
Asia Pacific Network Information Centre

APNIC

# DNS root server deployments

***George Michaelson***

***DNS operations SIG***
***APNIC17/APRICOT 2004***
***Feb 23-27 2004***
***KL, Malaysia***

# Why deploy an anycast node?

- Increase resistance against DDoS Attacks on the root
- Improve service quality
  - Speed of root-transit DNS queries
  - Resiliance to loss of connectivity

# Increase resistance to DDoS

- Prime goal for existing deployments
  - Target sources, locations with rich interconnect
    - Higher risk of DDoS traffic
- Achieve rich path diversity
    - No single path to flood
    - No single points of failure
    - Distribute DDoS load over net

# Improve service quality

- For us, in AP region
  - CN node has improved DNS RTT to root 15x
  - Invest in emerging/developing Internet nations
    - Encourage good routing practices
    - avoid expensive offshore transit to critical infrastructure
  - Improve knowledge/cooperation
  - HKIX node serves very diverse paths in AP region
- Protect countries against loss of external connectivity
    - Eg undersea cable failure

# APNIC's Role in root services

- Facilitate improved root services in AP region
  - Leverage APNIC member (ISP) resources
- Provide coordination point in AP region
  - Coordinate with root operators (F,K,I,M)
  - Host discussions during APNIC meetings
- Fund and/or coordinate sponsorship
  - Hardware, hosting, maintenance costs
  - According to individual circumstances
- Undertake formal agreements
  - MoUs with root operators (F and I so far)
  - MoUs with hosts
  - Long standing relationship with RIPE NCC
- Currently no "root operator" responsibility

# Timeline of APNIC deployments

- Nov 2002
  - APNIC announces MoU with ISC to deploy root nameservers in AP region, calls for EOI
- Jan 2003
  - Node deployed at HKIX, hosted by CUHK
- Sep 2003
  - Node deployed in Seoul, hosted by KRNIC
- Oct 2003
  - Signed MoU with Autonomica (I-Root)
- Nov 2003
  - Node deployed in Beijing, hosted by CNNIC, China Telecom and China Netcom Corporation (CNC)
- Dec 2003
  - Node deployed in Taipei Hosted by HANET
  - Node deployed in Singapore Hosted by NUS/SOX
- Jan 2004
  - Node deployed in Brisbane Hosted by PIPE networks

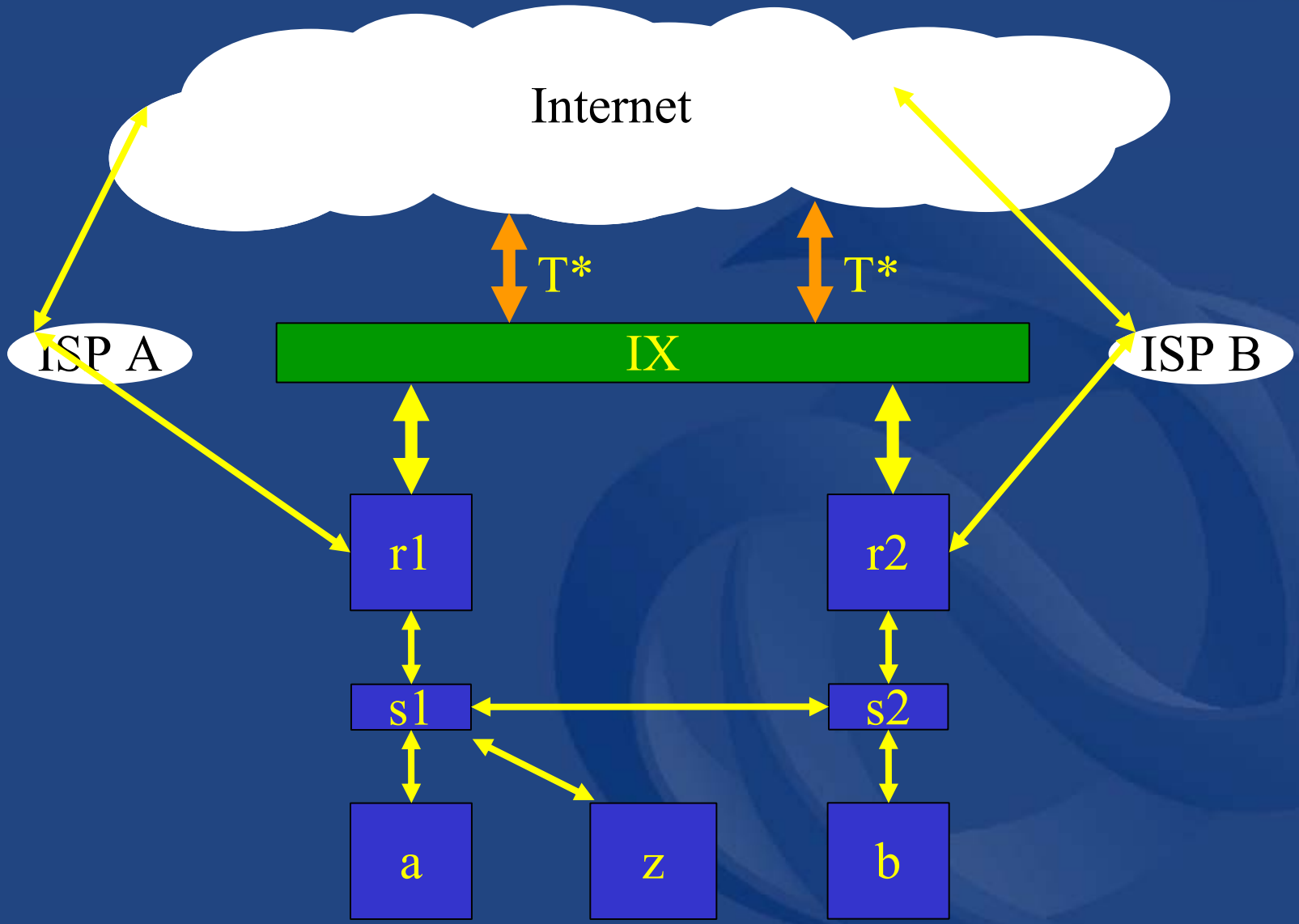
# Plans for 2004

- Further deployments planned
  - with F (ISC) ,I (autonomica) & K (RIPE NCC)
    - Balancing goals, locations, size issues
    - K-Root interested in AP region 'global' node
  - I-Root to be deployed at HKIX, Mar 2004
- EOI/CFP to be re-issued. Goals:
  - Regional development
  - Improve resiliency at existing PoP
- Continued coordination in region
  - Report to APNIC dns ops sig regularly

# Where is my root coming from?

- For F:
  - Find current nodes from http://f.root-servers.org
  - `dig @f.root-servers.net. HOSTNAME.BIND chaos txt`
  - Should show sensible path to local node
    - Eg In NZ should show path to F-root in Auckland
    - Eg In KR should show path to F-root in Seoul
- For any root:
  - `traceroute i.root-servers.net`
- APNIC encourages participation in BGP peering with critical infrastructure

# F-Root Node Overview



# F-Root Node Hardware

**Routers**

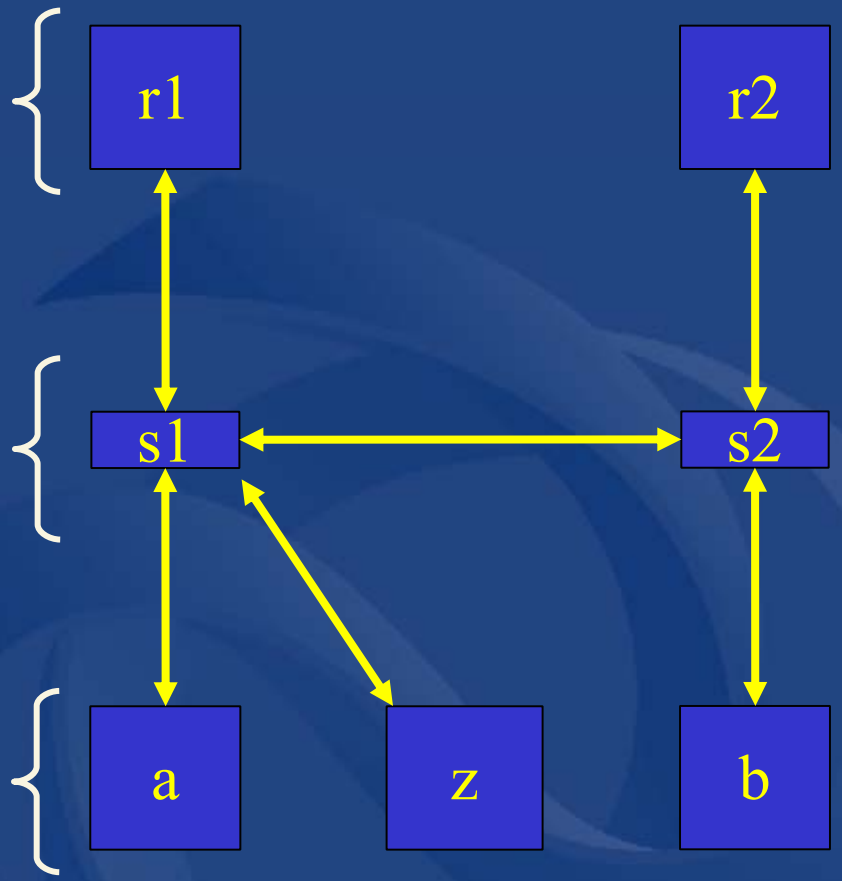
- Cisco 7xxx or Juniper M5
  - High filtered packet rate
  - Multi Gbps throughput
  - At least two external ports

**Switches**

- Cisco 29xx/35xx
  - 100mbit FD
  - gig-E interconnect to routers

**Hosts**

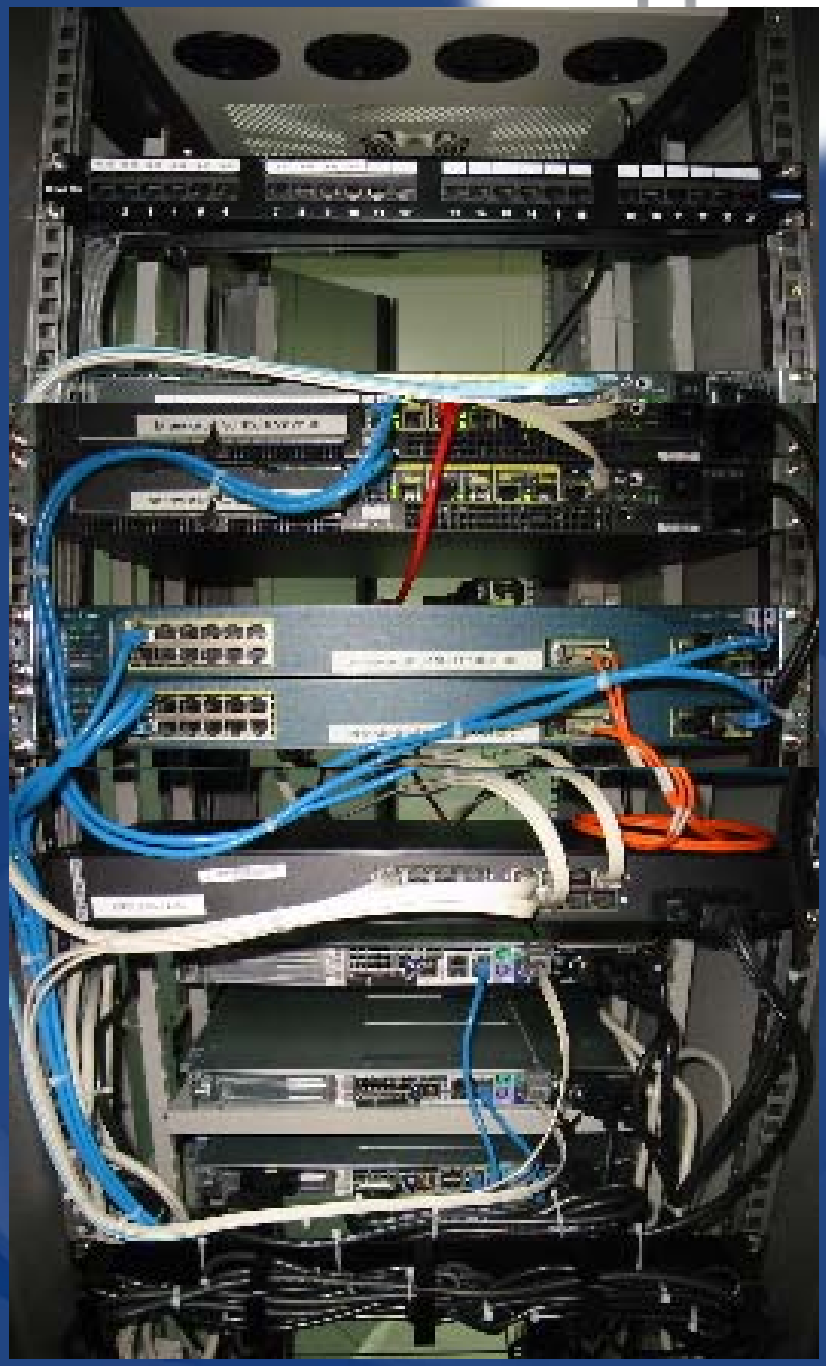
- Dell 1750
  - 2.4Ghz/1Gb/40Gb
  - Dual 10/100/gig-E NIC
  - All FreeBSD 4/5-Stable
  - z (monitor) node has RAID
  - a & b run Bind 9



# Node Behaviour

- Two independent Routing paths, one switch fabric on two switches
  - Management (z) host has consoles for all devices, modem
- Each Router connected to IXP directly
  - Also has independent off-exchange transit
- No cross-connect from the DNS hosts (a,b) or z host
  - Additional benefit marginal, set against added complexity
- Node is capable of handling local overcommitted load
  - Both network and CPU/memory bandwidth will scale to meet future trends
  - Local Nodes do not take failover service from each other, failure mode is to global nodes only at this time
- Nodes can fail: but there are many of them

Front of Node
Routers, Switches,
(serial console), hosts



Back of Node
Local media conversion
(ZX to SX), hosts

# Routing Architecture

- One AS for service, route announced to IX participants
  - Each node announces 192.5.5.0/24 with a node-specific peer AS but with a consistent Origin AS across all anycast nodes of 3557
- Additional Management AS
  - A node specific route is announced to two or more transit providers which allows the node to be managed remotely
    - Management paths avoid IX fabric, DDoS risk if IX flooded.
- Two connects to IX fabric, Two management paths
  - Can engineer routing changes, service changes with no loss of service from site as a whole

# Routing Architecture

- Prefix announced with 'no export' community
- Propagates firstly to IX participants
  - One or more may provide limited transit
  - Direct customer routes possible (see APNIC)
- Limits horizon of visibility
  - Wider visibility can be organized (see APNIC)
  - During failure (eg DDoS) load shed is always to Global node(s)
    - So far, all ISC Deployments in AP region local nodes only
  - Visibility is not necessarily limited to one country
- Anycast prefix should not be announced for transit

# ISC information pages

- Current list of anycast nodes
  - http://f.root-servers.org
- Hierarchical Anycast Architecture
  - http://www.isc.org/tn/isc-tn-2003-1.html
- Peering with ISC
  - http://www.isc.org/peering

# Questions?

***Thank you***

***ggm@apnic.net***