Henya

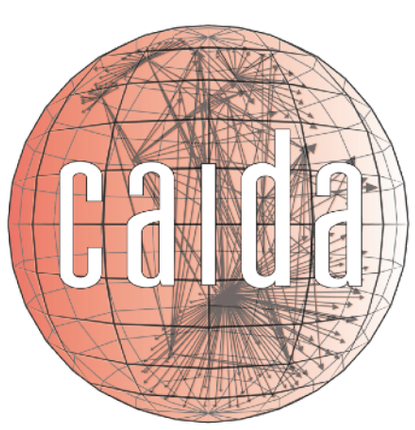
Large-Scale Internet Topology Query System

Young Hyun

CAIDA

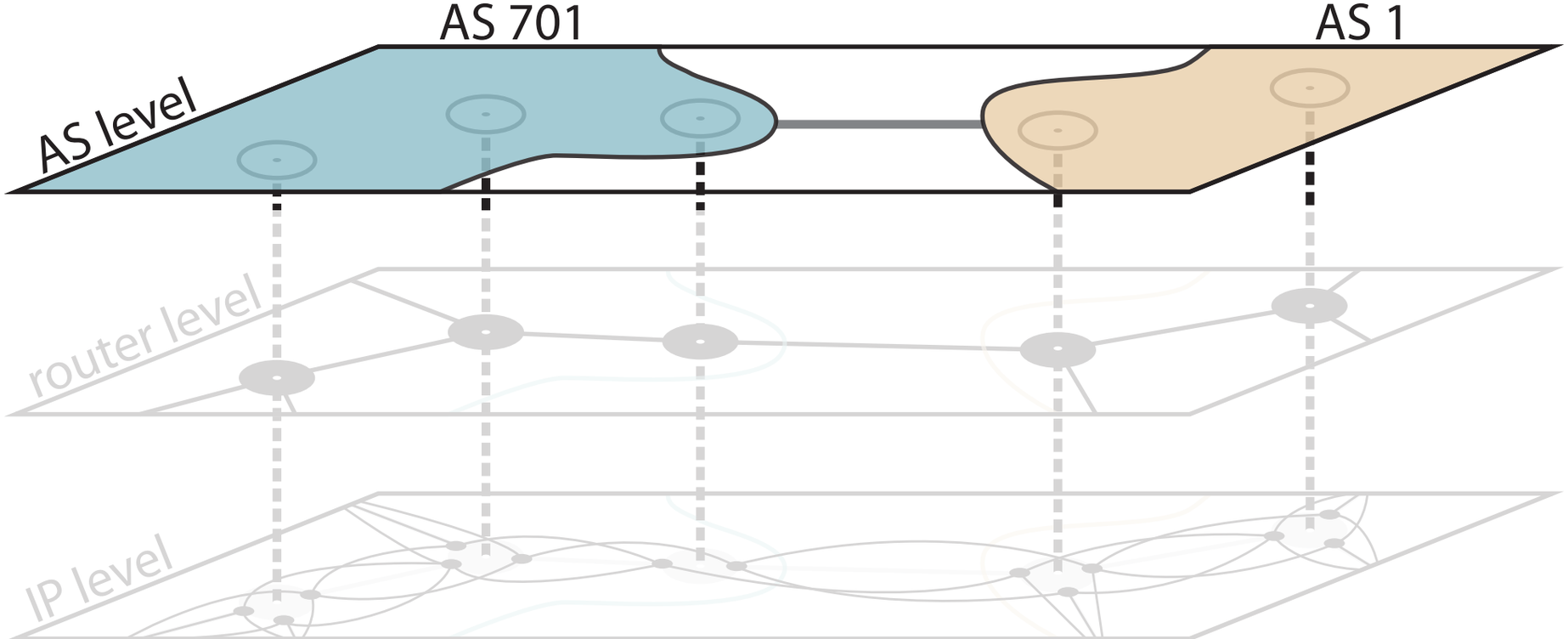
SDSC/UCSD

Nov 2016



layered view of the Internet

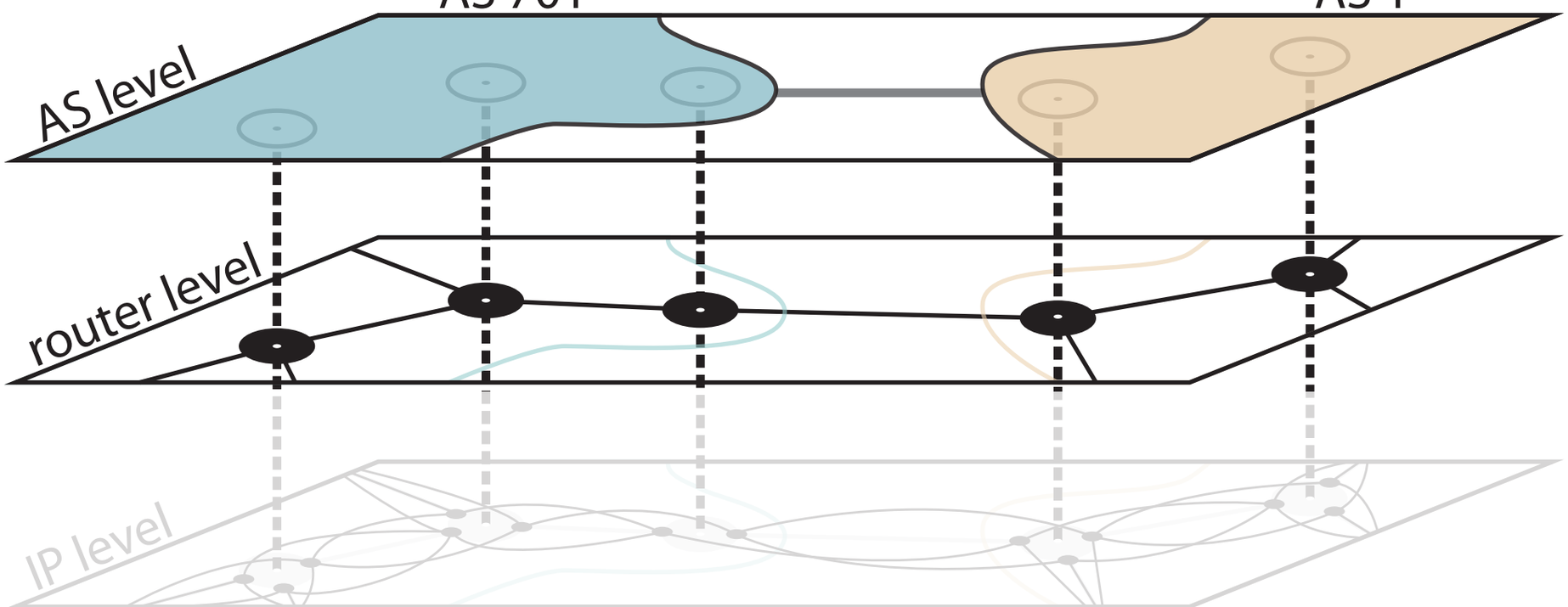
autonomous systems (AS) – network providers



layered view of the Internet

autonomous systems (AS) – network providers

routers – connected by network links

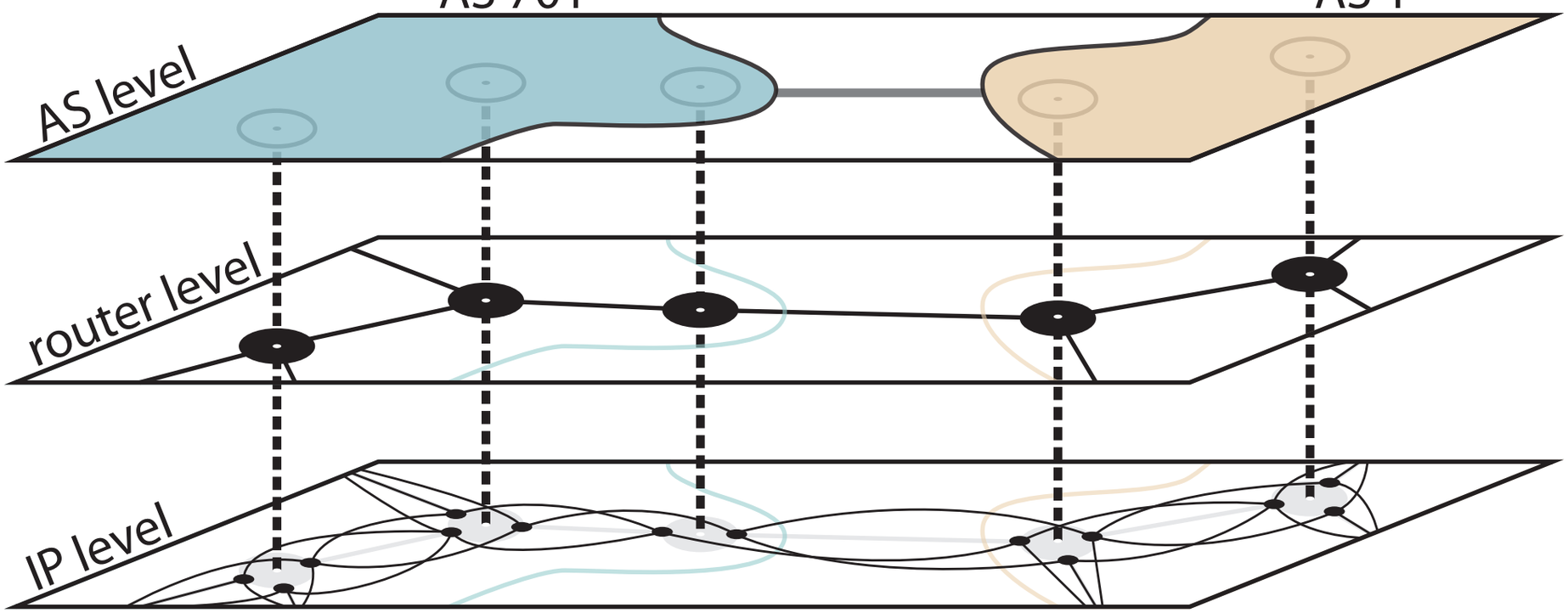


layered view of the Internet

autonomous systems (AS) – network service providers

routers – connected by network links

IP addresses – network interfaces on routers

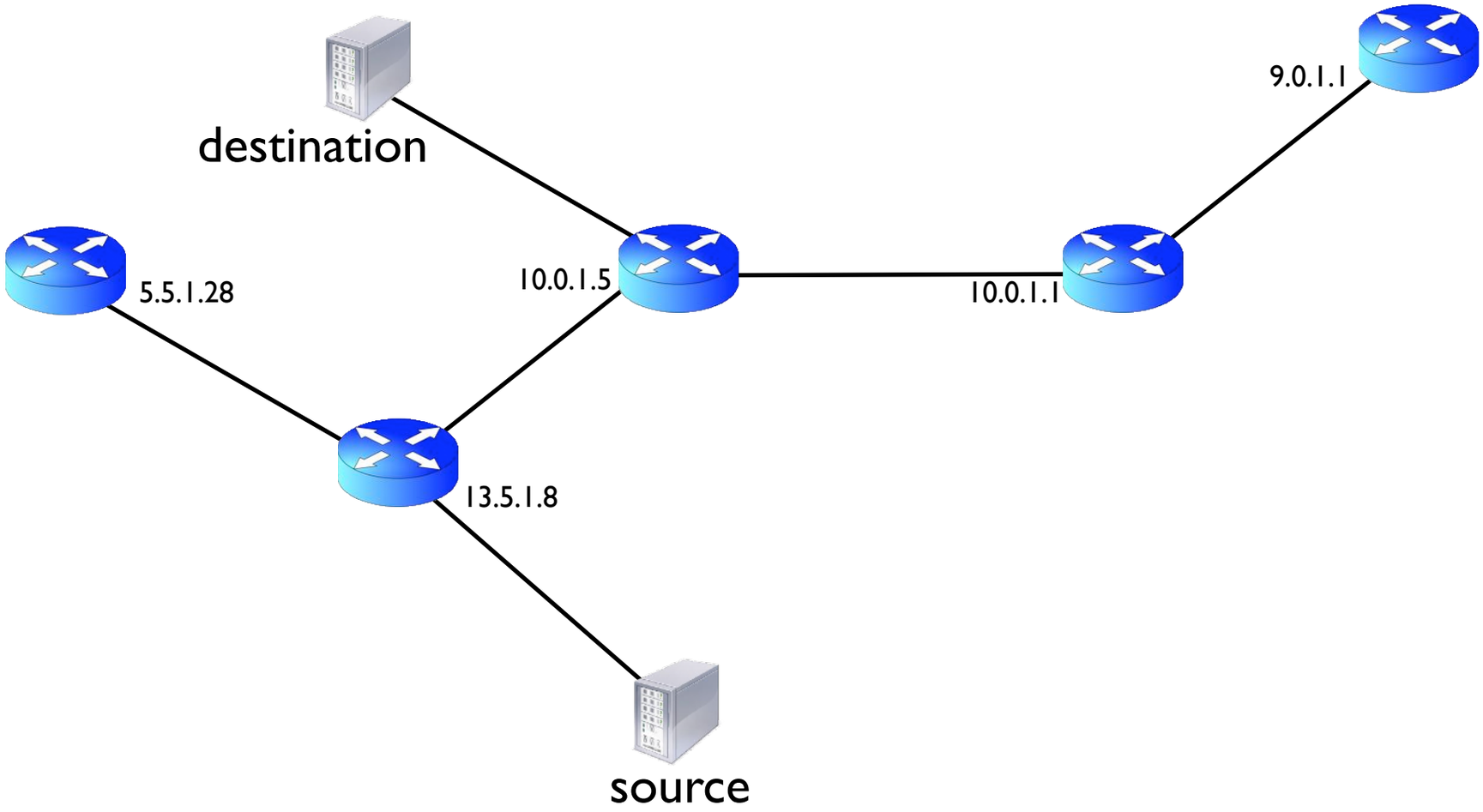


traceroute paths

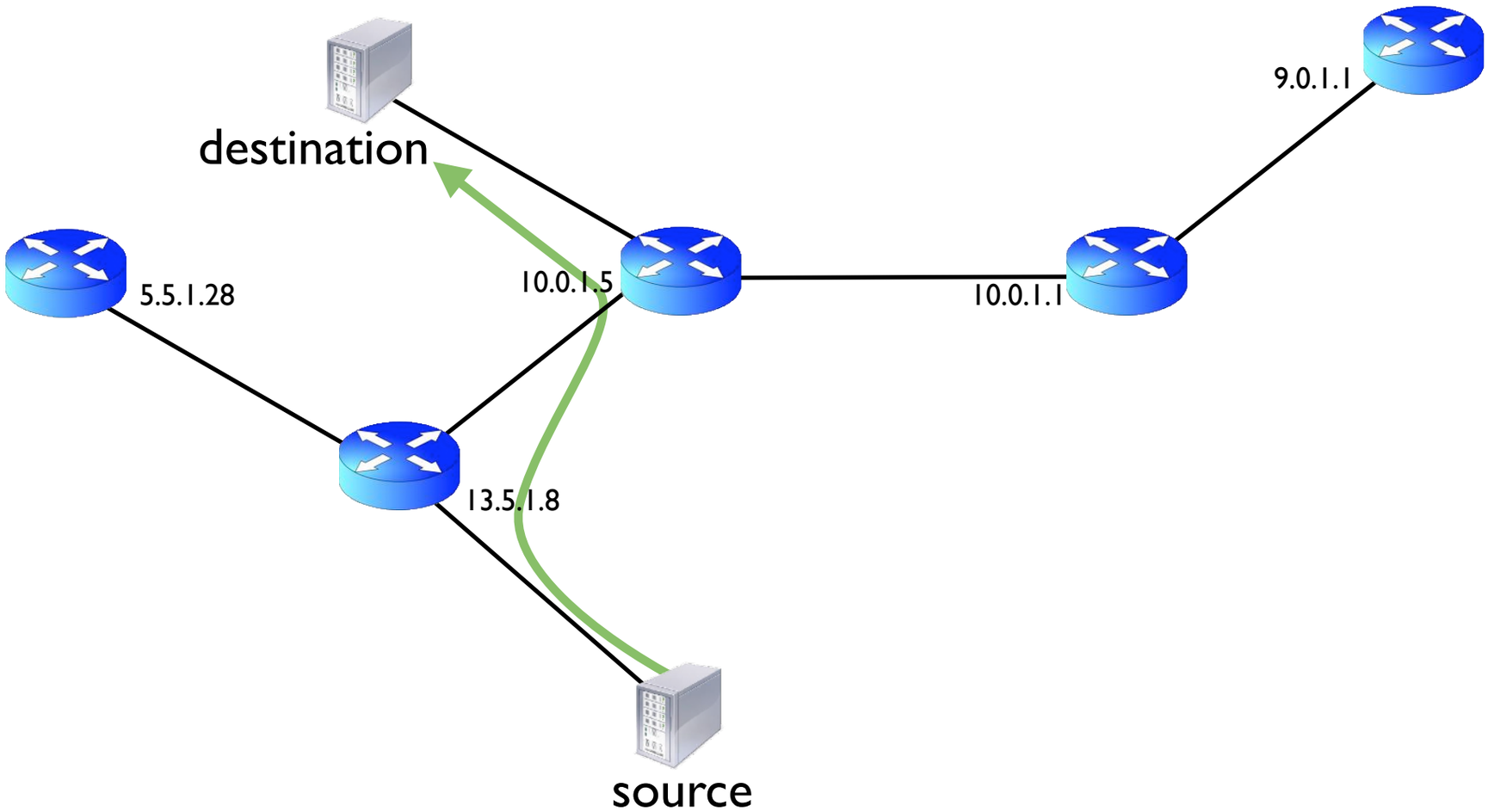
destination

source

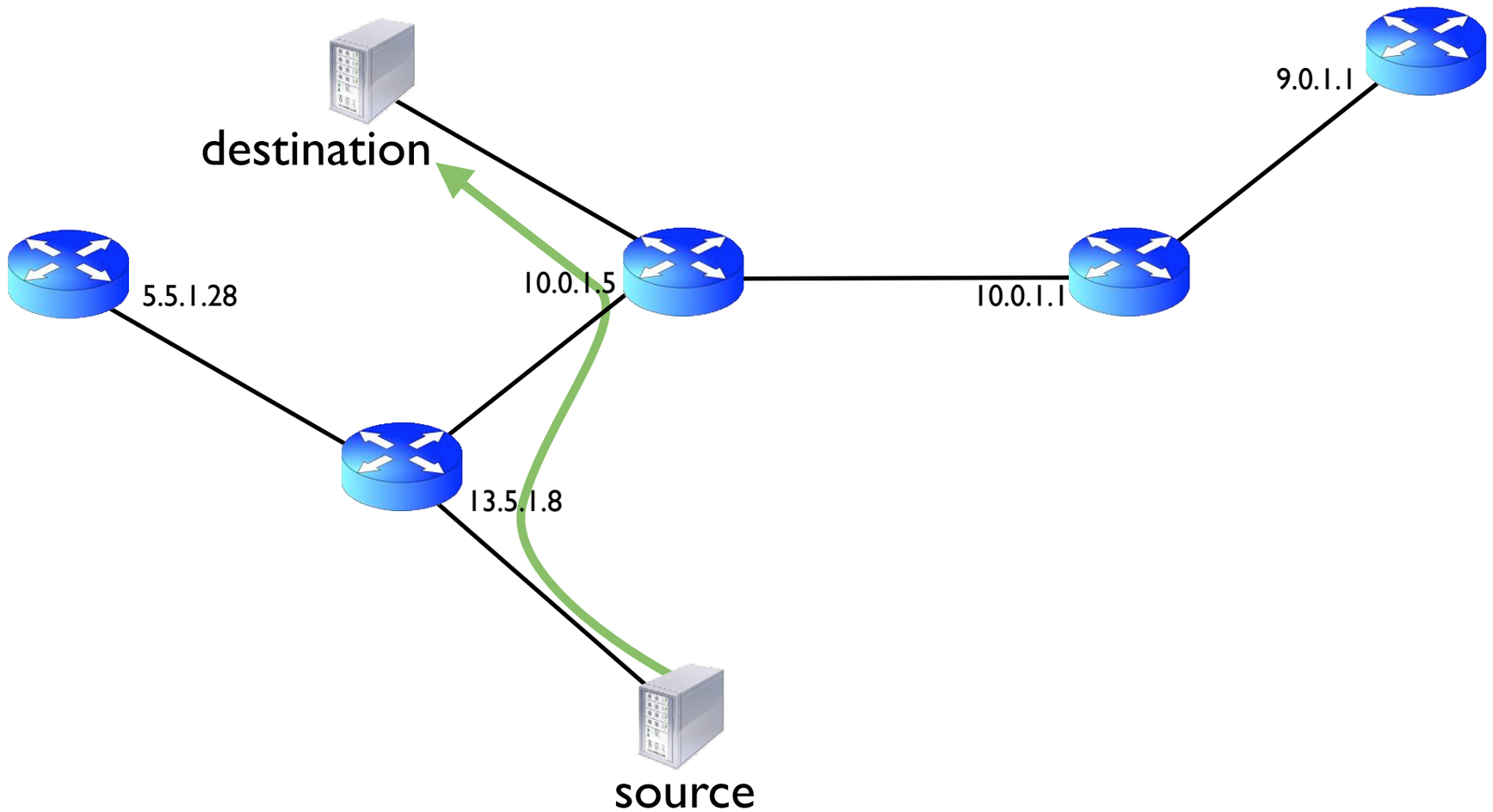
traceroute paths



traceroute paths

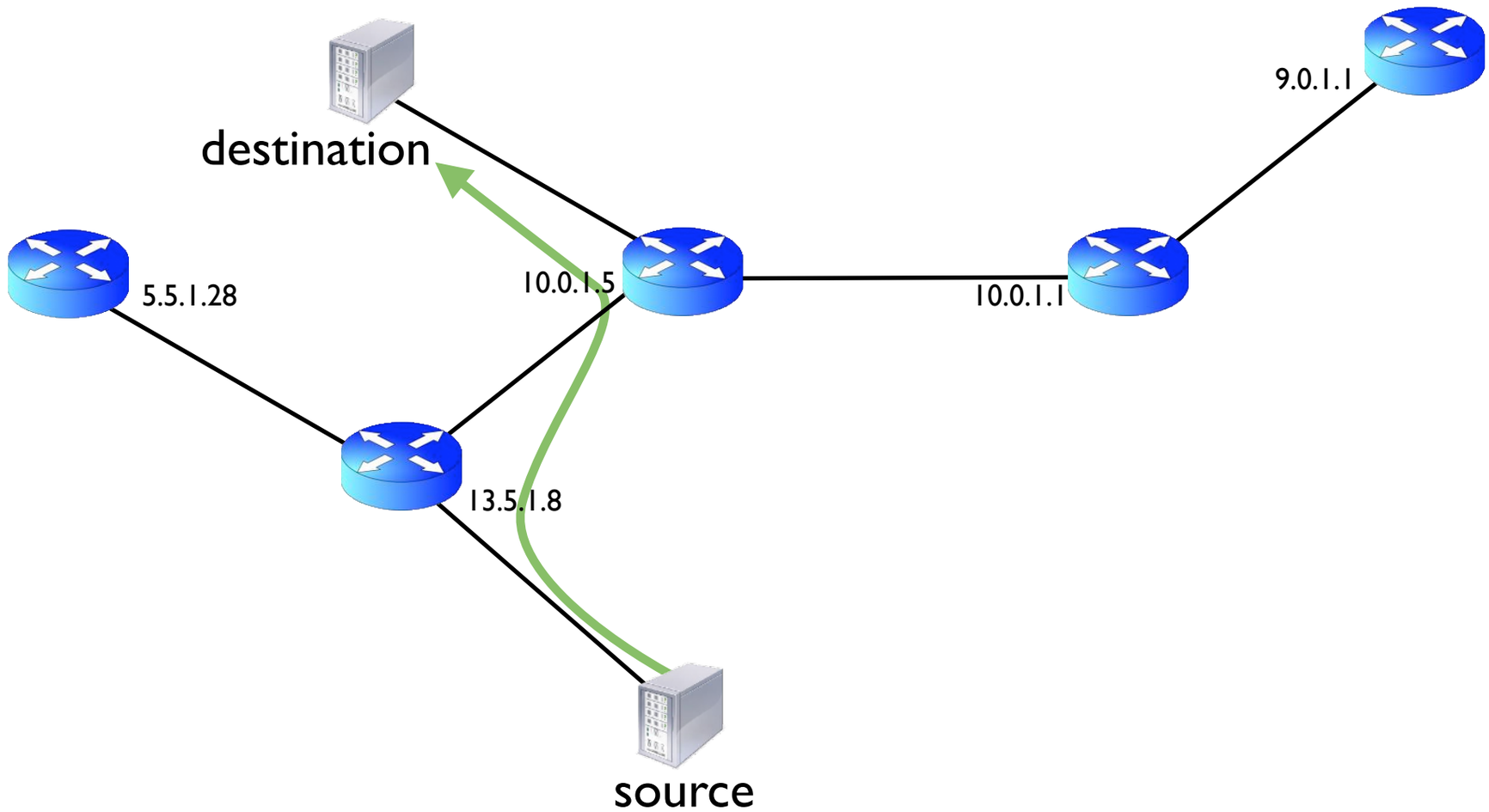


traceroute paths



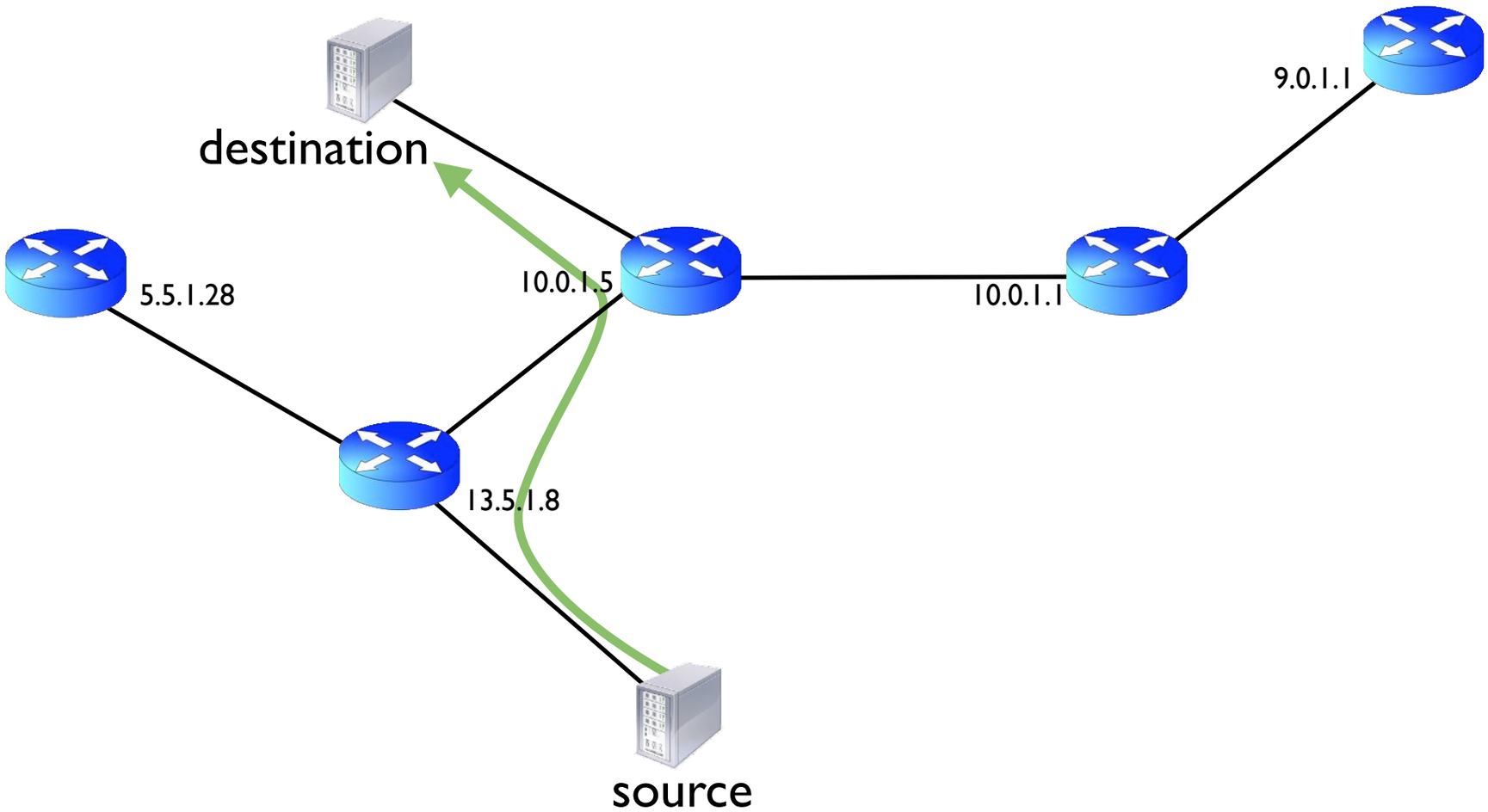
source → 13.5.1.8

traceroute paths



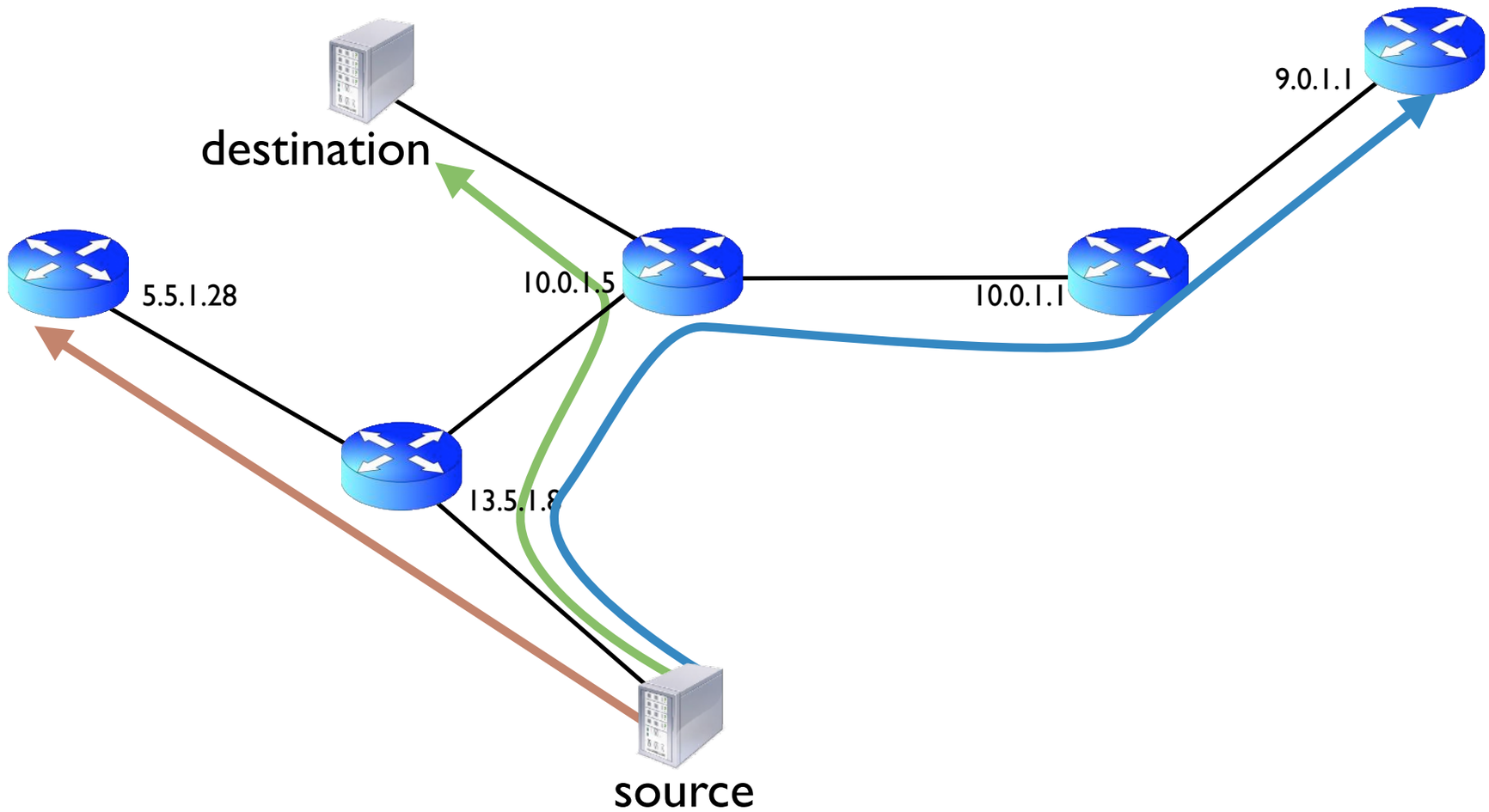
source → 13.5.1.8 → 10.0.1.5

traceroute paths



source → 13.5.1.8 → 10.0.1.5 → destination

traceroute paths



source → 13.5.1.8 → 5.5.1.28

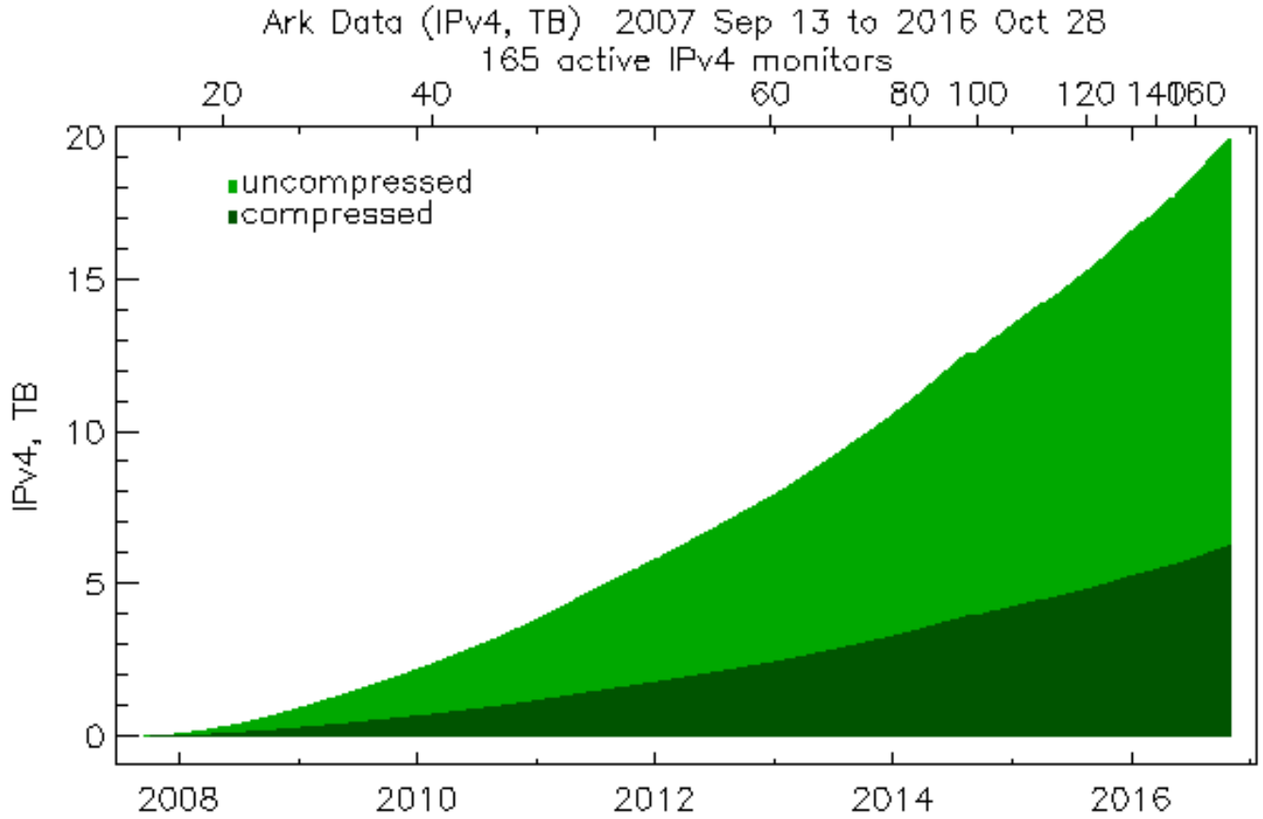
source → 13.5.1.8 → 10.0.1.5 → destination

source → 13.5.1.8 → 10.0.1.5 → 10.0.1.1 → 9.0.1.1

Ark monitors

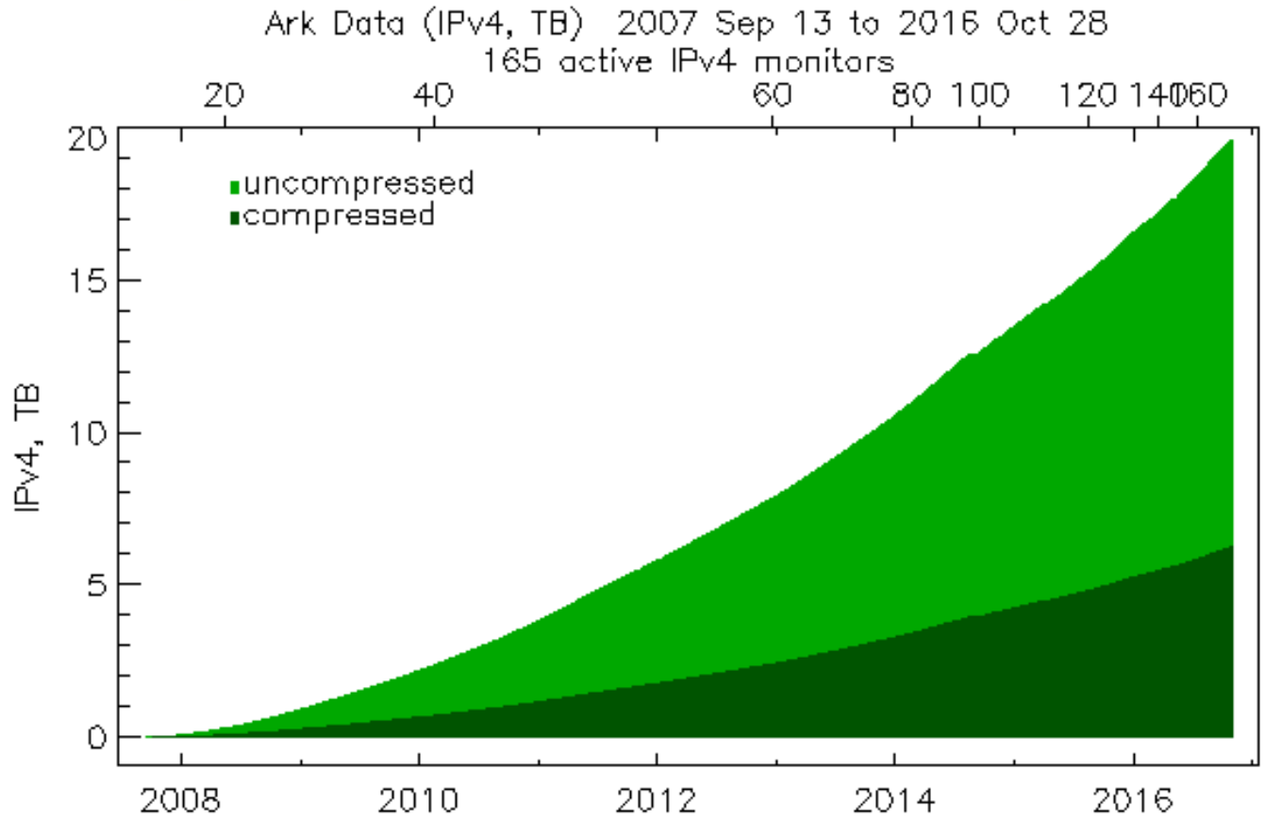
- Ark launched in Sep 2007 with 8 monitors
- now at 165 monitors in 57 countries

data growth



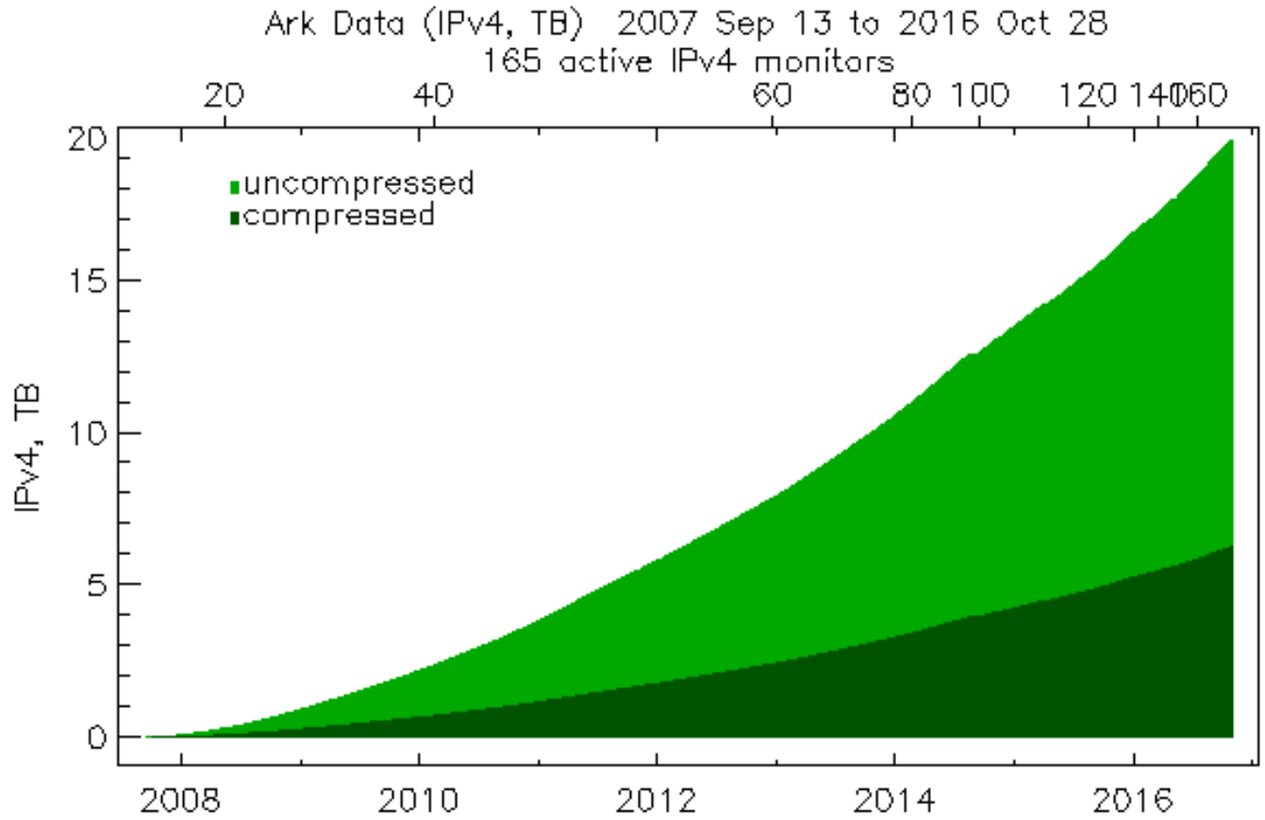
- 9 years of "Routed /24" traceroutes
  - 47 billion traces in 20 TB of files
  - growing yearly by 10 billion traces

data growth



- 9 years of "Routed /24" traceroutes
  - 47 billion traces in 20 TB of files
  - growing yearly by 10 billion traces
- 1 year of "Prefix Probing" traceroutes
  - growing yearly by 9 billion traces

data growth



- 9 years of "Routed /24" traceroutes
  - 47 billion traces in 20 TB of files
  - growing yearly by 10 billion traces
- 1 year of "Prefix Probing" traceroutes
  - growing yearly by 9 billion traces
- combined dataset growing by **19 billions traces/year**

traceroute uses

Ukraine Emerges as Bogus Routing Source

MARCH 14, 2016 DOUG MADORY

Last fall, the Interior Minister of Ukraine announced the creation of a national Cyberpolice (Кіберполіцію) to protect the country from everything from credit card fraud to malware. Here's something that would be great to add to their list: fraudulent BGP routing out of Ukraine. Last year, we reported on an incident in which Ukrainian ISP Vega hijacked routes from British Telecom (including that of the UK's Atomic Weapons Establishment), an event that could *perhaps* be chalked up to an innocent mistake. However, the fraudulent routing we're now seeing from Ukraine is deliberately designed to go unnoticed. We'll review some of this new behavior in this blog.

Governments take note

The profile of this issue has grown in the past year as governments have had to respond to their address space being fraudulently used. Last July, the Dutch Minister of Foreign Affairs (pictured right) was confronted with parliamentary questions concerning an incident where "attackers" had commandeered IP address space belonging to the Ministry of Foreign Affairs the previous year. In that incident, on 18 November 2014, Decision Marketing (AS62228) out of Sofia, Bulgaria began globally announcing eleven BGP routes that did not belong to them.

traceroute uses

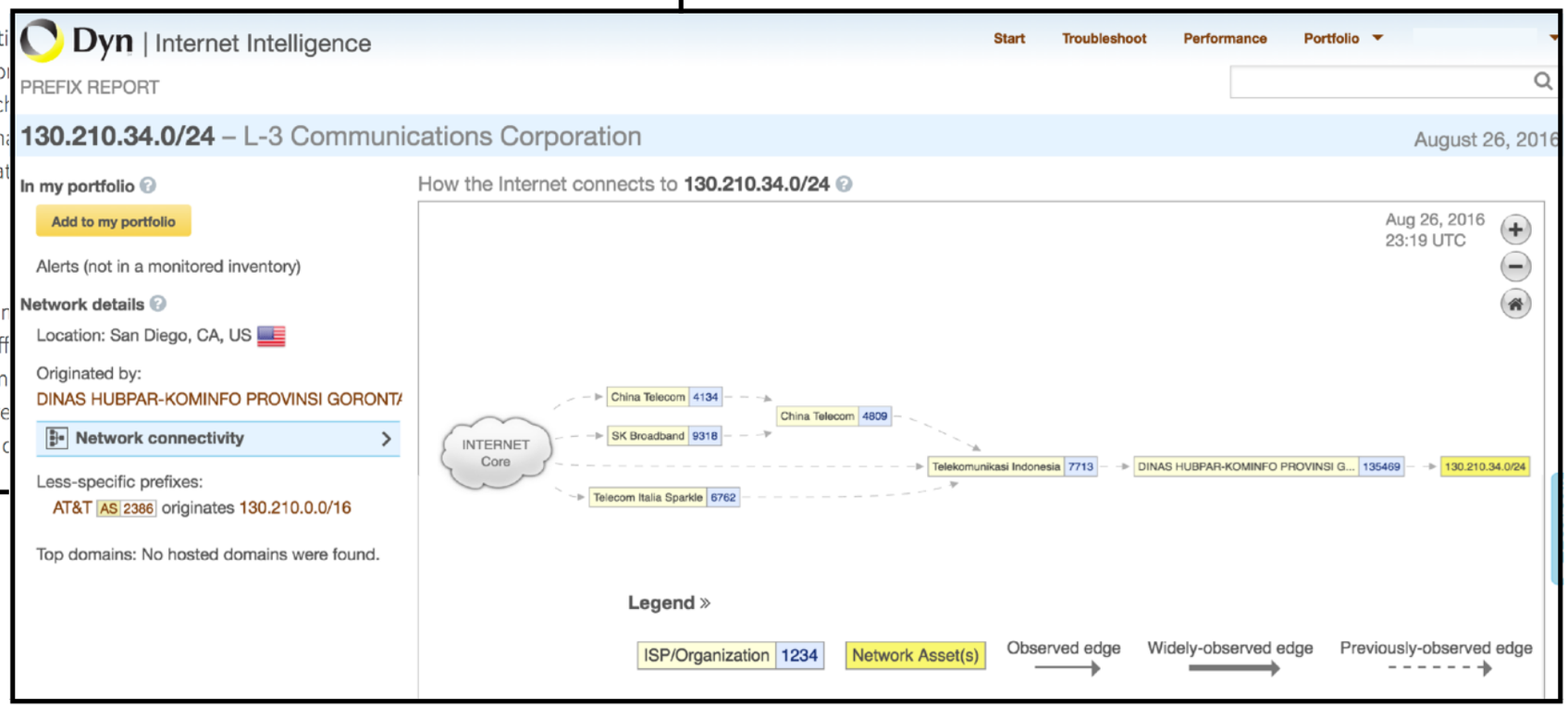
Ukraine Emerges as Bogus Routing Source

MARCH 14, 2016 DOUG MADORY

Last fall, the Interior Minister of Ukraine announced the creati
from everything from credit card fraud to malware. Here's so
out of Ukraine. Last year, we reported on an incident in which
that of the UK's Atomic Weapons Establishment), an event tha
fraudulent routing we're now seeing from Ukraine is deliberat
in this blog.

Governments take note

The profile of this issue has grown in the past year as governr
fraudulently used. Last July, the Dutch Minister of Foreign Aff
questions concerning an incident where "attackers" had com
Foreign Affairs the previous year. In that incident, on 18 Nove
Bulgaria began globally announcing eleven BGP routes that c



traceroute uses

Ukraine Emerges as Bogus Routing Source

MARCH 14, 2016 DOUG MADORY

Last fall, the Interior Minister of Ukraine announced the creati
from everything from credit card fraud to malware. Here's so
out of Ukraine. Last year, we reported on an incident in which
that of the UK's Atomic Weapons Establishment), an event tha
fraudulent routing we're now seeing from Ukraine is deliberat
in this blog.

Govern

The pro
fraudule
questio
Foreign
Bulgaria

Dyn | Internet Intelligence

Start Troubleshoot Performance Portfolio

PREFIX REPORT

130.210.34.0/24 – L-3 Communications Corporation

August 26, 2016

In my portfolio

How the Internet connects to **130.210.34.0/24**

Aug 26, 2016
23:19 UTC

DINAS HUBPAR-KOMINFO PROVINSI G... 135469 130.210.34.0/24

edge Widely-observed edge Previously-observed edge

Jürgen Jaritsch

July 16, 2015 at 11:16 PM

To: nanog@nanog.org

Details

Hi,

does anyone else see some prefix hijacks from AS7514? They started to announce some of our /24 ....

Thanks & best regards

Jürgen Jaritsch
Head of Network & Infrastructure

ANEXIA Internetdienstleistungs GmbH

traceroute uses

Ukraine Emerges as Bogus Routing Source

MARCH 14, 2016 DOUG MADORY

Last fall, the Interior Minister of Ukraine announced the creati
from everything from credit card fraud to malware. Here's so
out of Ukraine. Last year, we reported on an incident in which
that of the UK's Atomic Weapons Establishment), an event tha
fraudulent routing we're now seeing from Ukraine is deliberat
in this blog.

Govern

The pro
fraudule
questio
Foreign
Bulgaria

Dyn | Internet Intelligence

Start Troubleshoot Performance Portfolio

PREFIX REPORT

130.210.34.0/24 – L-3 Communications Corporation

August 26, 2016

In my portfolio

How the Internet connects to **130.210.34.0/24**

Aug 26, 2016
23:19 UTC

Jürgen Jaritsch

To: nanog@nanog.org

July 16, 2015 at 11:16 PM

Details

Hi,

does anyone else

Thanks & best rega

Jürgen Jaritsch
Head of Network &

ANEXIA Internetdie

Ronald F. Guilmette nanog November 11, 2016 at 3:50 AM

AS37135, AS6560, AS32714, AS14029 - Squatted or not? You be the judge.

To: nanog@nanog.org

At least one person has now asserted to me in private email that
my suggestion that AS30186 was being squatted on was in fact accurate.
Thus, I now feel confident enough to provide here the rest of the story
which goes along with that.

In a nutshell, AS30186 and also two other ASNs, together appear to all
be parts of a single large multi-ASN squat.

Henya

- CAIDA's large-scale topology query system
- provides **remote search** of traceroute data without requiring data downloads

Henya

- CAIDA's large-scale topology query system
- provides **remote search** of traceroute data without requiring data downloads
- built-in **analyses and visualizations**
  - for commonly occurring needs

Henya

- CAIDA's large-scale topology query system
- provides **remote search** of traceroute data without requiring data downloads
- built-in **analyses and visualizations**
  - for commonly occurring needs
- **responsive** enough for interactive data exploration
  - goal: query latency of 30 seconds or less

topology queries

- find occurrences of traceroute path elements
- ⟪*targets*⟫ = IP addreses, prefixes, ASes, or countries
- queries:
  - traceroutes **toward** ⟪*targets*⟫
  - traceroutes **containing** one or more ⟪*targets*⟫

topology queries

- find occurrences of traceroute path elements
- ⟪*targets*⟫ = IP addreses, prefixes, ASes, or countries
- queries:
  - traceroutes **toward** ⟪*targets*⟫
  - traceroutes **containing** one or more ⟪*targets*⟫
- parameters:
  - measurement vantage points
  - data collection time periods
  - position of ⟪*targets*⟫ in path
  - hop distance between sets of ⟪*targets*⟫

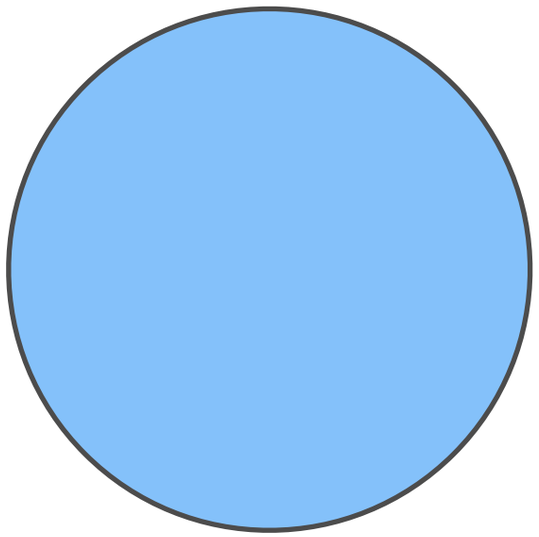
query complexity

- the most complex case:
  - traceroutes containing **two or more** ⟪*targets*⟫
    - precisely: traceroutes containing some hop $h_1 \in$ ⟪*targets*$_1$⟫, $h_2 \in$ ⟪*targets*$_2$⟫, $\cdots$
  - example: traceroutes containing hops in both ⟪Germany⟫ and ⟪Japan⟫

query complexity

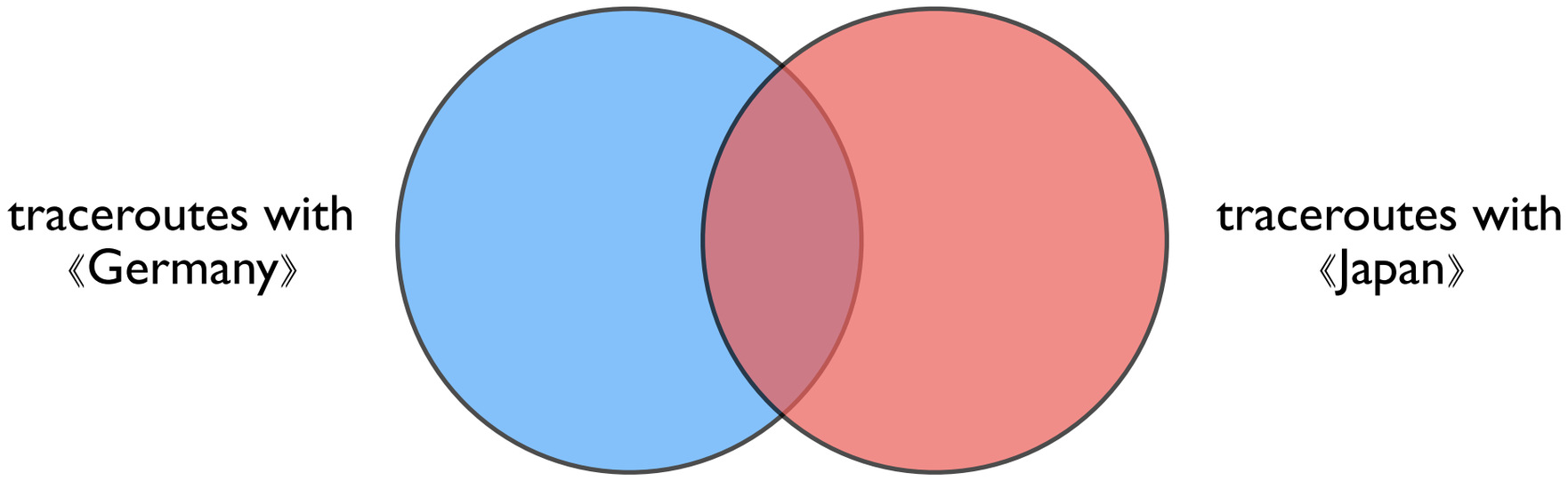
- the most complex case:
  - traceroutes containing **two or more** ⟪*targets*⟫
    - precisely: traceroutes containing some hop $h_1 \in$ ⟪*targets*$_1$⟫, $h_2 \in$ ⟪*targets*$_2$⟫, $\cdots$
  - example: traceroutes containing hops in both ⟪Germany⟫ and ⟪Japan⟫

traceroutes with ⟪Germany⟫



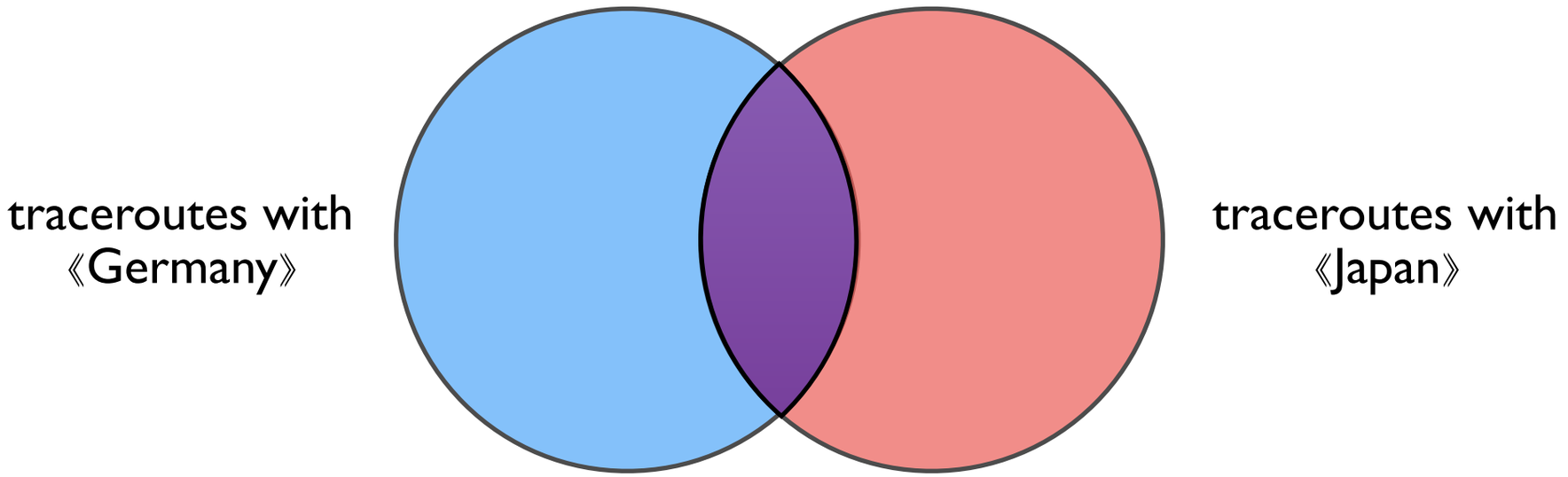
query complexity

- the most complex case:
  - traceroutes containing **two or more** ⟪*targets*⟫
    - precisely: traceroutes containing some hop $h_1 \in$ ⟪*targets*$_1$⟫, $h_2 \in$ ⟪*targets*$_2$⟫, $\cdots$
  - example: traceroutes containing hops in both ⟪Germany⟫ and ⟪Japan⟫



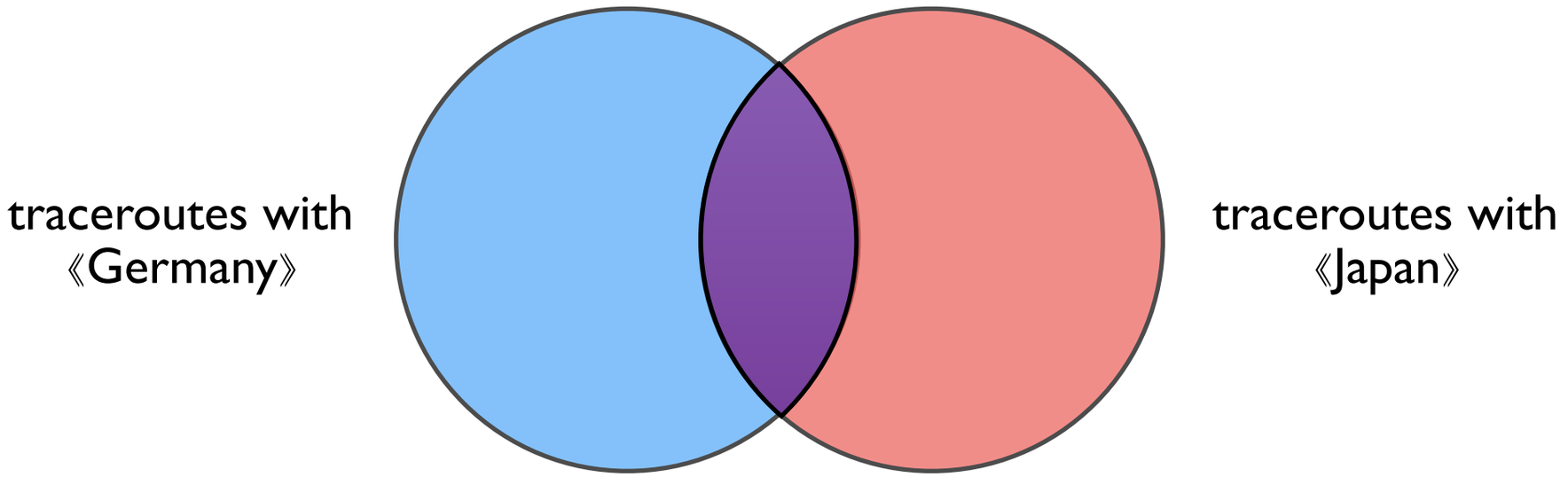
query complexity

- the most complex case:
  - traceroutes containing **two or more** ⟪*targets*⟫
    - precisely: traceroutes containing some hop $h_1 \in$ ⟪*targets*$_1$⟫, $h_2 \in$ ⟪*targets*$_2$⟫, ⋯
  - example: traceroutes containing hops in both ⟪Germany⟫ and ⟪Japan⟫



query complexity

- the most complex case:
  - traceroutes containing **two or more** ⟪*targets*⟫
    - precisely: traceroutes containing some hop $h_1 \in$ ⟪*targets*$_1$⟫, $h_2 \in$ ⟪*targets*$_2$⟫, ⋯
  - example: traceroutes containing hops in both ⟪Germany⟫ and ⟪Japan⟫



- harder:
  - traceroutes with hops in ⟪Germany or UK or France⟫ and hops in ⟪ATT or Level3 network⟫ and hops in ⟪Amsterdam Internet Exchange⟫

challenges

- large target sets
  - 《Germany》 = 9,906 BGP prefixes = 92,239,360 target IP addresses
  - 《Japan》 = 8,769 BGP prefixes = 154,025,984 target IP addresses
- multiple 《*targets*》 in a single query
  - need the **intersection** of subqueries for 《$targets_1$》 and 《$targets_2$》 and ...

challenges

- large target sets
  - 《Germany》 = 9,906 BGP prefixes = 92,239,360 target IP addresses
  - 《Japan》 = 8,769 BGP prefixes = 154,025,984 target IP addresses
- multiple 《*targets*》 in a single query
  - need the **intersection** of subqueries for 《$targets_1$》 and 《$targets_2$》 and ...

- these challenges poorly met by existing database systems
  - relational databases not designed/optimized for multi-key searches
    - can't always use column indexes; may need to do table scans on separate columns
  - not a good fit for existing NoSQL databases
    - schema-less document stores (JSON/XML) come the closest

design

- implemented **custom index data structures**
  - highly tailored and tuned to the characteristics of our data and workload
    - efficiently supports large numbers of targets and subquery intersections
  - gave up generality and flexibility for speed

design

- implemented **custom index data structures**
  - highly tailored and tuned to the characteristics of our data and workload
    - efficiently supports large numbers of targets and subquery intersections
  - gave up generality and flexibility for speed

- built on **RocksDB** key-value store
  - persistent hash table
  - maps binary string (key) to binary string (value)
    - can also traverse keys in sorted order
  - stores both traceroute data and custom indexes

design

- implemented **custom index data structures**
  - highly tailored and tuned to the characteristics of our data and workload
    - efficiently supports large numbers of targets and subquery intersections
  - gave up generality and flexibility for speed
- built on **RocksDB** key-value store
  - persistent hash table
  - maps binary string (key) to binary string (value)
    - can also traverse keys in sorted order
  - stores both traceroute data and custom indexes
- **custom query engine**
  - written in Python
  - running on 64 cores; may use HPC facilities in future

ad-hoc queries

Query Traces for IP Paths

Displays traceroute paths.

Query

Target Address/Prefix/AS/Country: 192.168.0.0/24

Second Target for *neigh* Query:

Separate multiple targets with commas.
Example: 1.2.3.4, 10.0.0.0/8, as1234, .sy

Start Date: 2016-01 End Date: 2016-02

Dates can be *YYYY*, *YYYY-MM*, or *YYYY-MM-DD*. End date is exclusive.
Leave start/end (or both) blank for an open-ended range.

Query Method: dest addr neigh

dest — search by trace *destination* address
addr — search for *responding address* (hop or responding destination address)
neigh — search for *neighboring* addresses (responding hop or destination)

Target Position/Neighbor Separation: 0 Max Traces: 10 Reverse Order

positive position — hop distance relative to *beginning* of trace
negative position — hop distance relative to *end* of trace
neighbor **separation** — hop distance *between* neighboring targets

Vantage Point

ams-nl * By Continent By Country By Org Type

Monitors with IPv6 have an asterisk next to their name.

Submit Reset

ad-hoc queries

Neighbor query of 206.223.119.0/24 and as6939 from *bma-se*

Download JSON results

1. Traceroute to 173.218.24.1 on **2016-01-01 00:26:24**

Hop	Address	Target Match	Prefix	AS	Location	RTT (ms)
1	*					
2	*					
3	95.143.207.173		95.143.192.0/20	49770	hudiksvall swe	5.8
4	**MX-CORE1.internetport.se** 95.143.207.229		95.143.192.0/20	49770	hudiksvall swe	5.4
5	**CO-RO2.internetport.se** 95.143.207.186		95.143.192.0/20	49770	hudiksvall swe	5.5
6	**gige-g2-1.core1.sto1.he.net** 192.121.80.162				stockholm swe	18.8
7	**v991.core1.slc1.he.net** 72.52.92.81	**72.52.64.0/18** (as6939)	72.52.92.0/24	6939	fremont, ca usa	30.0
8	**100ge5-2.core1.par2.he.net** 72.52.92.13	**72.52.64.0/18** (as6939)	72.52.92.0/24	6939	fremont, ca usa	40.2
9	**100ge10-1.core1.nyc4.he.net** 184.105.81.77	**184.104.0.0/15** (as6939)	184.104.0.0/15	6939	new york, ny usa	117.4
10	**100ge5-1.core1.chi1.he.net** 184.105.223.161	**184.104.0.0/15** (as6939)	184.104.0.0/15	6939	chicago, il usa	132.2
11	**equinix-chi.suddenlink.NET** 206.223.119.72	**206.223.119.0/24** (A)			chicago, il usa	127.7
12	**173-219-231-169.suddenlink.net** 173.219.231.169		173.216.0.0/14	19108	lufkin, tx usa	164.7

pre-made analysis

Query Traces for RTT Time Series

Plots an RTT time series for target destinations, an RTT histogram, and a time series of target unreachability.

Query

Target Address/Prefix/AS/Country:

Separate multiple targets with commas.
Example: 1.2.3.4, 10.0.0.0/8, as1234, .sy

Start Date: End Date:

Dates can be *YYYY*, *YYYY-MM*, or *YYYY-MM-DD*. End date is exclusive.
Leave start/end (or both) blank for an open-ended range.

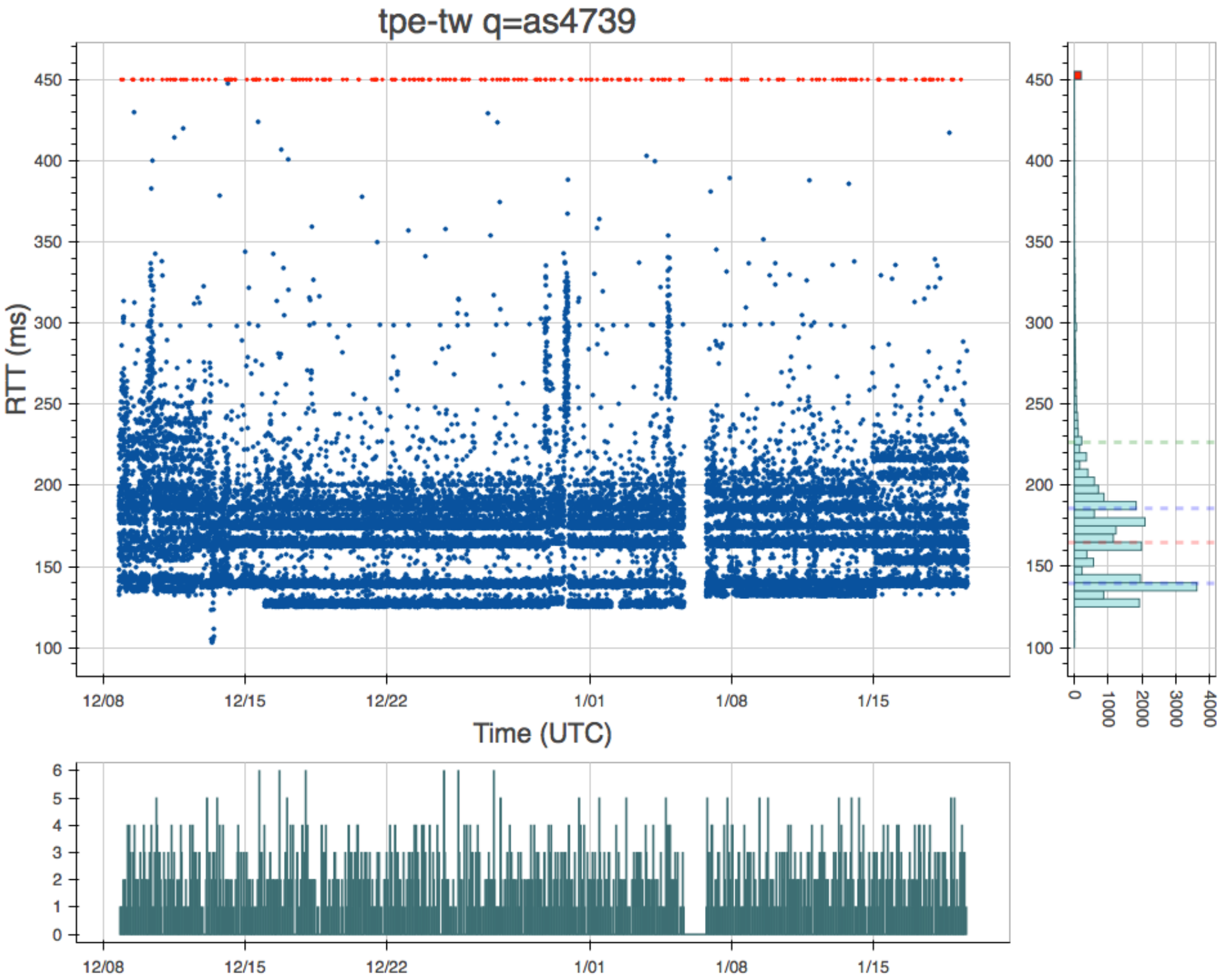
Vantage Point

By Name | By Continent | By Country | By Org Type

Monitors with IPv6 have an asterisk next to their name.

Submit Reset

pre-made analysis



conclusions

- Henya **opens up** our vast data archive to researchers

- Henya **broadens the base** of potential users with built-in analyses and visualizations

- Henya integrates available data into a **whole that's greater than the parts**

acknowledgments

The work was funded by the Department of Homeland Security (DHS) Science and Technology Directorate, Cyber Security Division (DHS S&T/CSD) Broad Agency Announcement 11-02 and SPAWAR Systems Center Pacific via contract number N66001-12-C-0130, and by Defence Research and Development Canada (DRDC) pursuant to an Agreement between the U.S. and Canadian governments for Cooperation in Science and Technology for Critical Infrastructure Protection and Border Security. The work represents the position of the authors and not necessarily that of DHS or DRDC.