Building and operating a global DNS anycast network

ENOG 14

Gael Hernandez

Packet Clearing House (PCH)

Minsk, 10 October 2017

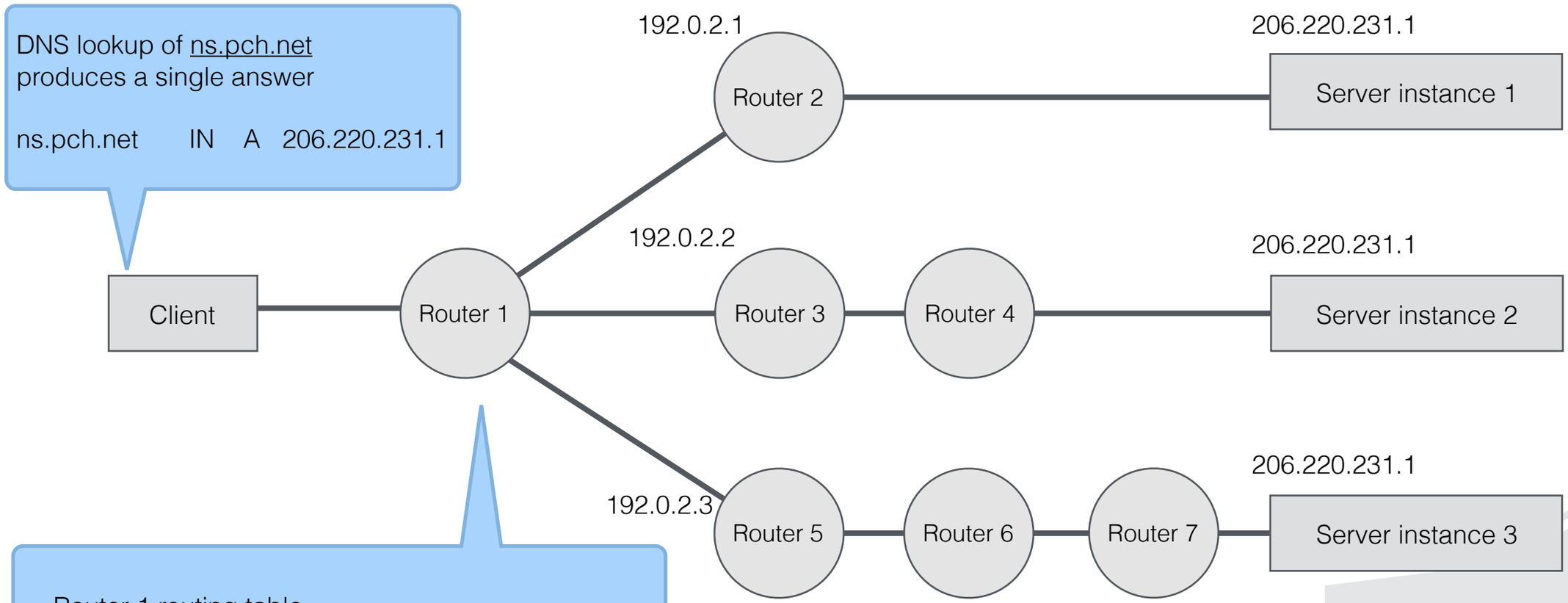
Anycast technology

- An anycast cloud is a distributed cluster of identical instances of a server, each typically containing identical data, and capable of servicing requests identically.

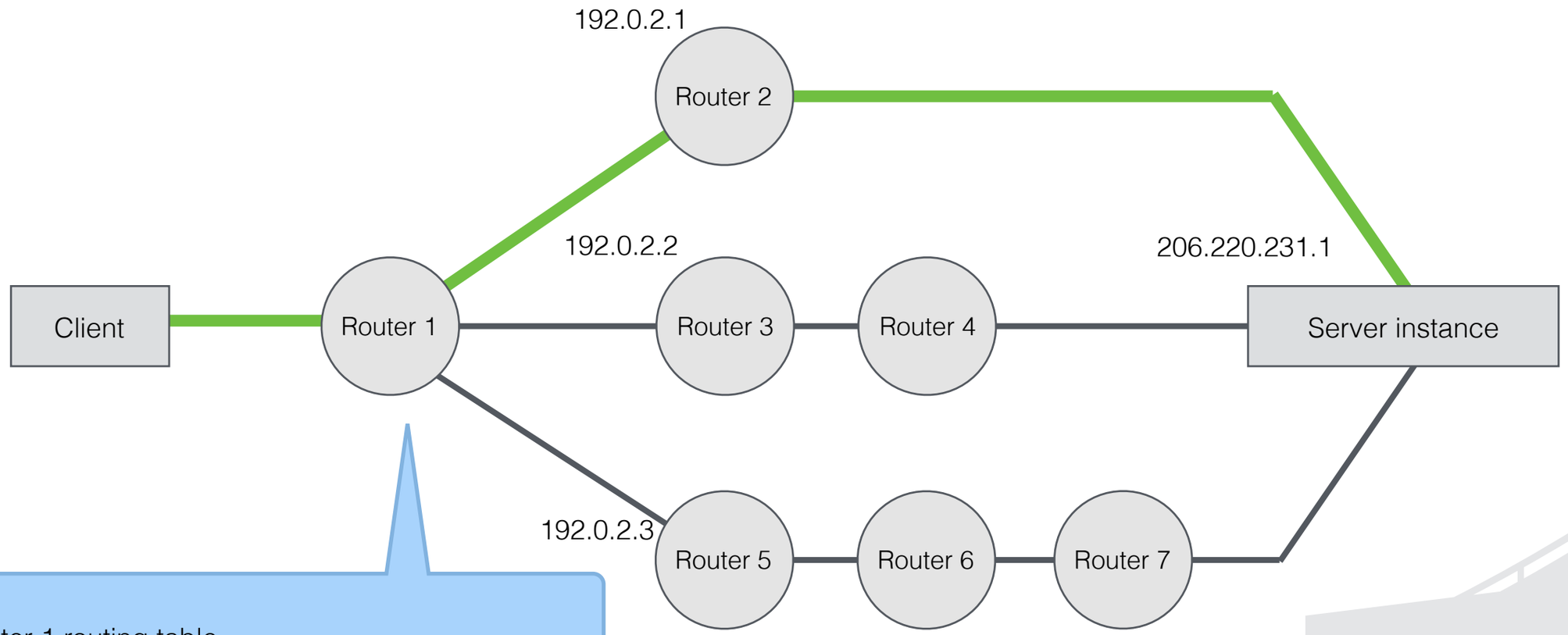
- Each instance has a regular unique globally routable IP address for management purposes, but… each instance also shares an IP address in common with all the others.

- The Internet's global routing system (BGP) routes every query to the instance of the anycast cloud that is closest in routing terms to the user who originated the query.

Anycast technology (ii)



Anycast technology (iii)



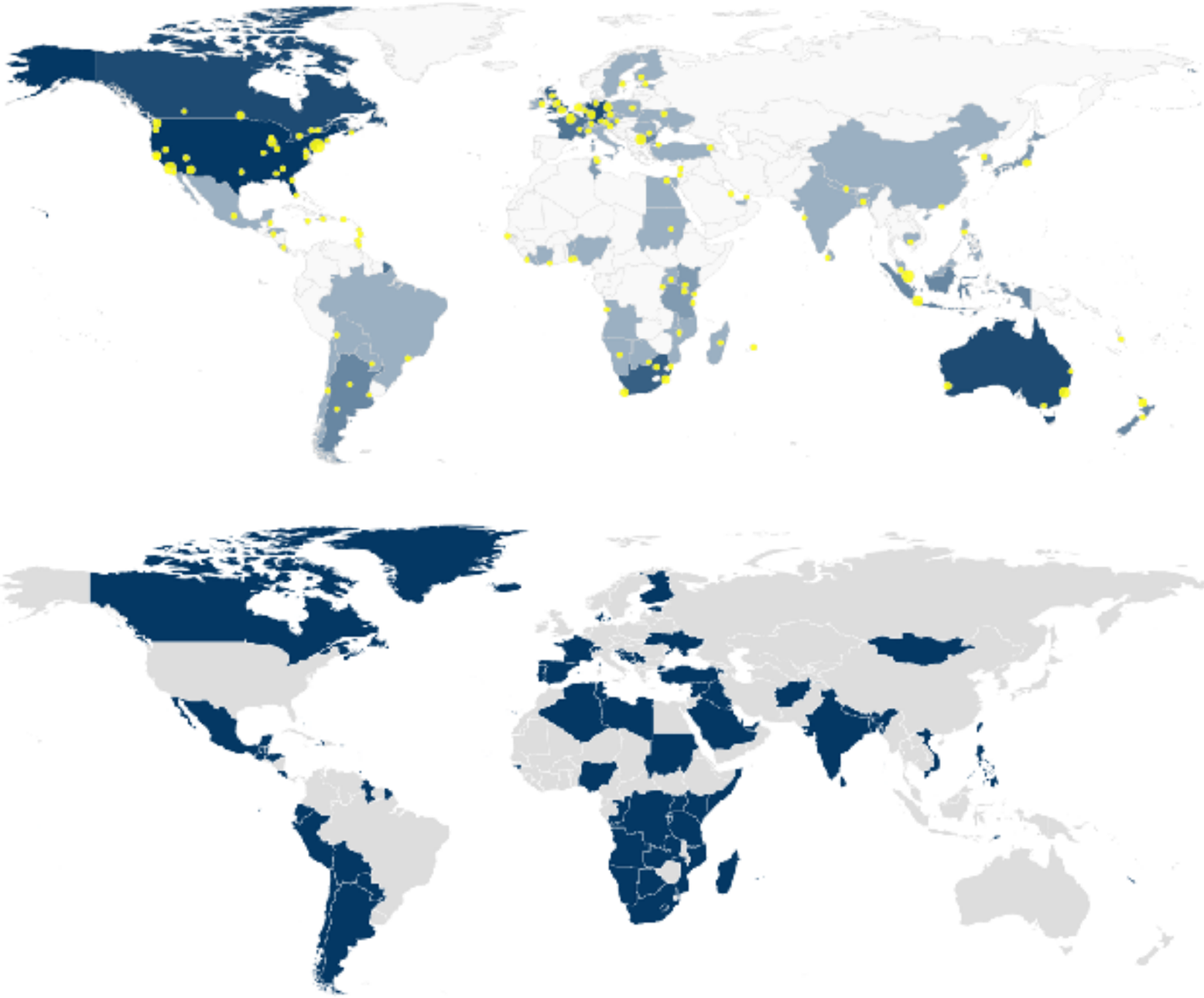
Router 1 routing table

Destination	Mask	Next-Hop	Distance
192.0.2.0	/29	127.0.0.1	0
206.220.231.1	**/32**	**192.0.2.1**	**1**
206.220.231.1	/32	192.0.2.2	2
206.220.231.1	/32	192.0.2.3	3

Anycast for DNS

- PCH and its precursors have run production anycast services since 1989.
- Bill Woodcock (PCH) and Mark Kosters (then at Verisign) first proposed the idea of anycasting authoritative root and TLD DNS at the Montreal IEPG in 1995.
- PCH began operating production anycast for ccTLDs and in-addr zones in 1997.
- PCH first hosted an anycast production of a root name server in 2002.
- We operate services through IPv6 since 2000.

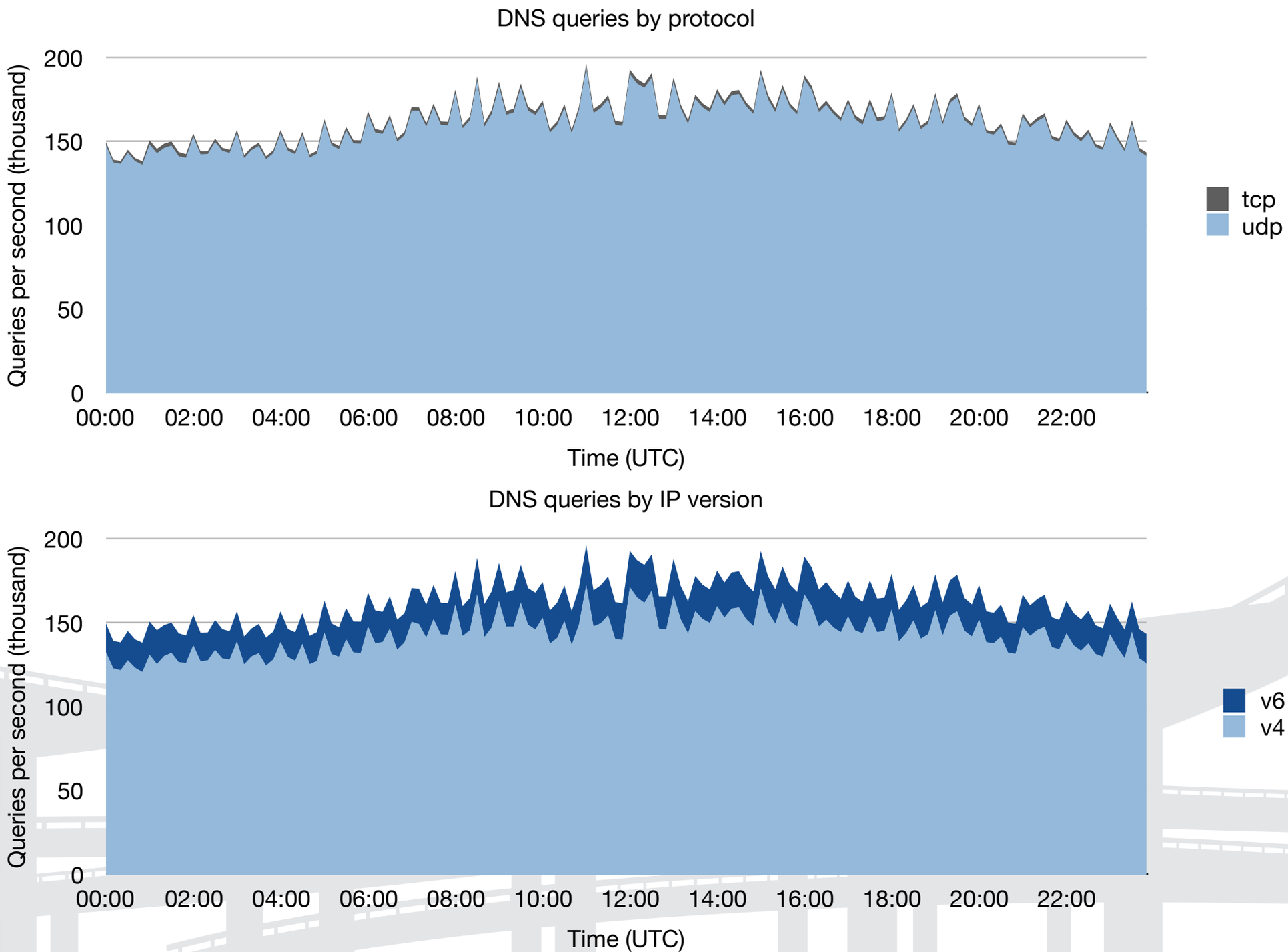
PCH's Anycast Cloud (AS42)



- 118 anycast nodes in all five continents
  - 14 global nodes + 4 high traffic nodes
- 152 locations in five continents
  - 33 in ARIN region
  - 28 in RIPE region
  - 25 in AFRINIC region
  - 18 in APNIC region
  - 14 in LACNIC region
- 2,691 unique ASN peers
  - 150 route-servers ASN
- Secondary authoritative service to 400+ TLDs and two letters of the DNS root.
  - ~105 ccTLDs
  - ~120 million resource records

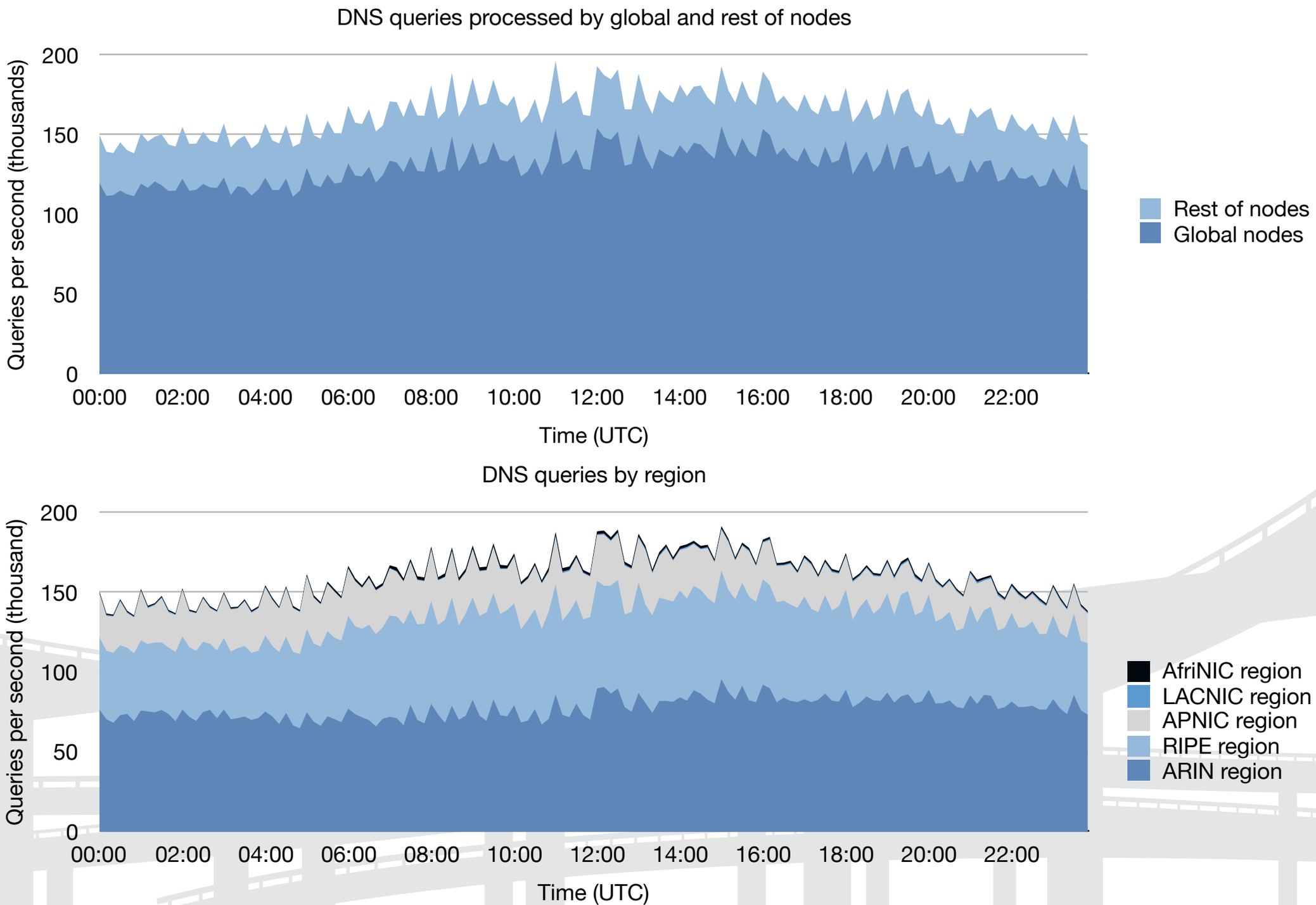
Packet Clearing House

A day in PCH's anycast network



Packet Clearing House

A day in PCH's anycast network (ii)



Planning Anycast Nodes

- Anycast is a robust and well-proven technology: it works!
  - **E-root** is the fastest in the U.S., South Africa, Poland, Ireland, and Malaysia and **D-root** is the fastest in the U.K., Netherlands, Austria and Thailand (Thousand Eyes, June 2017)

- Considerations when planning for new sites
  - Invitation from an IX operator to host a DNS node
  - Traffic levels, number of participants and prefixes at the IX
  - Availability of our transit providers
  - Relative location of other nodes

- Delivering content in some regions is challenging
  - Less developed interconnection market in emerging economies
  - Absence of open and neutral exchanges with public peering
  - Large networks won't be peering at small exchanges

Operations

- Services run in separated virtual machines
  - Dedicated VMs for root servers, TLDs and monitoring services.

- Depending on the type of deployment (small/medium/large) and type of node (local/global), we announce via BGP a full or a partial set of services:
  - Small sites: anywhere in the world, local-only and partial service announcements.
  - Medium sites: medium to high-volume locations, local-only and partial service announcements.
  - Full sites: global nodes in high volume locations, with full service announcements via our transit providers (NTT and Level3).

- A failure in the DNS service triggers the removing of the node from the routing table by stopping its BGP announcement

Monitoring

- Multiple layers of monitoring to proactively detect issues that could be leading to a degradation of the service
  - Hardware layer: CPU levels, temperature, RAM.
  - Interconnection layer: ports and traffic levels.
  - Routing layer: AS-PATH and prefix announcements.
  - Service layer: queries per second, replies per second.

- Passive monitoring tools
  - Nagios with custom plugins for DNS and DNSSEC
  - Netflow monitoring traffic levels

- Active monitoring of global performance using RIPE Atlas and RIPE DNSMon measurements on a regular basis

Questions?

Thanks for your attention

Gael Hernandez
Senior Manager, Interconnection Policy and Regulatory Affairs
gael@pch.net