DNS, NAT, Congestion Control

29 January 2025
Lecture 12

Slides Credits: Steve Zdancewic (UPenn)

Topics for Today

- DNS
- Network Address Translation (NAT)
- Congestion Control
  - Queuing

- Sources:
  - DNS: PD 9.3.1
  - NAT: PD 4.3
  - Congestion Control PD 6.1-6.2

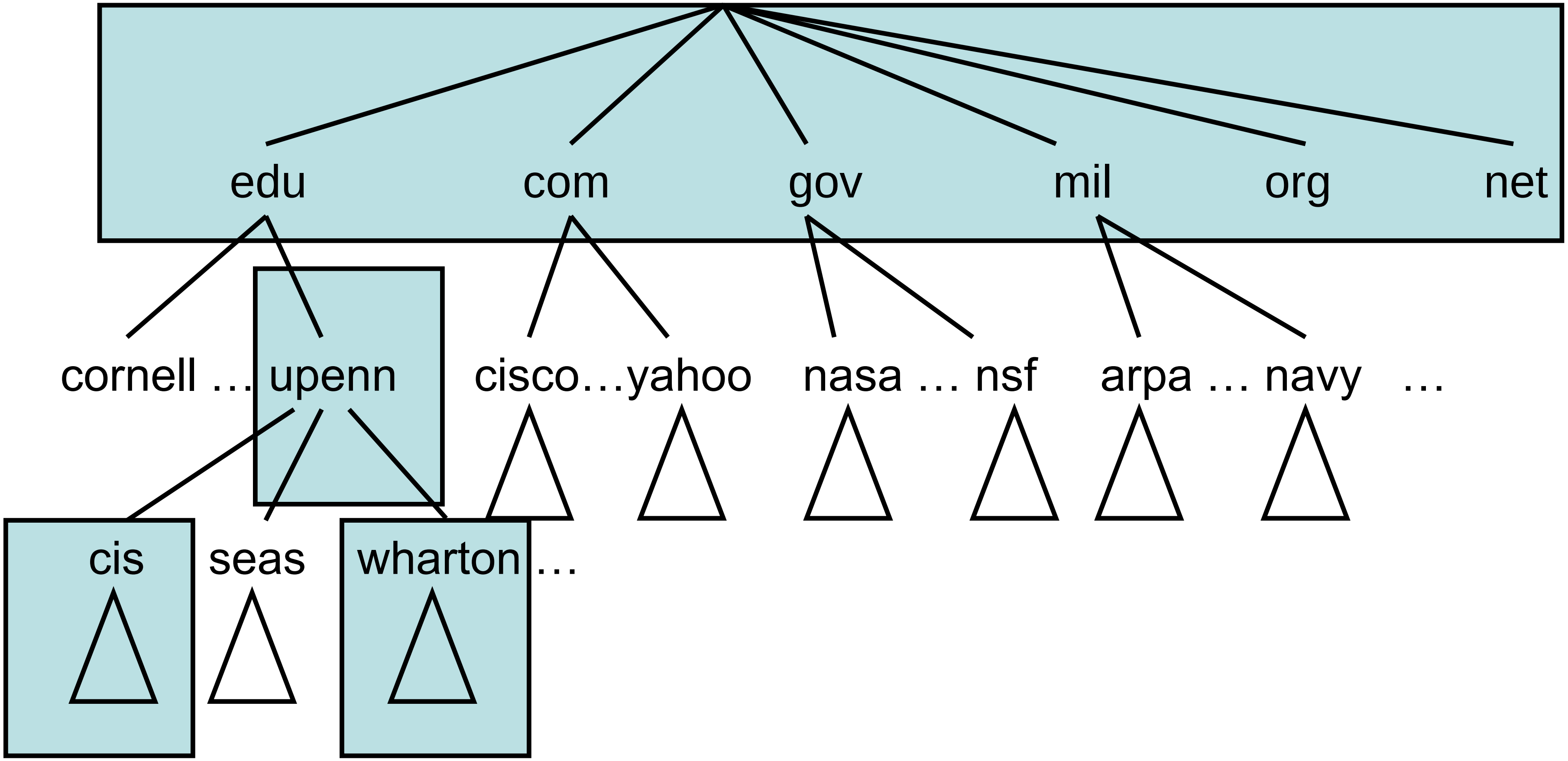
Domain Name System

- System for mapping mnemonic names for computers into IP addresses.

  softwares.kinneret.ac.il → 172.67.23.145

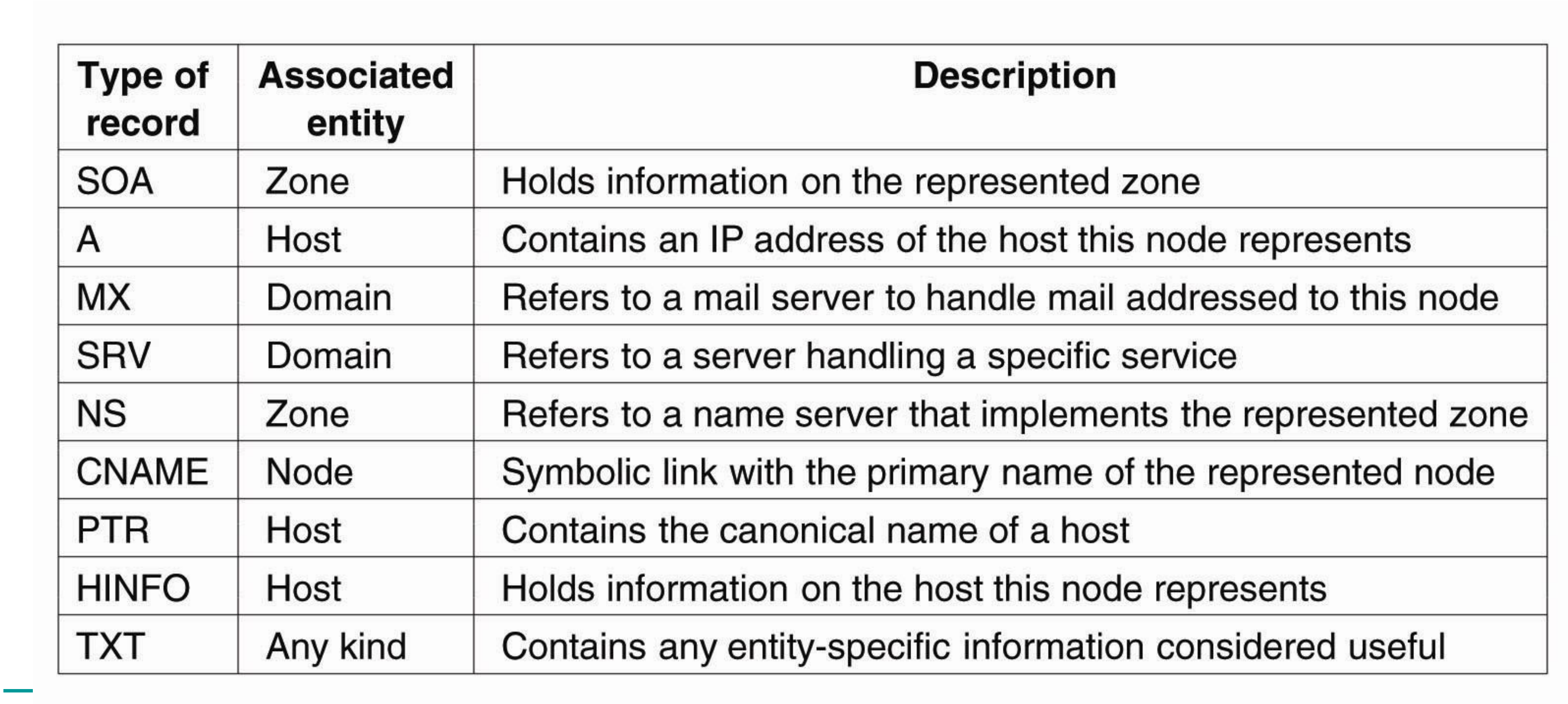
- Domain Hierarchy
- Name Servers
  - 13 Root servers map top-level domains such as ".com" or ".net"
- Name Resolution
  - Protocol for looking up hierarchical domain names to determine the IP address
  - Protocol runs on UDP port 53

Domain Name Hierarchy



DNS Records

- The most important types of resource records forming the contents of nodes in the DNS name space.



Type of record	Associated entity	Description
SOA	Zone	Holds information on the represented zone
A	Host	Contains an IP address of the host this node represents
MX	Domain	Refers to a mail server to handle mail addressed to this node
SRV	Domain	Refers to a server handling a specific service
NS	Zone	Refers to a name server that implements the represented zone
CNAME	Node	Symbolic link with the primary name of the represented node
PTR	Host	Contains the canonical name of a host
HINFO	Host	Holds information on the host this node represents
TXT	Any kind	Contains any entity-specific information considered useful

Excerpt from the DNS database for the zone *cs.vu.nl*.

Name	Record Type	Record Value

cs.vu.nl.	SOA	primary name server = star.cs.vu.nl
		responsible mail addr = hostmaster.cs.vu.nl
		serial = 2022112500
		refresh = 7200 (2 hours)
		retry = 3600 (1 hour)
		expire = 2419200 (28 days)
		default TTL = 7200 (2 hours)

cs.vu.nl	TXT	"v=spf1 redirect=vu.nl"
cs.vu.nl	TXT	"google-site-verification=Hgkj69rep7_FHZsXaTOoO8JxO6e9XUpK1aeNqPKUo7I"
cs.vu.nl	NS	ns1.labs.vu.nl
cs.vu.nl	NS	ns0.labs.vu.nl
cs.vu.nl	NS	ns2.labs.vu.nl
cs.vu.nl	NS	new-ns1.vu.nl
cs.vu.nl	NS	new-ns2.vu.nl

Excerpt from the DNS database for the zone *cs.vu.nl*.

Name	**Record Type**	**Record Value**
ns0.labs.vu.nl	A	192.31.231.42
ns1.labs.vu.nl	A	130.37.192.252
ns2.labs.vu.nl	A	130.37.192.254
new-ns1.vu.nl	A	130.37.164.20
new-ns2.vu.nl	A	130.37.164.22
ns0.labs.vu.nl	AAAA	2001:610:110:6e0::2a
ns1.labs.vu.nl	AAAA	2001:610:110:6e0::1:0
ns2.labs.vu.nl	AAAA	2001:610:110:6e0::1:2
cs.vu.nl	MX	0 cs-vu-nl-mail.protection.outlook.com
star.cs.vu.nl	A	192.31.231.42
zephyr.cs.vu.nl	HINFO	“CPU = Sun OS = Unix”
ftp.cs.vu.nl	CNAME	soling.cs.vu.nl

Excerpt from the DNS database for the zone *cs.vu.nl*.

Name	Record Type	Record Value
www.cs.vu.nl	CNAME	papac022.vu.nl
papac02.vu.nl	A	130.37.164.171
inkt.cs.vu.nl	A	192.168.4.3
inkt.cs.vu.nl	HINFO	“CPU = OCE OS = Proprietary”
pen.cs.vu.nl	HINFO	“CPU = OCE OS = Proprietary”
pen.cs.vu.nl	A	192.168.4.2

Kinneret DNS Records (1/2)

- An excerpt from the DNS database for zone kinneret.ac.il

kinneret.ac.il	NS	kineret.kinneret.ac.il
kineret.kinneret.ac.il	A	212.150.112.60
kinneret.ac.il	NS	ns2.kinneret.ac.il
ns2.kinneret.ac.il	A	212.150.112.59
kinneret.ac.il	A	88.218.117.88

Kinneret DNS Records (2/2)

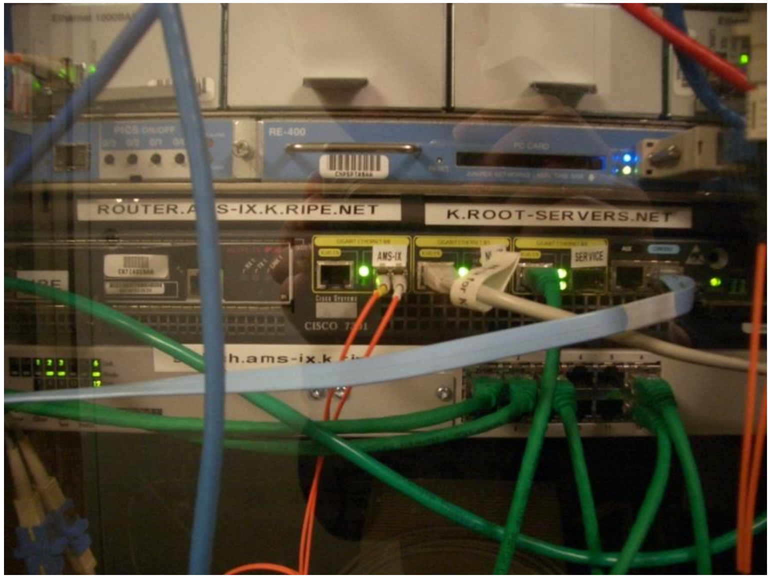
- An excerpt from the DNS database for zone kinneret.ac.il

kinneret.ac.il	MX	10 mail-secure.kinneret.ac.il

kinneret.ac.il	SOA	origin = kineret.kinneret.ac.il
		mail addr = mordo.kinneret.ac.il
		serial = 2024053124
		refresh = 7200
		retry = 3600
		expire = 2419200
		minimum = 3600

mail-secure.kinneret.ac.il	A	172.25.1912.1

DNS Roots



Root server K in Amsterdam, Holland (Wikipedia)

ICANN is responsible for managing roots and top level domains

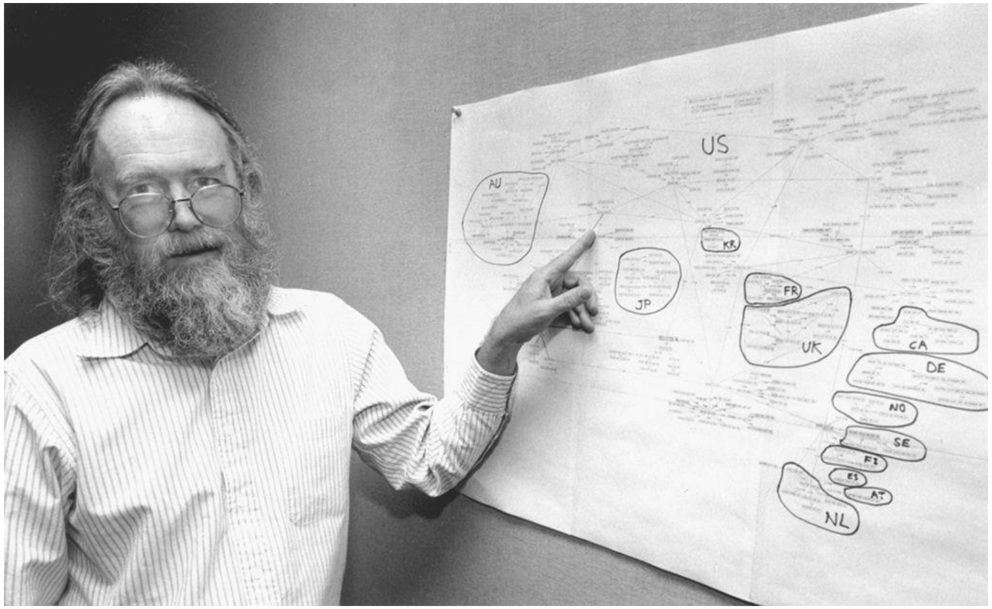
- 13 DNS root servers heavily replicated around the world
  – 12 independent orgs run the roots

Distributed Control (DNS)

Jan 1998: Jon Postel of IANA told 8 of the 12 roots at the time to contact IANA’s root copy instead of the US government’s root copy (Network Solutions, Inc. in Herndon, VA)

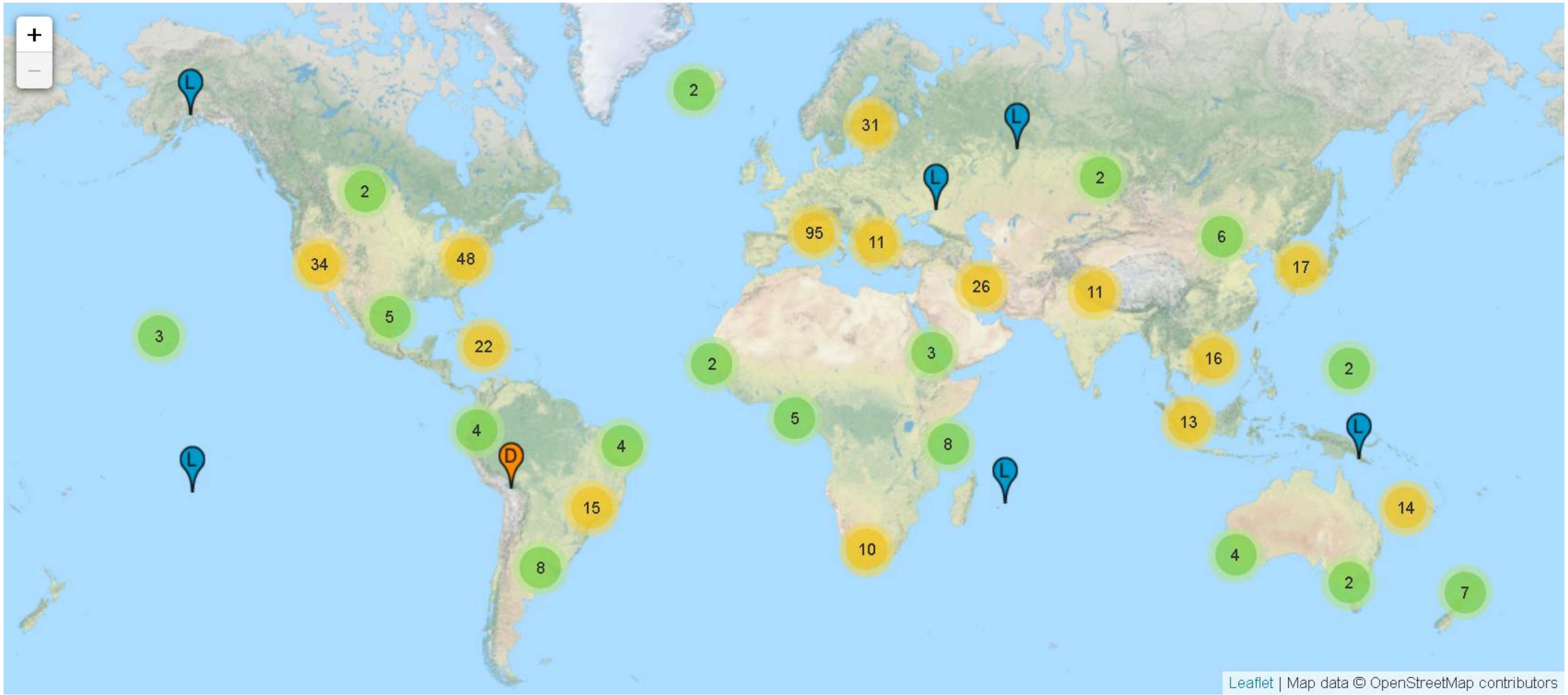
– Postel said it was a test and changed it back when asked (?)

– Sept 1998 – ICANN is formed and takes over IANA’s job

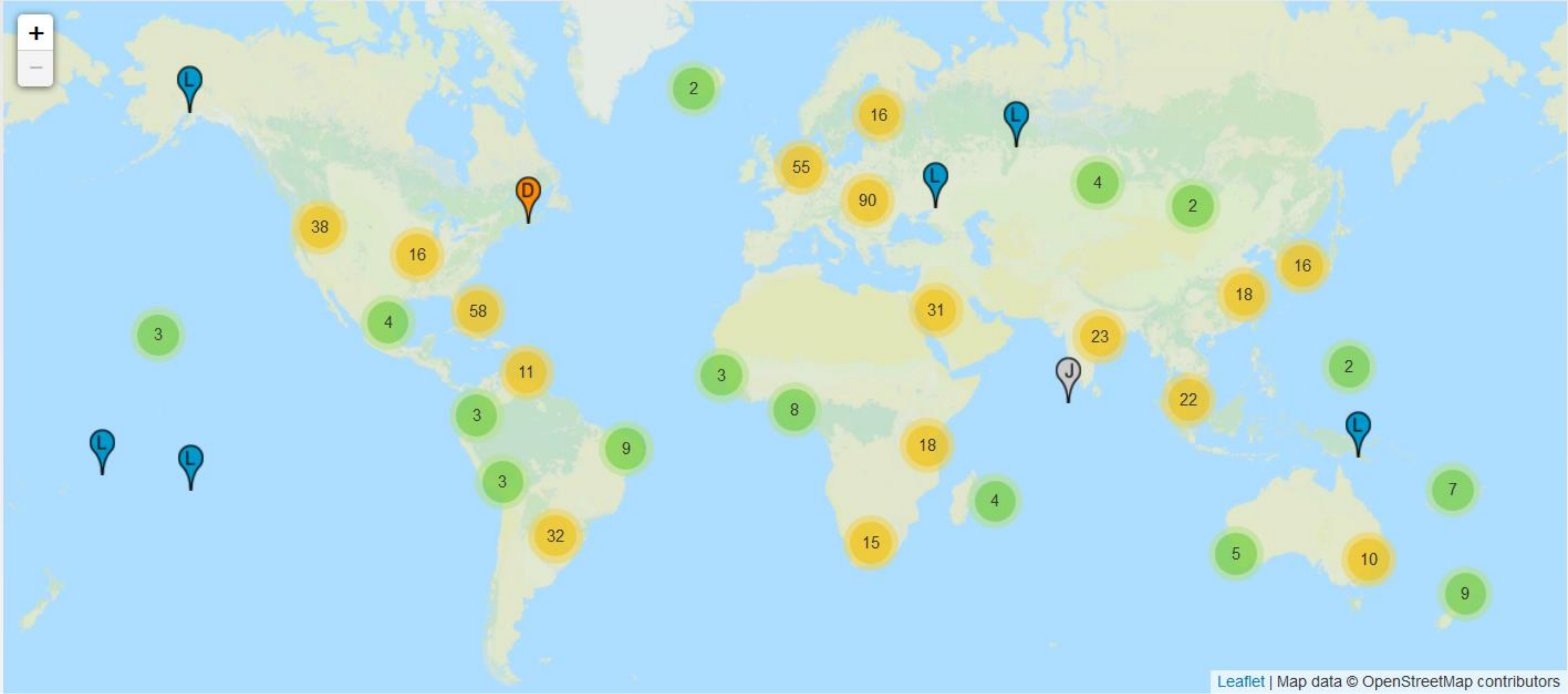


From http://www.postel.org/pr.htm: Photo by Irene Fertik, USC News Service. © 1994, USC. Permission granted for free use and distribution, conditioned upon inclusion of the above attribution and copyright notice.

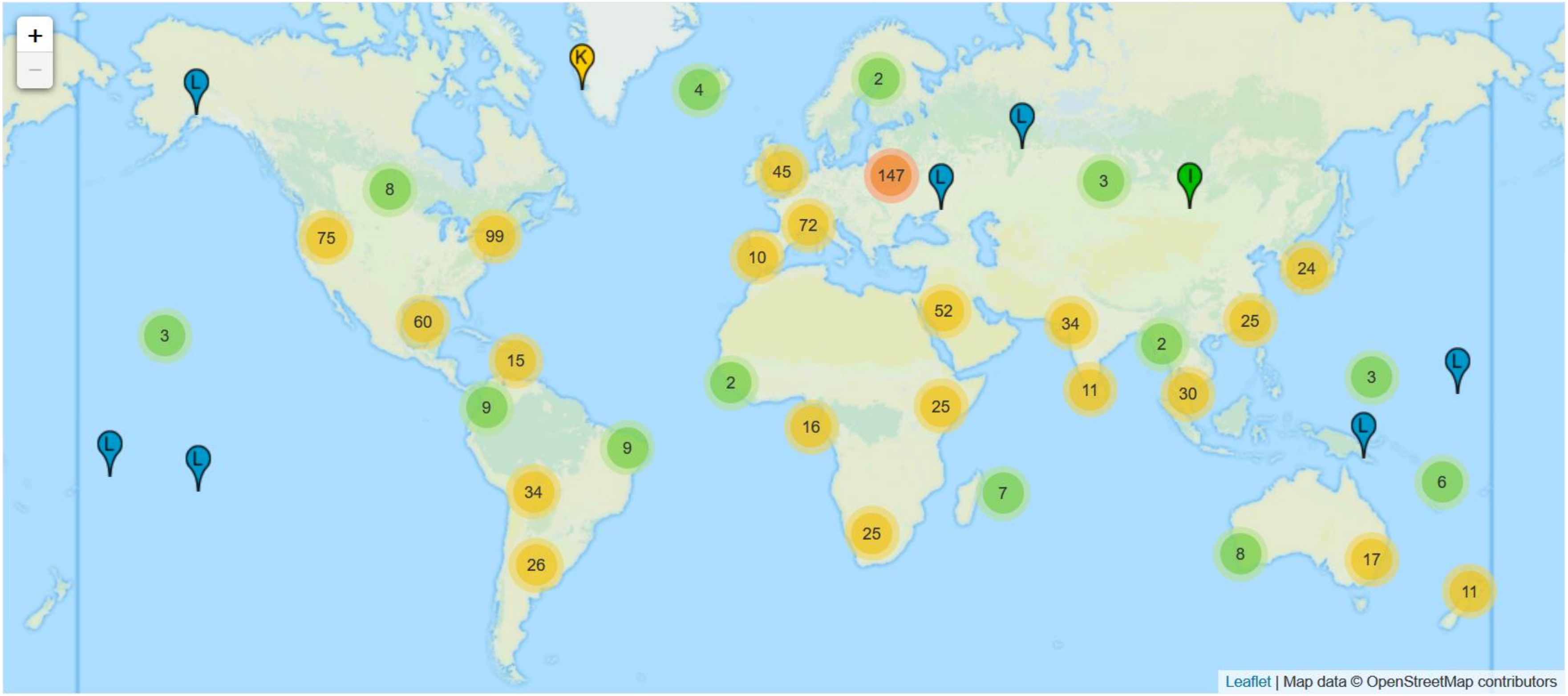
DNS Roots Worldwide (2015)



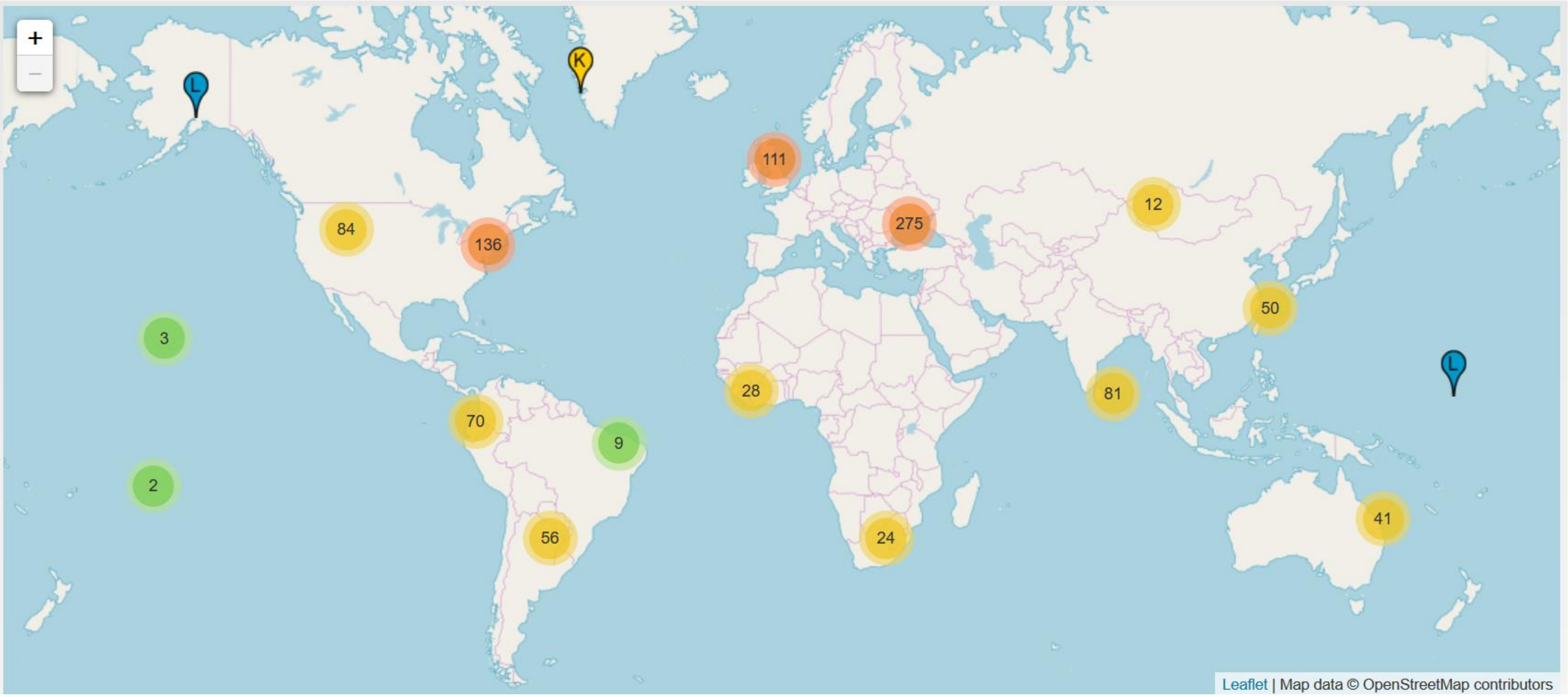
DNS Roots Worldwide (2016)



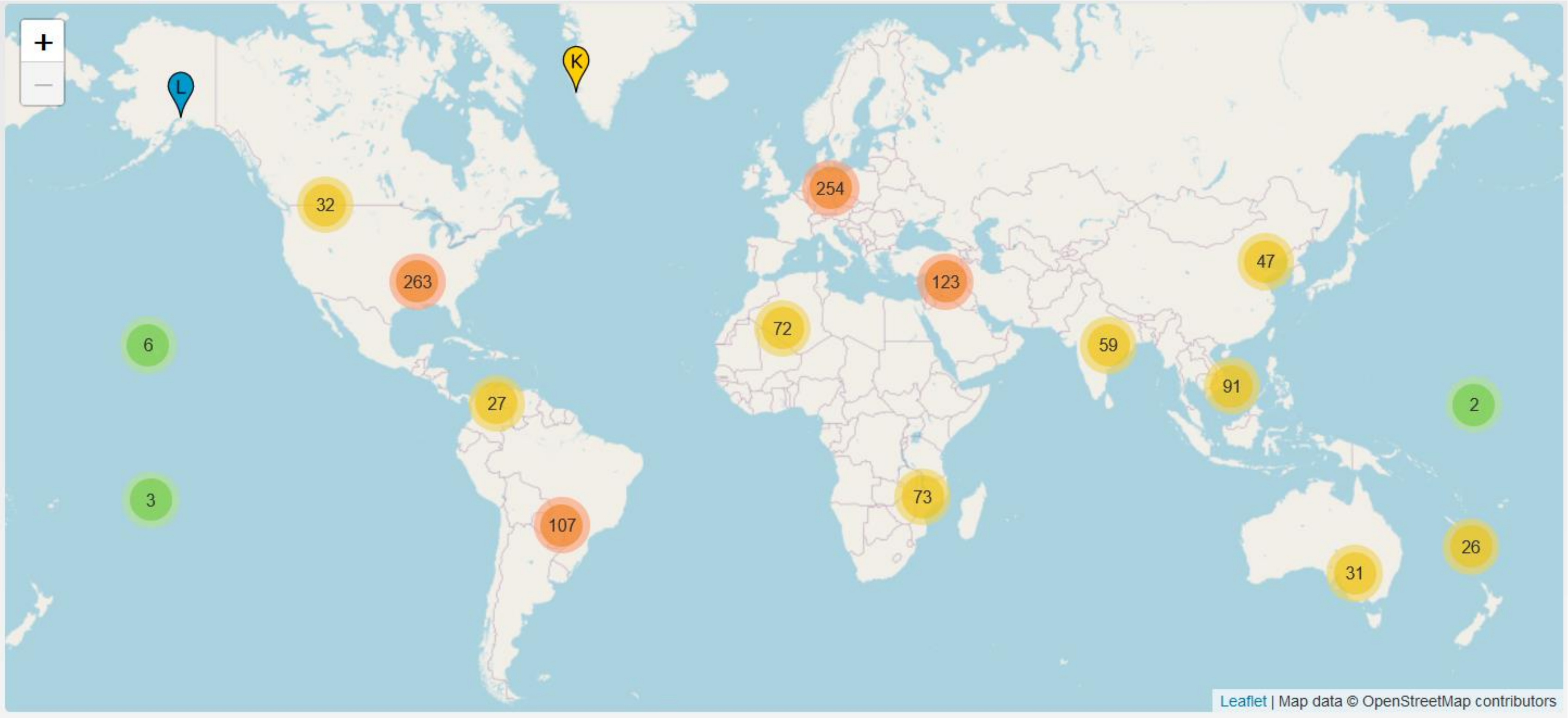
DNS Roots Worldwide (2018)



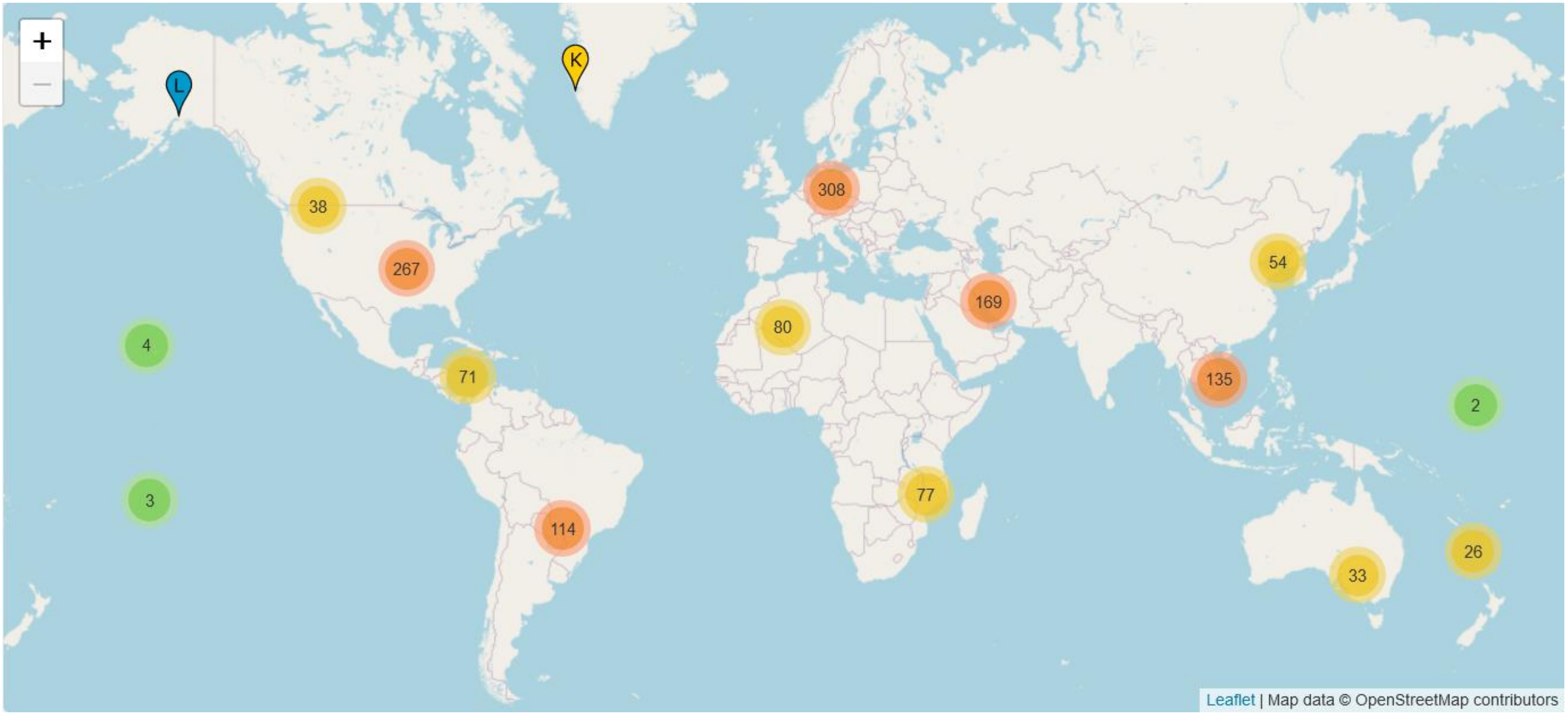
DNS Roots Worldwide (2019)



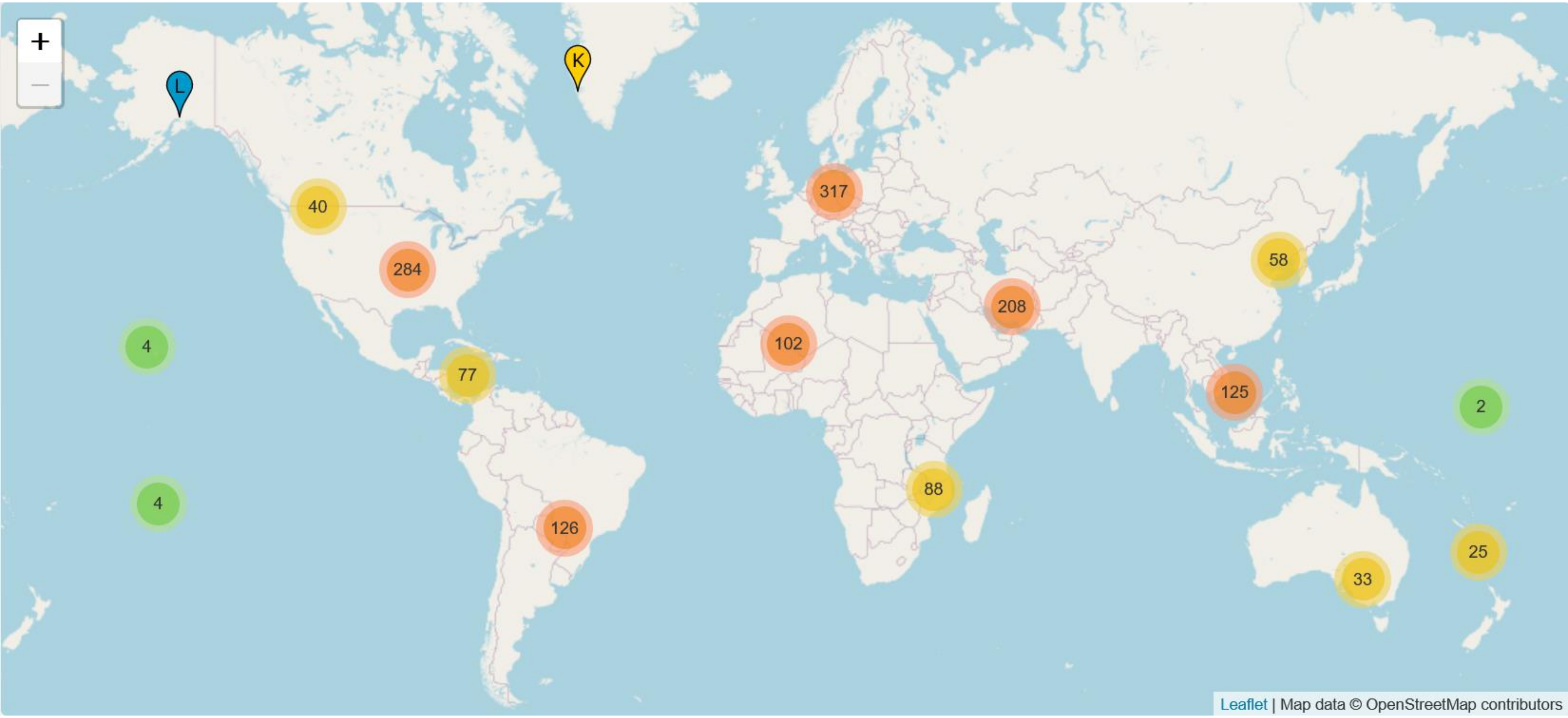
DNS Roots Worldwide (2021)



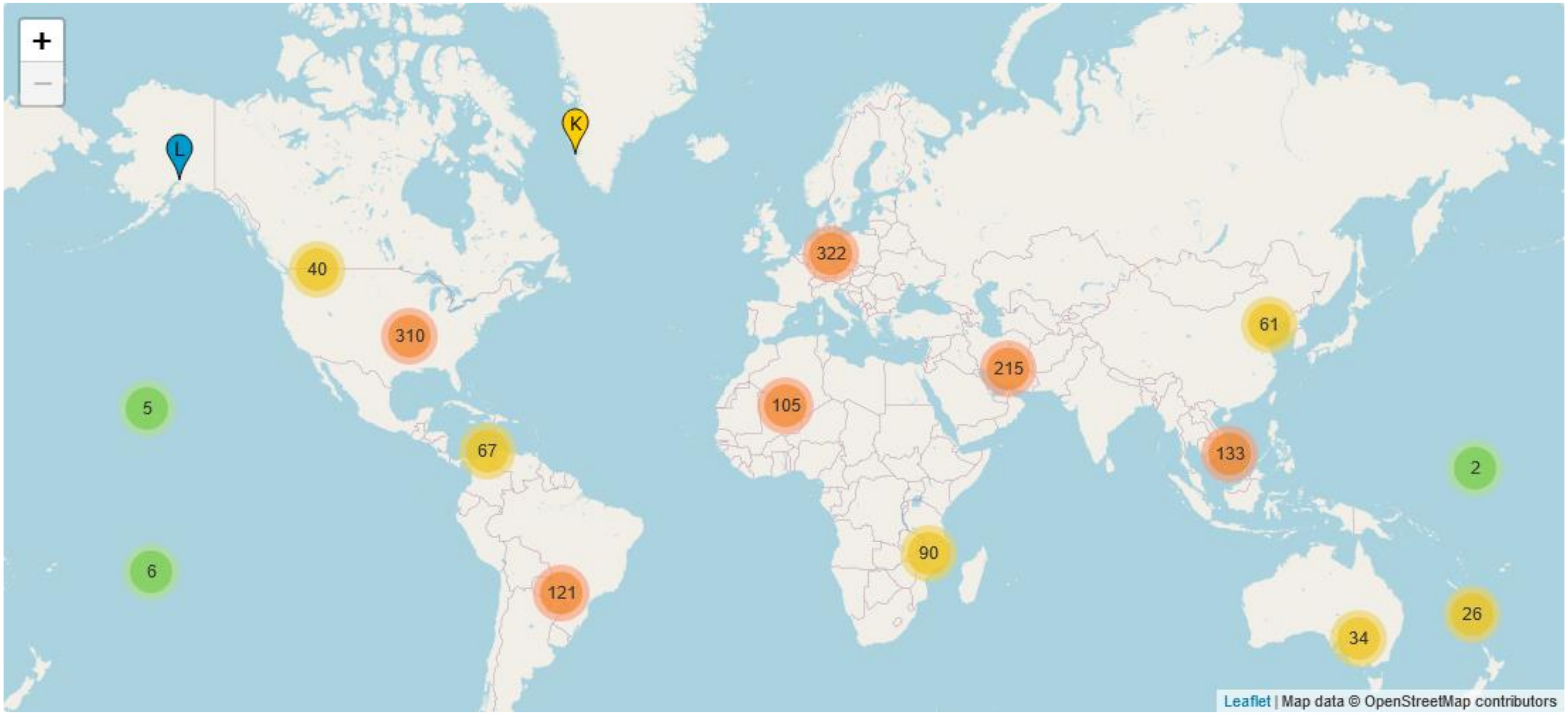
DNS Roots Worldwide (2022)



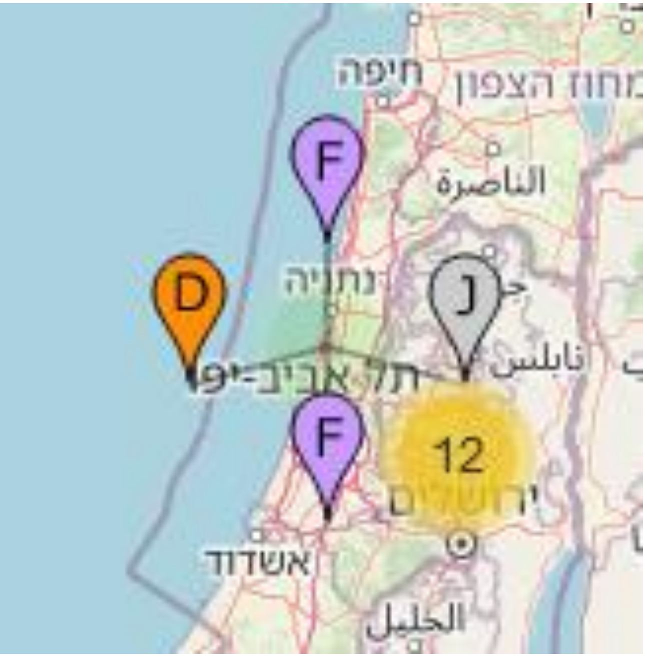
DNS Roots Worldwide (2023)

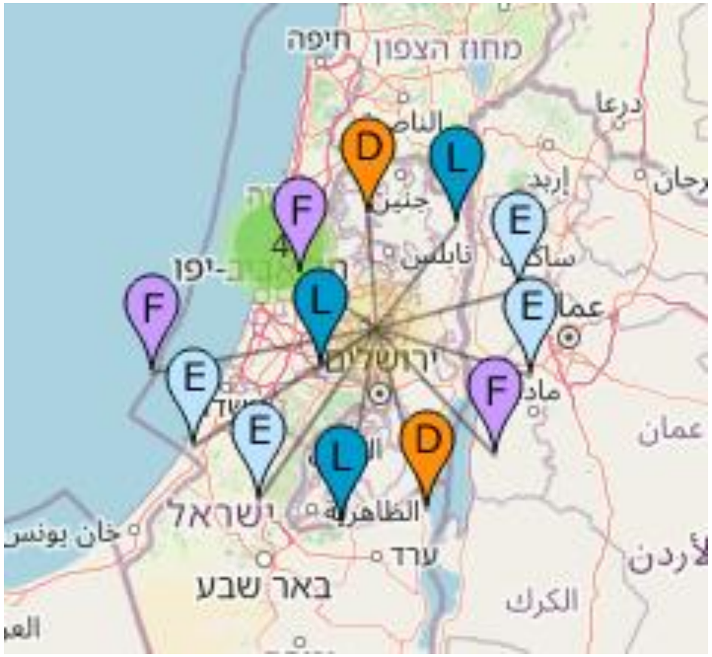


DNS Roots Worldwide (2024)



DNS Roots in Israel





Map includes some in Jordan, Ramallah and Gaza.
Total of 7 in Petah Tikvah and Tel Aviv.

DNS TLDs

1,445 TLDs (Top Level Domains) are maintained by private networking companies and organizations (Jan 2025)

- Private registrars sign up customers

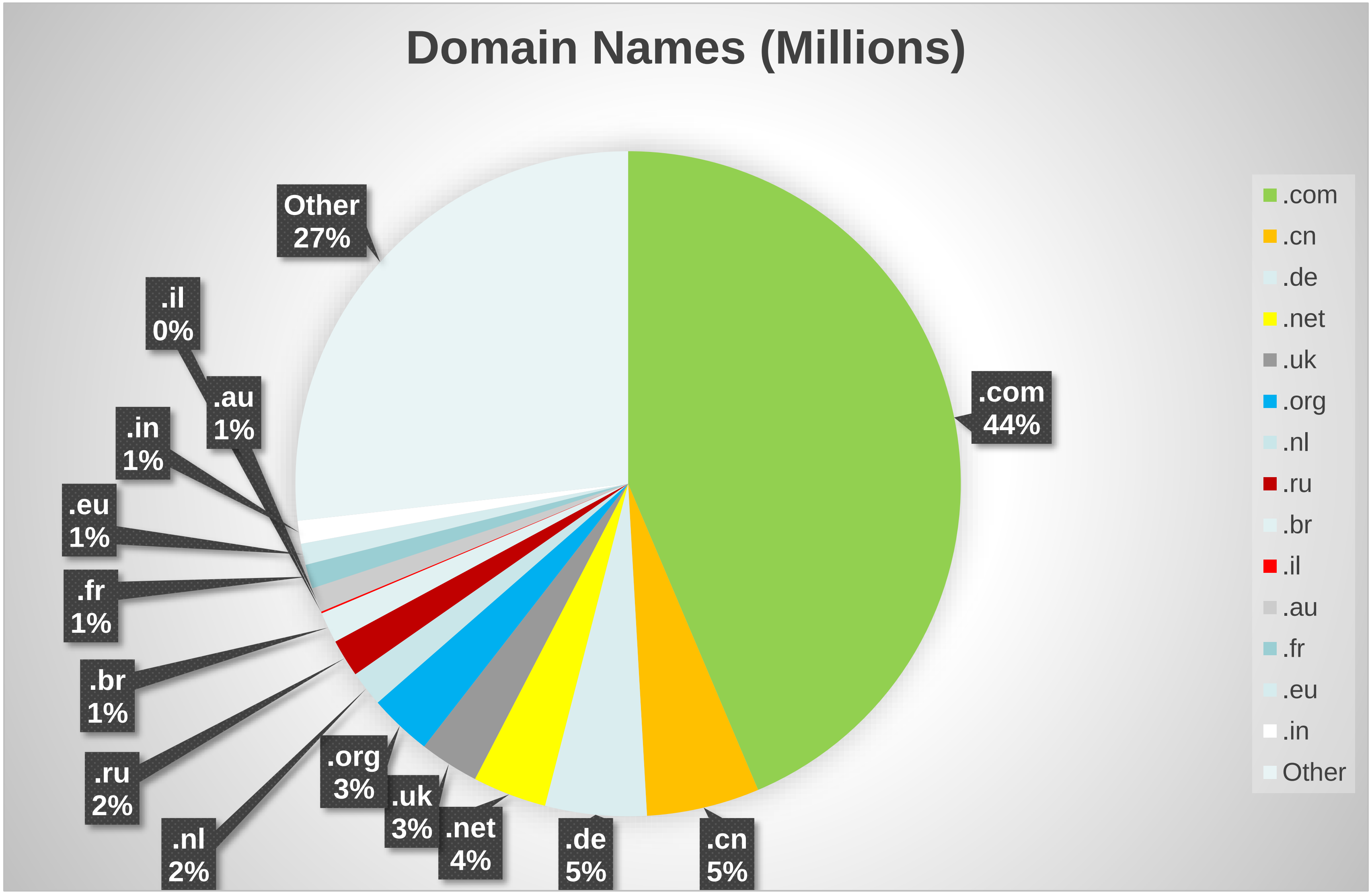
TLDs are

- By business sector (ex. .bike, .clothing, .plumbing)
- By country (ex. .us, .il, .ca, .uk)
- By organization type (ex. .org, .ac.il, .edu, .co.uk)
- By language (ex. XN--1QQW23A (Chinese), XN--3E0B707E (Korean), XN--45BRJ9C (Hindi), XN--4GBRIM (Arabic – Saudi Arabia))
- Generic (ex. .info, .xyz, .center, .cards)

Notable TLDs:

- .com used to be run by US DoD, now by Verisign – 160.9 million domains (Dec 2022)
- .edu run by Educause (contracted to Verisign)
- .il is run by ISOC Israel - 285K domains (2025)
  .ישראל is also run by ISOC 9K domains (2025)

Domain Name Distribution



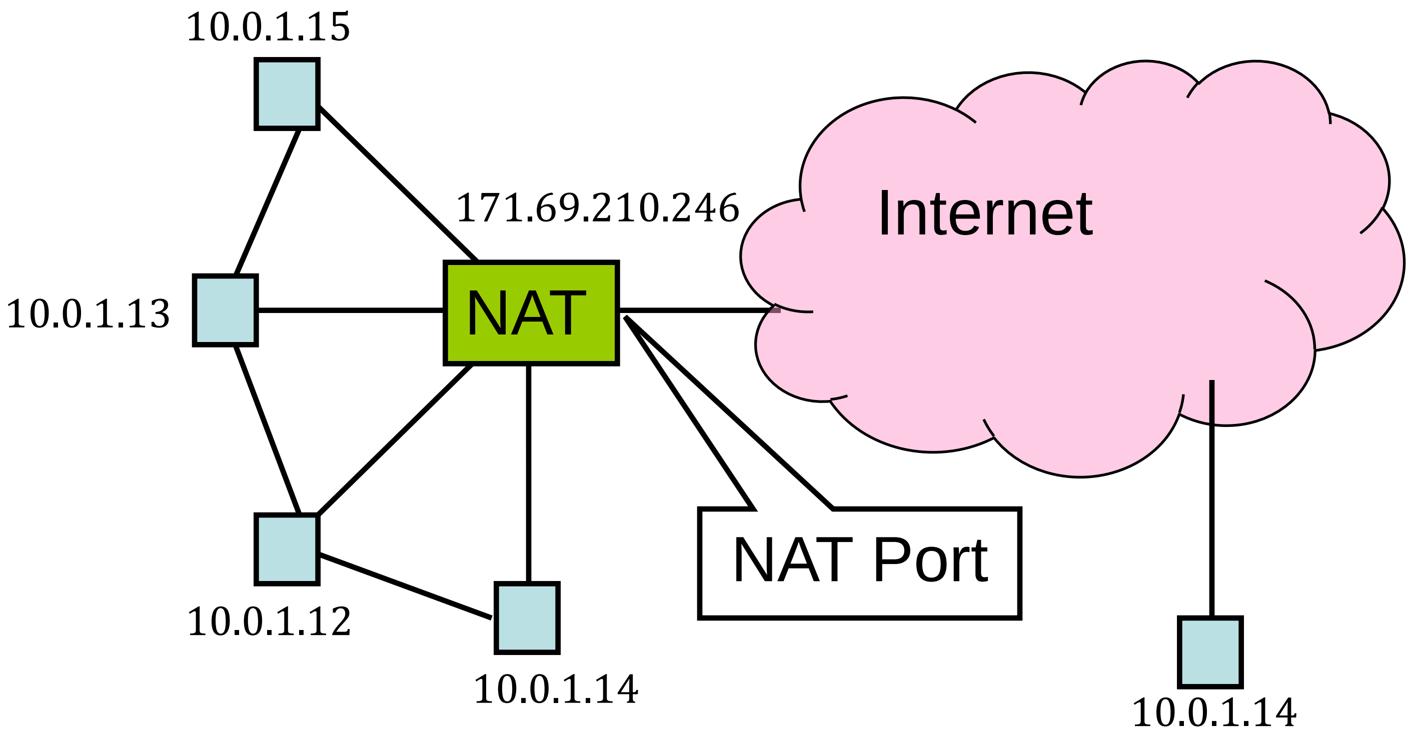
Data source: Domain Name Industry Brief Q3 2024
(https://dnib.com/articles/the-domain-name-industry-brief-q3-2024)

So Far

- DNS
- Network Address Translation (NAT)
- Congestion Control
  - Queuing

Network Address Translation

- Idea: Break the invariant that IP addresses are globally unique



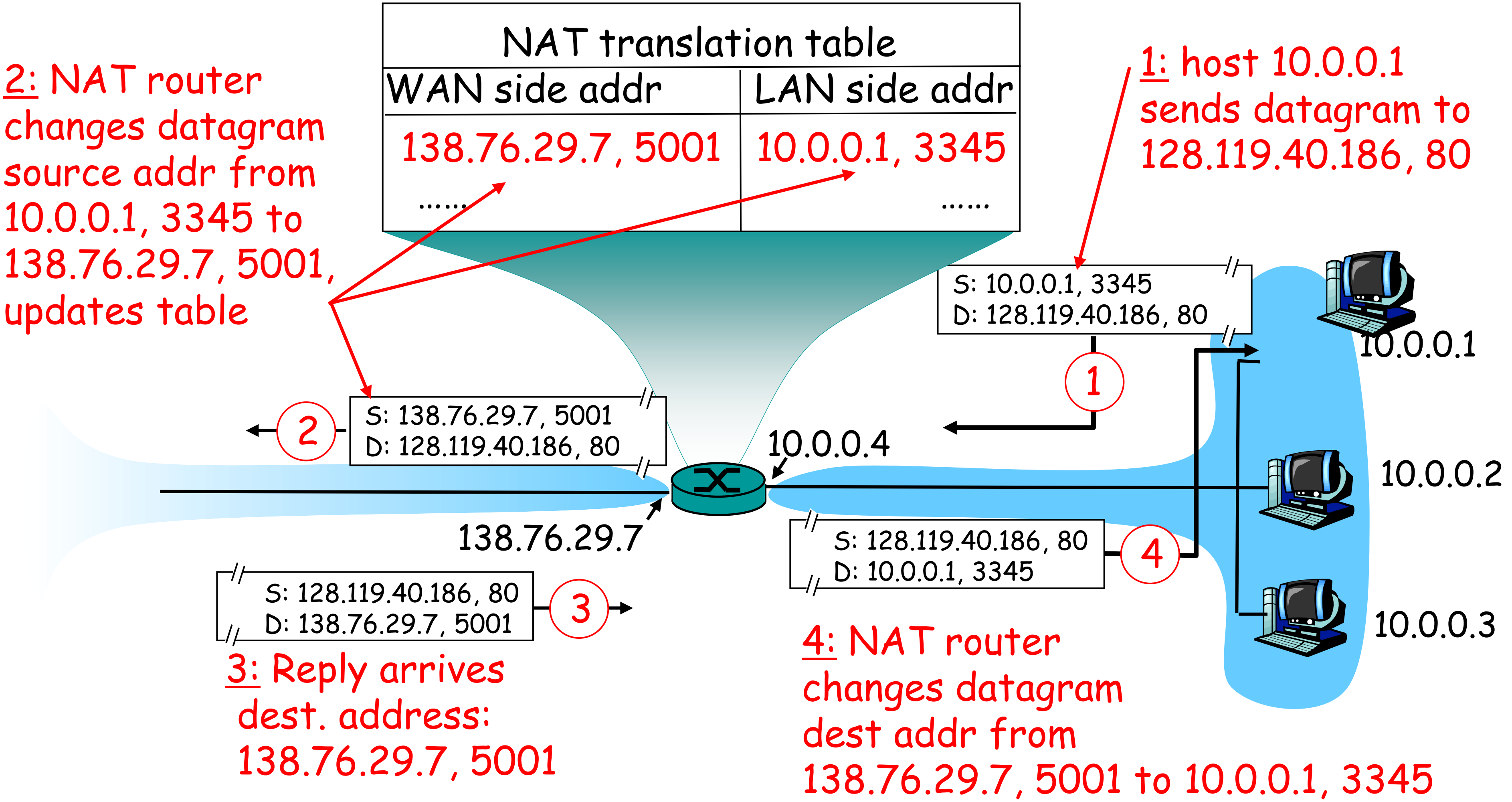
NAT Behavior

- NAT maintains a table of the form:

  $\langle client\ IP\rangle\langle client\ port\rangle\langle NAT\ ID\rangle$

- Outgoing packets (on non-NAT port):
  - Look for client IP address, client port in the mapping table
  - If found, replace client port with previously allocated NAT ID (same size as PORT #)
  - If not found, allocate a new unique NAT ID and replace source port with NAT ID
  - Replace source address with NAT address

NAT: Network Address Translation



NAT translation table

WAN side addr	LAN side addr
138.76.29.7, 5001	10.0.0.1, 3345
……	……

1: host 10.0.0.1 sends datagram to 128.119.40.186, 80

2: NAT router changes datagram source addr from 10.0.0.1, 3345 to 138.76.29.7, 5001, updates table

3: Reply arrives dest. address: 138.76.29.7, 5001

4: NAT router changes datagram dest addr from 138.76.29.7, 5001 to 10.0.0.1, 3345

NAT Behavior

- Incoming Packets (on NAT port)
  - Look up destination port number as NAT ID in port mapping table
  - If found, replace destination address and port with client entries from the mapping table
  - If not found, the packet is not for us and should be rejected

- Table entries expire after 2-3 minutes to allow them to be garbage collected

- "Private" IP addresses:
  - $192.168.x.x$
  - $172.16.x.x$-$172.31.x.x$
  - $10.x.x.x$

Benefits of NAT

- Only allows connections to the outside that are established from *inside.*
  - Hosts from outside can only contact internal hosts that appear in the mapping table, and they’re only added when they establish the connection
  - Some NATs support firewall-like configurability

- Can simplify network administration
  - Divide network into smaller chunks
  - Consolidate configuration data

- Traffic logging
- Load balancing
- Robust failover

Drawbacks of NAT

Rewriting IP addresses isn’t so easy:

- Must also look for IP addresses in other locations and rewrite them (may have to be protocol-aware)
- Potentially changes sequence number information
- Must validate/recalculate checksums

Limited filtering of packets / change packet semantics

- For example, NATs may not work well with encryption schemes that include IP address information

May not work with all protocols

- Clients may have to be aware that NAT translation is going on

Hinders throughput

Slow the adoption of IPv6?

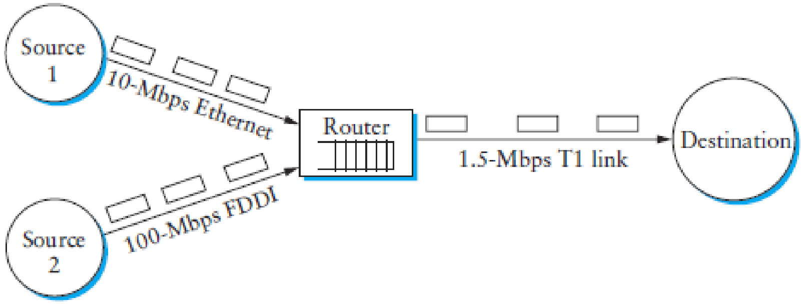
So Far

- DNS
- Network Address Translation (NAT)
- Congestion Control
  - Queuing

Resource Allocation

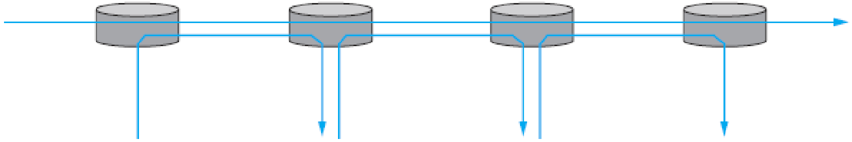
- When we have a real network we must deal with contention and congestion
  - Too many users, not enough resources
- We'll talk about packet switched networks for now

- Congestion can come from:
  - Too many users trying to make small connections
  - A few users making huge connections
  - Fast links that must pass over a slower link

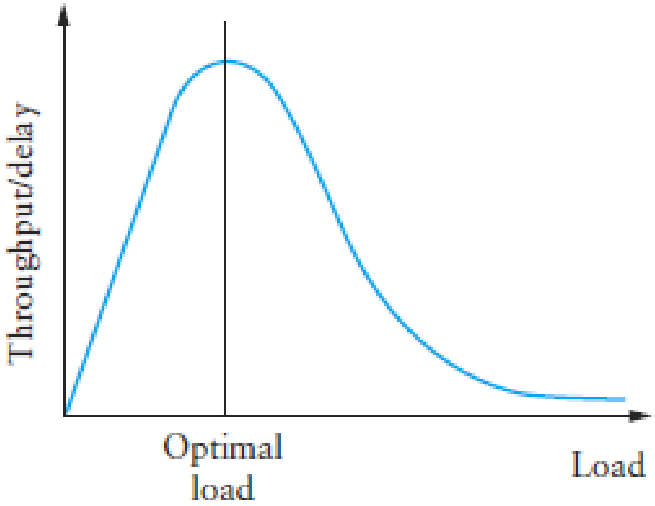


What is the Goal?

- Fairness



- Utilization



What are we Managing?

- Connectionless Flows
  - Data sent between sender and receiver
  - The routers sees them as moving between addresses (ignore ports)
- Routers maintain *soft state* about connections
  - Detected automatically
  - Lives and dies as the connection does
  - Helps the router make better routing decisions
- Flows can be *explicit* or *implicit*
  - Difference is whether the end points tell the routers before they start
  - Datagram versus Virtual Circuits

What is the Network Offering?

- The basic model: Best Effort
  - Try, but no guarantee
  - All packets are created (more or less) equal

- More advanced: Quality of Service (QoS)
  - Senders and receivers *request* the routers to guarantee a minimum amount of resources
  - Some protocols: RSVP, ATM

How are we Managing?

- Router Centric vs. Host Centric
  - Who is doing most of the decision making?
  - Router Centric – the router tells the hosts how fast they can send
  - Host Centric – the hosts decide how fast to send based on their experiences

- Reservation Based vs. Feedback Based
  - Reservation: send request before
    - Requires Router Centric
  - Feedback: change based on what happens
    - Explicit – Router more involved
    - Implicit – Host more involved

- Window Based vs. Rate Based

What is Common?

- With Best Effort:
  - Feedback - since we can't reserve, and therefore…
  - Host centric, and often…
  - Window based
- With QoS:
  - Reservation – normally, and therefore…
  - Router centric, and therefore often…
  - Rate based

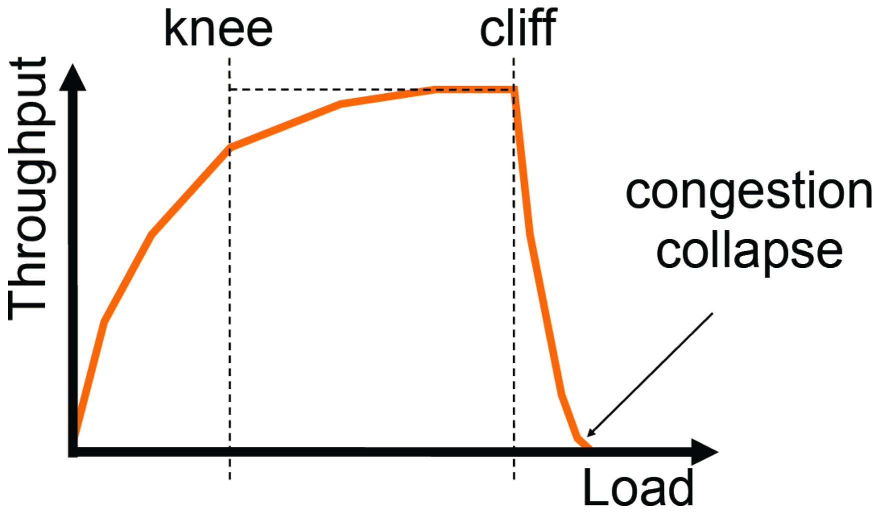
So Far

- DNS
- Network Address Translation (NAT)
- Congestion Control
  - Queuing

Congestion Control vs. Avoidance

Congestion **control**

goal: Stay left of cliff



Congestion **avoidance**

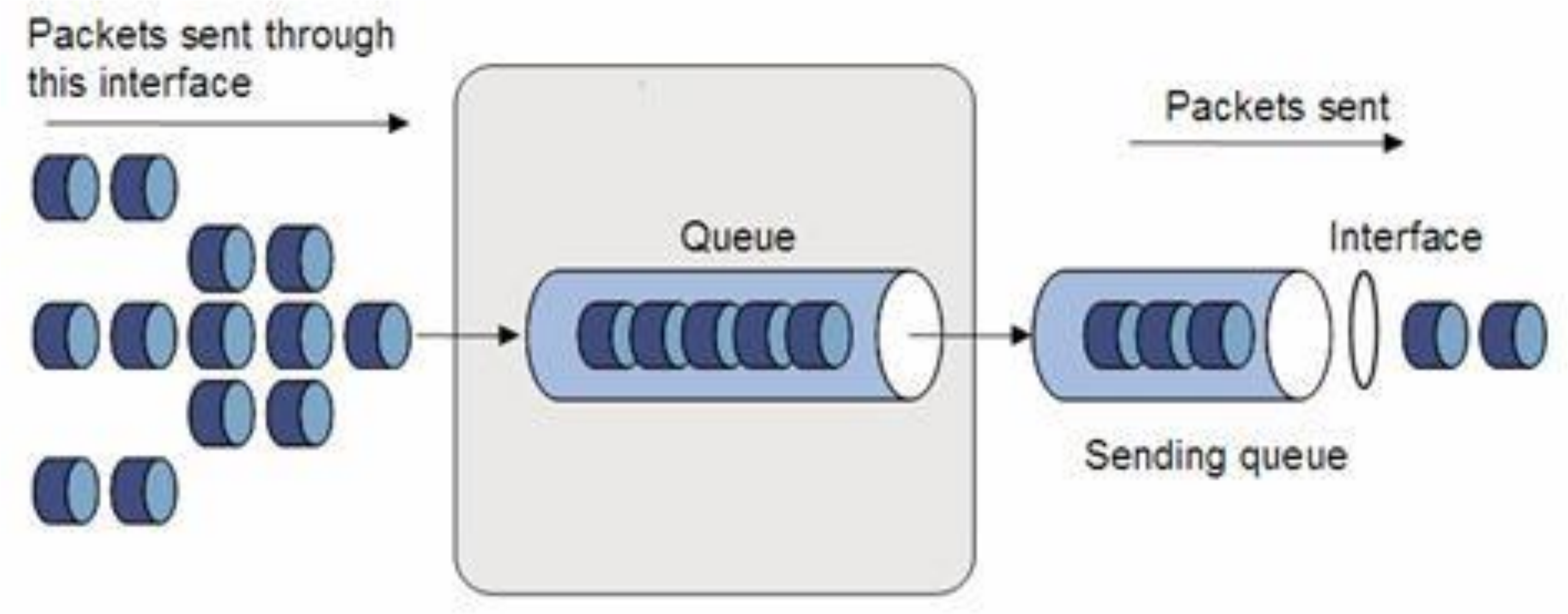
goal: Stay left of knee

Queuing Techniques

- First In First Out (FIFO)
- Priority Queuing (PQ)
- Fair Queuing (FQ)
- Weighted Fair Queuing (WFQ)

First In First Out

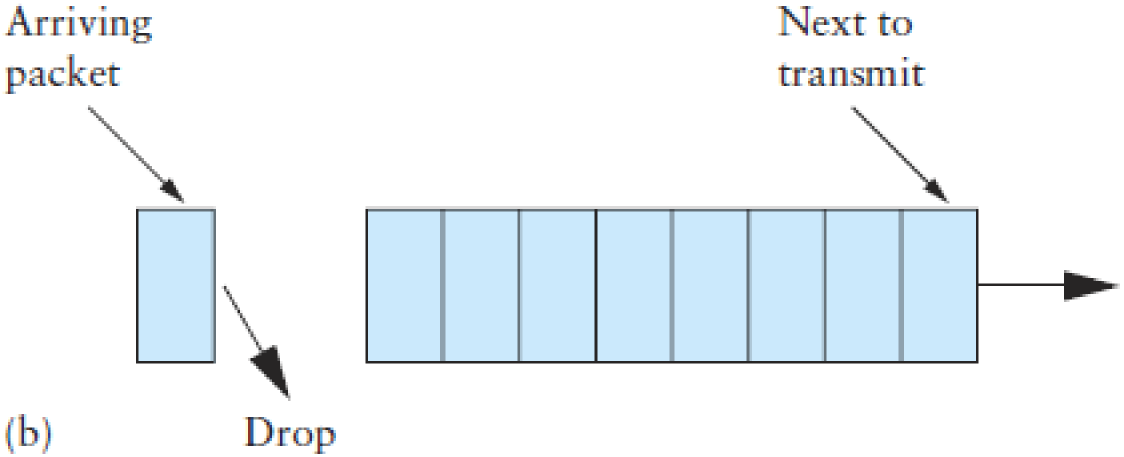
- Rule: Packets are sent out of the router as they arrive



FIFO and Dropping

- What if the queue is full?

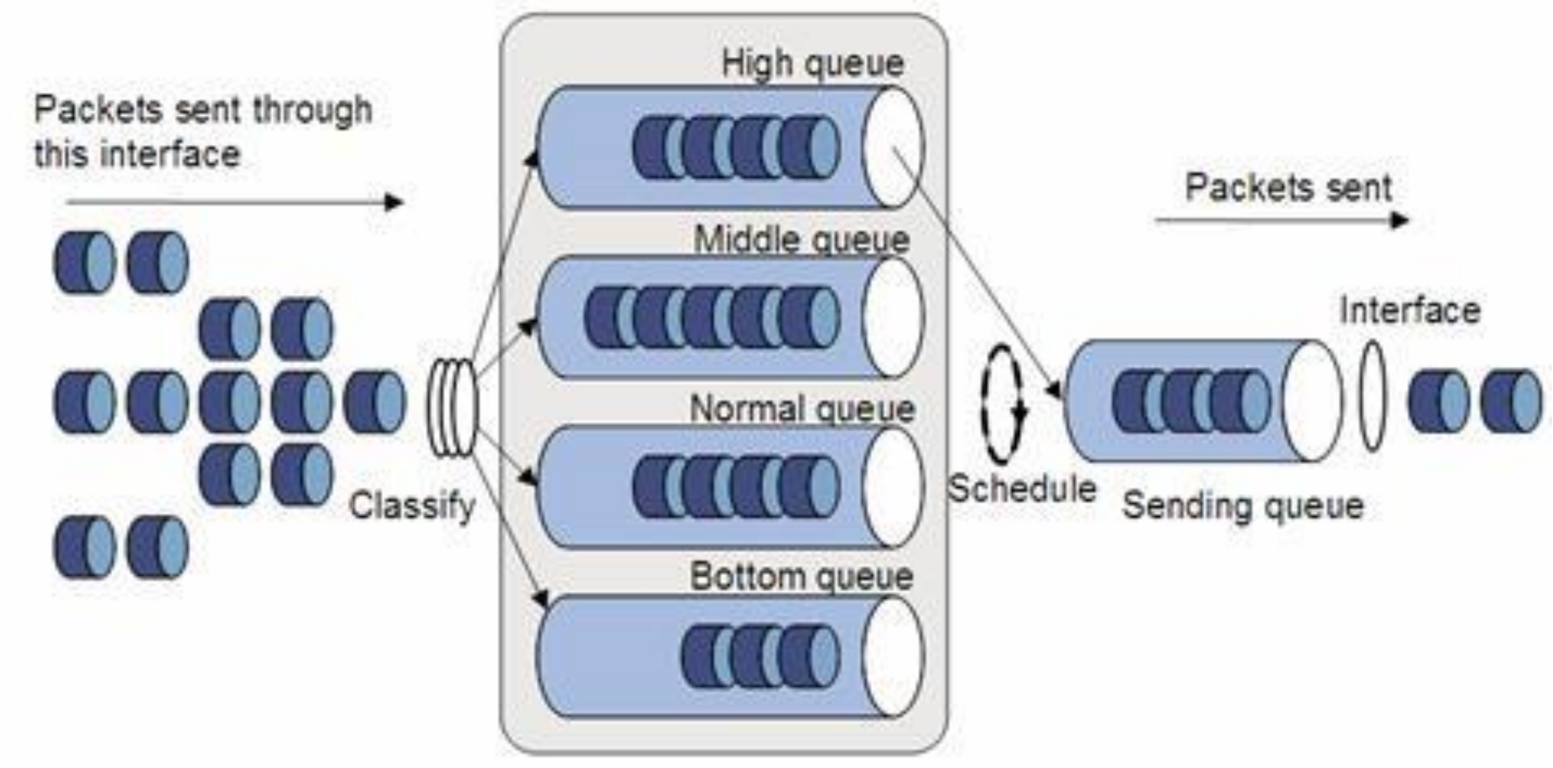
- Drop somebody:
  - Tail Drop



  - Random Drop
    - Why?

Priority Queuing

- Put a strict order on the queues
  - Highest priority first, then secondary ones
  - Advantages? Disadvantages?



Conclusion

- DNS
- Network Address Translation (NAT)
- Congestion Control
  - Queuing