DNS Infrastructure Distribution

Steve Gibbard
Packet Clearing House
http://www.pch.net/
scg@pch.net

Introduction

- Previous talk on importance of keeping critical infrastructure local.
- Without local infrastructure, local communications are subject to far away outages, costs, and performance.
- Critical infrastructure includes DNS.
- If a domain is critical, so is everything above it in the hierarchy.
- Sri Lanka a case in point.

Root server placement

- Currently 110 root servers(?)
  - Number is a moving target.
- Operated by 12 organizations.
- 13 IP addresses.
  - (At most) 13 servers visible from any one place at any one time.
  - Six are anycasted.
  - Four are anycasted in large numbers.
- All remaining unicast roots are in the Bay Area, Los Angeles, or Washington, DC.

Distribution by continent

- 34 in North America:
  - 8 each in Bay Area/DC Area, 5 in Los Angeles.
  - Only non-costal roots in US are in Chicago and Atlanta.
- 34 in Europe:
  - Clusters of 4 each in London and Amsterdam, Europe's biggest exchanges.
  - Even throughout rest of Europe.

Distribution by continent…

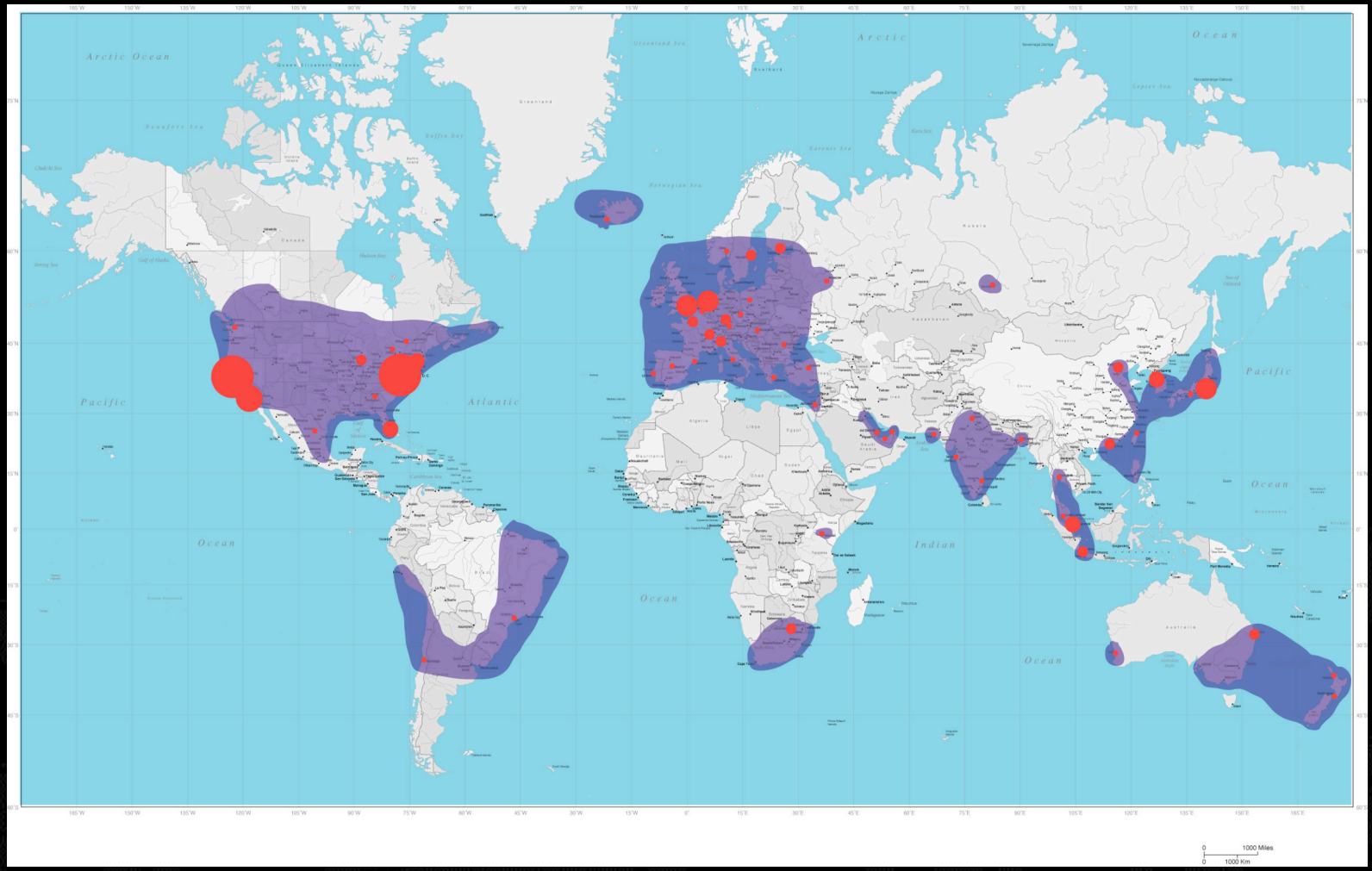
- 26 in Asia (excluding Middle East):
  - 5 in Japan.
  - 3 each in India, Korea, and Singapore.
  - 2 each in Hong Kong, Jakarta, and Beijing.
  - South Asia an area of rapid expansion.
- 6 in Australia/New Zealand:
  - 2 in Brisbane.
  - 1 each in Auckland, Perth, Sydney, and Wellington.

Distribution by continent…

- 5 in Middle East:
  - 1 each in Ankara, Tel Aviv, Doha, Dubai, and Abu Dhabi.
- 3 in Africa:
  - 2 in Johannesburg
  - 1 in Nairobi -- 1 more being shipped.
  - Very little inter-city or inter-country connectivity.
- 2 in South America:
  - Sao Paolo.
  - Santiago de Chile.

Global root server map

Redundant root coverage



Root server expansion

- Four of twelve root server operators actively installing new roots.
- 110 root servers is a big improvement over the 13 that existed three years ago.
- Two operators (Autonomica and ISC) are installing wherever they can get funding.
  - Funding sources are typically RIRs, local governments, or ISP associations.
  - Limitations in currently unserved areas are generally due to a lack of money.

Fs and Is

- In large portions of the world, the several closest roots are Is and Fs.
  - At most two root IP addreses visible locally; others far away.
  - Does this matter?
    - Gives poorly connected regions less ability to use BIND's failure and closest server detection mechanisms.
    - Non-BIND DNS implementations may default to far away roots.
  - Should all 13 roots be anycasted evenly?
    - CAIDA study from 2003 assumed a maximum of 13 locations -- not really relevant anymore.

Big clusters

- Lots of complaints about uneven distribution.
- Only really a concern if resources are finite.
- Large numbers in some places doesn't prevent growth in others.
- Bay Area and DC clusters seem a bit much, but sort of match topology.
- Western Europe's dense but relatively even distribution exactly right.
- Two per city perhaps a good goal for everywhere.

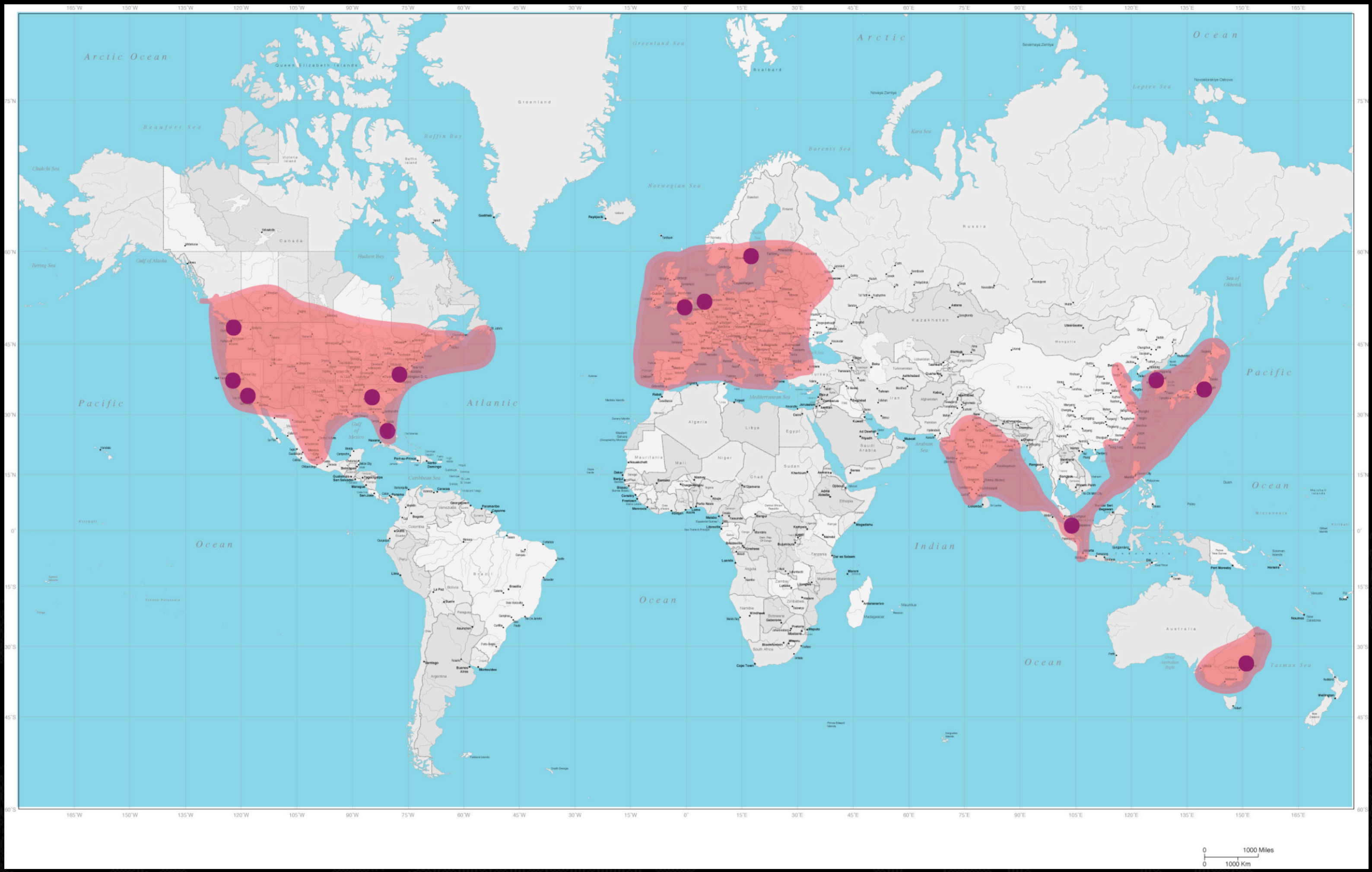
TLD Distribution

- Like the root, Locally used TLDs need to be served locally.
  - Locally used TLDs: Local ccTLD; any other TLDs in common use.
  - Regions don’t need ALL TLDs.

gTLD Distribution: .Com/.Net

- .Com/.Net:
  - Well connected to the “Internet Core.” Servers in Japan, Korea, Netherlands, Sweden, UK; US states of California, Florida, Georgia, Virginia, and Washington.
  - Non-Core locations -- Sydney.

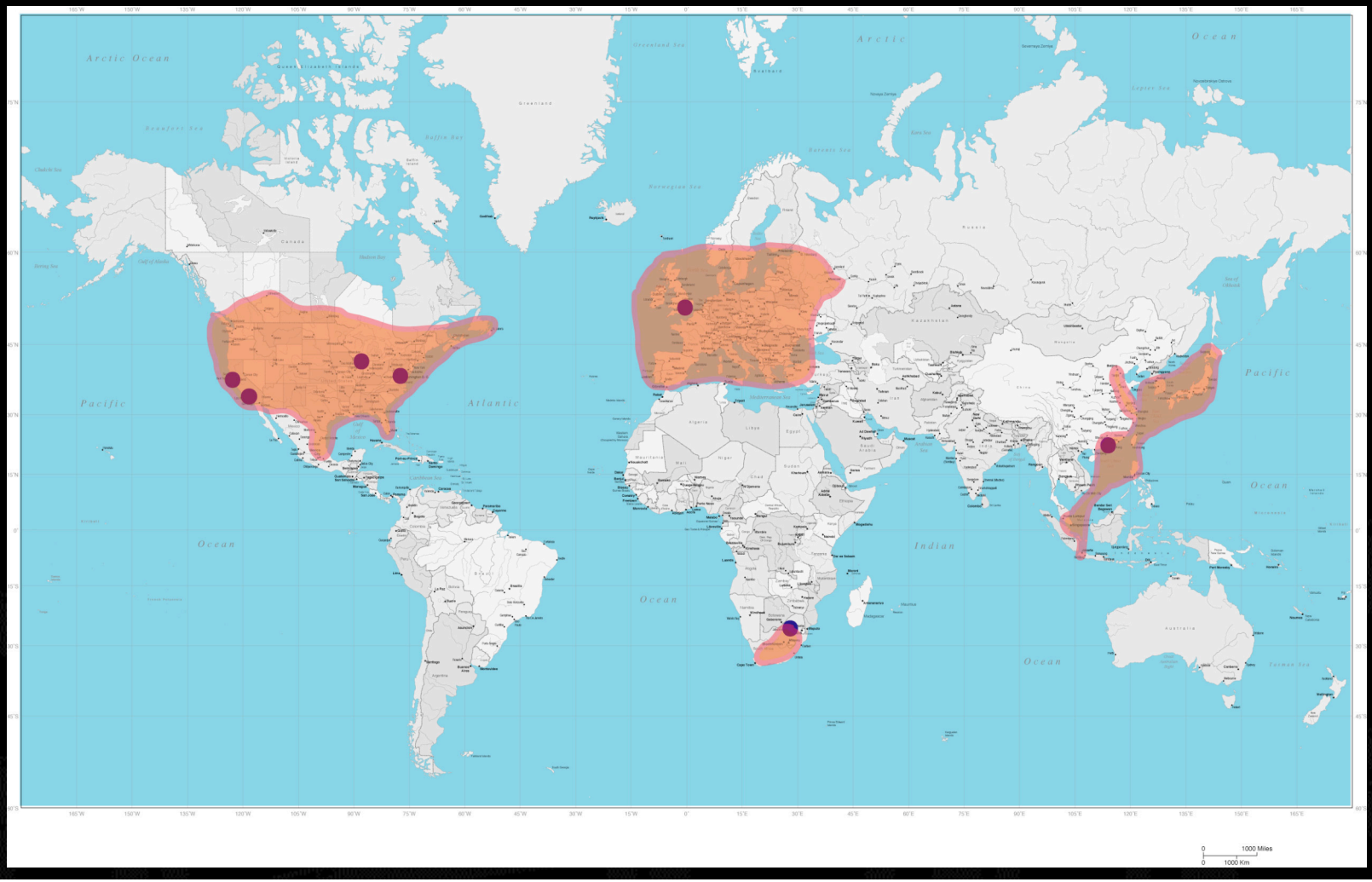
.Com/.Net map


Arctic Ocean
Pacific
Atlantic
Indian
Ocean
1000 Miles
1000 Km

gTLD Distribution: .Org/.Info/.Coop

- .Org/.Info/.Coop:
  - Considered confidential. Data may be incomplete.
  - Significantly fewer publicly visible servers, almost all in “Internet Core:” Hong Kong, UK, South Africa; US: California, Illinois, and Virginia.
  - Only one public location in each of Asia and Europe. No Australia/New Zealand.
  - South Africa outside “Internet Core.”
  - Claims locations reachable only by caching resolvers of some major ISPs. Unspecific claims. Impact hard to judge.

.Org/.Info/.Coop Map



A few other gTLDs:

- .Gov -- US Government: Canada, Germany; US states of California, Florida, New Jersey, Pennsylvania, Texas.
- .Edu -- Universities, mostly US: Netherlands, Singapore, US states of California, Florida, Georgia, Virginia.
- .Int -- International treaty organizations: Netherlands, UK, California.
- .Biz -- .Com competitor: Australia, Hong Kong, Netherlands, New Zealand, Singapore, UK, US states of California, Florida, Georgia, New York, Virginia, Washington.
- Complete listing in the paper.

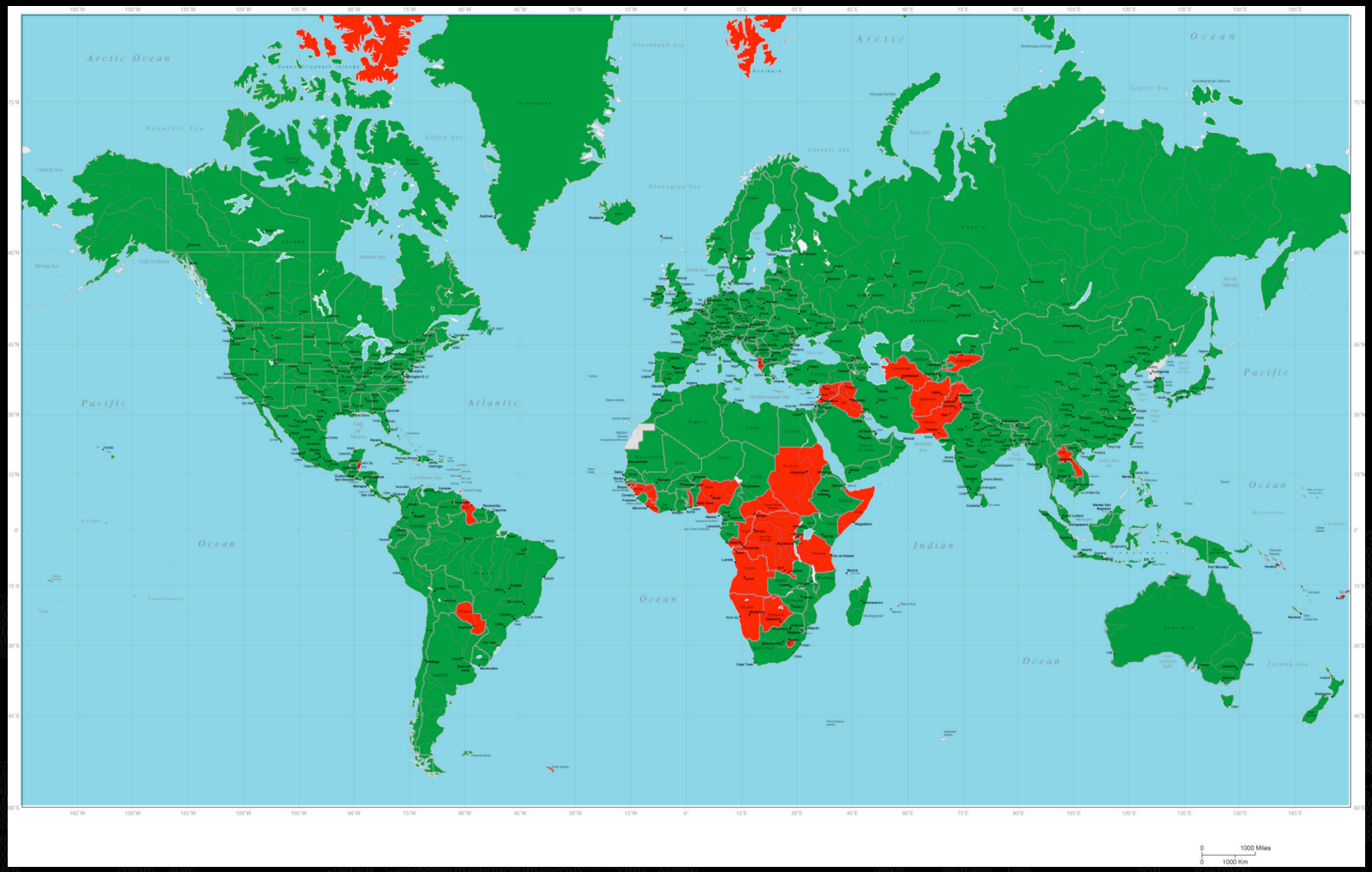
Where should gTLDs be?

- Presumably depends on their market.
- If it's ok for large portions of the world to not use the gTLDs, it's ok for those gTLDs to not be hosted there.
- Really a question for ICANN and the registries.
- .Int's lack of international coverage seems strange.

ccTLD Distribution:

- The answers to where various ccTLDs should work seem much more obvious.
  - Working in their own regions a must.
  - Working in the Internet core, and in regions they communicate with a big plus.
- Just over 2/3 of ccTLDs are hosted in their own countries.
  - (but a lot of those that aren't are for really tiny countries).

Countries with local ccTLDs



ccTLDs not slaved in core

- 18 ccTLDs aren't slaved in the global core.
- If their regions get cut off, those ccTLDs won't be visible to the rest of the world.
- Is this an issue?
  - Certainly, if these ccTLDs are used to address resources outside their regions or not connected to the core the same way.
  - A cause of misleading failure modes for incoming communications. A clear RFC 2182 violation.
  - Not an issue if communications from outside don't matter.

ccTLDs not hosted in core

- .BB -- Barbados
- .BD -- Bangladesh
- .BH -- Bahrain
- .CN -- China
- .EC -- Ecuador
- .GF -- French Guiana
- .JM -- Jamaica
- .KG -- Kyrgyzstan
- .KW -- Kuwait
- .MP -- Northern Mariana Islands
- .MQ -- Martinique
- .MV -- Maldives
- .PA -- Panama
- .PF -- French Polynesia
- .QA -- Qatar
- .SR -- Suriname
- .TJ -- Tajikistan
- .ZM -- Zambia

Example countries

- Kenya
  - Exchange point, root server, ccTLD server, all external connectivity by satellite.
- Pakistan:
  - Root server, no exchange point, no TLDs.

Kenya

- Kenya:
  - Local exchange point in Nairobi.
  - Local root server in Nairobi.
  - Local .ke ccTLD servers.
  - No external fiber.
  - Local users accessing local services in the .ke domain have their queries stay local and should be reliable. Queries to non-local TLDs depend on satellite connectivity, which may not be working.

Pakistan

- Pakistan:
  - Local root server (for at least one ISP).
  - No TLDs.
  - .pk hosted entirely in the US.
  - Root queries may get answered locally, but get followed by long distance queries for .pk, ten timezones away.
  - .Com queries go to Singapore or Europe, a bit closer.
  - Single fiber connection, so if that breaks, no TLD lookups are possible. Root server not a huge benefit.

Local peering caveat

- Local traffic has to be kept local before keeping DNS local is much of an issue.
  - If DNS queries have to leave the region and come back, that doubles the problems created by queries merely needing to leave.
  - This generally requires either a local exchange point or monopoly transit provider.
- Examples used here have already taken care of that.
- I haven't done that research on the rest of the world yet.

Methodology

- Get name server addresses for TLDs
- Assume everything in a /24 is in the same place or set of places.
  - Bad assumption for UUNet servers. Didn't find any other problems. May have missed some.
  - 625 /24s contain name servers for TLDs. 135 host multiple TLDs; over 60 in RIPE's case.
- Figure out where those subnets are:
  - Do lots of traceroutes, and ask lots of questions.
  - Automated geolocation systems tended to be wrong.

Subnets with 10+ TLDs

193.0.12/24	RIPE	62	Amsterdam
192.36.125/24	SUNET/NS.SE	38	Stockholm
204.152.184/24	ISC	36	Palo Alto
198.6.1/24	UUNet	33	Various US locations
137.9.1/24	UUNet	27	Various US locatoins
193.0.0/24	RIPE	26	Amsterdam
147.28.0/24	PSG	23	Seattle
204.74.112/24	UltraDNS	20	Anycast
192.93.0/24	NIC.FR	19	Paris
204.74.113/24	UltraDNS	18	Anycast
192.134.0/24	NIC.FR	15	Paris
204.61.216/24	PCH	14	Anycast
202.12.28/24	APNIC	13	Tokyo

Other sources

- www.root-servers.org had root server data. I assume it to be accurate.
- UltraDNS considers its locations confidential. Got info from Afilias's .Net application. Verified with traceroutes. I'm told I missed some sites.
- In general, TLD operators were very helpful. Thanks!

Thanks!

Full paper at
http://www.pch.net/resources/papers/infrastructure-distribution/

Corrections and updates would be appreciated

Steve Gibbard
Packet Clearing House
scg@pch.net