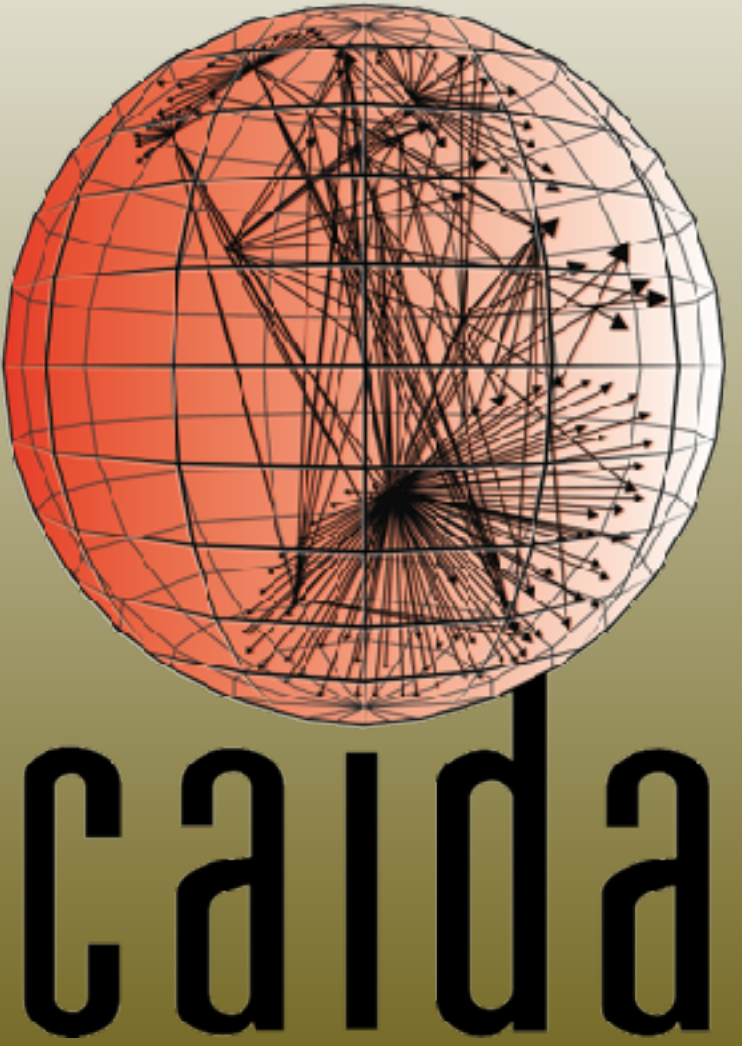
CAIDA/UCSD

TTA#1 PI k claffy, CAIDA
SRI
Menlo Park, CA
7-8 December 2017



Overview

- Statement of Work
  - **Supporting Cybersecurity Research through Network Data Collection and Curation** (TTA #1)
    1. Data Provider - U.S. backbone bidirectional traffic data
    2. Data Host - manage, maintain, distribute
    3. New Data - to support detection and analysis of connectivity disruptions incidents (see Alberto's talk)
    4. IMPACT team activities
- Related activities
  - New data infrastructure
  - New data sets
  - External data statistics

http://www.caida.org/funding/impact/

[will be updated for new contract in Jan 2018]

Task1.1 Data Provider

- Curate, archive data to support security R&D
  - Internet Topology Measured from Ark Platform
  - Internet Topology Data Kits
  - UCSD Real-time Network Telescope Data
  - (all of above co-funded by NSF)

  - Seek to develop instrumentation to collect U.S. backbone bi-directional traffic data
    - plan to acquire a 100GB packet capture monitor (Year 2)
    - capture packet headers
    - currently conducting PcapDB testing, mostly to help LANL with pilot existing technology (DHS supported)

Task1.2 Data Host

- Manage, maintain, share CAIDA data with vetted security researchers via IMPACT portal
  - expand hosting capabilities
  - index new datasets in IMPACT
  - upgrade compute servers

Task1.3 Generate New Data Sets

- Generate new data sets to enable effective detection and analysis of incidents of connectivity disruptions
  - logs of detected outages inferred from BGP
  - darknet traffic
  - active measurements from Ark
  - crowd-sourced measurements of networks vulnerable to IP source address spoofing
  - Above data will serve as input to TTA#2

Task1.4 Contribute to IMPACT

- Contribute to improvement of IMPACT portal utility, convenience, and overall user experience
  - work with IMPACT team
  - update MoAs
  - host project meetings
  - provide documentation
  - outreach materials
  - marketing efforts
  - annual summary of research outcomes from use of our data through IMPACT

Data Collection Infrastructures

- Ark Platform (as of December 2017)
  - 203 monitors in 63 countries
  - 94 IPv6-enabled
  - 158 Raspberry PIs
- UCSD Network Telescope
  - As of January 2017, captures more than 1-1.5 TB of compressed traffic trace data per day.
  - 37 TB: last full month (Nov 2017)
  - 323 TB: 2016
  - 360 TB: YTD 2017 (as of 12/3/17)
  - 393 TB: last 12 months at NERSC (as of 12/3/17)
  - 1162 TB: total archived at NERSC
- New compute platform (Thor 2.0)
  - 2x E5-2630 v4 CPUs (10 core each @ 2.2 GHz).
  - 512GB of RAM.
  - 12x 4TB HDDs (+2 OS drives)

New and Upcoming Data Sets

1. BGP Communities (CIRI deliverable) dictionary
    1. more complete partly manually produced static version (used in SIGCOMM Kepler paper)
       V. Giotsas, C. Dietzel, G. Smaragdakis, A. Feldmann, A. Berger, and E. Aben, "Detecting Peering Infrastructure Outages in the Wild", in ACM SIGCOMM, Aug 2017.
    2. less complete automatically generated version of data

- (What is a BGP Community Dictionary?)
  - BGP Communities have format X:Y, each 16-bit values
    - X contains ASN of the operator that sets the community
    - Y contains arbitrary value to denote specific information, e.g, ingress location of a route,
  - Two types of communities:
    - inbound communities: applied when receiving prefix advertisement
    - outbound communities applied when sending prefix advertisement
  - Only BGP attribute with no standardized semantics or encoding

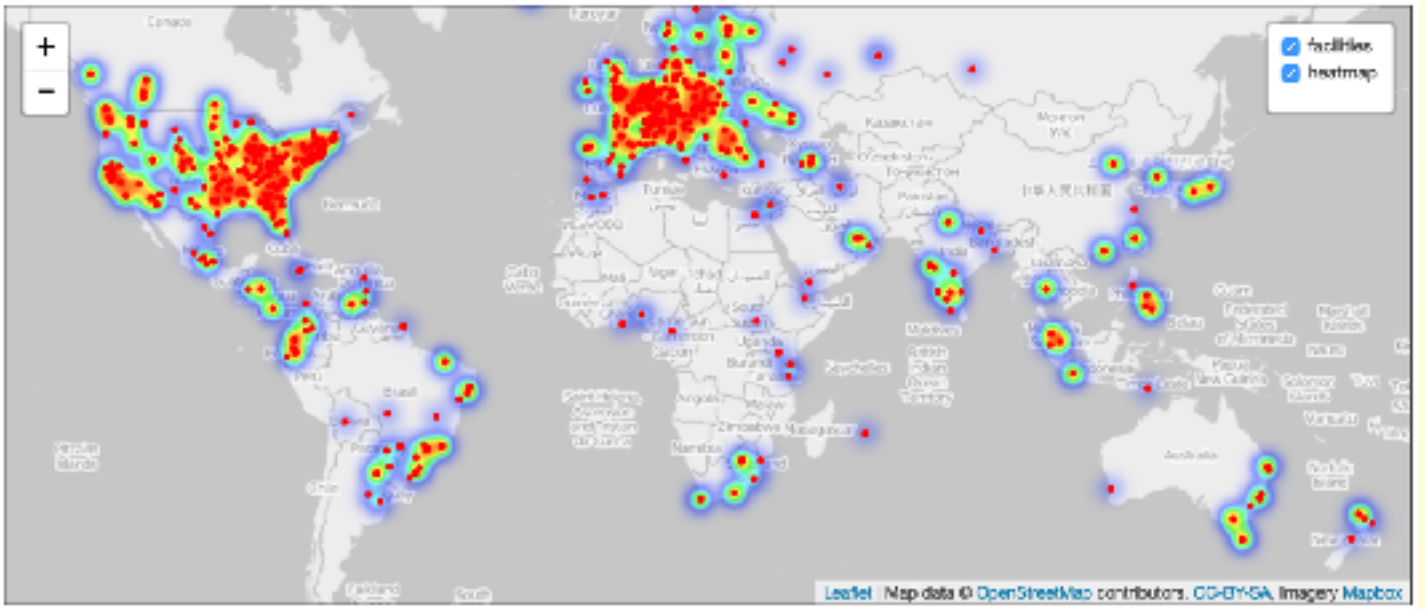
New and Upcoming Data Sets (cont)

2. Mapping ASes to Facilities (CIRI/SISTER deliverable)

- access mode Quasi-Restricted (QR)

• The AS Facilities Dataset

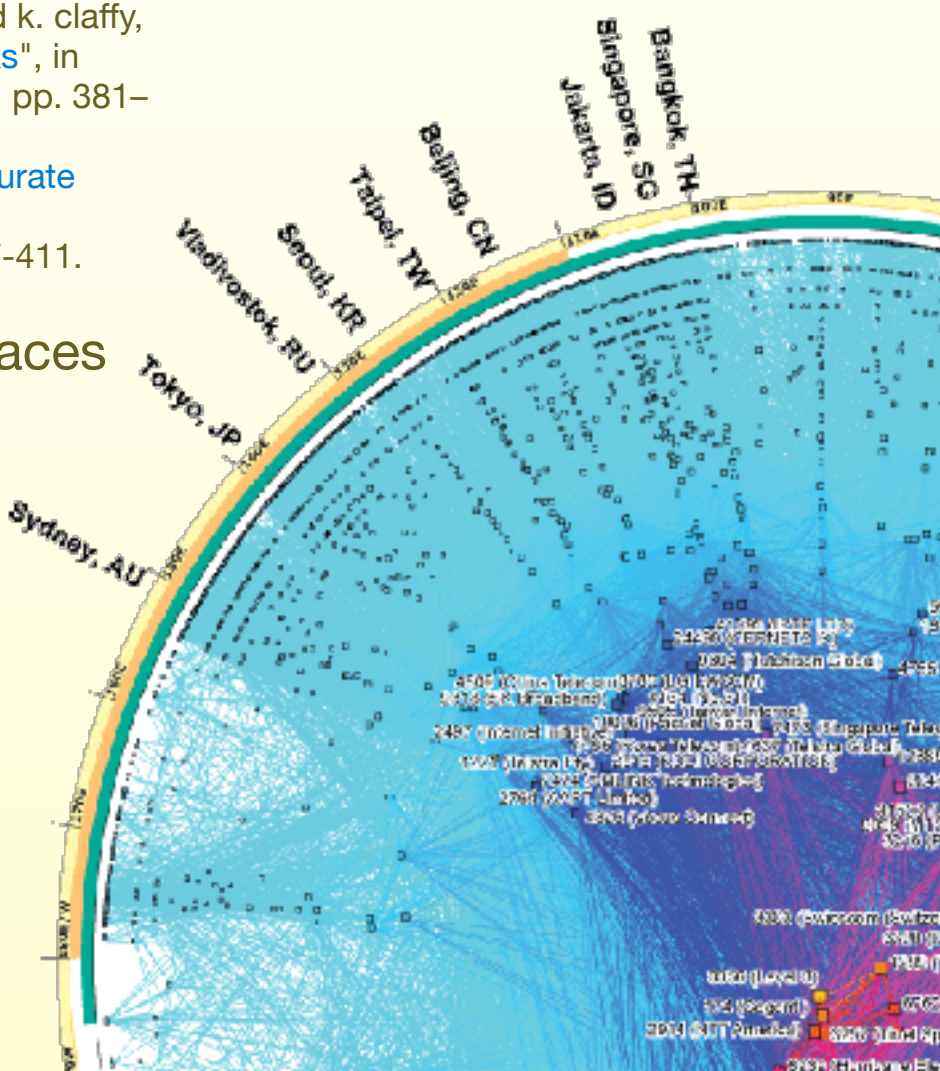
- geographic locations of interconnection facilities
- autonomous systems (ASes) that interconnect at those facilities
- maps peering interconnections to interconnection facilities and, in some cases, to an Internet eXchange (IX) point.
- interconnection method: Private Peering with Cross-connect, Public Peering, Private Interconnects over IXP, or Remote Peering



New and Upcoming Data Sets (cont)

3. Internet Topology Data Kit (ITDK) 2017-08

- access mode Restricted
- First version to use bordermapIT
  - M. Luckie, A. Dhamdhere, B. Huffaker, D. Clark, and k. claffy, "bdrmap: Inference of Borders Between IP Networks", in Internet Measurement Conference (IMC), Nov 2016, pp. 381–396.
  - A. Marder and J. M. Smith, "MAP-IT: Multipass Accurate Passive Inferences from Traceroute." in Internet Measurement Conference (IMC), Nov 2016, pp. 397-411.
  - source code complete; not yet published
- First version to include RIPE Atlas traces



New and Upcoming Data Sets (cont)

4. Software Systems for Surveying Spoofing Susceptibility (Spoofer) data
   - decision required on what we can submit and conditions
   - access mode Restricted ?

- DHS S&T funded project (until July 31, 2018) that seeks to minimize Internet's susceptibility to spoofed DDoS attacks
- Goal: develop, build, and operate multiple open-source software tools to assess and report on the deployment of source address validation (SAV) best anti-spoofing practices.
- https://spoofer.caida.org/ <— plz download now!

Software Systems for Surveying Spoofing Susceptibility

Recent Tests

Result filters:

ASNs: Country codes: Exclude NAT Only show spoofing Change filters

Session	Timestamp	Client IP	ASN	Country	NAT	Spoof Private	Spoof Routable	v4 Adjacency Spoofing	Results
73442	2016-09-28 11:57:32	82.195.64.x	6830 (LGI-UPC)		yes	rewritten	rewritten	none	Full report
73440	2016-09-28 11:57:10	37.235.60.x	57169 (EDIS AS EU)		no	blocked	received	/8	Full report
73439	2016-09-28 11:57:07	84.59.214.x	3209 (VODANET)		yes	blocked	blocked	none	Full report
73438	2016-09-28 11:51:56	95.90.233.x	31334 (KABELDEUTSCHLAND-AS)		yes	blocked	blocked	none	Full report
		2a02:8109::x	31334 (KABELDEUTSCHLAND-AS)		no	blocked	blocked		
73437	2016-09-28 11:49:27	91.14.132.x	3320 (DTAG)		yes	blocked	blocked	none	Full report
73435	2016-09-28 11:47:31	79.237.172.x	3320 (DTAG)		yes	rewritten	rewritten	none	Full report
		2003:66::x	3320 (DTAG)		no	blocked	blocked		
73434	2016-09-28 11:43:39	94.214.191.x	9143 (ZIGGO)		yes	blocked	blocked	none	Full report
73431	2016-09-28 11:35:16	70.196.30.x	22394 (CELLCO)	usa (United States)	yes	blocked	rewritten	none	Full report
		2600:100e::x	22394 (CELLCO)		no	blocked	blocked		
73429	2016-09-28 11:30:12	213.221.216.x	15600 (FINECOM)	che (Switzerland)	yes	blocked	blocked	none	Full report
73428	2016-09-28 11:21:08	122.252.250.x	24560 (RAILTEL-AS-IN)	ind (India)	yes	unknown	unknown	none	Full report
73424	2016-09-28 11:09:37	37.201.192.x	6830 (LGI-UPC)	deu (Germany)	yes	blocked	blocked	none	Full report
		2a02:908::x	6830 (LGI-UPC)		no	blocked	blocked		
73423	2016-09-28 11:08:43	128.151.13.x	20 (UR)	usa (United States)	no	unknown	unknown	none	Full report
73421	2016-09-28 11:08:25	91.154.254.x	719 (ELISA-AS)	fin (Finland)	no	unknown	unknown	none	Full report
73420	2016-09-28 10:56:58	47.29.88.x	55836 (RELIANCEJIO-IN)	ind (India)	yes	rewritten	rewritten	none	Full report
73419	2016-09-28 10:46:13	66.88.134.x	1136 (KPN)	nld (Netherlands)	yes	blocked	blocked	none	Full report
		204.235.144.x	2452 (TW-CABLE)	usa (United States)	yes	unknown	unknown		

http://spoofer.caida.org/recent_tests.php

New and Upcoming Data Sets (cont)

5. Grey market IPv4 transferred prefixes (DHS SISTER deliverable)

- current static data based on BGP inferences
  - I. Livadariu, A. Dhamdhere, and A. Elmokashfi, "On IPv4 transfer markets: Analyzing reported transfers and inferring transfers in the wild", Elsevier Computer Communications Journal, vol. 111, pp. 105--119, Oct 2017.
- eventually will produce regular list of candidate transferred prefixes that also makes use of traceroute prefix inferences

External Data Distribution Stats

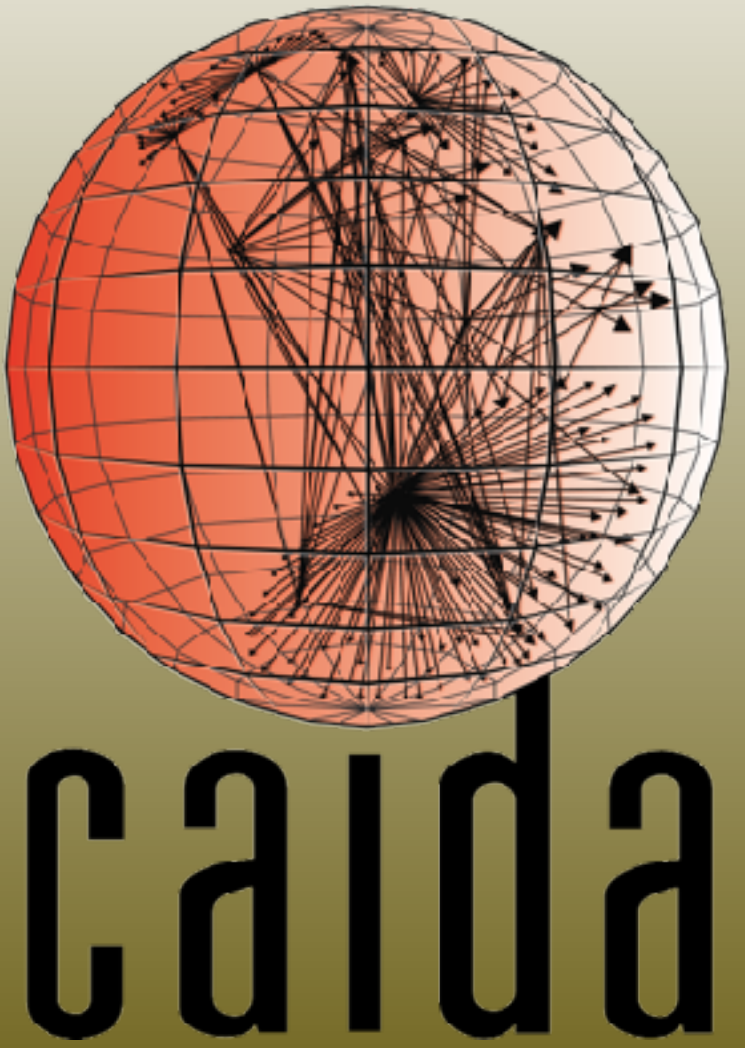
September - November 2017: Data Requests Statistics						
Data Set Name	Requests from IMPACT	Approved requests received through CAIDA	Requests eligible for both CAIDA and IMPACT access	Requests not eligible for CAIDA access but eligible for Impact	Requests from China	Requests not eligible for IMPACT access (and not China)
IPv4 Routed /24 Topology	1	13	8	0	4	1 (Italy)
Anonymized Internet trace data	N/A	95	44	0	22	29*
DDoS 2007 Attack Dataset	4	N/A				
Real-time Network Telescope	1	N/A				
UCSD Telescope Darknet Scanners	3	N/A				
Geolocated Router Dataset	1	N/A				
IPv4 2013 census dataset	1	N/A				

*Brazil, Egypt, Germany,Greece, India, Korea, Nigeria, Pakistan, Sweden, Switzeland, Turkey, Russia,NG, SA, Qatar

Publications

- M. Jonker, A. King, J. Krupp, C. Rossow, A. Sperotto, and A. Dainotti, "Millions of Targets Under Attack: a Macroscopic Characterization of the DoS Ecosystem", in Internet Measurement Conference (IMC), Nov 2017.
- k. claffy and D. Clark, "The 9th Workshop on Active Internet Measurements (AIMS-9) Report'', ACM SIGCOMM Computer Communication Review (CCR), Oct 2017.
- M. Gharaibeh, A. Shah, B. Huffaker, H. Zhang, R. Ensafi, and C. Papadopoulos, "A Look at Router Geolocation in Public and Commercial Databases'', in Internet Measurement Conference (IMC), Nov 2017.
- I. Livadariu, A. Dhamdhere, and A. Elmokashfi, "On IPv4 transfer markets: Analyzing reported transfers and inferring transfers in the wild", Elsevier Computer Communications Journal, Oct 2017

Contact Information



TTA#1 PI: k claffy, CAIDA
kc@caida.org
http://www.caida.org/