Dynamics of the ego-centred views of the internet topology

Clémence Magnien

Complex Networks Team
LIP6 - CNRS - Université Pierre et Marie Curie

September 15th, 2011
Outline

1. Measurements
2. Analysis
3. Modeling
Outline

1. Measurements
2. Analysis
3. Modeling
Internet topology – IP level

- IP addresses
- IP-level hops

Internet considered as a graph, study its structure
Measurement

No available map $\rightarrow$ Need for measurement
Measurement

No available map $\rightarrow$ Need for measurement

Route from $S$ to $D$ $\rightarrow$ traceroute

\[ S \quad ? \quad D \]
Measurement

No available map $\rightarrow$ Need for measurement

Route from $S$ to $D$ $\rightarrow$ traceroute
Measurement

No available map $\rightarrow$ Need for measurement

More information

Add

- Destinations ($\sim$ 1 million)
- Sources ($\sim$ a few dozen)

Costly, Long!
Properties: observed vs real

- bias in the observed structure (degrees, ...)

→ Theoretical and empirical work for:
  - Evaluating bias
  - Evaluating given properties with certainty

No perfect method
A radar for the internet

Studying dynamics?

**Orthogonal approach: ego-centred view**

- one monitor
- several destinations
- well-defined object
- high frequency: radar
A radar for the internet

Studying dynamics?

Orthogonal approach: ego-centred view

- one monitor
- several destinations
- well-defined object
- high frequency: radar
Measurements

- New tool design: tracetree

<table>
<thead>
<tr>
<th>Measurements</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>~ 100 monitors</td>
<td>several months</td>
</tr>
</tbody>
</table>

Tool and data publicly available
Outline

1. Measurements
2. Analysis
3. Modeling
Dynamics of the ego-centered views of the internet

Expected Dynamics

- load-balancing
- topology time-evolution
- events

Observed IP address set evolves with time

→ characterisation, description?
Dynamics: stabilisation?

Number of distinct IP addresses seen since measurement beginning

[Magnien et. al, 2009]

new IP addresses continuously discovered at an important rate
Observations vs consecutive observations

5 observations, 2 consecutive observations
Observations vs consecutive observations

Triangle: $\text{Nb consecutive observations} \leq \min(\text{nb observations, nb non-observations})$
Observations vs consecutive observations

Parabola: Number of expected cons. observations for this number of observations: $n \times (x/n) \times ((n - x)/n)$

→ load-balancing
Observations vs consecutive observations

Two different classes

parabola and above: load-balancing
close to x axis: stable addresses
Outline

1. Measurements
2. Analysis
3. Modeling
Explain observed behaviours with simple model

Need to model:

- topology
- topology dynamics
- routes
- load balancing
Explain observed behaviours with simple model

Need to model:

- topology
- topology dynamics
- routes
- load balancing

random graph
randomly replace $k$ links
shortest paths
random order in BFS
Preliminary results

- $n = 500,000$
- $m = 1,000,000$
- 3000 destinations
- 5 link changes at each step
Preliminary results

- $n = 500,000$
- $m = 1,000,000$
- 3000 destinations
- 5 link changes at each step
Preliminary results

Differences

- upper triangle: per-destination load-balancing, tool artefact
- parabola: same
- lower triangle: present but less visible in real data
Conclusion and Future Work

Conclusion

- Analysis of ego-centered views dynamics
- Identified different factors
- Proposed a model that reproduces observed behaviours

Suites

- Modeling: reproduce other observable properties
- Couple with long-term evolution models
- Simulations: varying measurement frequency → Quantify the impact of dynamics factors
- Formal analysis