

# BGP path 'hinting' proposal

Brent Sweeny  
Global Research NOC at Indiana University  
*sweeny@grnoc.iu.edu*

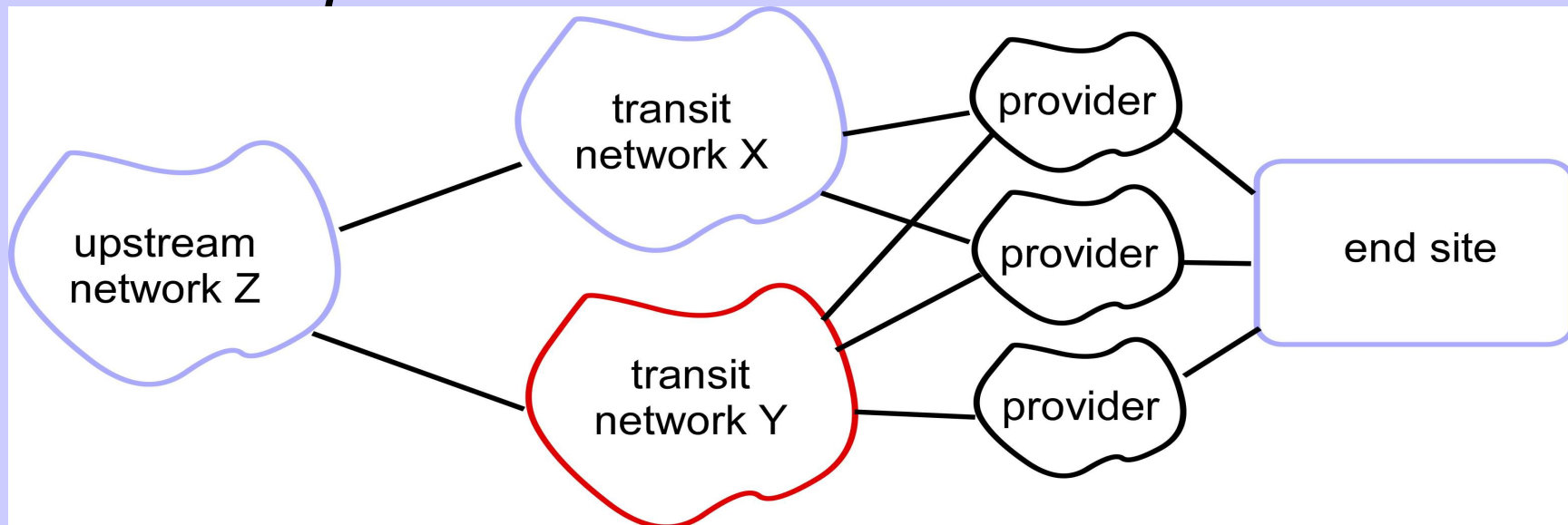
*to:*  
Internet2 Member Meeting (San Diego), 9 October 2007

# Topics

- Purpose (what's this all about?)
- Why? (why might this be a *good* idea?)
- Why not? (why might this be a *bad* idea?)
- How? (how might it be done?)
- How? (possible variations)
- What's necessary to make it work?
- Possible hint values (sample table)
- What next?

# Purpose

- Allow end sites to 'hint' or suggest to intermediate networks a path the end site would *prefer* traffic take toward them



- Use well-known BGP community values
- Its use is optional to intermediate networks

# Why?

- Unequal paths toward user, user wants to direct
- Intermediate networks may use (default) criteria for path selection different than end-site's but may take requests into consideration
- Existing alternative methods (MEDs, AS-prepend, single-network local-pref hinting) aren't sufficient
- R&E network community approach may work (as with jumbo-MTU BCP)
- Common, documented approach easier to debug through net than current potpourri
- 'regularized' approach could allow for selection to be programmed for general cases, not one-off exceptions to your normal routing policies

# Why not?

## What problems could it cause?

- End-sites influencing path selection
- Wrong people making the decision about traffic flow
- Could it cause loops?
- How do you know the actual end-site really requested this? (the “rogue transit network”: could it be a DoS?)
- Do the end-sites know something about topology and policy (esp for intermediate networks) that the networks directly involved don't? (answer: sometimes, yes, but sufficient to override?)
- Added complexity to routing, troubleshooting
- Network sections won't scale to lots of networks

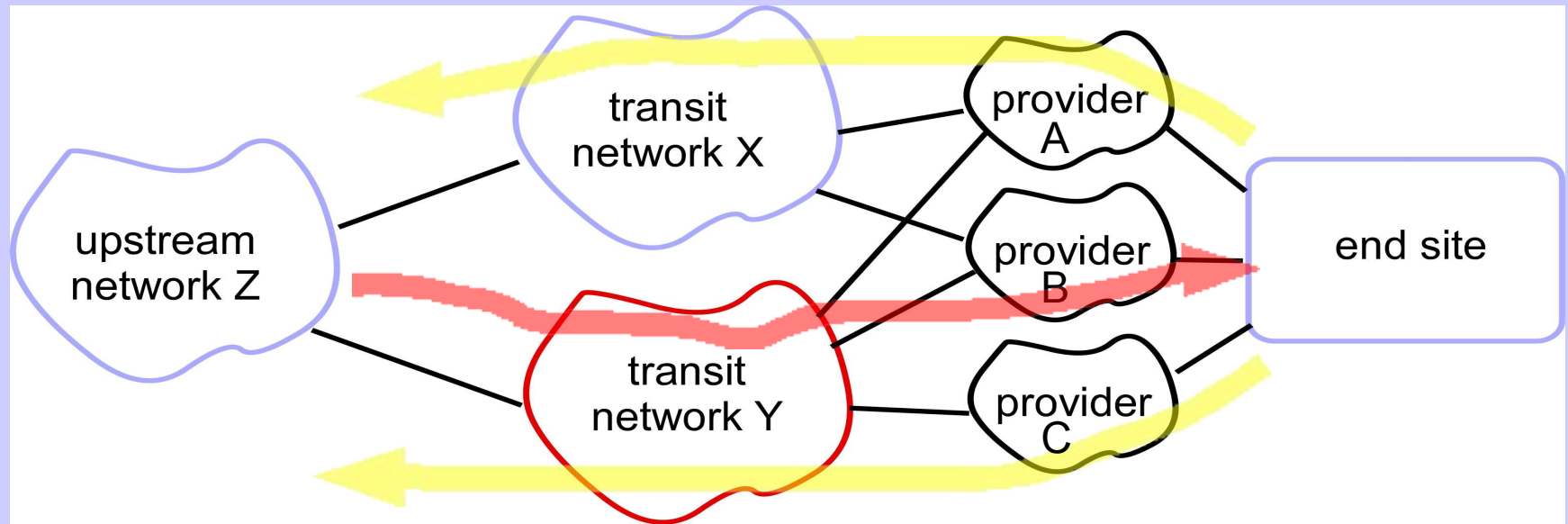
# How?

- Propose a 'well-known' consensus-agreed-upon set of unique BGP communities that work in the same way among any transits who choose to participate
- Set by originator of the BGP announcement, who is the owner of the 'hinted' destination
- Format: (transit-network) : (AS-path-hint-value)
  - e.g. 11537:64101
  - Where *transit-network* is the intermediate network you're hinting TO, and *AS-path-hint-value* is the key to what you're asking (currently a value in a list)
  - Read as “please, I2, when you're deciding which available path to take, I'd prefer you use AARnet”.
- All networks in the path must pass the communities
- Participating transit-nets modify their policy to implement hinting *any way they choose*

# How? (possible variations)

- Omit *transit-network* altogether (or optionally, so it applies to all)
- Only send a preference hint on the path you want followed *back* to you. For instance, if you're connected by networks X and Y and want X preferred, send a transitive well-known value in your announcements to X that networks *beyond* X could use to prefer the *X-path* back to you.
  - Use a special ('well-known') BGP community to signal request, e.g. nnnnn:7500
  - Some Pros: could scale better, no hint-list req'd
  - Some Cons: less deterministic and explicit

# How? (example)



1. End site sends 'hint' upstream: prefer path via transit network 'Y'
2. Upstream network 'Z' honors the request and prefers 'Y' path *for end site's requested prefixes*
3. Endsite could change preference at any time (*and so could 'Z'!*)

# What's necessary to make it work?

- (1) General agreement on a method (“critical mass”)
- (2) Willingness at some user sites to use this method for hinting (does this condition need to exist before #2 will happen?)
- (3) Agreement by some R&E transit networks—ideally those who carry traffic for users in #2 above—to implement that method and to honor at least *some* requests

# Possible hint values (example)

aarnet	64101	nii/sinet
apan/transpac	64102	nisn
asnet	64103	nlr
canet	64104	nren
clara	64105	reannz
cudi	etc...	sinet
dren		scinet (SC0x)
esnet		singaren
geant		surfnet
gemnet		tein
gloriad		transitrail
internet2		twanet(tanet)
kreonet		ultralight
LHCnet		usgs

# What next?

- Further discussion by community:
  - Is this a worthwhile idea? Is it really needed?
  - What's the best way to do it?
  - *How to improve scaling?*
  - Discuss where? RENOG.org list?
  - How do we know when we agree? Consensus? Lack of argument? Lack of interest? Someone willing to try?
- More discussion: member mtg, JET, Joint Techs
- Agreement on 'well-known' approach
- Public documentation of the consensus scheme
- Provide example policies for netOSs (e.g. IOS, JunOS)
- Some trial adopters
  - Need both requesting-sites and transit-nets
- Re-evaluation: Is it working as desired?
- Revision path: publicizing & implementing tweaks