Yarp'ing the IPv6 Internet

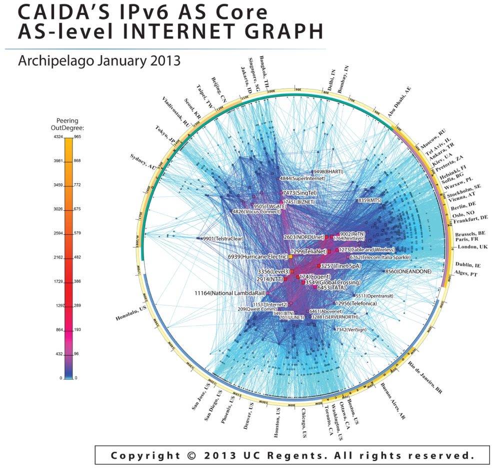
Eric Gaston
Robert Beverly

Naval Postgraduate School

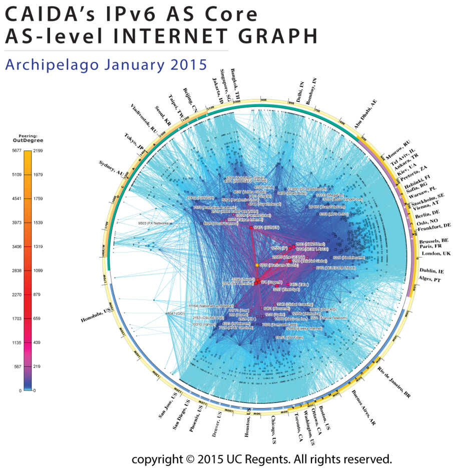
CMAND

AIMS 2017
March 2, 2017

IPv6 Active Topology Discovery


CAIDA'S IPv6 AS Core
AS-level INTERNET GRAPH
Archipelago January 2013

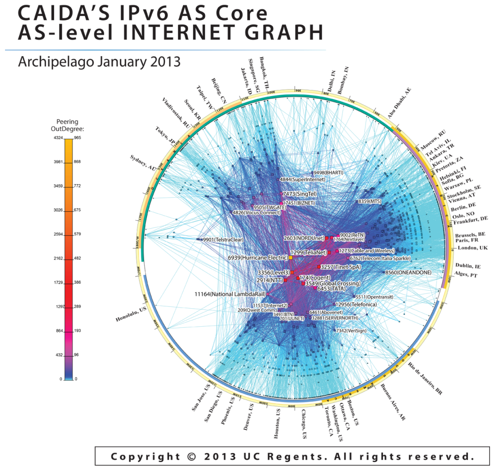
Copyright © 2013 UC Regents. All rights reserved.

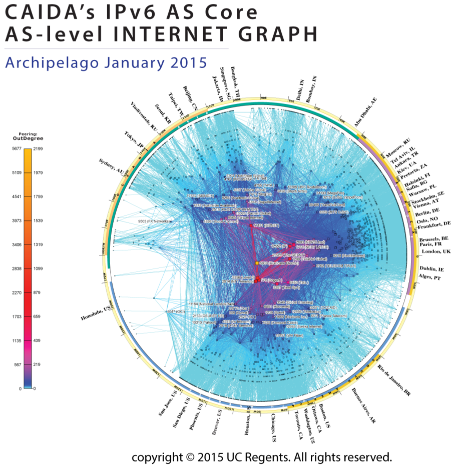

CAIDA's IPv6 AS Core
AS-level INTERNET GRAPH
Archipelago January 2015

copyright © 2015 UC Regents. All rights reserved.

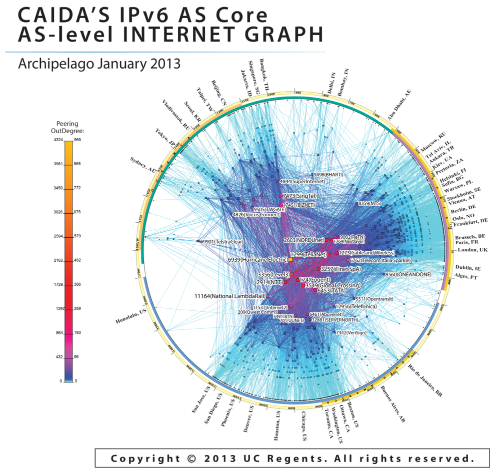
- Goal: Discover IPv6 Internet's interface-level topology
- But, completeness is a challenge with 2^{128} (~3.4 X $10^{38)}$ unique addresses
- And, rate limiting in IPv6 is more aggressive than in IPv4
- Current state-of-the-art: scan small number of prefixes slowly.

cMAND

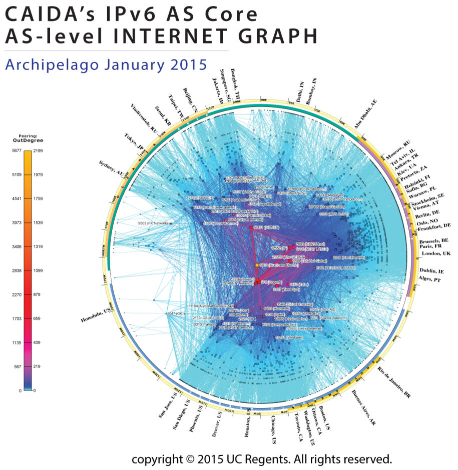




IPv6 Topology Mapping Today

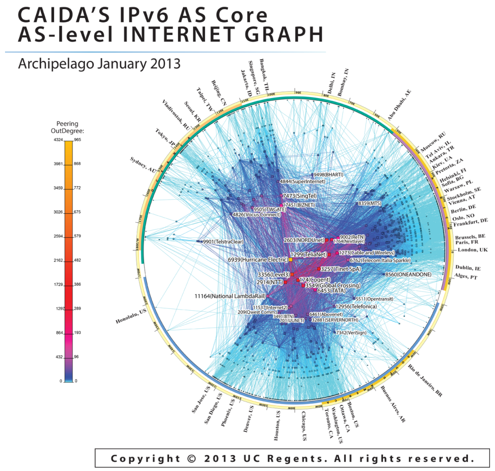


CAIDA IPv6 Topology Probing

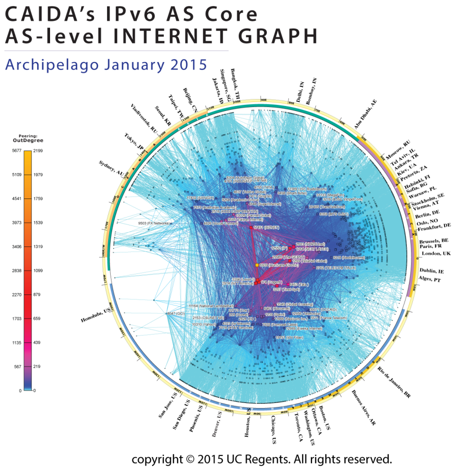


- Send probes toward each globally announced /48 or shorter prefix once every 48 hours
- 37,797 prefixes as of February 12, 2017
- From 46 globally distributed Ark VP
- Each VP scamper icmp-paris traceroutes toward ::1 and a random address in each prefixes.

CMAND

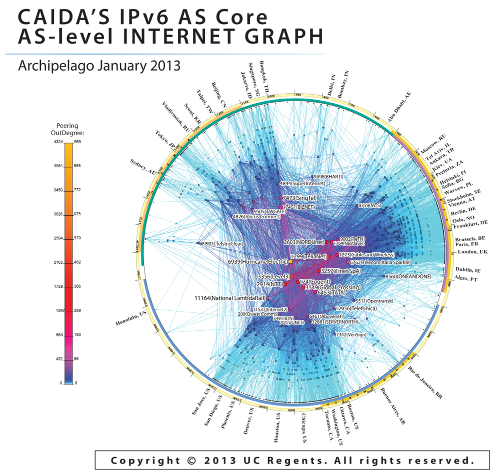


Rohrer et al: IPv6 Scans

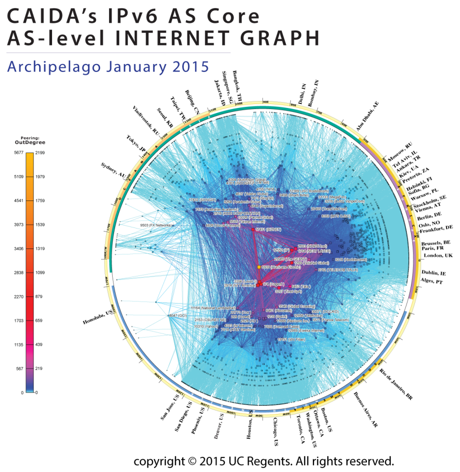


- Used Ark
- Largest scan to date probing ~406 million prefixes
- (Data publicly available)
- Traceroute to the ::1 in each /48 in all /32's
- Scan took 4 months to complete (Nov 14 – Mar 15)
- Current routing table contains ~536 million prefixes
- Increase of 32% in 2 years

CMAND

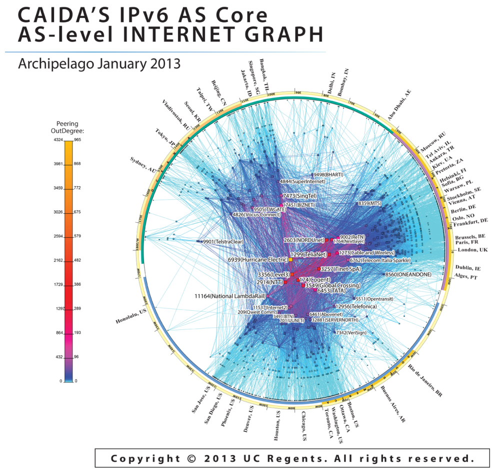


Foremski et al: Entropy/IP



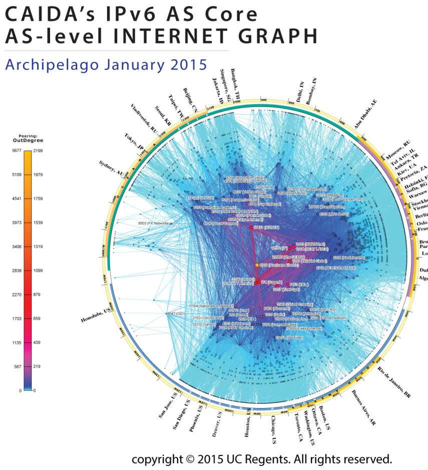
- IMC 2016 study to find active portions of IPv6 Internet
- Combines information theory and machine learning to probabilistically model IPv6 addresses
- Ability to generate candidate address list for active scanning can be used to reduce the target space

cMAND


CAIDA'S IPv6 AS Core AS-level INTERNET GRAPH

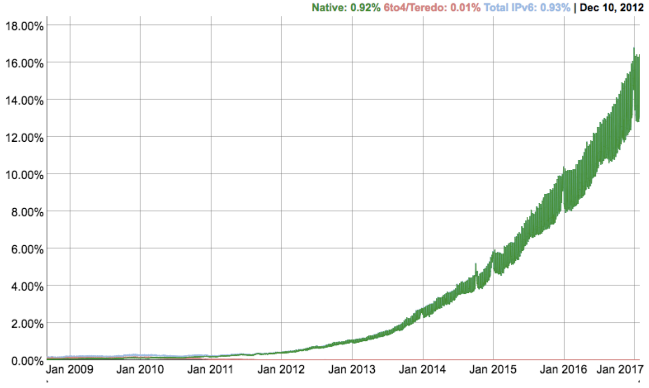
Archipelago January 2013

Why is mapping IPv6 Important?

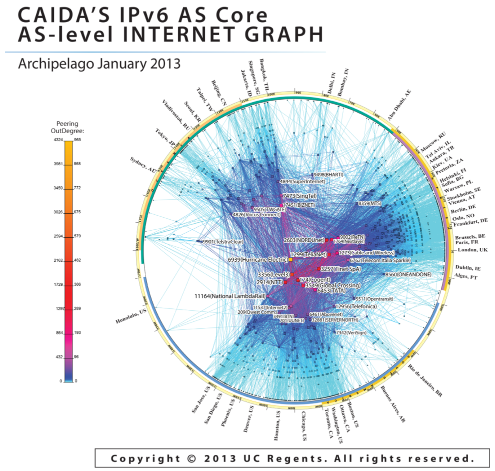

CAIDA's IPv6 AS Core AS-level INTERNET GRAPH

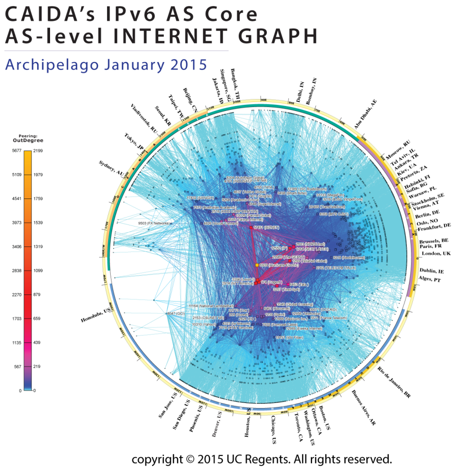
Archipelago January 2015

- IPv6 Topology mapping crucial to:
  - Security
  - Policy
  - Research
- IPv6 use has doubled every year since 2012
- Measurement community needs:
  - Better visibility into IPv6 topology
  - Better tools



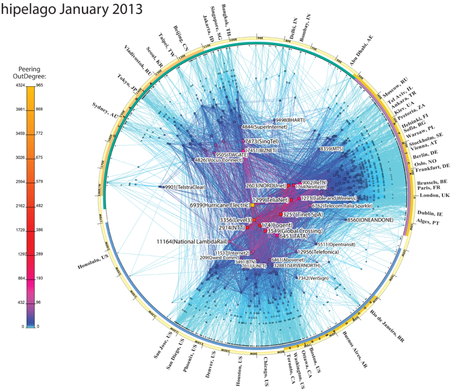
cMAND



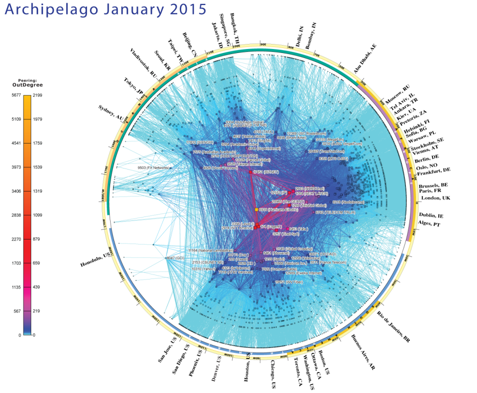


Our approach: Yarrp6

CAIDA'S IPv6 AS Core
AS-level INTERNET GRAPH
Archipelago January 2013



CAIDA's IPv6 AS Core
AS-level INTERNET GRAPH
Archipelago January 2015

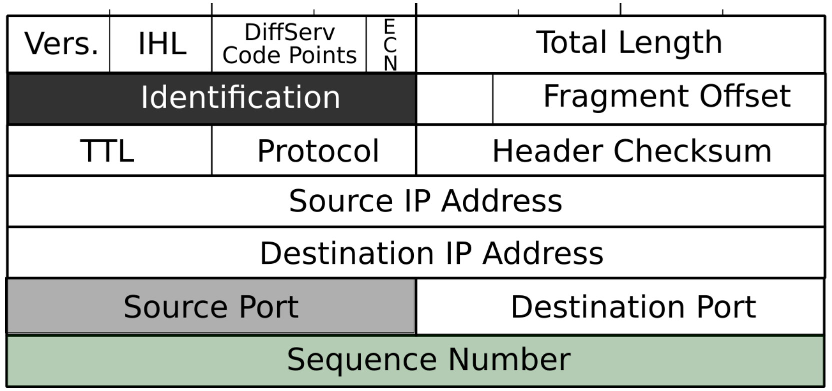


What is Yarrp?

https://www.cmand.org/yarrp/

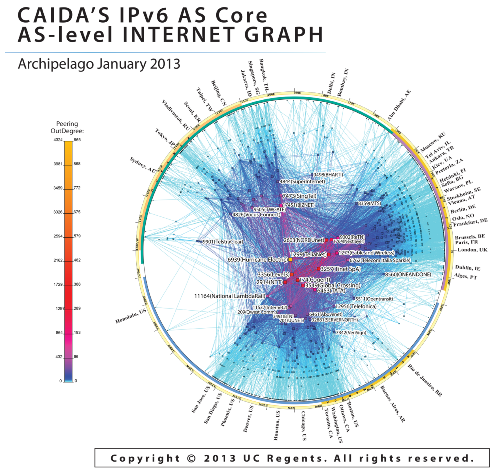
- A new <u>high-speed</u> <u>stateless</u> traceroute technique (IMC 2016 demonstrates topo discovery @100K pps)
- Reconstructs states from data encoded in IP and TCP headers of ICMP quotation
- Currently only supports IPv4 and TCP probes
- (Presently working w/ CAIDA to deploy in production)

cMAND



Vers.	IHL	DiffServ Code Points	ECN	Total Length	
Identification					Fragment Offset
TTL	Protocol			Header Checksum	
Source IP Address					
Destination IP Address					
Source Port				Destination Port	
Sequence Number					

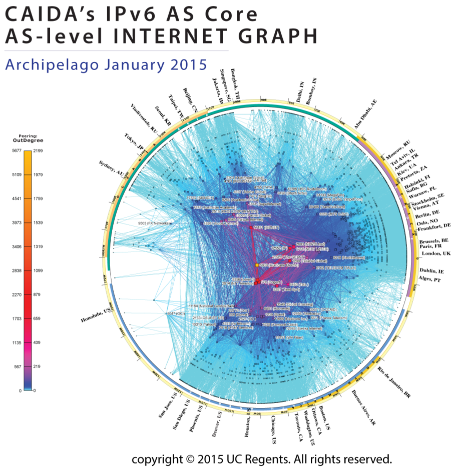
■ Send TTL
▨ cksum(IP Dest)
🟩 Send Elapsed Time



CAIDA'S IPv6 AS Core AS-level INTERNET GRAPH

Archipelago January 2013

What is Yarrp6?

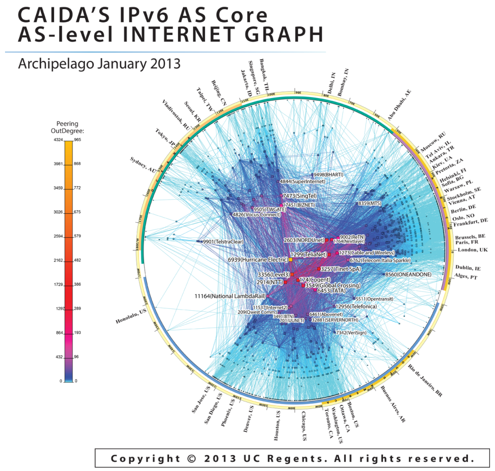


CAIDA's IPv6 AS Core AS-level INTERNET GRAPH

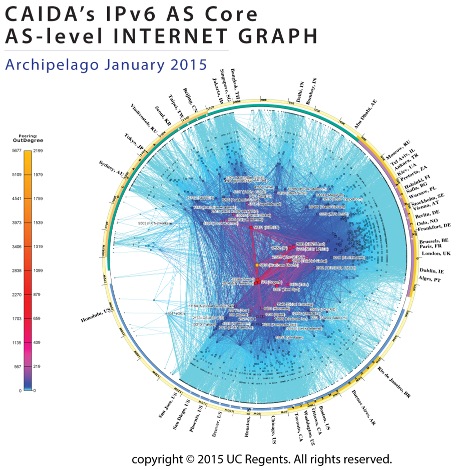
Archipelago January 2015

- Yarrp6 is a port of Yarrp for IPv6
- Also stateless and randomized
- But encodes state in a different manner
- Maintains Paris traceroute method for all scan
- Adds the capability to do ICMPv6 and UDP scans as well as the TCP SYN and TCP ACK provided by Yarrp

CMAND

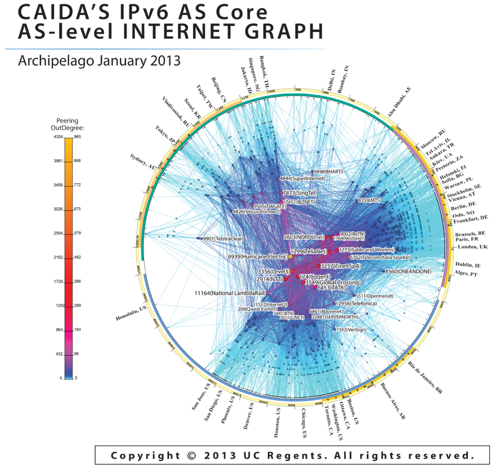


Porting Yarrp to IPv6

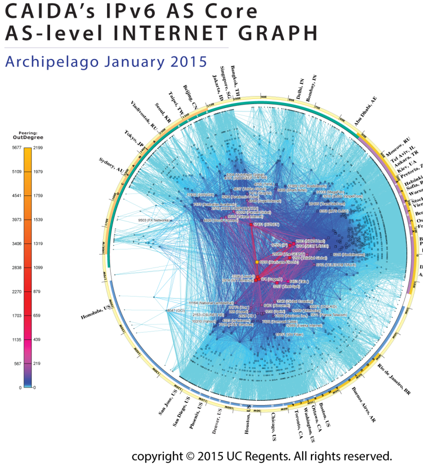


- Extending Yarrp to IPv6 is not a trivial task
- Issues:
  - How to encode state
  - Yarrp permutation library's 32-bit block size too small for IPv6
  - Raw sockets in IPv6 do not allow for full control of packet headers
  - Rate-Limiting of ICMPv6 error messages
  - Unable to detect responses to TCP probes from targets

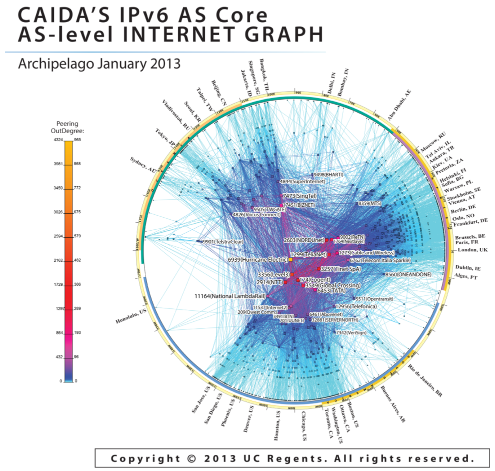
CMAND



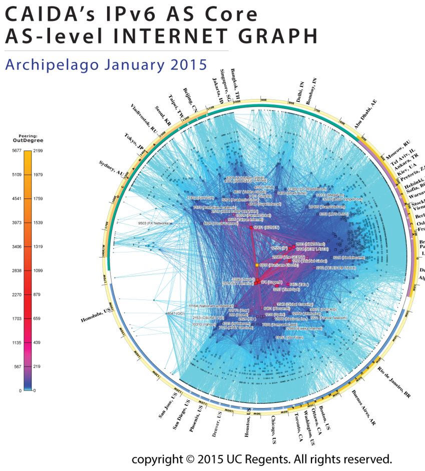
Initial Experiments

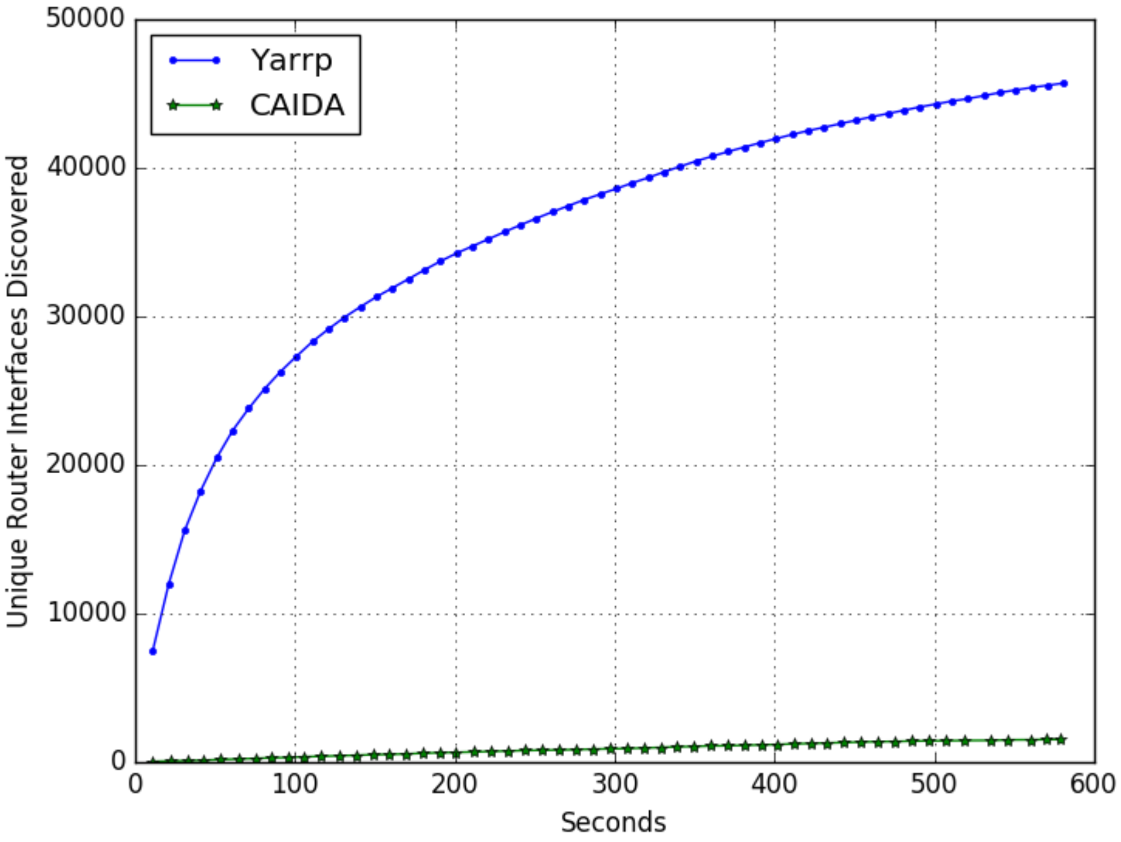


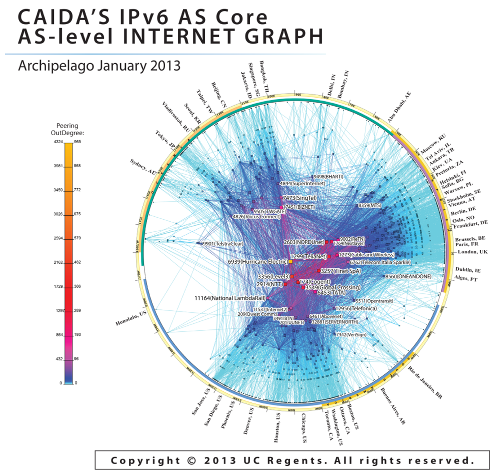
- Sought to validate and compare Yarrp to current state-of-the-art:
  - Recall of Yarrp6 vs. CAIDA v6 probe cycle
  - Speed of Yarrp6 vs. CAIDA v6 probe cycle
- Compared using CAIDA's IPv6 data from san-us VP scans done on February 12, 2017
- Same target list containing 75,594 addresses



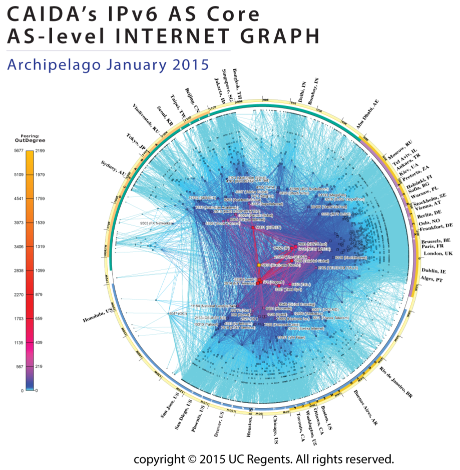
Yarrp6 vs. CAIDA (cont.)



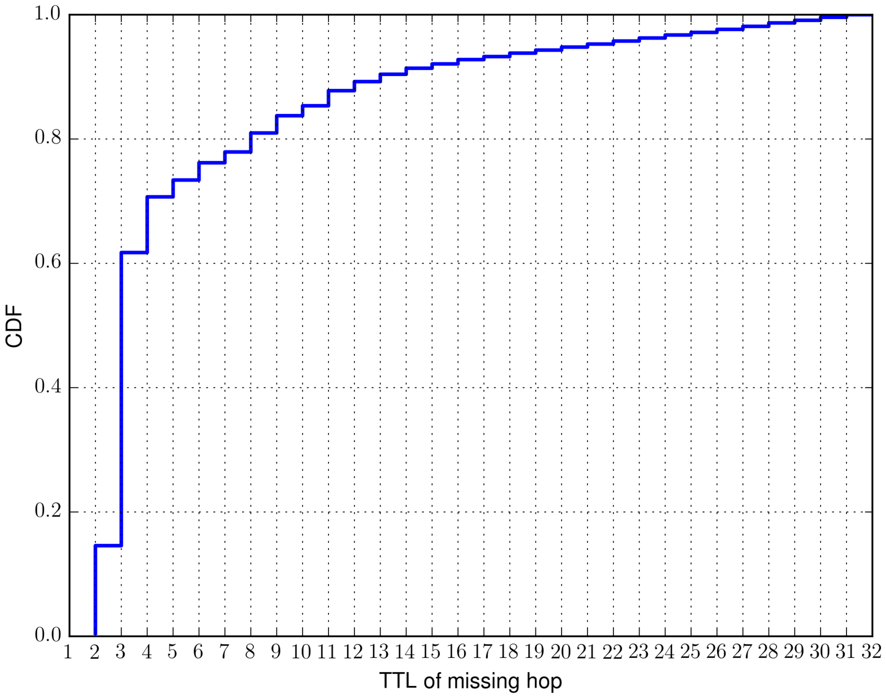




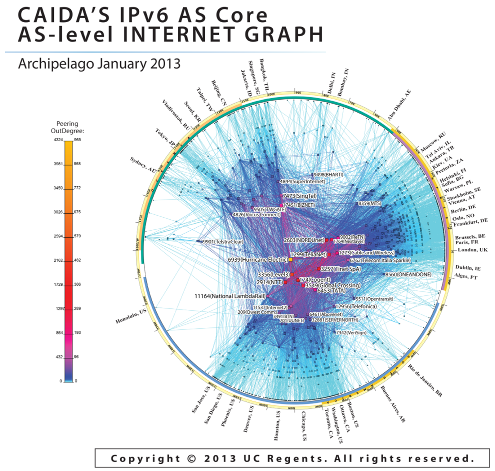
Rate Limiting of IPv6



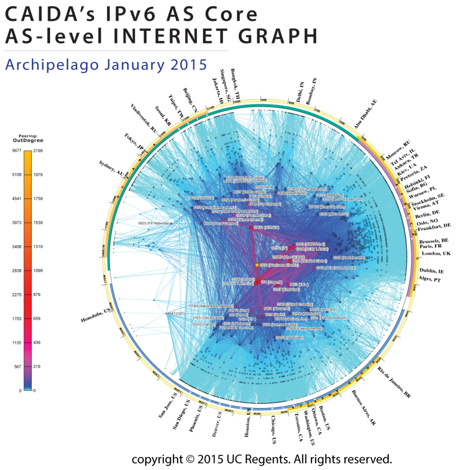
- "an IPv6 node MUST limit the rate of ICMPv6 error messages it originates." – RFC 4443
- We did observe rate-limiting on IPv6
- Hops 1-4 accounted for ~75% of all missing hops
- Only 57 unique addresses missing from these hop



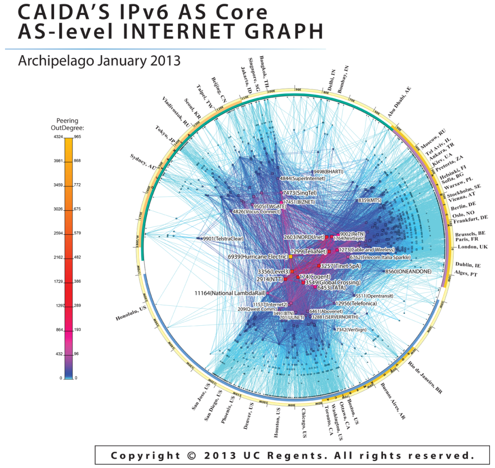
cMAND



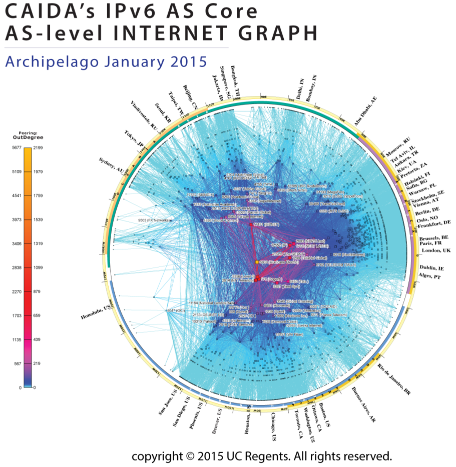
Comparison of Transport Protocols



- Used yarrp6 to compare probe protocol
- Comparison of Transport Protocol on forward IP path inference.
- Used ICMPv6, UDP, TCP SYN, and TCP ACK Paris traceroute probes
- 3 metrics used for comparison:
  - Destination Reached
  - Complete Paths
  - Unique IP Links



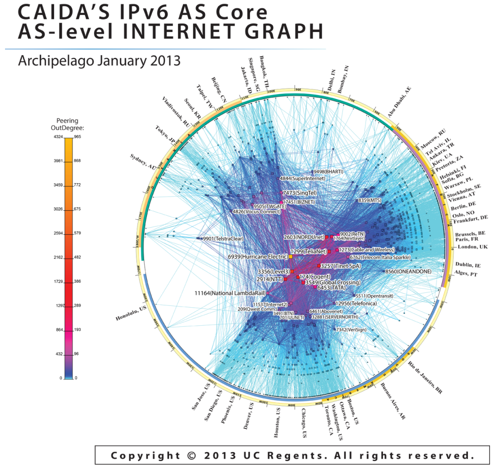
Comparison of Transport Protocols (cont.)



Probe Method	Unique Interface	Destinations Reached	Complete IP Paths	Unique IP Links
ICMPv6	45,706	9,535	3,562*	57,667
UDP	34,567	4,455	1,776*	37,514
TCP SYN	34,879	N/A#	N/A#	37,655
TCP ACK	35,178	N/A#	N/A#	38,262

* Hop 3 skipped in determination of complete path
# Unable to retrieve encoded information from TCP responses

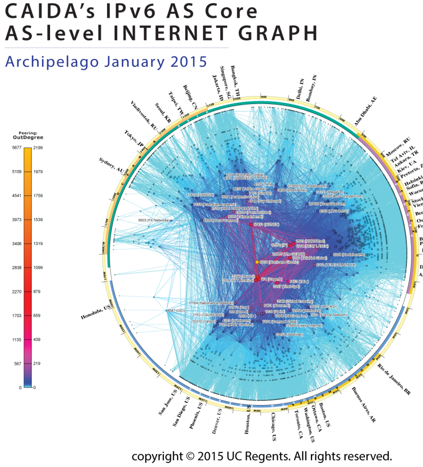
CMAND



CAIDA'S IPv6 AS Core AS-level INTERNET GRAPH

Archipelago January 2013

Future Work

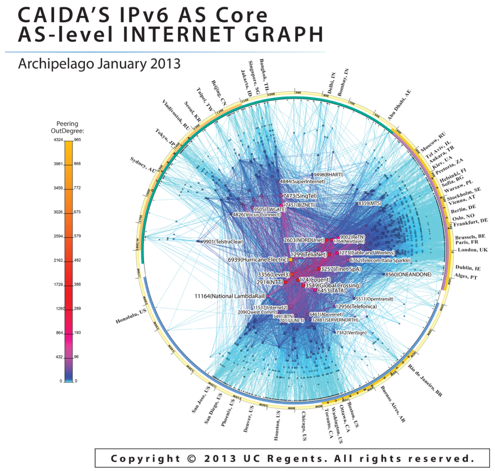


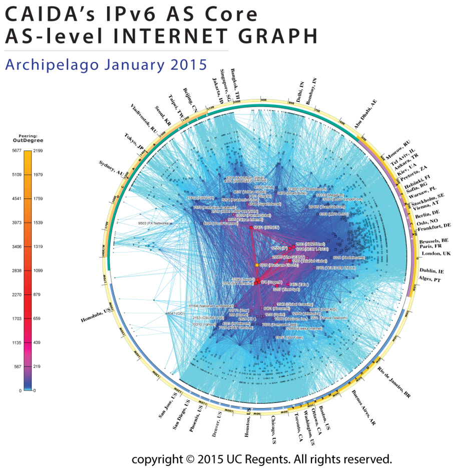
CAIDA's IPv6 AS Core AS-level INTERNET GRAPH

Archipelago January 2015

- Working w/ Dave Plonka: Use Entropy/IP to generate target list for Yarrp6 to scan.
- Comparison of Yarrp6 to larger dataset such as Rohrer et al. dataset
- Running scans in rapid succession to allow for study into dynamics of IPv6 Internet.
- Yarrp available now; Yarrp6 real soon now. Contact us to beta!

https://www.cmand.org/yarrp/





Questions?

cMAND

https://www.cmand.org/yarrp/