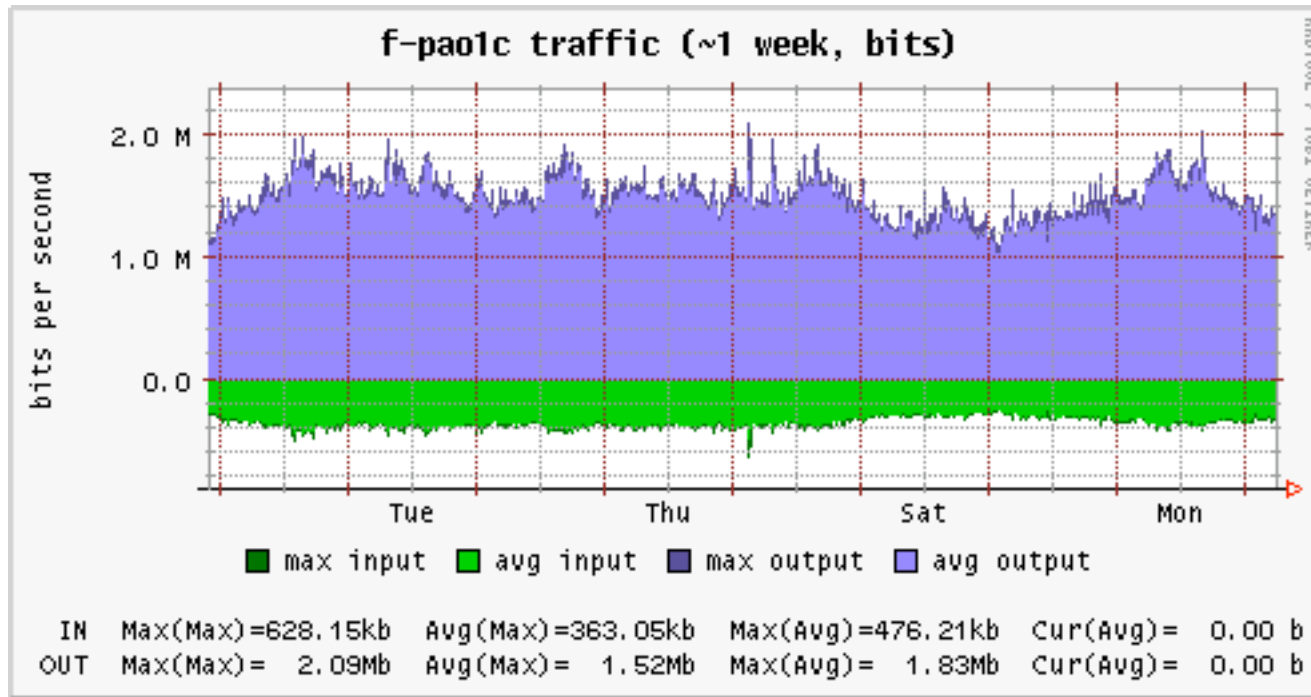INTERNET SOFTWARE CONSORTIUM

f.root-servers.net

Joe Abley
jabley@isc.org

Root Nameservers

- Serve the root zone

- Provide delegations for TLDs

- 13 root nameservers throughout the world, named A through M

- F is operated by the ISC

Query Load

- For the F-root complex as a whole:

  - 2 Mbit/s of queries in, on average

  - about 6 Mbit/s out

Local Load Sharing

- A layer-4 switch with health monitoring can distribute query load across a cluster

- This "extra powered box" is a failure point

- Sometimes requires that all TCP lands 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 Sharing

- Using routers and switches that are already present in the data path

- Use gated or Zebra on hosts

- Assign service address as a "lo0" alias

- OSPF stub logic advertises it

- Modern cisco (CEF) and Juniper (IP-II) routers will do flow hashed load sharing

Cluster Architecture
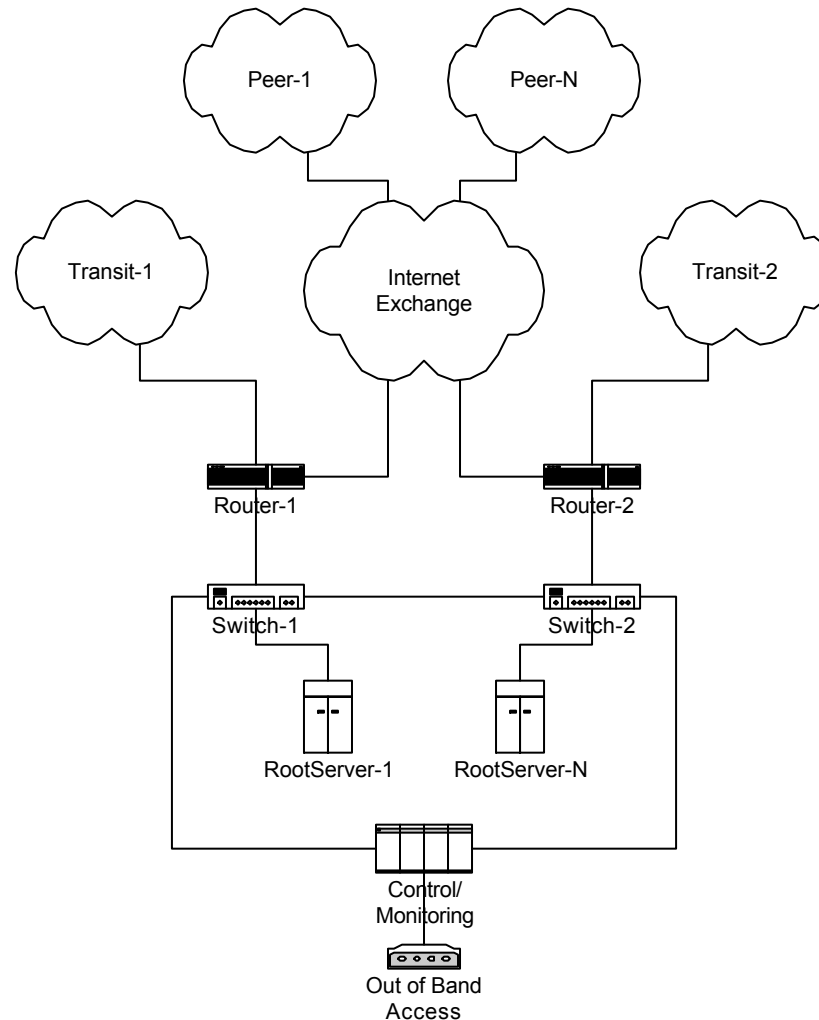
Distributed Load Sharing

- Core Internet routing protocol is BGP, which is loosely distance-vector based

- Where multiple paths are present, exactly one is chosen

- Not useful for "load balancing"

  - geography and topology are different

  - depends on ISP policies

Distributed Load Sharing

- Restricted radius of propagation of 192.5.5.0/24

    - no-export, other local arrangements

- Each cluster can collect queries from a well-known and predictable catchment

- e.g. all peers at an exchange point

- Especially useful for partitioning a DoS attack, and keeping it "local"

Global F-Root Deployment

- We are deploying many F-Root clusters around the world

- Each cluster provides a termination point for traffic directed at 192.5.5.0/24

  - f.root-servers.net is 192.5.5.241

  - restricted IPv4 "anycast", similar to that used in the AS 112 project (and others)

Global F-Root Deployment

- Each F-Root Cluster has its own AS number, and it's own management /24

- Each cluster gets transit for its management /24 from two or more local ISPs, over dedicated ethernet cross-connects

- The F /24 is advertised to peers over an independent fabric (e.g. an exchange point)

- Attacks are localised to the F peering fabric, and do not impact management traffic

F-Root Cities

F-cities
12/19/2002

San Francisco — Global BGP — Madrid

Palo Alto — Global BGP — Hong Kong
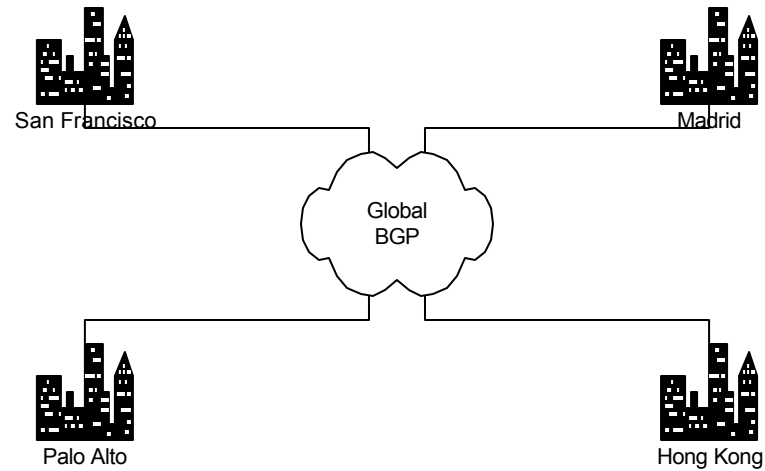
Advantages

- Flexibility: add or drop servers at will

  - to upgrade a host or city, drop then add

  - failures are local and meaningless

  - add capacity or shift load during attacks

- Measurement: triangulate on DDoS sources

  - source routing and source spoofing don't mix

New F-Root Cities

- Adding 7-10 new cities in 2003

- Aim to add a further 20 cities by the end of 2004

- Each city facilitated by a local sponsor

  - equipment, deployment and operational assistance, liason with peering and transit partners