



Iniziative Europee per lo Sviluppo di un'Infrastruttura Distribuita per Misure su Reti

Il progetto **mPlane** per monitorare lo stato di Internet



Stefano Traverso
stefano.traverso@polito.it





Cosa è Internet oggi

- La più grande **rete telematica** mondiale
 - 2.3 miliardi di utenti residenziali e 4.5 miliardi di utenti mobile nel 2016 [CISCO VNI '13]
- Trasporta **zettabytes** di infomazioni con una vastissima gamma di **servizi**
- E' basata su un paradigma "**open**"
 - Non esiste una singola organizzazione che governa Internet
 - Ognuno può creare un proprio servizio con un proprio modello di business



*“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 – Google Exec. Chairman



Cerchiamo di misurare “The anarchy”?

Dietro l'anarchia ci sono

Gli utenti

- Qualità del servizio
- Network Neutrality & Privacy

anno obiettivi diversi

Providers



Fornitori di contenuto

- Raggiungere gli utenti
- Offrire nuovi servizi

operatori telefonici

- Identificare malfunzionamenti
- Migliorare l'ingegnerizzazione della rete
- Qualità del servizio

Lo scenario è
complicato!

Social Web



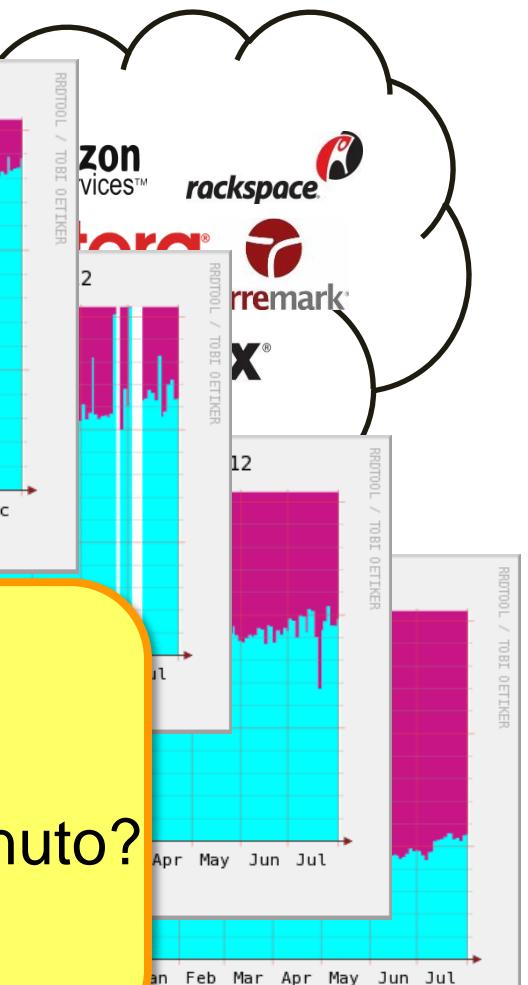
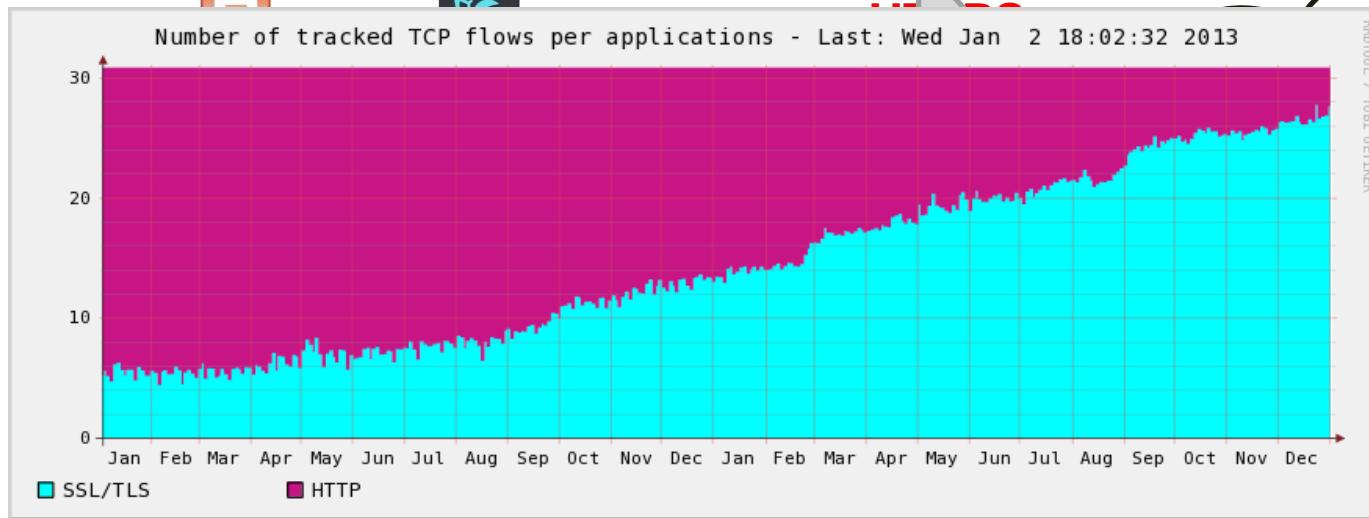
Agenzie di regolamentazione

- Service Level Agreement (SLA)
- Definire e far rispettare politiche di comportamento





Dinamiche sempre più oscure



Stiamo perdendo la **visibilità** del traffico

- Quali servizi gli utenti usano?
- Chi sta servendo un determinato contenuto?
- Come posso bloccare I giochi Zynga?



Capire Internet

Come?

- Misurando e classificando il traffico della rete – Misure **passive**
- Testando le performance della rete – Misure **attive**

Dove?

- **Software/plugin** installati dagli utenti
- **Sonde** nella rete
- Misure su **dispositivi di rete**

Applicazioni

- Procedure automatiche di **Troubleshooting**
- **Controllo del traffico**
- **Anomaly Detection**
- ...



Cosa è stato fatto finora

Progetto	Obiettivo			Metodo			
	Nome	Network Mapping	Performance	Troubleshooting	SW plugin	Sonde nella Rete	Misure su Dispositivi di Rete
•Atlas •Archipelago •Merlin		✓				✓	
•Bismark •Dasu •M-Lab •Netalyzr			✓		✓		
•NetViews •RouteViews •TopHat •ASP		✓					✓
perfSONAR		✓	✓	✓			✓
CCAMP			✓		✓		✓
DIMES		✓			✓		
MOMENT			✓				✓



Iniziative Europee

perfSONAR Performance focused Service Oriented Network monitoring ARchitecture – 2007-ancora attivo
GÉANT



Monitoring and Measurement in the Next generation Technologies – **STREP**, 2007-2013



From global measurements to local management – **STREP**, 2012-ancora attivo





mPlane

- Sigla: **mPlane**
- Nome completo: “mPlane – an Intelligent Measurement Plane for Future Network and Application Management”
- Tipo di progetto: FP7 Integrated Project
- Grant agreement no: 318627
- Inizio: November 1st 2012
- Fondo Totale: 11,2MEuro
- Durata: 3 anni
- Partners: 16





Chi siamo



Prof. Marco Mellia (Polito - IT)

Project Coordinator

Dr Saverio Niccolini (NEC - UK)

Technical Coordinator

Prof. Pietro Michiardi (Eurecom - FR)

Prof. Dario Rossi (ENST - FR)

Prof. Laurent Mathy (Univ.Liege - BE)

Dr. Brian Trammel (ETH - CH)

Prof. Fabio Ricciato (FTW - AU)

Dr. Dina Papagiannaki (TID - ES)

Dr. Giovanna Carofiglio (ALU - FR)

Dr. Dimitri Papadimitriou (ALU – BE)

Dr. Andrea Fregosi (FW -IT)

Dr. Alessandro Capello (TI -IT)

Dr. Rolf Winter (Hochschule Augsburg - DE)

Ing. Luigi Magnanini (SSB progetti - IT)

Ing. Francesco Matera (FUB -IT)

Dr. Arpad Bakay (Netvisor -HU)

...





Le motivazioni di mPlane

- Internet è una rete globale di reti pubbliche e private
 - Non esiste una singola entità che la governi
 - È robusta, ma anche fragile
- In caso di “failure”, chi può dire cosa è andato storto?
 - Ogni ISP (forse) sa cosa succede nella propria rete
 - Ma se il guasto dipende da altri ISP? O dal content provider? O dal device dell’utente?
- Il web è un labirinto
 - Nessuno sa esattamente cosa succede oggi su Internet
 - Come possiamo predire cosa succederà in futuro?
- Serve un sistema che **raccolga misure** sulla rete, le **analizzi** e **migliori la visibilità**: un oracolo che fornisca **risposte**



mPlane – Obiettivi ambiziosi...

Progetto	Obiettivo			Metodo			
	Nome	Network Mapping	Performance	Troubleshooting	SW plugin	Sonde nella Rete	Misure su Dispositivi di Rete
 mPlane							

Come?

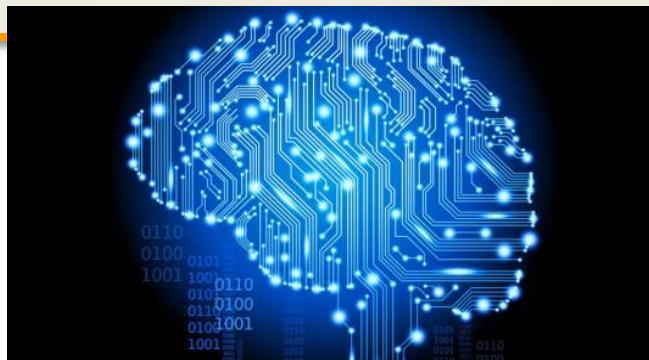




mPlane – Architettura

Repository

- Storage di grandi quantità di dati
- Alta velocità di trasferimento
- Analisi/mining sui dati
- Compatibilità con sistemi di calcolo parallelo (BigData Analysis)



e)
nue

Supervisor/Reasoner

- Coordina sonde e repository
- Analizza i dati e li sintetizza
- Individua le anomalie
- Comanda ulteriori analisi (Root cause analysis)



mPlane – I players e le loro interazioni

Internet Service Provider

Sonda

Supervisor

Repository

Supervisor

Repository

Perchè Skype non
Funziona?

Sonda

App Provider

skype™





L'esperienza di TNG



- **TNG** ha una lunga esperienza su misure **passive** e caratterizzazione del traffico
- Stiamo lavorando a
 - Sonde passive a 40Gb/s
 - Grossi repository ($\pm 50\text{TB}$)
 - Utilizzo di sistemi BigData per l'analisi (Hadoop MapReduce)
- Casa di **Tstat**



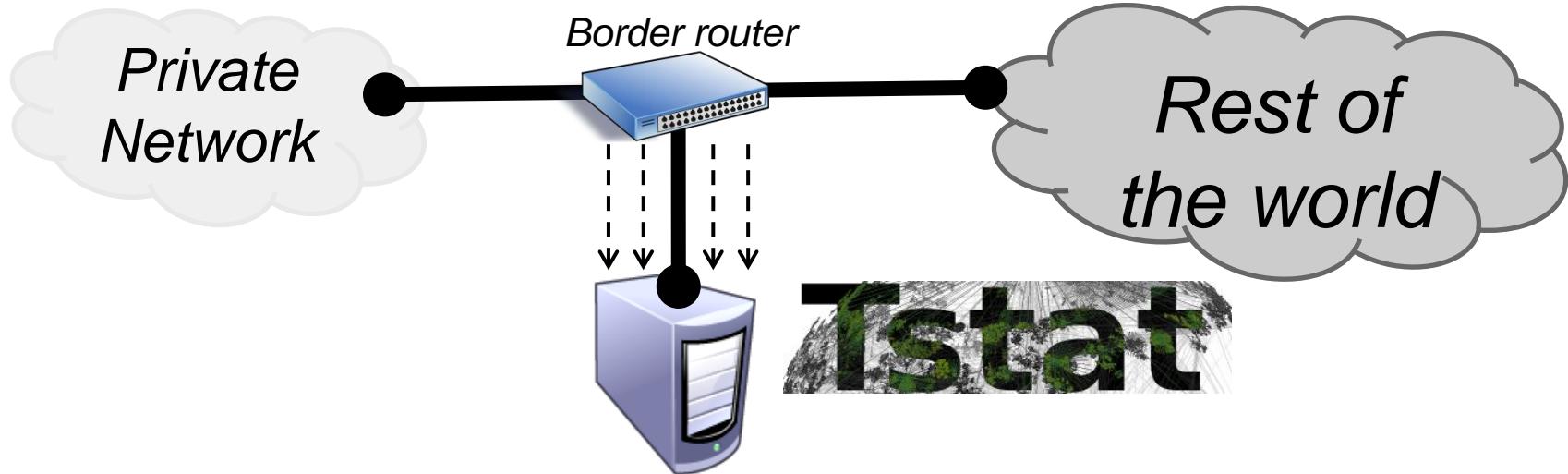
Esempio di Sonda ad alta velocità: Tstat

- Tstat è un tool **Open Source** per il monitoring passivo della rete
 - Nato come branch di **tcptrace** e specializzato per TCP
 - Si è evoluto per analizzare traffico da L3 a L7
 - 10 anni di sviluppo (e bachi ☺)
- Funziona su Linux/FreeBSD/NetBSD/MacOS
 - svn co <http://tstat.polito.it/svn/software/tstat/trunk/>
- *libtstat* per l'integrazione con altri tool di monitoring





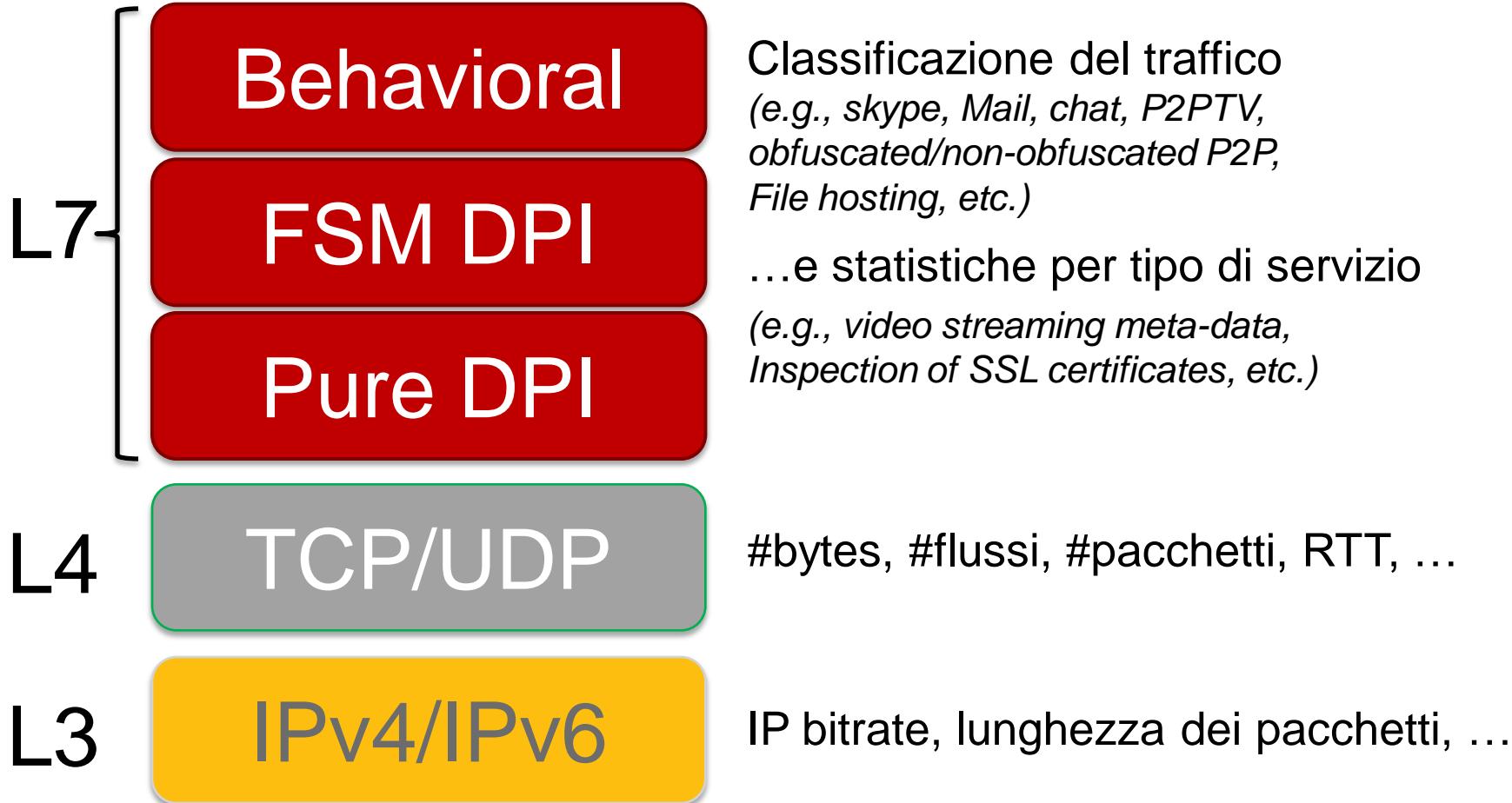
Esempio di Sonda ad alta velocità: Tstat



- **Tstat ricostruisce i flussi TCP/UDP**
 - Cattura il traffico sulla scheda di rete e lo processa in **real-time**
 - Da dump files (pcap, DAGERF, etc.)
- **Classificatore di traffico**
 - Deep packet inspection
 - Algoritmi statistici
 - Metodologie di classificazione
- **Output scalabile**
 - Round Robin Database (RRD)
 - Histograms
 - Logs
- **Funziona su HW off-the-shelf**
 - Qualsiasi NIC abilitata alla cattura
 - Fino a 4Gb/s
 - Prossima meta: 40Gb/s
- **Oproduce grandi quantità di dati**
 - Da post-processare



Tstat workflow





Formati output di Tstat

light

RRD – Round Robin Database

- Un formato compatto per rappresentare il traffico su diverse scale temporali
- Sul sito si può provare una CGI <http://tstat.polito.it/web.shtml>

Iistogrammi

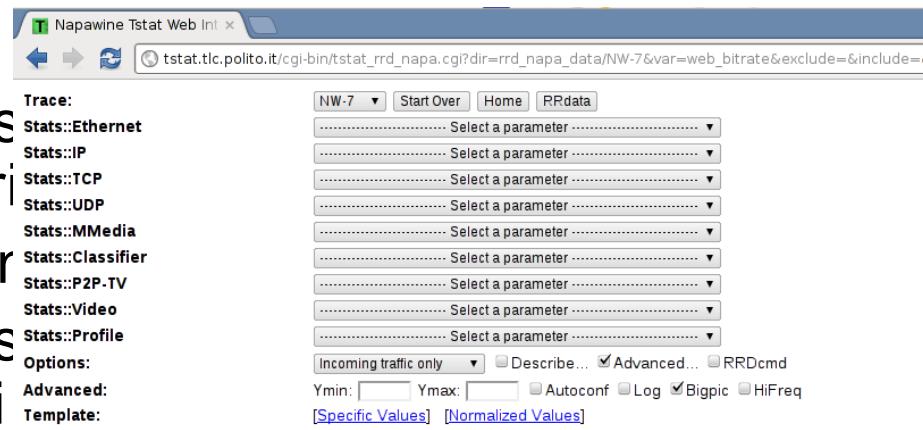
- File di testo parametri

Logs per-corso

- File di testo osservati

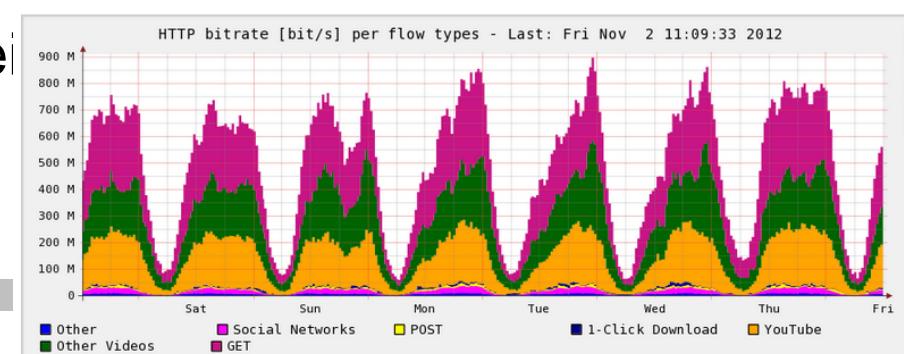
Tracce a live

- Dump dei



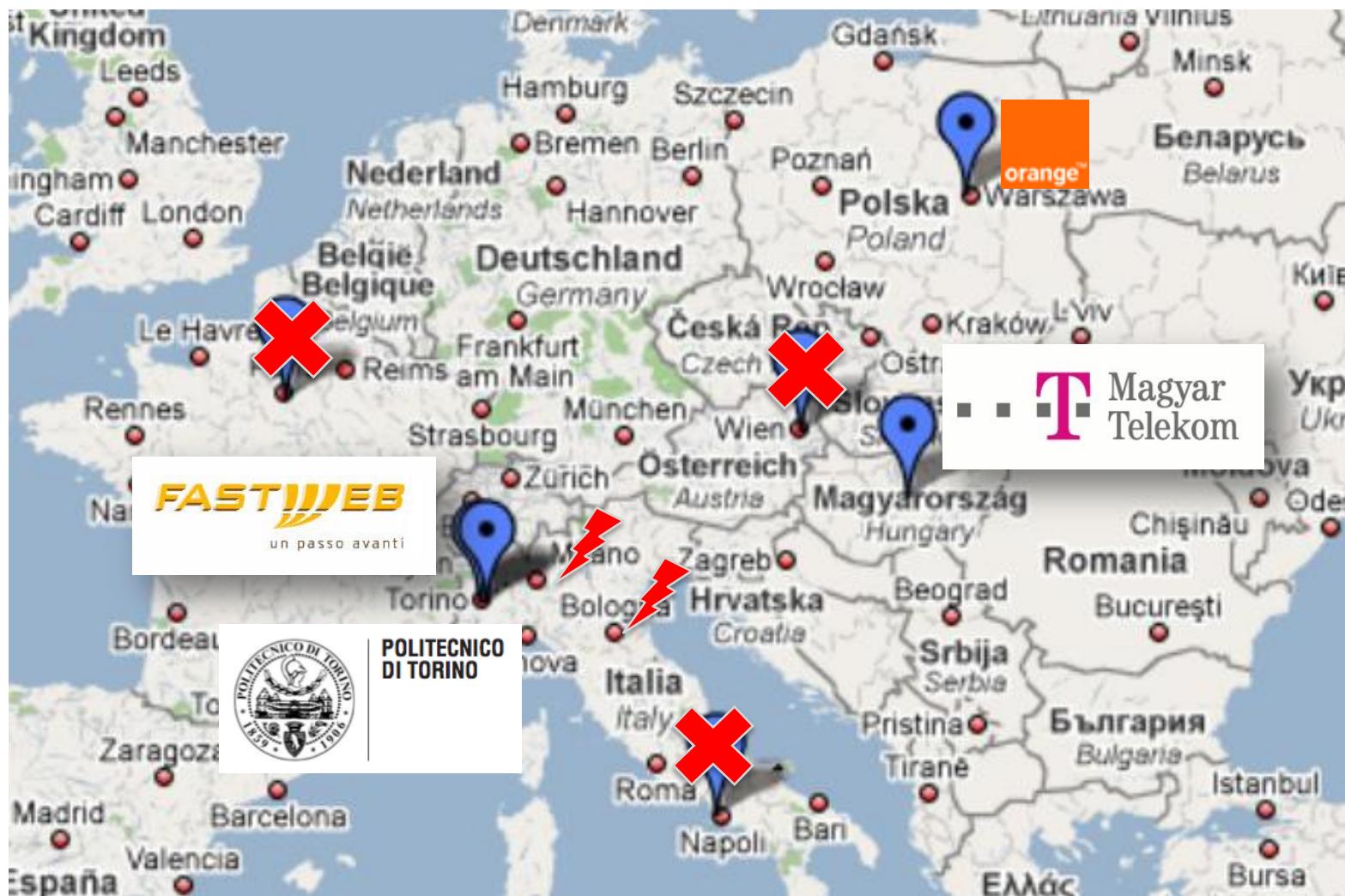
determinati

nti per i flussi





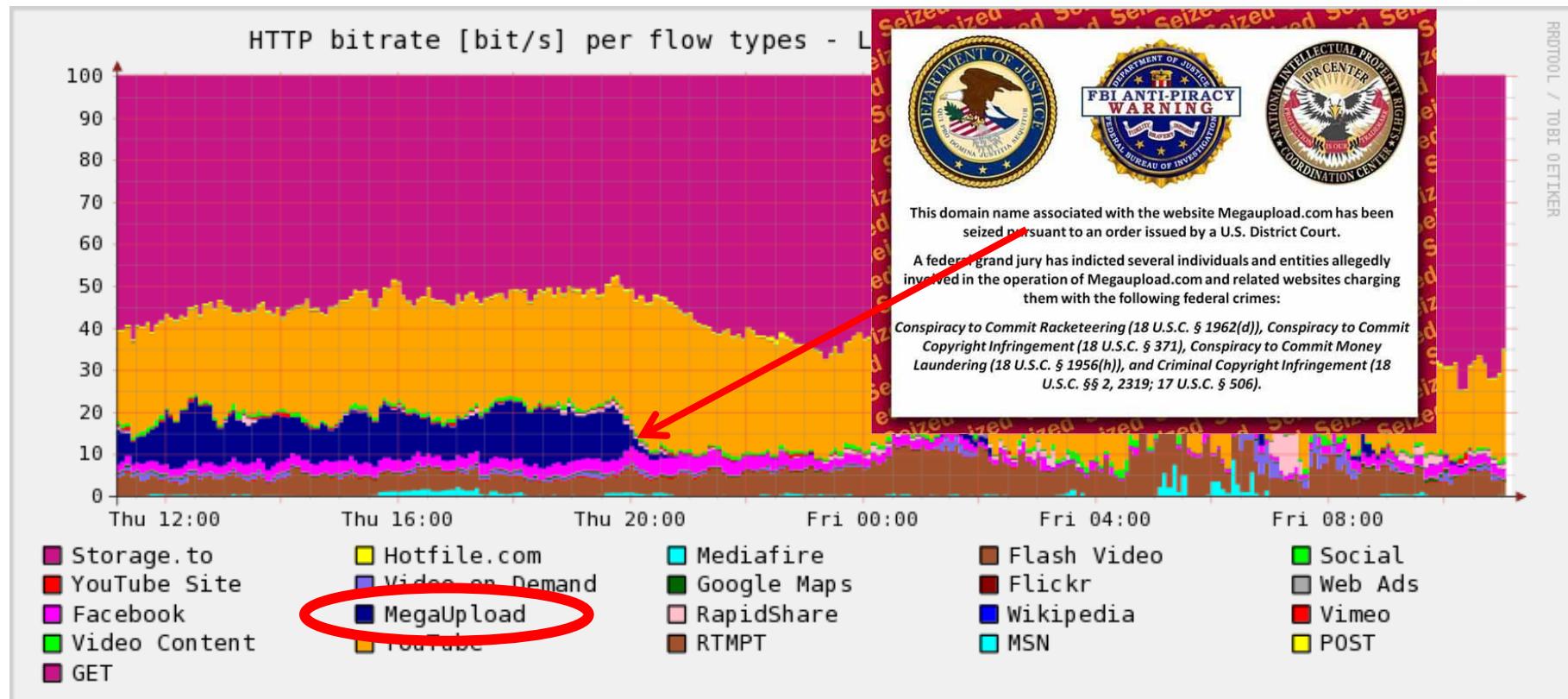
Dove vive Tstat





Megaupload chiuso il 19/01/12

MEGAUPLOAD

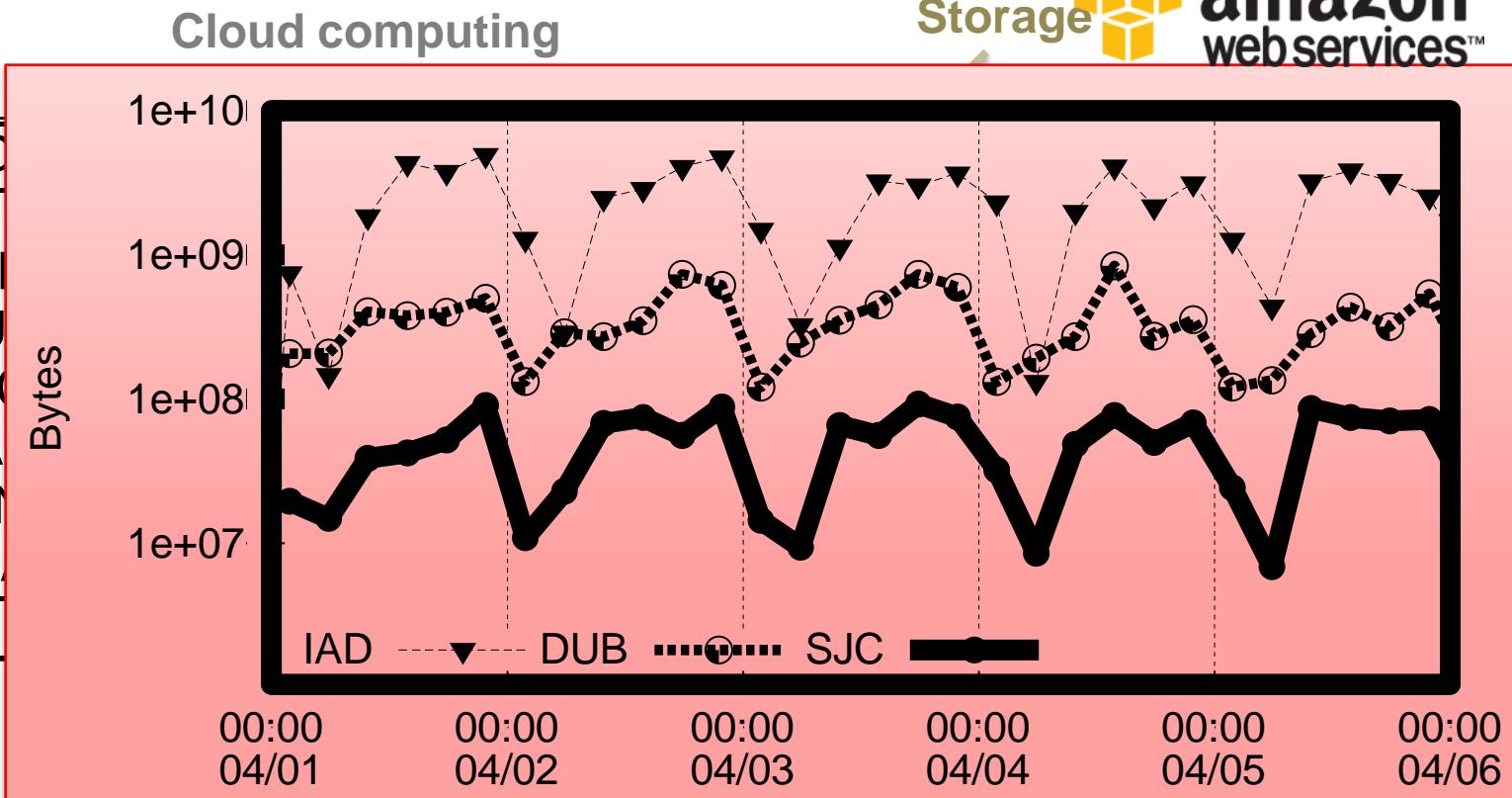




Quali sono i DataCenter Amazon più utilizzati dagli utenti Italiani?

Codici degli aeroporti

ID	IAD
IAI	DUB
DUB	SJC
SJC	NA
NA	SIN
SIN	SEA
SEA	



1. Il DC di
2. Solo il **13%~35%** dei dati sono inviati al dc in **Europa**
3. La **latenza** (RTT) è **critica** per applicazioni come i **giochi multiplayer real time**



Link utili

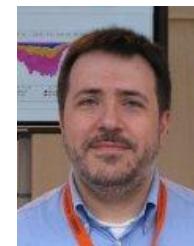
- <http://www.ict-mplane.eu/>
- <http://tstat.tlc.polito.it/index.shtml>



Contatti



Marco Mellia
mellia@tlc.polito.it



Maurizio Munafò
munafo@tlc.polito.it

Alessandro Finamore
alessandro.finamore@polito.it



Ignacio Bermudez
ignacio.bermudez@polito.it



Stefano Traverso
stefano.traverso@polito.it





Perguntas Fragen Domande Galdera Frågor Questions Otázky Spørgsmål Pertanyaan kysymykset вопросы Preguntes Sorular Plane Pytania



DN-Hunter

- Correlating flows IPs with DNS queries will provide a natural way of mapping content and traffic
 - Registered names usually carry some semantic
 - Many web/client-server applications use DNS to get the IP address of the target host
- For simplicity, it is implemented with
 - single buffer to store FQDN (no need to handle TTL)
 - access based on client and server IP

