

Anycasting DNS infrastructure in .CZ

CZ.NIC z.s.p.o.
Jaromir Talir
jaromir.talir@nic.cz
6. 5. 2010

Agenda

- DNS infrastructure, Anycast
- The past and the future
- Steps we had to make
- Issues we had to face
- Simple node distribution measurement
- Conclusions and plans

DNS Infrastructure

- 680 000 domains
- 8 000 qps
- DNS service fully maintained by us
- 9 localities for slave servers
 - Server hosting of two servers (AMD, Intel, Sparc)
 - Transit connectivity + peering (optional)
 - DNS server (NSD, BIND)
 - DSC for statistic collection

Anycast

- Routing technology – announcing one IP address from several different places
 - Special pool/policy for these address in RIRs
- Advantages
 - Reliability
 - Speed

Those were the days...

Location/IP version	IPv4	IPv6
CZ-1	UNI-1	UNI-1
CZ-2	UNI-2	UNI-2
CZ-3	ANY-1	ANY-1
SE	UNI-3	
GB	UNI-4	UNI-4
AT	UNI-5	UNI-5
DE	ANY-1	ANY-1
US	ANY-1	ANY-1
CL	ANY-1	ANY-1

...our plans

Location/IP version	IPv4	IPv6
CZ-1	ANY-1	ANY-1
CZ-2	ANY-4	ANY-4
US-2	ANY-2	ANY-2
SE	ANY-2	ANY-2
GB	ANY-3	ANY-3
AT	ANY-3	ANY-3
DE	ANY-1	ANY-1
US-1	ANY-1	ANY-1
CL	ANY-4	ANY-4

First steps – change of policy

- Former policy:
 - 1 Anycast IP address per TLD
 - Proof of necessity
- Current policy:
 - Initiated by CZ.NIC and Nominet
 - Accepted on RIPE-58 at Lisbon
 - 4 Anycast addresses per TLD or ENUM operator
 - Less strict

Next steps - updating servers

- Flexibility with BGP daemons (BIRD, Quagga, OpenBGPD)
- Negotiation with connectivity partners (BGP session, filters...)
- Redundancy using BGP (AS path prepend)
- Monitoring scripts to kill BGP in case NS die
- Finally updating root zone through IANA

Current state

Host Name	IP Address(es)
a.ns.nic.cz.	194.0.12.1 2001:678:f:0:0:0:0:1
b.ns.nic.cz.	194.0.13.1 2001:678:10:0:0:0:0:1
c.ns.nic.cz.	194.0.14.1 2001:678:11:0:0:0:0:1
d.ns.nic.cz.	193.29.206.1 2001:678:1:0:0:0:0:1
f.ns.nic.cz.	193.171.255.48 2001:628:453:420:0:0:0:48

Current state

Location/IP version	IPv4	IPv6
CZ-1	ANY-1	ANY-1
CZ-2	ANY-4	ANY-4
US-2	ANY-2	ANY-2
SE	ANY-2	ANY-2
GB	ANY-3	ANY-3
AT	UNI-1	UNI-1
DE	ANY-1	ANY-1
US-1	ANY-1	ANY-1
CL	ANY-4	ANY-4

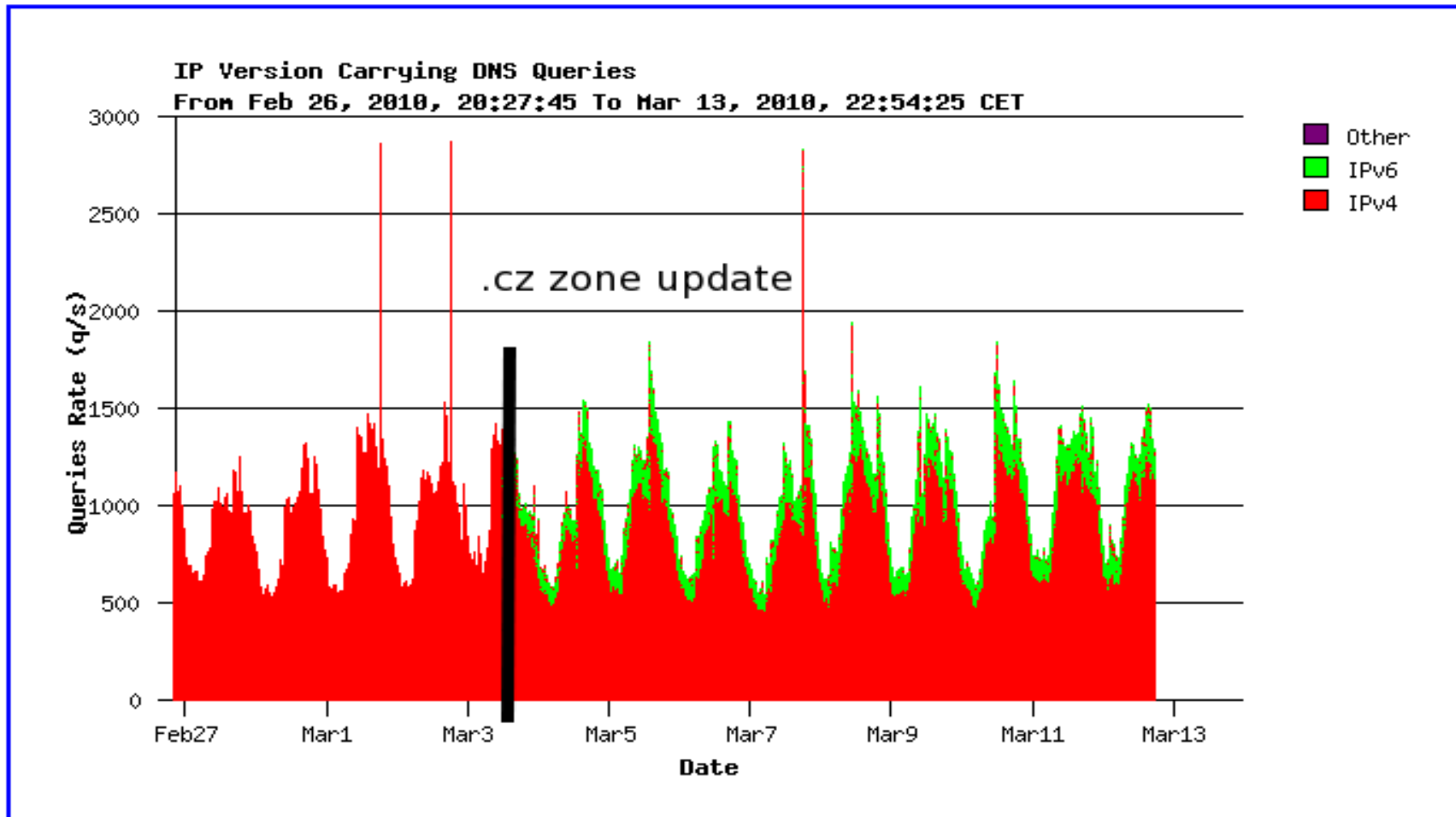
Current state



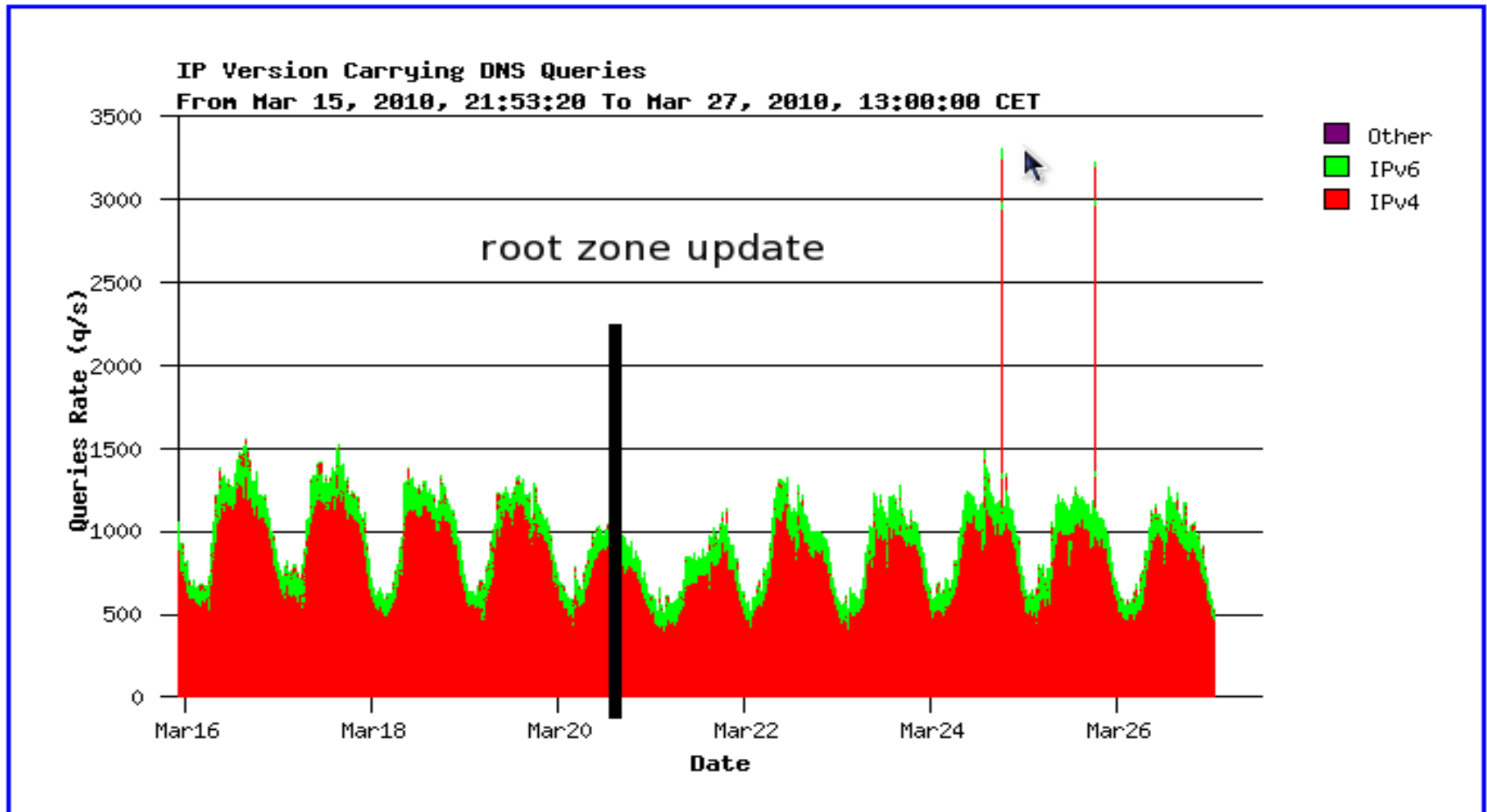
Change to full IPv4/IPv6

- Sweden node had just IPv4
- Fear of bad resolvers that will just go to next nameserver in case IPv6 problems
- Queries to mailing lists, checking root zone
- Should be OK.
- We were looking for decrease of Ipv4 query load

Change to full IPv4/IPv6



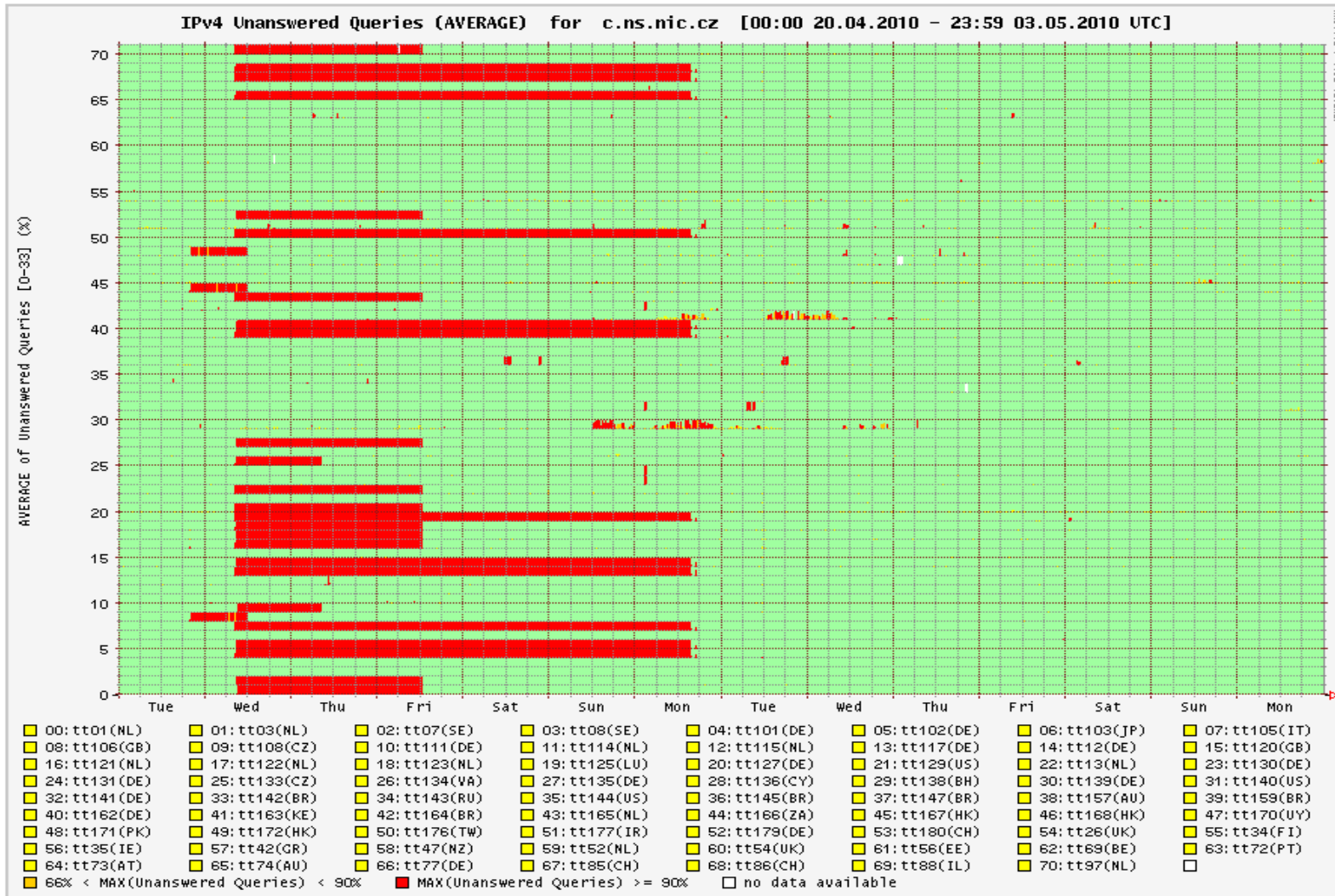
Change to full IPv4/IPv6



Anycast address filtering

- How to be sure address is propagated correctly?
- Ping from every DNS location succeeded
- Put address into zonefile and look to DNSMON
- The game begin: get rid of red color in favor of green – we had to contact 3 ASs to win...

Anycast address filtering



Ping RTT distance measurement

- Most simple approach
 - Choosing several IP addresses from each country from DSC
 - Each nameserver ping with source anycast address to these addresses
 - If it returns, store minimum RTT
- What do we get?
 - Approximate “distance” of our nameservers from different countries
 - Which nameserver is “responsible” for that country

Find businesses, addresses and places of interest.

Get Directions My Maps

Save to My Maps

Displaying content from www.nic.cz

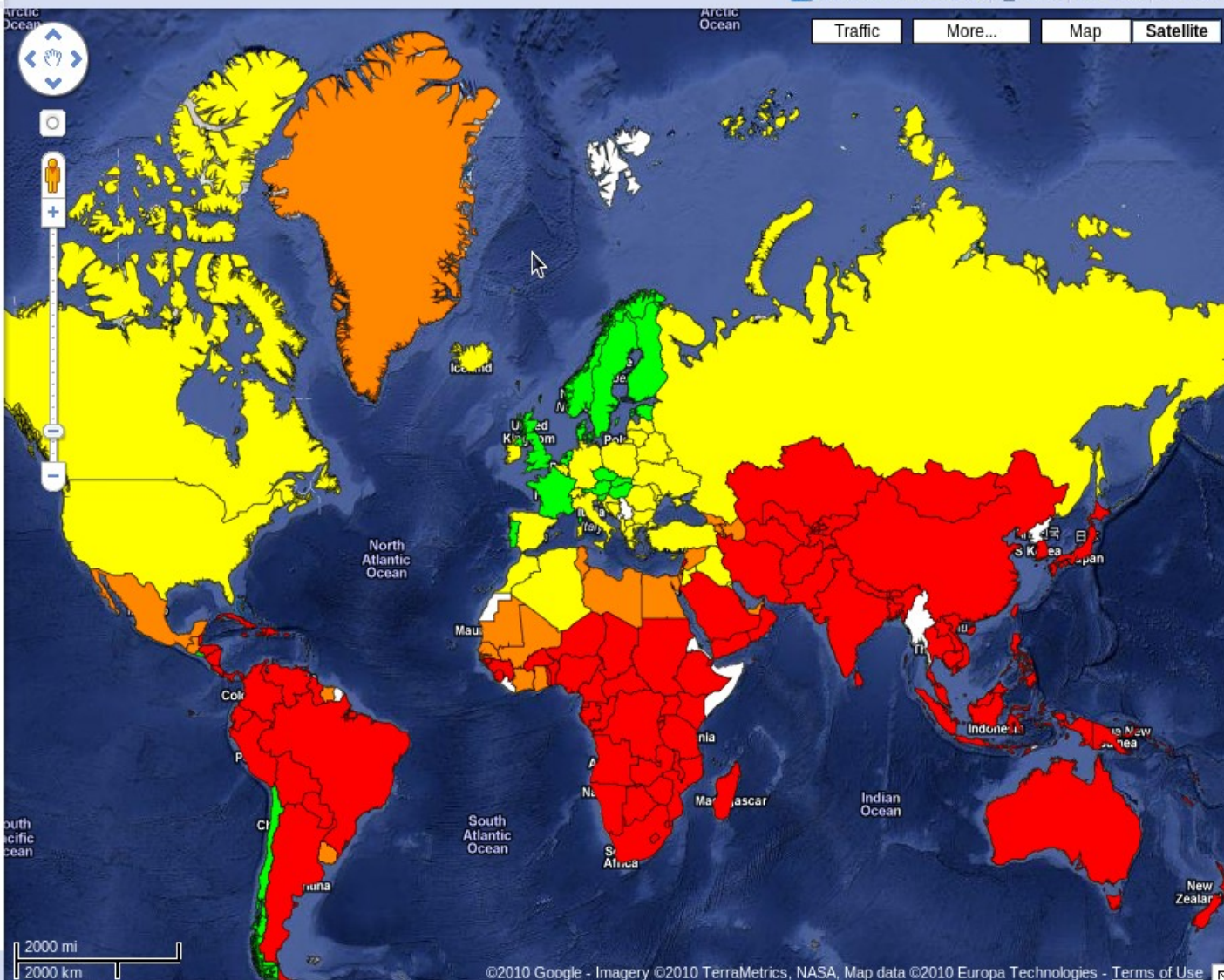
The content displayed below and overlaid onto this map is provided by a third party, and Google is not responsible for it. Information you enter below may become available to the third party.

Node with the lowest pingtime

Legend

- unknown
- below 10 ms
- 10-50 ms
- 50-100 ms
- over 100 ms

Countries



Find businesses, addresses and places of interest.

Get Directions My Maps

Save to My Maps

View in Google Earth Print Send Link

Traffic More... Map Satellite

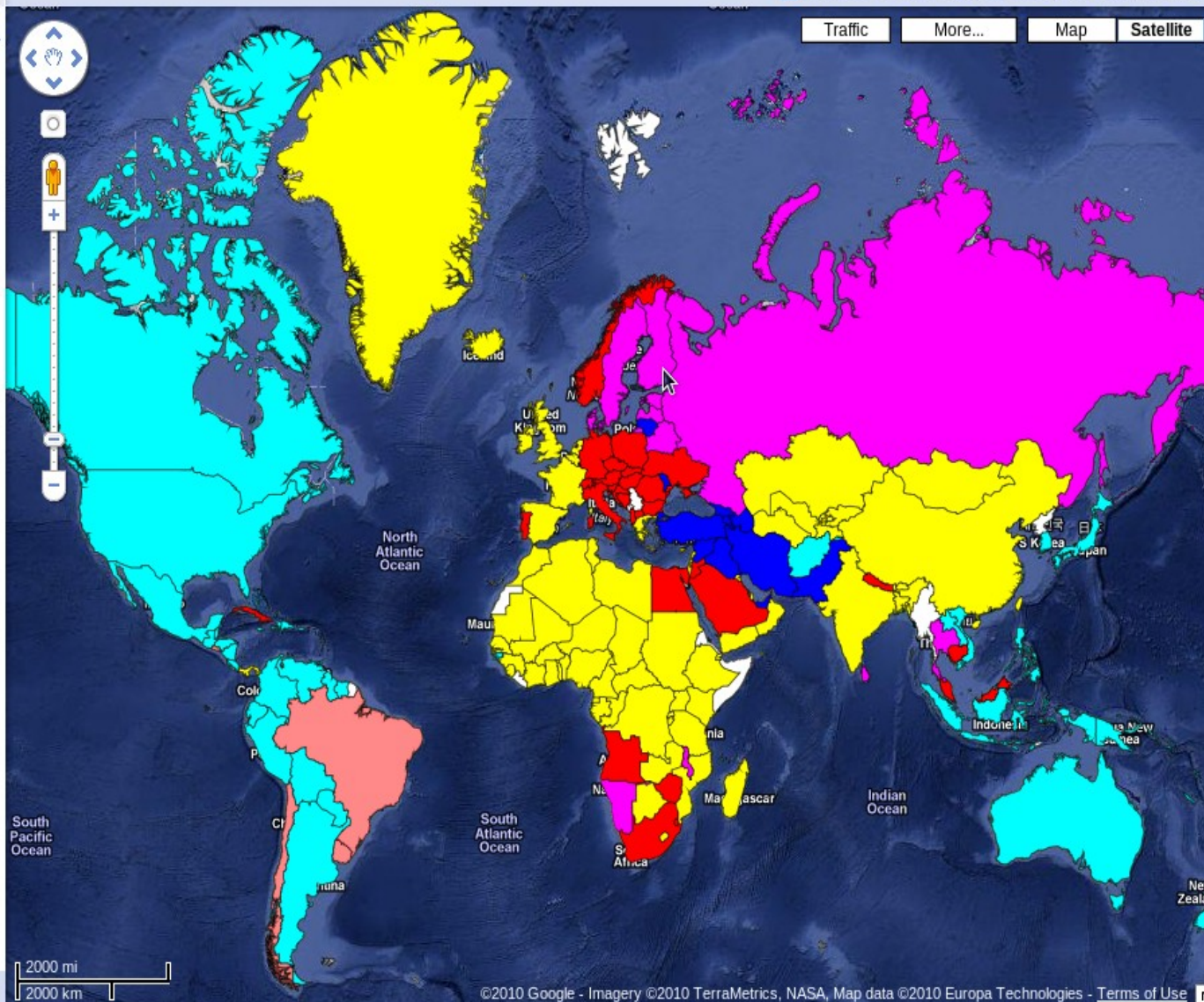
Displaying content from www.nic.cz

The content displayed below and overlaid onto this map is provided by a third party, and Google is not responsible for it. Information you enter below may become available to the third party.

Location with the lowest pingtime

Legend

- unknown
- CZ
- DE
- UK
- SE
- US/ISC
- US/ICANN
- CL
- Countries



Node with the lo... | Lowest pingtime ...

Conclusions

- There are still some places not accessible using anycast
 - Let last unicast exists and explore it for a while
- IANA technical checks
 - Would it be nice to enable technical checks to TLD operators as a web service?

Plans

- Close to start our new anycast node in Tokyo
- Explore more seriously distribution of anycast addresses
- Support developing countries by free DNS hosting service (currently .tz, .ao)



Questions??

jaromir.talir@nic.cz