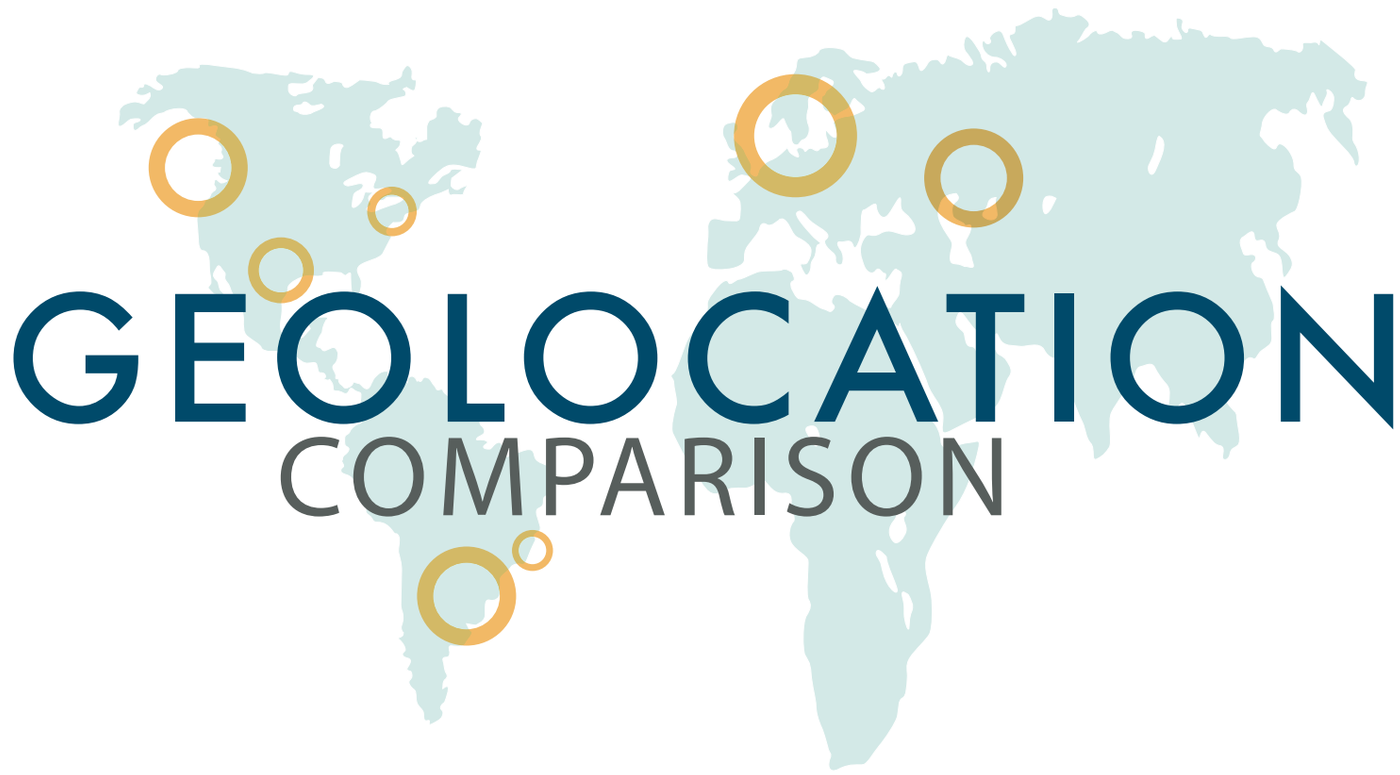


GEOLOCATION COMPARISON

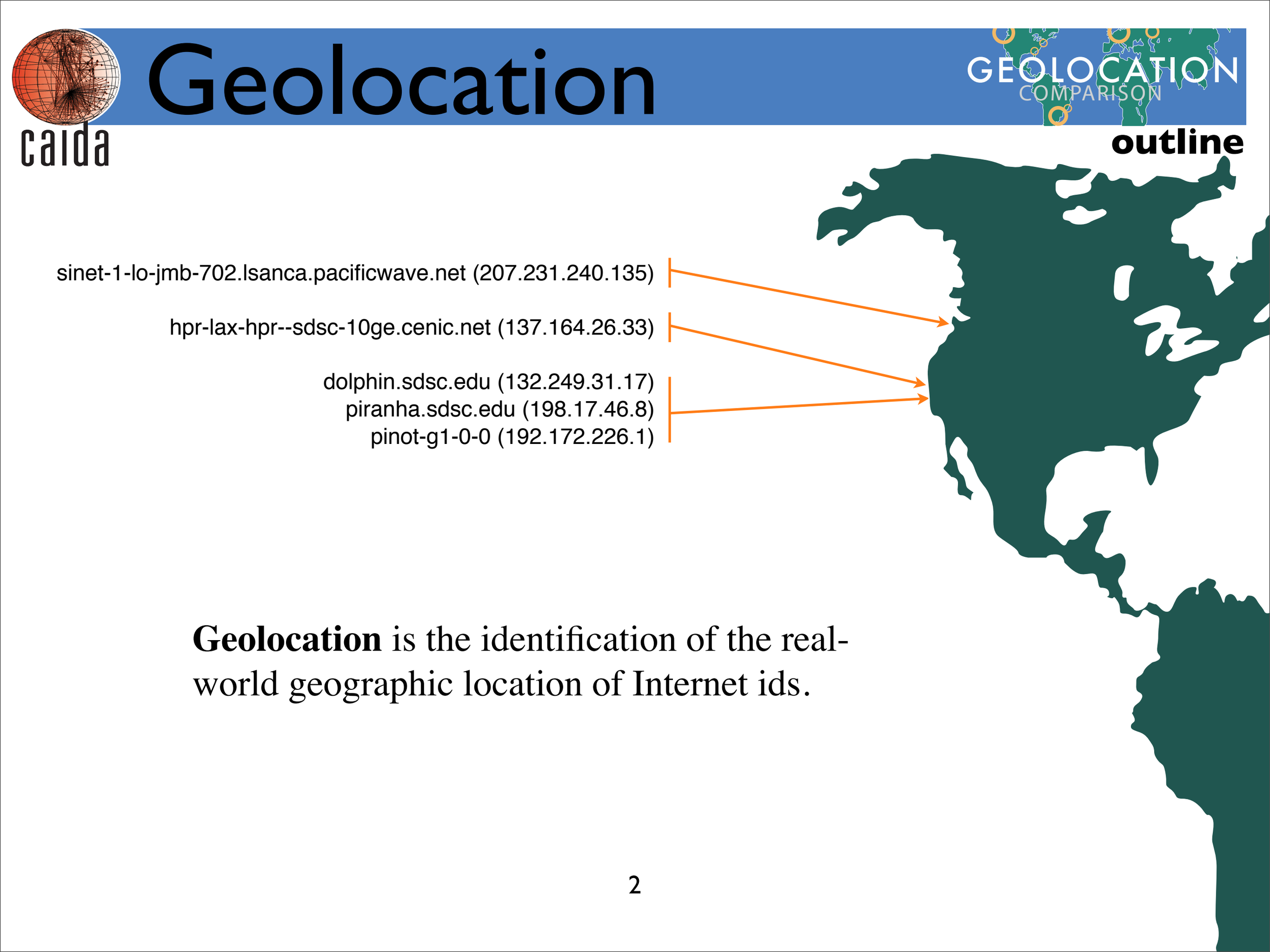
CAIDA's Geolocation Tools Comparison

Bradley Huffaker
bradley@caida.org

CAIDA
University of California
at San Diego, La Jolla, CA

AIM Workshop -- February 2010

Geolocation



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

Outline

- timeline
- NANOG feedback
- data
- proposed process

Timeline

GEOLOCATION COMPARISON

timeline

Jan.	public request for feedback to NANOG (12 responses)
Feb.	discuss feedback at CAIDA's AIMS 2010 workshop
Mar.	development
May	run comparison
Dec.	publish report

Most mentioned Serivces

GEOLOCATION COMPARISON

NANOG feedback

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

- Smaller Services
  - Quova (Quova On Demand)
  - IP2Location (IP2Location)

Interest (NANOG)

NANOG feedback

- content localization
- credit card verification
- taxation purposes
- legal terms of service applicability
- ad targeting
- data privacy requirements
- DRM restricted content
- nearest datacenter

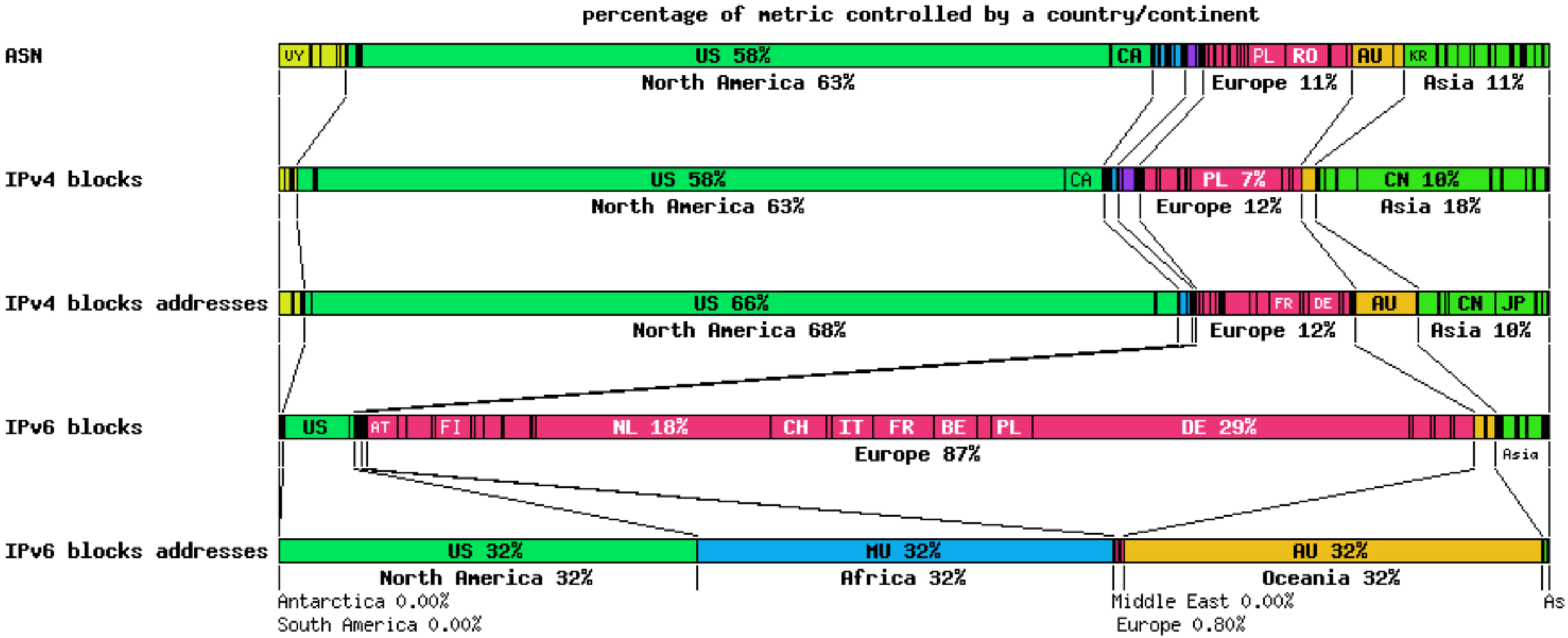
Whois data

- whois allocations
  - for each geolocation service, break each block down into largest continuous block of addresses that have the same location
  - whois provides finer breakdown then BGP prefixes
- whois country level "ground truth"ish
  - not really accurate for large organizations
  - registries only provide country as separate field

Whois breakdown

GEOLOCATION COMPARISON

data



AFRINIC, APNIC, ARIN, LACNIC,RIPE database dumps
Jan. 29, 2010

Topology Data

GEOLOCATION COMPARISON

data

- BGP (Routeviews/RIPE)
  - IPv4 Prefix to AS mapping
- ark router graph
  - IPv4 prefixes / hostnames
- CAIDA's AS relationships
  - classify AS's into categories
    - large transit provider
    - small transit provider
    - enterprise
    - content provider

Ground Truth data

GEOLOCATION COMPARISON

data

	number of IPv4 addresses
Tier I	26k
Tier 2	519
GEANT	299
I-Light	265
Internet 2	317
National LambdaRail	183
CANET	96

If you have ground data, please let me know!
bradley@caida.org

Proposed Process

steps

1. Run each service against the whois allocations, subdivide until all IP addresses with in the suballocation map to the same region.
2. annotate blocks by their AS class

   IP → prefix → AS → class

analysis

- ground truth
  - + compare against ISP locations
  - + compare against whois country
- AS class
  - + how does AS class affect accuracy
  - + how does AS class affect the frequency of block subdivision
- hostnames vs location
  - + check if hostname changes effects geolocation