CAIDA Activities, 2009-2010

Marina Fomenkov

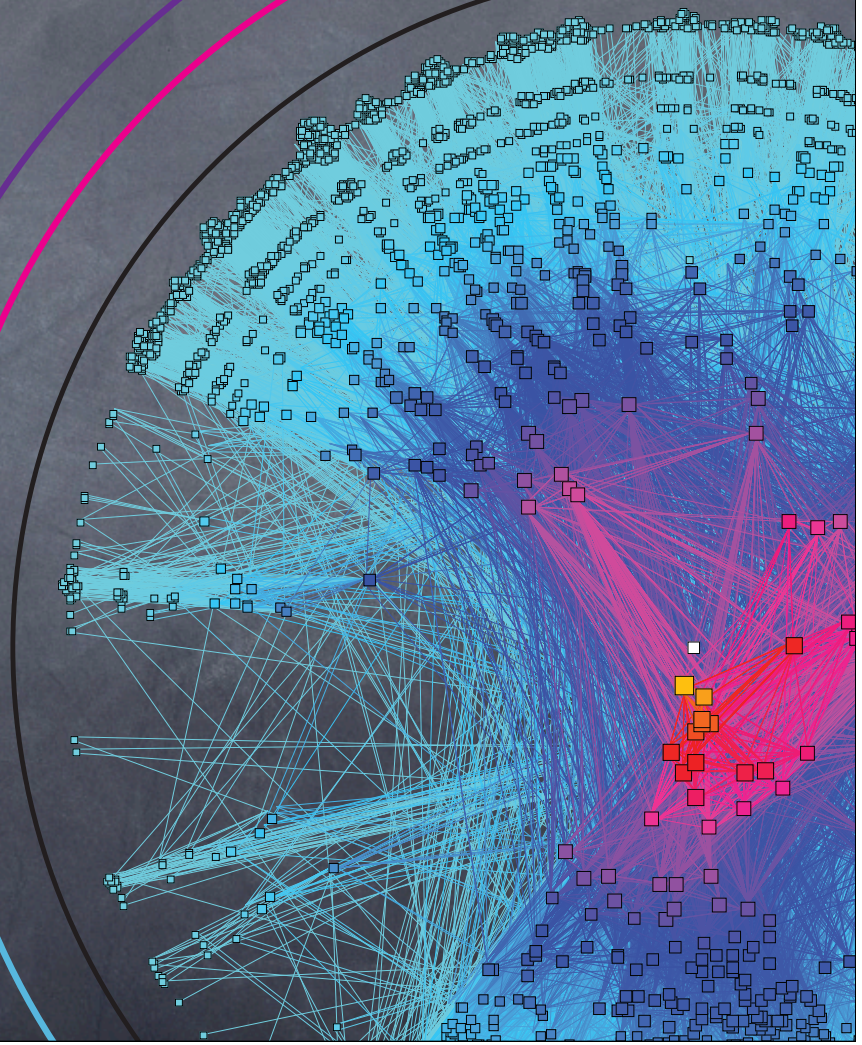
marina@caida.org

CAIDA/UCSD

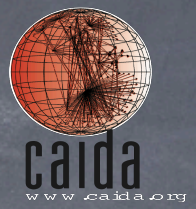
The Quilt Fall 2010
Member Meeting

Sept 29th, 2010

Boston, MA

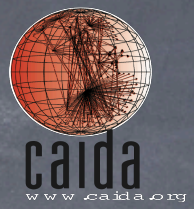


CAIDA



- The Cooperative Association for Internet Data Analysis
- Independent research group
- kc claffy - Founder and Director
- Based at the University of California San Diego/ San Diego Supercomputer Center
- Supported by NSF, DHS, Cisco, and members

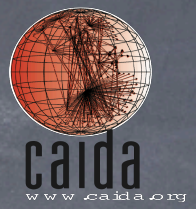
CAIDA



Resources

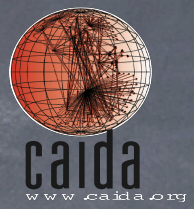
- 2009 annual report
  http://www.caida.org/home/about/annualreports/2009/
- 2010-13 program plan
  http://www.caida.org/home/about/progplan/progplan2010/
- blog
  http://blog.caida.org/best_available_data/
- publications
  http://www.caida.org/publications/

CAIDA mission



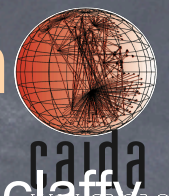
- conducting research
- building infrastructure
- collecting and curating data
- developing tools
- informing policy
- promoting public outreach

Research



- Mapping of Internet Topology
- Modeling of Complex Networks
- Internet Economics
- Traffic Analysis and Classification
- DNS

Internet Economics - previous research



- S. Shakkottai, M. Fomenkov, R. Koga, D. Krioukov, and kc claffy, "Evolution of the Internet AS-Level Ecosystem", (Complex'2009) European Physical Journal B, v. 74, pp.271-278 (2010) http://www.caida.org/publications/papers/2009/AS_evolution/
  - analytically tractable model of Internet evolution at the AS level, based on multiclass preferential attachment
  - closes "measure-model-validate-predict" loop

- A. Dhamdhere, C. Dovrolis, F. Pierre, "A Value-based Framework for Internet Peering Agreements", 22nd International Teletraffic Congress, 2010 http://www.caida.org/publications/papers/2010/framework_peering_agreements/
  - quantitative framework for settlement-free and paid-peering links based on the benefit of a peering link

Internet Economics

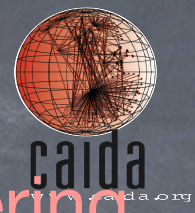
- Workshop on Internet Economics WIE2009
  - virtual, 1 day
  - 15 participants (researchers, policy makers, providers)
  - made use of Caltech's EVO teleconferencing infrastructure
- Goal: *frame an agenda for the emerging but empirically stunted field of Internet infrastructure economics*
- Report published in CCR, v. 40, pp. 55-59 (2010)

  http://www.caida.org/publications/papers/2010/wie_report/

Internet Economics - WIE2009

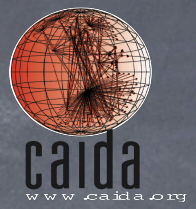
- Economics of Internet identifiers (IP and ASN)
  - How will pricing of IPv4 addresses affect consumers?
  - Which factors define the market for IPv6?
  - How can we track and model the IPv6 penetration process?
  - Will deaggregation impact reachability?
- Economics of Internet peering and interconnection
  - How do settlements occur between ISPs?
  - How can we get more data about peering relationships?
  - What are the underlying cost structures for carrying traffic and expanding capacity?
- The lists to be continued...

Internet Economics - current research



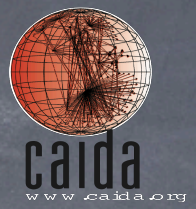
- NSF NetSE: The Economics of Transit and Peering Connections in the Internet
  - collaborative effort of CAIDA and Georgia Tech
  - started in August 2010
- Objective: understand the structure and dynamics of the Internet ecosystem from an economic perspective
  - capture relevant interactions between network business relations, internetwork topology, routing policies, and resulting interdomain traffic flow
  - create a scientific basis for modeling Internet interdomain interconnection and dynamics

Internet Economics



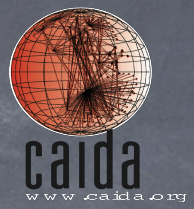
- Studying the evolution of address allocation/ advertisement
  - use historical BGP and WHOIS data
  - measure changes in ownership over time
  - implications for ISP profitability&competition
- Studying peering link dynamics
  - use historical BGP data
  - study ASes providing feeds to RouteViews/RIPE
  - infer AS relationships, track changes over time
- Characterizing peering policies
  - use peeringDB database
  - daily peering snapshots (ISPs volunteer this information)

ITER Model



- Computational model of the Internet ecosystem
  - capture economic, topological, and traffic interactions
- Components
  - different network types
    - Enterprise customers, Transit Providers, Content Providers
  - transit, peering, local operational costs
  - geographical presence
  - provider and peer selection strategies
- Find parameters from measurements
- Calculate equilibrium state
  - no network has incentive to change its providers or peers

Data needs



- **routing policies**
- **topology dynamics**
- **peering practices**
- **pricing/cost structures**
- **interdomain traffic characteristics**

Data are next to impossible to get from commercial providers.

We have a project to support IRNC networks with data collection that will take some baby steps toward this goal.

Very interested in your feedback/suggestions for how to make progress.

NSF IRNC-SP

NSF IRNC-SP: Sustainable data-handling and analysis methodologies for the IRNC networks

Contribute to the IRNC community's measurement efforts:

1. foster and distill discussion of how to best make IRNC data and statistics available
2. adapt two CAIDA measurement technologies (Coralreef and Ark) for IRNC community needs
3. experiment with novel data-handling procedures applied to existing IRNC measurements

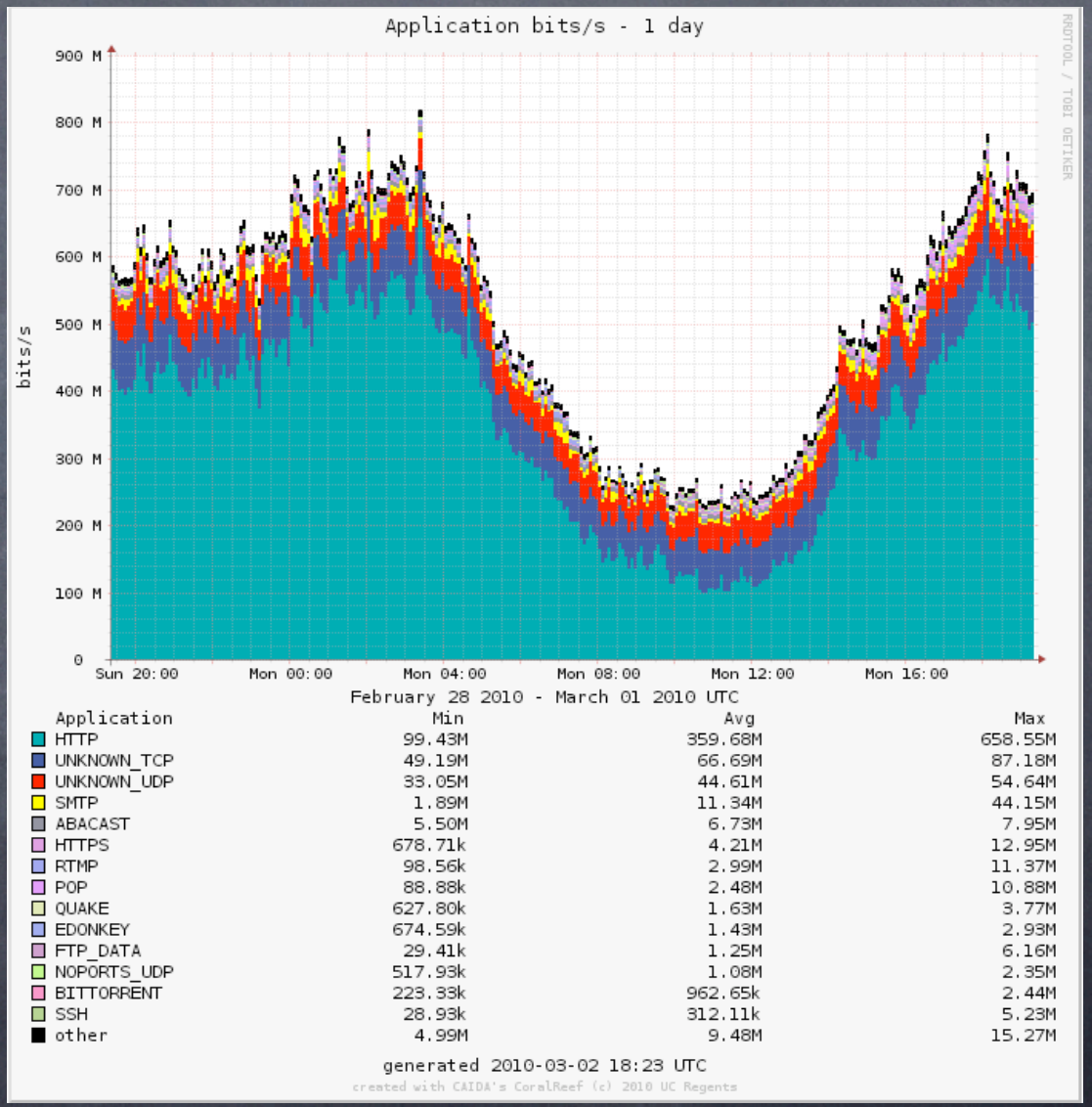
Coralreef software

http://www.caida.org/tools/measurement/coralreef/

- open source
  - maintained and supported by CAIDA
  - available for download
- versatile
  - traffic capture
  - real-time flow analysis
  - data anonymization
  - IPv6 support
  - filtering, aggregation, sampling
  - port-based application classification
  - traffic report generator

Coralreef capabilities

- Report Generator: Chicago-Seattle oc192 link
- bits, packets, flows
- day, week, month

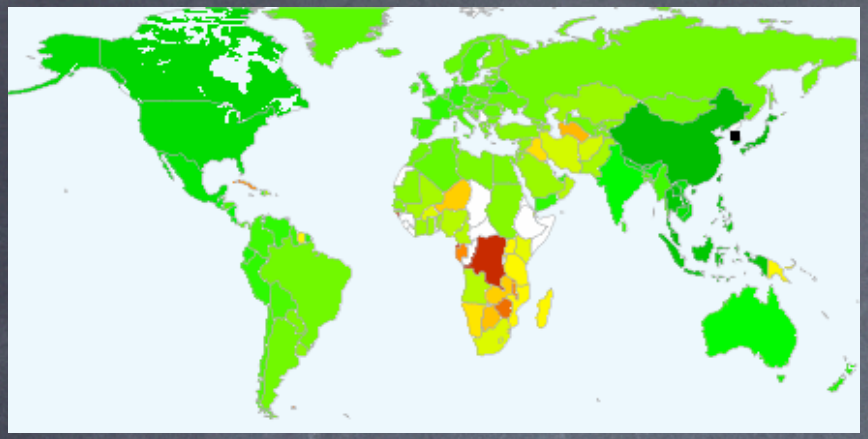


CAIDA active measurement infrastructure

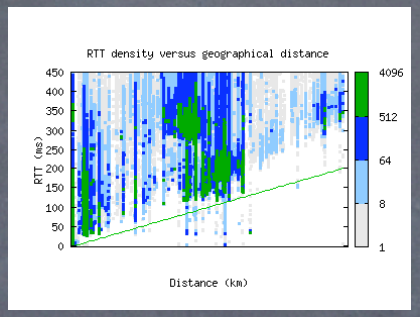
Archipelago

- 47 monitors – growing 1 or 2 per month
- 13 with IPv6 connectivity

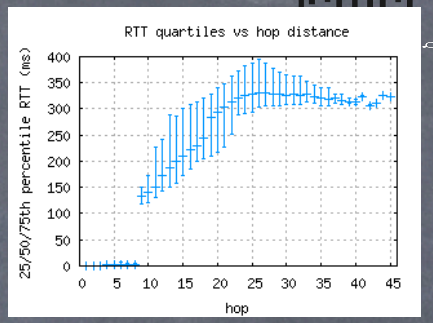
Ark statistics



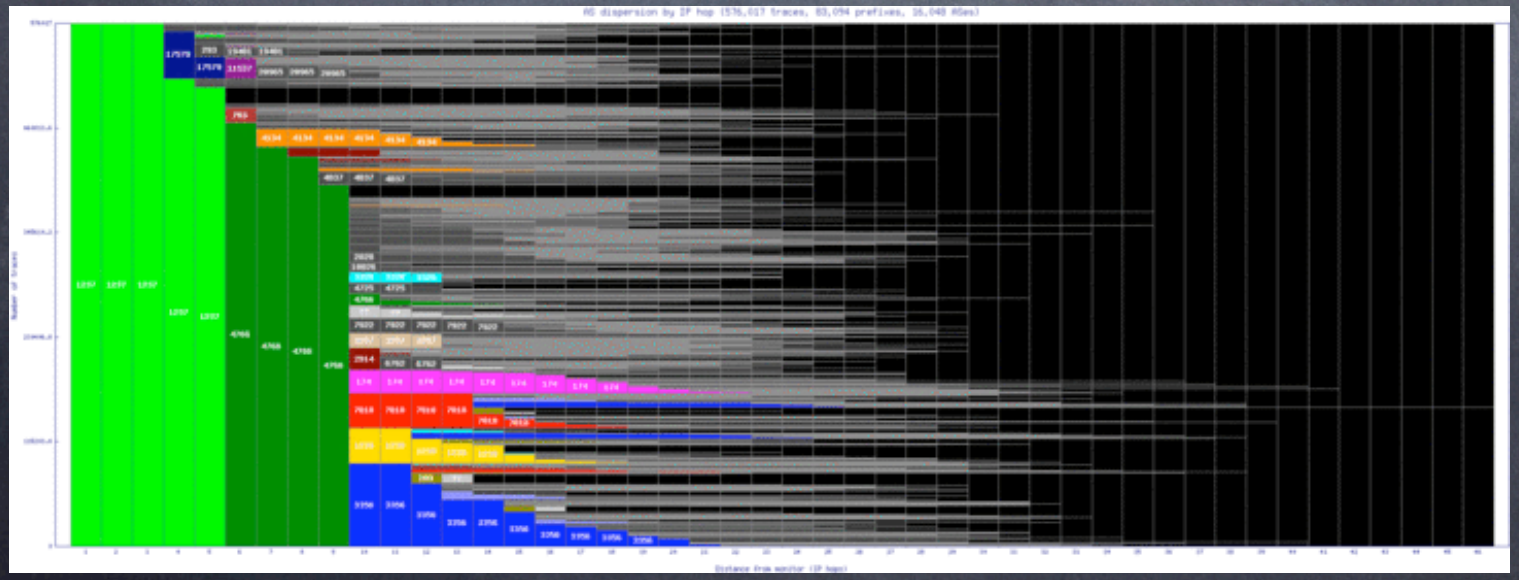
Median RTT per country and US state



RTT density vs distance

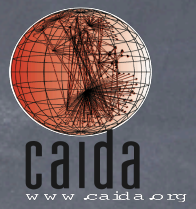


CCDF of dest. RTTs



AS Path Dispersion (by IP Hop)

NSF IRNC-SP Awardee List



- Internet2: Eric Boyd, eboyd@internet2.edu
  - Dynamic Gateway for International Research (DyGIR)
  - International Research Instrumentation System (IRIS)
- University of California, Los Angeles: Lixia Zhang, lixia@cs.ucla.edu
  - 6Watch: Routeviews Infrastructure for Monitoring, Tracking and Diagnosing IPv6 Deployment
- University of Oregon: Dale Smith, dsmith@oregon.edu
  - Cultivating the International Research and Education Network Fabric: An Essential Underpinning of Cyberinfrastructure

Data <-> BTOP

Broadband Technology Opportunities Program (BTOP)

NTIA recommended the following reporting requirements:

i. The terms of any interconnection agreements entered into during the reporting period;

ii. Traffic exchange relationships (e.g., peering) and terms;

iii. Broadband equipment purchases;

iv. Total & peak utilization of access links;

v. Total & peak utilization on interconnection links to other networks;

vi. IP address utilization & IPv6 implementation;

vii. Any changes or updates to network management practices.

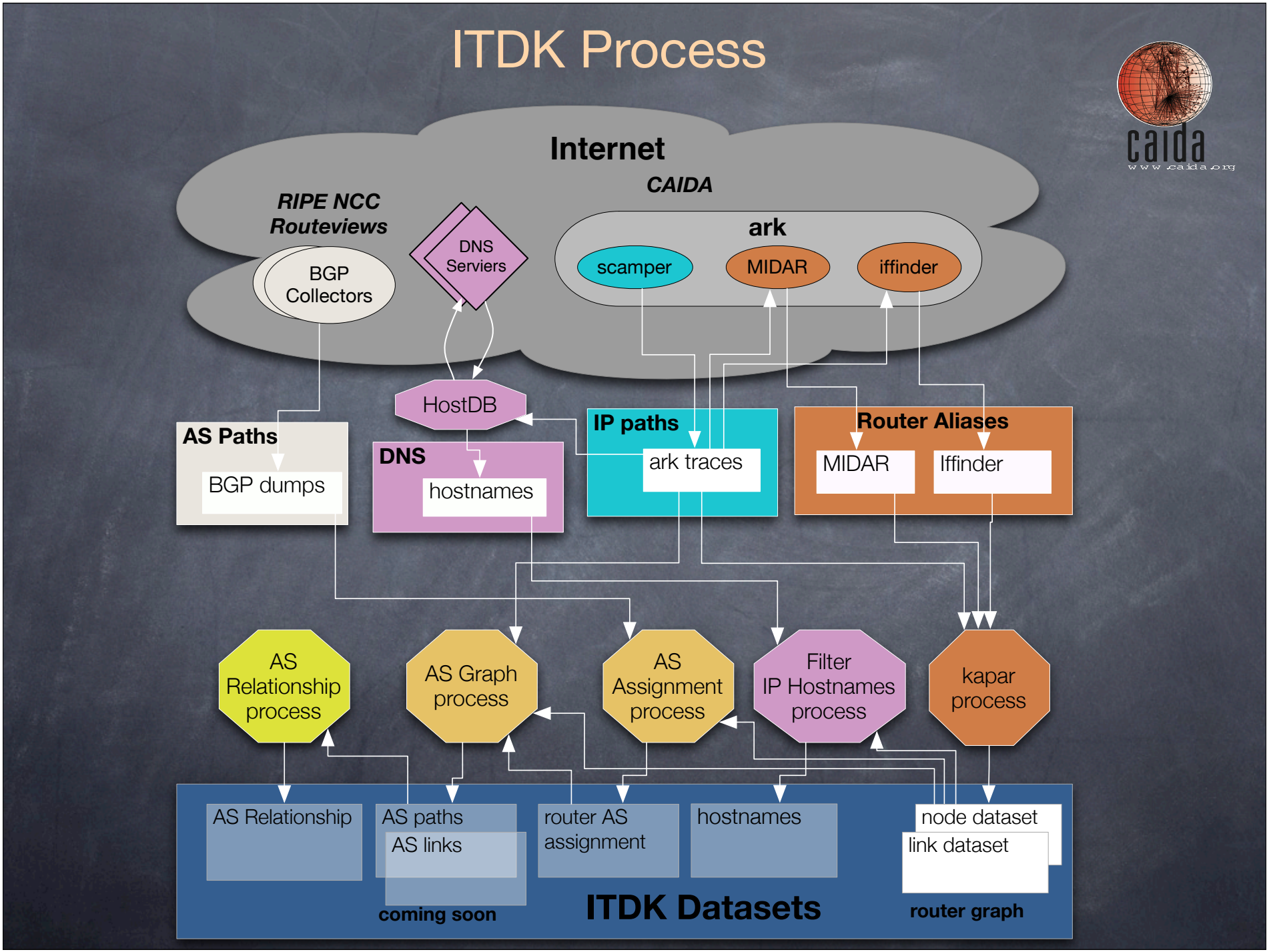
Mapping Internet Topology

- IPv4 and IPv6 topology discovery
- hostname collection
- alias resolution
- router to AS assignment
- AS relationships

Internet Topology Data Kit - ITDK

http://www.caida.org/data/active/internet-topology-data-kit/

ITDK Process



Mapping Internet Topology

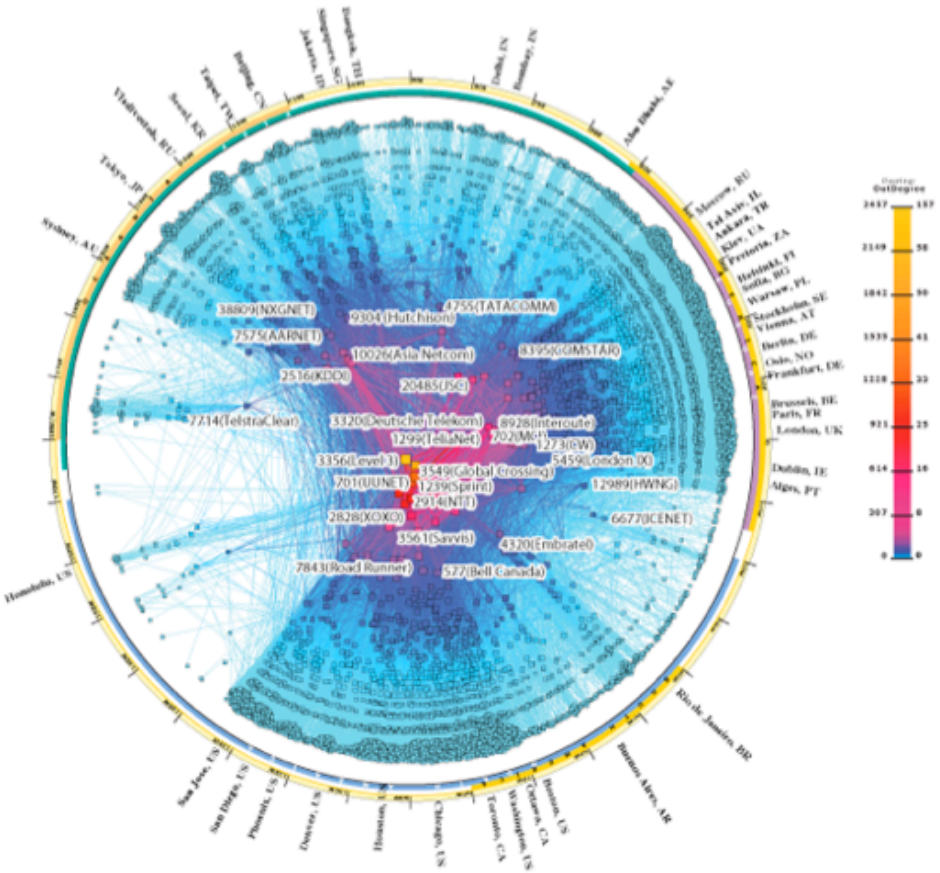
- IPv4 and IPv6 topology discovery
  - continuously collected
  - traceroute-like probing
  - IPv4: random destination in each routed /24 network
    - about 3 days for a complete cycle
  - random destination in each assigned IPv6 prefix, /48 or shorter
    - every 48 hours
- hostname collection
  - try to resolve for every observed IPv4 address
  - release per cycle

IPv4 & IPv6
INTERNET TOPOLOGY MAP
JANUARY 2009

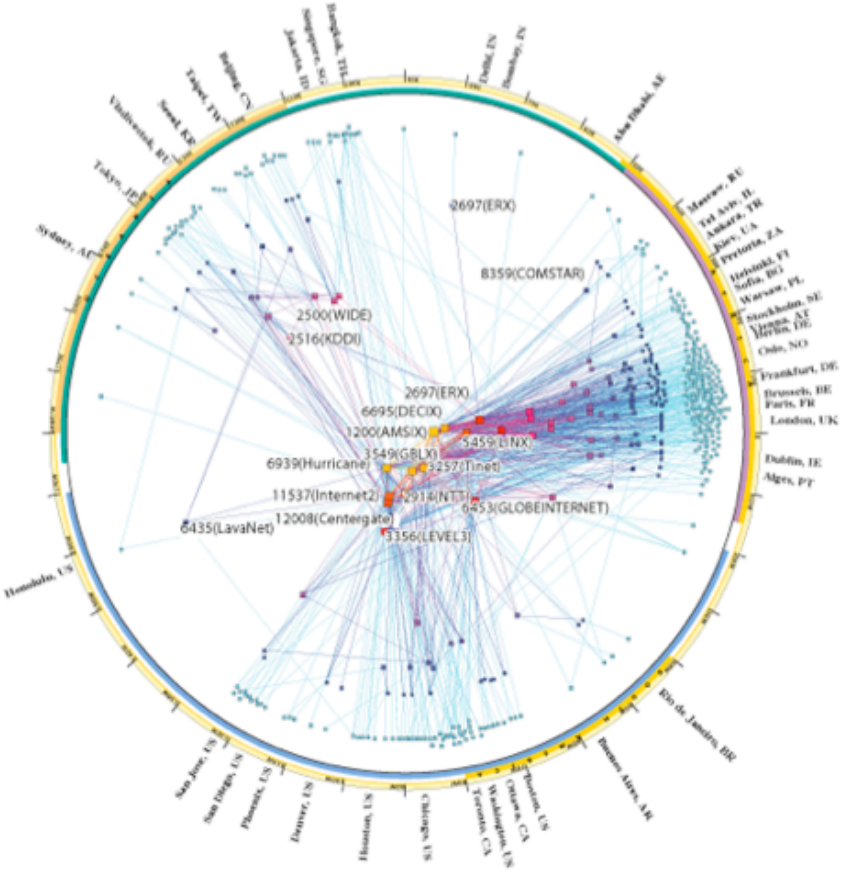
caida

AS-level INTERNET GRAPH

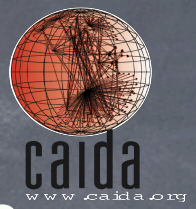
IPv4



IPv6



Mapping Internet Topology



- Alias Resolution = collapse interfaces to produce router-level graph
  - conduct additional measurements
  - use two weeks of topology data for analysis
- CAIDA developed alias resolution tools:
  - iffinder, kapar, MIDAR
  - K. Keys "Internet-Scale IP Alias Resolution Techniques", CCR, v. 40, pp-50-55, 2010. http://www.caida.org/publications/papers/2009/as_assignment/
- MAARS - Multi-Approach Alias Resolution System

Mapping Internet Topology

- router to AS mapping
  - empirically based heuristics
  - validated using Tier-1 network and public data
  - result: dual graph of the Internet
  - B. Huffaker, A. Dhamdhere, M. Fomenkov, kc claffy, "Toward Topology Dualism", PAM 2010. http://www.caida.org/publications/papers/2009/as_assignment/
- AS relationships
  - customer-provider and peering links
  - maximize the number of valid paths
  - validation by providers (work in progress)

AS Rank

Autonomous Systems rank by “customer cone”

http://as-rank.caida.org/ - BETA version

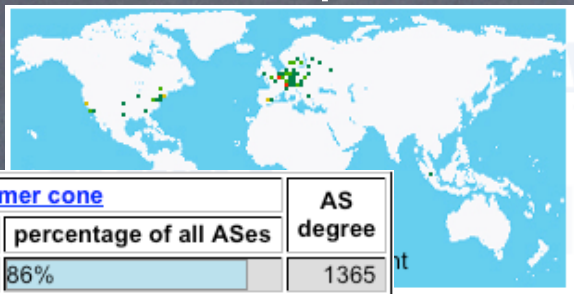
rank	AS number	AS name	customer cone		AS degree
			customer cone size	percentage of all ASes	
1	3356	LEVEL3 Level 3 Commu	31112	92%	2632
2	7018	AT&T WorldNet Servic	29978	89%	2283
3	701	MCI Communications S	29820	88%	2066
4	174	Cogent/PSI	29328	87%	2533
5	3549	Global Crossing Ltd.	29035	86%	1365
6	1239	Sprint	29012	86%	1381
7	209	Qwest Communications	28983	86%	1387
8	6939	Hurricane Electric,	27227	81%	1552
9	4323	tw telecom holdings,	27198	81%	1291
10	1299	TeliaNet Global Netw	27117	80%	561

data sources

country	ASN allocation	2010.04.22	IANA
	delegated	2010.08.19	AFRINIC,APNIC,ARIN,IANA,LACNIC,RIPENCC
	whois	2010.04.01	AFRINIC,APNIC,ARIN,LACNIC,RIPE
name	ASN allocation	2010.04.22	IANA
	autnum.txt	2010.08.19	potaroo.net
	whois	2010.04.01	AFRINIC,APNIC,ARIN,LACNIC,RIPE
topology	BGP	2010.01.29	Ripe NCC RCC12,routeviews2

AS Rank

Tabular views of individual ISP info, rank, degree, customer cone size, customers, peers, and providers



AS number:	1299
AS name:	TeliaNet Global Network
rank:	10
customer cone size:	27117
degree:	561

rank	AS number	AS name	customer cone		AS degree
			customer cone size	percentage of all ASes	
5	3549	Global Crossing Ltd.	29035	86%	1365
6	1239	Sprint	29012	86%	1381
7	209	Qwest Communications	28983	86%	1387
8	6939	Hurricane Electric,	27227	81%	1552
9	4323	tw telecom holdings,	27198	81%	1291
10	1299	TeliaNet Global Netw	27117	80%	561
11	2914	NTT America, Inc.	26832	79%	650
12	6453	TATA Communications	26236	78%	530
13	3561	Savvis	25690	76%	
14	9002	ReTN.net Autonomous	25146	74%	

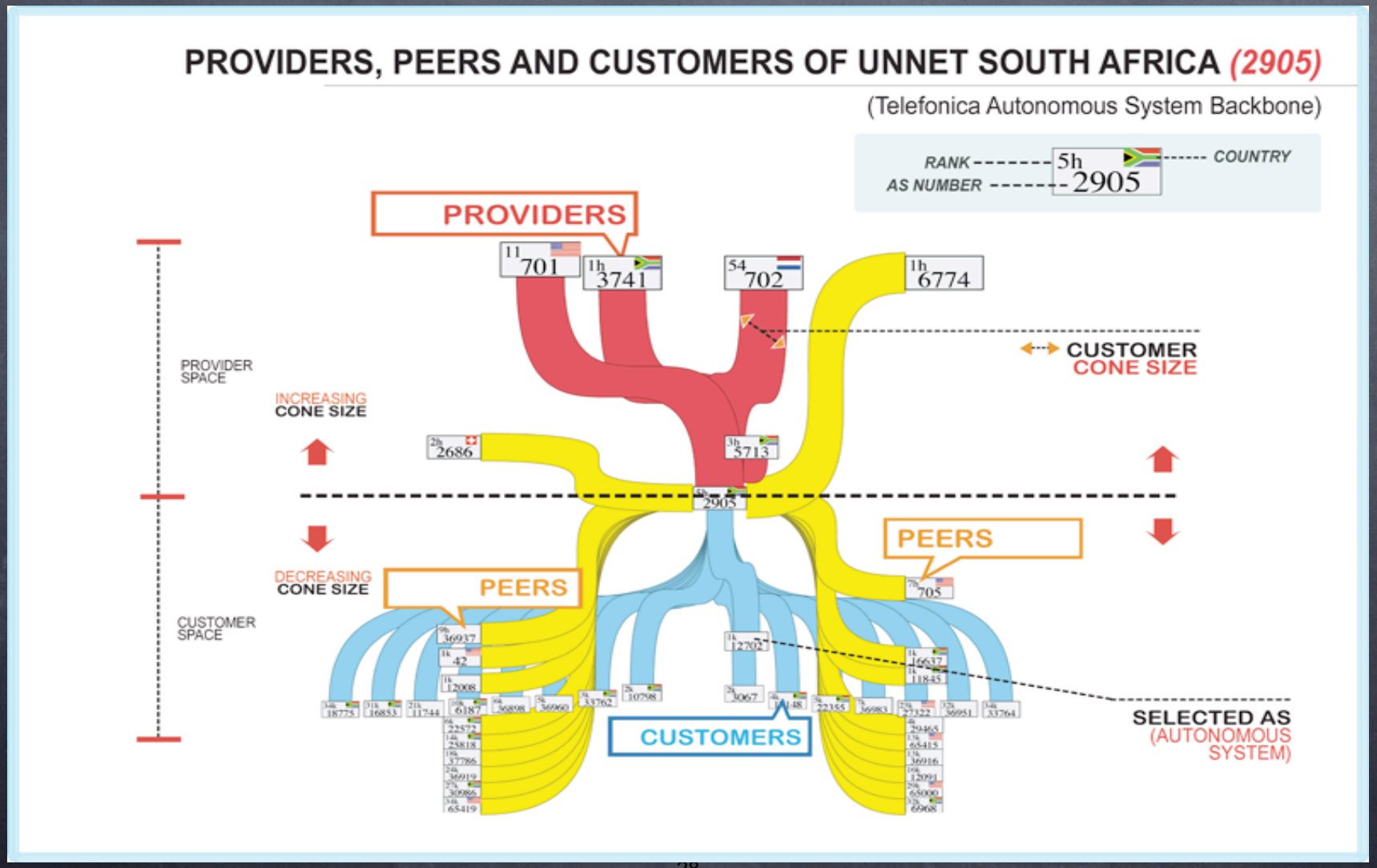
Ranking ←

Customers, providers, and peers →

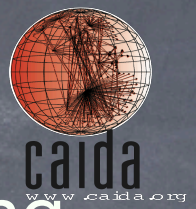
rank	neighbor AS	neighbor name	type
4	174	Cogent/PSI	↑ provider
6	1239	Sprint	↔ peer
5	3549	Global Crossing Ltd.	↑ provider
7	209	Qwest Communications	↑ provider
8	6939	Hurricane Electric,	↔ peer
9	4323	tw telecom holdings,	↔ peer
11	2914	NTT America, Inc.	↓ customer
12	6453	TATA Communications	↓ customer
13	3561	Savvis	↓ customer
15	1273	Cable and Wireless p	↓ customer

AS Visualization

Graphical view of customers, providers and peers.

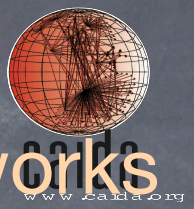


Collaborative Research using Ark



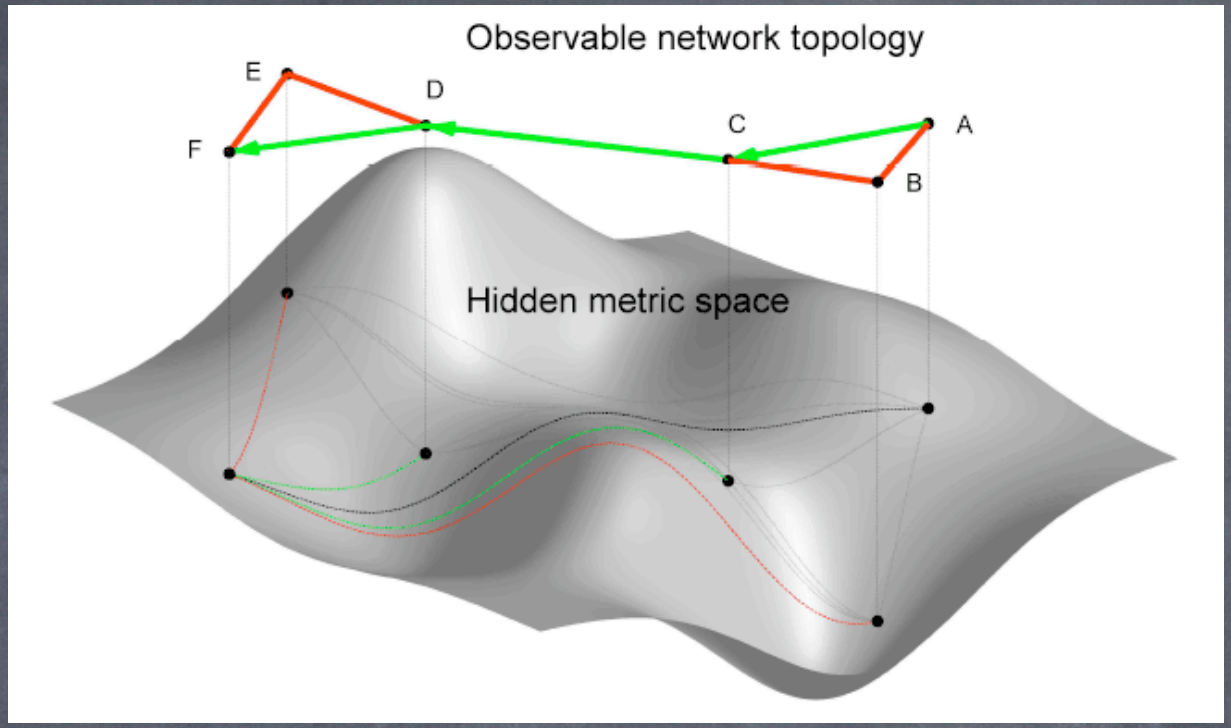
- Using Ark to examine source address spoofing
  - how many networks allow packets with spoofed IP addresses to leave their network
  - R. Beverly, A. Berger, Y. Hung, kc claffy "Understanding the Efficacy of Deployed Internet Source Address Validation Filtering", IMC 2009.
    http://www.caida.org/publications/papers/2009/imc_spoofer/
  - working on adding IPv6
- Improving the efficiency of topology probing
  - implemented Doubletree using Marinda (tuple space)
  - M. Luckie and A. King (Univ. of Waikato, New Zealand)

Modeling Complex Networks



- Explain observed properties of complex networks
  - scale-free
  - highly clustered
  - navigable
- Multidiscipinary research problem
  - physics
  - biology
  - social sciences
  - lots of math (graph theory, group theory, topology, symmetry...)
- Potential outcome: enable more efficient and scalable routing in the Internet

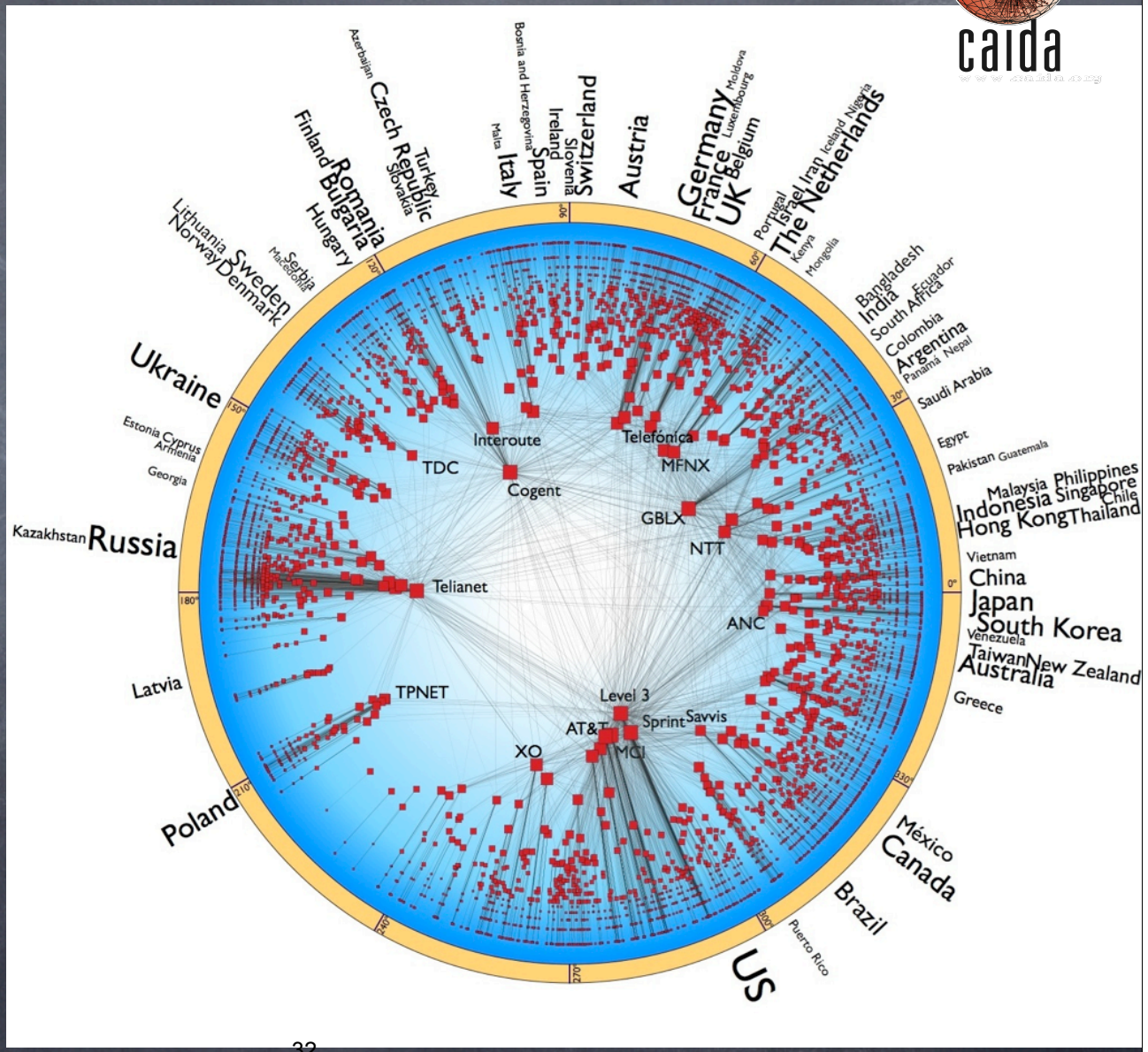
Hidden Metric Spaces



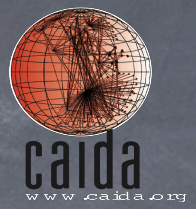
- Find geometry underlying complex networks
  - distances abstract similarities between nodes
  - connection probability decreases with distance
  - explains interaction between network structure and function
  - enables distance-based greedy forwarding

Optimal geometric routing

- hyperbolic metric spaces (exponentially expanding)
- most congruent with observed Internet topology
- embed AS-level graph into a hyperbolic space
- routing reduces to geometric forwarding
- published in Nature Communications, v1, September 2010



Complex networks - publications



- **Scalable Routing with Hyperbolic Mapping**

  M. Boguñá, F. Papadapoulos, D. Krioukov, "Sustaining the Internet with Hyperbolic Mapping", Nature Communications, v1, September 2010.
  http://www.caida.org/publications/papers/2010/sustaining_internet_hyperbolic/

- **Curvature and Temperature**

  D. Krioukov, F. Papadapoulos, A. Vahdat, M. Boguna, "Curvature and temperature of complex networks", Physical Review E, v80, 035101(R), 2009.
  http://www.caida.org/publications/papers/2009/curv_temp_complex_nets/

- **Greedy Forwarding**

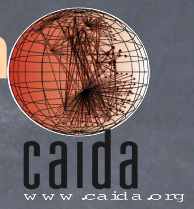
  F. Papadapoulus, D. Krioukov, M. Boguna, A. Vahdat, "Greedy forwarding in scale free networks embedded in hyperbolic metric spaces", in ACM SIGMETRICS Performance Evaluation Review, vol. 37, pp. 15-17, Oct 2009.
  http://www.caida.org/publications/papers/2009/greedy_forwarding_embedded/

- **Navigability**

  M. Boguna, D. Krioukov, kc claffy, "Navigability of complex networks", Nature Physics, v 5, pp-74-80, January 2009.
  http://www.caida.org/publications/papers/2009/navigability_complex_networks/

  M. Boguna, D. Krioukov, "Navigating ultrasmall worlds in ultrashort time", in Physical Review Letters, vol 102, no 058701, 2009.
  http://www.caida.org/publications/papers/2009/navigating_ultrasmall/

Policy to support Empirical Internet Research



- Privacy-Sensitive Sharing (PS2) Framework

  E. Kenneally, kc claffy "Dialing privacy and utility: a proposed data sharing framework to advance Internet research", IEEE Security and Privacy, v. 8, pp. 31-39, 2010.
  http://www.caida.org/publications/papers/2010/dialing_privacy_utility/

  - data sharing: research benefits vs. privacy risks?
    - data seekers and data providers
  - PS2 offers consistent evaluation methodology for risk-benefit determinations
    - analytical tool for assessing risk posture
    - basis for establishing privacy management controls
    - template for developing operational solutions
  - uses AUPs and disclosure control techniques
    - explicit agreement, transparency, oversight, purpose adherence

Policy to support Empirical Internet Research

- Developing Ethical Guidelines for Internet Research

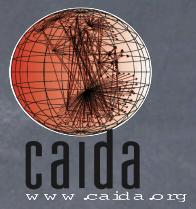
  E. Kenneally, M. Bailey, D. Maughan, "A Framework for Understanding and Applying Ethical Principles in Network and Security Research", WECSR January 2009.
  http://www.caida.org/publications/papers/2010/framework_ethical_research/

  - Grew from the DHS sponsored PREDICT (Protected Repository for the Defense of Infrastructure against Cyber Threats)
  - Inspired by Belmont report
    - authoritative guide on ethical standards for human subject research
  - Goal: develop a workable Ethical Impact Assessment Framework
  - Issues to address:
    - What are user's perceptions of privacy and confidentiality?
    - What are the legal prohibitions to collecting/disclosing network data?
    - How to identify at risk population in a network trace?
    - Is it possible to receive implied consent from users observed in network measurements?
  - Work in progress, feedback/comments are welcome

Policy

- Advising regulators on "Network Neutrality"
  - kc claffy, "Historical and Architectural Context for Traffic Management Needs Today", presented at the FCC Technical Advisory Process workshop in December 2009. http://www.caida.org/publications/presentations/2009/traffic_historical_context/
  - Recommend in short term to protect private property rights as well as consumer/citizen/user rights: transparency, objective data obligations
  - Recommend in longer term to guide FCC, an Interdisciplinary advisory function
  - Segmenting technology, policy, and economic advice is a recipe for failure: the Internet connects it all.
- Empirical Internet science should guide policy

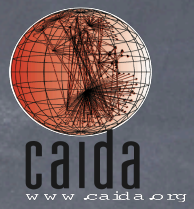
Workshops



- Active Internet Measurement Systems (AIMS 1 and 2)
  - AIMS-3 in February 2011
- Workshop on Internet Economics (WIE)
  - WIE-2 some time in 2011
- Joint workshop with WIDE (Japan)/CASFI (Korea)

http://www.caida.org/workshops/

Conclusion



How The Quilt community might interact with CAIDA

- contribute traffic statistics to help build a model of interdomain interconnection and dynamics
- AS Rank feedback interface to provide corrections to our AS relationships model
- alias resolution validation (router-level topology)
- become CAIDA members

For more information please contact: info@caida.org

http://www.caida.org/