Scalable flow-level measurements and traffic analysis

Fredi Raspall
David Rincón

UPC UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

i2cat FUNDACIÓ

GORE 8 - Barcelona, October 6th 2011

Contents

- What do we do?
  - Our research interests
- How do we do it?
  - Techniques, application
- Why are we here telling you?

Our research interests

Our research interests

- Traffic monitoring at several levels
  - Flow, link, traffic matrix
  - Analysis & prediction
- Applications
  - Planning / dimensioning / provisioning
  - Anomaly detection
  - Optimization
    - Load balancing, power consumption, protection...

Traffic monitoring

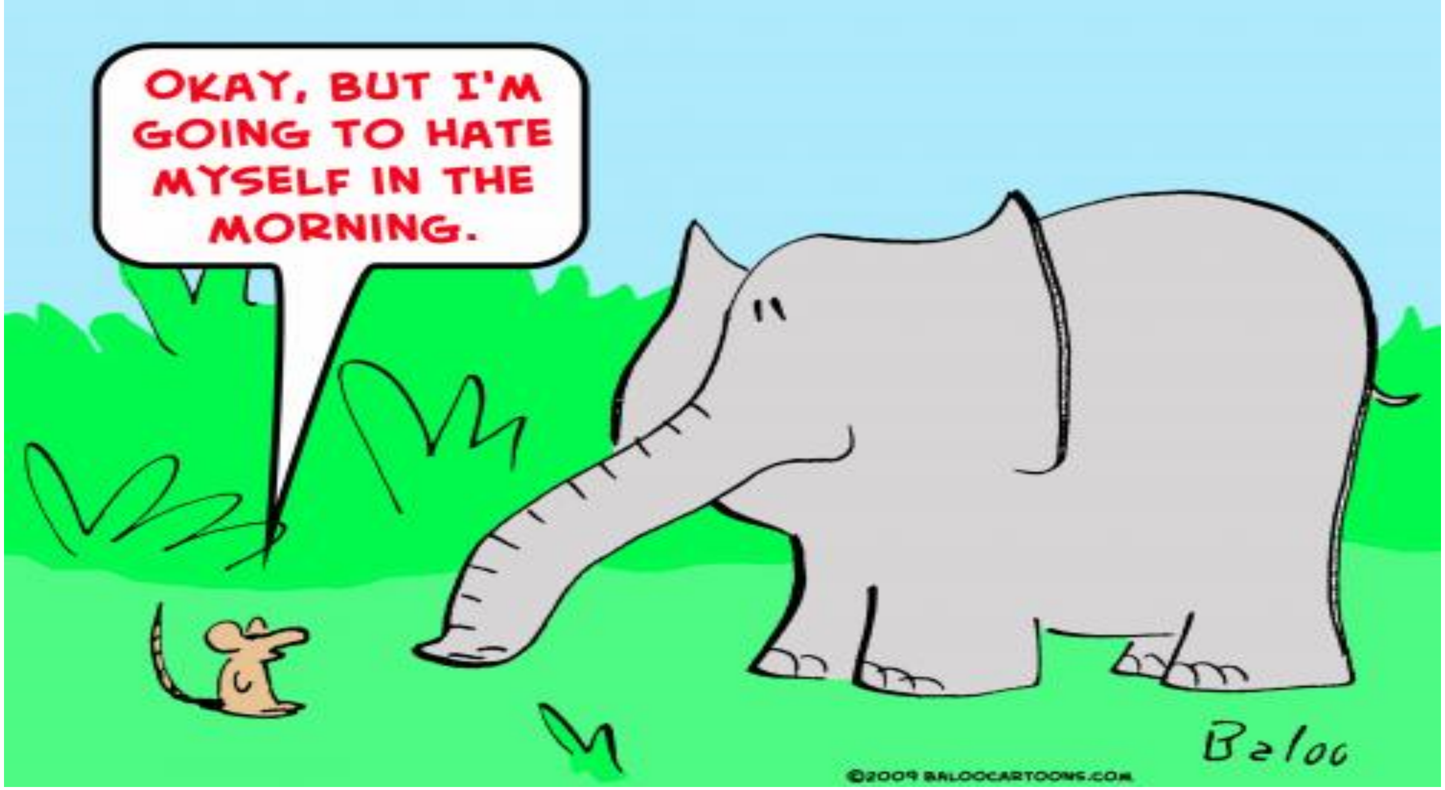
- Netflow records

```
Date flow start          Duration Proto      Src IP Addr:Port          Dst IP Addr:Port    Packets    Bytes Flows
2010-01-17 09:15:48.372     0.000 TCP      222.35.136.xxx:45034 ->   193.144.51.xxx:22             1       68     1
2010-01-17 09:15:15.094     0.000 TCP      222.35.136.xxx:34767 ->   193.144.51.xxx:22             1       56     1
2010-01-17 09:15:19.949     0.000 TCP      222.35.136.xxx:47396 ->   193.144.51.xxx:22             1       72     1
2010-01-17 09:15:24.389     0.000 TCP      222.35.136.xxx:58955 ->   193.144.50.xxx:22             1       52     1
2010-01-17 09:15:34.299     0.000 TCP      222.35.136.xxx:56388 ->   193.144.59.xxx:22             1       52     1
2010-01-17 09:15:15.293     0.000 TCP      222.35.136.xxx:41868 ->   193.144.59.xxx:22             1      104     1
2010-01-17 09:15:30.299     0.000 TCP      222.35.136.xxx:39707 ->   193.144.59.xxx:22             1      104     1
2010-01-17 09:15:36.669     0.000 TCP      222.35.136.xxx:38977 ->   193.144.59.xxx:22             1      204     1
2010-01-17 09:15:41.622     0.000 TCP      222.35.136.xxx:43512 ->   193.144.59.xxx:22             1       68     1
2010-01-17 09:15:22.033     0.446 TCP     194.169.201.xxx:80    ->   193.144.79.xxx:26505          8    12000     1
2010-01-17 09:15:38.694     0.000 TCP        81.184.8.xxx:51260 ->   193.146.43.xxx:38787          1       54     1
2010-01-17 09:15:06.122    40.614 TCP        85.56.18.xxx:26215 ->    193.146.38.xx:64613          4     4516     1
2010-01-17 09:15:22.033     0.000 TCP        85.56.18.xxx:26215 ->    193.146.38.xx:64613          1      465     1
2010-01-17 09:15:02.196    55.011 TCP        82.60.18.xxx:57563 ->   193.144.56.xxx:11629         16      652     1
2010-01-17 09:15:01.676     0.000 UDP      188.128.29.xxx:24370 ->    193.146.32.xx:53             1       73     1
2010-01-17 09:15:05.943    51.187 TCP      193.144.56.xxx:11629 ->      85.179.88.x:1951           7    10178     1
2010-01-17 09:15:17.911    39.219 TCP      193.144.56.xxx:11629 ->      85.179.88.x:1951           4     2899     1
```

- Limitations of Netflow
  – Accuracy, router resources (CPU / memory)
  – Sampled Netflow

Per-flow measurements

- Elephant / mice principle

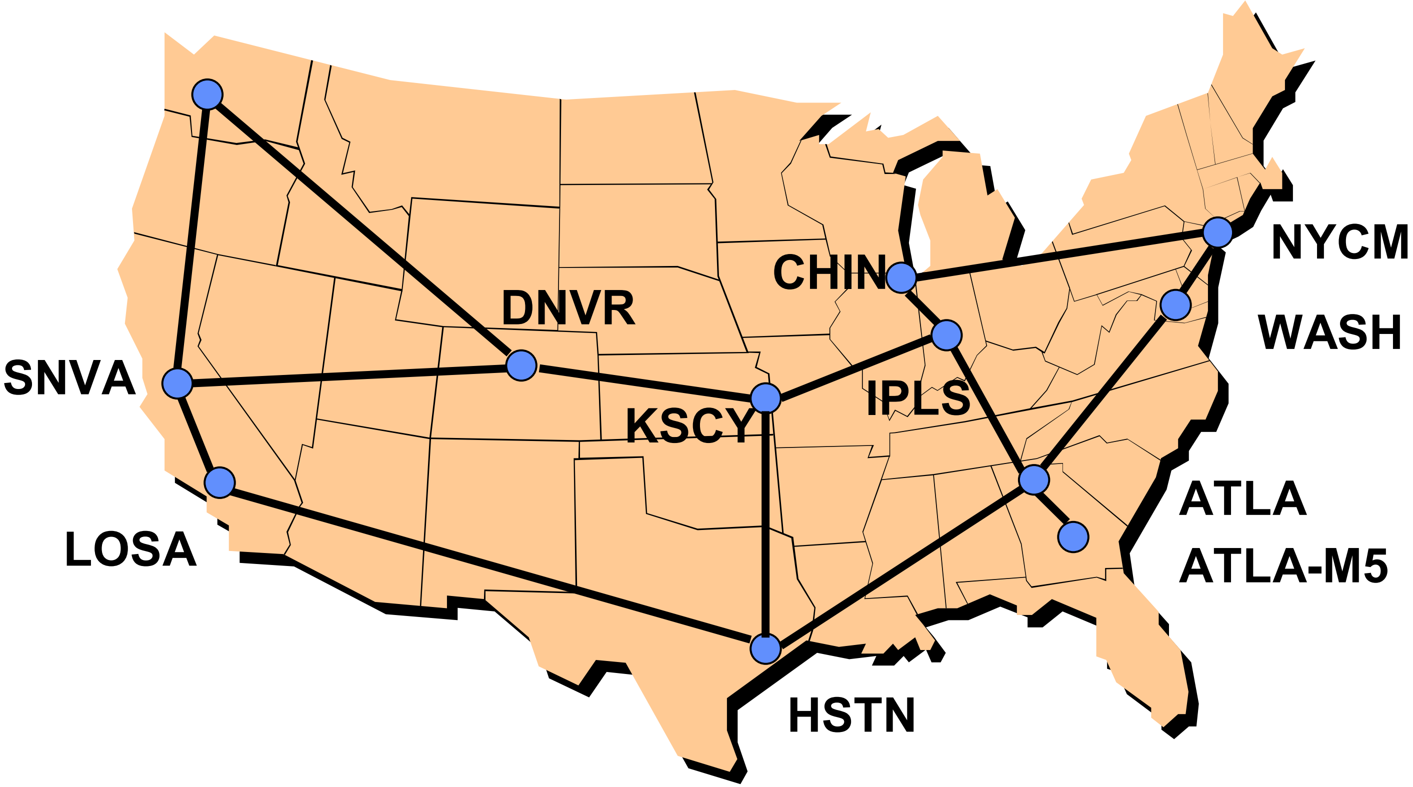


Sampling the largest flows

- Let’s not capture everything – **only the largest flows**
  - Useful for traffic engineering, for example
  - Definition of **heavy-hitter**: those who exceed a given threshold (absolute, a fraction of the link capacity, a fraction of the volume) during some measurement interval, or “so far”.
- Combining ideas: **sampling** & identifying **largest flows**
  - Sampling: lightweight and flexible, but inaccurate
  - Largest flows: memory-efficient but heavyweight, often require memory operations per packet and housekeeping tasks
- Contribution: Sampling algorithms for identification of large flows, that scale **both** in memory and speed.

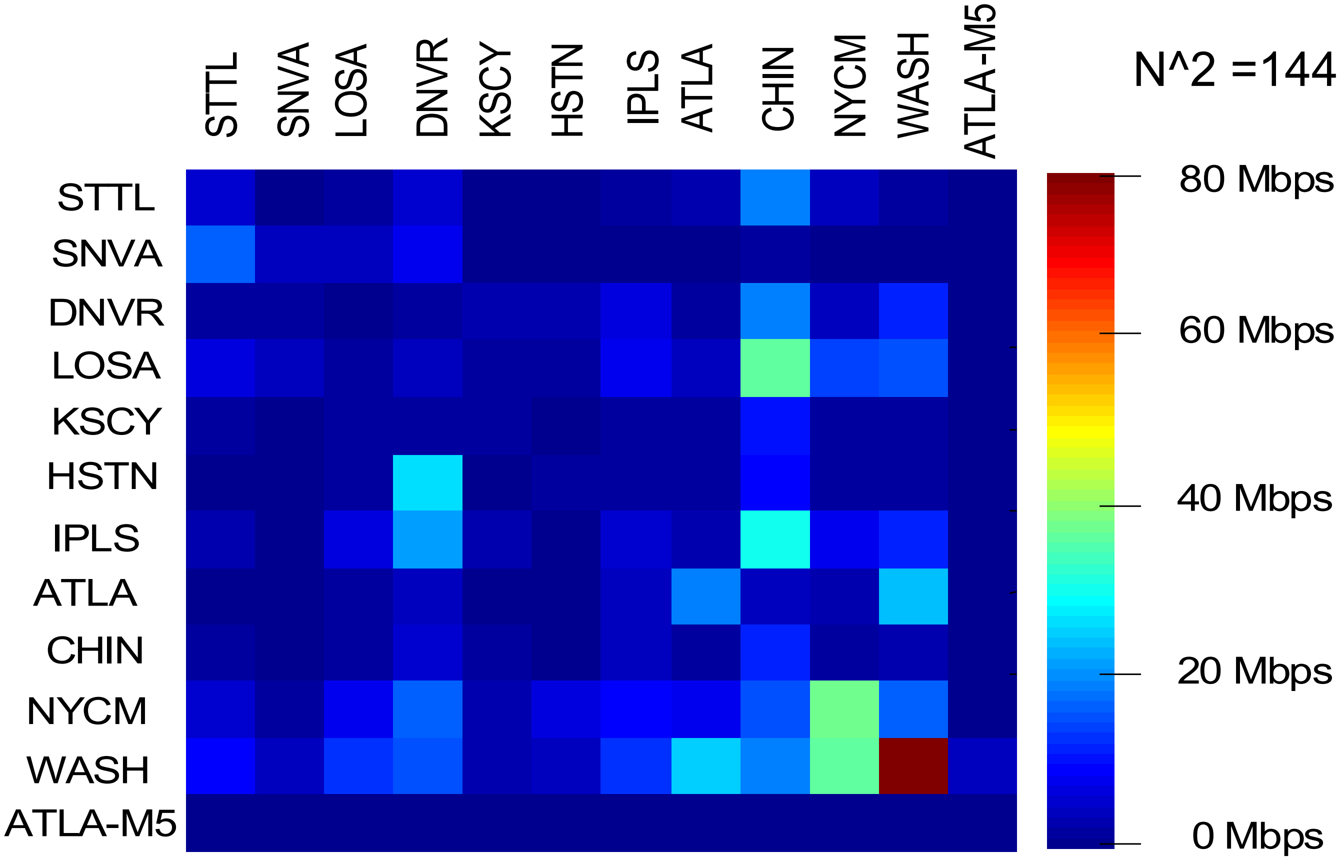
Traffic matrices

N^2 =144 routes



Abilene topology (2004)

Traffic matrices



N^2 =144 entries

Traffic matrix from Abilene (March 2nd 2004, 12:00-12:05)

TMs: measurement & modelling

- Direct measurement with Netflow and others
  - Difficult to scale → sampling (correctly?)
  - Difficult to synchronize measurements

- Indirect inference from SNMP counters
  - Ill-posed problem: N^2 unknowns from O(N) link loads: $y=Ax$
  - **Gravity model / Tomogravity**:
    - Traffic exchanged between two nodes is proportional to the total traffic entering/exiting the nodes

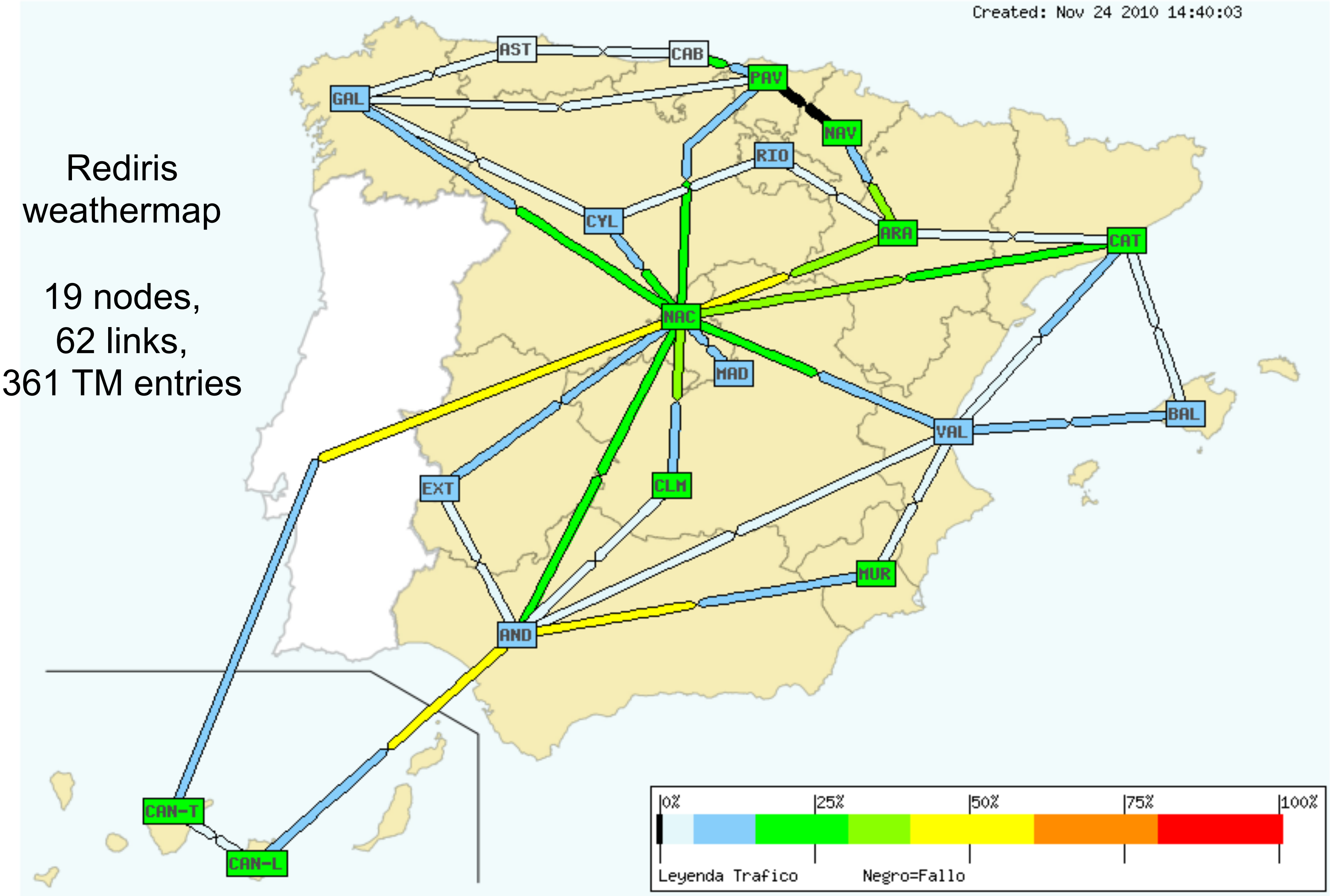
$$TM_{grav} = T_{total} \times p_{in} \times p_{out}{}^{T}$$

    - Only 2N parameters: in/out traffic fractions
    - Many possible solutions – projection to solution plane

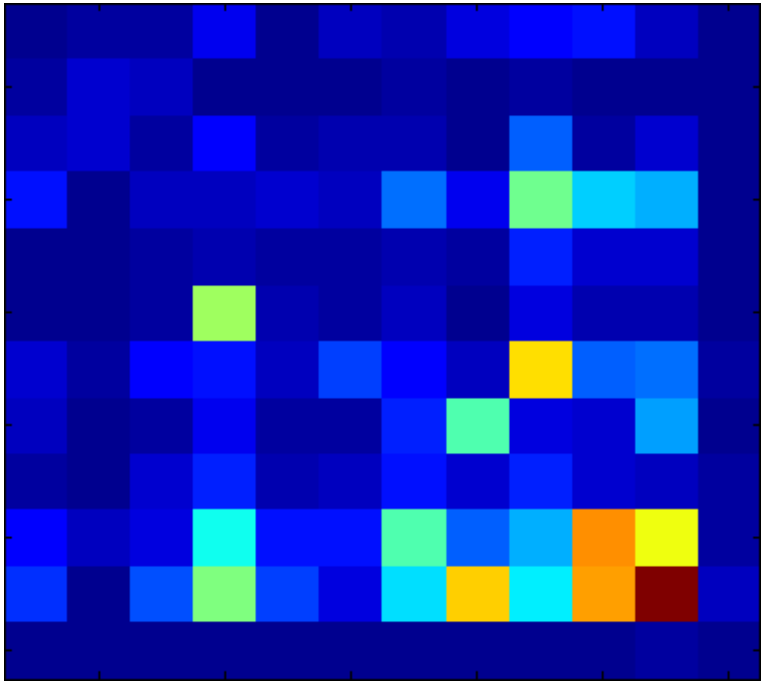
TM inference – example

Rediris weathermap

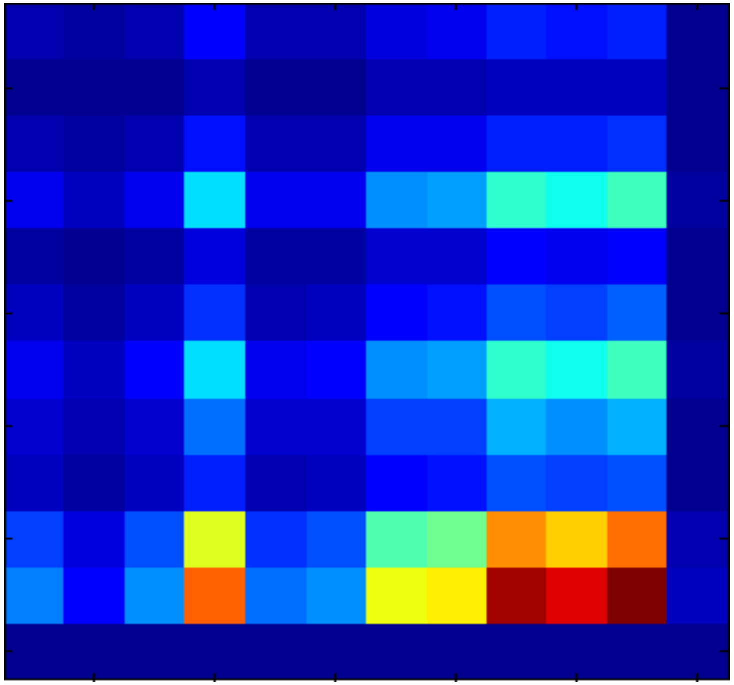
19 nodes,
62 links,
361 TM entries



TM inference – example



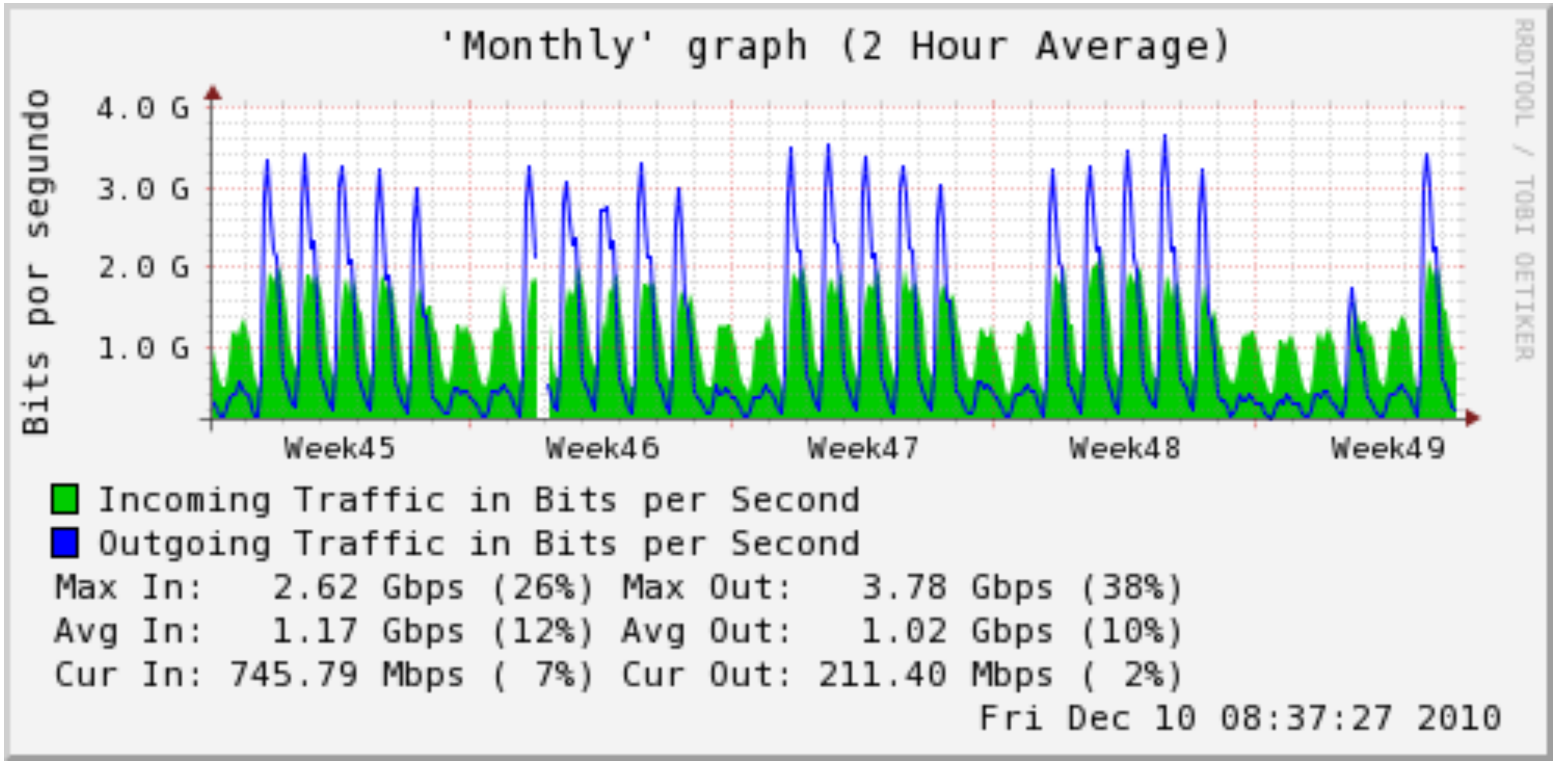
Original TM



Gravity model

Prediction – link load

Rediris AND-NAC link



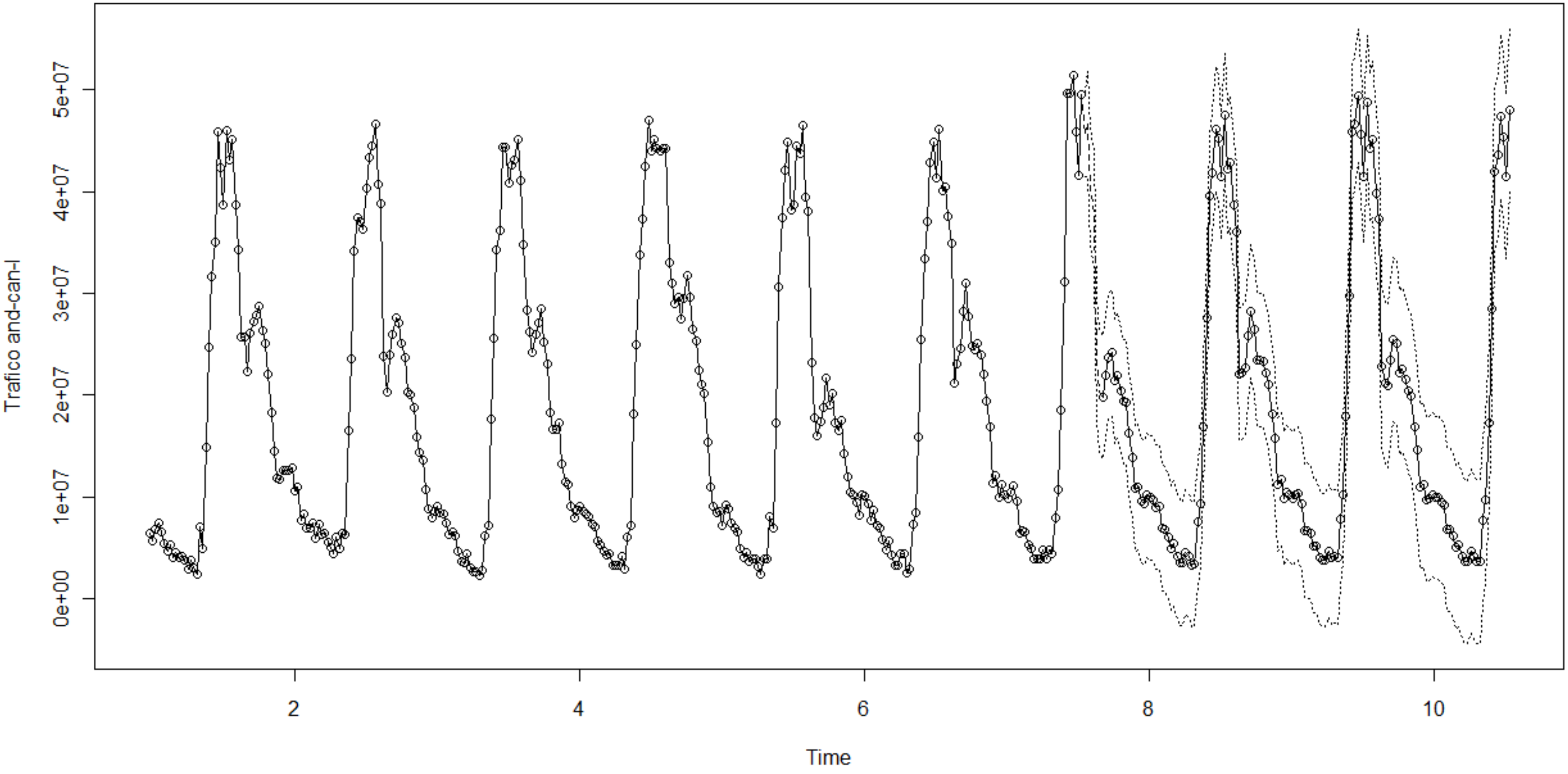
🟩 Incoming Traffic in Bits per Second

🟦 Outgoing Traffic in Bits per Second

```
Max In:    2.62 Gbps (26%) Max Out:    3.78 Gbps (38%)
Avg In:    1.17 Gbps (12%) Avg Out:    1.02 Gbps (10%)
Cur In: 745.79 Mbps ( 7%) Cur Out: 211.40 Mbps ( 2%)
                                      Fri Dec 10 08:37:27 2010
```

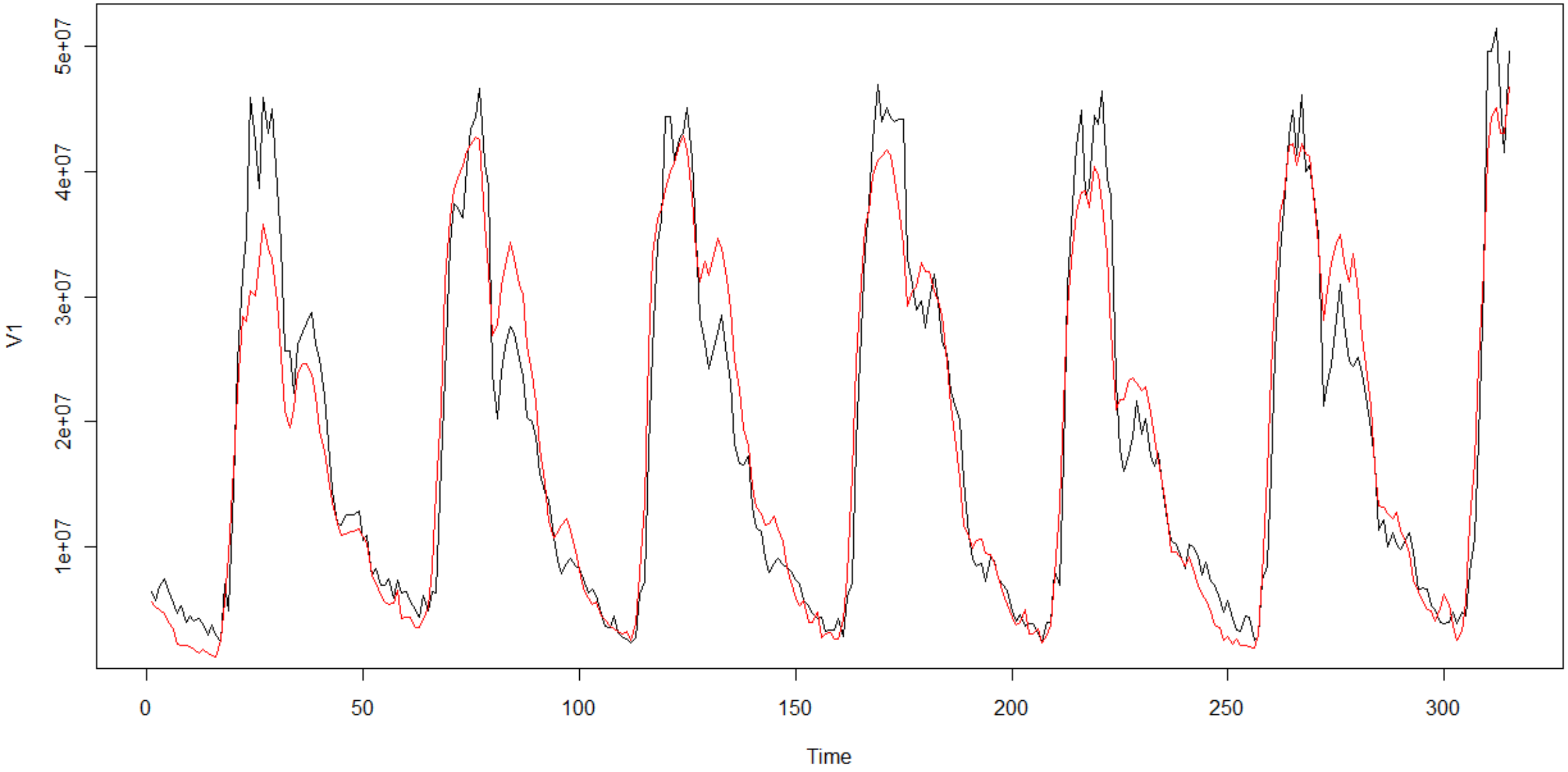
Prediction – link load

Rediris AND-CAN-L link, april 2010 – ARMA (1,1) model



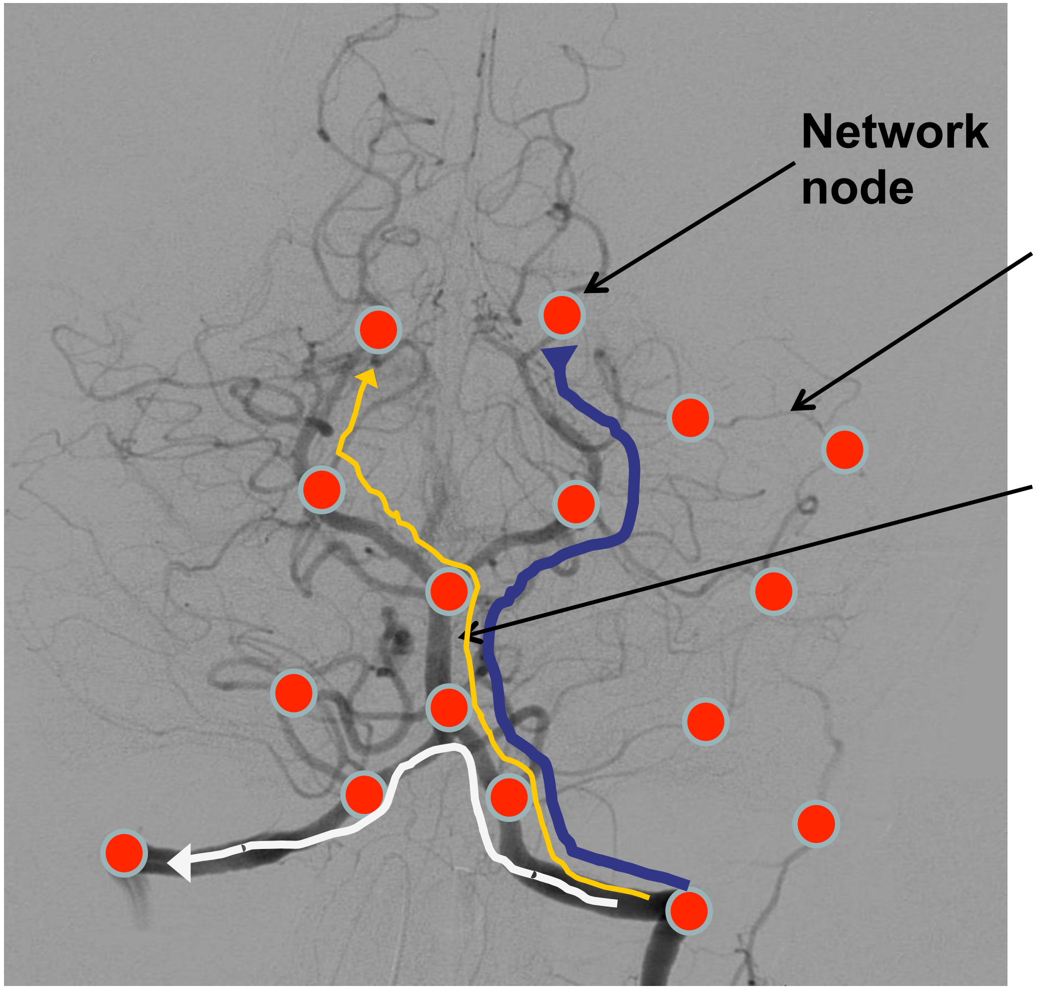
Principal Component Analysis

Rediris AND-CAN-L link, april 2010, PC1



Monitoring tool

A system able to discover / measure the largest trajectories in a network (ISPs, ASes, groups of them)



Network node

Underutilized link

Big path, several highly-utilized links

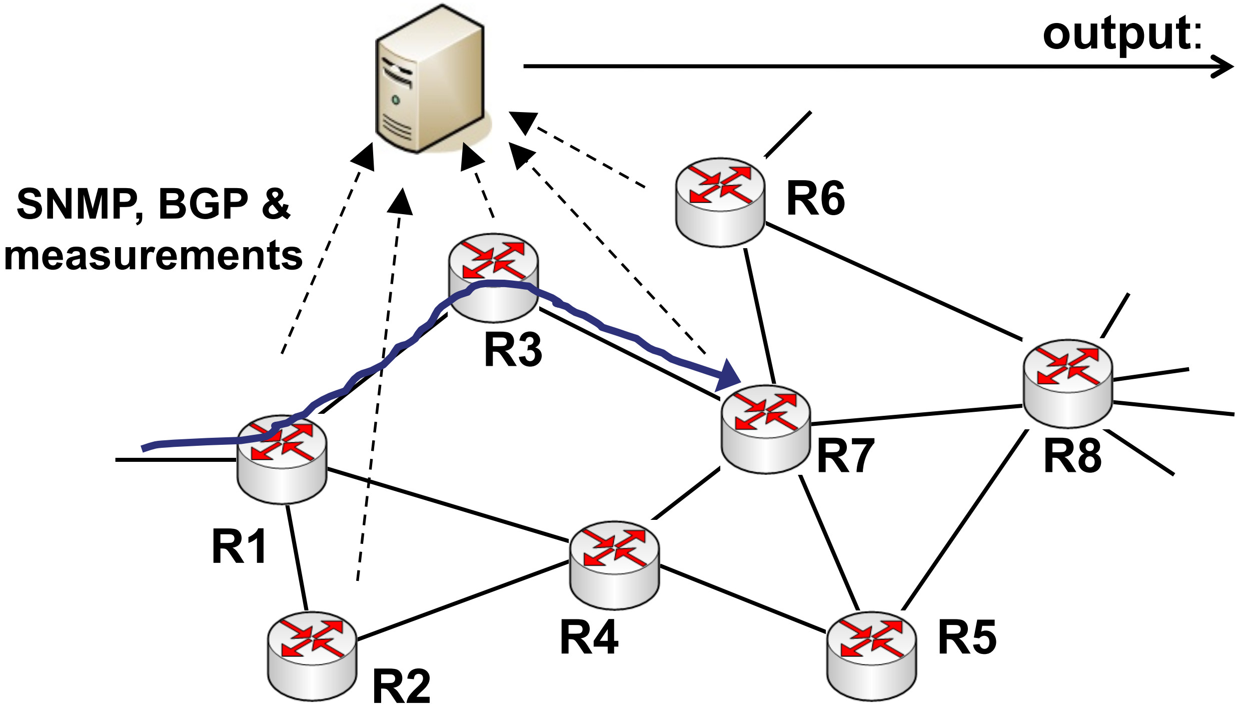
Who's the responsible for this?

What is the composition of that traffic? One flow? several? Same trajectory?

Approach

- To combine
  - traffic measurement data
  - routing information acquired via BGP
  - routing information acquired via SNMP
    - routers’ list of IP **addresses**:
    - routers’ **forwarding tables**
    - MPLS LSPs & TE tunnels
- To obtain real-time view of
  - topology
  - traffic matrix
  - flow trajectories

Approach



R1-R3-R7:
- volume
- rate (min/avg/max)
- k-top largest flows:
  - flow #1
  - flow #2
  - flow #3 ….

+ plots

R2-R4-R7-R8: …

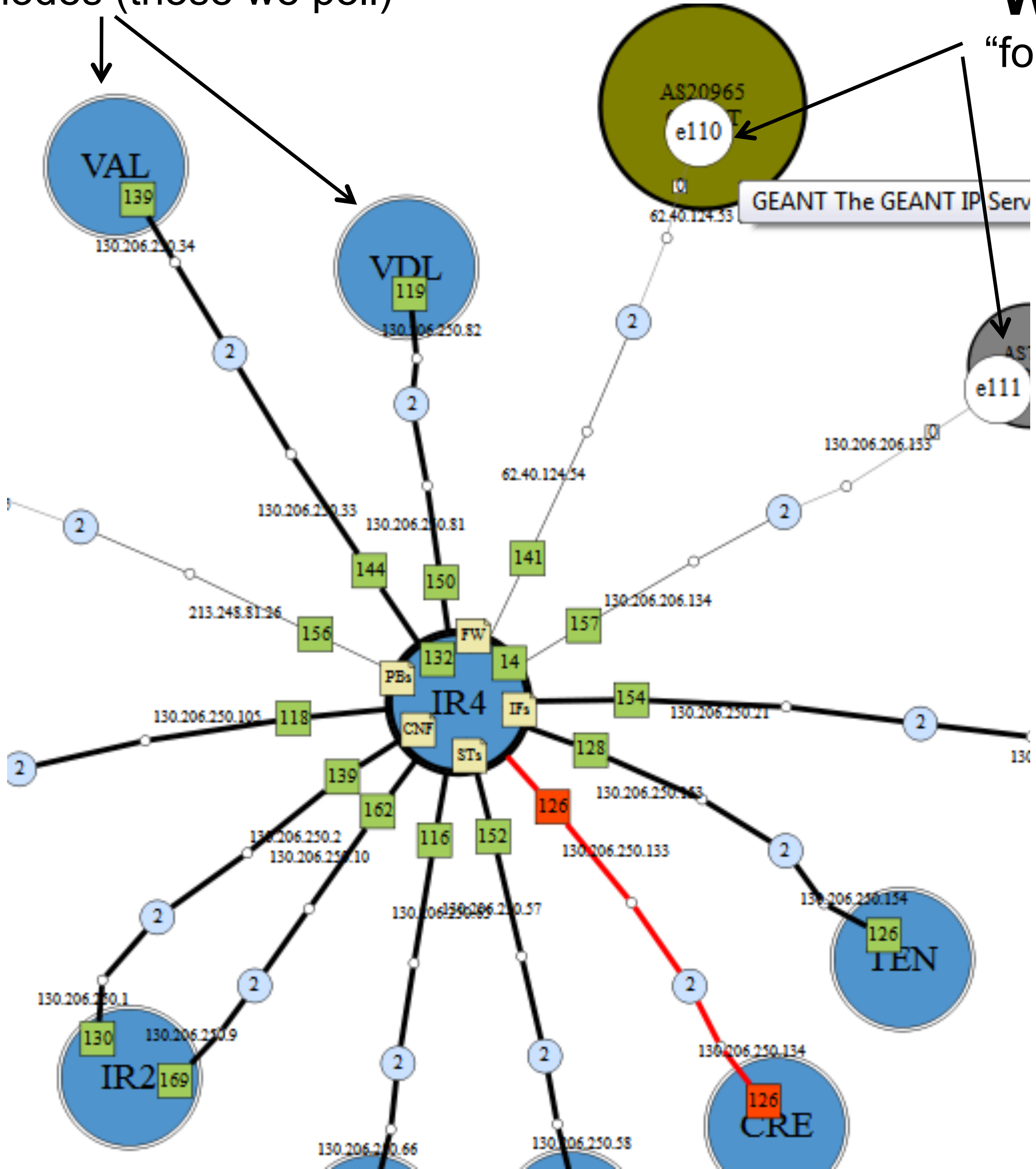
Problems / challenges / issues

- **Retrieving large forwarding tables** (> $3x10^5$ routes/router)
    - SNMP versus BGP approaches
- **Storing large forwarding tables** for a large number of routers→ scalability
- Further, storage must be such that:
    - **lookups** are efficiently handled
    - **route changes** are easily performed (updates)
- How to keep all the forwarding data **up-to-date**
    - periodically? Consider traps? State of interfaces? ….
- **Accuracy**: how to be sure that traffic actually goes where records indicate?
    - Load balancing, per packet or per-flow
    - Still, tool may reveal very useful information.

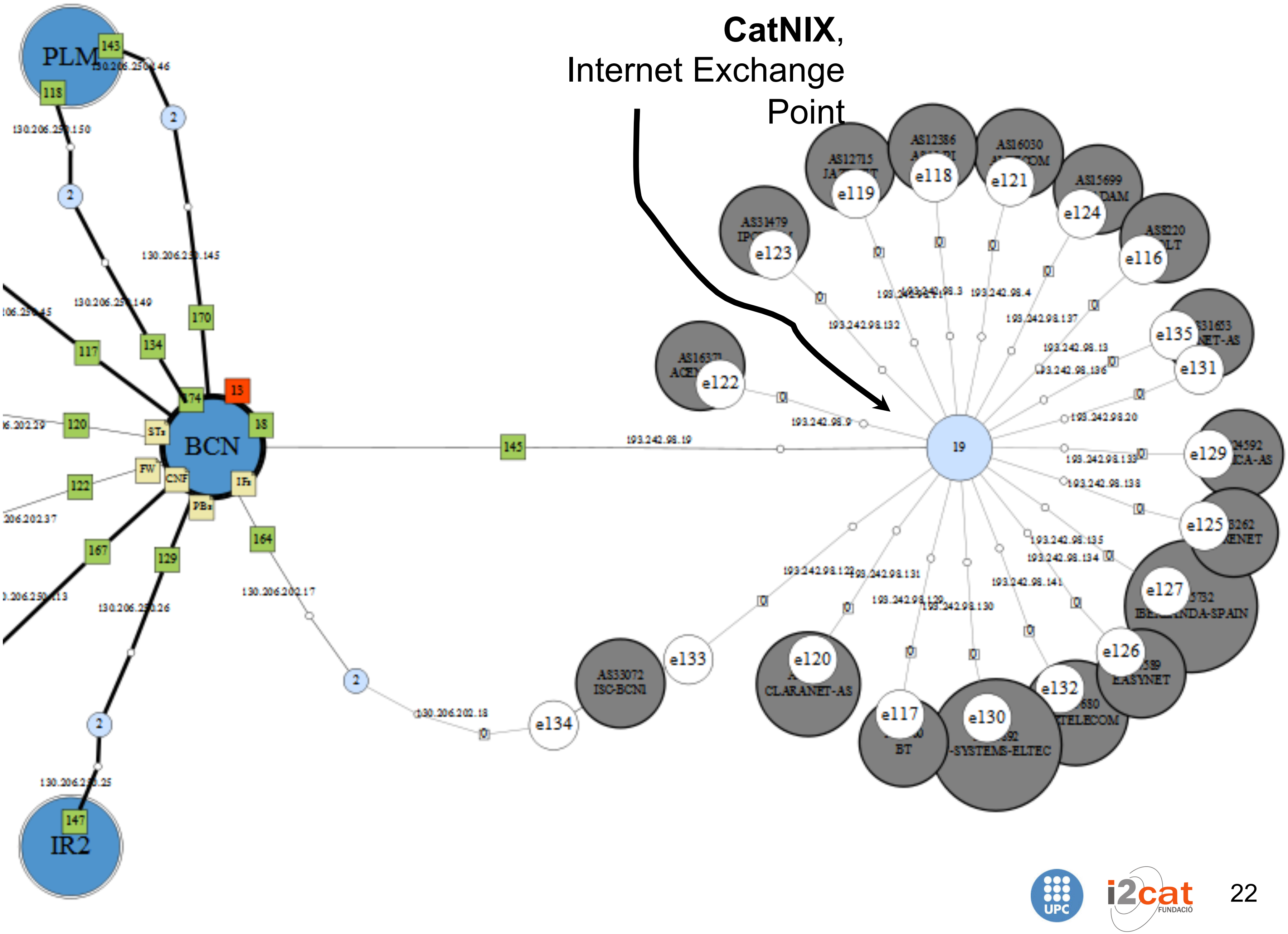
Our “little RedIris”

Blue nodes are "local" nodes (those we poll)

White nodes are "foreign": nodes that we discover



UPC i2cat FUNDACIÓ



CatNIX, Internet Exchange Point

Why are we here?

Our motivation to come

- We are interested in
  - Monitoring applications/procedures you use
  - Learning about your needs
  - Deploying and testing the monitoring tool
  - Analyzing real data
- Other activities
  - OpenFlow
  - Routing / resource optimization
    - Green(er) networking

Scalable flow-level measurements and traffic analysis

Fredi Raspall
David Rincón

UPC
UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

i2cat
FUNDACIÓ

GORE 8 - Barcelona, October 6th 2011