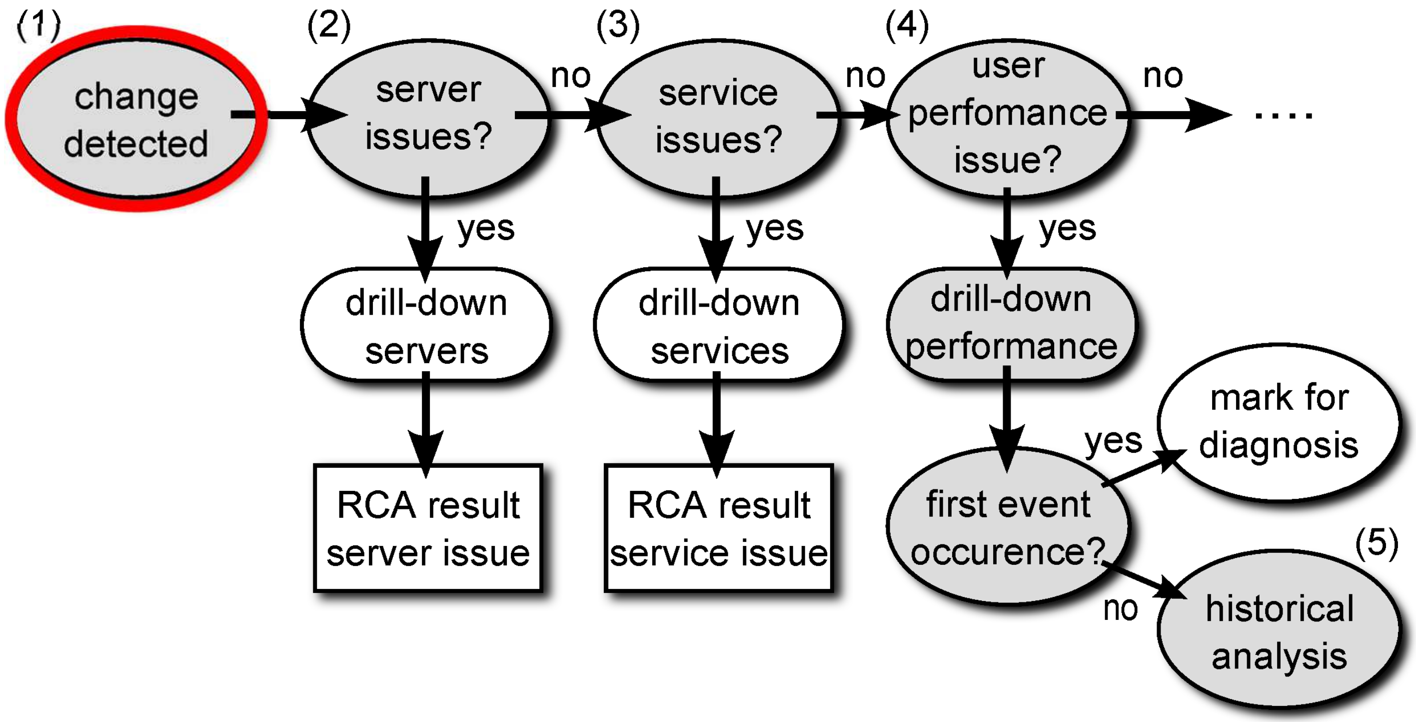
Reasoner + DBStream + Tstat:

- Continuous big data analytics
- Flexible processing language
- Full SQL processing capabilities
- Processing in small batches
- Storage for post-mortem analysis



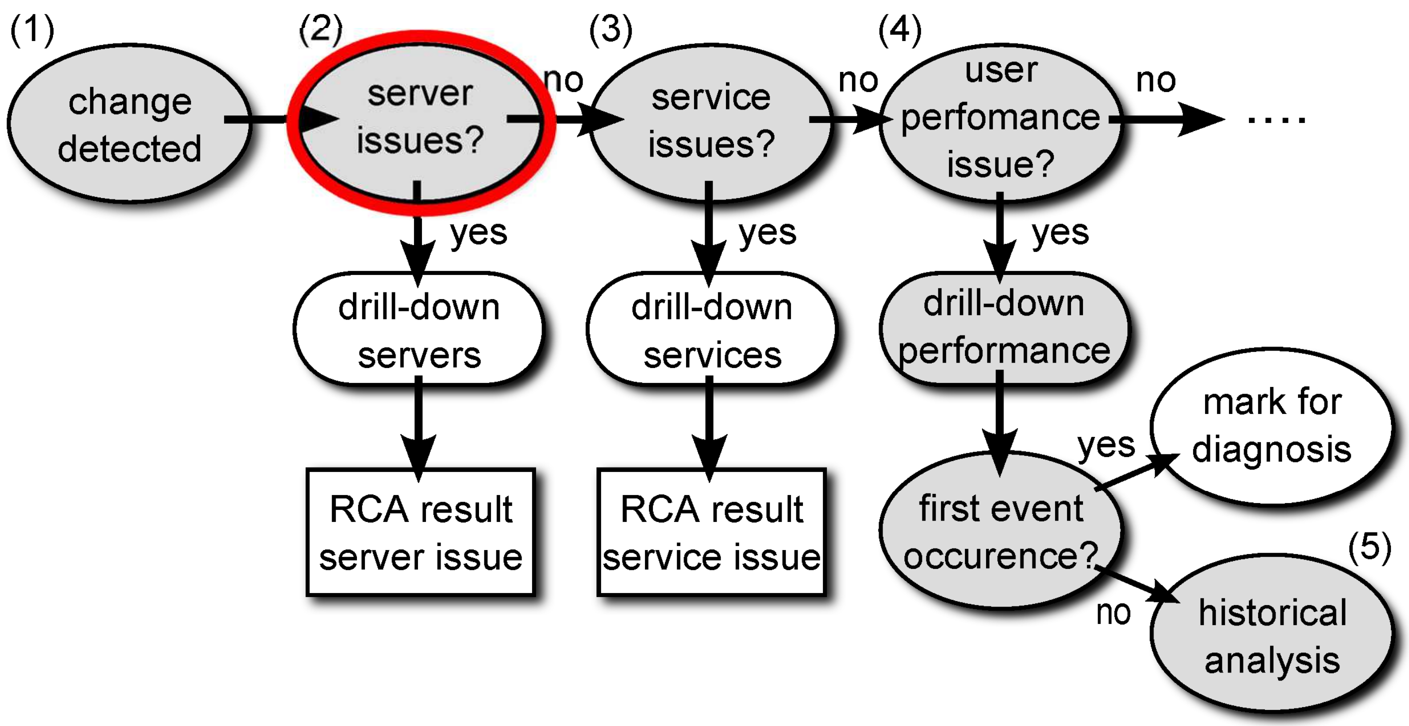
Shift in the Akamai served traffic

- Iterative analysis performed by the reasoner
  - Following a tree-like structure



Shift in the Akamai served traffic

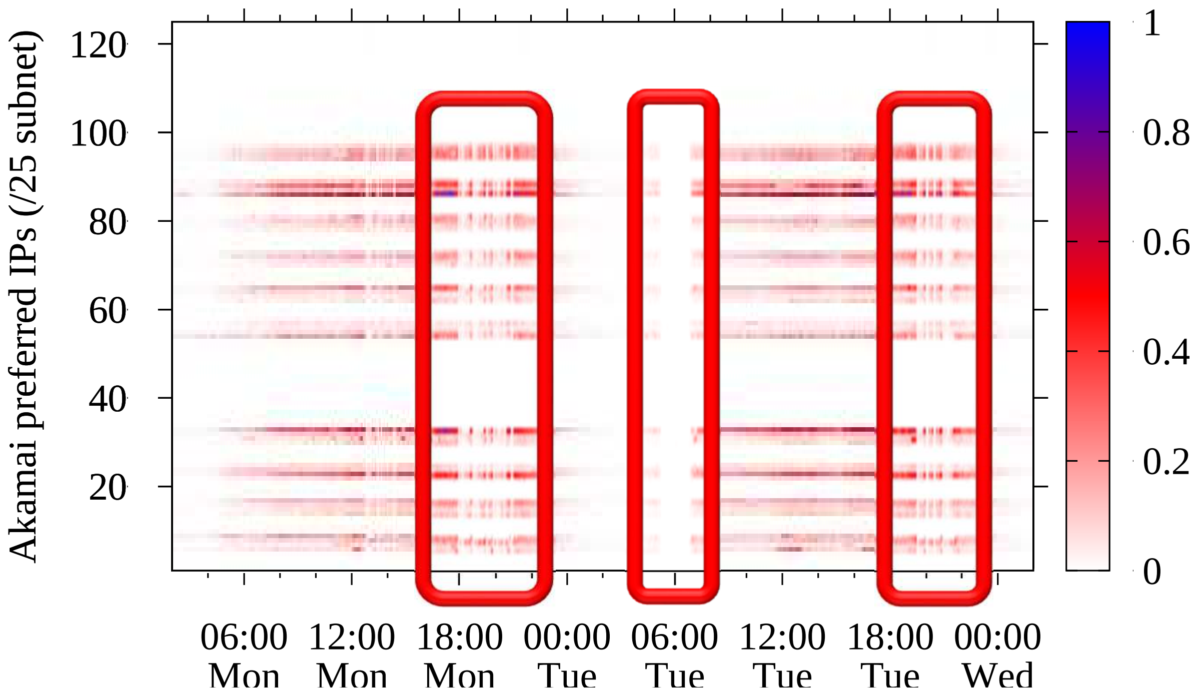
- Iterative analysis performed by the reasoner
  - Following a tree-like structure



Q1: Are the variations due to "faulty" servers?

NO

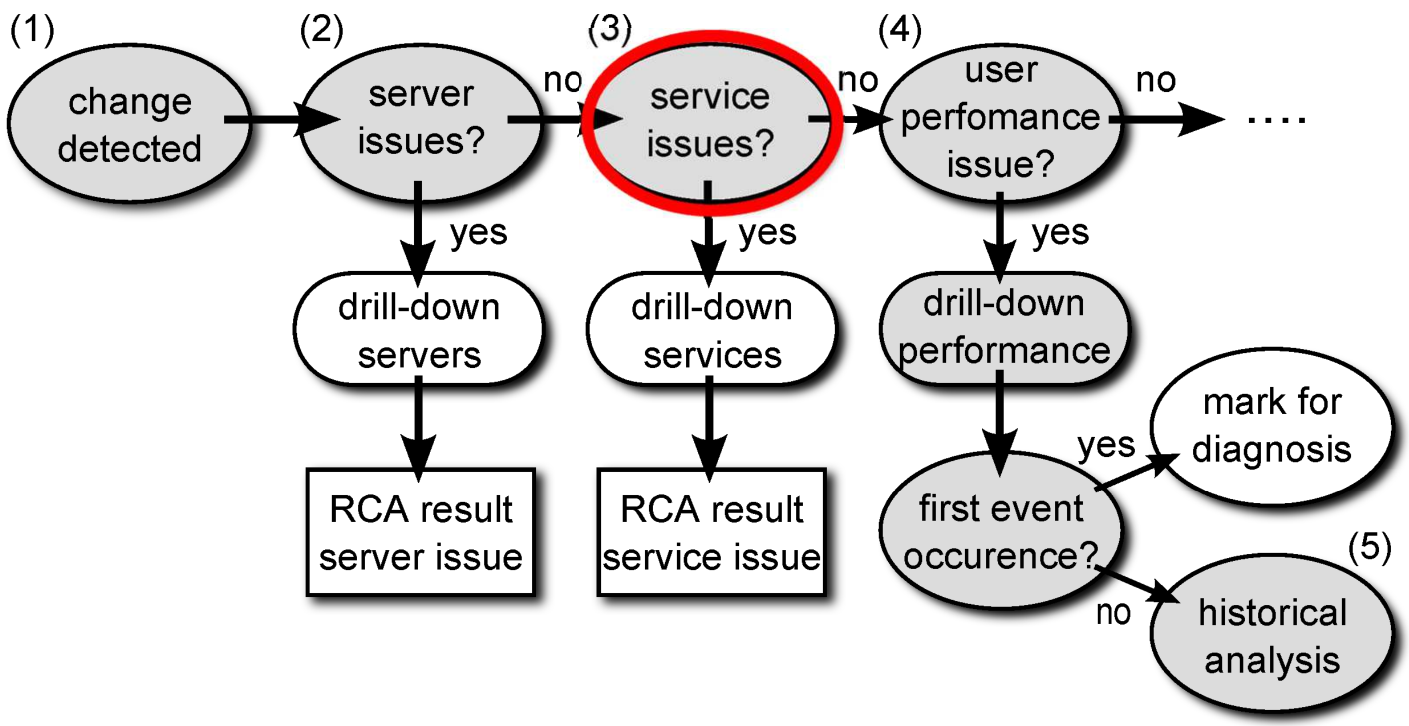
- Compute the traffic volume per IP address
- Check which are the active IPs during the disruption
- Repeat each 5 min
- 40 servers always active handle 62% of traffic



mPlane

Shift in the Akamai served traffic

- Iterative analysis performed by the reasoner
  - Following a tree-like structure

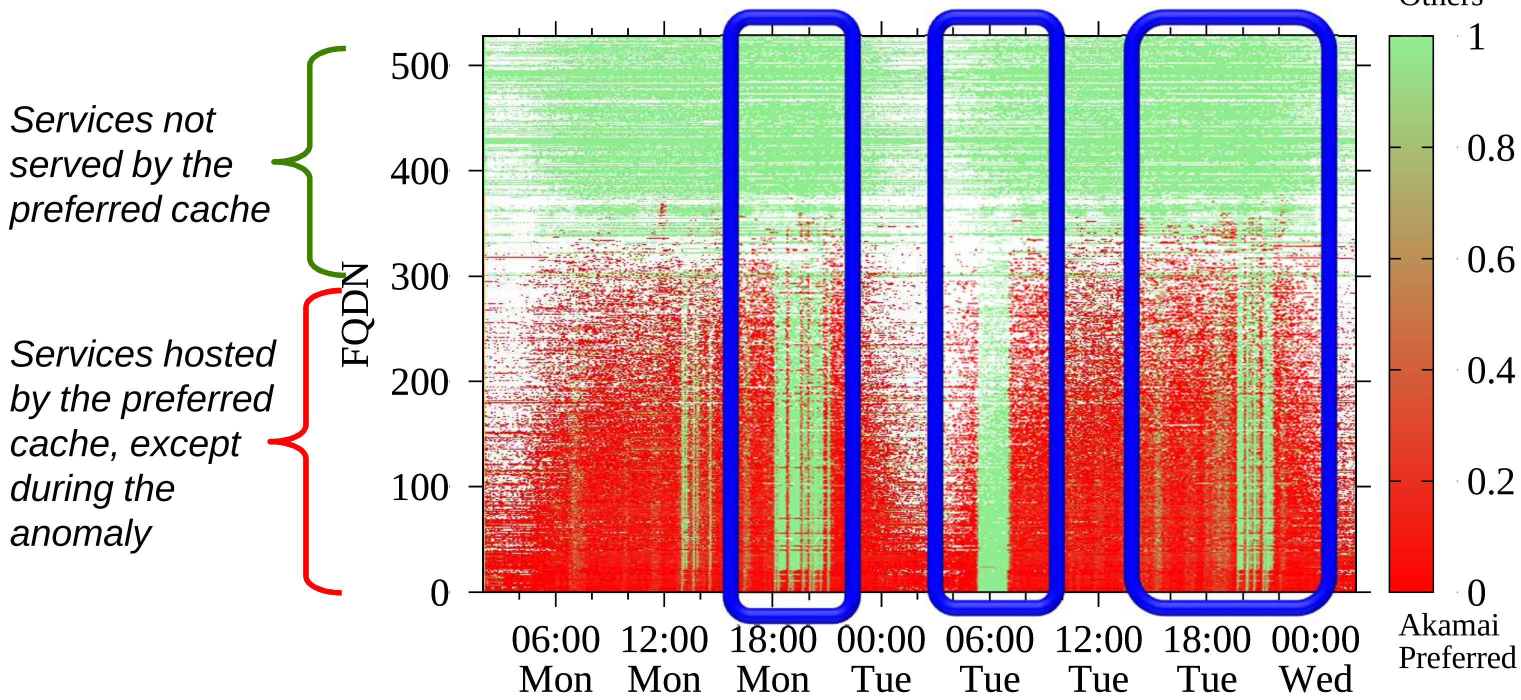


Q2: Is this affecting a specific service?

NO

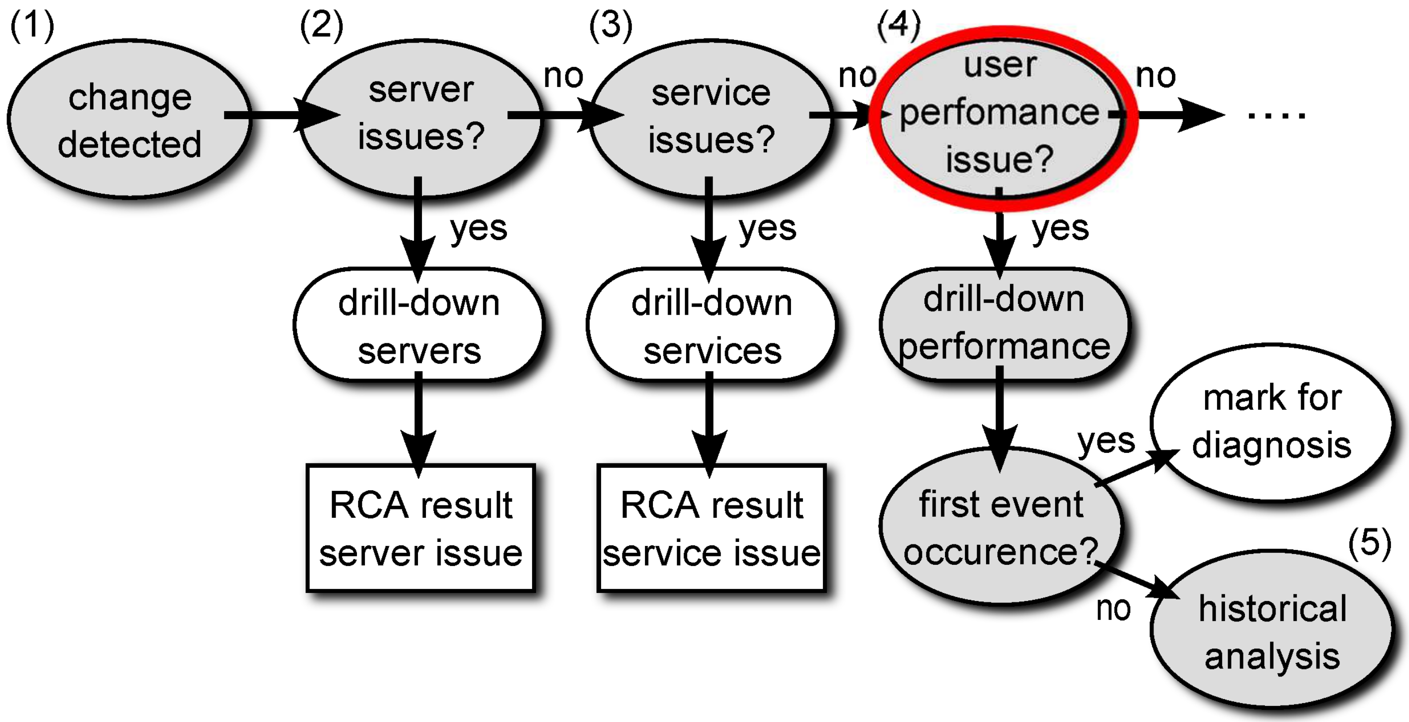
- Select the top 500 Fully Qualified Domain Names (FQDN) served by Akamai
- Check if they are served by the preferred cache
- Repeat every 5 min

The anomaly is not related to individual services



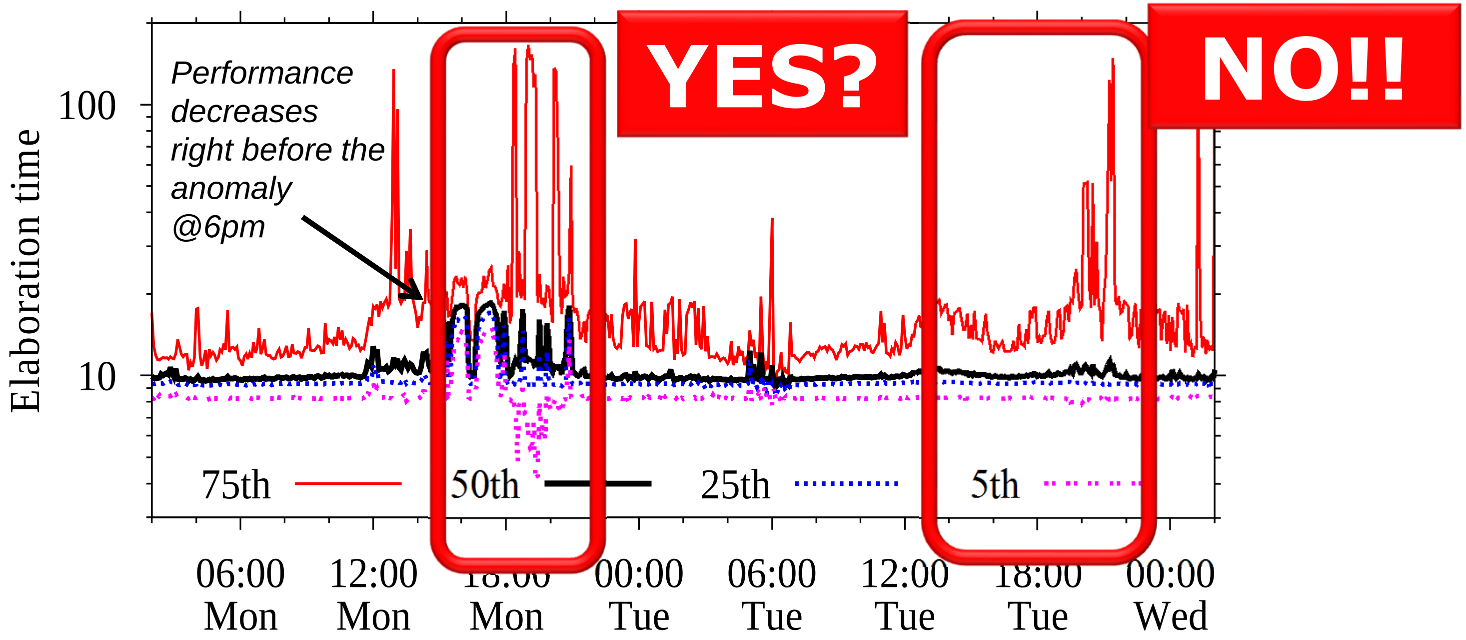
Shift in the Akamai served traffic

- Iterative analysis performed by the reasoner
  - Following a tree-like structure



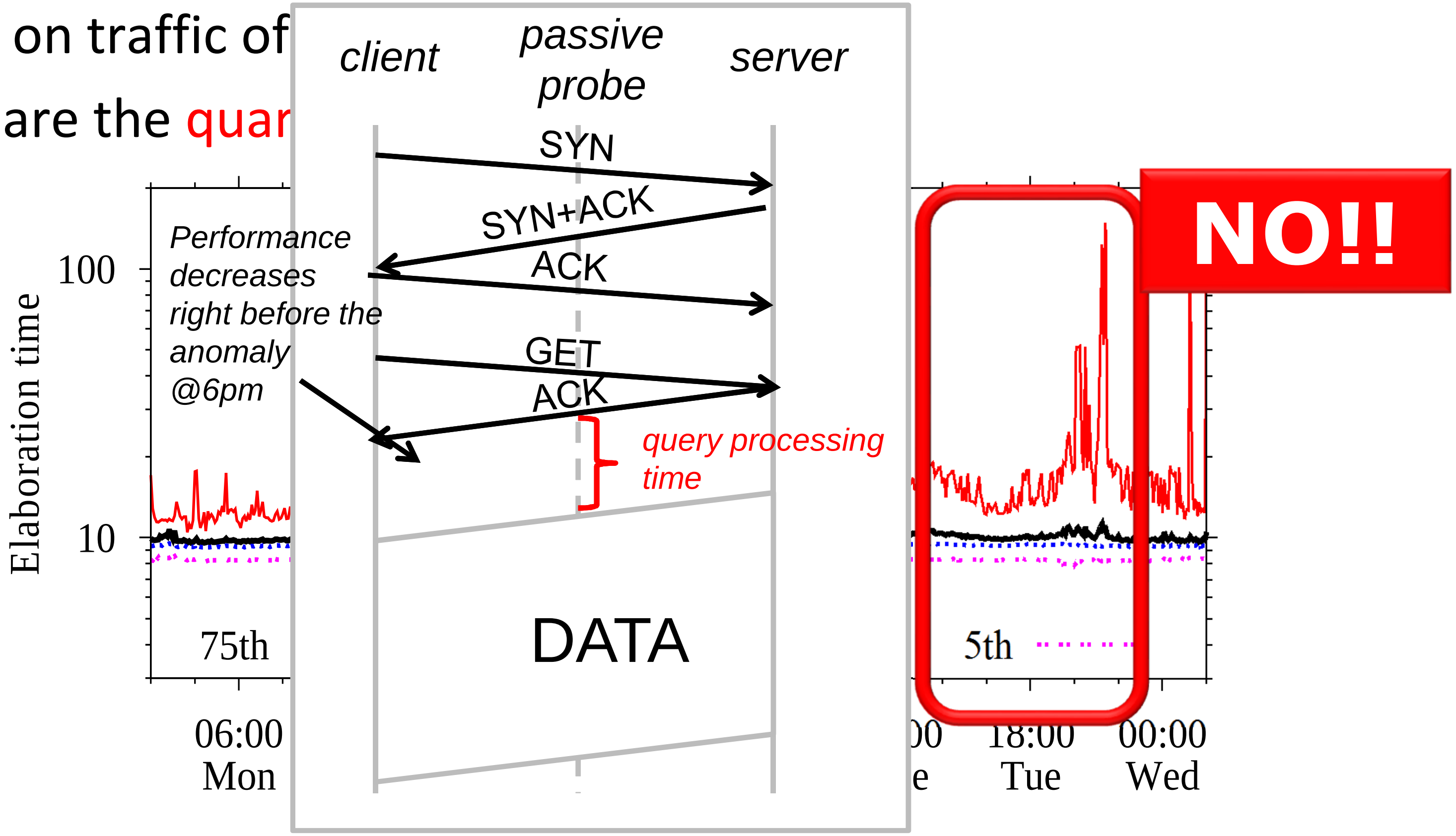
Q3: Was this triggered by CDN **performance issues?**

- Compute the distribution of server elaboration time
  - It is the time between the TCP ACK of the HTTP GET and the reception of the first byte of the reply
- Focus on traffic of the /25 preferred subnet
- Compare the quartiles every 5 min



Q3: Was this triggered by CDN performance issues?

- Compute the distribution of server elaboration time
  - It is the time between the TCP ACK of the HTTP GET and the reception of the first byte of the reply
- Focus on traffic of
- Compare the quar

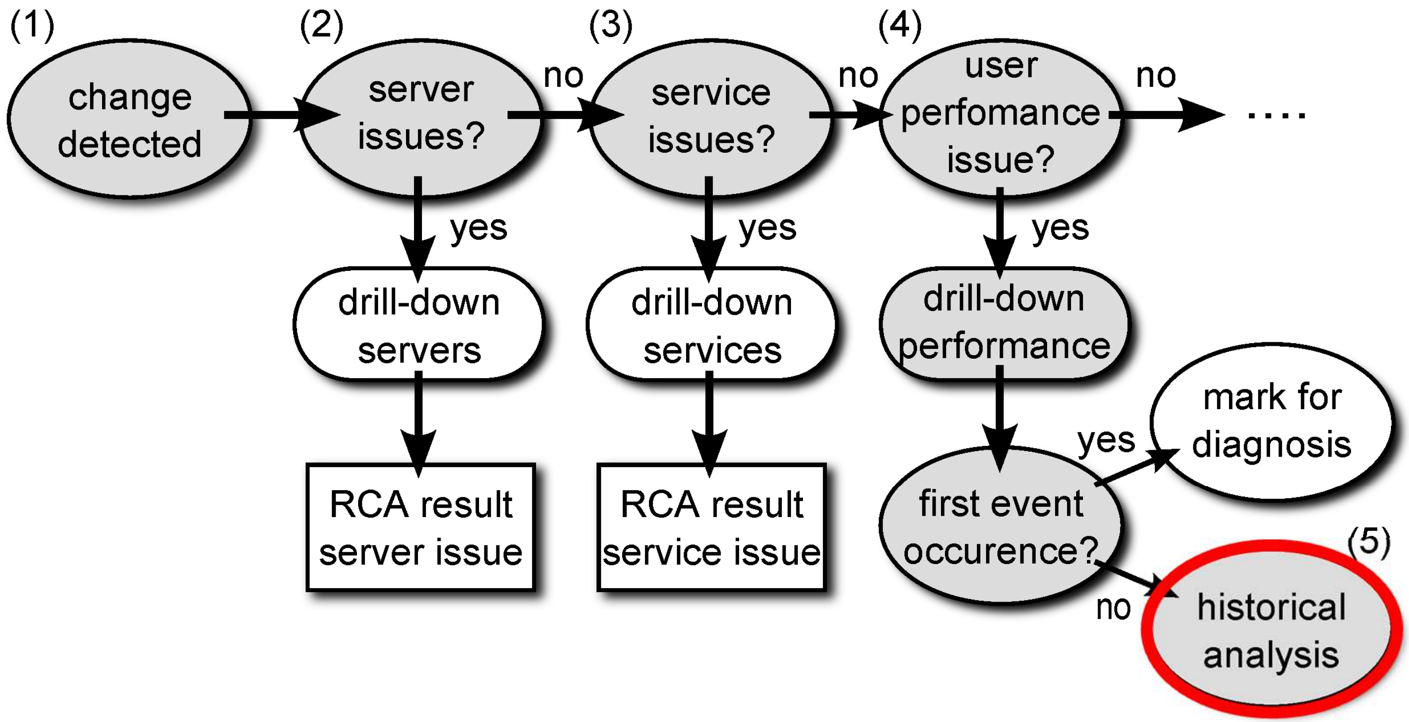


Reasoning about the problem

- Q1: Are the variations due to "faulty" servers? **NO**
- Q2: Is this affecting only specific services? **NO**
- Q3: Was this triggered by CDN performance issues? **NO**
- What else?
  - Other vantage points report the same changes? YES!
  - What about extending the time period?
    - The anomaly is present along the whole period we considered
    - Extension of the analysis on more recent data sets (possibly exposing also other effects/anomalies)
  - Routing? Not in this example → Integrating Route Views
  - DNS mapping? → Integrating Ripe Atlas + ISP active probing infrastructure

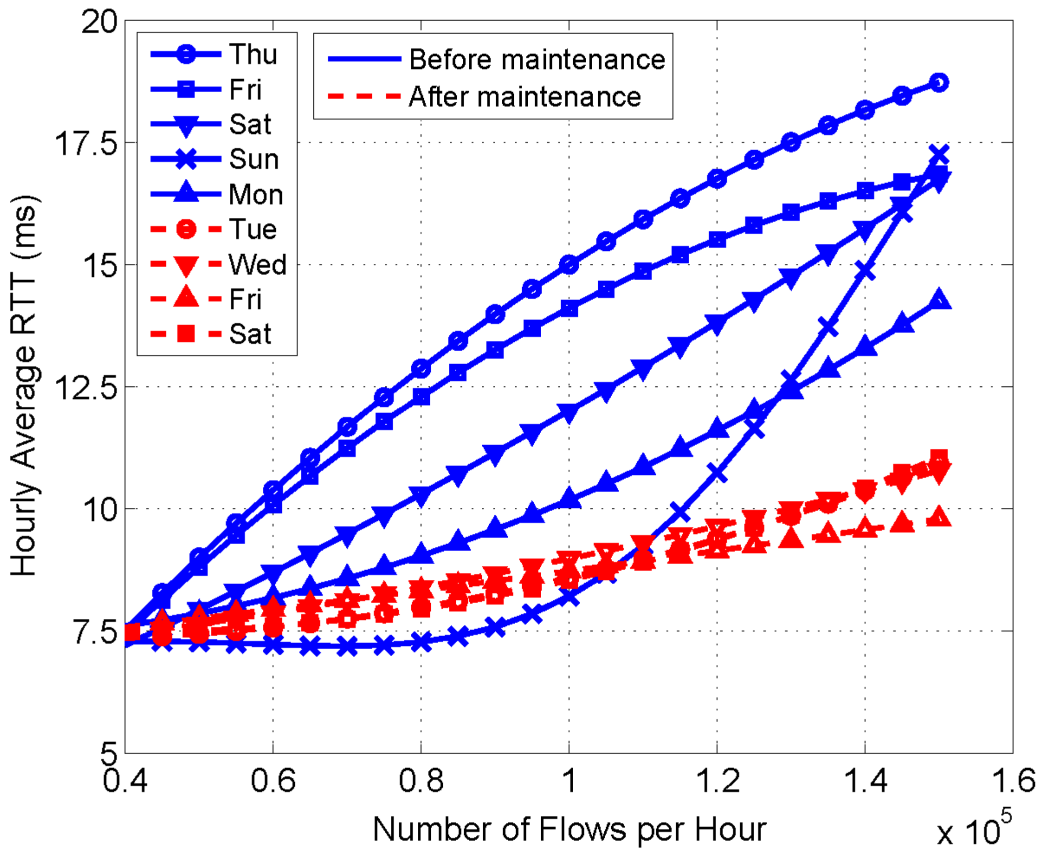
Shift in the Akamai served traffic

- Iterative analysis performed by the reasoner
  - Following a tree-like structure



Impact on performance: historical analyis

- Analysis a week before/after the maintenance reveals:
  - Shift of 50th percentile on all the days before the maintenance
  - No shift in the days following the maintenance intervention
  - Preferred cache shifts are still present → **difficult to engineer for the ISP**

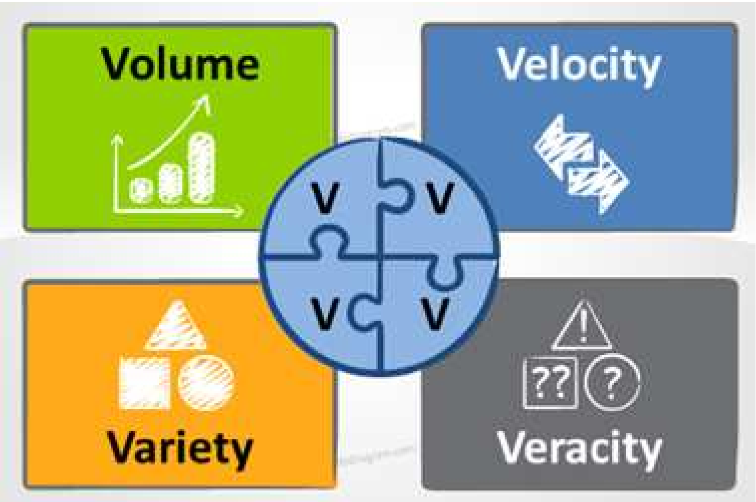


Big Monitoring Data

How to process and anlayze it?

Big data in Network Traffic Monitornig

- Network traffic monitoring generates LOT'S of data!
- e.g., at the local mobile operator
  - DBStream running online since more than one year
  - 160 queries online, 40 input streams
  - **2.5 TB per day**, 77 TB disk space, **38 TB used at the moment**
- The 4 Vs of Big data (or 5 Vs, considering the potential **V**alue)
- All of them are highly relevant for TMA
  - Some applications require results NOW!
  - Some others need to go through large amo
    extract useful knowhow
- Which kind of system should I use?



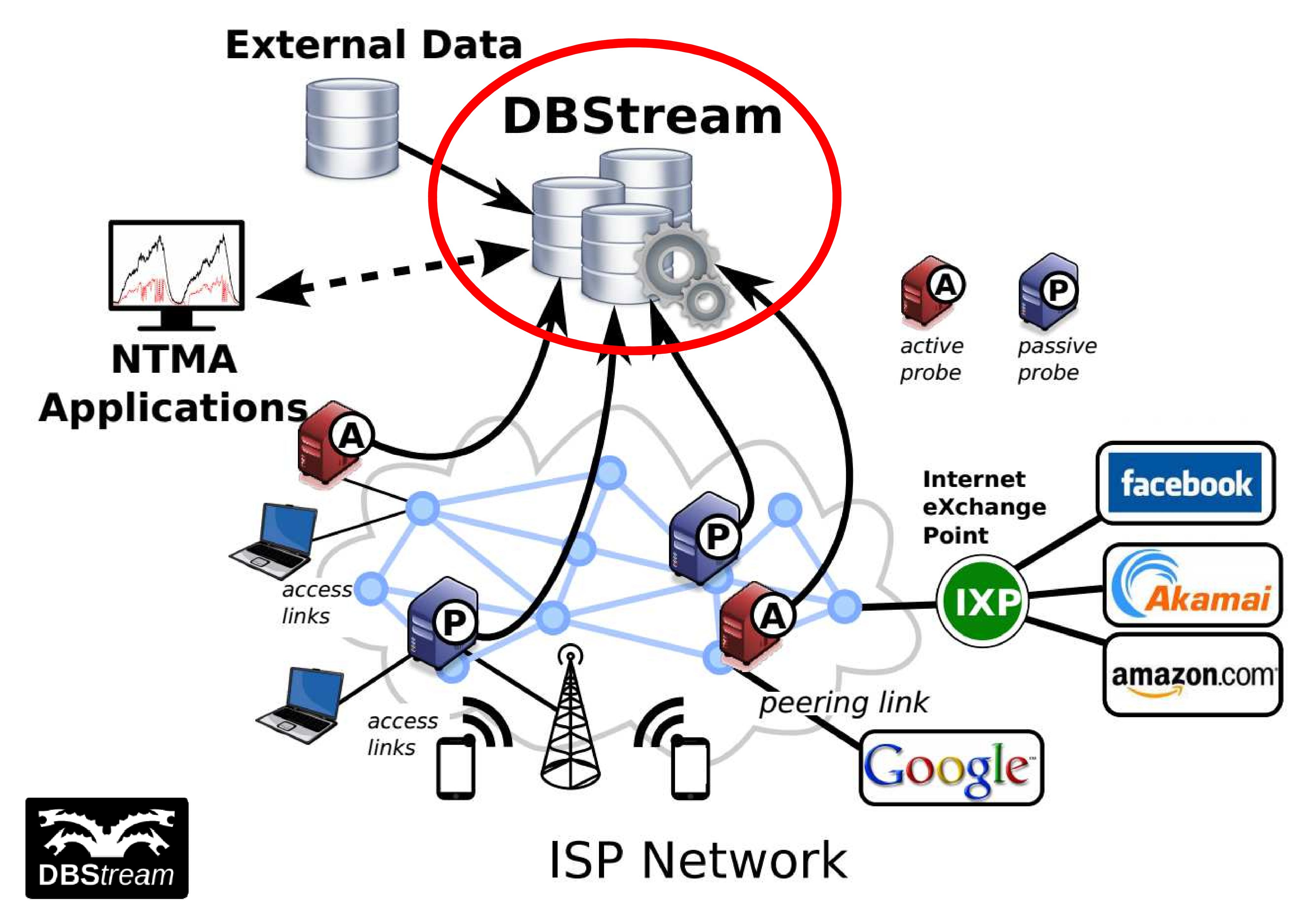
Big data in Network Traffic Monitornig

- Ne
- e.
- Th
- Al
- W

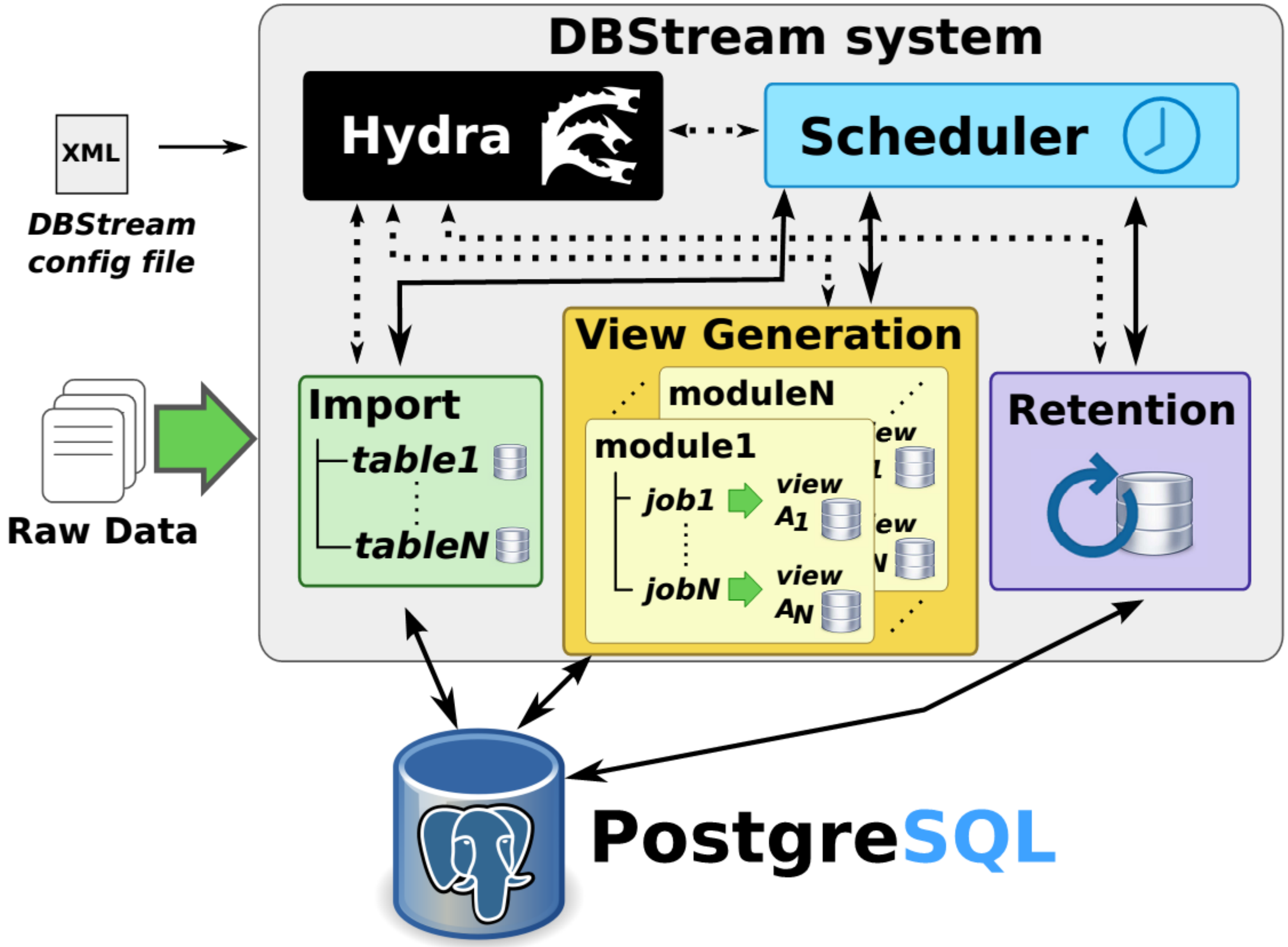
DBStream

an Online Aggregation, Filtering and Processing System for Big Network Traffic Monitoring

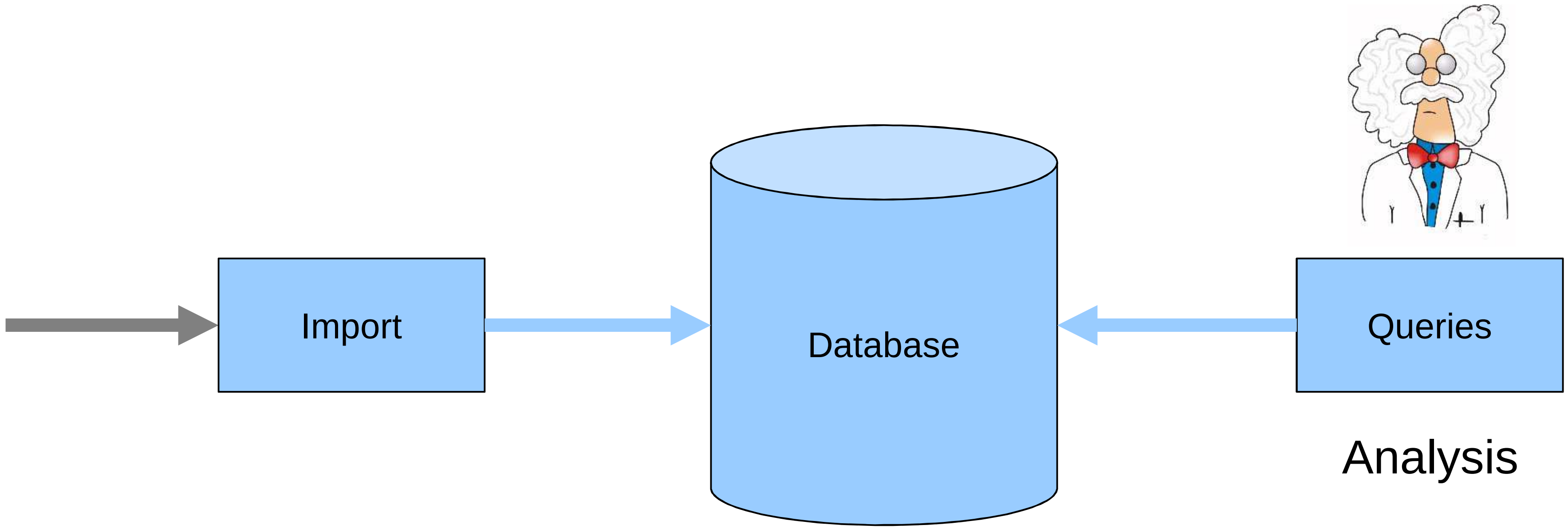
DBStream


External Data
DBStream
NTMA
Applications
active
probe
passive
probe
Internet
eXchange
Point
IXP
facebook
Akamai
amazon.com
Google
access
links
access
links
peering link
ISP Network
DBStream

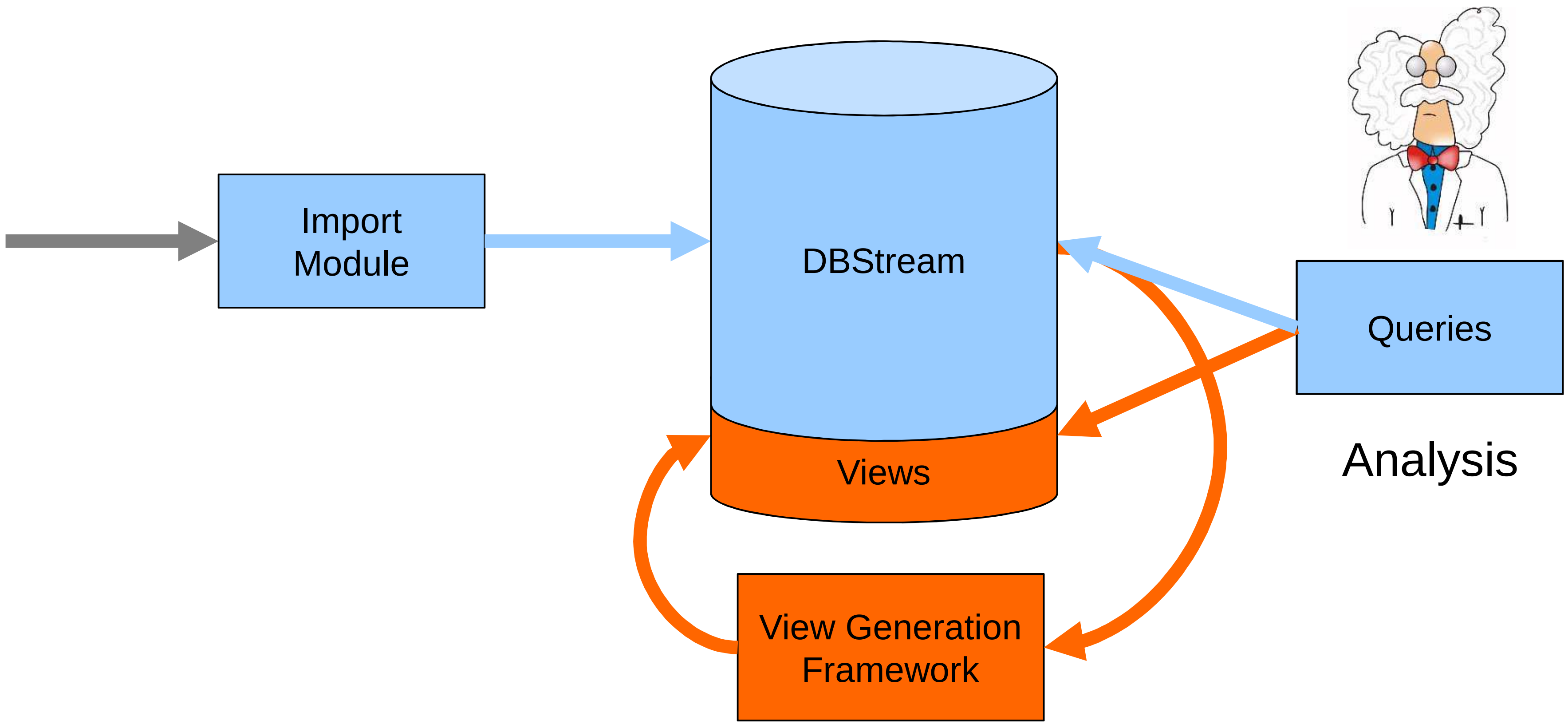
DBStream Middleware Overview



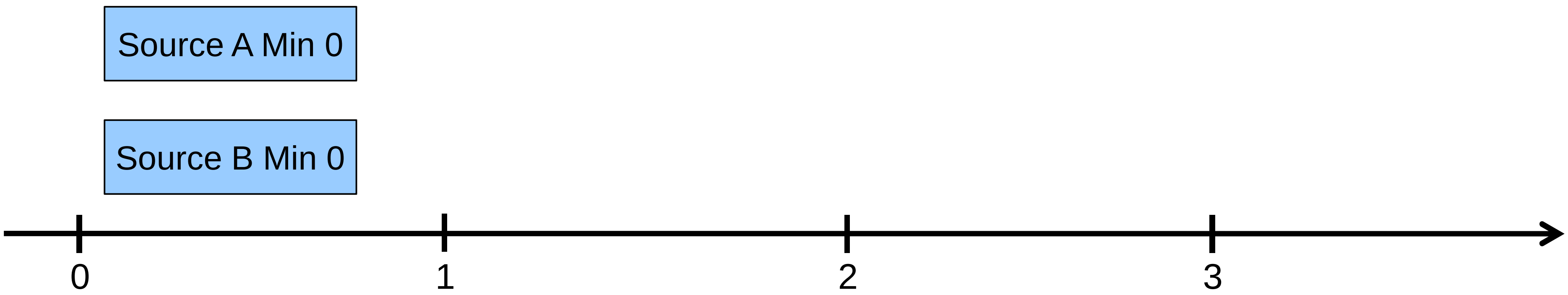
General Database Approach



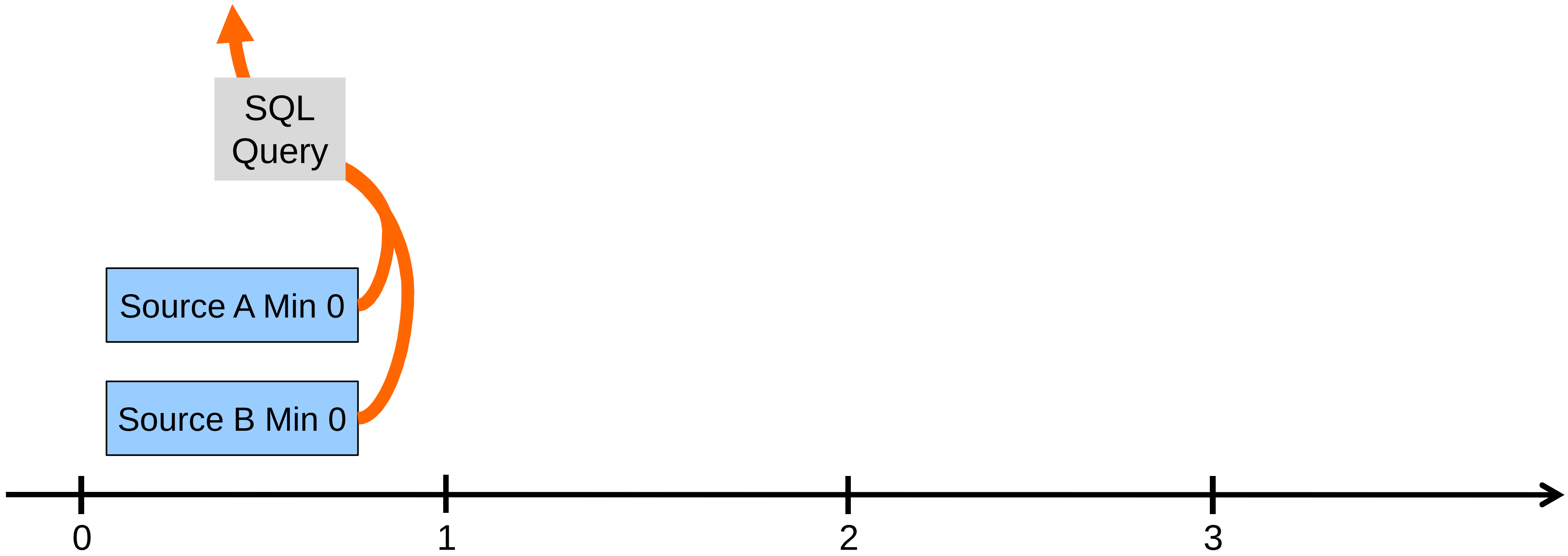
Our Approach: DBStream
Short-time scale batch processing



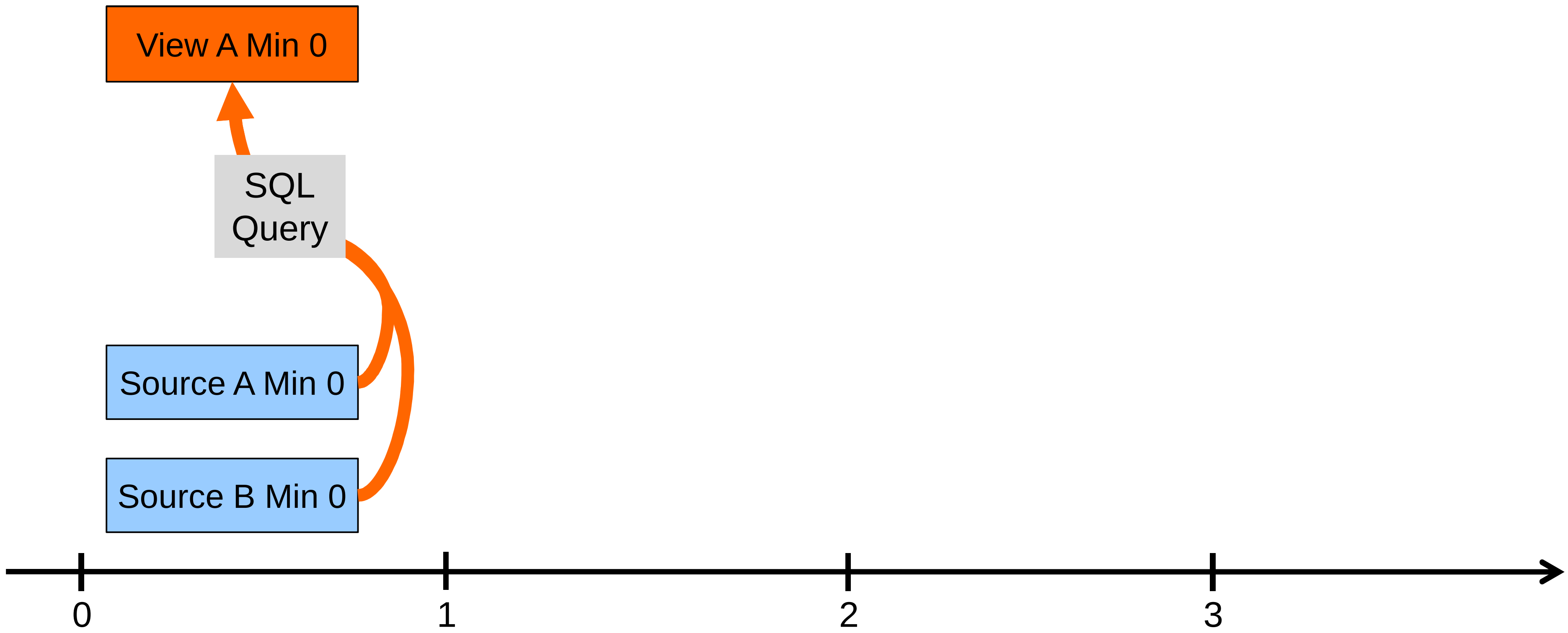
DBStream – View Generation



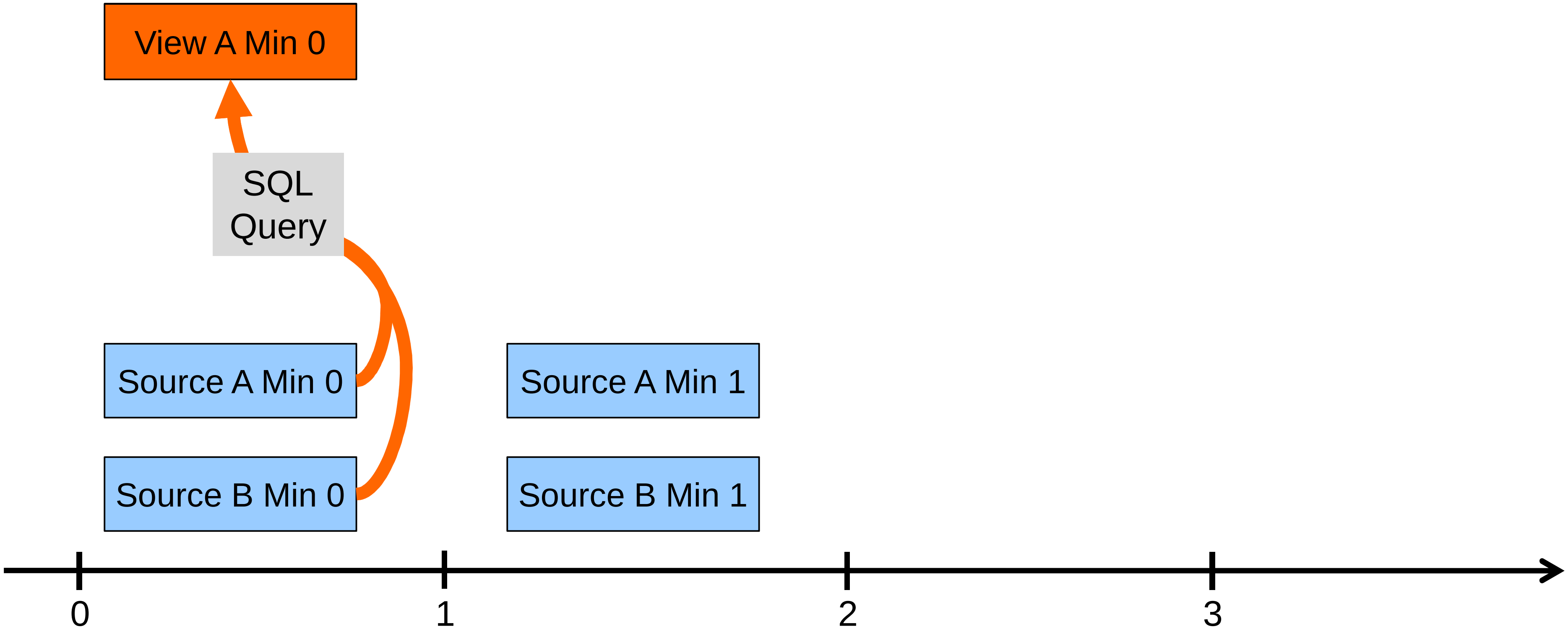
DBStream – View Generation



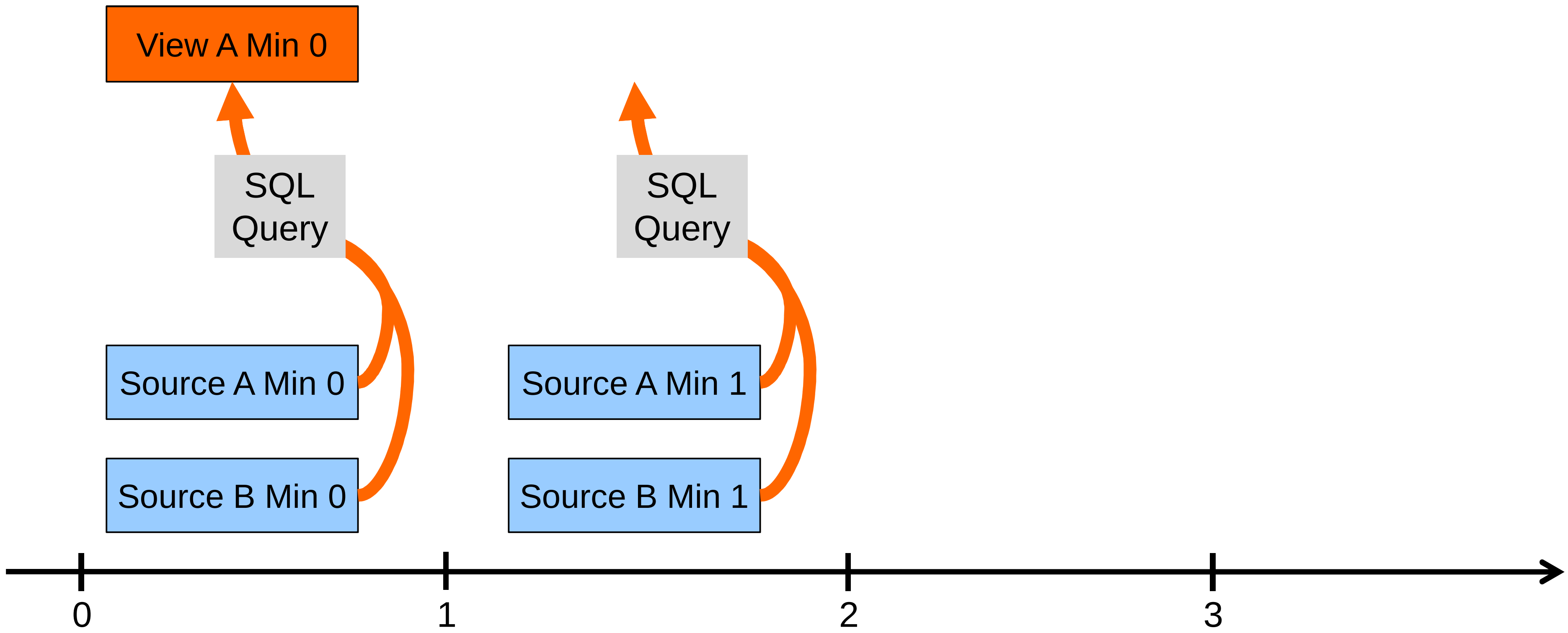
DBStream – View Generation



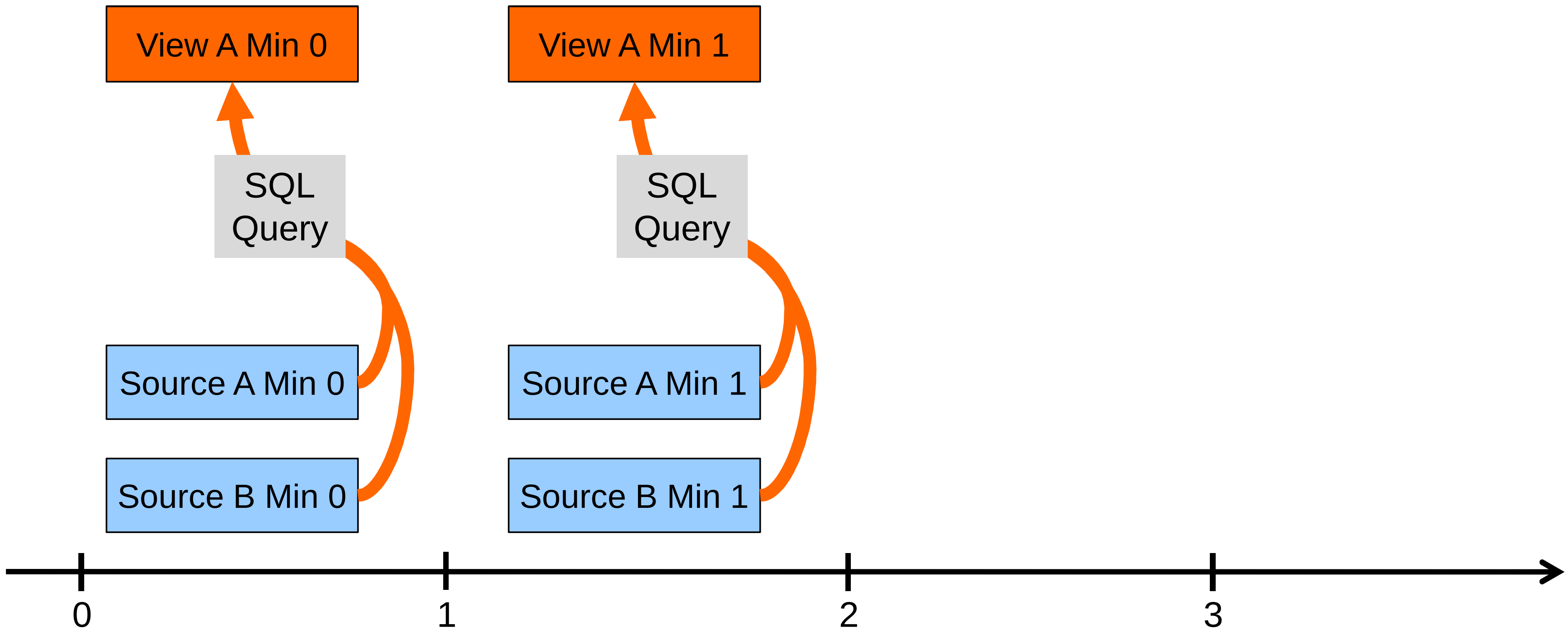
DBStream – View Generation



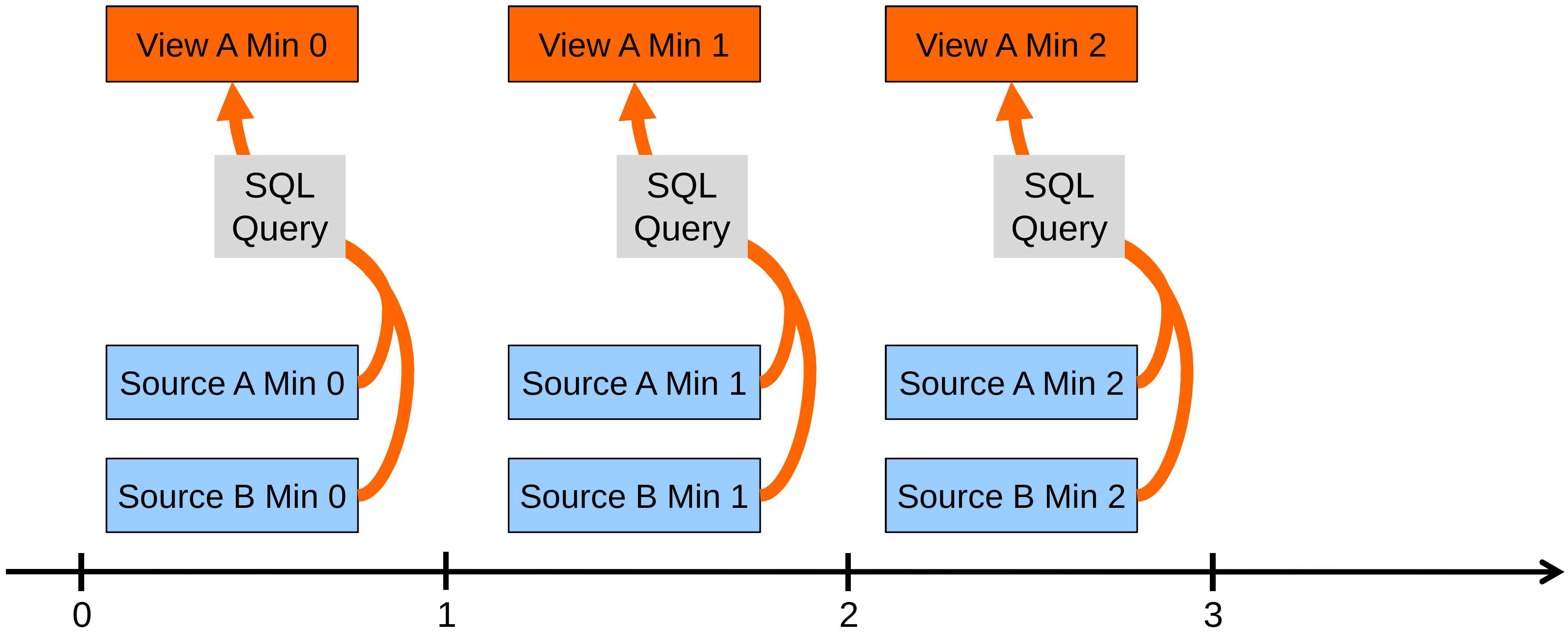
DBStream – View Generation



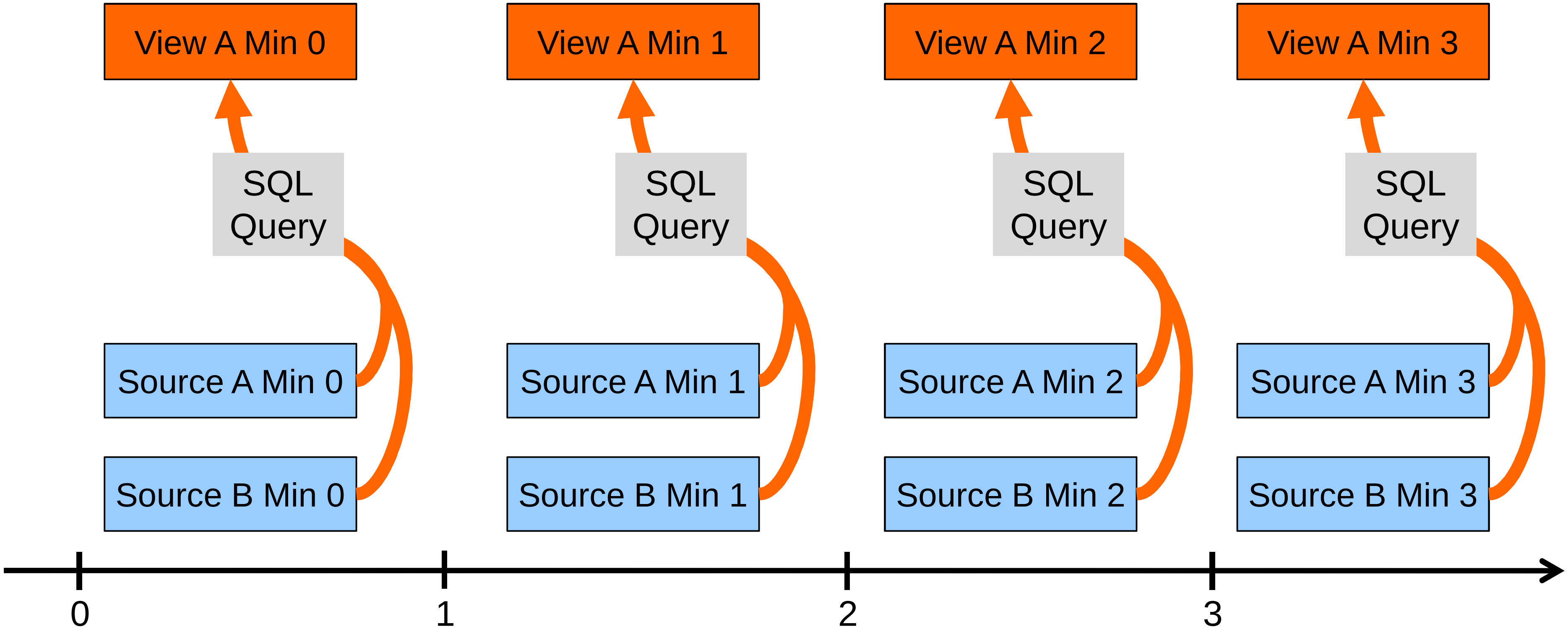
DBStream – View Generation



DBStream – View Generation



DBStream – View Generation



DBStream Query Language (1/5)

→Continuous query processing

- Flexible
- SQL based

```
<job inputs="A (window 15min primary)"
     output="B (window 15min)"
     schema="serial_time int4,
             device_class int4,
             count int4"
     query="select serial_time, device_class,
            count(*) from A
            group by serial_time, device_class"/>
```

DBStream Query Language (2/5)

→Multiple inputs

- Window definition per input
- Multiple inputs possible

```
<job inputs="A (window 15min primary)"
     output="B (window 15min)"
     schema="serial_time int4,
             device_class int4,
             count int4"
     query="select serial_time, device_class,
            count(*) from A
            group by serial_time, device_class"/>
```

DBStream Query Language (3/5)

→Single output

- Table name for storing results
- Window defines partition size

```
<job inputs="A (window 15min primary)"
     output="B (window 15min)"
     schema="serial_time int4,
             device_class int4,
             count int4"
     query="select serial_time, device_class,
            count(*) from A
            group by serial_time, device_class"/>
```

DBStream Query Language (4/5)

→Data format definition

- First column is time
- Other columns can be any PostgreSQL type

```
<job inputs="A (window 15min primary)"
     output="B (window 15min)"
     schema="serial_time int4,
             device_class int4,
             count int4"
     query="select serial_time, device_class,
            count(*) from A
            group by serial_time, device_class"/>
```

DBStream Query Language (5/5)

→Processing query

- Defines how data is aggregated
- Example: number of packets per device class

```
<job inputs="A (window 15min primary)"
     output="B (window 15min)"
     schema="serial_time int4,
             device_class int4,
             count int4"
     query="select serial_time, device_class,
            count(*) from A
            group by serial_time, device_class"/>
```

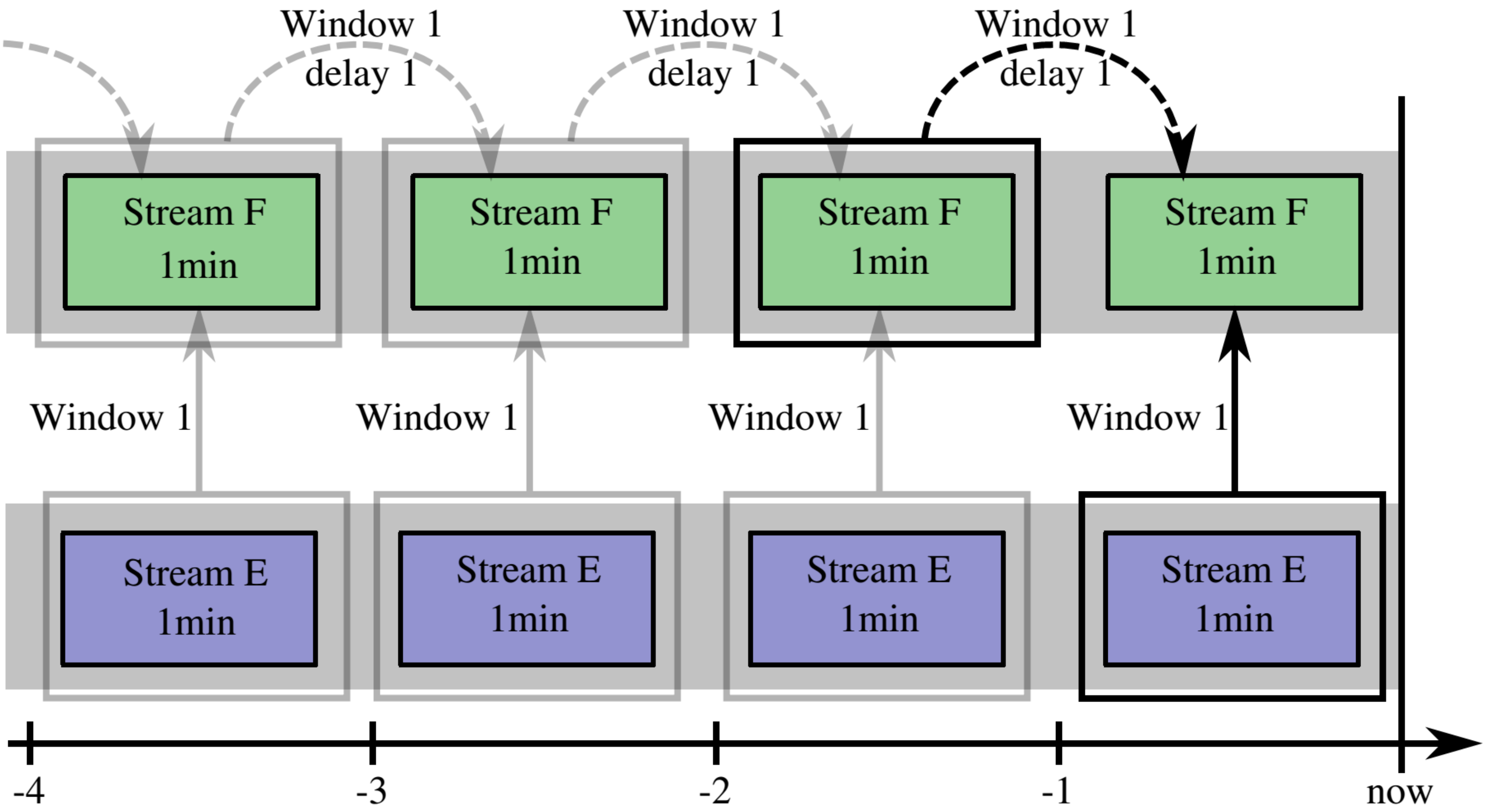
Complex Incremental Query

→Rolling Set Query

- IPs active in the last hour, updated every minute
- Past output is used as input for the next batch

```
<job inputs="E (window 1min primary),
             F (window 1min delay 1min)"
     output="F"
 schema="serial_time int4, last int4, ip inet">
<query>
select _STARTTS, max(last), ip
from (
 select _STARTTS as last, ip
  from E group by 1,2 union all
 select last, ip
  from F where last <= _STARTTS-60 group by 1,2
) t group by 1,3
</query></job>
```

Incremental Query Processing



Experimental Benchmarking – Setup

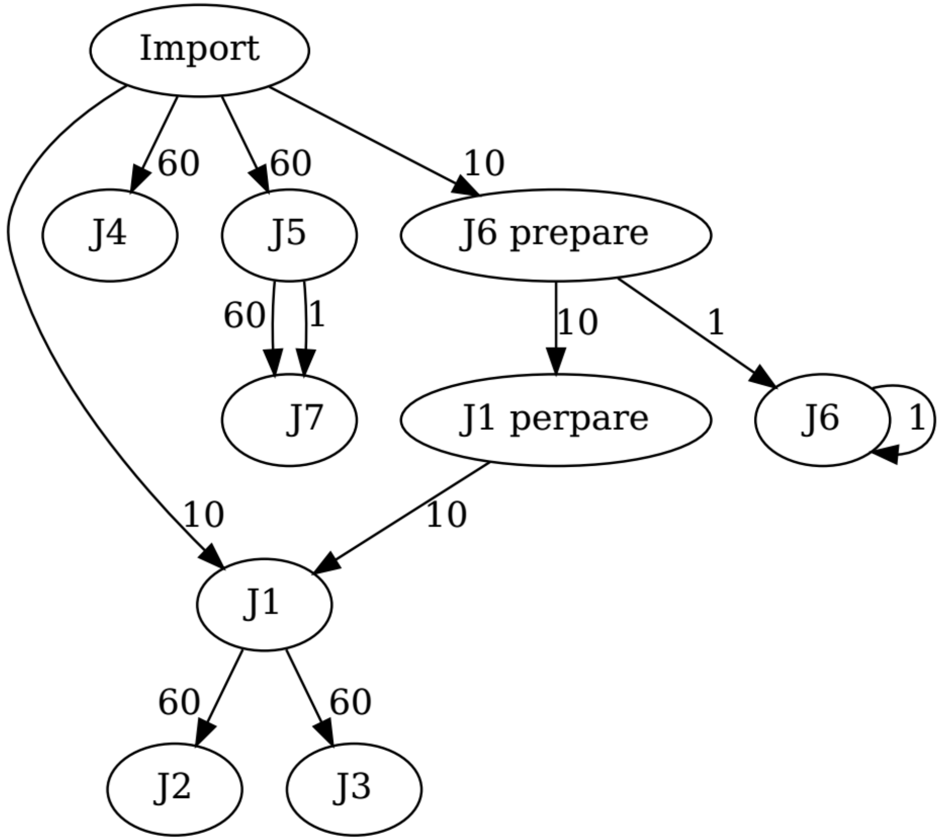
Creating Communication Technologies

- Hardware
  - 10 nodes cluster
  - 6 core XEON E5 2640
  - 32 GB of RAM
  - 5 HD of 3TB each

- Dataset
  - Flow based Tstat data with **about 100 fields**
  - Collected at **4 Vantage Points** (VP), **1 Gbit/s** each
  - Each 162 GB, approx. **650 GB in total**

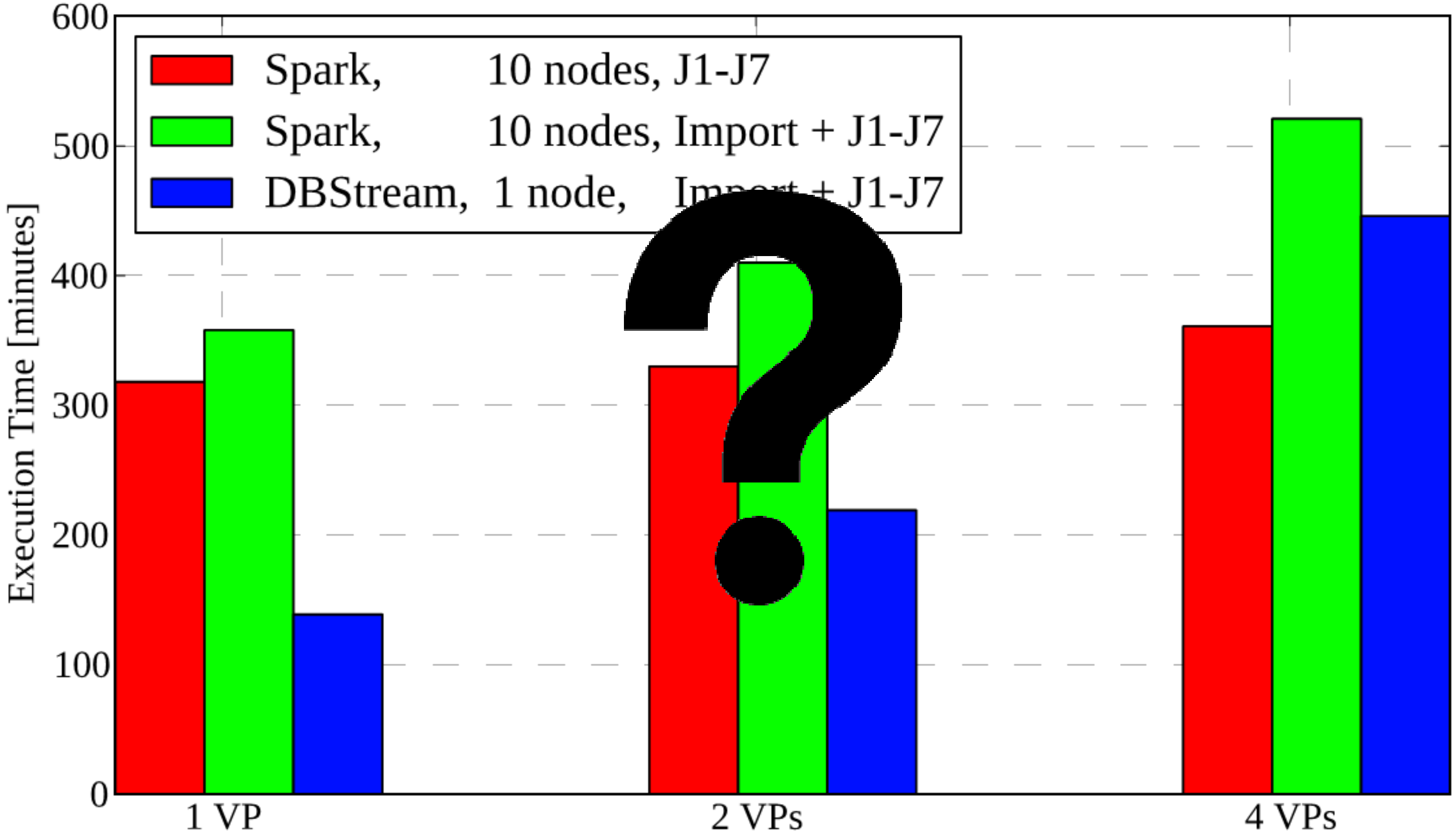
Query Workload – Analysis Jobs

- J1: RTT stats per Orgname
- J2: Akamai stats
- J3: Top 10 Orgname
- J4: Top 10 /24 subnets
- J5: Up/download per source IP
- J6: IPs active in the last hour
  - Updated every minute
- J7: Avg. up/download last hour
  - Updated every minute



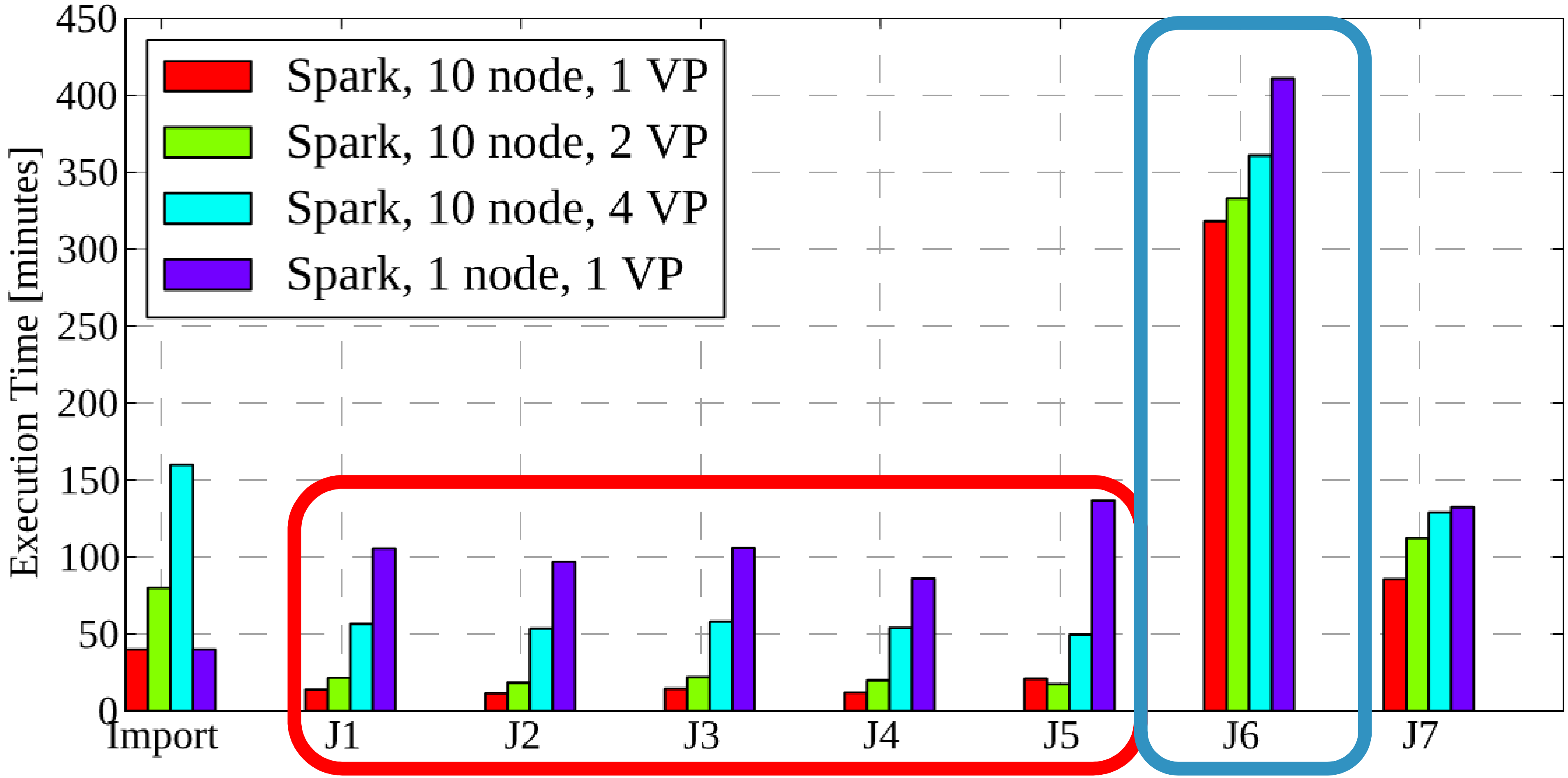
Performance comparison with Spark

ftw Creating Communication Technologies



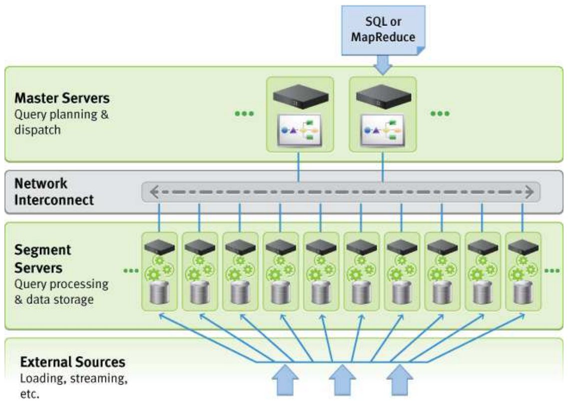
Spark
Lightning-Fast Cluster Computing

Spark Performance Details



Performance Summary

- Performance
  - 1 node DBStream up to 2.6 faster than 10 node Spark for specific analysis jobs
- Result Projections
  - 446 minutes for 4 VP → 12 VP in one day
  - Each VP is 5 days
    - → **DBStream can process a equivalent of 60 VP or 1 VP with 60 GBit/s**
  - HW can be updated, more disks, SSDs?
  - Running on top of parallel databases (e.g., Greenplum)
- Operational DBStream @mobile operator
  - Running online since more than one year
  - 160 queries online, 40 input streams
  - 2.5 TB per day, 77 TB disk space, 38 TB used



Thanks You for Your Attention!

Pedro Casas, casas@ftw.at

ftw Creating Communication Technologies