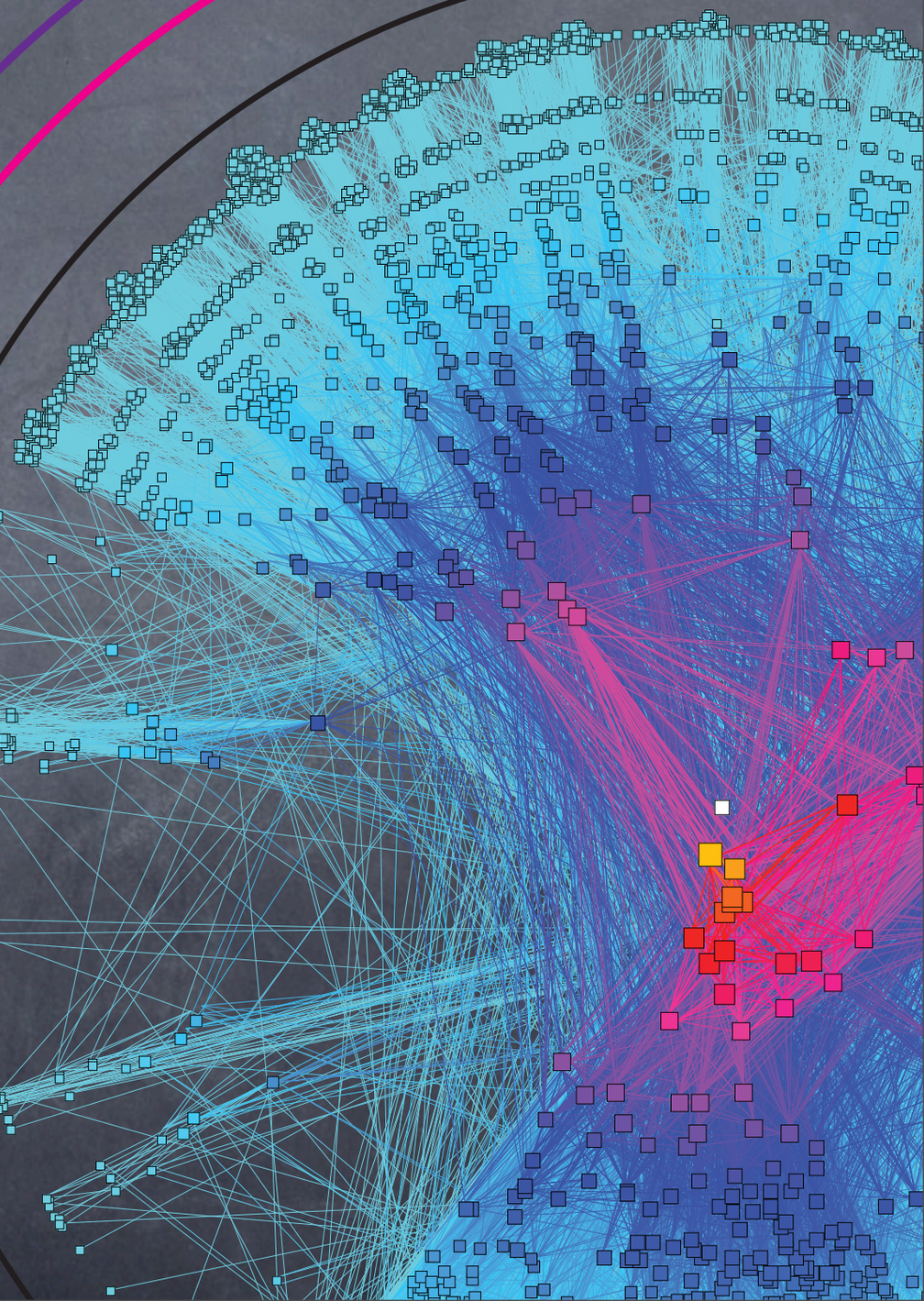
Inferring Geolocation Ownership of Internet Identifiers

Bradley Huffaker
CAIDA/UCSD

Spring 2010 ECTF
26 May 2010
Moffett Field, CA



Geolocation/Ownership

Pacificwave
sinet-1-lo-jmb-702.lsanca.pacificwave.net (207.231.240.135)

Cenic hpr-lax-hpr--sdsc-10ge.cenic.net (137.164.26.33)

SDSC dolphin.sdsc.edu (132.249.31.17)
piranha.sdsc.edu (198.17.46.8)

CAIDA pinot-g1-0-0 (192.172.226.1)

Internet Identifiers are strings used to label resources on the Internet. (ASN, hostnames, IP addresses, domain names, ...)

Geolocation is the identification of the real-world geographic location of Internet ID.

Ownership determining who owns or controls the resource connected to those Internet IDs.

resources (Methods)

Commercial Services

several companies provide turn-key systems for geolocation/ownership

Domain Name System (DNS)

hierarchical naming system for IP addresses

WHOIS

public database maintained by the Regional Internet Registries (RIRs) and National Internet Registries (NIRs)

Border Gateway Protocol (BGP) archives + WHOIS

organizations that maintain historical BGP routing information

Commercial Services

What

Companies that provide a geolocation service. Typical local database/ server against which to send queries.

How

Pay the commercial vender; they do most setup

Pros

- someone else does the hard part
- uniformity of data

Cons

- no historic data
- cost

Commercial Services

(from NANOG responses to CAIDA's geolocation inquiry January 2010)

Major Services

- MaxMind (GeoIP, GeoLite)
- Akamai (EdgePlatform)
- Google (Google Gears)
- Digital Envoy (Netacuity)

Small Services

- Quova (Quova On Demand)
- IP2Location (IP2Location)

WHOIS -- number resources

What

Query/response protocol used to collect registrant/assignee of Internet resources from RIRs and NIRs.

How

Run a WHOIS client against one of the RIRs or NIRs databases.

whois -h whois.arin.net 10.0.2.1

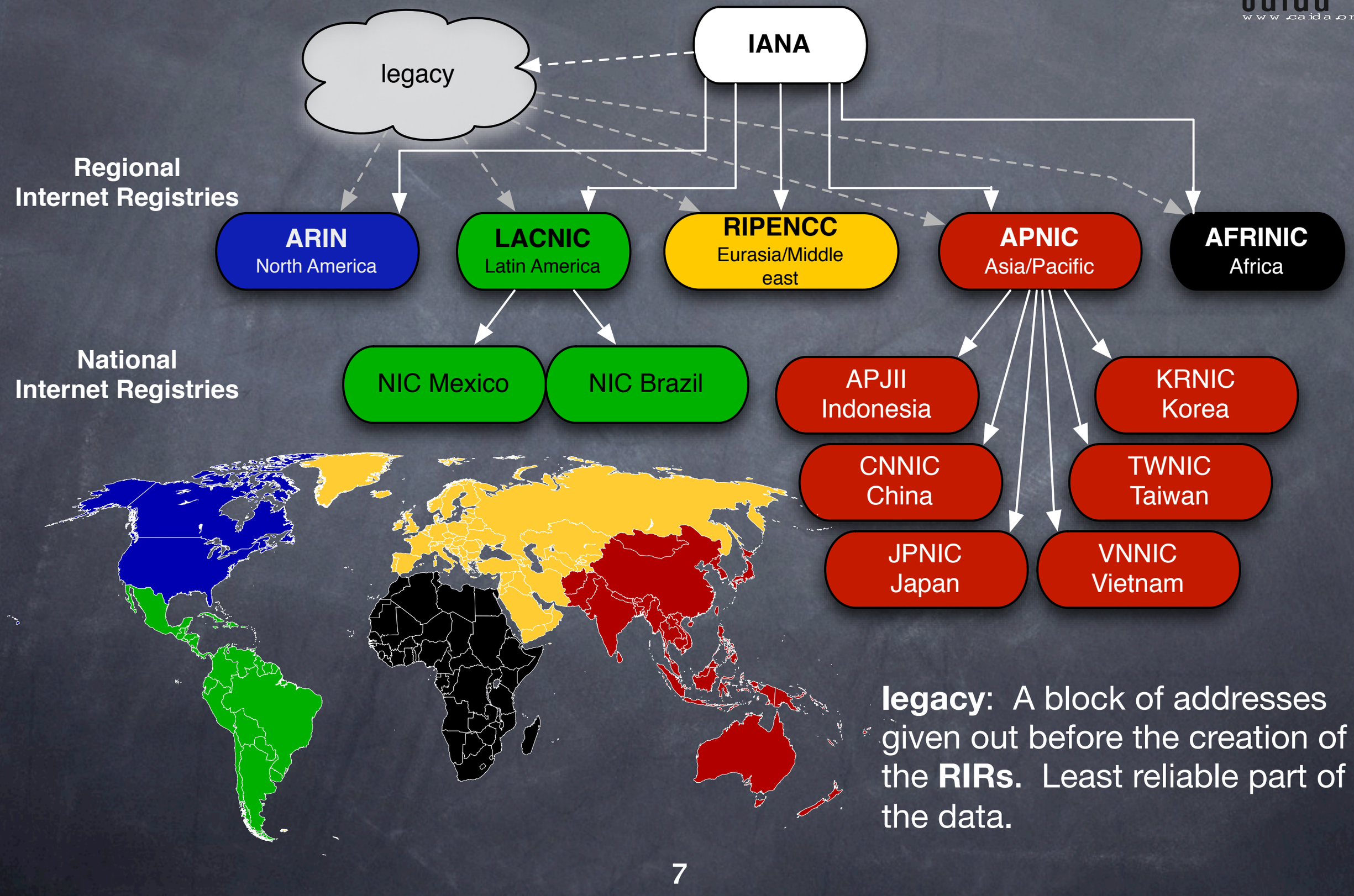
Pros

- free
- provides contact address directly

Cons

- stale entries, low incentive for organizations to maintain new information.
- non-uniform data formats (Some groups provide parsed data)
- no historic data (CAIDA collections dumps every 6 months)

WHOIS database



legacy: A block of addresses given out before the creation of the **RIRs.** Least reliable part of the data.

Domain Name System

What

Hierarchical naming system that provide a mapping between IP addresses and symbolic strings.

How

Run DNS client against a IP address, may get multiple names.

nslookup 10.0.3

Pros

- free
- organization is often apparent
- well-maintained

Cons

- no historic data (CAIDA does maintain for some IP addresses)
- not available for all IP addresses
- user would have to go from organization name to contact
- geographic location must be inferred

BGP + WHOIS

What

BGP collectors provide current and historic data on Internet paths.

How

- Collect BGP tables for period of interest.
- Map IP address to Autonomous System that announced them in those tables.
- Use WHOIS to find ownership information on the Autonomous Systems and IP addresses

Pros

- free
- provides historic data
- provides full path information (more organizations you can contact)

Cons

- complicated, no easy to use tools

conclusion

	difficulty	cost	historic
commercial	easy	low~high	no
DNS	easy	free	no
WHOIS	moderate	free	limited
BGP+WHOIS	hard	free	yes