Challenges of deploying anycast:
The F-root and local experience

Sebastián Castro Avila

DNS Administrator

NIC Chile

Background information

- 34 Internet providers.
- No national IX point.
- Several partial IX points, among some providers.
  - An some IX connected among them also.
- Government efforts to
  - Encourage some minimum levels of Internet connectivity among providers.

A F-root for Chile

- Promoted by LACNIC project called "+Raices"
- Deployed by ISC
- LACNIC pays the maintenance fee for three years
- Who pays for the initial setup: rackspace, bandwidth, servers?

Looking for sponsorship

- Search in government/private organizations
  - What is a root server anyway?
  - What do we get having one?

- No positive answer for weeks
  - Self sponsoring was decided.

- A friend on the way
  - Adexus provided bandwidth and rackspace for a very affordable price.

http://www.nic.cl

NIC Chile as transit provider

- Due to time constraints (one person from ISC traveled to the country), the link was requested to be available within three days

- Using NIC Chile providers and bandwidth, we became transit ASN for F root.

Getting connectivity for F

- **The lobby for independent connectivity**
  - Several meetings with people from major ISP
  - The common questions were
    - What is a root server?
    - What do we get getting connected?
    - Who's going to pay for the link?
    - Do we save money?

DNS-Ops Meeting

Getting connectivity for F

- ## The poker strategy
  - This provider is going to be connected, what about you?
  - How would you look in the media if your competitor cooperates with this project?

- ## The results
  - Those organizations became sensitive to the importance of the project decided to support

http://www.nic.cl

The anomalies

- The F-root instance installed in Santiago is considered as local
  - The prefix is announced with NO_EXPORT attribute
- This condition let us detect
  - One provider using lots of static routes.
  - One provider peering to a second ASN of his own: no one prefered that route because the AS-Path became artificially long.

http://www.nic.cl

The current state

- Four providers connected: two of them arrive on first place.

- Two more on the way (still working on it after months)

- Our gratitude to:
  - Carlos Buzio, from Adexus, from sponsoring the project.
  - Edgardo Krell, Operations Director at NIC Chile, for his political efforts.
  - Eduardo Mercader and Rodrigo Arenas, Network Engineers at NIC Chile, for dealing with the providers and explaining what F-root is intended for.

http://www.nic.cl

The future

- Due to all the problems arised, some research seems to be needed.

- Develop a topology map of chilean Internet could be useful to detect incorrect routing configuration.

- Conduct rutinary measurements would help detecting routes stability, connectivity shortages and so on.

http://www.nic.cl

Anycast deployment at .CL

- Previous to F-root
- From January to June 2004.
- Objective: Harden national DNS architecture
- First goal: Physical diversity: Santiago, Valparaíso and Concepción.
  - Valparaiso and Santiago are 150 km away. Santiago and Concepcion, 650km.
- Second goal: Topological diversity: Three different providers
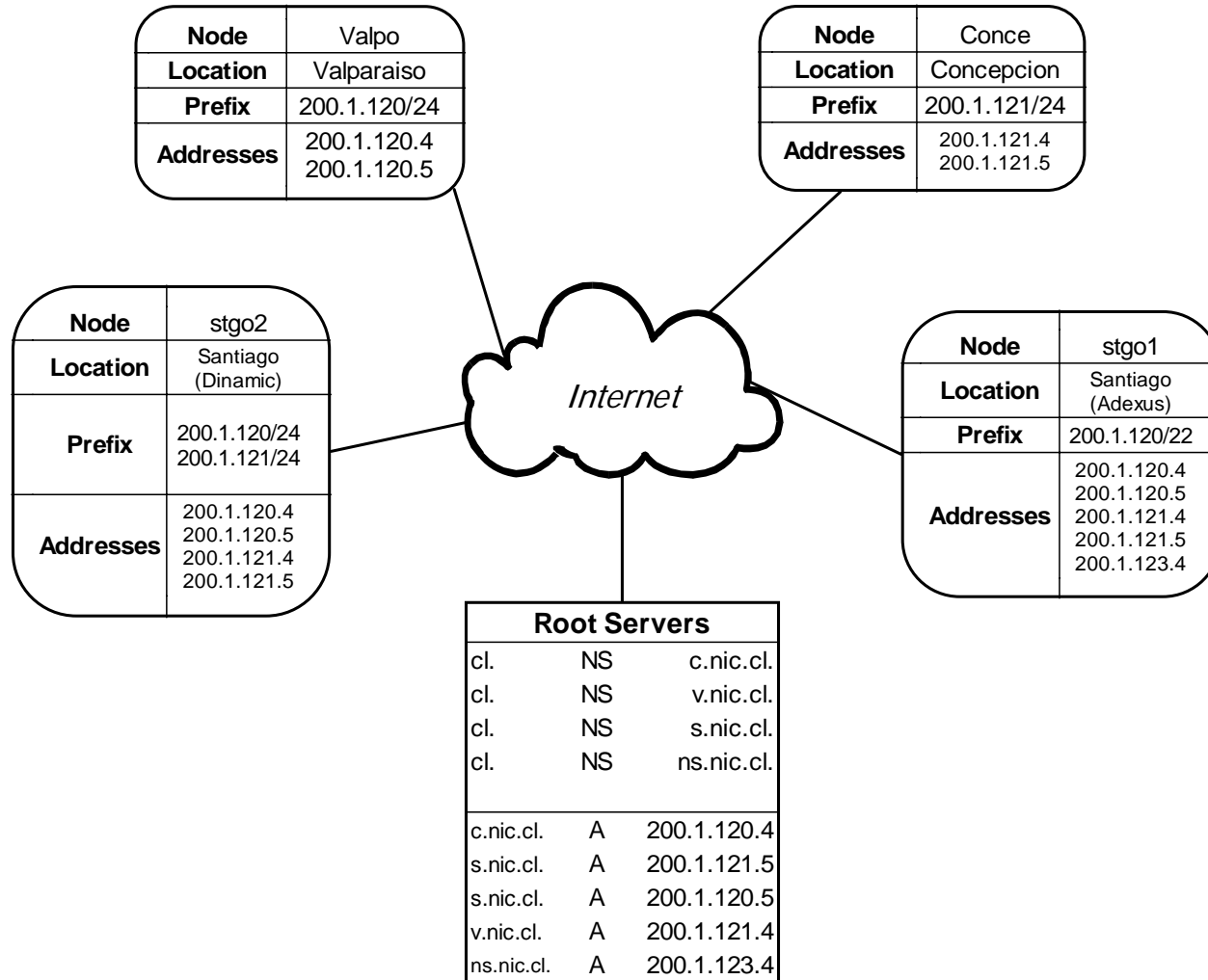- Third goal: Software diversity. FreeBSD and Linux, NSD and BIND9.

Challenges

- Getting understood: Explain to the providers what anycast is.

- Trouble on deployment: two physical locations share the same provider. Impossible to annouce the same prefix from those two locations.

- The "mesh" experiment
  - Three NS records ([**c**|**s**|**v**].nic.cl), four addresses (200.1.12[*0*|*1*].[4|5]), two prefixes (200.1.12[0|1]/24) each announced from three locations, one ASN (27678).
  - In some way forced by the situation described above.
  - Looking for model validation.
  - Expectation: doing this we keep more traffic in the country.

The "mesh" diagram

Node	Valpo
Location	Valparaiso
Prefix	200.1.120/24
Addresses	200.1.120.4 200.1.120.5

Node	Conce
Location	Concepcion
Prefix	200.1.121/24
Addresses	200.1.121.4 200.1.121.5

Node	stgo2
Location	Santiago (Dinamic)
Prefix	200.1.120/24 200.1.121/24
Addresses	200.1.120.4 200.1.120.5 200.1.121.4 200.1.121.5

Internet

Node	stgo1
Location	Santiago (Adexus)
Prefix	200.1.120/22
Addresses	200.1.120.4 200.1.120.5 200.1.121.4 200.1.121.5 200.1.123.4

Root Servers		
cl.	NS	c.nic.cl.
cl.	NS	v.nic.cl.
cl.	NS	s.nic.cl.
cl.	NS	ns.nic.cl.
c.nic.cl.	A	200.1.120.4
s.nic.cl.	A	200.1.121.5
s.nic.cl.	A	200.1.120.5
v.nic.cl.	A	200.1.121.4
ns.nic.cl.	A	200.1.123.4

The questions

- Time for switching: How long does it take to a source to be "forwarded" to other location, when a node is shut down.

- Which node gets the higher query rate.

- If one node fails, how the load is balanced among the others.

- Do all the NS receive the same load?

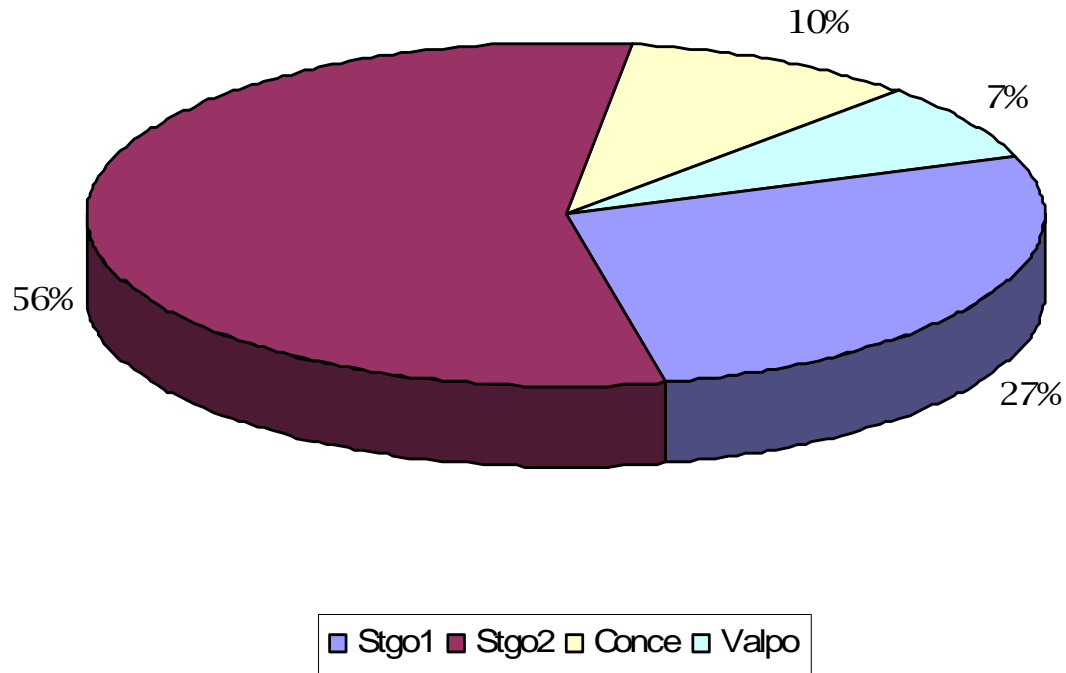- Do the foreigner NS receive more load?

http://www.nic.cl

The experiment

- Query capture using tcpdump, for 24 hours on every node.

- Analysis over origin addresses, origin AS, AS-Path length (using a BGP peer)

- Analysis over destination address among all nodes (to find out which one from NS list was selected) and on each node (to find out which one finally got the query).
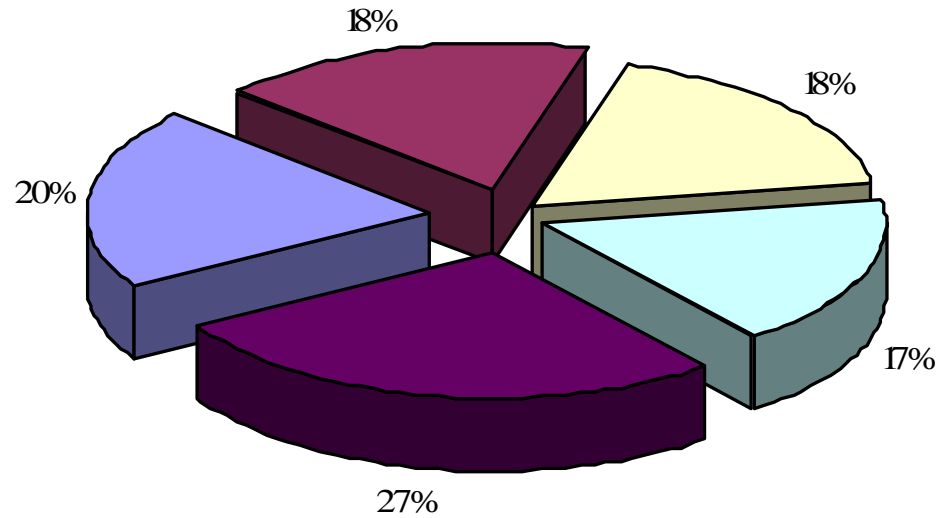
http://www.nic.cl

DNS-Ops Meeting

Preliminary results

Total queries by node

10%

7%

56%

27%

Stgo1 Stgo2 Conce Valpo

http://www.nic.cl

Preliminary results

Total queries by NS

Legend:
- c.nic.cl 200.1.120.4
- v.nic.cl 200.1.121.4
- s.nic.cl 200.1.120.5
- s.nic.cl 200.1.121.5
- ns.nic.cl 200.1.123.4

Percentages: 18%, 18%, 20%, 17%, 27%

Preliminary results

Top 5 ASN origin by node **Conce** (peering with ASN 6471)			
ASN	**Count**	**Description**	**AS-Path Length**
6471	1,056,236	ENTEL CHILE S.A.	2
22047	484,646	VTR BANDA ANCHA S.A.	4
14259	293,442	Gtd Internet S.A.	4
5089	164,980	NTL NTL Group Limited	2
17147	95,576	Gobierno de Chile	4

DNS-Ops Meeting

http://www.nic.cl

Preliminary results

Top 5 ASN origin by node **Valpo** (peering with ASN 6471)			
ASN	**Count**	**Description**	**AS-Path Length**
6471	633,726	ENTEL CHILE S.A.	2
22047	409,687	VTR BANDA ANCHA S.A.	4
5089	174,259	NTL NTL Group Limited	2
14259	159,502	Gtd Internet S.A.	4
15169	73,296	GOOGLE - Google Inc.	2

Preliminary results

Top 5 ASN origin by node			
Stgo1			
(peering with ASN 6471, 33075, 6505)			
ASN	**Count**	**Description**	**AS-Path Length**
6471	1,036,849	ENTEL CHILE S.A.	2
14259	774,604	Gtd Internet S.A.	4
22047	513,432	VTR BANDA ANCHA S.A.	4
8151	256,241	Uninet S.A. de C.V.	6
6429	191,662	Core Internet AT_T Chile	4

DNS-Ops Meeting

http://www.nic.cl

Preliminary results

Top 5 ASN origin by node **Stgo2** (peering with ASN 10778)			
ASN	**Count**	**Description**	**AS-Path Length**
15311	1,049,743	Telefonica Empresas	3
8151	1,048,240	Uninet S.A. de C.V.	5
7418	590,263	Terra Networks Chile S.A.	4
6429	519,878	Core Internet AT_T Chile	4
18747	515,276	IFX-NW - IFX Communication Ventures, Inc.	4

Preliminary results

- ## Partial Map of ASN
  - Having three out of four nodes peering with the same ASN is not a good idea
  - ASN 8151 (Mexico) strangely goes to Chile for DNS.
  - Two providers use two ASN for transit, growing AS-Path length artificially.
  - Due to volume, could be a good a idea install a node peering with ASN 19228
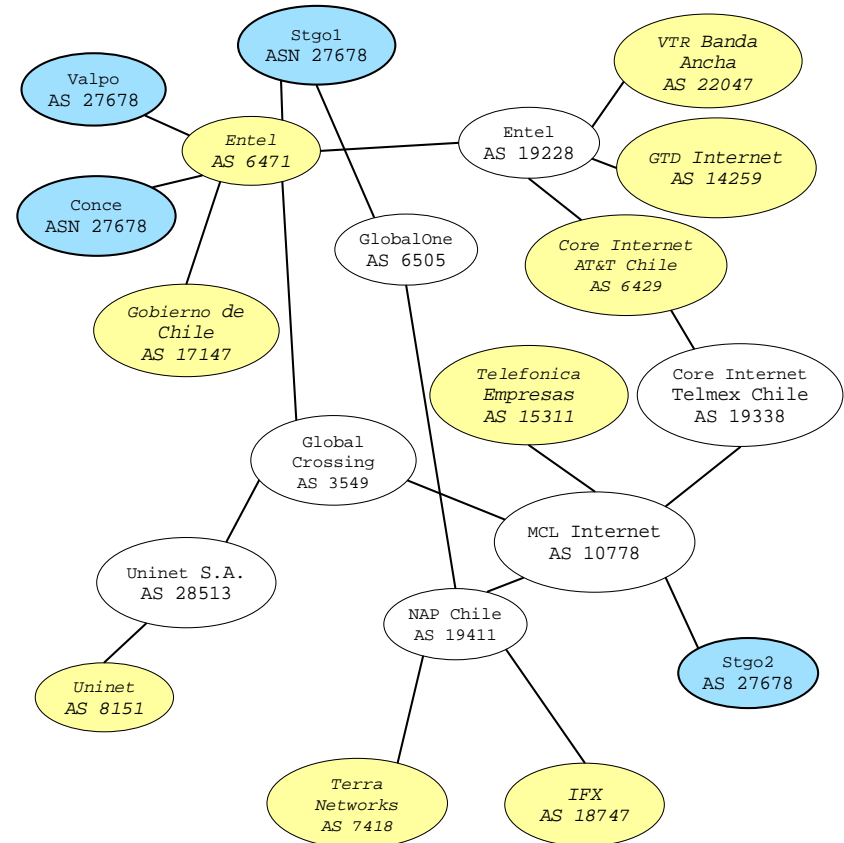
Work to be done

- Find a way to identify queries select by topology (routing process) and queries selected by lower RTT (Resolver selection).

- Shut down one or more nodes and check how the flow changes.

- Try to use a "normal" anycast setup (one NS, several nodes) and check the load.

DNS-Ops Meeting