IAB Routing Workshop
Traffic Engineering Case Studies

Jason Schiller

Business Requirements for TE

- TE requirements are usually driven by business rules
- TE policy set for the organization by the person who purchases the Internet connectivity
- TE policy currently configured in a few places (Internet routers)

- Enterprise end sites business requirements for TE
- Small ISP business requirements for TE
- Large ISP business requirements for TE

Multi-homing TE Building Blocks

- Primary / Backup
- Load sharing across all paths
- Best path

Additional requirements to shift traffic between links with any of the three options, such as pushing traffic away from over utilized links

- Complex combinations of the three cases

Case 1: Primary / Backup

Requires the ability to designate a link or set of links as the primary link to use for all traffic to one or more destination prefixes. Primary link should carry all traffic for the designated prefixes. The backup link should only carry traffic if the primary goes down.

Common reason for primary / backup configuration is one link is more expensive than the other

- Cost
- Latency
- Performance
- Bandwidth

Case 1: Primary / Backup – Implementation

If all links are to the same AS, inbound traffic manipulated by setting higher MED on backup link(s)

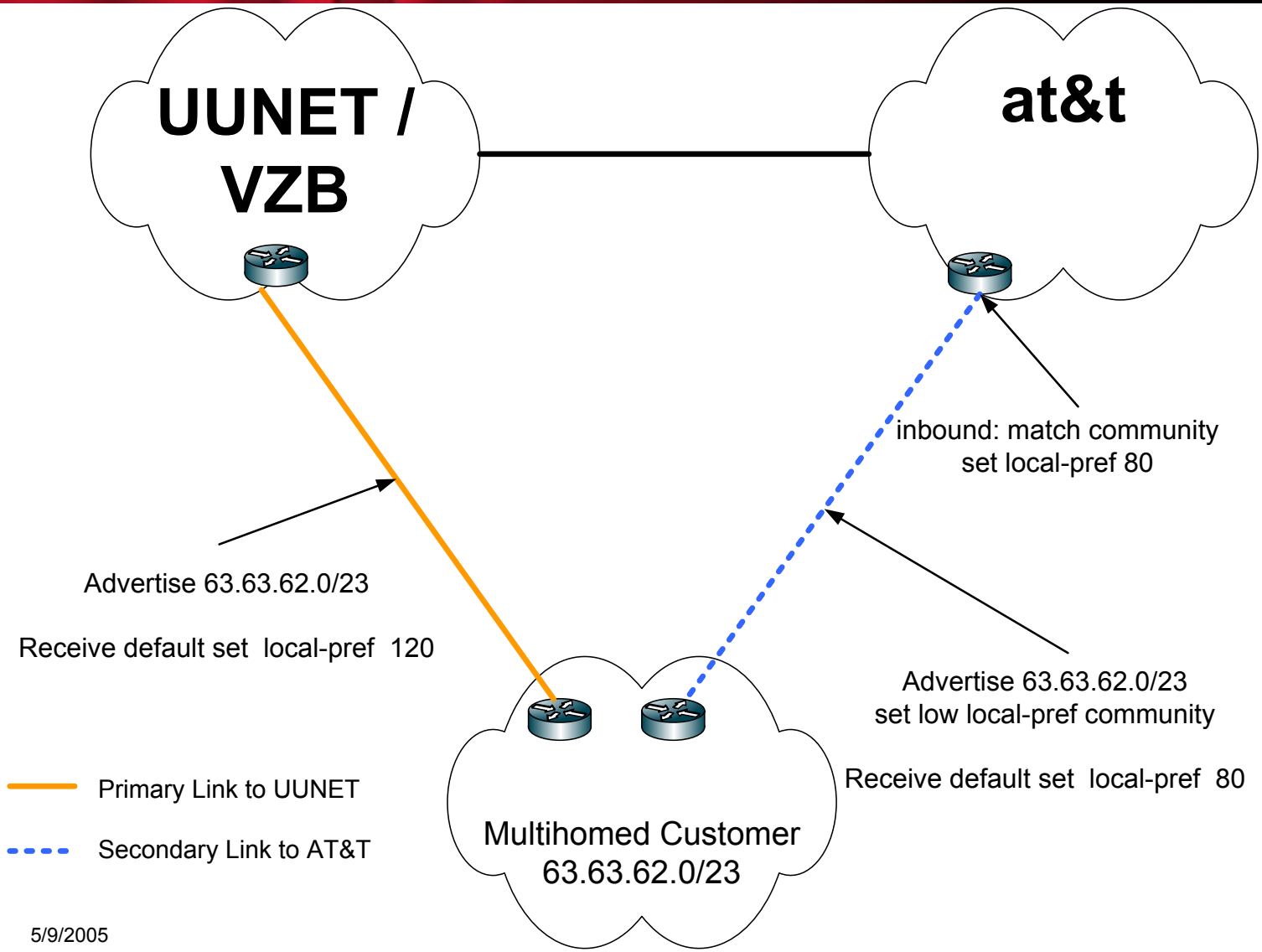
If links are to different ASes, inbound traffic manipulated by having the provider set low local-preference on backup link(s)

Outbound traffic manipulated by learning a default route and setting higher MED or lower local-preference on the backup link.
(can set lower MED or higher local-preference on primary link)

Outbound traffic can be manipulated by weighted static default routes.

Can be configured to have multiple level of backup links (secondary, tertiary, etc..)

Case 1: Primary / Backup



Case 2: Load Sharing

Requires loading traffic on all links. The goal is to load the links as evenly as possible without negative impact on traffic flows.

Common reason for load sharing is to squeeze as much bandwidth out of multiple links as possible. This is often the case where larger links are cost prohibitive such as for small companies or locations where circuit cost is high.

Case 2: Load Sharing – Implementation

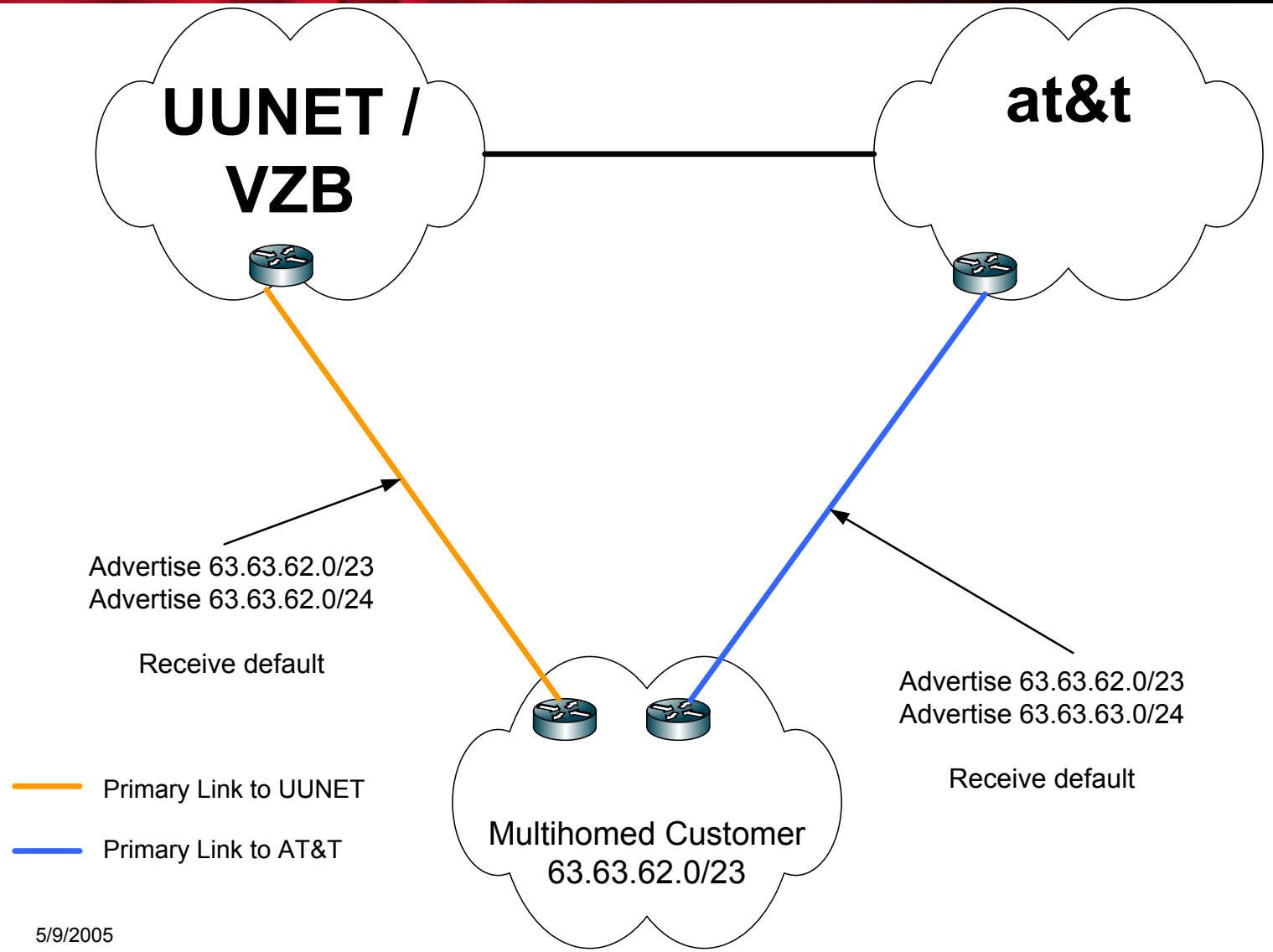
Inbound traffic manipulated by dividing IP space and making particular more specific route announcements across different links in addition to the aggregate.

Outbound traffic manipulated by depreferencing certain inbound route announcements by setting a MED value or local-pref inbound.

Outbound traffic manipulated by adjusting IGP metrics to make certain hosts closer to certain exit points.

Outbound traffic can be manipulated by equal cost static default routes.

Case 2: Load Sharing



Case 3: Best Path

Requires the ability to use a non-random “best” path (For some definition of best).

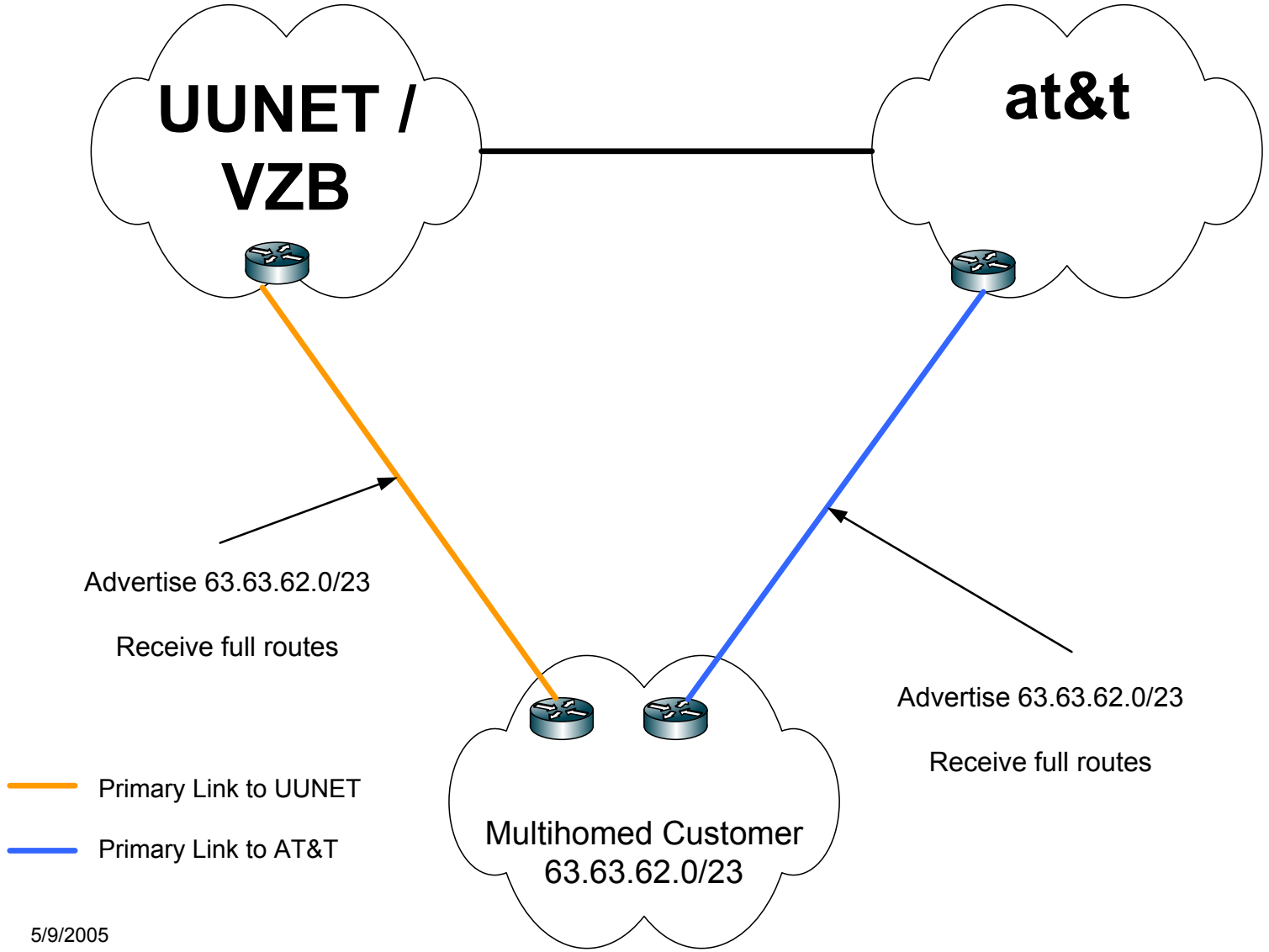
Current best path is based on routing information based on BGP path selection algorithm (LocalPref, AS-path, origin code, MED, eBGP over iBGP, IGP distance, RR cluster length, RID, lowest neighbor)

Best path approximates “shortest” path to host

Inbound traffic manipulated by source BGP table best path selection

Outbound traffic manipulated by learning full BGP routes from all upstream ISPs

Case 3: Best Path



Dialing Traffic

BGP lacks a congestion control mechanism (I'm not suggesting it be added).

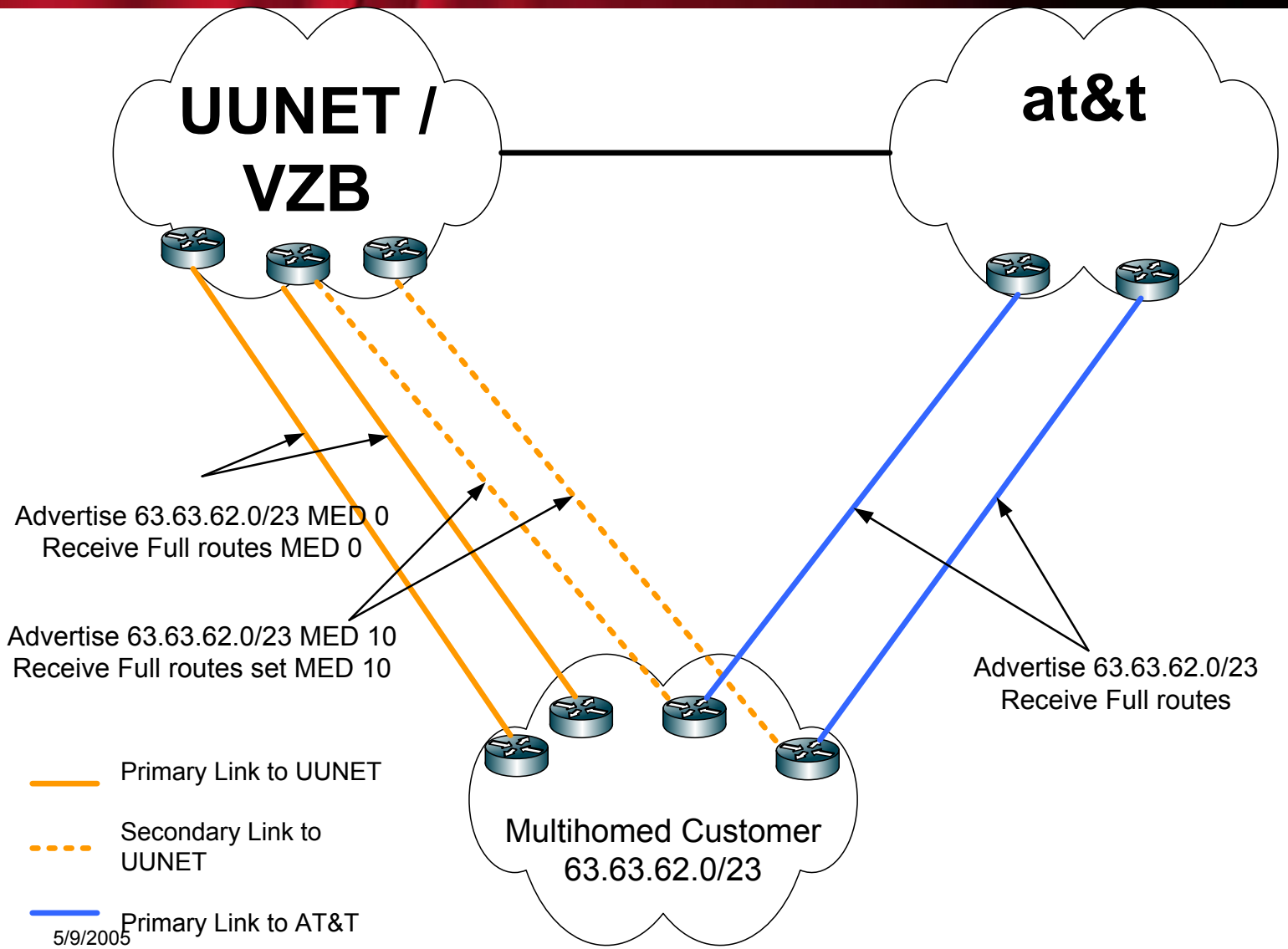
- To avoid congestion operators will shift traffic away from over-utilized outbound links, and attempt to draw traffic to under-utilized inbound links. This is a manual process to avoid congestion.
  - Primary – backup move a small amount of traffic to the backup link
  - Load sharing move traffic away from congested links and toward non-conjested links
  - Best path move some sources from the best path to then next best path if the difference between the paths is small

Complex Combination of Cases

Inter-AS BGP traffic engineering can be a combination of the 3 cases and further refined by dialing the traffic up or down.

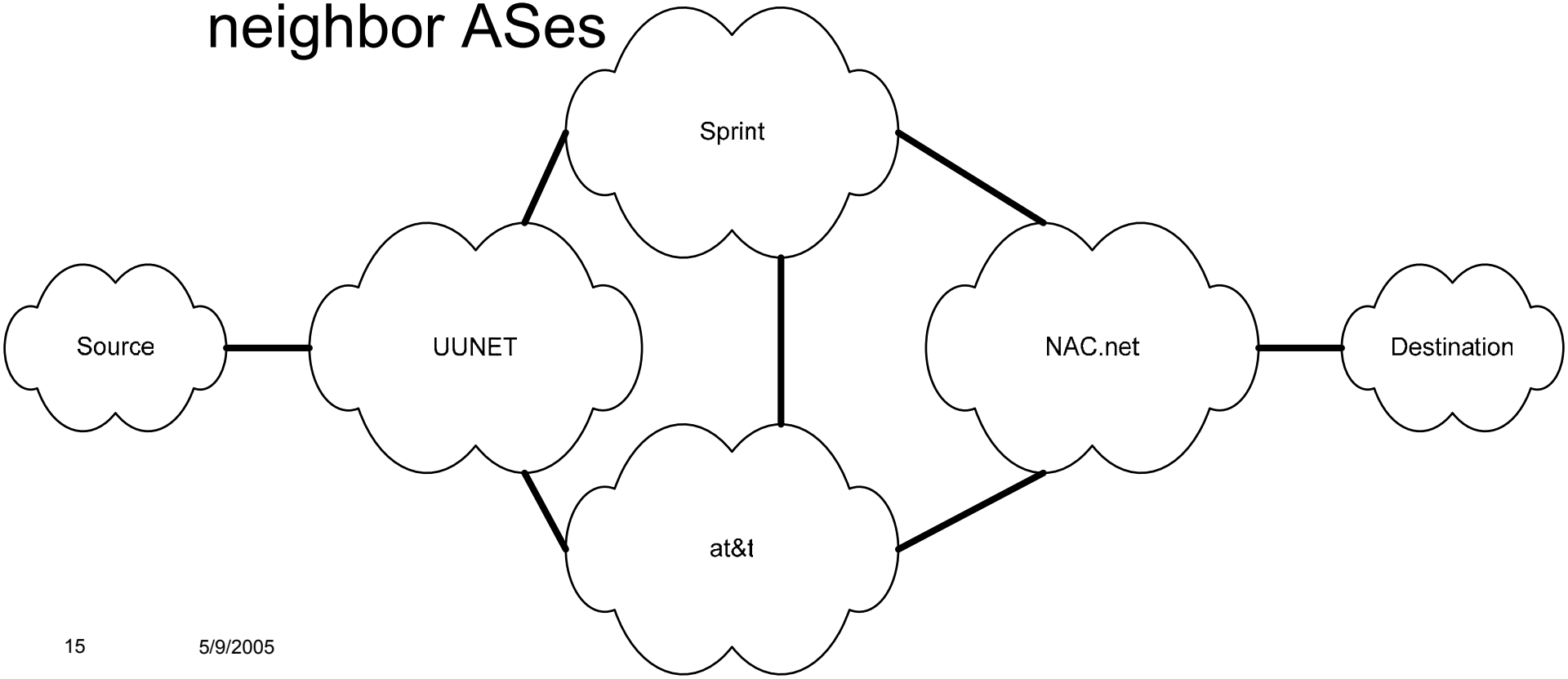
You could imagine a customer with links to two ISPs say UUNET / VZB and at&t, where best path is used inbound and outbound between the customer and both upstream ISPs. Also imagine the connection to UUNET / VZB consists of a pair of primary links and a pair of backup links while the links to at&t consist of a pair of primary links.

Complex Combination



Transit AS TE

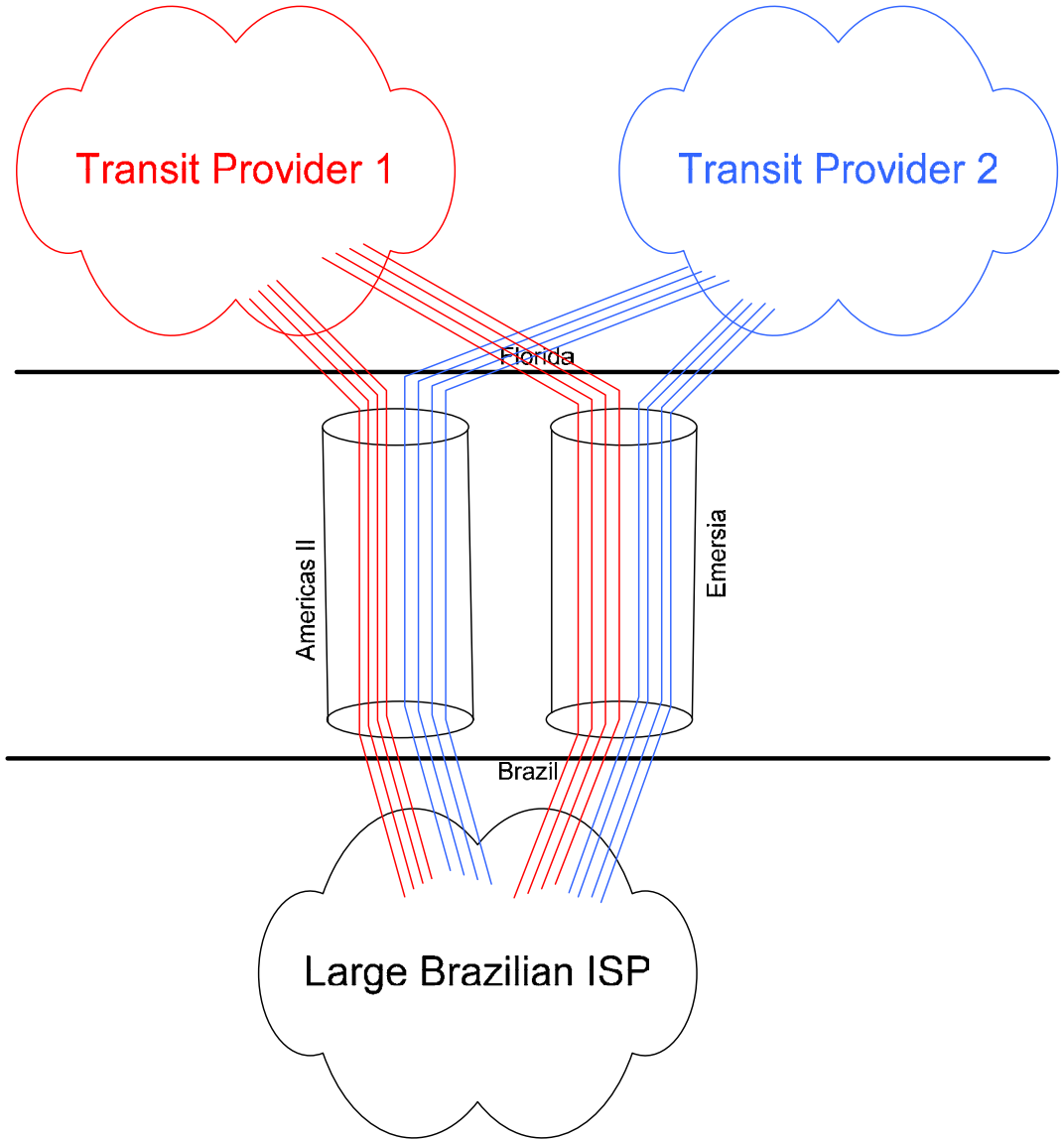
Transit ASes have TE choices when the destination is not directly connected and reachable equally as well through multiple neighbor ASes



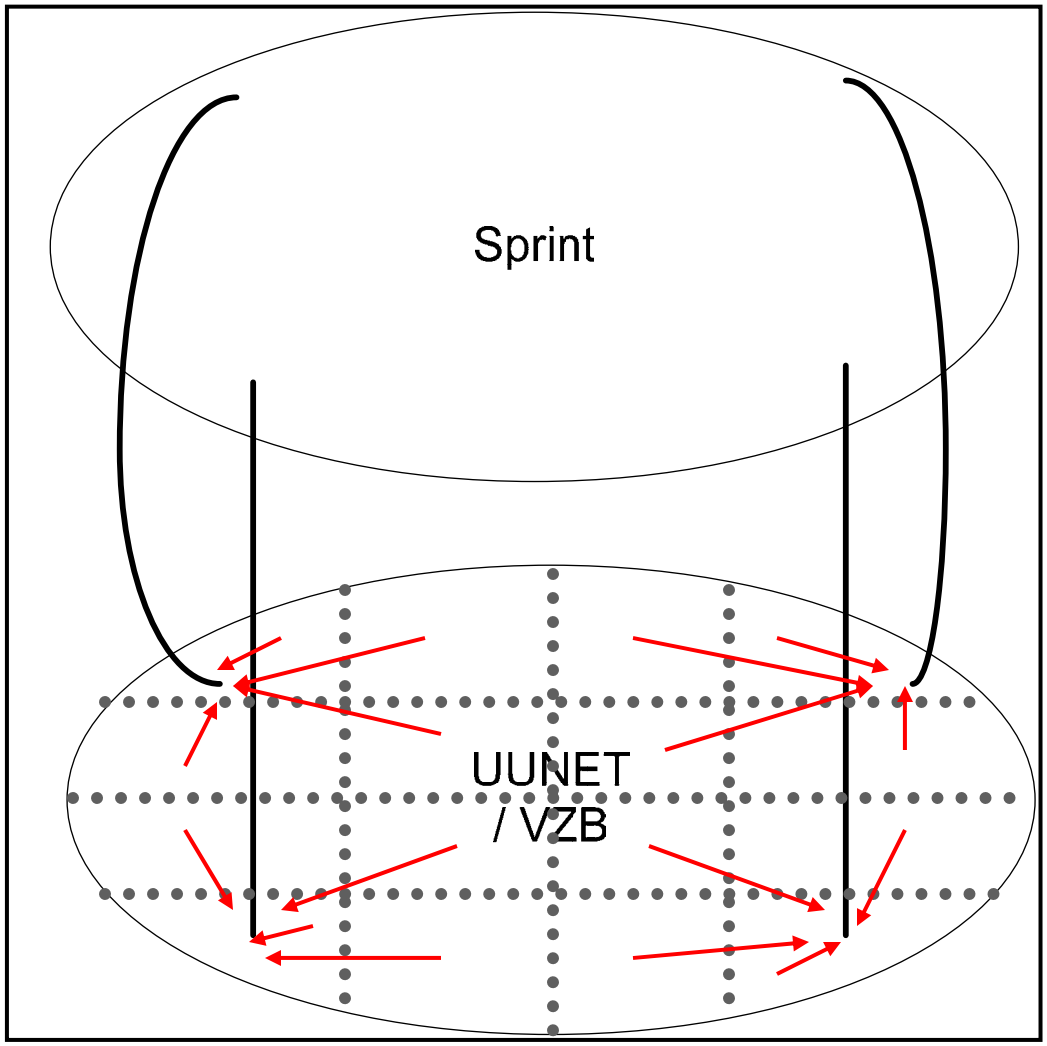
Small ISP Transit TE

- Small ISPs usually use TE towards their transit providers to get better performance / utilize most bandwidth
- Consider a large ISPs in South America
  - Two transit providers
    - transit provider 1 is better connected
    - Transit provider 1 offers better DoS mitigation
  - 16 x STM-1 to each transit provider split evenly across two oceanic cable systems
  - One oceanic cable is lower latency
  - Need to load up all oceanic links fully due to cost

Small ISP Transit TE



Transit TE importance to Large ISPs



Transit TE importance to Large ISPs