

Route policies are accomplished by constraining which BGP routes are selected and exported



Selection

which path to use?



Export

which path to advertise?

Route policies are defined by constraining which BGP routes are selected and exported



Selection

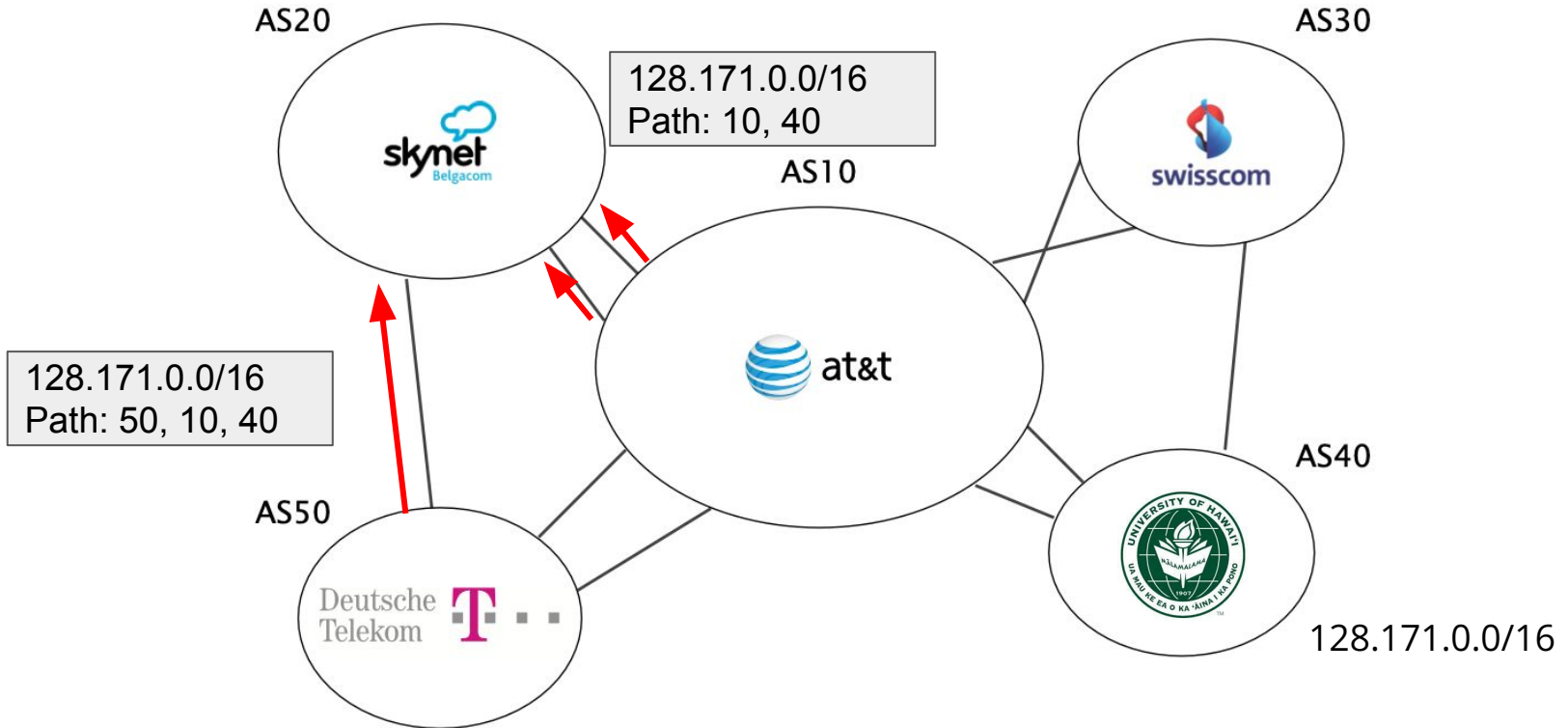
which path to use?

Control outbound traffic

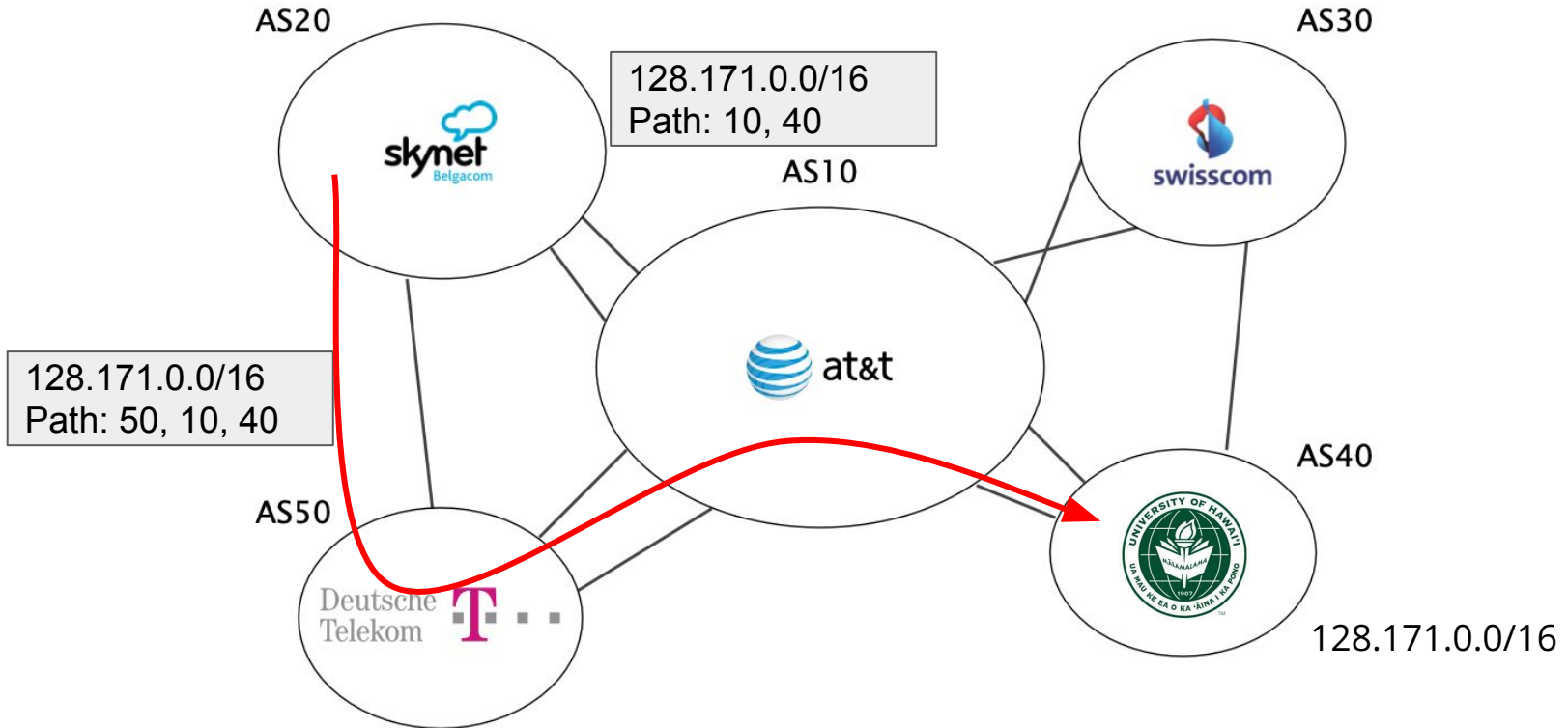
Export

which path to advertise?

Always prefer DT over AT&T



Always prefer DT over AT&T



Business relationships condition route selection

For a destination p , prefer routes coming from

- customers over
- peers over
- providers



route type

Route policies are defined by constraining which BGP routes are selected and exported

Selection

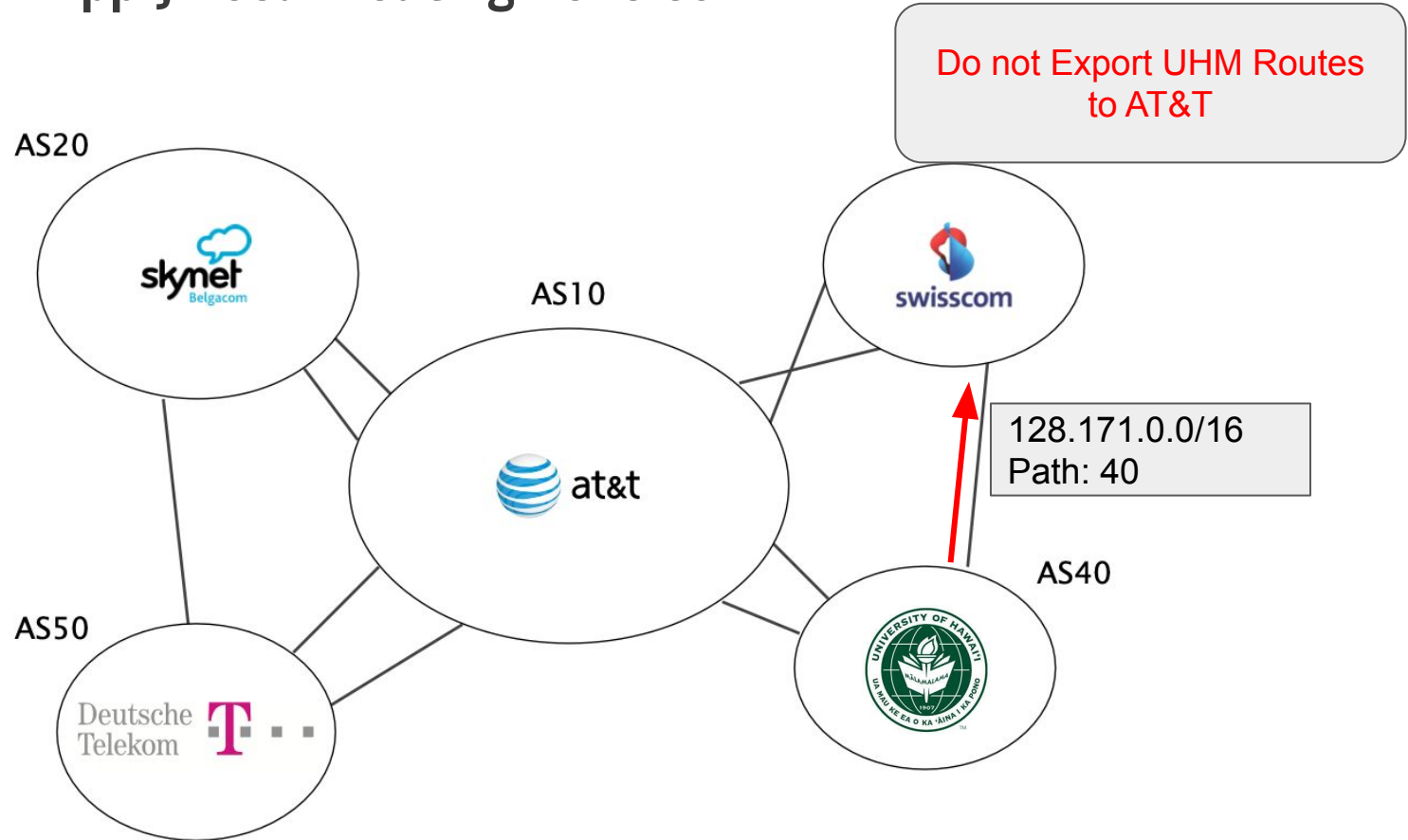
which path to use?

Export

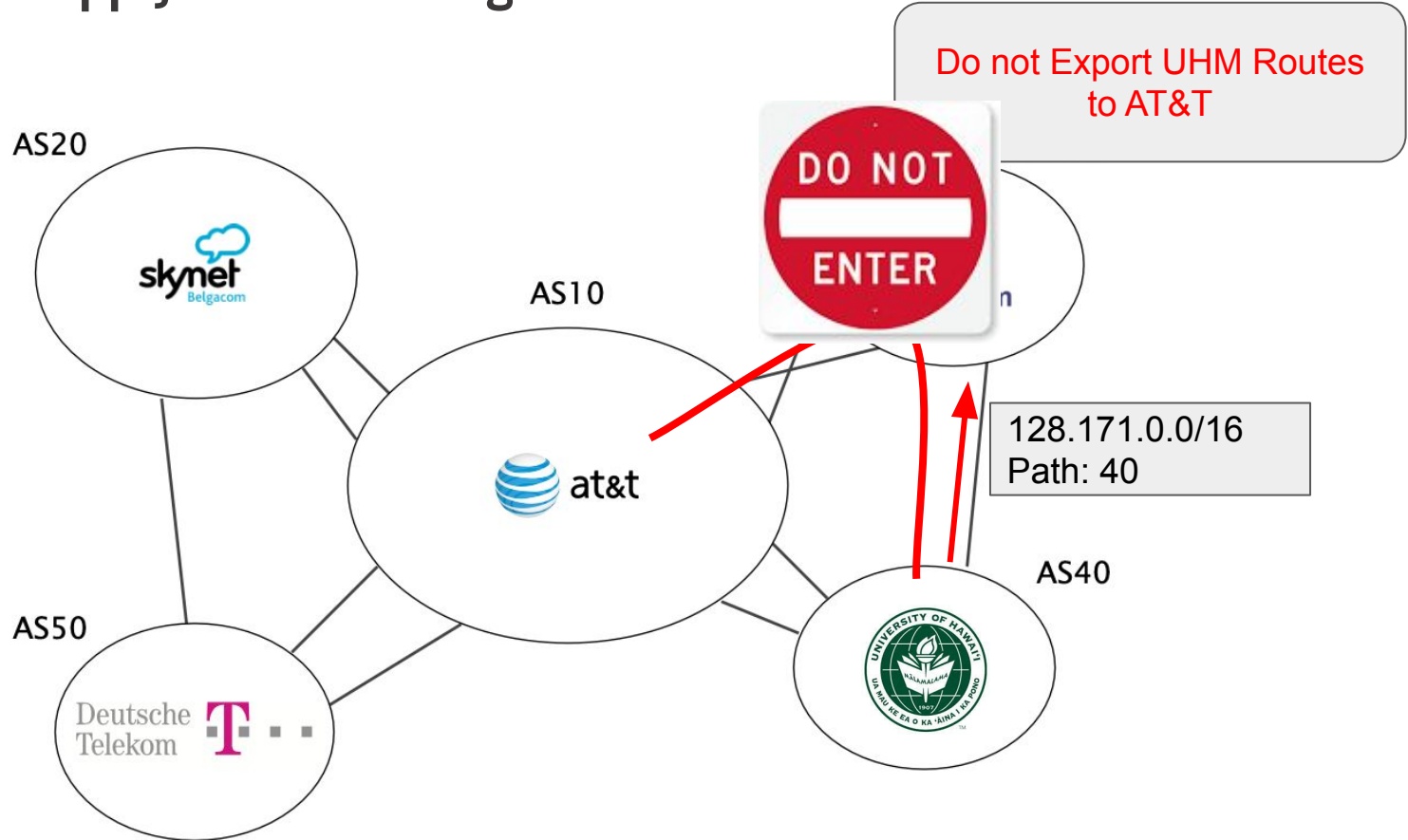
which path to advertise?

Control inbound traffic

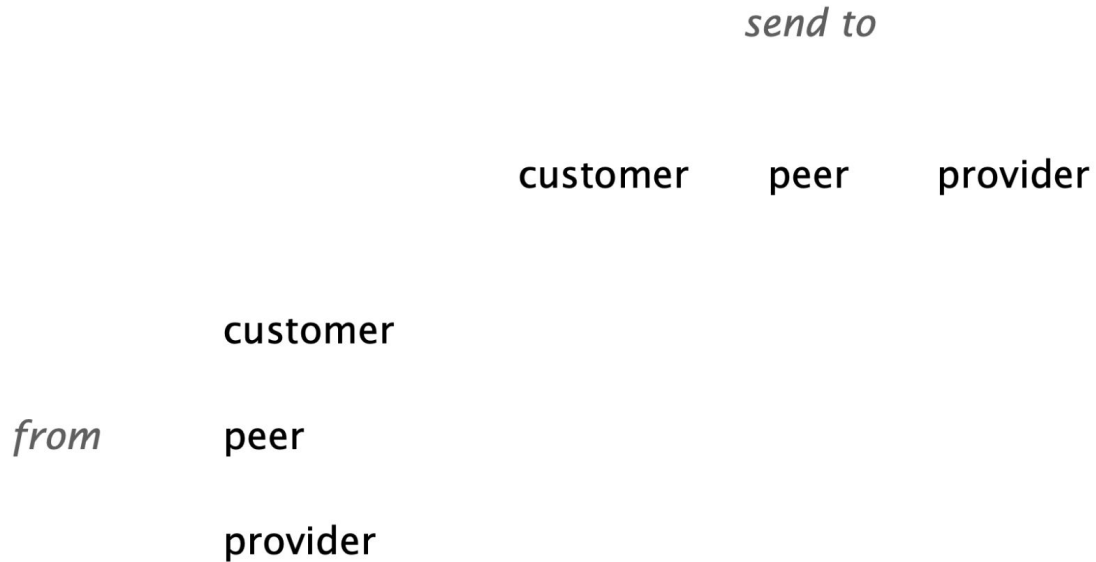
Each AS Can Apply Local Routing Policies






Each AS Can Apply Local Routing Policies




Business relationships constrain route exportation



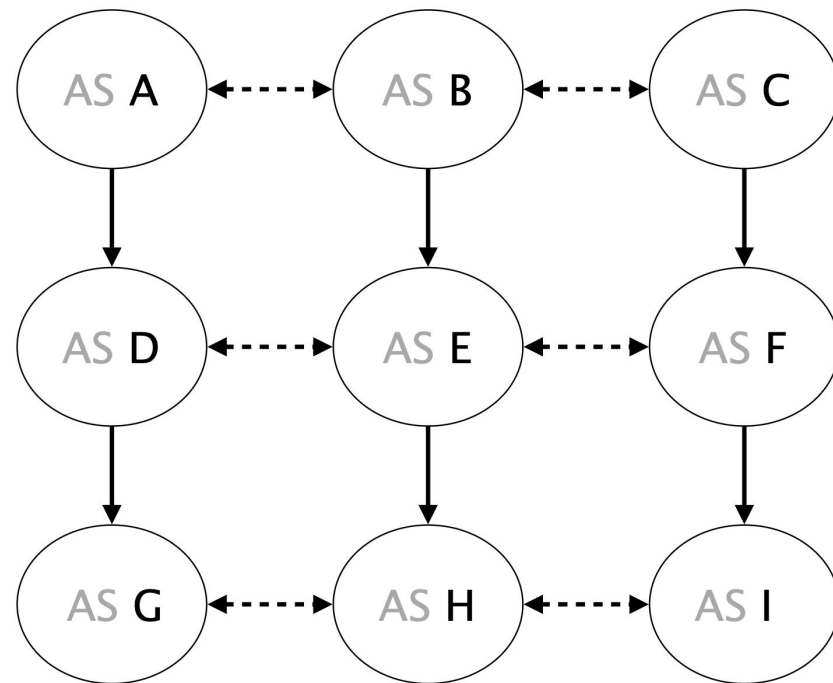
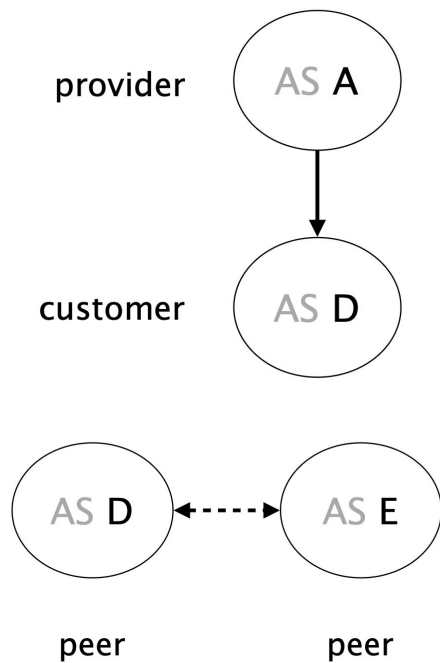
Routes from customers are propagated to everyone else

		<i>send to</i>		
		customer	peer	provider
<i>from</i>	customer			
	peer			
	provider			

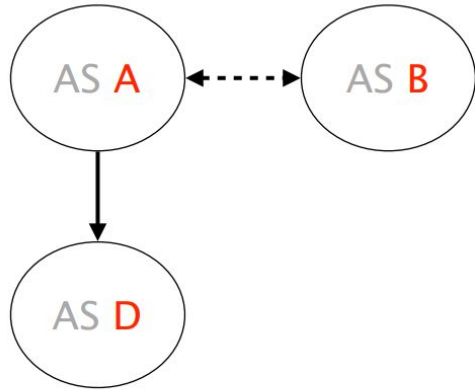
Routes from peers and providers are only propagated to customer

		<i>send to</i>		
		customer	peer	provider
<i>from</i>	customer			
	peer		-	-
	provider		-	-

Example

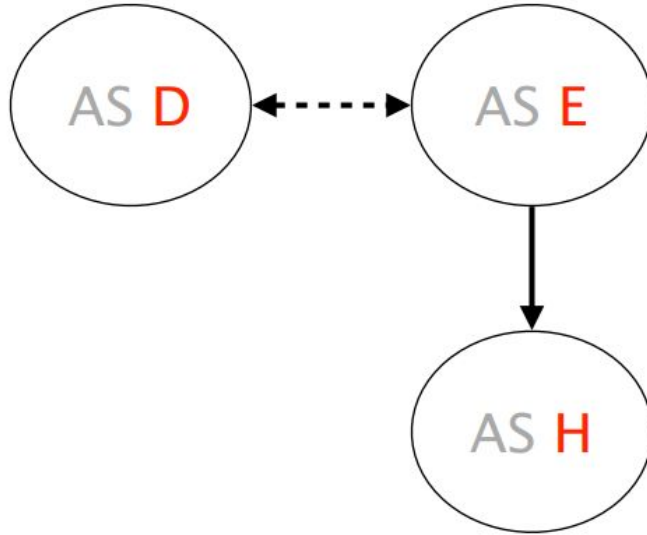


Example



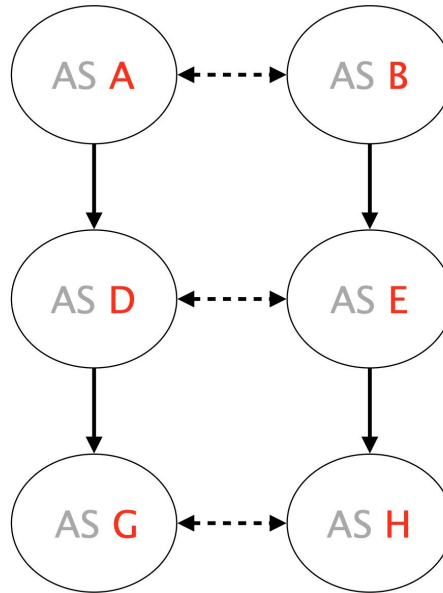
Is (B, A, D) a valid path?

Example



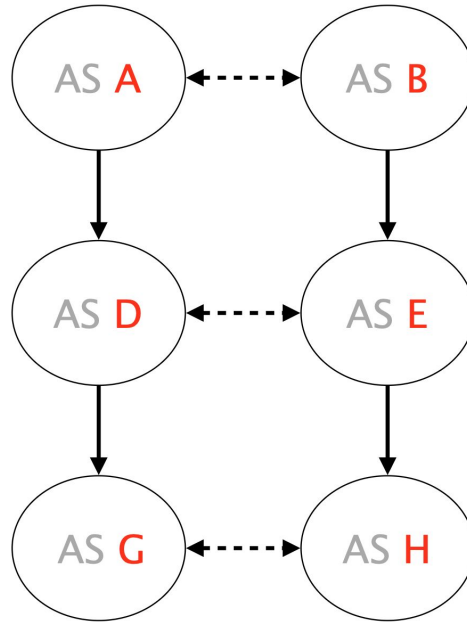
Is (H, E, D) a valid path?

Example



Is (G,D,A,B,E,H) a valid path?

Example



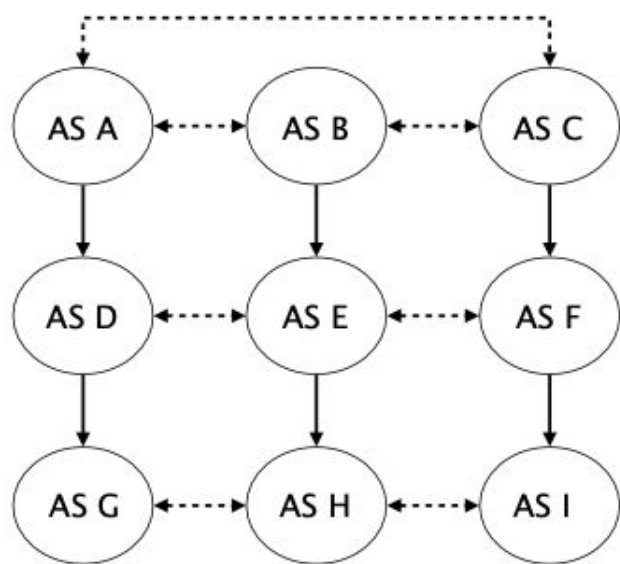
Will (G,D,A,B,E,H) actually see packets?

Questions

Consider now the network depicted on the right. Single-headed plain arrows point from providers to their customers (AS A is the provider of AS D), while double-headed dashed arrows connect peers (AS D and AS E are peers). Each AS in the network originates a unique prefix that it advertises to all its BGP neighbors. Each AS also applies the default selection and exportation BGP policies based on their customers, peers and providers.

What path (sequence of ASes) is followed when AS G sends packets destined to the prefix originated by AS E?

What path (sequence of ASes) is followed when AS F sends packets destined to the prefix originated by AS G?



A simple BGP network

Questions

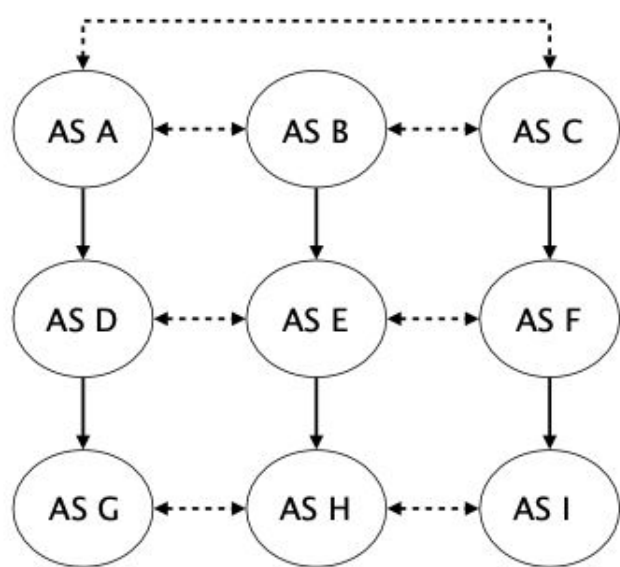
Consider now the network depicted on the right. Single-headed plain arrows point from providers to their customers (AS A is the provider of AS D), while double-headed dashed arrows connect peers (AS D and AS E are peers). Each AS in the network originates a unique prefix that it advertises to all its BGP neighbors. Each AS also applies the default selection and exportation BGP policies based on their customers, peers and providers.

What path (sequence of ASes) is followed when AS G sends packets destined to the prefix originated by AS E?

Solution: [G, D, E]

What path (sequence of ASes) is followed when AS F sends packets destined to the prefix originated by AS G?

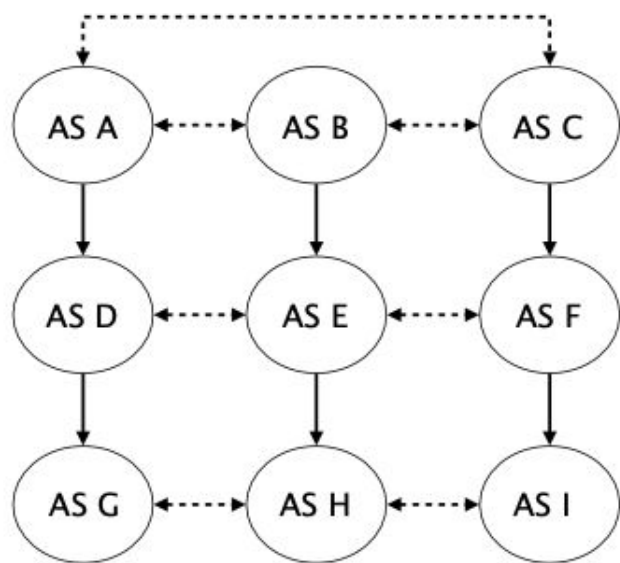
Solution: [F, C, A, D, G]



A simple BGP network

Questions

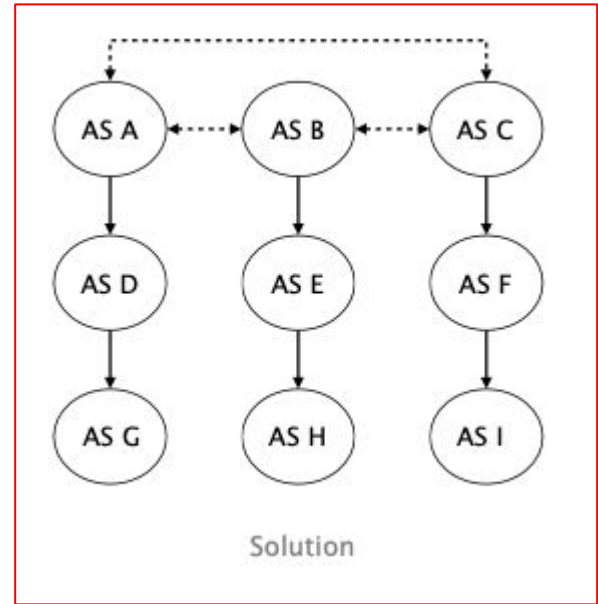
Suppose AS A and AS C give you a “dump” of all the BGP routes they learn for every destination. You then extract all links from the AS paths seen in those “dumps” and use them to construct a view of the AS-level topology. Draw the resulting AS-level topology.



A simple BGP network

Questions

Suppose AS A and AS C give you a “dump” of all the BGP routes they learn for every destination. You then extract all links from the AS paths seen in those “dumps” and use them to construct a view of the AS-level topology. Draw the resulting AS-level topology.



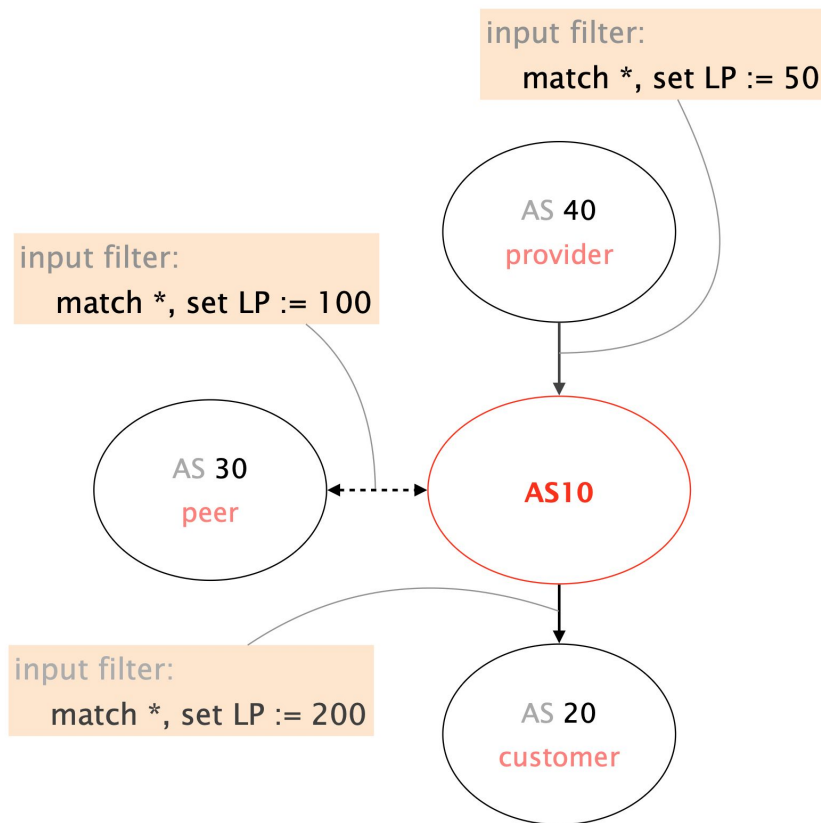
BGP Policy Wrapup

To implement their selection policy, operators define input filters which manipulates the LOCAL-PREF

For a destination p , prefer routes coming from

- customers over
 - peers over
 - providers
- 
- route type*

BGP Policy Wrapup - LP is how this is done



BGP Problems

BGP has numerous problems

Problems

Reachability

Security

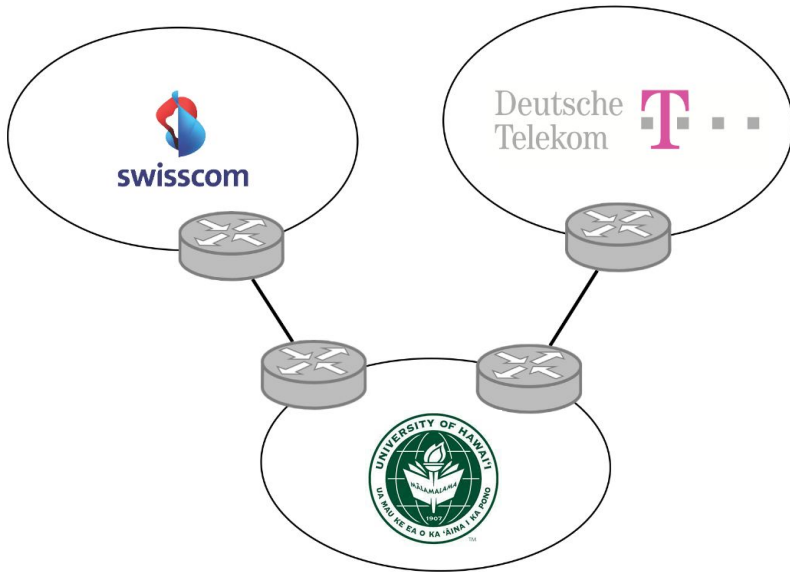
Convergence

Performance

Anomalies

Relevance

Unlike normal routing, policy-based routing does NOT guarantee reachability even if the graph is connected



Because of policies,
Swisscom cannot reach DT
even if the graph is connected

BGP attacks on underlying TCP

- BGP session runs over TCP
 - TCP connection between neighboring routers
 - BGP messages sent over TCP connection
 - Makes BGP vulnerable to attacks on TCP
- Main kinds of attacks
 - Against confidentiality: eavesdropping
 - Against integrity: tampering
 - Against performance: denial-of-service
- Main defenses
 - Message authentication or encryption
 - Limiting access to physical path between routers
 - Defensive filtering to block unexpected packets

BGP confidentiality attacks

- Eavesdropping
 - Monitoring the messages on the BGP session
 - ... by tapping the link(s) between the neighbors
- Reveals sensitive information
 - Inference of business relationships
 - Analysis of network stability
- Reasons why it may be hard
 - Challenging to tap the link
 - Often, eBGP session traverses just one link
 - ... and may be hard to get access to tap it
 - Encryption may obscure message contents
 - BGP neighbors may run BGP over IPSec

BGP message integrity attacks

- Tampering
 - Man-in-the-middle tampers with the messages
 - Insert, delete, modify, or replay messages
- Leads to incorrect BGP behavior
 - Delete: neighbor doesn't learn the new route
 - Insert/modify: neighbor learns bogus route
- Reasons why it may be hard
 - Getting in-between the two routers is hard
 - Use of authentication (signatures) or encryption
 - Spoofing TCP packets the right way is hard
 - Getting past source-address packet filters
 - Generating the right TCP sequence number

BGP denial of service attacks

- Overload the link between the routers
 - To cause packet loss and delay
 - ... disrupting the performance of the BGP session
- Relatively easy to do
 - Can send traffic between end hosts
 - As long as the packets traverse the link
 - (which you can figure out from traceroute)
- Easy to defend
 - Give higher priority to BGP packets
 - E.g., by putting packets in separate queue

BGP denial of service attacks - part 2

- Third party sends bogus TCP packets
 - FIN/RST to close the session
 - SYN flooding to overload the router
- Leads to disruptions in BGP
 - Session reset, causing transient routing changes
 - Route-flapping, which may trigger flap damping
- Reasons why it may be hard
 - Spoofing TCP packets the right way is hard
 - Difficult to send FIN/RST with the right TCP header
 - Packet filters may block the SYN flooding
 - Filter packets to BGP port from unexpected source
 - ... or destined to router from unexpected source

Exploiting the IP TTL field

- BGP speakers are usually one hop apart
 - To thwart an attacker, can check that the packets carrying the BGP message have not traveled far
- IP Time-to-Live (TTL) field
 - Decrement once per hop
 - Avoids packets staying in network forever
- Generalized TTL Security Mechanism (RFC 3682)
 - Send BGP packets with initial TTL of 255
 - Receiving BGP speaker checks that TTL is 254
 - ... and flags and/or discards the packet others
- Hard for third-party to inject packets remotely

Many security considerations are absent from the BGP specification

ASes can advertise any prefixes
even if they don't own them!

ASes can arbitrarily modify route content
e.g., change the content of the AS-PATH

ASes can forward traffic along different paths
than the advertised one

BGP's (terrible) security

- #1 BGP does not validate the origin of advertisements
- #2 BGP does not validate the content of advertisements

BGP's (terrible) security

#1 BGP does not validate the origin of advertisements

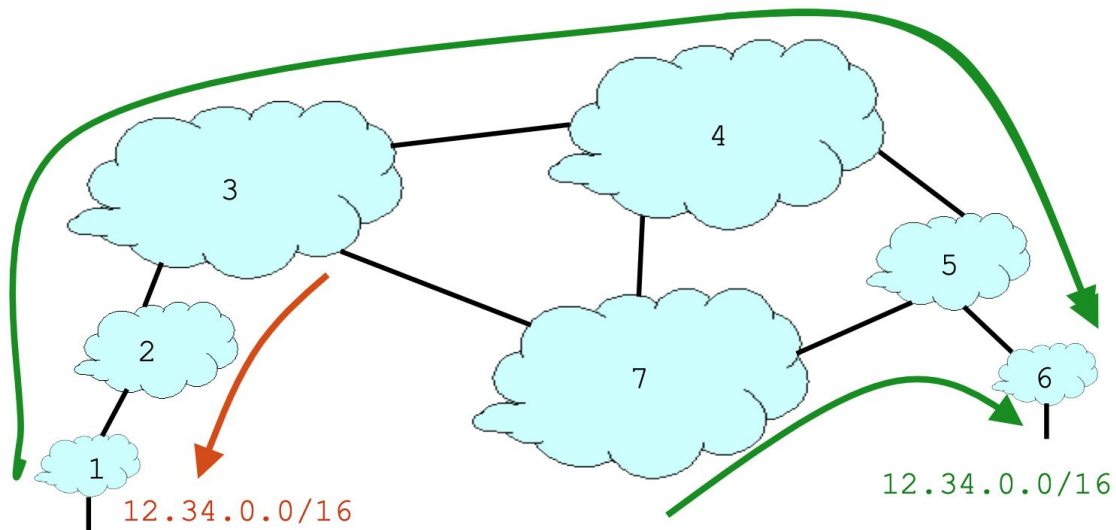
#2 BGP does not validate the content of advertisements

IP Address Ownership / Hijacking

- IP address block assignment
 - Regional Internet Registries (ARIN, RIPE, APNIC)
 - Internet Service Providers
- Proper origination of a prefix into BGP
 - By the AS who owns the prefix
 - ... or, by its upstream provider(s) in its behalf
- However, what's to stop someone else?
 - Prefix hijacking: another AS originates the prefix
 - BGP does not verify that the AS is authorized
 - Registries of prefix ownership are inaccurate

Prefix Hijacking

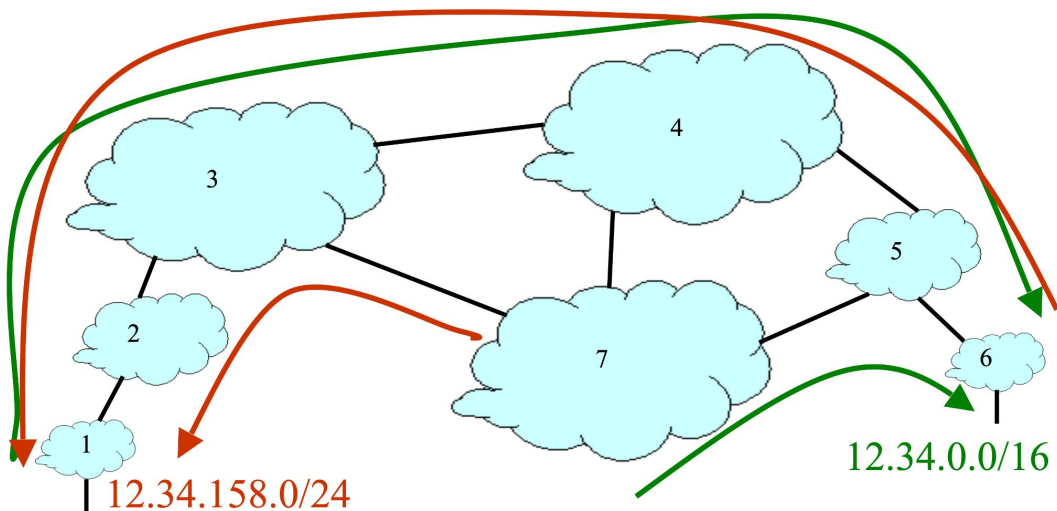
- **Blackhole:** data traffic is discarded
- **Snooping:** data traffic is inspected, then redirected
- **Impersonation:** traffic sent to bogus destinations



Hijacking is not easy to debug

- The victim AS doesn't see the problem
 - Picks its own route, might not learn the bogus route
- May not cause loss of connectivity
 - Snooping, with minor performance degradation
- Or, loss of connectivity is isolated
 - E.g., only for sources in parts of the Internet
- Diagnosing prefix hijacking
 - Analyzing updates from many vantage points
 - Launching traceroute from many vantage points

Sub-Prefix Hijacking



- Originating a more-specific prefix
 - **Every** AS picks the bogus route for that prefix
 - Traffic follows the longest matching prefix