

# Lost in Space: Improving Inference of IPv4 Address Space Utilization

**Alberto Dainotti**, Karyn Benson, Alistair King,  
Bradley Huffaker, Eduard Glatz,  
Xenofontas Dimitropoulos, Philipp Richter,  
Alessandro Finamore, Alex C. Snoeren  
***alberto@caida.org***



UC San Diego

**ETH** zürich



# MAPPING USE OF IPV4 SPACE

*Why do we care?*

## **Security**

- inform host reputation and access control
  - e.g., to mitigate network abuse
- detect BGP hijacking attacks

## **Policy / Social / Economics**

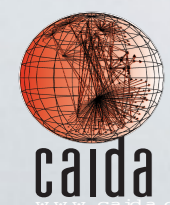
- inform policy on address space allocation
- estimate Internet usage over time
  - e.g., policy, political and social science, economics

## **Better Data Analysis**

- identify homogeneous address aggregates
  - e.g., for IP geolocation
- data normalization
  - e.g., per-AS or per-Country normalization

## **Operations**

- detect changes in network operation
- select targets for active measurements
  - e.g., traceroutes, vulnerability scans, and reachability surveys



# MAPPING USE OF IPV4 SPACE

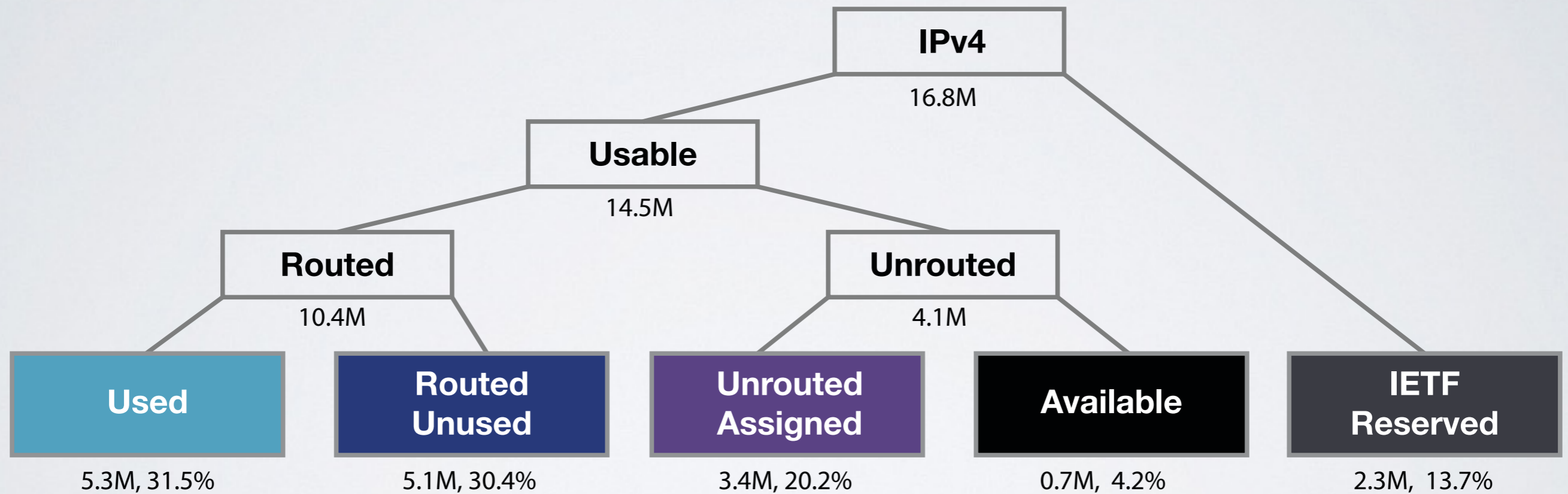
## methodology & 2013 datasets

Dataset	Source type	Data format	Period
UCSD-NT [2]	Traffic: Darknet	full pkt traces	July 23 to August 25, 2013
SWITCH [59]	Traffic: Live Academic Net.	Netflow logs	July 23 to August 25, 2013
IXP [8]	Traffic: IXP	sFlow packet samples	July 8 to July 28, August 12 to September 8, 2013
R-ISP [25]	Traffic: Residential ISP	Tstat [24] logs	July 1 to September 31, 2013
ISI [41]	Active Probing: ICMP ping	logs	July 23 to August 25, 2013
HTTP [28]	Active Probing: HTTP GET	logs	October 29, 2013
ARK-TTL [34]	Active Probing: traceroute	logs	July to September, 2013
BGP [6], [57]	BGP announcements	RIBs	July to September, 2013
Available Blocks [27]	IANA/RIRs	IP ranges	October 1, 2013
NetAcuity Edge [22]	IP Geolocation	IP ranges	July 2013
prefix2AS [16]	BGP announcements	prefix to ASN	July 2013

- active + passive measurement approaches
- passive: main challenge is filtering out *spoofed* and *scanning* traffic

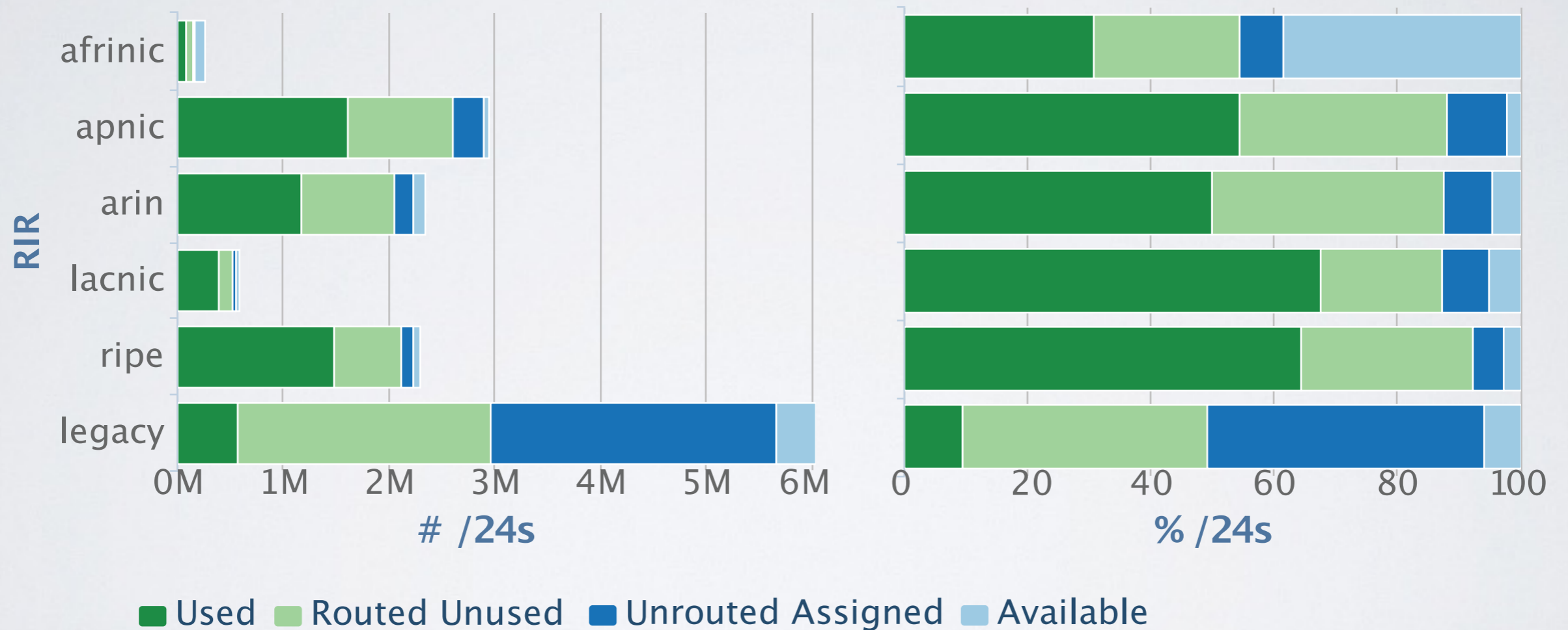
# A TAXONOMY OF SPACE USE

*announced on BGP does NOT imply it is used*



# A PEEK AT RESULTS

*paper/website are packed with graphs and tables*



# YOU ARE HERE



# CONTRIBUTE

*how you can help us*

- **data data data..**

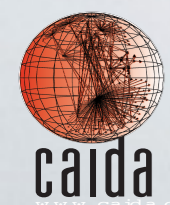
- ***anonymized, no timestamps, no content***

- just /24 blocks you observe (*NOT IPs of your network*) within a large time frame (e.g., *3 months or even a year*)

- 1) existing: server logs, NetFlow records...

- 2) collect: host a simple box running a modified Tstat. opensource code (*we don't need access to the machine*)

- we share our resulting dataset through the DHS' *Protected Repository for the Defense of Infrastructure Against Cyber Threats (PREDICT)* [www.predict.org](http://www.predict.org)



THANKS  
*questions?*

*alberto@caida.org*

