Logistics

- Webinar is 1 hour long
- A recording will available by May 12
  - http://www.isc.org/webinars
- Participants are muted
- Questions should be entered into the WebEx Q&A tab for the presenter
  - The presenter may defer some questions until the end of the presentation
Agenda

• Define Anycast
• Examine use cases
• Explore the impact on Internet protocols
• Explore Anycast and DNS
• Share ISC’s operational experience
• Answer questions
Define

ANYCAST
What is Anycast?

- Anycast describes a method of using the same IP address on multiple servers
- Fundamentally, Anycast is a *routing scheme*
- Anycast is more about the configuration of routers and routing than servers
  - Server admins have to understand what’s going on in order to properly operate the service
Routing Schemes Compared

Baseline

Unicast

Broadcast

Multicast

Anycast

Diagrams from http://en.wikipedia.org/wiki/Anycast, and are public domain.
Properties of Anycast

• Each packet sent to an Anycasted IP address may reach a different server

• Packets are routed to the IP address with the best *network metric*
  – This is often the nearest server, but not always. Metrics could be set based on other factors, such as bandwidth, cost, load or reliability

• Servers with an Anycast address must also have a Unicast IP address
  – Management functions can’t be done to the Anycast address as they would only reach one server!
Examine

USE CASES
Use Cases

• Local Anycast
  - Distributes load across multiple servers on same subnet
  - Eliminates need for load balancer by making the network (router) distribute traffic

ONE ROUTE!
Reduces routing issues

A ➔ 1
A ➔ 2
A ➔ 3
A ➔ 4

Routes may originate via any supported protocol
- static/RIP/OSPF/ISIS/EIGRP/BGP
- dynamic routing handles most failure cases
- active service probing from the router is an option
Global Anycast

- Distributes load across multiple locations
- Provides redundancy
Use Cases

• Most popular things to Anycast:
  1. DNS, recursive servers
     • Configured by IP address on clients
     • Latency is important
     • Distribute load across multiple devices
  2. DNS, authoritative
     • Limited number of authority IP’s can be listed in a single reply packet
     • Latency to the server is important
     • Redundancy a large concern
     • Distribute load across multiple devices
  3. NTP
     • Generally only in ISP’s that have a large amount of CPE that requires
       configuring NTP by IP address and not name, or enough clients that load
       distribution is required.
  4. HTTP Redirect Servers
     • HTTP servers that redirect a user to another local instance.
Explore

IMPACT ON PROTOCOLS
Impact on Protocols: ICMP

- Global, stateless options work fine
  - Ping request/reply
  - ICMP Traceroute
    - Network instability can produce some odd results with traceroute

- Avoid LAN options
  - Router Advertisement/Solicitation
  - Address Mask Request/Reply
  - Redirect
  - A unicast address on the server can mitigate these issues
  - It’s easy to avoid all of these ICMP options
Impact on Protocols: ICMP

- Transmission failure messages are a problem
  - Destination {network,host,protocol,port} {unreachable,unknown}
  - Fragmentation required
  - Source route failed
  - Source host isolated
  - Network administratively prohibited
  - {Network,Host} unreachable for TOS
  - Communications administratively prohibited
  - Source quench
  - Time exceeded
Impact on Protocols: ICMP

ICMP Source Quench

ICMP packet discarded here; other server never receives notification
Impact on Protocols: ICMP

• Operationally, what really matters?
• Losing “packet too big” breaks PMTU
  – Packets from an Anycast host should *never* be sent with the DF bit set
    • Options are to accept packets being fragmented mid-stream, or to send with the minimum MTU
  – *IPv6 does not allow for intermediate routers to fragment, all packets must be sent with the minimum MTU of 1280*

• Lost messages prevent orderly teardown
  – Timeouts for end users, may be long waits!
  – Resources consumed on the servers waiting to tear down connections
Impact on Protocols: UDP

- Stateless, which is good for Anycast
- Works well when the query is one packet, and the response is 1-n packets, and there is no state between queries
  - Sounds like the majority of DNS queries!
- If the query is more than one packet, or there is state between queries, the behavior tends to be the same as TCP
Impact on Protocols: TCP

• Only works when the network path is stable.
  – This is *never true in the long term*, but is often true for short periods of time

• The Unicast sender has to reach the same Anycast destination for the duration of the connection
  – One packet to the wrong device causes it to generate a TCP Reset, which generally tears down the connection
Impact on Protocols: TCP

TCP SYN
TCP SYN/ACK
TCP ACK/Data
TCP Reset
Path Instability: Sources

1. Load Balancing
   - Per-packet load balancing directs each packet to a different link and possibly server
   - Per-flow load balancing typically hashes on a 5-tuple, which creates a stable path for many topologies, but there are topologies where even this sort of hash won’t be stable

2. Route Churn
   - \{Link, Router, Server\} failures
   - User configuration; sessions added/removed, metrics changed

3. Middle Boxes
   - “Route optimizers” and load balancers do all sorts of interesting things to packet flows!
Impact on Protocols: TCP

• Operationally, what does it mean?
  – The location of the Anycast servers is important, and depends on the network topology and configuration
  – When properly deployed, there is a high success rate for short duration connections
  – The longer the connection, the greater the risk of failure

• For Internet services it’s not just your network, but **every network the packet traverses** to the Anycast server!

• Avoid Anycasting TCP services when there are good alternatives
Explore DNS & ANYCAST
DNS & Anycast

• Most common queries are a single UDP packet, with 1-3 UDP packets of response
• TCP queries are extremely short lived
  – User->Server: SYN, ACK w/query, ACK/FIN
  – Server->User: SYN/ACK, ACK w/Data, ACK/FIN
    • Maybe an additional data packet
  – The FIN can be lost in some implementations and the data still be received

• Zone transfers are long lived TCP queries
  – Length depends on zone size
  – Some zones don’t allow, mitigating the issue
End User Resolvers

Regional Hub

Backbone

Regional Hub

Pop #1
Users

Pop #2
Users

Pop #3
Users

Pop #4
Users

User queries
Stay local

Failure reroutes
No user outage
Anycast & DNS

- Authority servers across an ISP/Enterprise provide redundancy, load distribution and hitless maintenance

Queries stay local
Pop Failure, service still up
Anycast & DNS

- Authority servers across multiple networks

- ISP redundancy
- Lower latency, Keep traffic local

Queries stay local
Anycast & DNS: Advanced

• Inconsistent content
  – Part of the secret sauce in some CDN’s
    • Each Anycast server is loaded with a slightly different data set, and returns answers that direct users to specific servers or to names or IP’s that provide some information about the name server the user queried
    • Keep in mind the user generally queries a resolver, so the Anycast Authority server hit was the **one closest to the resolver**, not the end user
    • That may be good enough

• Routing mechanisms can be used to direct traffic in interesting ways
  – Using multiple super/subnets
  – Metrics that alter dynamically
  – Cisco’s “IP SLA” to add/remove routes
ISC’S OPERATIONAL EXPERIENCE
SNS@ISC

• ISC’s authoritative hosting product
• Present on 3 different ISP networks
  – Cogent, Hurricane Electric, Tata Communications
• Anycast *inside* of each ISP
  – IP address space is used from each ISP inside their own network
  – A minimum of 3 locations on each ISP’s network
• By including 3 NS records in a zone the zone is available across 9 locations worldwide on 3 different ISP networks!
F-Root

• Three levels of Anycast
  – Local LAN
    • Each deployment has a minimum of 2 servers on the local network for redundancy, more where necessary
  – Local Nodes
    • A typical F-Root deployment at a exchange point or inside of an ISP network
    • Announces 192.5.5.0/24 and 2001:500:2f::/48 with NO_EXPORT set
      – Because of the NO_EXPORT settings these routes will not be visible to all end users
  – Global Nodes
    • Larger nodes, with significant transit capacity
    • Announce 192.5.4.0/23 and 2001:500:2e::/47, supernets of the local node prefixes
    • These networks should be visible to all end users on the Internet
F-Root

192.5.4.0/23  2001:500:2E::/47

Global Node

Local Node

IX

ASN

NO_EXPORT
192.5.5.0/24
2001:500:2F::/48

192.5.4.0/23
192.5.5.0/24
2001:500:2E::/47
2001:500:2F::/48

Customer

Peers w/ F

Does not peer

Customer

Customer

Customer

Customer

Customer

NO_EXPORT
192.5.5.0/24
2001:500:2F::/48

192.5.4.0/23
192.5.5.0/24
2001:500:2E::/47
2001:500:2F::/48
F-Root

• Why 3 levels?
  – A strong desire to keep local traffic local
    • Local nodes may be deployed in bandwidth starved areas, like behind satellite links, and thus shouldn’t draw in queries from far away
    • Provide an incentive for local ISP’s to peer with the local F-Root instance
  – Diversity in the Root Server ecosystem
    • Root operators believe that having different parties deploy in different models allows for more effective service of different user communities, and provides a more difficult attack surface
    • No one else uses this method!

• This does create some confusion
  – ISP’s think that because the local route has NO_EXPORT their customers won’t see F-Root, but this isn’t true due to the covering supernet
F-Root

• Zone transfers are not officially supported, but allowed
  – If the long lived TCP connections fail ISC does not consider it an outage

• Prior to IPv6 and DNSSEC deployment TCP queries were extremely rare
  – 0.00%, before DNSSEC
  – 0.2-0.4% after DNSSEC
  – Most DNS implementations handle a non-responsive server in an intelligent fashion by using other servers

• It may not be wise to have 100% of the authority servers for a domain Anycasted
F-Root
Summarize

ANYCAST
Summary

• Anycast is a routing scheme that can be useful when deploying some applications
• There are some protocol level implications that must be considered when designing an Anycast deployment
• DNS is generally well suited to Anycast deployments, and is one of the most popular services to Anycast
• Lots of other folks are doing it, don’t be afraid!
Learn

ISC EVENTS
Events and Trainings

www.isc.org/webinars

- Despliegue y Experiencia Operativa con Anycast
  - 15 May 2012
- Cyber Crime Remediation
  - 22 May 2012
- IPv6 Lessons Learned
  - 12 June 2012

www.isc.org/support/training

- 3-Day IPv6 Fundamentals
  - 4-6 June 2012, Amsterdam
- 2-Day DHCP Workshop
  - 7-8 June 2012, Amsterdam
- 2-Day Intro DNS & BIND
  - 18-19 June 2012, Virginia
- 5-Day Adv DNS & BIND
  - 18-22 June 2012, Virginia
- 2-Day Intro DNS & BIND
  - 2-3 July 2012, Amsterdam
- 5-Day Adv DNS & BIND
  - 2-6 July 2012, Amsterdam
SPECIAL OFFER

18% discount on any training sessions, now to 30 September 2012!

A thank you for attending!

Coupon code in the follow up e-mail you will receive from this webinar.

www.isc.org/support/training
Ask

QUESTION AND ANSWER
Keep in Touch

www.facebook.com/InternetSystemsConsortium

www.linkedin.com/company/internet-systems-consortium

www.twitter.com/ISCdotORG