

BGP “Path-Hinting” is real!

Brent Sweeny
GRNOC at Indiana University
sweeny@grnoc.iu.edu

March 2009 APAN27 Kaohsiung

(very) brief summary

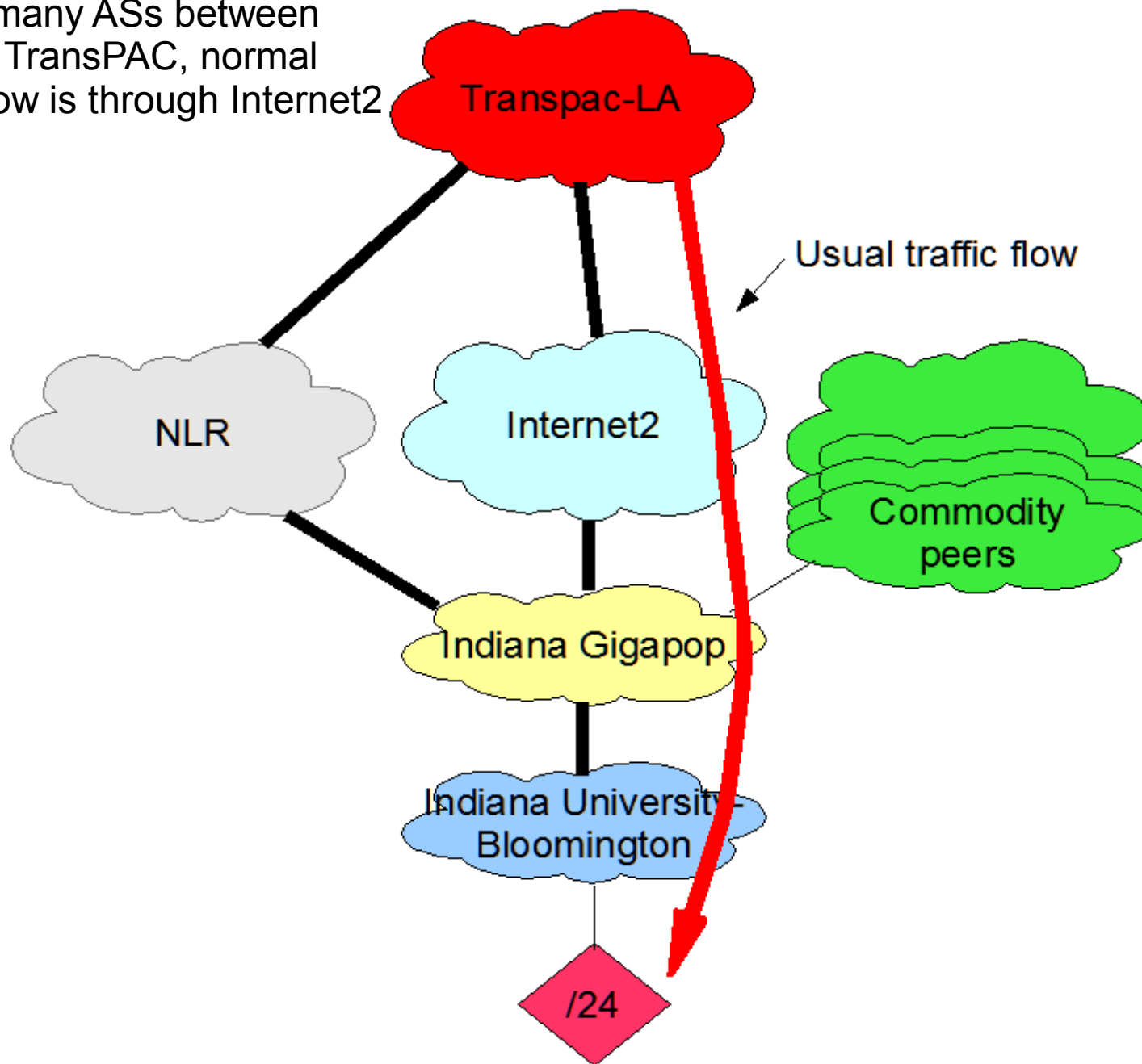
- Purpose is to allow end-sites to signal preference for return traffic path (back toward them)
- Available 'tools' for influencing return path-selection are ineffective for various reasons
- Particularly useful in cases where multiple paths through upstreams exist & some paths are equal-length but much more desirable (eg SC08)
- Uses simple site-controllable mechanism, a well-known BGP community with tags, simple for upstreams to 'programmize'
- Optionally implemented by upstreams



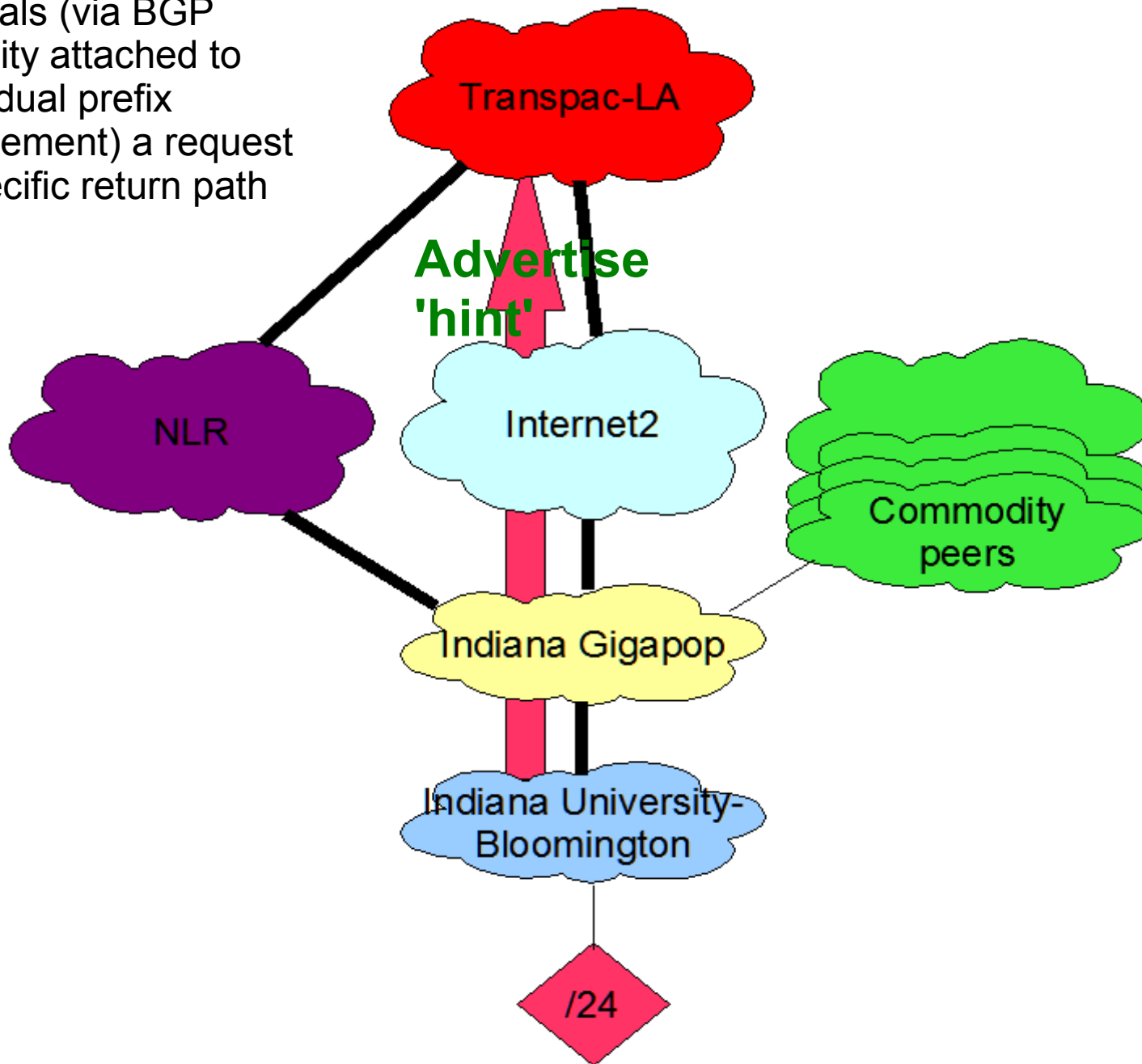
How does it work?

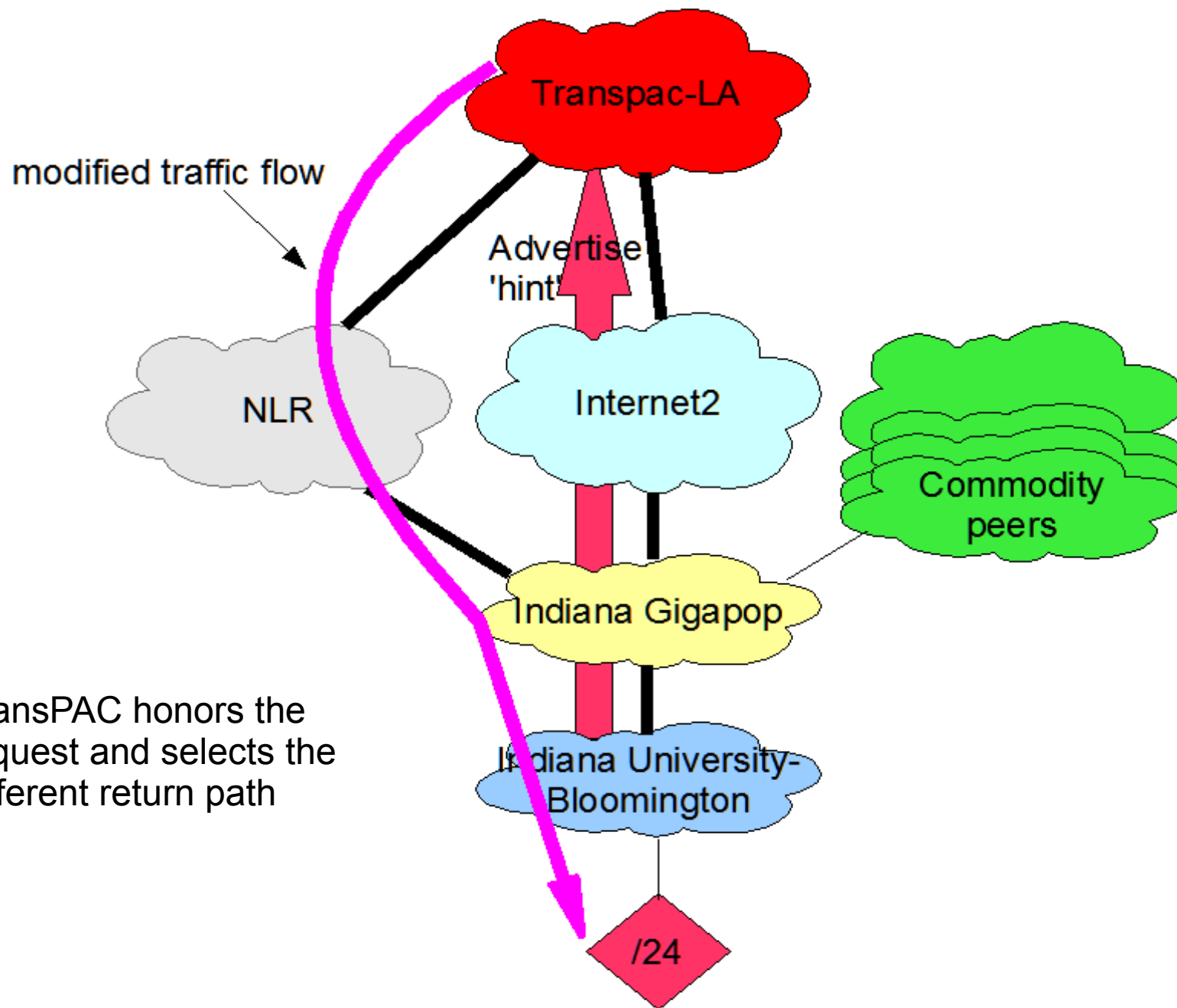
(bird's-eye view)

Before: many ASs between IUB and TransPAC, normal Traffic flow is through Internet2



IUB signals (via BGP community attached to an individual prefix announcement) a request for a specific return path





TransPAC honors the request and selects the different return path



How does it work?

(the details)

Requesting site marks traffic with 'hint'

```
from {  
    route-filter 129.79.9.1/32  
    exact;  
    # could also use prefix-list  
}  
then {  
    community add 27198:19401;  
}
```

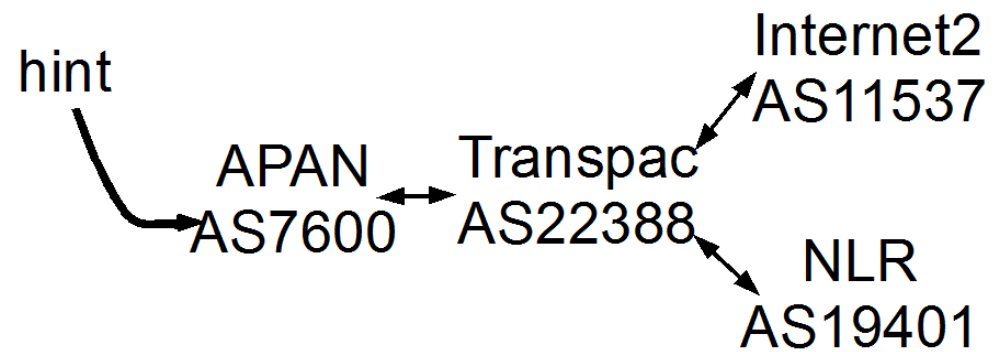
Upstream site honors 'hint'

(JunOS configuration in the TransPAC router):

```
community HINT-I2 members 27198:11537;    (11537 is I2's AS)
community HINT-NLR members 27198:19401;    (19401 is NLR's AS)
```

```
policy-statement HINT-IN {                    (add to BGP import policies)
  term hint-I2 {
    from community HINT-I2;
    then {
      local-preference 5000;
      next-hop 207.231.240.131;
      next term;
    } ..
  term hint-NLR {
    from community HINT-NLR;
    then {
      local-preference 5000;
      next-hop 207.231.241.14;
      next term;
    } ..
  } ..
}
```

Watch out for loops!



- Possible to introduce a loop IFF:
 - 'signaling' site requests hints to multiple networks
 - Transit site is connected to >1 of them
- To avoid this:
 - Be careful initiating multiple hints
 - Transit site may be able to avoid with more complicated policy, by examining AS-path or by examining hint-requests for directional

.. ..

Current status

- Did proof-of-concept in early 2008 from Indiana University through Indiana Gigapop, NLR & Internet2, to TransPAC in Los Angeles to change return path
- Worked with ARIN 2H 2008 to procure unique AS27198 used for 'signaling'; received Feb'09
- Several expressions of interest from R&E nets; need to convert those to action
- Document as draft RFC
- Goal to have in place in several R&E networks

Selected References

- Chandra, Traina, Li, RFC1997, “BGP Communities Attribute” (1996).
- Chen & Bates, RFC1998 “An Application of the BGP Community Attribute in Multi-home Routing” (1996).
- Sangli, Tappan, Rekhter, RFC4360, “BGP Extended Communities Attribute” (2006).
- Meyer, RFC4384/BCP114, “BGP Communities for Data Collection” (2006).
- IANA, “Data Collection Standard Communities per RFC4360” (2007).
- Olivier Bonaventure et al., Internet Draft (Draft-bonaventure-bgpredistribution),
“Controlling the redistribution of BGP Routes” (2002).
- Jin Tanaka (JP-NOC/KDDI), “BGP Routing with Communities”, presented at the 24 APAN Meeting Network Engineering Workshop, August 2007 and the October 2007 Internet2 Member Meeting RENOG session with additional suggestion from Akira Kato.
- Brent Sweeny, “BGP path 'hinting' proposal”, presented as it evolved to JET, Internet2/ESnet Joint Techs, and Internet2 Member meetings 2006-2008.