“Anycasting” f.root-servers.net

Suzanne Woolf

January 2003

Threat or Menace?

- “Anycasting a root name server” sounds radical. It’s not.
  - Natural extension of simple redundancy techniques
  - Technology is familiar to the community
  - Response to serious but familiar threats: DDoS and network degradation

Necessary Distinctions

- Local load balancing
  - Sometimes called clustering
  - Maybe using an appliance
- Distributed load balancing
  - Might just be a diverse set of NS RRs
  - Or it might be that a single NS RR is global
- Policy based (directed) load balancing
  - Different answers in different regions
  - We call this “Stupid DNS tricks” (don’t do it!)

Local Load Balancing (1)

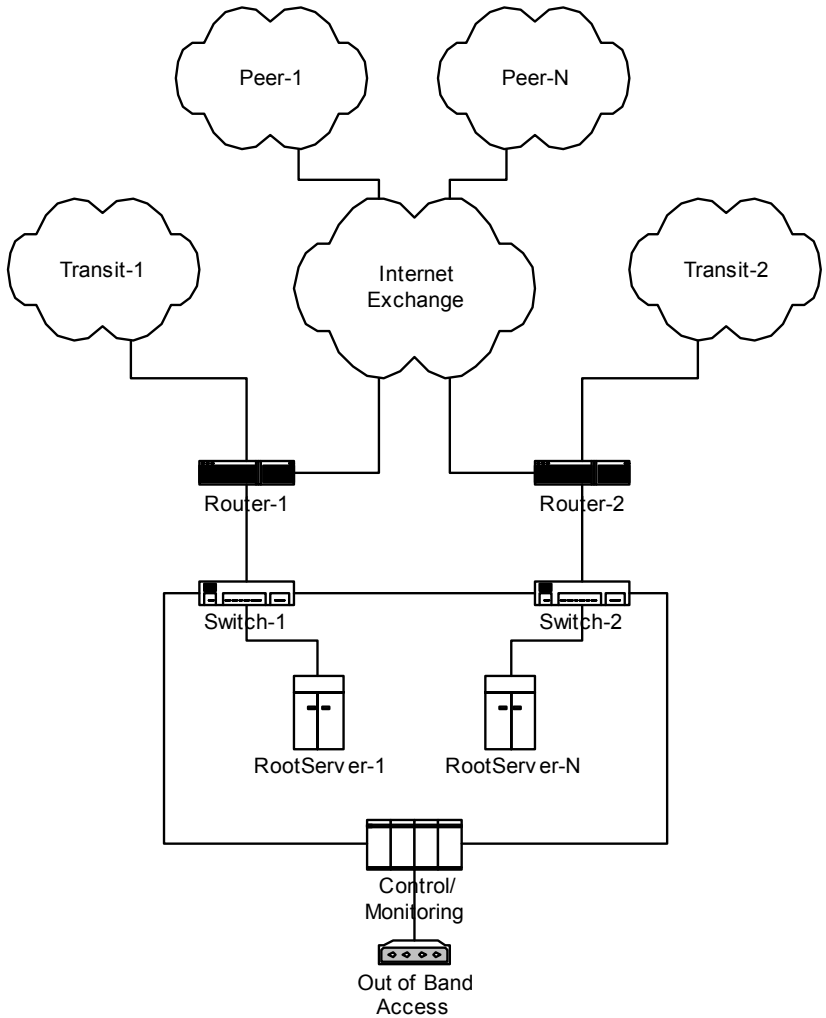
- An L4 switch with health monitoring can distribute query load across a cluster
- This “extra powered box” is a failure point
- Sometimes requires that all TCP land on a single host instance (with fallback)
- Sometimes requires that a single MAC address be used by all cluster members
- This is really the wrong approach

Local Load Balancing (2)

- Using routers and switches that you probably already have in the data path…
- Use GateD/Zebra for host-based OSPF
- Assign a single service address as an “lo0” alias on all members of the cluster
- OSPF “stub host” logic advertises it
- Modern Cisco (CEF) and Juniper (IP-II) routers will do flow hashed load sharing

F-Root

1/19/2003



Distributed Load Balancing (1)

- Core internet routing protocol is BGP, which is loosely distance-vector based
- When multiple paths exist, one is chosen, usually based on AS-path length
- This is not useful for actual load balancing:
  - Geography != Topology
  - Too coarse-grained
  - Depends on other ISPs' policies

Distributed Load Balancing (2)

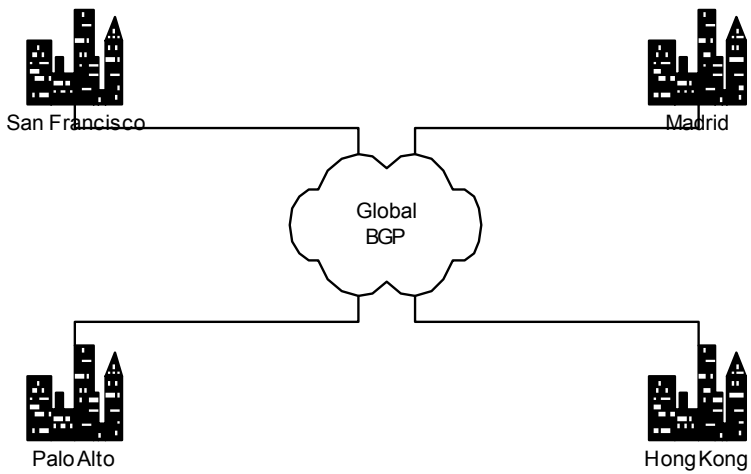
- BGP routes can be tagged “no-export” to ensure that there is no “accidental transit”
  - F-root only has transit at PAIX Palo Alto
- Thereby one can collect traffic from a deliberately restricted part of the topology
- For example, all peers at an exchange point
- This is especially useful for partitioning DDoS attacks and keeping them “local”

Distributed Load Balancing (3)

- Each wide area F-root has its own AS number and its own “management” /24
- The management /24 gets transit from multiple ISPs over private crossover-ether
- The “F” /24 is advertised through the public exchange point, tagged with “no-export”
- Attacks are localized, and do not interfere with network management
- Exchange point fabric means no “DDoS bottleneck”

F-cities

1/19/2003



Other Advantages

- Fluidity: add or drop servers or cities at will
  - To upgrade a host or city, drop then add
  - Failures are local and meaningless
  - Add capacity or shift load during attacks
  - Headroom, headroom, headroom!
- Measurement: triangulate on DDoS sources
  - Source routing and source spoofing don’t mix

Other Root Servers

- Others are also looking at wide-area distributed load-balancing, but
- Diversity is a major strength of the root nameserver system, so
- Each server operator has their own strategy

Others (2)

- K.root-servers.net (RIPE-NCC) is working on a detailed plan with RIPE members
- C.root-servers.net (Cogent) has instances located in several Cogent datacenters
- J.root-servers.net (Verisign) has instances colocated with their gTLD servers
- Some others are considering costs and architectures

FAQ

- Do you retain administrative control of the F-root instances? *YES*
- Have you provisioned OOB monitoring and troubleshooting? *YES*
- Why not let ISPs do it?

  ISPs run networks. We run root name servers, including responsibility for integrity of the data.

Any Questions?

….Or contact us at ISC:

Paul Vixie
paul@vix.com