Anycast, anywhere..!

APTLD72

Gael Hernandez

Packet Clearing House

Tbilisi, 15 September 2017

Who are we?

- Packet Clearing House (PCH) is the global non-profit organisation providing operational support and security to critical Internet infrastructure, including Internet exchange points and the core of the DNS, since 1993.

- Funded by government grants, service-provision fees from the Internet operations industry and specialised consultancies on IXP construction/operation and capacity building.

- Global footprint with head office in San Francisco (US) and regional offices in Buenos Aires, Johannesburg, Kathmandu and Dublin.

Our work

- Participated in the construction of more than half of the Internet exchange points worldwide.

- Operational support and security to the core of the Domain Name System through our anycast platform and DNSSEC signing platform.

- Training and capacity building in the areas of routing, Internet economics, DNS operations, policy and regulatory, Internet governance, etc.

- Cyber-security co-ordination: operation of the global internet infrastructure protection system INOC-DBA (looking for sponsorship!).

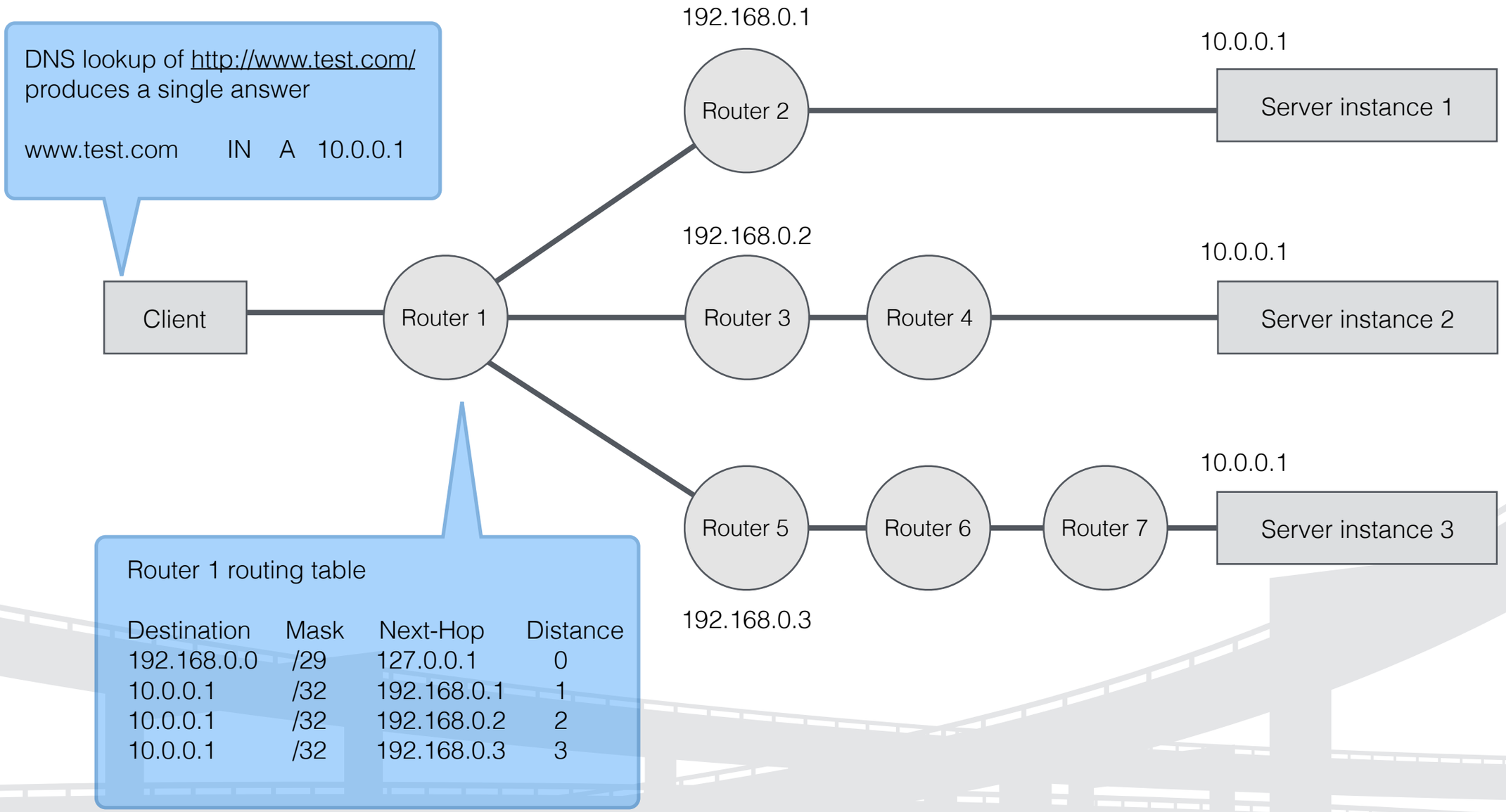
Anycast technology

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

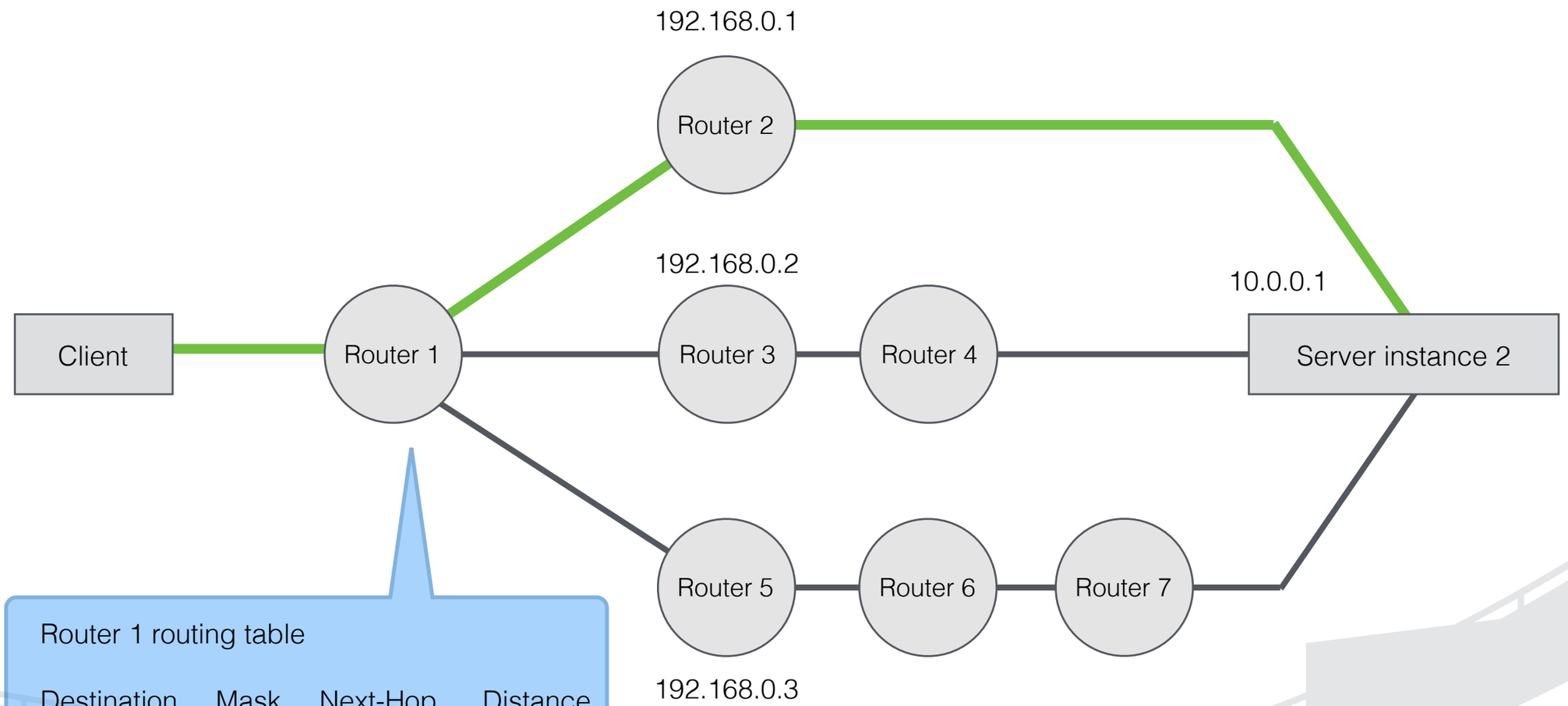
- Each instance has a regular unique globally routable IP address for management purposes, but… each instance also shares an IP address in common with all the others.

- The Internet's normal 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
192.168.0.0	/29	127.0.0.1	0
10.0.0.1	**/32**	**192.168.0.1**	**1**
10.0.0.1	/32	192.168.0.2	2
10.0.0.1	/32	192.168.0.3	3

Anycast for DNS

- PCH and its precursors have run production anycast services since 1989.

- Bill Woodcock (PCH) and Mark Kosters (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-critical infrastructure in-addrs in 1997, with 100% up-time over more than eleven years.

- PCH first hosted an anycast production of a root nameserver in 2002. We operate services through IPv6 since 2000.

PCH's Anycast Network in figures

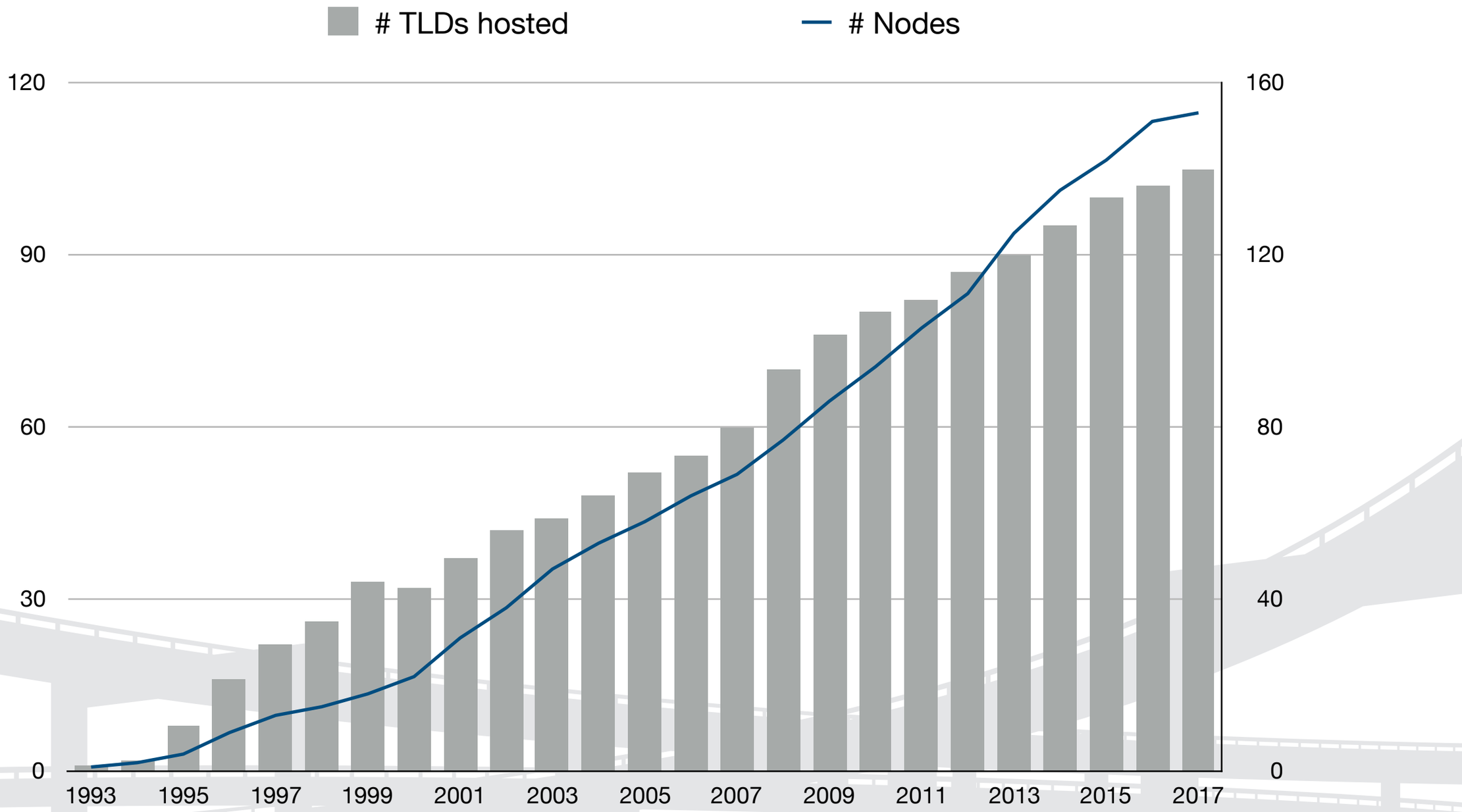
- PCH operates the world's largest anycast server cloud with 28 years of production experience.

- We operate 118 anycast nodes in 152 locations in all five continents
  - 14 global nodes + 4 high traffic nodes

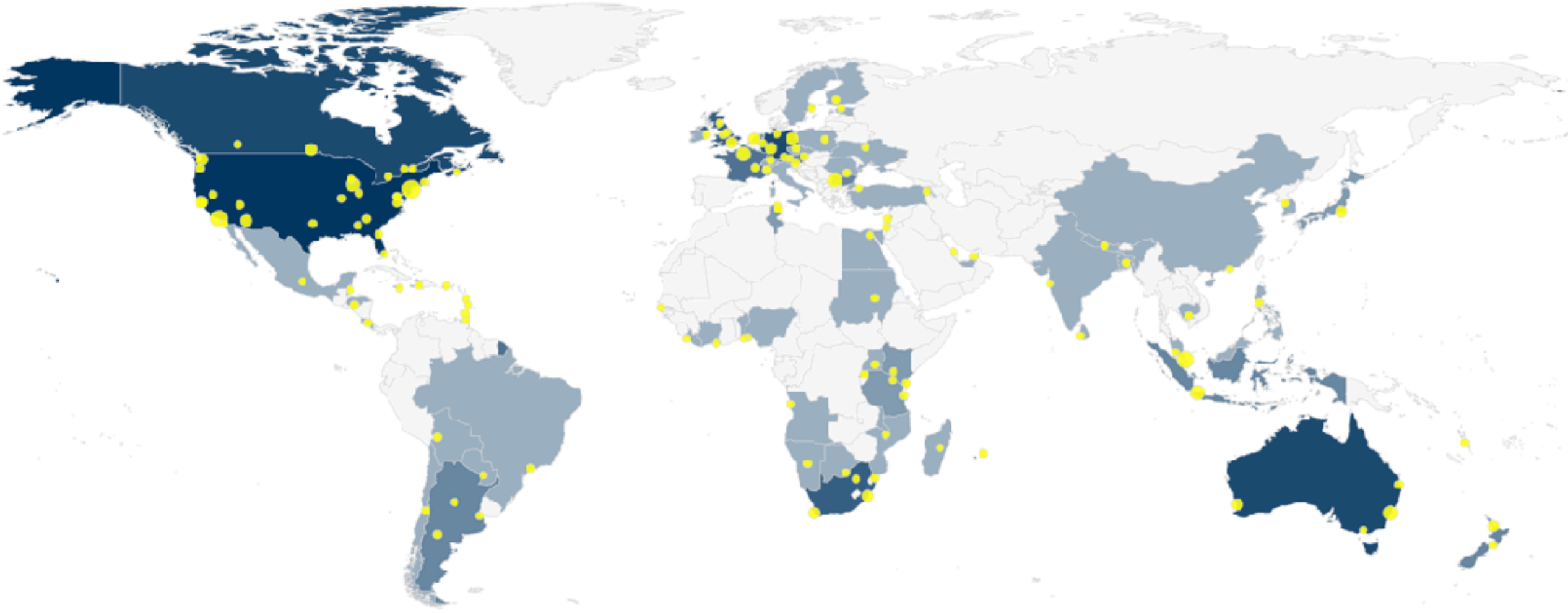
- Our infrastructure provides secondary service to gTLDs, ccTLDs, in-addr zones and two letters of the DNS root.
  - 400+ TLDs and ~105 ccTLDs
  - ~120 million resource records (RR).

Packet Clearing House

Evolution of PCH's Anycast Network



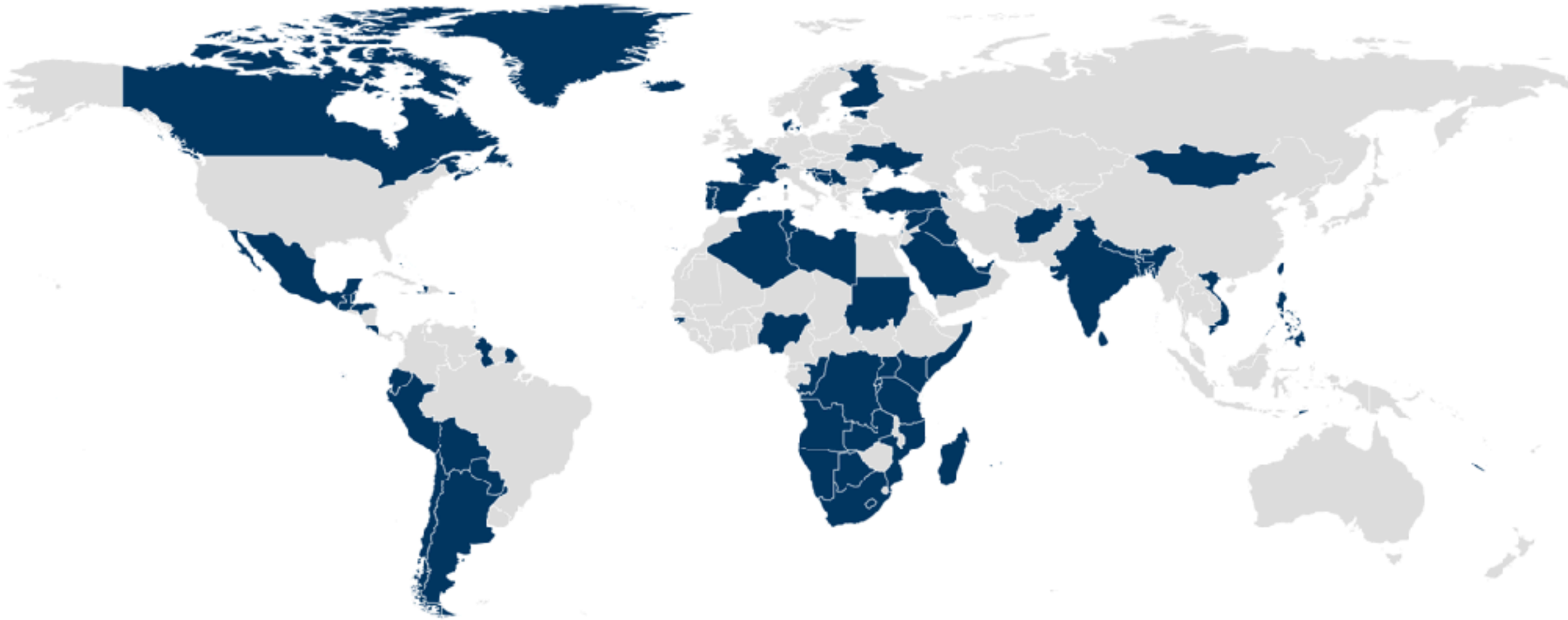
Current global footprint



PCH
Packet Clearing House

PCH
Packet Clearing House

Our ccTLDs customers



PCH
Packet Clearing House

PCH 8th-Generation Architecture

- Small (2U), medium (5U) and full (7U) installations
- Routing vendor redundancy: Cisco and Quagga.
- Cisco servers with 48G, 192G y 256G of RAM memory.
- VMware ESX clusters, supporting any X86 32-bit or 64-bit OS.
- Hosted servers fully integrated with BGP routing architecture.
- OS redundancy: Solaris and CentOS.
- DNS redundancy: BIND and NSD.

- Long-term strategic relationships with all involved vendors: Cisco, AMD, Sun, VMware, ISC, and NLNet Labs.

Site selection and planning

- Anycast is a robust and well-proven technology: it just works great!

- Load-balancing in some regions can be challenging
  - A less developed interconnection market in emerging economies
  - Absence of neutral and open IXPs
  - Large networks wont be peering at small IXs

- Considerations when planning anycast nodes:
  - 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

Operations

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

- Depending on the type of deployment (small/medium/large) and type of node (local/global), we will be announcing a full or partial set of services:

  - Small sites (~58): anywhere in the world, local-only and partial service announcements.
  - Medium sites (~38): medium to high-volume locations, local-only and partial service announcements.
  - Full sites (~22): high volume locations, mostly global sites with full service announcements.

Monitoring

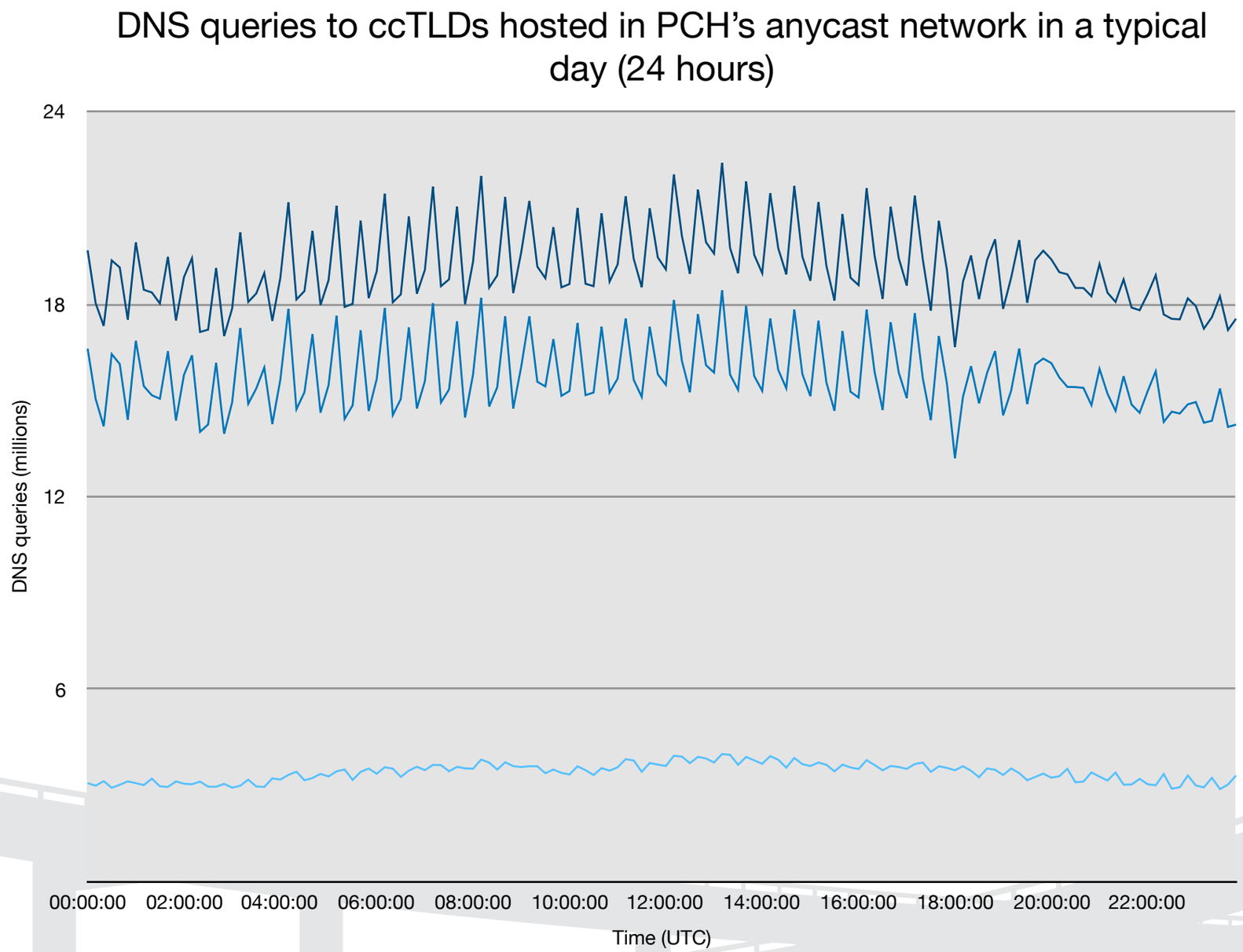
- Multiple layers of monitoring to proactively detect issues that could be leading to a degradation on the service
  - Hardware layer: CPU levels and interfaces.
  - Interconnection layer: ports and traffic levels
  - Routing layer: AS-PATH and prefix announcements
  - Service levels: queries per second,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

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

- Active monitoring of service performance via RIPE Atlas and RIPE DNSMon measurements

Packet Clearing House

A day in PCH's anycast network



What keeps us awake?

- UDP spoofing and networks not implementing BCP38
- Network operators doing too much traffic engineering
- Critical zero-day exploits affecting name servers or other critical software

Packet Clearing House

Things we're working on…

- Better statistics dashboard for our customers
- Research lab work benchmarking alternatives for our current name server software: for example Knot

Questions?

Thanks for your attention

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