On the Feasibility of Measuring the Internet Through Smartphone-based Crowdsourcing

Adriano Faggiani, Enrico Gregori,
Luciano Lenzini, Simone Mainardi
and Alessio Vecchio

UNIVERSITÀ DI PISA

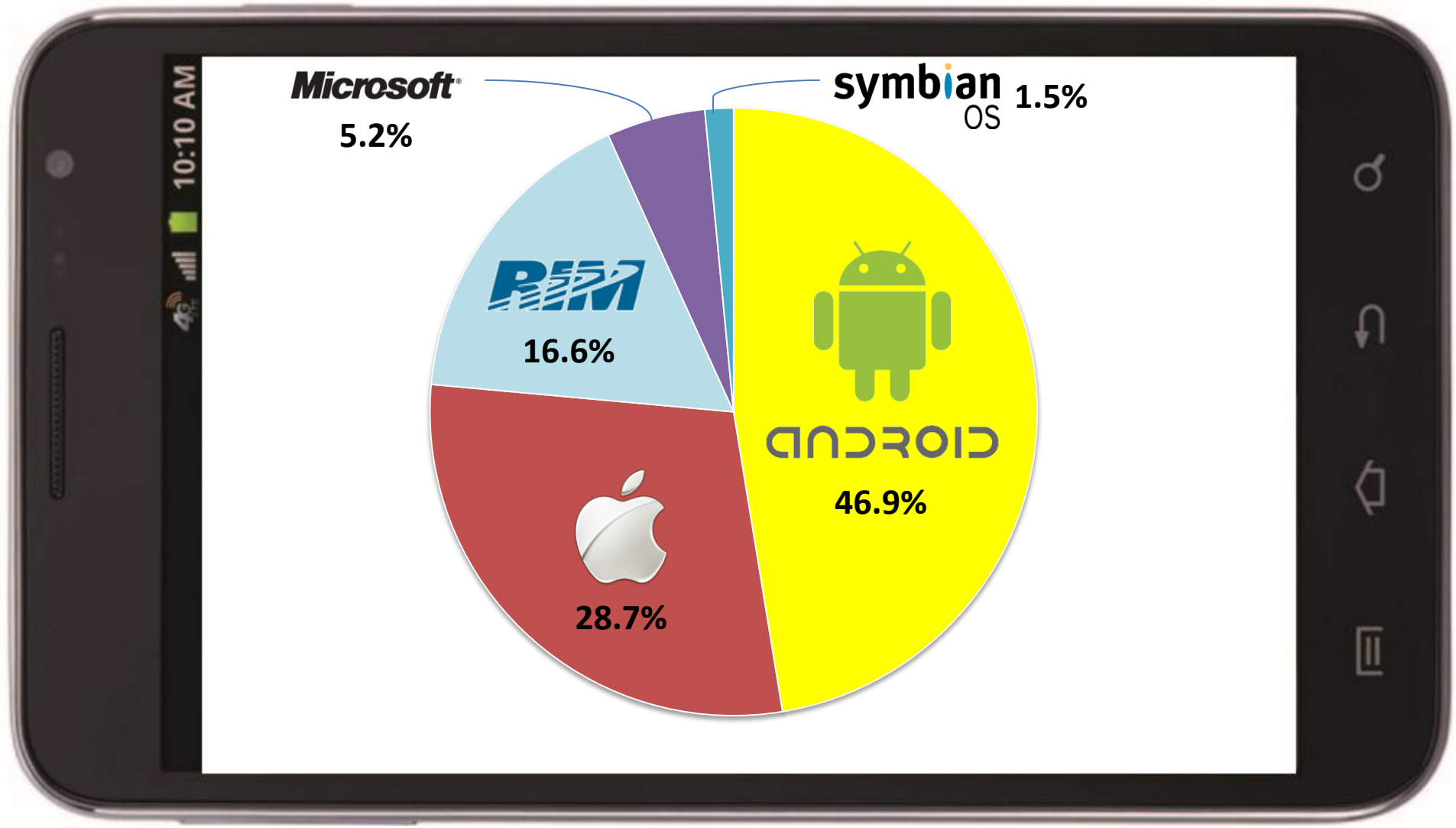
(Smart-)phones diffusion in the World

5 Billion mobile phones in the world …

… out of which **1.08 Billion** are smartphones

Source: *go-gulf.com*

Smartphone Platform Market Share



Source: *go-gulf.com*

How Much Firepower We Could Have!

- Approx. **500M** Android smartphones

Internet enabled

Multiple wireless network access technologies

GPS enabled

Firepower for What Purpose?

Internet + GPS + Multiple network access technologies…

Excellent opportunity for

- A **fine-grained observation** of the Internet structure
- A **quantitative evaluation** of its characteristics

Measuring Internet Periphery: Why?

Valuable information for Carriers and ISPs

- Resource provisioning and allocation
- Bottlenecks identification

Valuable Information for end-users

- QoS received from their ISP (e.g. latency, bandwidth)

Measuring Internet Periphery: How?

- Bottom-Up Approach (start the measurement campaigns from the end-users)
- As many end-users as possible
- Heterogeneous ISPs, Carriers and network access technologies

Measuring Internet Periphery: Why Existing Projects are not Enough?

CAIDA Archipelago
(**Ark**)

Distributed Internet Measurements
(**DIMES**)

- Few, fixed measurement agents (Ark 60, DIMES ~300)
- Uneven geographical distribution
- Non-mobile agents

Measuring Internet Periphery: What We Can Get

- Up to millions of measurement agents
- Widespread geographical distribution
- Mobile agents

Why Users Should Install Our Application

1. Encourage users by providing them a nice representation of their surrounding network properties (e.g. `speedtest.net`)

2. We can avoid telling them and add our code to a popular application

System Architecture Overview

- A (large) set of agents (i.e. the smartphones) that actually carry out the measurements
- A server that orchestrates measurement campaigns

Clients

- Register to the server and periodically send it updates with their position/network/battery level etc.
- Execute micro, short-range measurements according to the jobs received from the server
- Send the results back to the server

Server

- Receives the specification of a measurement campaign, divides it into a set of jobs and assign them to the agents
  - Job assignment could be based on agents geographic position, network, battery level, etc.

- Collects, refines and stores the measures received from the agents.

Implementation

- Android agent-side **traceroute**
- UDP-probes
- **User-space** measurements by using datagram sockets IP_RECVERR option

```
struct sock_extended_err {
    /* error number */
    u_int32_t ee_errno;
    /* error origin */
    u_int8_t ee_origin;
    /* type */
    u_int8_t ee_type;
    /* code */
    u_int8_t ee_code;
    …
};
```

Experiments

- **Traceroutes** to 141 targets within the **GARR**, the Italian academic and research network

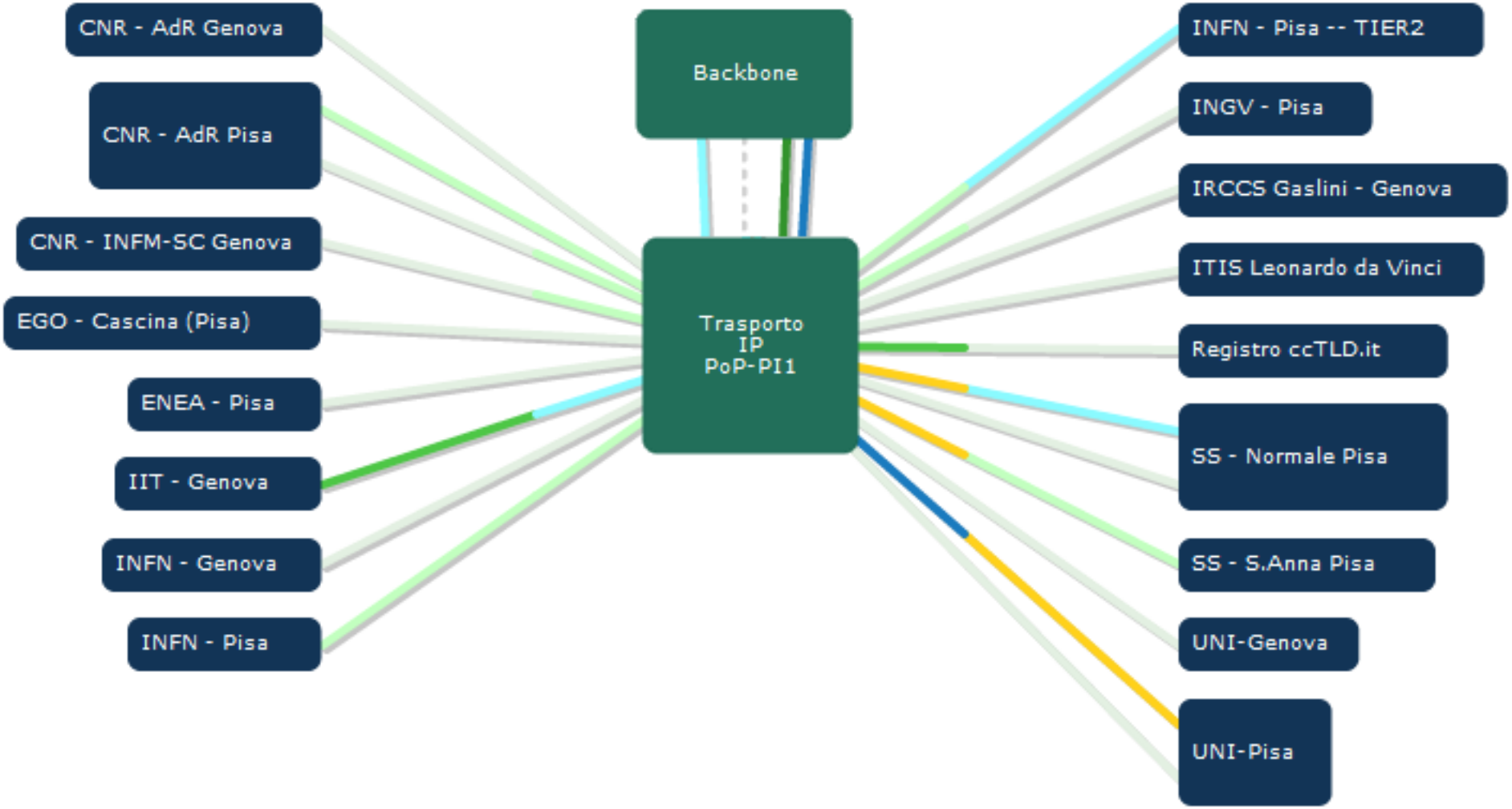
- From the **CNR** of Pisa

- **GARR** network topology and statistics publicly available

Validation & Evaluation

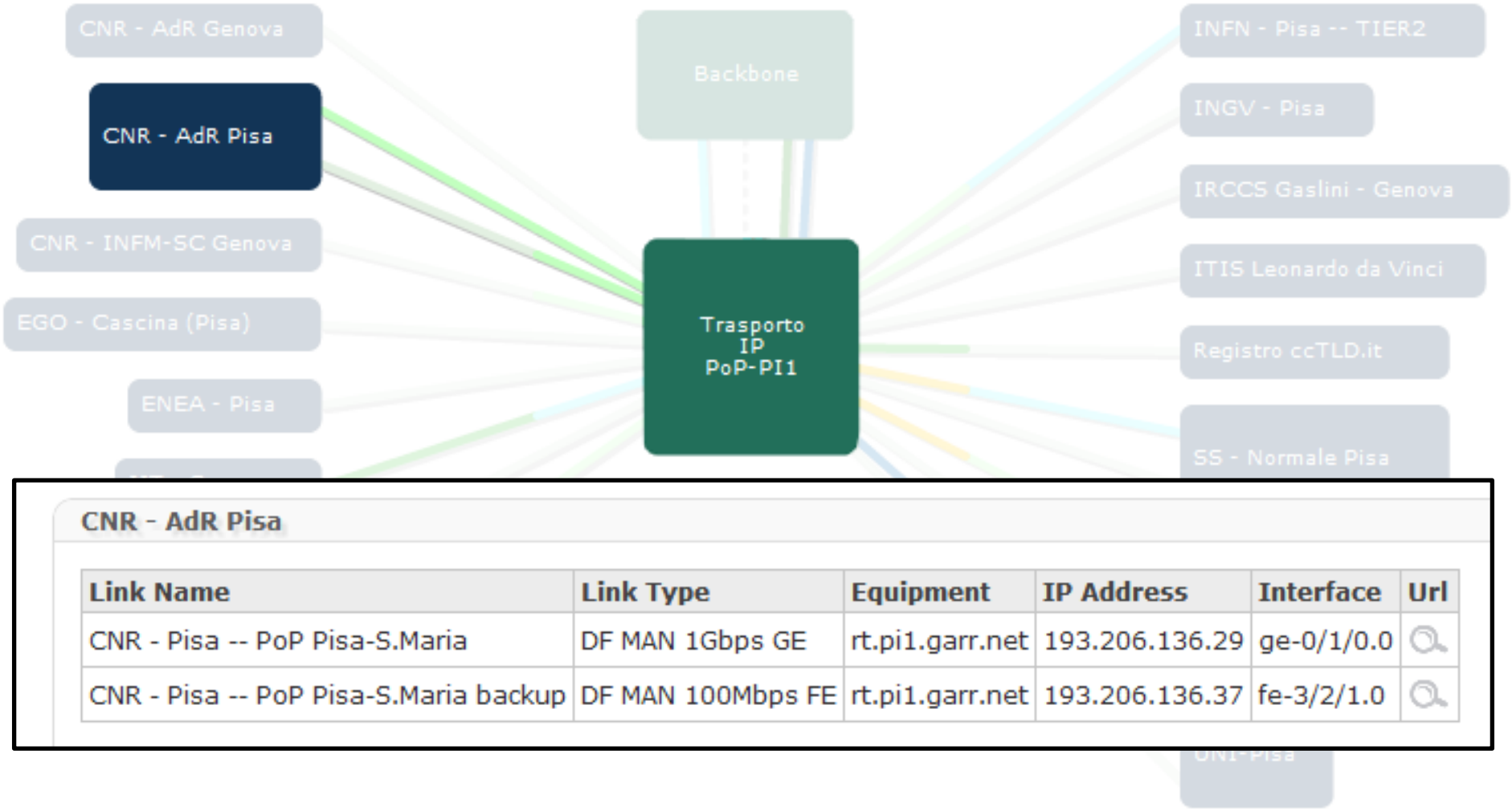
- Validation by comparing discovered interfaces with:
  - Those made available by the **GARR**
  - Those discovered by `paris traceroute`
- Evaluation by:
  - Determining to what extent discovered interfaces are reported by **CAIDA Macroscopic Internet Data Kit (ITDK)**
  - Measuring battery consumption

The Importance of Measuring the Internet from its Periphery: An Example



GARR PoP PI1

The Importance of Measuring the Internet from its Periphery: An Example



CNR - AdR Pisa

Link Name	Link Type	Equipment	IP Address	Interface	Url
CNR - Pisa -- PoP Pisa-S.Maria	DF MAN 1Gbps GE	rt.pi1.garr.net	193.206.136.29	ge-0/1/0.0	
CNR - Pisa -- PoP Pisa-S.Maria backup	DF MAN 100Mbps FE	rt.pi1.garr.net	193.206.136.37	fe-3/2/1.0	

GARR PoP PI1

The Importance of Measuring the Internet from its Periphery

CNR - AdR Pisa

Link Name	Link Type	Equipment	IP Address	Interface	Url
CNR - Pisa -- PoP Pisa-S.Maria	DF MAN 1Gbps GE	rt.pi1.garr.net	193.206.136.29	ge-0/1/0.0	
CNR - Pisa -- PoP Pisa-S.Maria backup	DF MAN 100Mbps FE	rt.pi1.garr.net	193.206.136.37	fe-3/2/1.0	

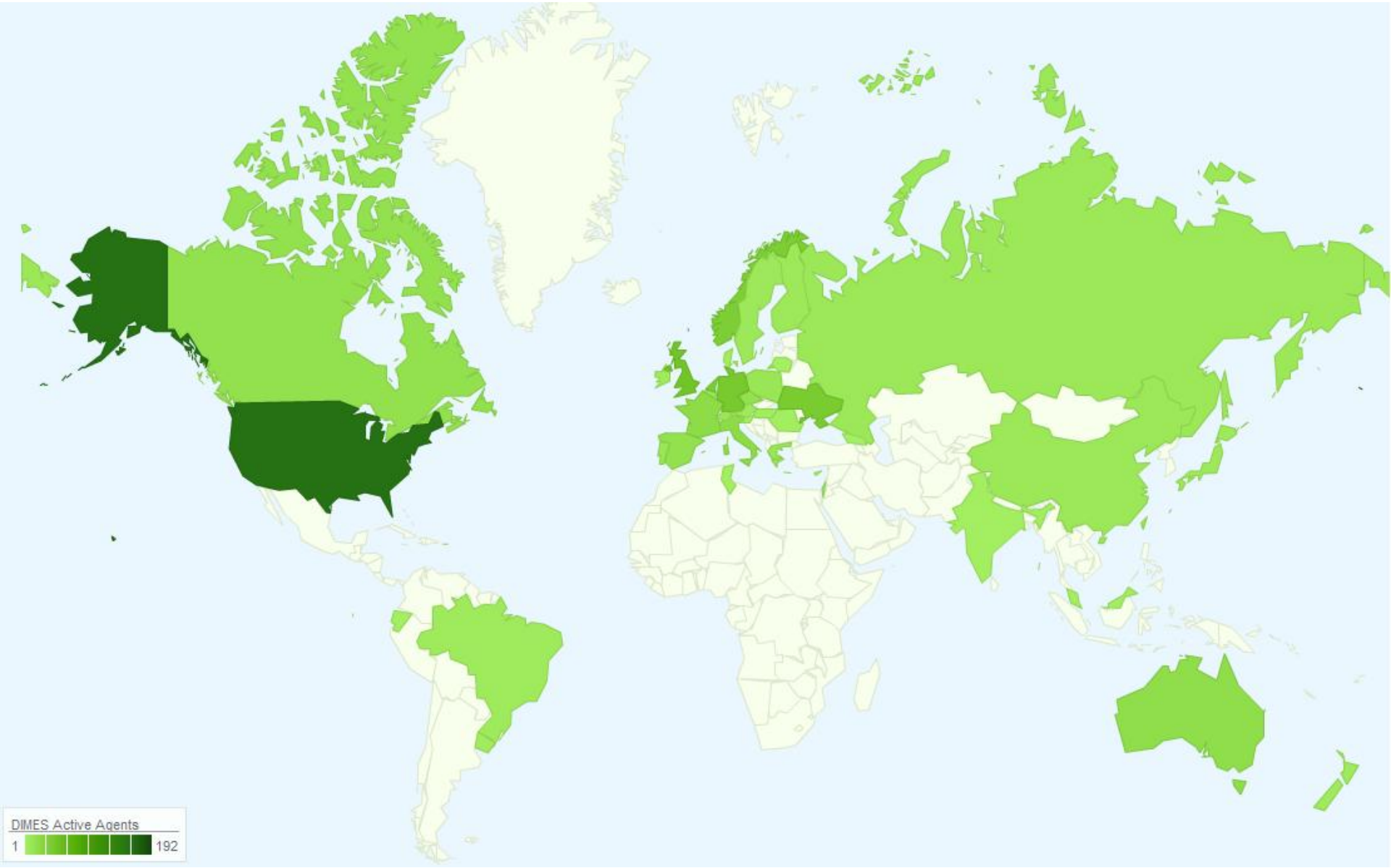
- Interface $192.206.136.29$ is present in our traceroutes but is missing in **CAIDA ITDK**

- **Fundamental** to traverse links in both directions

Thank You!

Questions?

DIMES – Agents Distribution



Ark – Agents Distribution