Exploring (a bit of the) Internet Infrastructure

Lars Eggert

lars@eggert.org

2023-6-18

About

- Distinguished Engineer for Internet Standards at NetApp
- Current chair of the Internet Engineering Task Force (IETF)
  - Many other roles since starting at IETF in 2000
- Ph.D. in Computer Science from the University of Southern California (USC) in 2003
- Principal Scientist at Nokia and served on the corporation’s CTO and CEO Technology Councils
- 2009-2014, Adjunct Professor at Aalto University
- 2003-2006, senior researcher at NEC Labs

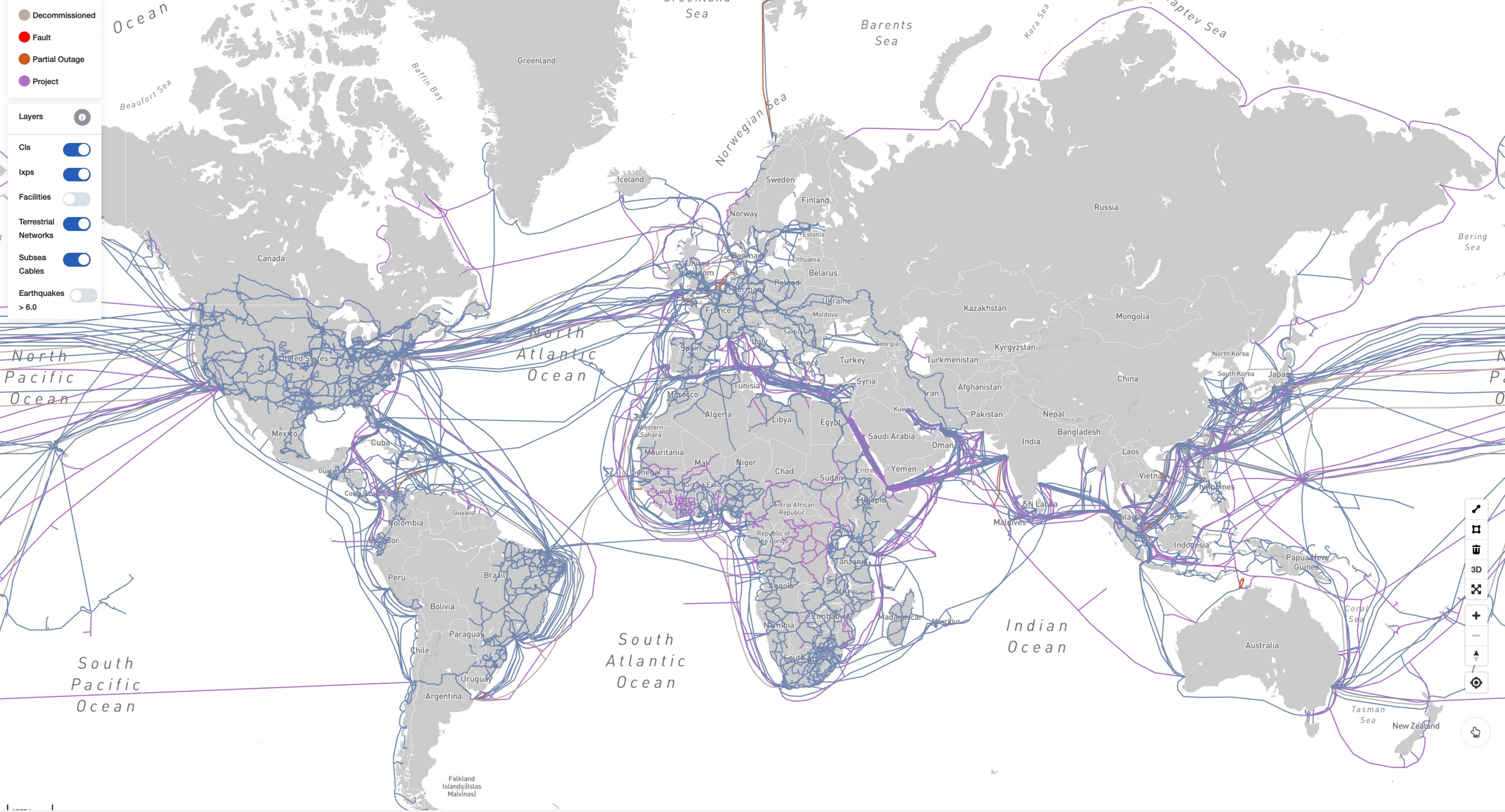
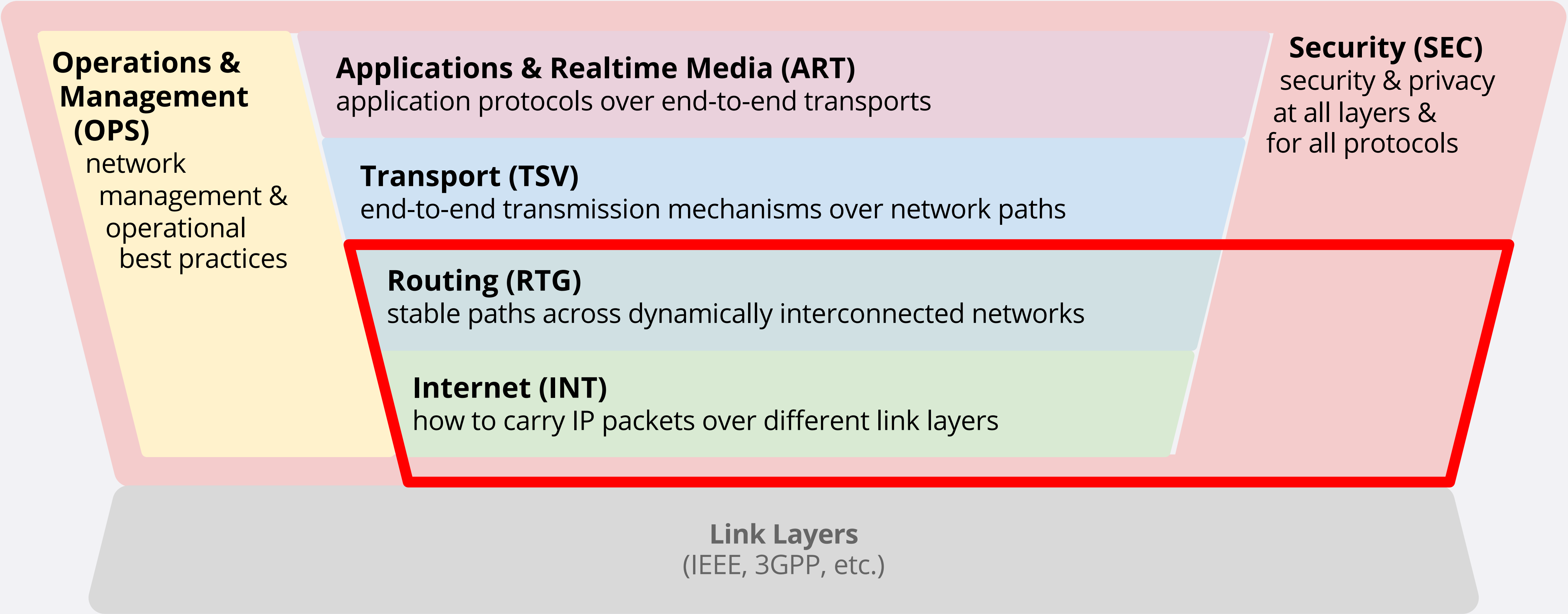
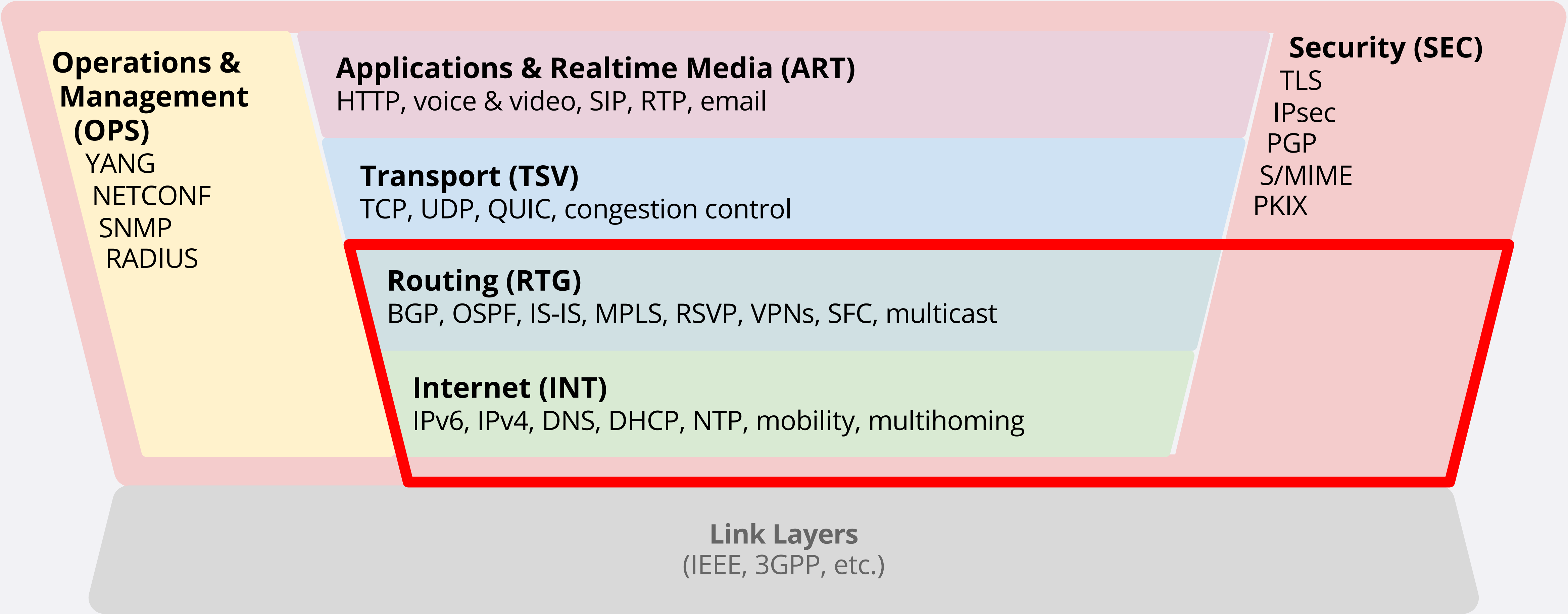


Image: https://www.infrapedia.com/

IETF
Work Areas



IETF
Key Technologies and Protocols



Internet-Infrastructure-Related Organizations

- Internet Engineering Task Force (**IETF**)
  - Develops and maintains Internet standards (RFCs) and protocols
- Internet Assigned Numbers Authority (**IANA**)
  - Coordination of DNS root, IP addresses, & other Internet resources
- Regional Internet Registry (RIR)
  - Manages allocation and registration of IP addresses within a region
- Internet Corporation for Assigned Names and Numbers (**ICANN**)
  - Coordinates DNS functions; contracts with registries (ccTLDs & others) and registrars (sellers of DNS names)

lars@dev ~ route -n

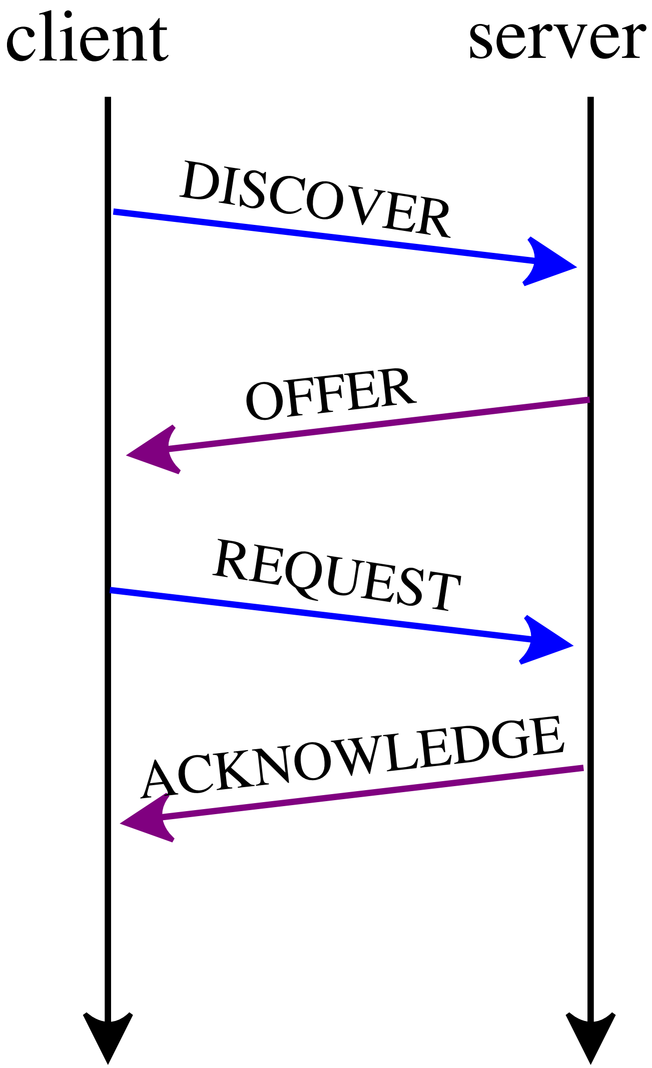
```
lars@dev ~ route -n
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use Iface
0.0.0.0         172.24.0.1      0.0.0.0         UG    100    0        0 ens3
172.24.0.0      0.0.0.0         255.248.0.0     U     100    0        0 ens3
172.24.0.1      0.0.0.0         255.255.255.255 UH    100    0        0 ens3
lars@dev ~
```

```
lars@dev ~ sudo tcpdump -i ens3 -v -n src port bootps
```

```
lars@dev  ~  sudo tcpdump -i ens3 -v -n src port bootps
tcpdump: listening on ens3, link-type EN10MB (Ethernet), snapshot length 262144 bytes
17:38:19.820045 IP (tos 0x0, ttl 64, id 56276, offset 0, flags [none], proto UDP (17), length 328)
    172.24.0.1.67 > 172.24.0.100.68: BOOTP/DHCP, Reply, length 300, xid 0xa6ef96d3, secs 3323, Flags [none]
          Client-IP 172.24.0.100
          Your-IP 172.24.0.100
          Client-Ethernet-Address 00:a0:98:11:cc:4f
          Vendor-rfc1048 Extensions
            Magic Cookie 0x63825363
            DHCP-Message (53), length 1: ACK
            Server-ID (54), length 4: 172.24.0.1
            Lease-Time (51), length 4: 300
            Subnet-Mask (1), length 4: 255.248.0.0
            Default-Gateway (3), length 4: 172.24.0.1
            Domain-Name-Server (6), length 4: 172.24.0.1
            Hostname (12), length 3: "dev"
            Domain-Name (15), length 10: "eggert.org"
```

Dynamic Host Configuration Protocol (**DHCP**)

- Each of the (many) DHCP deployment is independent
- No coordination between deployments
- Hence, not typically thought of as Internet infrastructure



"The **Dynamic Host Configuration Protocol** (**DHCP**) is a network management protocol used on Internet Protocol (IP) networks for automatically assigning IP addresses and other communication parameters to devices connected to the network using a client–server architecture."

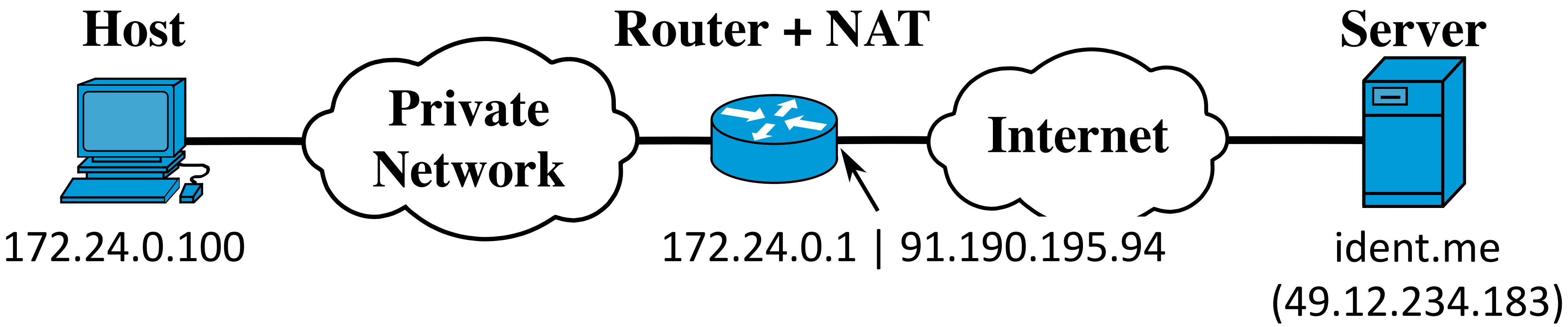
[Wikipedia]

```
lars@dev ~ curl -4 ident.me
```

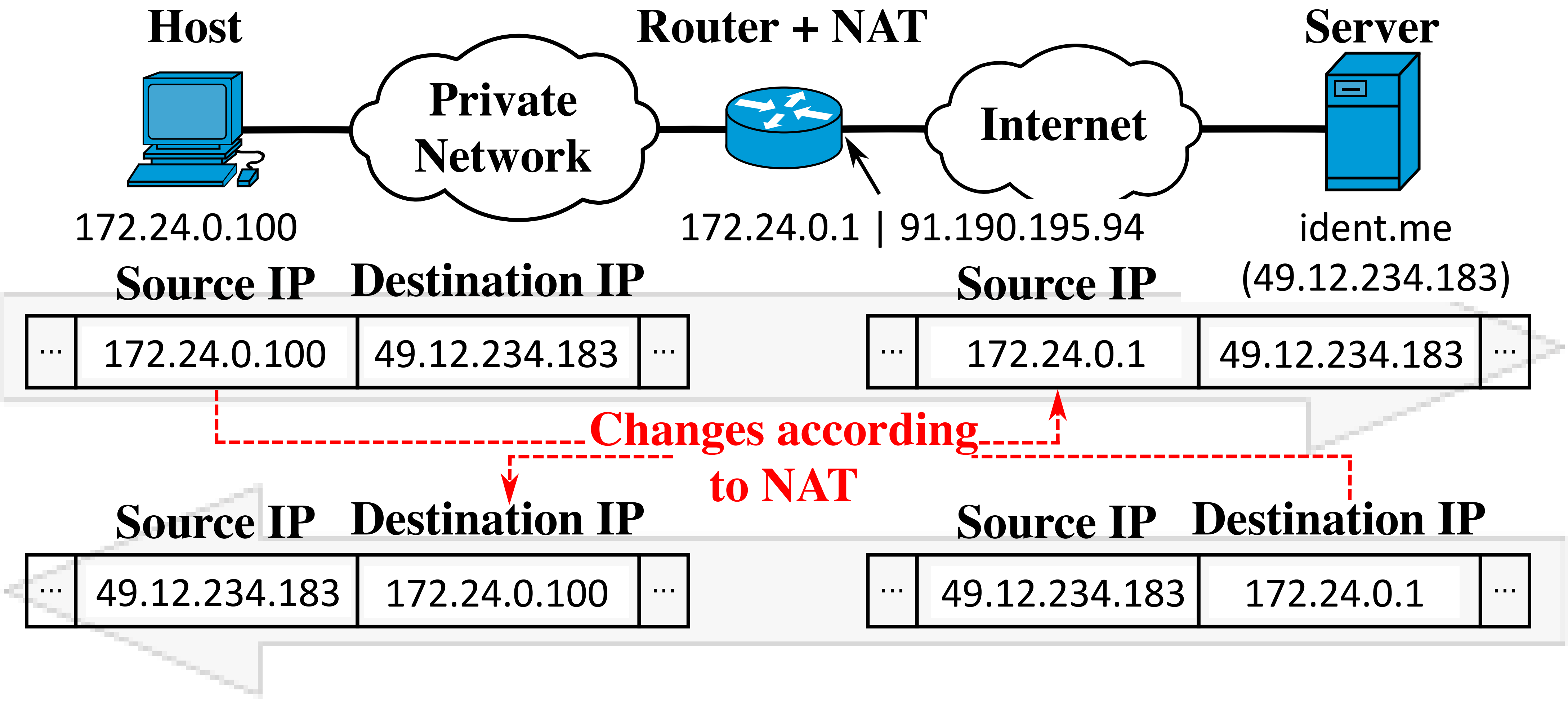
```
lars@dev ~ curl -4 ident.me
91.190.195.94⏎
```

```
lars@dev  ~  curl -4 ident.me
91.190.195.94⏎
lars@dev  ~  ip -4 addr show ens3
2: ens3: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP group default qlen 1000
    altname enp0s3
    inet 172.24.0.100/13 metric 100 brd 172.31.255.255 scope global dynamic ens3
       valid_lft 161sec preferred_lft 161sec
lars@dev  ~  
```

Network Address Translation (**NAT**)



Network Address Translation (**NAT**)



Network Address Translation (NAT)

- Each of the (many) NAT deployment is independent
- No coordination between deployments
  - Exception: NAT inside/behind NAT
  - This requires extensive coordination/configuration and is difficult to operate
- Servers inside/behind NAT require careful configuration
  - Port forwarding
  - Reverse proxy
- Not typically thought of as Internet infrastructure

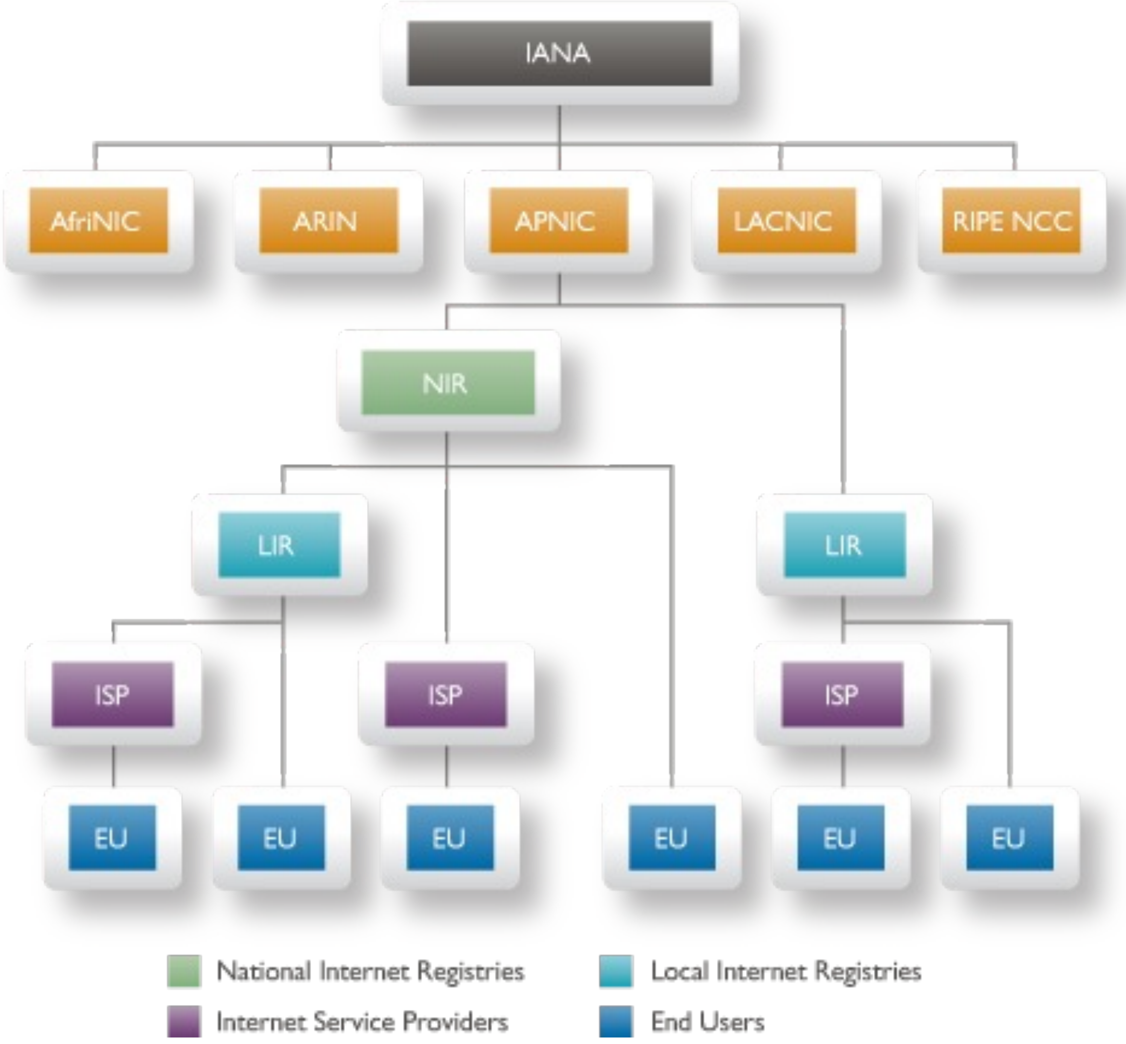
```
lars@dev ~ ip -6 addr show ens3
```

```
lars@dev ~ ip -6 addr show ens3
2: ens3: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP gr
oup default qlen 1000
    altname enp0s3
    inet6 2a00:ac00:4000:400:2a0:98ff:fe11:cc4f/64 scope global dynamic mngtmp
addr noprefixroute
       valid_lft 85885sec preferred_lft 13885sec
    inet6 fe80::2a0:98ff:fe11:cc4f/64 scope link
       valid_lft forever preferred_lft forever
```

```
lars@dev ~ ip -6 addr show ens3
2: ens3: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP group default qlen 1000
    altname enp0s3
    inet6 2a00:ac00:4000:400:2a0:98ff:fe11:cc4f/64 scope global dynamic mngtmpaddr noprefixroute
       valid_lft 85885sec preferred_lft 13885sec
    inet6 fe80::2a0:98ff:fe11:cc4f/64 scope link
       valid_lft forever preferred_lft forever
lars@dev ~ curl -6 ident.me
2a00:ac00:4000:400:2a0:98ff:fe11:cc4f⏎
lars@dev ~
```

IP Address Space Management

- **IANA** (Internet Assigned Numbers Authority) allocates address space to RIRs
- **RIR** (Regional Internet Registry) redistributes in its geographic region
- **Customers** (ISPs and end users) obtain address space from their RIR



APNIC. Understanding address management hierarchy. https://www.apnic.net/manage-ip/manage-resources/address-management-objectives-2/address-management-objectives/

```
lars@dev ~ whois -I 91.190.195.94 | grep -Ev '^%|^$' | head -n 14
```

```
lars@dev ~ whois -I 91.190.195.94 | grep -Ev '^%|^$' | head -n 14
refer:        whois.ripe.net
inetnum:      91.0.0.0 - 91.255.255.255
organisation: RIPE NCC
status:       ALLOCATED
whois:        whois.ripe.net
changed:      2005-06
source:       IANA
inetnum:        91.190.192.0 - 91.190.199.255
netname:        FI-SELTIMIL-20101014
country:        FI
org:            ORG-SO31-RIPE
admin-c:        ST5534-RIPE
tech-c:         ST5534-RIPE
status:         ALLOCATED PA
lars@dev ~
```

```
lars@dev ~ sudo traceroute -A -T -n ietf.org
traceroute to ietf.org (104.16.44.99), 30 hops max, 60 byte packets
 1  91.190.195.93 [AS51728]  6.541 ms  6.415 ms  6.377 ms
 2  * * *
 3  * * *
 4  * * *
 5  * * *
 6  62.115.44.164 [AS1299]  11.150 ms  11.307 ms  11.193 ms
 7  62.115.122.147 [AS1299]  12.151 ms  12.548 ms  12.447 ms
 8  213.248.94.67 [AS1299]  12.403 ms  12.338 ms  12.561 ms
 9  104.16.44.99 [AS13335]  12.193 ms  12.432 ms  12.071 ms
lars@dev ~
```

Autonomous System (**AS**)

- The Internet is a network of networks
- Each such network is also sometimes called an **Autonomous System**
  - Because they retain internal autonomy
- These ASs connect together, often at Internet Exchange Points (**IXPs**)
- ASs can be peers or have provider/customer relationships
  - This is individually negotiated

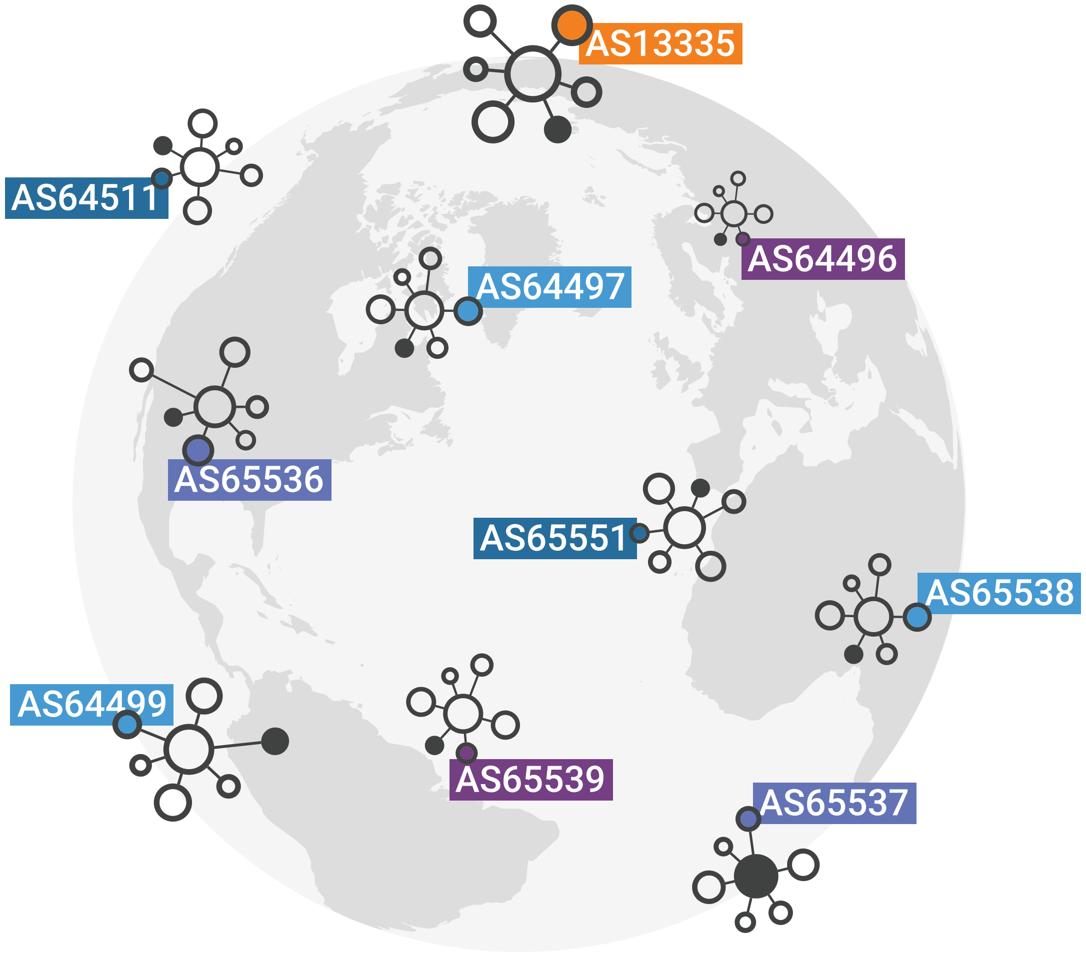


Image: https://www.cloudflare.com/learning/network-layer/what-is-an-autonomous-system/

```
lars@dev ~ asn AS51728

ASN lookup for AS51728

AS Number      ──> 51728
AS Name        ──> SELTIMIL-AS, FI
Organization   ──> Seltimil Oy (FI-SELTIMIL-20101222)
Abuse contact  ──> abuse@seltimil.fi
AS Reg. date   ──> 2010-10-03 12:42:11
Peering @IXPs  ──> FICIX 2 (Helsinki): IPv4+IPv6 MTU 1500
```

```
lars@dev  ~  asn AS1299

ASN lookup for AS1299

AS Number      ──> 1299
AS Name        ──> TWELVE99 Arelion, fka Telia Carrier, SE
Organization   ──> Arelion Sweden AB (SE-TWELVE99-20040510)
Abuse contact  ──> abuse@twelve99.net
AS Reg. date   ──> 2020-12-06 01:41:55
Peering @IXPs  ──> NONE
```

```
lars@dev ~ asn AS13335

 ASN lookup for AS13335

 AS Number       ──> 13335
 AS Name         ──> CLOUDFLARENET, US
 Organization    ──> CLOUDFLARENET (Cloudflare, Inc.)
 Abuse contact   ──> abuse@cloudflare.com
 AS Reg. date    ──> 2010-07-15 02:12:53
 Peering @IXPs   ──> 1-IX UA • 48 IX • ABQIX • AKL-IX (Auckland NZ): AKL-IX • AM
S-IX • AMS-IX BA • AMS-IX Caribbean • AMS-IX Chicago • AMS-IX Hong Kong • AMS-
IX Lagos: Main • Any2Denver • Any2East • Any2West • APE • AR-IX Cabase • Balca
n-IX • BALT-IX: BALT-IX • BBIX Chicago • BBIX Dallas • BBIX Fukuoka • BBIX Hon
g Kong • BBIX London • BBIX Marseille • BBIX Osaka • BBIX Singapore • BBIX Tok
yo • BBIX US-West • BCIX: BCIX Peering LAN • Beirut IX • BelgiumIX: Peeringlan
 • BFIX Ouagadougou: BFIX Ouaga2000 Peering LAN • Bharat IX - Mumbai: Bharat I
X Peering LAN • BiX • B-IX • BIX.BG: Main • BIX Jakarta • BNIX • Borneo-IX • B
oston Internet Exchange • btIX: TTPL-LAN • CAS-IX: Main • CATNIX • CHC-IX (Chr
```

Border Gateway Protocol (**BGP**)

- ASs exchange IP address reachability information via BGP
- Routers participating in the global BGP exchange then compute preferred next hops
- End system traffic is then forwarded to the computed next hop at each router
- `traceroute` makes these paths visible

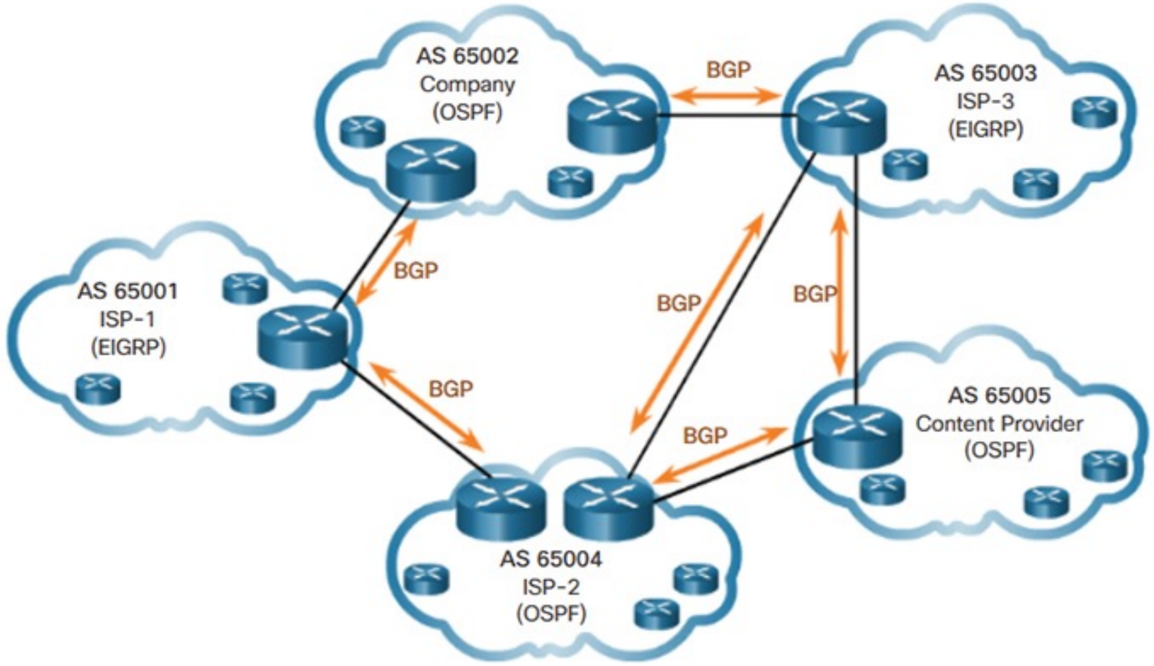
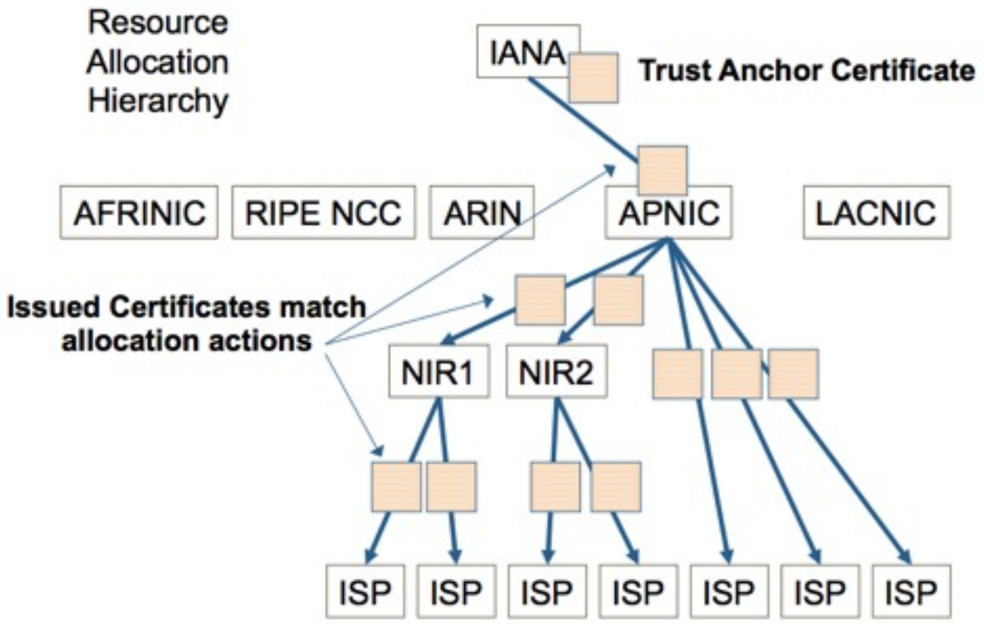
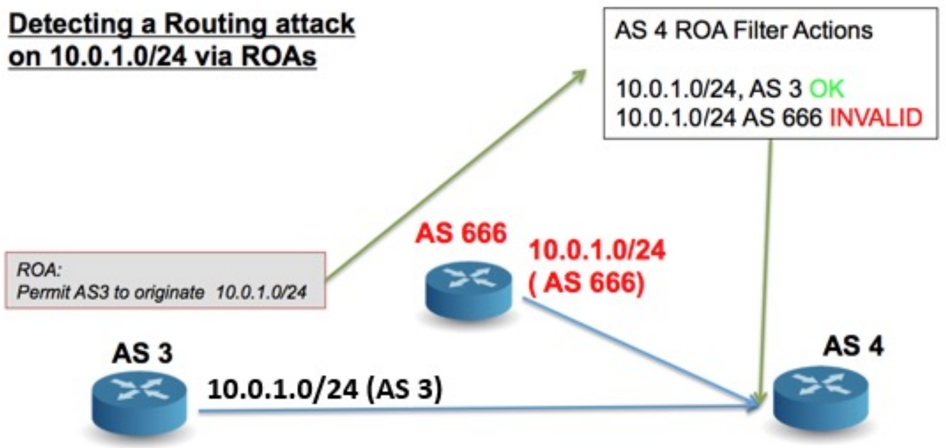


Image: https://devopedia.org/border-gateway-protocol

RPKI & BGPsec

- RPKI connects AS numbers (etc.) to a trust anchor
- Certificate structure mirrors the way in which Ass (etc.) are distributed
- BGPsec provides security for the path of ASes through which a BGP update message propagates





Images: https://www.potaroo.net/ispcol/2011-07/bgpsec.html

```
pi@raspberrypi:~ $ sudo traceroute -A -T -n ietf.org
```

```
pi@raspberrypi:~ $ sudo traceroute -A -T -n ietf.org
traceroute to ietf.org (104.16.45.99), 30 hops max, 60 byte packets
 1  192.168.2.1 [*]  0.550 ms  0.363 ms  0.335 ms
 2  62.155.246.159 [AS3320]  1.675 ms  1.621 ms  1.481 ms
 3  217.0.203.22 [AS3320]  4.835 ms 217.5.67.242 [AS3320]  4.988 ms  4.621 ms
 4  80.156.162.178 [AS3320]  15.259 ms  15.189 ms  15.067 ms
 5  * * *
 6  195.219.148.122 [AS6453]  5.221 ms * *
 7  162.158.108.2 [AS13335]  4.886 ms 162.158.84.53 [AS13335]  5.438 ms 172.70
.244.3 [AS13335]  9.078 ms
 8  104.16.45.99 [AS13335]  3.779 ms  4.520 ms  4.320 ms
pi@raspberrypi:~ $
```

Content Delivery Network (**CDN**)

- A CDN replicates content and services at many different points on the Internet
- Improves user experience, performance and resiliency
- Different types of CDNs
  - DNS-based, anycast
  - For web and other content
  - From hyperscalars and specialized providers

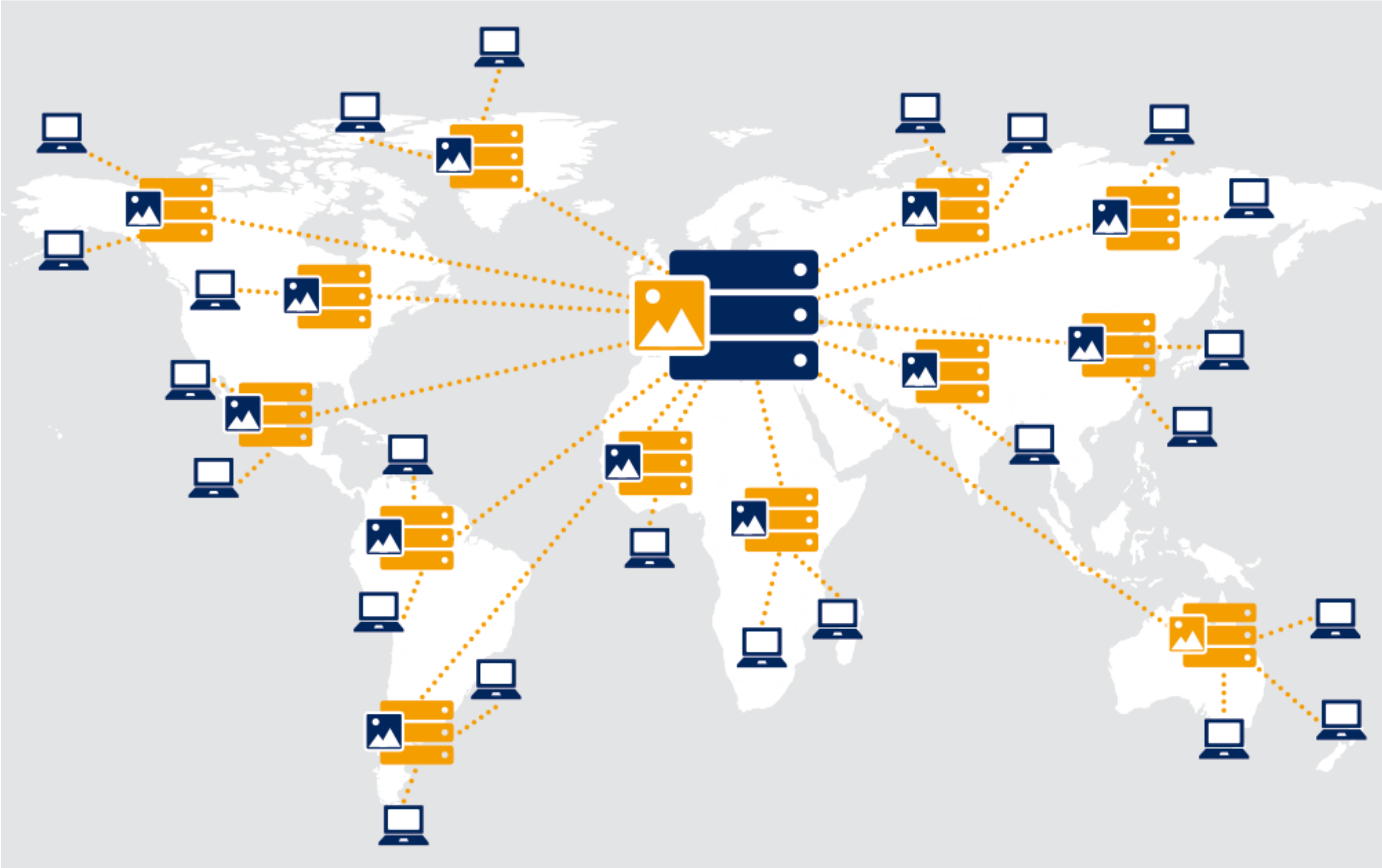


Image: https://blog.paessler.com/hubfs/cdn-orange.png

```
pi@raspberrypi:~ $ ping -4 -n -c 10 eggert.org
```

```
pi@raspberrypi:~ $ ping -4 -n -c 10 eggert.org
PING  (91.190.195.94) 56(84) bytes of data.
64 bytes from 91.190.195.94: icmp_seq=1 ttl=54 time=46.5 ms
64 bytes from 91.190.195.94: icmp_seq=2 ttl=54 time=46.0 ms
64 bytes from 91.190.195.94: icmp_seq=3 ttl=54 time=45.9 ms
64 bytes from 91.190.195.94: icmp_seq=4 ttl=54 time=45.8 ms
64 bytes from 91.190.195.94: icmp_seq=5 ttl=54 time=46.1 ms
64 bytes from 91.190.195.94: icmp_seq=6 ttl=54 time=46.0 ms
64 bytes from 91.190.195.94: icmp_seq=7 ttl=54 time=46.1 ms
64 bytes from 91.190.195.94: icmp_seq=8 ttl=54 time=46.2 ms
64 bytes from 91.190.195.94: icmp_seq=9 ttl=54 time=46.0 ms
64 bytes from 91.190.195.94: icmp_seq=10 ttl=54 time=45.9 ms

---  ping statistics ---
10 packets transmitted, 10 received, 0% packet loss, time 9013ms
rtt min/avg/max/mdev = 45.757/46.050/46.472/0.181 ms
pi@raspberrypi:~ $
```

```
lars@dev ~ cat /etc/resolv.conf | grep -v '^#'

nameserver 127.0.0.53
options edns0 trust-ad
search eggert.org
lars@dev ~ resolvectl status ens3
Link 2 (ens3)
    Current Scopes: DNS
         Protocols: +DefaultRoute +LLMNR -mDNS -DNSOverTLS
                    DNSSEC=yes/supported
Current DNS Server: 172.24.0.1
       DNS Servers: 172.24.0.1 2a00:ac00:4000:400::1
        DNS Domain: eggert.org
lars@dev ~
```

```
lars@dev ~ host -a ietf.org
```

```
lars@dev ~ host -a ietf.org
Trying "ietf.org"
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 22796
;; flags: qr rd ra; QUERY: 1, ANSWER: 5, AUTHORITY: 0, ADDITIONAL: 0

;; QUESTION SECTION:
;ietf.org.                      IN      ANY

;; ANSWER SECTION:
ietf.org.               294     IN      A       104.16.44.99
ietf.org.               294     IN      A       104.16.45.99
ietf.org.               294     IN      AAAA    2606:4700::6810:2c63
ietf.org.               294     IN      AAAA    2606:4700::6810:2d63
ietf.org.               105     IN      MX      0 mail.ietf.org.

Received 135 bytes from 127.0.0.53#53 in 4 ms
lars@dev ~
```

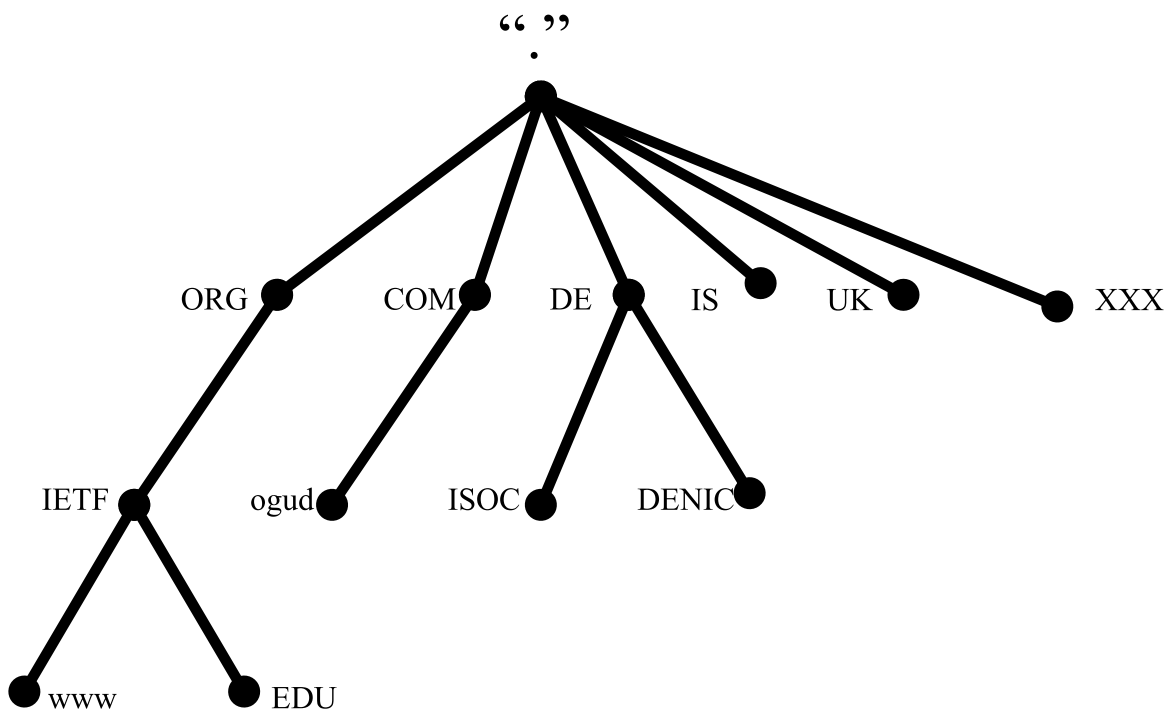
```
lars@dev ~ whois -I ietf.org | head -n 20
```

```
lars@dev ~ whois -I ietf.org | head -n 20
% IANA WHOIS server
% for more information on IANA, visit http://www.iana.org
% This query returned 1 object

refer:        whois.publicinterestregistry.org

domain:       ORG

organisation: Public Interest Registry (PIR)
address:      11911 Freedom Drive,
address:      10th Floor, Suite 1000
address:      Reston VA 20190
address:      United States of America (the)

contact:      administrative
name:         Director of Operations, Compliance and Customer Support
organisation: Public Interest Registry (PIR)
address:      11911 Freedom Drive,
address:      10th Floor, Suite 1000
```

What is DNS?

- DNS is one of the core Internet Protocols required for operation of the Internet
- Routing and DNS are the most important infrastructure protocols as without them nothing else will work
- DNS Provides:
  - Mapping from names to addresses
  - Mechanism to store and retrieve information in a global data store

DNS tree



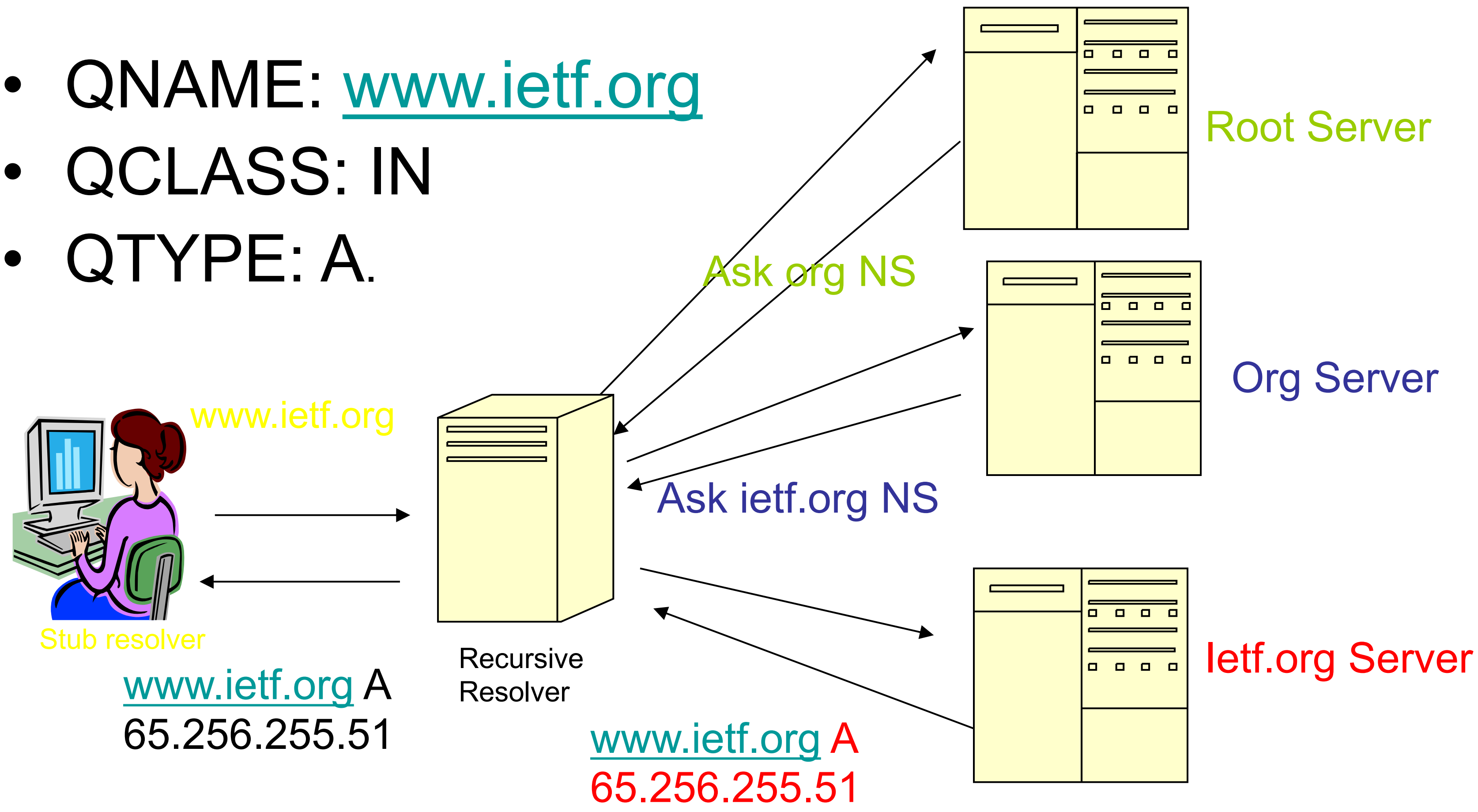
DNS Elements

- Resolver
  - stub: simple, only asks questions
  - recursive: takes simple query and makes all necessary steps to get the full answer,
- Server
  - authoritative: the servers that contain the zone file for a zone, one Primary, one or more Secondaries,
  - caching: A recursive resolver that stores prior results and reuses them
    - Some perform both roles at the same time.

ogud@ogud.com & pk@denic.de

DNS query

- QNAME: www.ietf.org
- QCLASS: IN
- QTYPE: A.



DNSSEC

- **DNSSEC** provides data authentication and integrity, and authenticated denial of existence
  - but not availability or confidentiality
- DNSSEC digitally signs records with public-key cryptography
- A DNSKEY record is authenticated via a chain of trust
  - Starting with verified public keys for the DNS root zone = the trusted third party
- Domain owners generate their own keys publish them to their registrar
- Which in turn pushes them to the zone operator who signs and publishes them in DNS

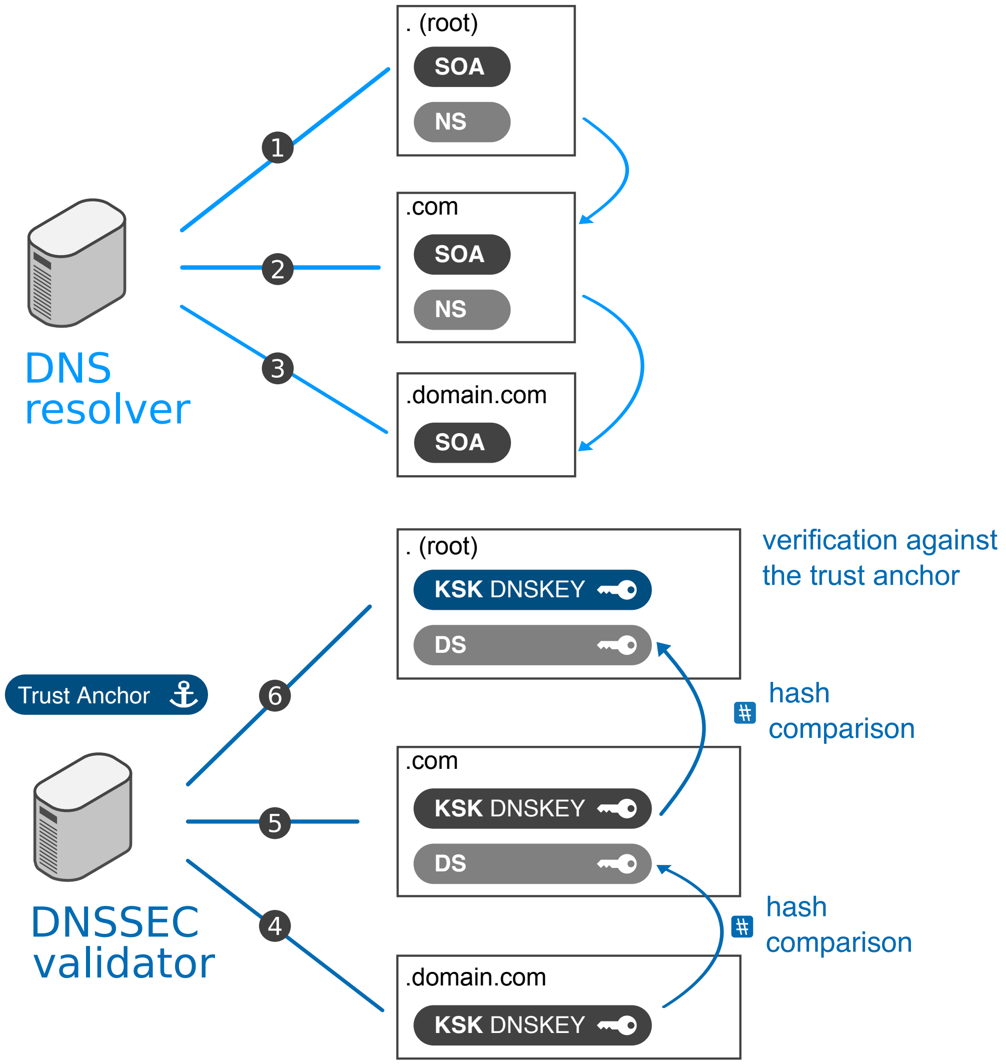
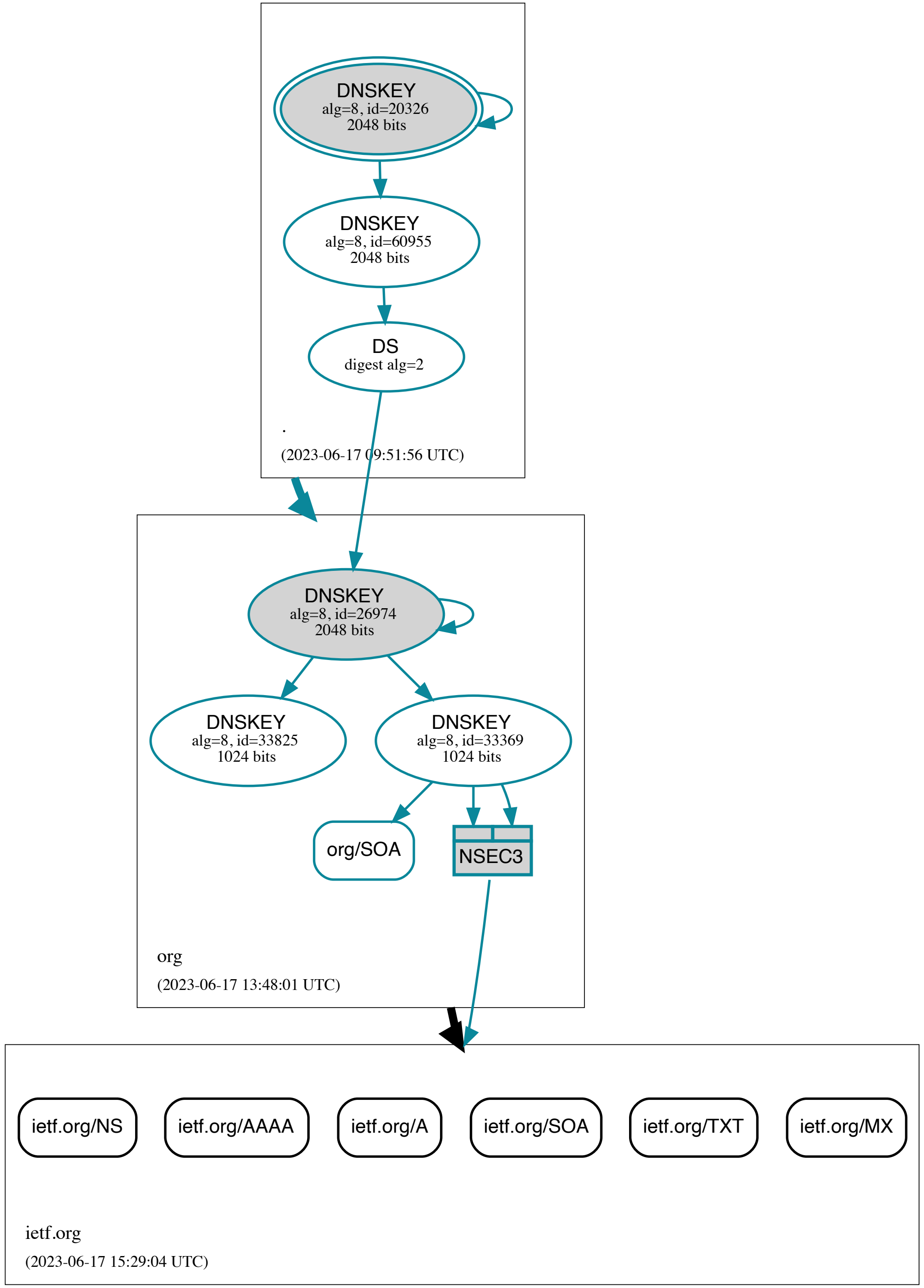
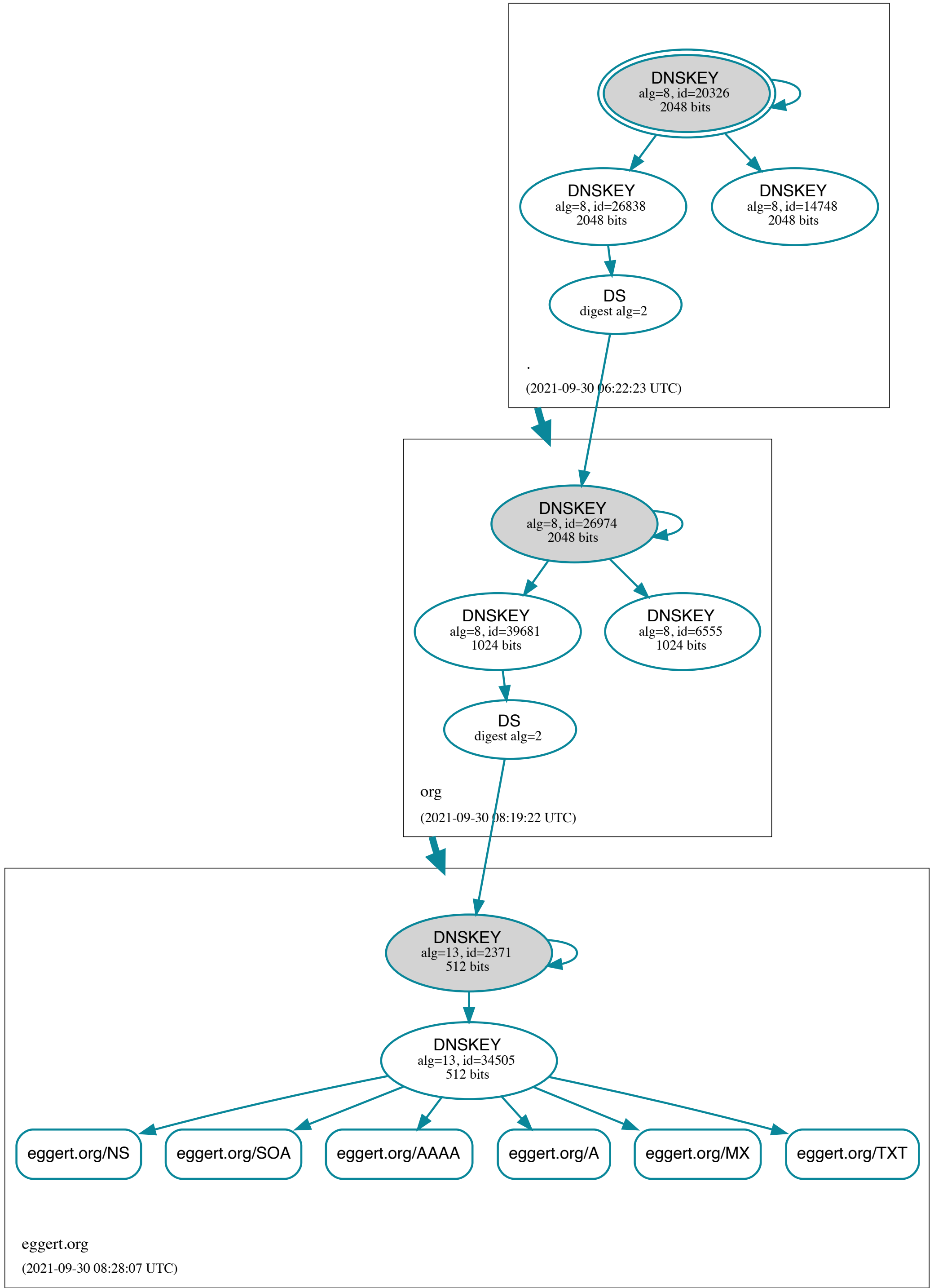


Image: https://efficientip.com/glossary/what-is-dnssec/

```
lars@dev ~ delv ietf.org
; unsigned answer
ietf.org.               300     IN      A       104.16.44.99
ietf.org.               300     IN      A       104.16.45.99
```

```
lars@dev ~ delv ietf.org
; unsigned answer
ietf.org.               300     IN      A       104.16.44.99
ietf.org.               300     IN      A       104.16.45.99
lars@dev ~ delv eggert.org
; fully validated
eggert.org.             49      IN      A       91.190.195.94
eggert.org.             49      IN      RRSIG   A 13 2 300 20230618202915 2023
0616182915 34505 eggert.org. LCnxqYebrTROnuYwCXaX91lrT8tDmfda11/AfWBB988XpwM31
ZORnHkG sWjtC6IrNXGWBlviuYW99IpCUbMGqA==
lars@dev ~
```


DNSKEY
alg=8, id=20326
2048 bits
DNSKEY
alg=8, id=60955
2048 bits
DS
digest alg=2
.
(2023-06-17 09:51:56 UTC)
DNSKEY
alg=8, id=26974
2048 bits
DNSKEY
alg=8, id=33825
1024 bits
DNSKEY
alg=8, id=33369
1024 bits
org/SOA
NSEC3
org
(2023-06-17 13:48:01 UTC)
ietf.org/NS
ietf.org/AAAA
ietf.org/A
ietf.org/SOA
ietf.org/TXT
ietf.org/MX
ietf.org
(2023-06-17 15:29:04 UTC)


DNSKEY
alg=8, id=20326
2048 bits
DNSKEY
alg=8, id=26838
2048 bits
DNSKEY
alg=8, id=14748
2048 bits
DS
digest alg=2
.
(2021-09-30 06:22:23 UTC)
DNSKEY
alg=8, id=26974
2048 bits
DNSKEY
alg=8, id=39681
1024 bits
DNSKEY
alg=8, id=6555
1024 bits
DS
digest alg=2
org
(2021-09-30 08:19:22 UTC)
DNSKEY
alg=13, id=2371
512 bits
DNSKEY
alg=13, id=34505
512 bits
eggert.org/NS
eggert.org/SOA
eggert.org/AAAA
eggert.org/A
eggert.org/MX
eggert.org/TXT
eggert.org
(2021-09-30 08:28:07 UTC)

Conclusion

Covered

- IP addresses
- ASs
- BGP, ~BGPsec
- DNS, ~DNSSEC
- WHOIS

- (DHCP, NAT, CDN)

Not Covered

- Network Time Protocol (**NTP**)
- HTTPS, TLS & WebPKI
- Global services at content layer
- Physical layer (fiber, cables, satellites, etc.)

Thank you!

Questions later? lars@eggert.org