Impact of power-law topology on IP-level routing dynamics: simulation results

Amélie Medem, Clémence Magnien, Fabien Tarissan
CNRS and UPMC Sorbonne Universités
LIP6 Laboratory

Netscicom 2012
30 mars 2012
Orlando FL, Florida

Laboratoire d'Informatique de Paris 6

Studying the Internet topology

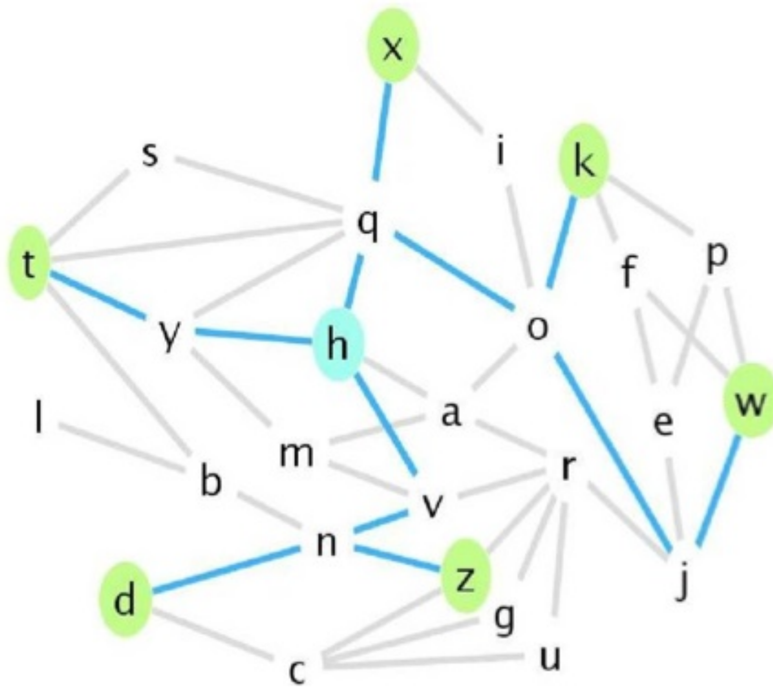
- Mapping the Internet topology is important
  - Future network protocols hard to test on real Internet
- Understanding how the Internet topology evolves in times is equally or even more important

→ Study its dynamics

Goal: IP-level routing topology and its dynamics

Ego-centered view

- **Focus**: the IP routing topology around a single node



Ego-centered view: what a node sees from the Internet

Outline

- (1) Measure the Internet IP topology around a node

What a machine sees from the internet ?

- (2) Extract dynamic behaviors

How the topology measured evolves in time ?

- (3) Confront them with the dynamics in artificial graphs

How to explain the observed dynamics ?

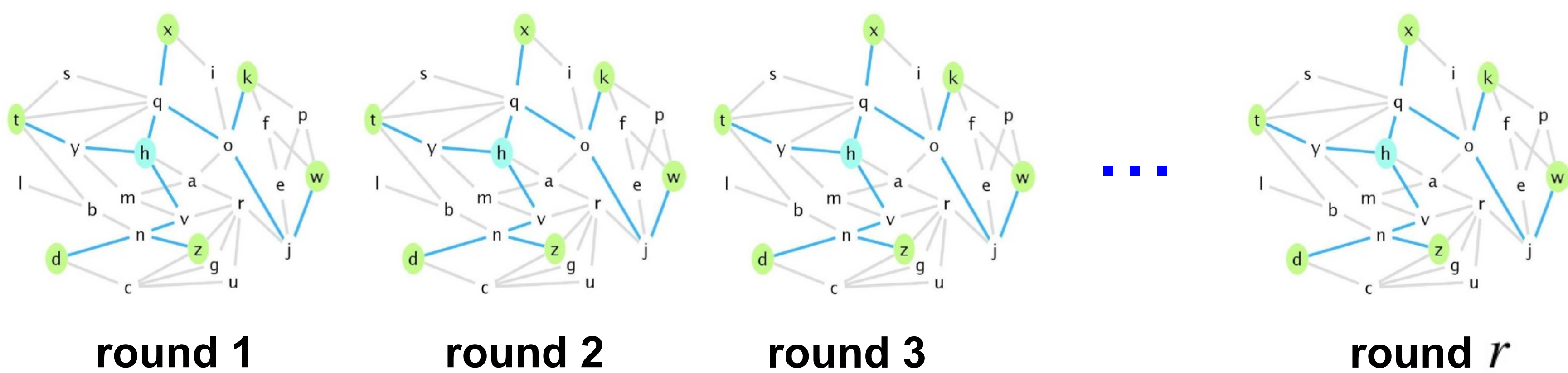
Outline

- (1) Measure the Internet IP topology around a node
- (2) Extract dynamic behaviors
- (3) Confront them with the dynamics of artificial graphs

Measures

- **Tracetree:** traceroute-like measurement tool
  - A routing tree of IP paths from 1 source to destinations
- **Radar:** periodic measure with **Tracetree**

Series of routing trees



Sources: mostly **PlanetLab** (> 150); Destinations : **random**; Measurement frequency : **15 mn.** or less

Web site: "A Radar for the Internet–Publicly available datasets."http://data.complexnetworks.fr/Radar/

Outline

- (1) Measure the Internet IP topology around a node
- (2) Extract dynamic behaviors
- (3) Confront them with the dynamics in artificial graphs

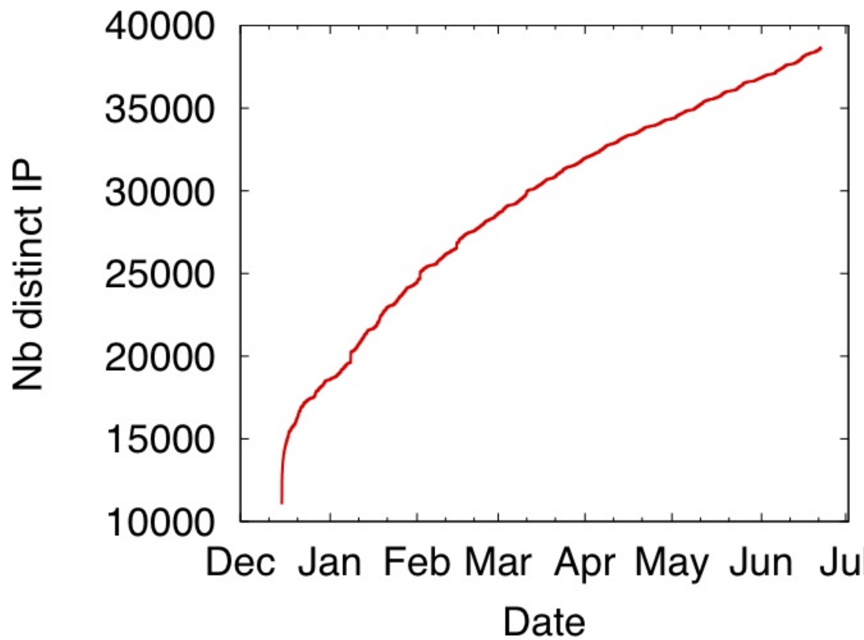
Dynamic behaviors

- Two dynamic behaviors
  - IPs discovery
  - Pattern of occurrence of IPs

IPs discovery

- Stabilization ?

Number of distinct IPs observed since the beginning



New IPs are persistently discovered

- **woolthorpe**
- 7 months
- 3,000 destinations
- 17,450 rounds

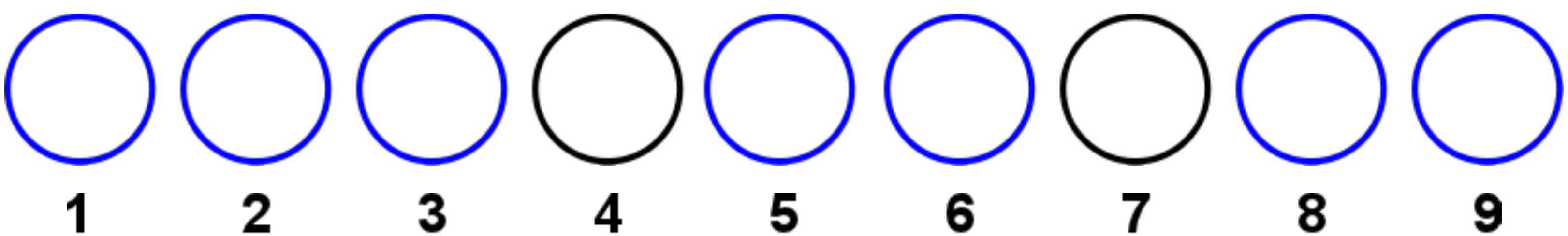
→ High rate of discovery: ~150 new IPs per day

→ Not due do measurement artifacts

Pattern of occurrence of IPs

- 2 values to quantify the occurrence of IPs
  - Observation number: The total of distinct rounds in which it occurs
  - Block number: The number of groups of consecutive rounds in which it is observed

9 rounds



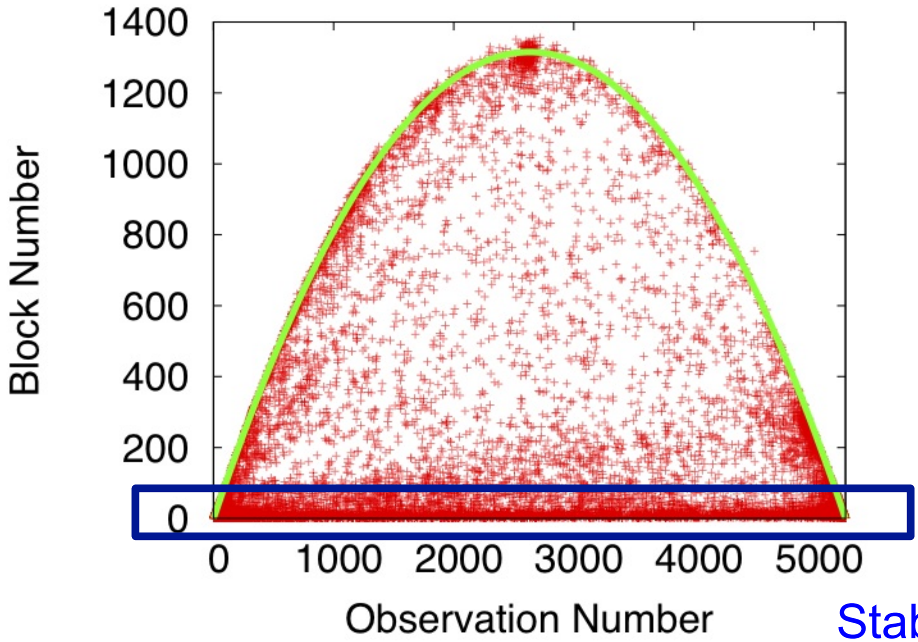
An **IP**$_a$ appears in rounds: 1, 2, 3, 5, 6, 8, 9

Observation number of **IP**$_a$ = 7

Block number of **IP**$_a$= 3

Pattern of occurrence of IPs

- Observation number *vs.* block number



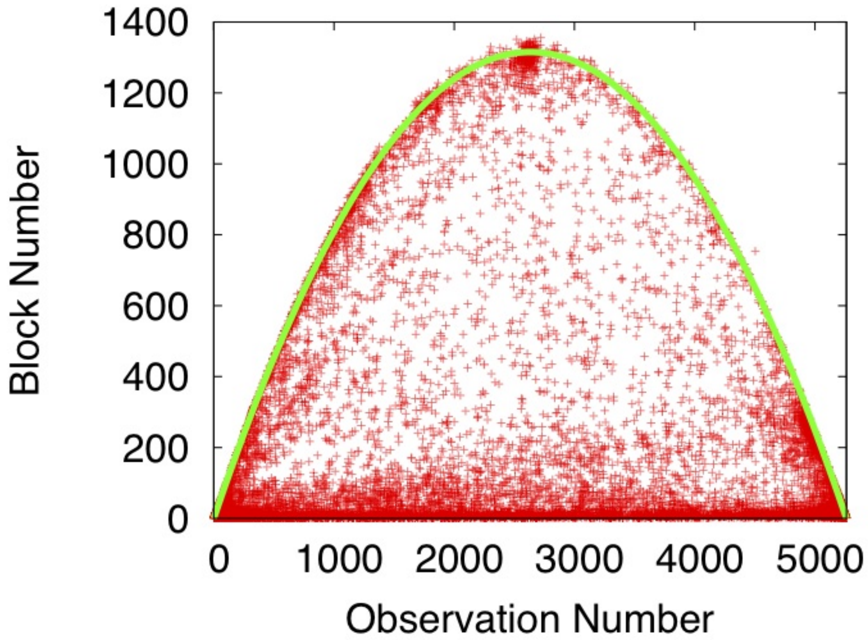
Stable IPs:
observed on consecutive rounds

(1) A parabola

(2) A large set of points close to the x-axis

Pattern of occurrence of IPs

- The parabola, explained through Load balancing



- Given r rounds of measurements
  - IPs on load-balancing paths has Prob p of being observed
  - Observation number: rp
  - Block number: $rp(1-p)$

Outline

- (1) Measure the Internet IP topology around a node
- (2) Extract dynamic behaviors

  → Hard to explain them without knowing the Internet

- (3) Confront them with the dynamics in artificial graphs

Our approach

- Goal: Simulate **Tracetree** measurements on artificial graphs

- Generate an initial undirected graph **G**
  - Power-law graph, Erdös-Rényi graph ...
- Simulate **Tracetree** on **G** to create a routing tree **T**
  - Shortest-path model, ...
- Incorporate on **G** well-known route change factors
- Repeat previous steps to simulate periodic **Tracetree**

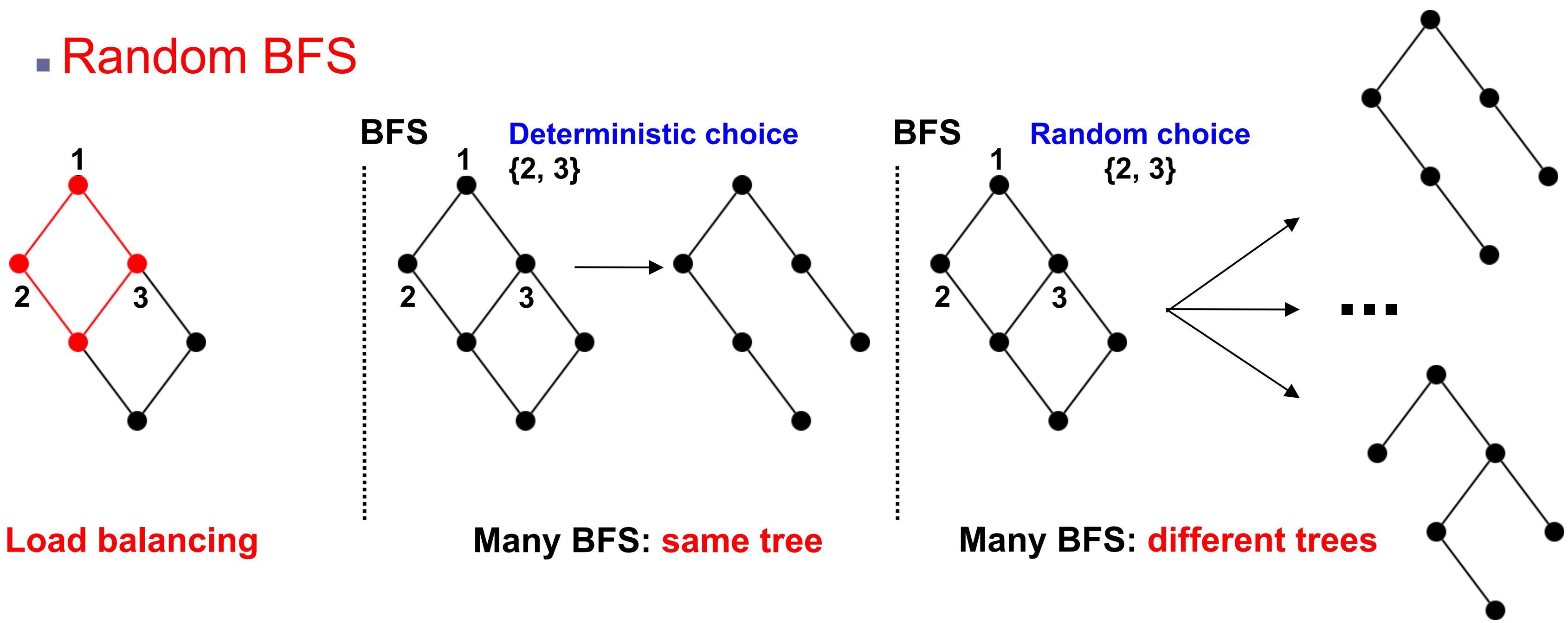
Route changes

- Well-known **factors of route changes**
  - Load balancing
  - Route evolution

How to model these factors ?

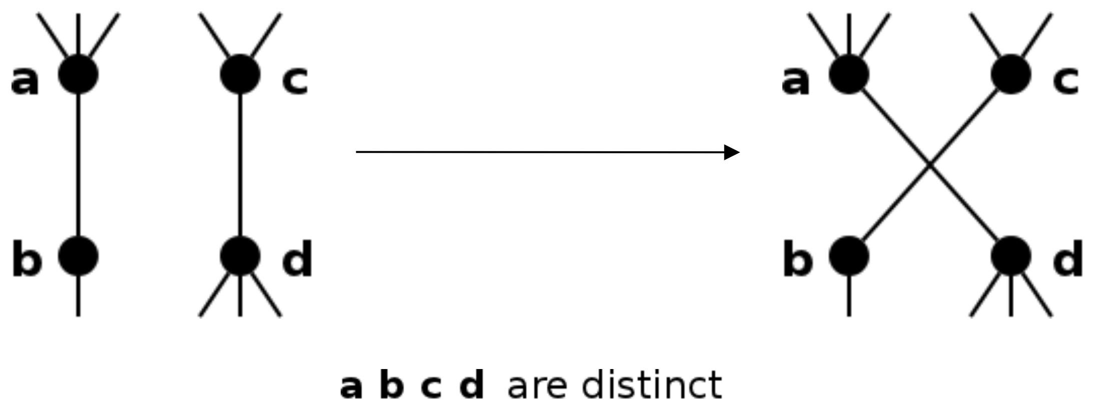
Modeling load balancing

- **How to model Load balancing ?**
- Tracetree measurement simulation
  - Shortest path model
  - Routing tree → BFS from the source
- Random BFS



Modeling route evolution

- **How to model route evolution ?**
  - Possible approaches:
    - Realistic or not: link rewiring, nodes added or removed
  - Link rewiring or *swap*



Main interest: the degree distribution of nodes is conserved

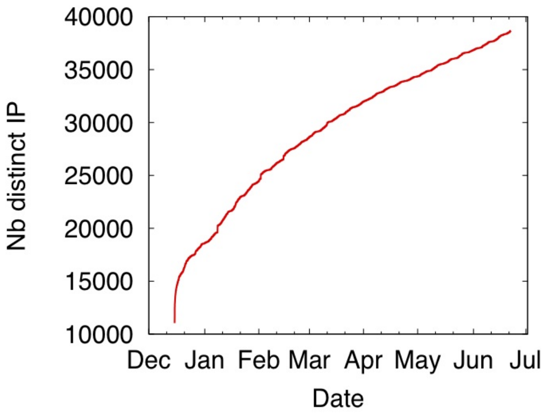
Simulations

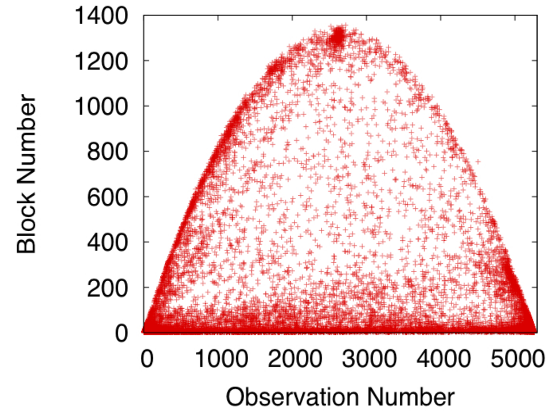
- **Goal:** Simulate **Tracetree** measurements on artificial graphs

- (1) Generate a Power-law (PL) graph **G = (V,E)**
    - Nodes: ***n***; Exponent: ***α***
- (2) Select 1 source and ***d*** destinations
    - Uniformly, at random
- (3) Simulate **Tracetree** from source to destinations in **G**
    - Shortest paths, Random BFS → $\mathbf{T_1}$
- (4) Simulate route evolution : Inject ***s*** link swap in **G**
- (5) ***Repeat steps 3 and 4*** → $\mathbf{T_2, T_3, T_4, \ldots, T_r}$

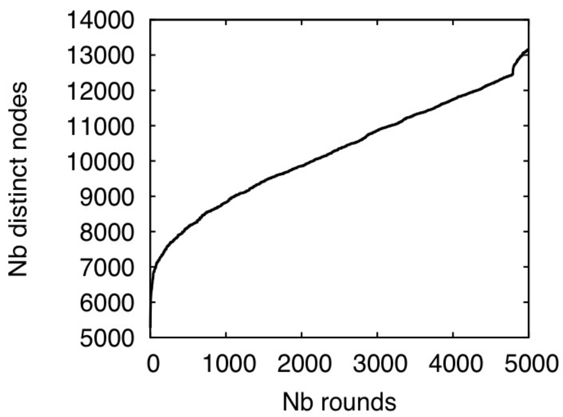
Reproducing the dynamics

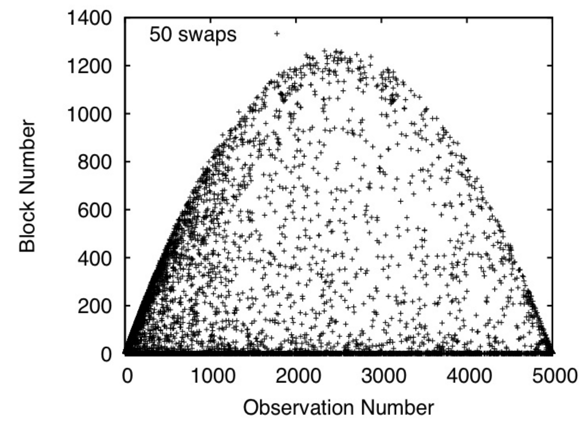
- Internet





- Simulations: Power graph with n=500,000; α=2.3; with d=3,000 **50 swaps**

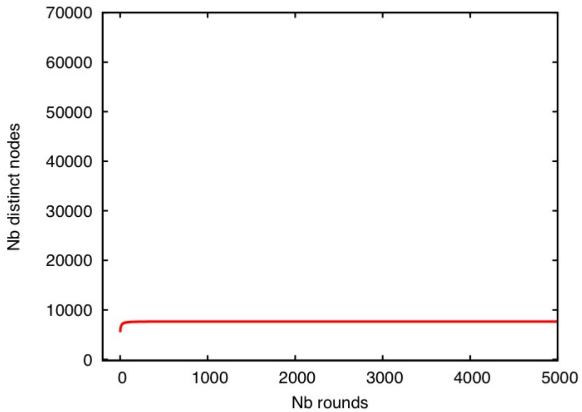




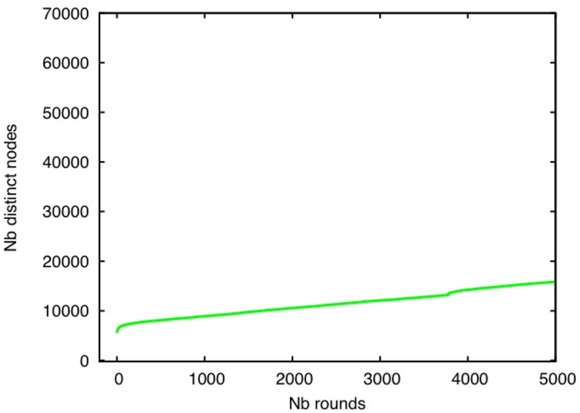
Qualitative similar behavior as observed in real Internet data !

Impact of simulation parameters

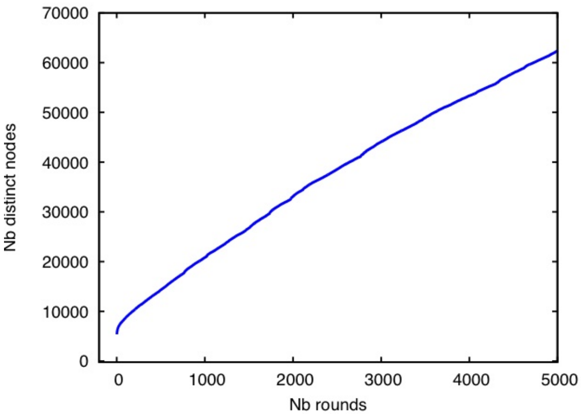
Impact of the number of swaps



0 swaps



100 swaps

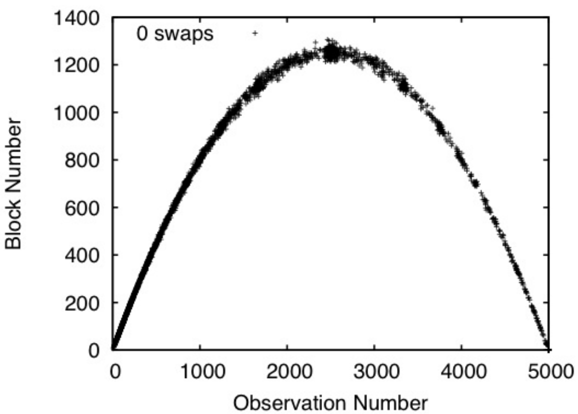


1,000 swaps

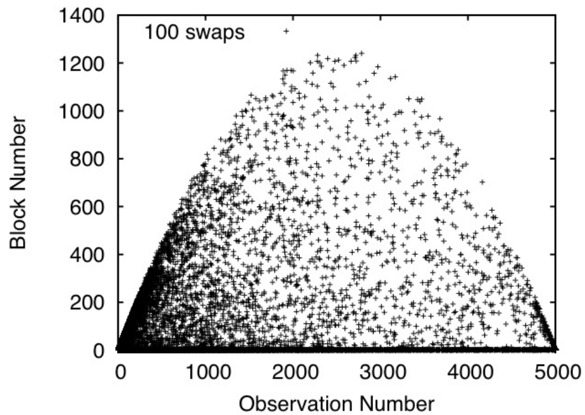
More swaps → Faster node discovery
No swaps → Stabilization
The persistent discovery of IPs: due to route evolution

Observation vs. block

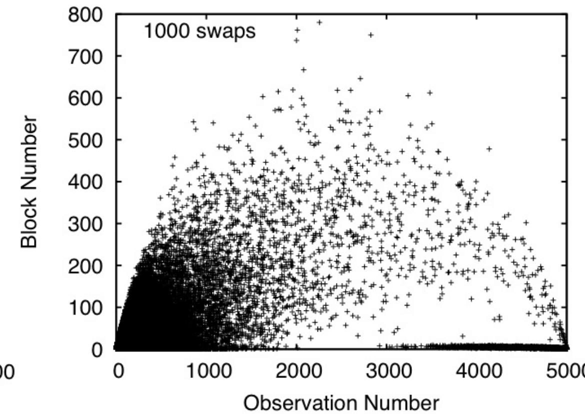
Impact of the number of swaps



0 swaps



100 swaps



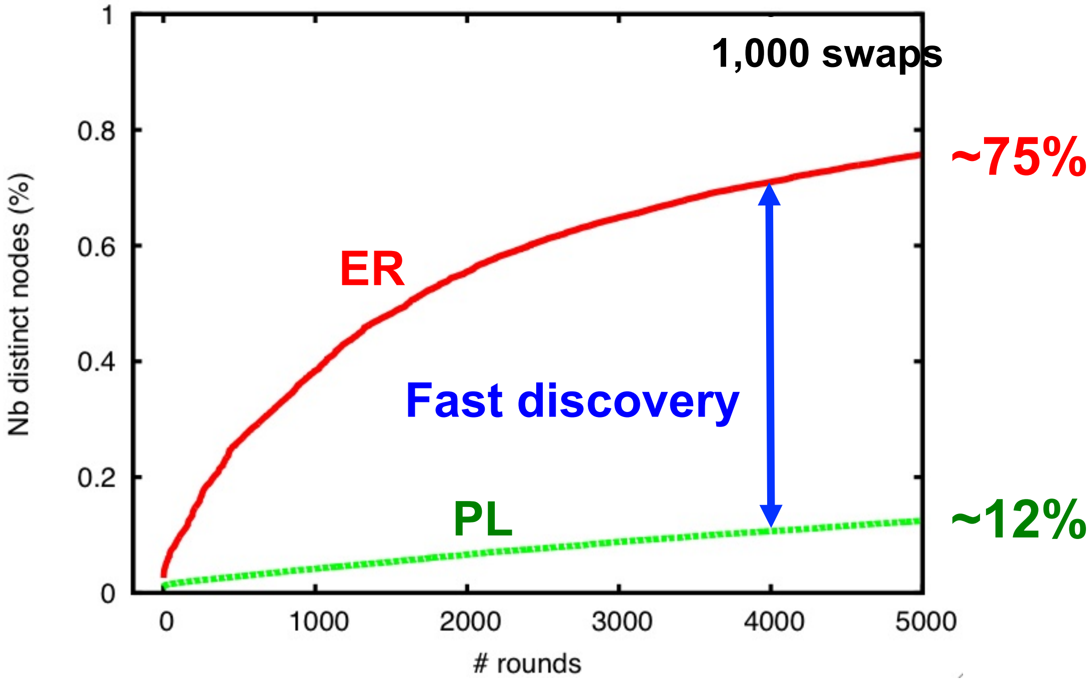
1,000 swaps

The parabola vanishes when the number of swaps increases

Points close to the parabola : due to load-balancing nodes

Power law vs. Erdös-Renyi

Quantitative difference between PL and ER graphs



500,000 nodes,
1,000,000 links,
1% for dest

Intuitions:

(1) Degree-1 nodes ?

(2) Average distance ?

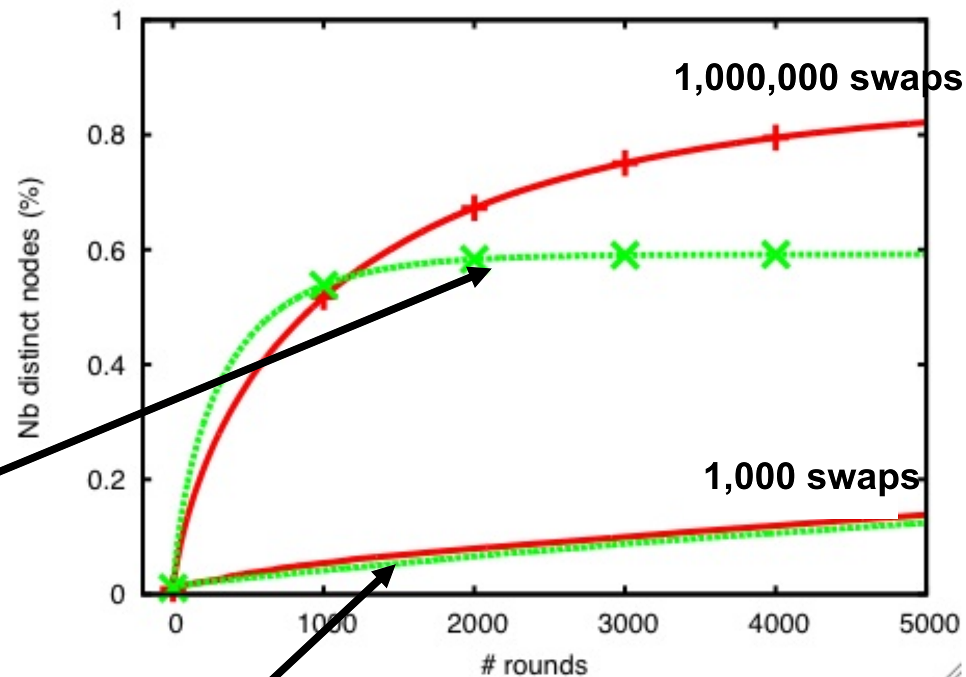
Power law *vs.* Erdös-Renyi

(1) Degree-1 nodes : Large fraction of nodes in PL graphs !!!

- Unless source/destinations, difficult for them to be discovered
- Not router nodes

PL original

PL wo degree-1 nodes



Flat phase, 99.9% of remaining nodes have degree 1

Same evolution

Power law *vs.* Erdös-Renyi

(2) The average distance is smaller in PL than in ER graphs !!!

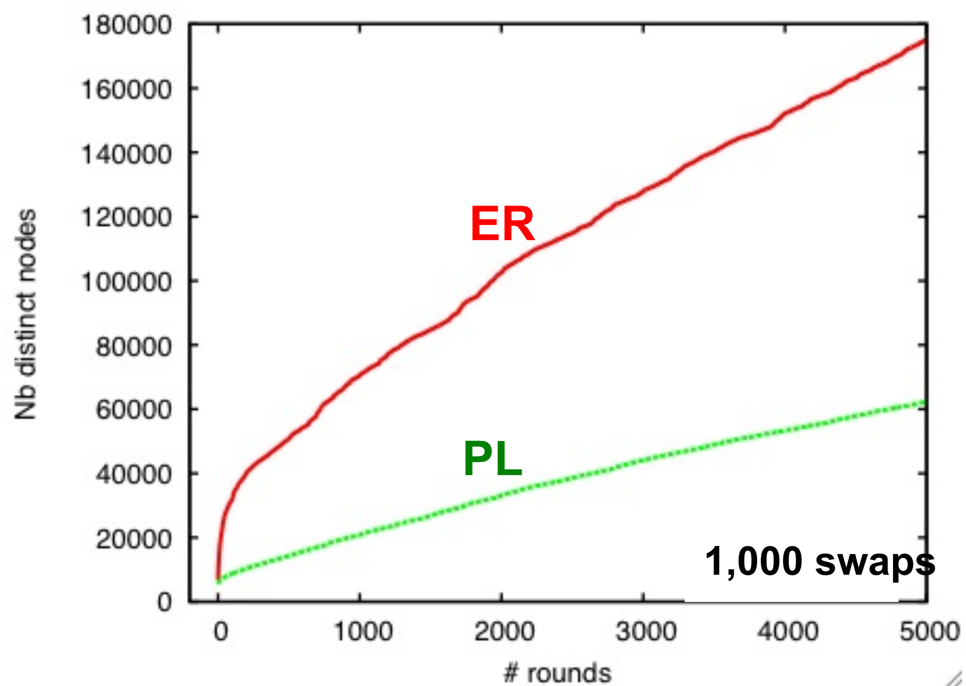
– PL graphs produce smaller routing trees than ER

- On avg: 5,363 *vs.* 12,868 (n=500,000; α=2.3/1,000,000 links)

Same average distance

ER (8,000,000 links)

PL (1,000,000 links)



Evolution, still faster in ER

Ongoing work !

Summary

- **Study the dynamics of Internet IP routing topology around a node**
  - Two main behaviors to characterize the dynamics
  - A model (IP topology, dynamics, **`Tracetree`**) for explanation
  - Observed dynamics reproduced on power-law graphs
  - Observations quantitatively different in Erdös-Rényi graphs
    - Degree-one nodes, Average distance
- **Perspectives**
  - Integrate other dynamics
    - Node adding/remove, link adding/remove
  - Test other topologies (realistic topologies)
  - Perform theoretical analysis (quantify the slopes of curves)

Dynamic IPs

Stable destinations