Communication Networks

Prof. Laurent Vanbever

Communication Networks

Spring 2021

Laurent Vanbever

nsg.ee.ethz.ch

ETH Zürich (D-ITET)

May 17 2021

Materials inspired from Scott Shenker & Jennifer Rexford

Last week on

Communication Networks

TCP: Reliable, in-order delivery

TCP provides a connection-oriented, reliable, bytestream transport service

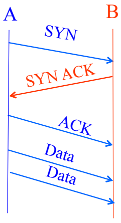
What UDP provides, plus:

- Retransmission of lost and corrupted packets
- Flow control (to not overflow receiver)
- Congestion control (to not overload network)
- "Connection" set-up & tear-down

TCP Header

Source port			Destination port
Sequence number			
Acknowledgment			
HdrLen	0	Flags	Advertised window
Checksum			Urgent pointer
Options (variable)			
Data			

Establishing a TCP Connection



Each host tells its ISN to the other host.

Three-way handshake to establish connection

- Host A sends a **SYN** (open; "synchronize sequence numbers")
- Host B returns a SYN acknowledgment (**SYN ACK**)
- Host A sends an **ACK** to acknowledge the SYN ACK

This week on

Communication Networks

Congestion Control	DNS	Introduction to 2nd project
	ethz.ch ⇨ 129.132.19.216	reliable transport starts *today*!

Congestion Control

DNS

Introduction to 2nd project

Congestion Control

DNS

Introduction to 2nd project

ethz.ch ⇨
129.132.19.216

Internet has one global system for

- addressing hosts — IP
  by design
- naming hosts — DNS
  by "accident", an afterthought

Internet has one global system for

- naming hosts — DNS
  by "accident", an afterthought

Using Internet services can be divided into four logical steps

step 1	A person has name of entity she wants to access	www.ethz.ch
step 2	She invokes an application to perform the task	Chrome
step 3	The application invokes DNS to resolve the name into an IP address	129.132.19.216
step 4	The application invokes transport protocol to establish an app-to-app connection	

The DNS system is a distributed database which enables to resolve a name into an IP address

name	DNS ↔	IP address
www.ethz.ch		129.132.19.216

In practice, names can be mapped to more than one IP

name	DNS ↔	IP address	
www.ethz.ch		129.132.19.216	
www.netflix.com		52.18.41.127	(load-balancing)
		34.248.59.175	
		52.209.236.203	+5 more!

In practice, IPs can be mapped by more than one name

name	DNS ↔	IP address
www.ethz.ch		129.132.19.216
www.vanbever.eu		82.130.102.71
www.route-aggregation.net		82.130.102.71
comm-net.ethz.ch		82.130.102.71

How does one resolve a name into an IP?

initially — *all* host to address mappings
were in a file called hosts.txt

in /etc/hosts

problem — scalability in terms of query load & speed
management

consistency

availability

When you need… more flexibility,

you add… a layer of indirection

When you need… more scalability,

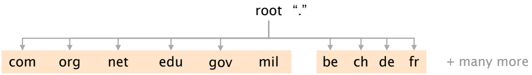
you add… a hierarchical structure

To scale,
DNS adopt three intertwined hierarchies

naming structure	hierarchy of addresses https://www.ee.ethz.ch/de/departement/
management	hierarchy of authority over names
infrastructure	hierarchy of DNS servers

naming structure — hierarchy of addresses
https://www.ee.ethz.ch/de/departement/

Top Level Domain (TLDs) sit at the top

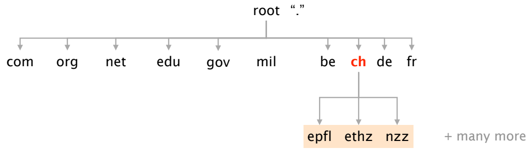


root "."

com org net edu gov mil be ch de fr

+ many more

Domains are subtrees



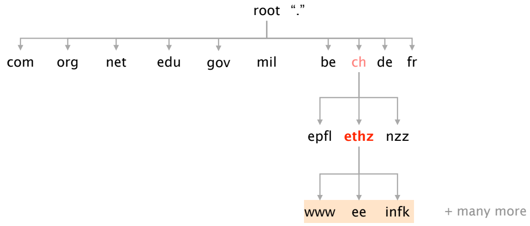
root "."

com org net edu gov mil be ch de fr

epfl ethz nzz

+ many more

A name, *e.g.* ee.ethz.ch, represents
a leaf-to-root path in the hierarchy



root "."

com org net edu gov mil be ch de fr

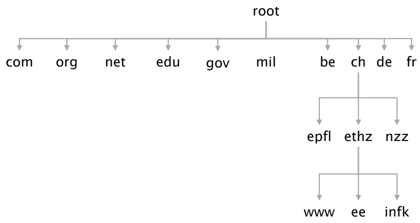
epfl ethz nzz

www ee infk

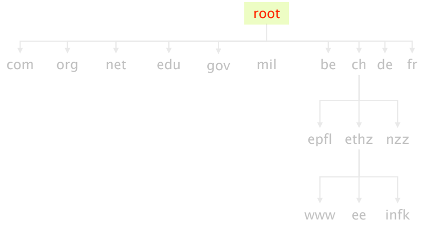
+ many more

management — hierarchy of authority
over names

The DNS system is hierarchically administered

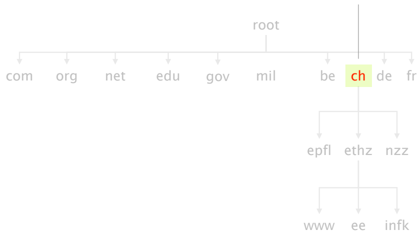


managed by IANA (*)

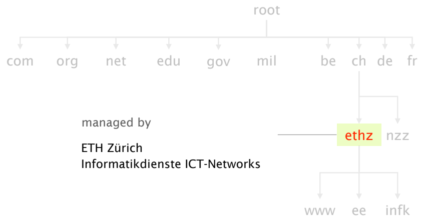


(*) see http://www.iana.org/domains/root/db

managed by The Swiss Education & Research Network (*)



(*) see https://www.switch.ch/about/id/



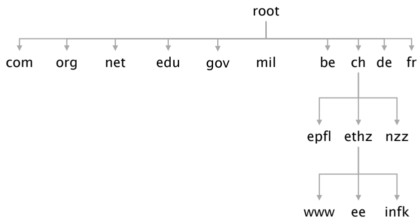
managed by

ETH Zürich
Informatikdienste ICT-Networks

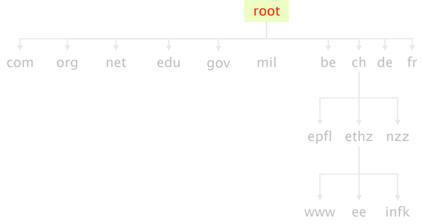
Hierarchical administration means that name collision is trivially avoided

infrastructure hierarchy of DNS servers

The DNS infrastructure is hierarchically organized

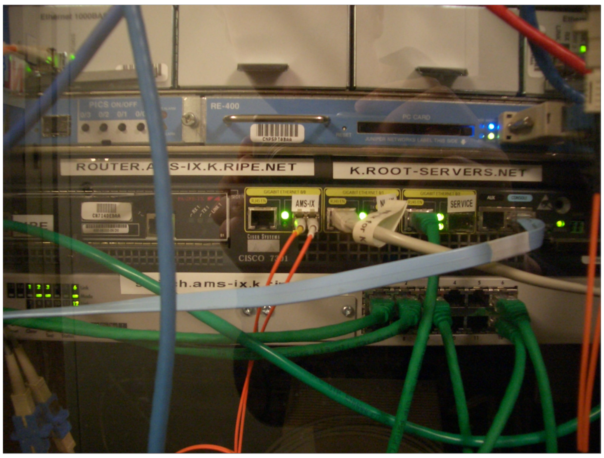


13 root servers (managed professionally) serve as root (*)



(*) see http://www.root-servers.org/

a. root-servers.net	VeriSign, Inc.
b. root-servers.net	University of Southern California
c. root-servers.net	Cogent Communications
d. root-servers.net	University of Maryland
e. root-servers.net	NASA
f. root-servers.net	Internet Systems Consortium
g. root-servers.net	US Department of Defense
h. root-servers.net	US Army
i. root-servers.net	Netnod
j. root-servers.net	VeriSign, Inc.
k. root-servers.net	RIPE NCC
l. root-servers.net	ICANN
m. root-servers.net	WIDE Project



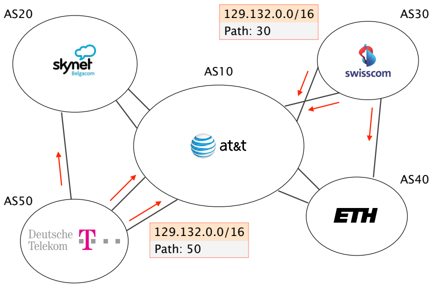
To scale root servers,
operators rely on BGP anycast

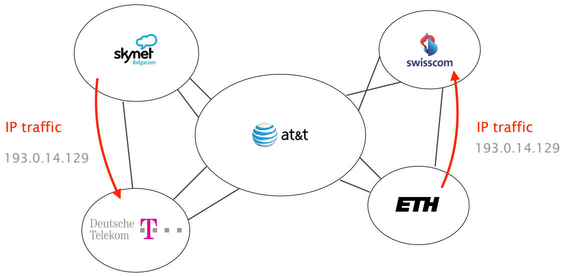
Intuition

Routing finds shortest-paths

If several locations announce the same prefix, then routing will deliver the packets to the “closest” location

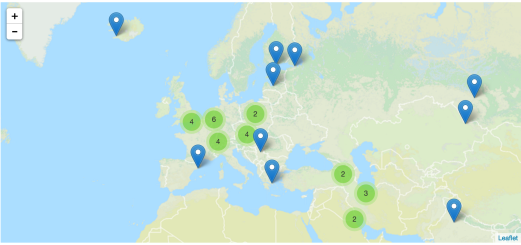
This enables seamless replications of resources





Do you see any problems in
performing load-balancing this way?

Instances of the k-root server (*) are hosted
in more than 40 locations worldwide

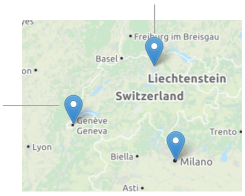


(*) see k.root-servers.org

Two of these locations are in Switzerland:
in Zürich and in Geneva

Swiss Internet Exchange
ns1.ch-zrh.k.ripe.net

CERN
ns1.ch-gva.k.ripe.net

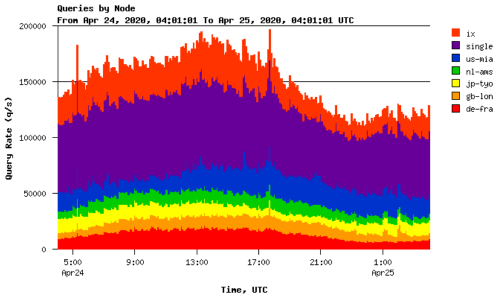


All locations announce 193.0.14.0/23 in BGP,
with 193.0.14.129 being the IP of the server

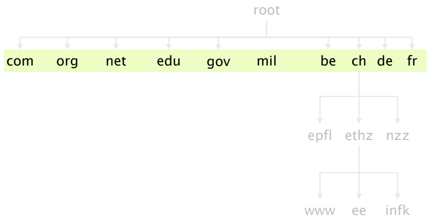
Two of these locations are in Switzerland:
in Zürich and in Geneva

Do you mind guessing which one we use, here… in Zürich?

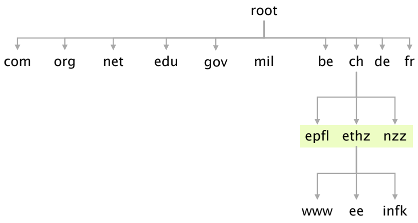
Each instance receives up to 70k queries per second
summing up to more than 4 billions queries per day



TLDs server are also managed professionally
by private or non-profit organization



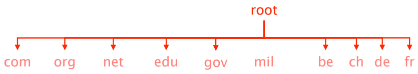
The bottom (and bulk) of the hierarchy is
managed by Internet Service Provider or locally



Every server knows the address of the root servers (*)
required for bootstrapping the systems

(*) see https://www.internic.net/domain/named.root

Each root server knows
the address of all TLD servers



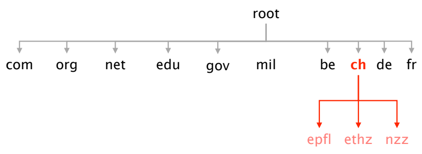
```
lvanbever:~$ dig @a.root-servers.net ch.

ch.                     172800  IN      NS      a.nic.ch.
ch.                     172800  IN      NS      b.nic.ch.
ch.                     172800  IN      NS      c.nic.ch.
ch.                     172800  IN      NS      d.nic.ch.
ch.                     172800  IN      NS      e.nic.ch.
ch.                     172800  IN      NS      f.nic.ch.
ch.                     172800  IN      NS      h.nic.ch.
```

also see https://www.iana.org/domains/root/db/ch.html

If you want to learn more on ".ch"
take a look at SWITCH's annual report

https://www.nic.ch/export/shared/.content/files/SWITCH_Report_Registry_2020.pdf

Any .ch DNS server knowns the addresses of the DNS servers of all sub-domains



```
lvanbever:~$ dig @a.nic.ch ethz.ch

ethz.ch.    3600  IN NS ns2.switch.ch.
ethz.ch.    3600  IN NS ns2.ethz.ch.
ethz.ch.    3600  IN NS ns1.ethz.ch.
```

Once arrived at the leaf of the hierarchy (*.ethz.ch), each DNS server knows the IP address of all children

```
lvanbever:~$ dig @ns1.ethz.ch comm-net.ethz.ch

comm-net.ethz.ch. 3600  IN CNAME virt07.ethz.ch.
virt07.ethz.ch.     3600  IN A  82.130.102.71
```

To scale, DNS adopt three intertwined hierarchies

naming structure — addresses are hierarchical
https://www.ee.ethz.ch/de/departement/

management — hierarchy of authority over names

infrastructure — hierarchy of DNS servers

To ensure availability, each domain must have at least a primary and secondary DNS server

Ensure name service availability
as long as one of the servers is up

DNS queries can be load-balanced
across the replicas

On timeout, client use alternate servers
exponential backoff when trying the same server

Overall, the DNS system is highly scalable, available, and extensible

scalable — #names, #updates, #lookups, #users,
but also in terms of administration

available — domains replicate independently of each other

extensible — any level (including the TLDs) can be modified independently

You've founded next-startup.ch and want to host it yourself, how do you insert it into the DNS?

You register next-startup.ch at a registrar *X*
e.g. Swisscom or GoDaddy

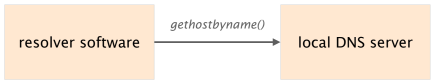
Provide *X* with the name and IP of your DNS servers
e.g., [ns1.next-startup.ch,129.132.19.253]

You set-up a DNS server @129.132.19.253
define A records for www, MX records for next-startup.ch...

A DNS server stores Resource Records composed of a (name, value, type, TTL)

Records	Name	Value
A	hostname	IP address
NS	domain	DNS server name
MX	domain	Mail server name
CNAME	alias	canonical name
PTR	IP address	corresponding hostname

Using DNS relies on two components



trigger resolution process
send request to local DNS server

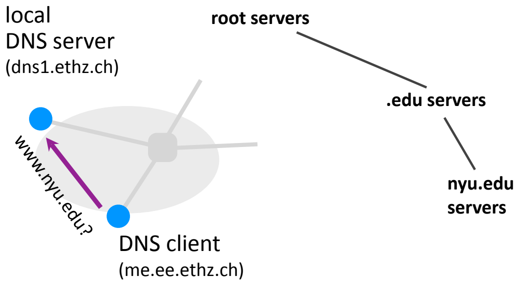
usually, near the endhosts
configured statically (resolv.conf)
or dynamically (DHCP)

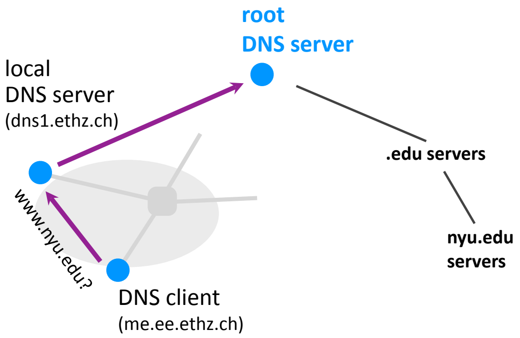
DNS query and reply uses UDP (port 53),
reliability is implemented by repeating requests (*)

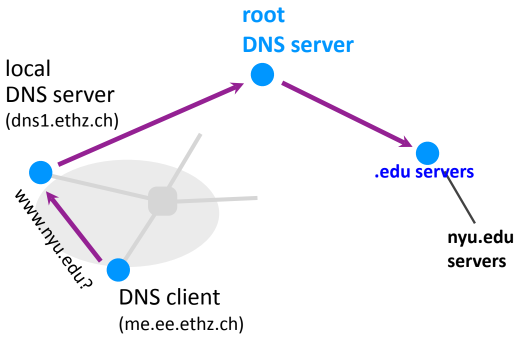
(*) see Book (Section 5)

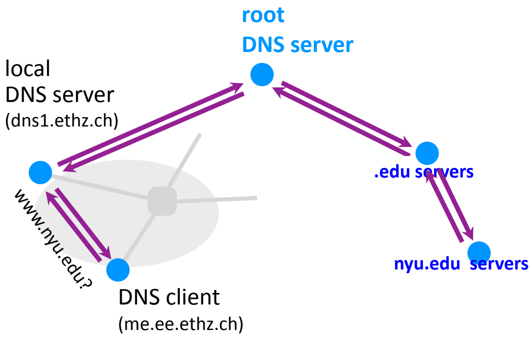
DNS resolution can either be
recursive or iterative

When performing a recursive query,
the client offload the task of resolving to the server

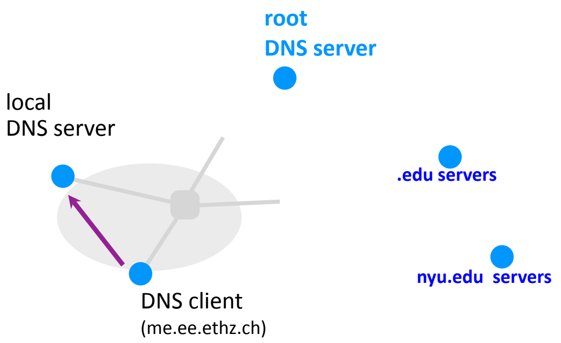


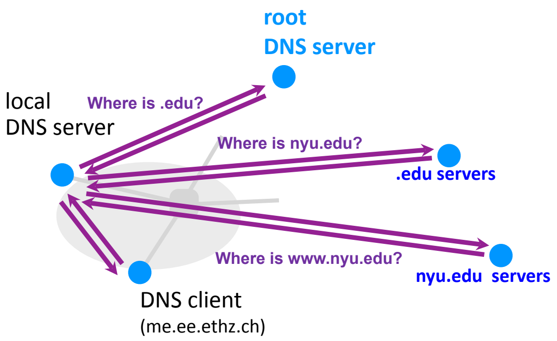






When performing a iterative query,
the server only returns the address of the "next server"





What about resolving speeds?
Waiting for servers all over the globe is not fast...

To reduce resolution times,
DNS relies on caching

DNS servers cache responses to former queries
and your client *and* the applications (!)

Authoritative servers associate a lifetime to each record
Time-To-Live (TTL)

DNS records can only be cached for TTL seconds
after which they must be cleared

As top-level servers rarely change & popular website visited often, caching is very effective (*)

Top 10% of names account for 70% of lookups

9% of lookups are unique
Limit cache hit rate to 91%

Practical cache hit rates ~75%

(*) see https://pdos.csail.mit.edu/papers/dns:ton.pdf

Congestion Control

DNS

Introduction to 2nd project

reliable transport
starts *today*!

Check Tobias' slides on
https://comm-net.ethz.ch