DNS anycast in CZ

Jaromir Talir • **jaromir.talir@nic.cz** • **7.3.2016**

Agenda

- IP structure

- Geographical structure

- Software and hardware diversity

- DNS hosting

- Thoughts about future

IP structure

- Since 2009 we have 4 anycast prefixes

  - IPv4 + IPv6

  - Change of policy in RIPE (only 1 anycast prefix allowed at that time)

  - Names are [abcd].ns.nic.cz

- Until 2013 we had 1 unicast using Vienna university IP address

  - Decision to drop it in favor of stronger anycast

  - Blocked by IANA - MinimumNetworkDiversityCheck

IP structure

- We asked RIPE for another AS

  - Got 200070

  - RIPE requirement is to have "new unique routing policy" for new AS – limitation for using this AS

  - [abd].ns.nic.cz distributed from AS25192

  - c.ns.nic.cz distributed from AS200070

- There are planned changes in IANA technical checks to relax this rule

Geographical structure – foreign

- Vienna (AT) – global + VIX indirectly

- Frankfurt (DE) – global + DECIX directly

- London (UK) – global + LINX indirectly

- Stockholm (SE) – global

- Tokyo (JP) – local in JPNAP directly

- Redwood City (US) – global

- Reston (US) – global

- Santiago de Chile (CL) - global

Geographical structure – foreign

- Two physical servers in hosted environment

- Both servers announces single anycast prefix
    - Each server has BGP connection with host
    - One server has lower metric
    - No load-balancing negotiated with host

- Host expect any of our prefixes
    - We can easily switch to another anycast prefix
    - We can potentially announce all prefixes

Geographical structure – foreign

- https://www.nic.cz/infrastructure/
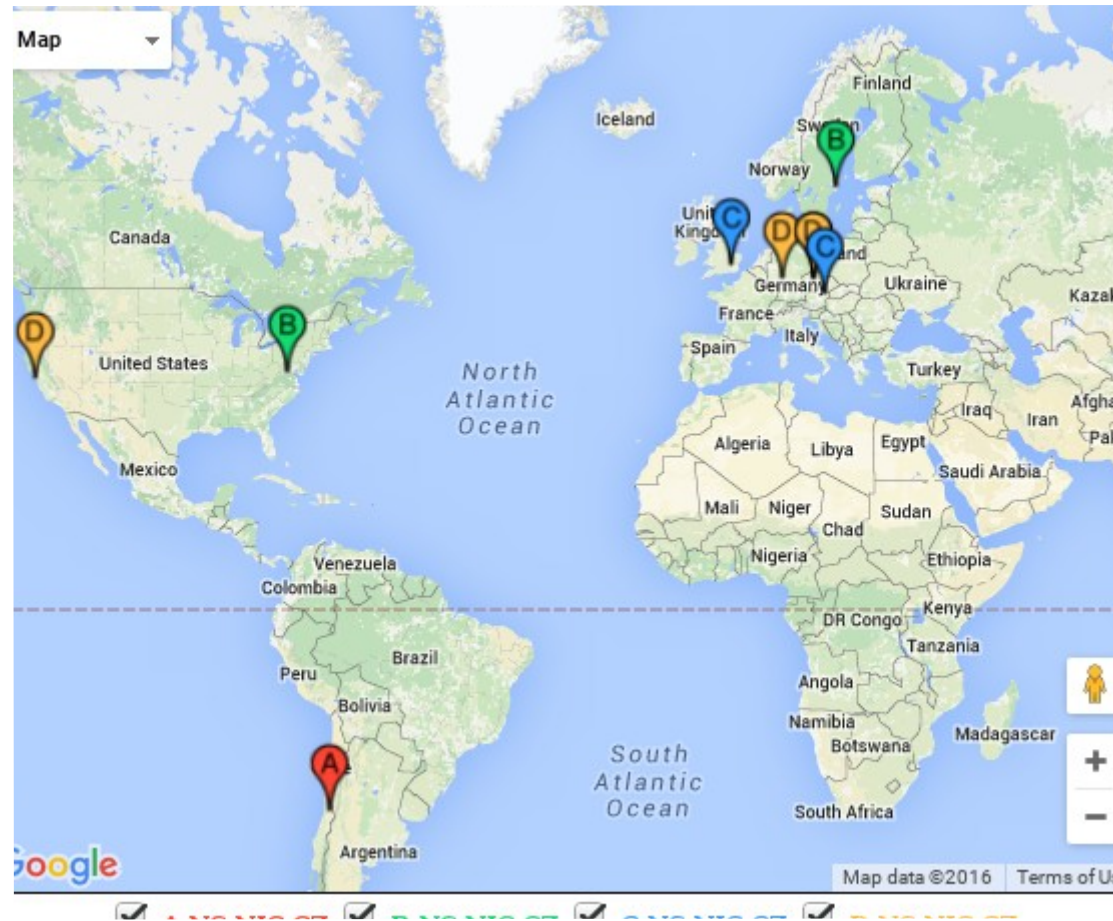
Geographical structure – foreign

- Measurement of reachability

  - Several pingable addresses selected in each country - based on DSC statistics

  - Ping from one anycast to this address

  - Listening on every instance – one should see the response.

- Currently RIPE Atlas is much better tool for such analysis

Geographical structure – local in CZ

- Upgrade in 2016

- 6→15 servers

- 1Gb→10Gb

  - based on CZ.NIC Labs performance tests we selected network cards Intel X520

- 8GB → 32GB RAM

Geographical structure – local in CZ

- Geographically distributed, but in a single VLAN

- Load balanced via BGP multipath

- Each server announce single prefix
  - 5x A, 5x B, 5x D

- 3 Tranzit providers (each 10Gb)

- Local IXPs

  - 2x 10Gb to NIX.CZ

  - 1x 10Gb to NIX.SK

cz.nic | CZ DOMAIN REGISTRY

Software diversity

- Operating systems
  - Ubuntu, Debian, OpenBSD
- DNS servers
  - **Knot**, NSD, Bind
  - All 3 installed on each server
- BGP servers
  - **Bird**, Quagga, OpenBGPD
- Orchestrated by Ansible

Hardware diversity

- Servers
  - Dell, HP, Intel
- Switches
  - Juniper EX4550, Cisco Nexus 5548
  - Tried Brocade ICX 6650 but not satisfied
- Routers
  - Juniper MX480, Cisco ASR9006
  - Tried Brocade CER2024C-4X-RT but not satisfied

DNS hosting

- We offer secondary DNS hosting for a few other registries

  - Angola – it.ao, co.ao

  - Tanzania – tz, ac.tz, go.tz, ne.tz, co.tz, or.tz

  - Macedonia – mk

- Free of charge

  - Partner registries that use our open source registry solution **Fred** – https://fred.nic.cz

  - Anyone interested?

Thoughts about future

- Better distribution of prefixes?

- Virtualization?

- More locations?
    - We are looking for partners
    - TLD at ISP?
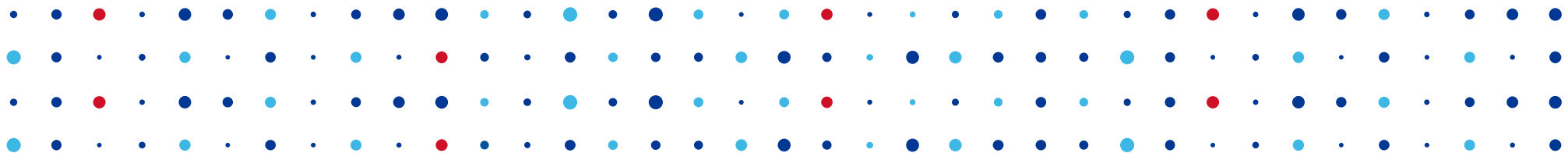
Thank You

Jaromir Talir • jaromir.talir@nic.cz