Level (3) COMMUNICATIONS

Connecting and Protecting the Networked World℠

A network operator perspective on PA multi-homing

Eric Osborne

5/31/2016

Background

In Buenos Aires I threw some context-free, impromptu stones at the mic

This presentation expands on that, now that I've had some time to digest the issue

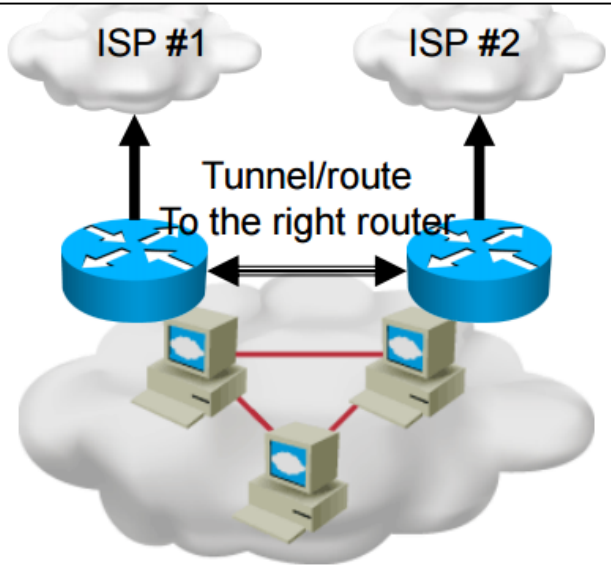
There's a lot of history here I haven't fully absorbed, and I may be asking dead horse questions.

Still going to ask them.

The trigger

(a clip of Fred Baker's slide from Buenos Aires)

- RFC 3704
- Concept:
  - Destination route within a network
  - At the egress, wonder what source prefix is in use
    - If the correct one for upstream, send upstream
    - Else, re-route to the correct egress router
- My question:
  - Why not route it to the right router in the first place?



Some solutions

There appear to be ~~four~~ **five** general approaches to this problem:

1) Don’t Do That

Provider control plane

2) Update BCP38 filters to allow PA space from those who don’t own it

Downstream control plane

3) Push src/dest routing to the origin (host)

4) Push src/dest routing to the ingress router

5) Tunnel between egress routers

Don’t Do That

It’s PA space for a reason.

Sending PA sources through other providers makes it PI space.

Table bloat, SWIP/IRR confusion, etc. Don’t Do That.

This is my favorite, but completely unworkable in practice.

Update BCP38 filters

Defeats the whole purpose of BCP38 filters.

And of PA space.

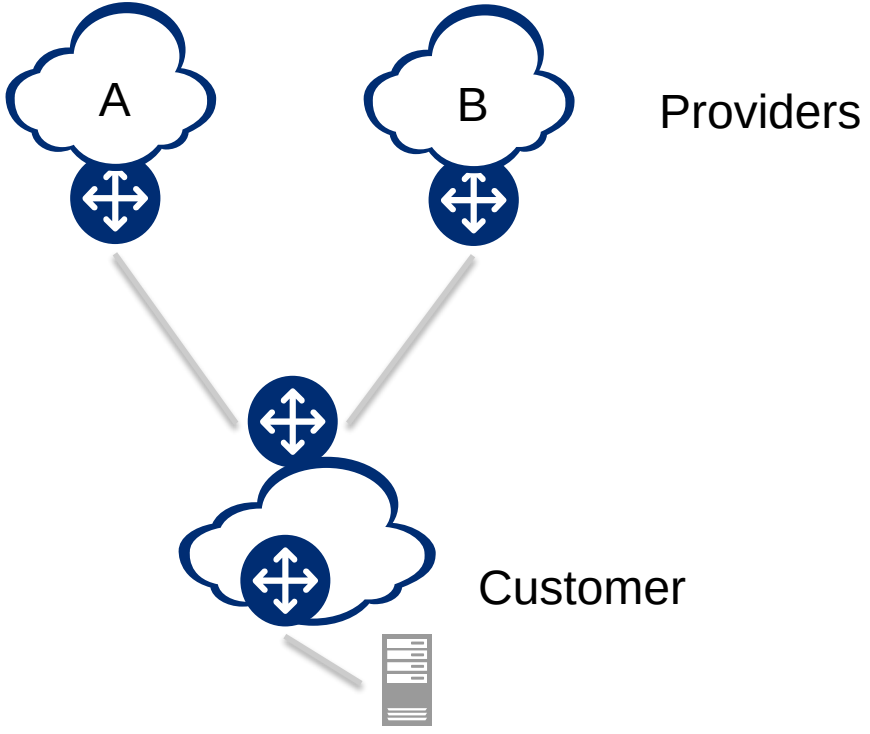
Hard to get right.

Also the most practical.

Simple-ish: Multihomed stub AS

End customer is an enterprise.

Host, ingress router, egress router can all handle src/dst routing.



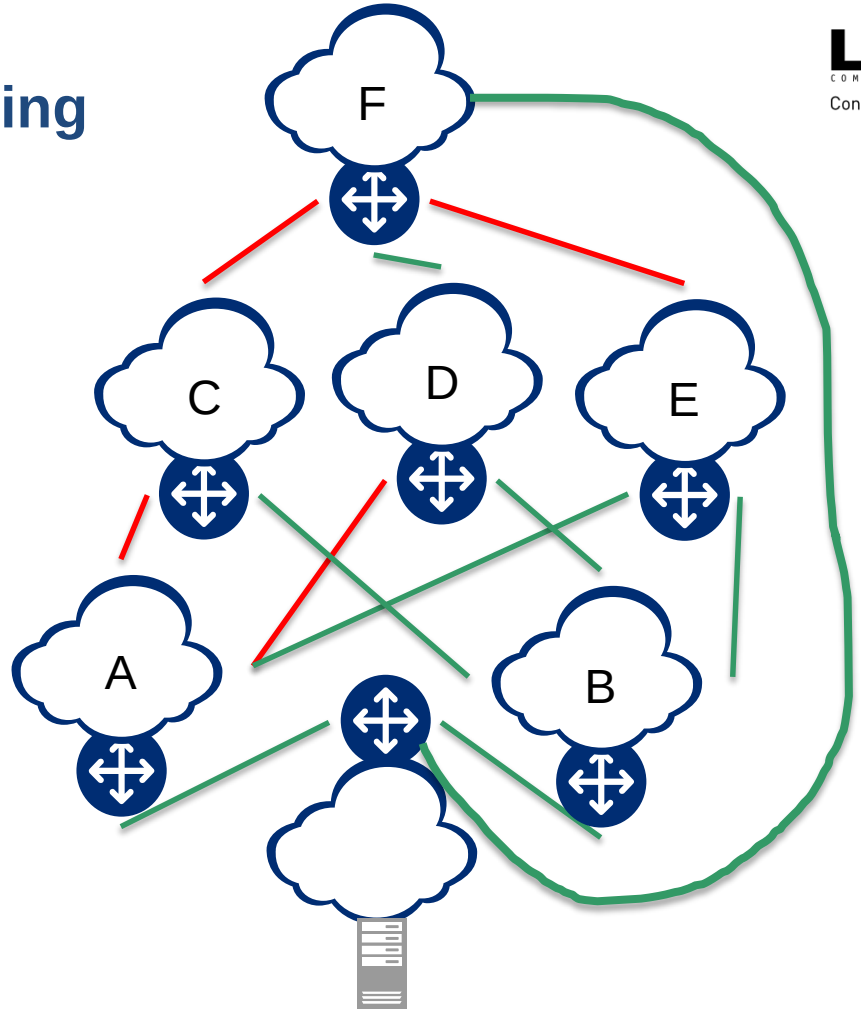
Harder: N-hop multihoming

Level 3 COMMUNICATIONS
Connecting and Protecting the Networked World℠

Customer has PA space from F, multihomes to F,A,B

Do [C,D,E] apply BCP38 to [A, B]?
Does F apply BCP38 to [C,D,E]?
What if Cust<->F goes down?

What is ‘PA space’ any more?



Push src/dest routing to the origin (host)

This is a terrible idea.

Requires hosts to have at least some kind of primitive routing.

Requires first-hop network to trust the host's routing (or verify+correct, in which case why have the host do it?)

"Internet routing to hosts" sounds like "ATM to the desktop"

Push src/dest routing to the ingress router

Level 3 COMMUNICATIONS
Connecting and Protecting the Networked World℠

If you're a transit AS, this works out to about the same as egress tunneling, since every ingress node is also an egress node.

If you're not, this may have a significant impact on the network. Ingress devices might be only ingress devices. Adding source/dest routing might be an uplift.

::0/0 via all exits is easy.

Specifics via some exits is harder - do you need v6 PA NAT?

Tunnel between egress routers

Second most practical.

Only involves coordination between egress routers.

- For an end AS, this may not be many routers.
- For a transit AS this could be thousands.

May involve new hardware (OpenFlow, anybody?) but only at exit points

Same question about specifics.