Interdomain Routing Protocols

MẠNG MÁY TÍNH NÂNG CAO

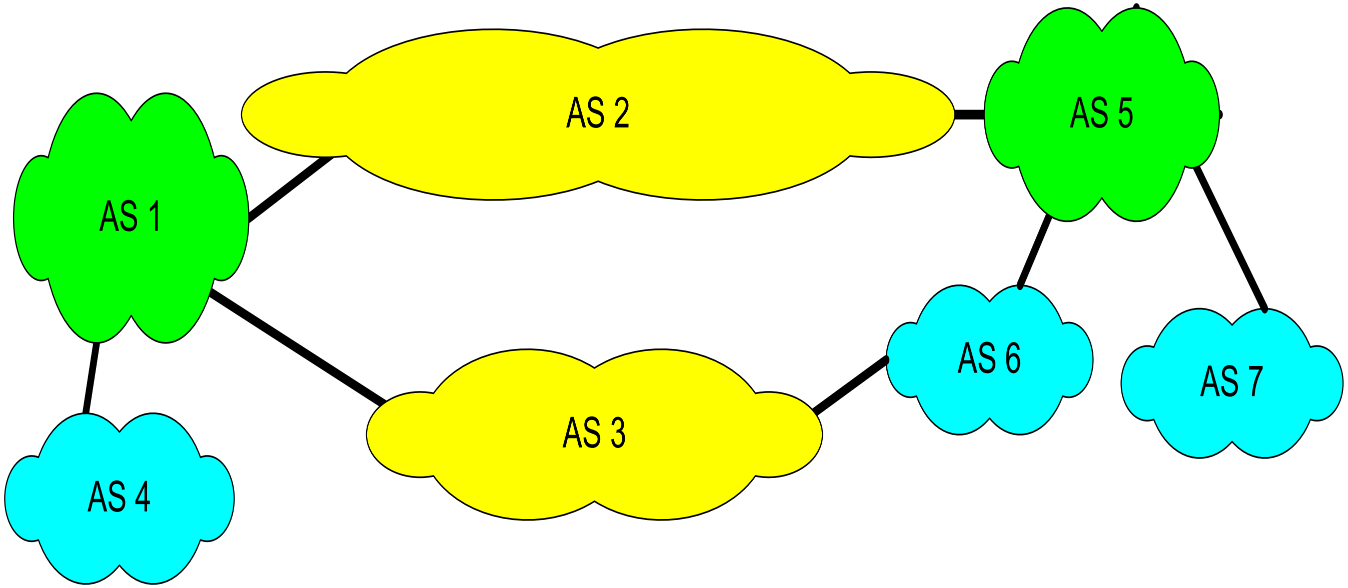
lnson@fit.hcmus.edu.vn

Autonomous Systems

- ❑ An **autonomous system (AS)** is a region of the Internet that is administered by a single entity and that has a unified routing policy

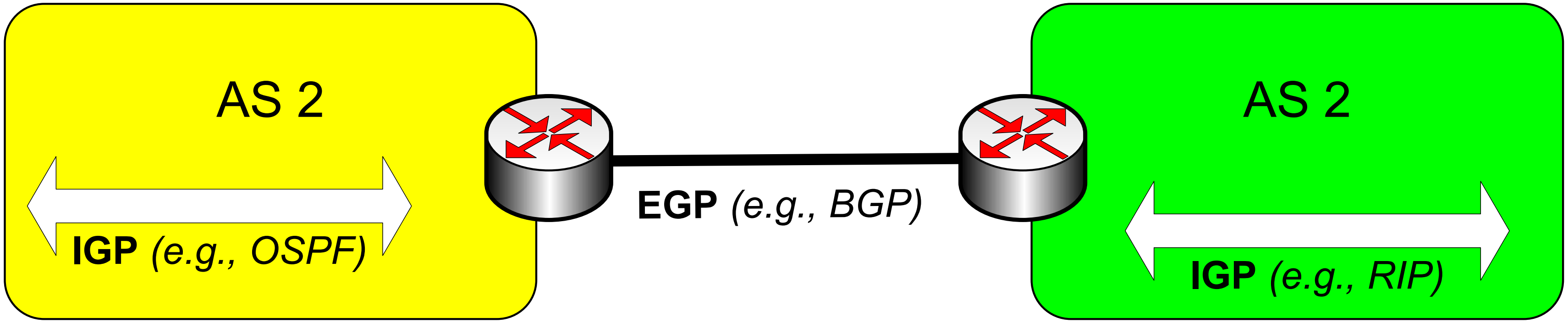
- ❑ Each autonomous system is assigned an Autonomous System Number (**ASN**).
  - UofT's campus network (AS239)
  - Rogers Cable Inc. (AS812)
  - Sprint (AS1239, AS1240, AS 6211, …)
- ❑ Interdomain routing is concerned with determining paths between autonomous systems **(interdomain routing**)
- ❑ Routing protocols for interdomain routing are called **exterior gateway protocols** (EGP)

Interdomain and Intradomain Routing



- Routing protocols for intradomain routing are called interior gateway protocols (IGP)
  - Objective: shortest path
- Routing protocols for interdomain routing are called exterior gateway protocols (EGP)
  - Objective: satisfy policy of the AS

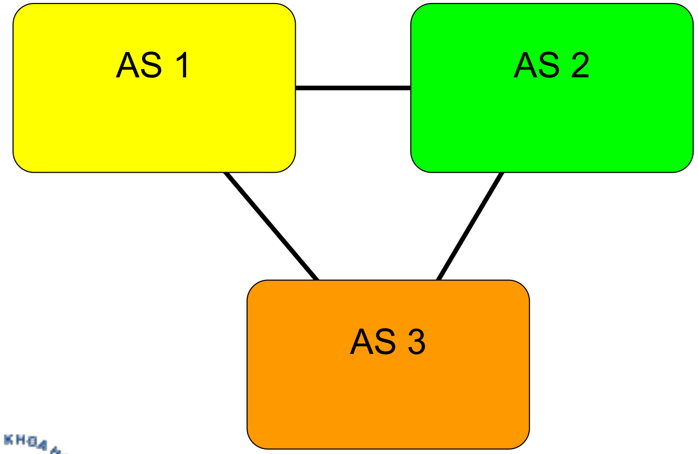
Interdomain vs Intradomain

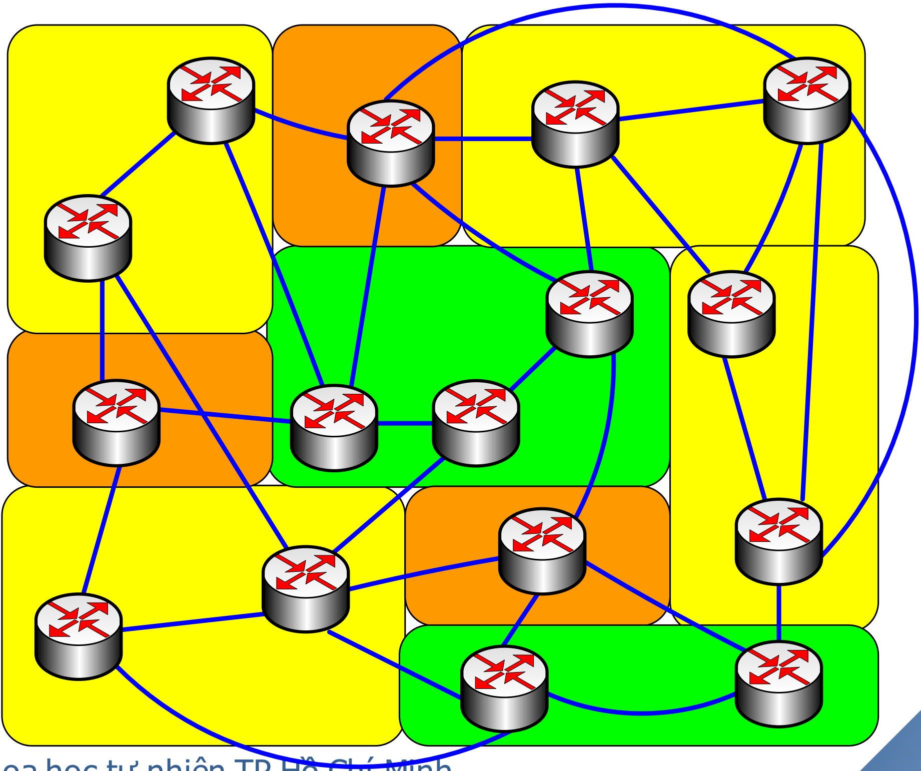


- ❑ Intradomain routing
  - Routing is done based on metrics
  - Routing domain is one autonomous system
- ❑ Interdomain routing
  - Routing is done based on policies
  - Routing domain is the entire Internet

Interdomain Routing

- ❑ Interdomain routing is based on connectivity between autonomous systems
- ❑ Interdomain routing can ignore many details of router interconnection

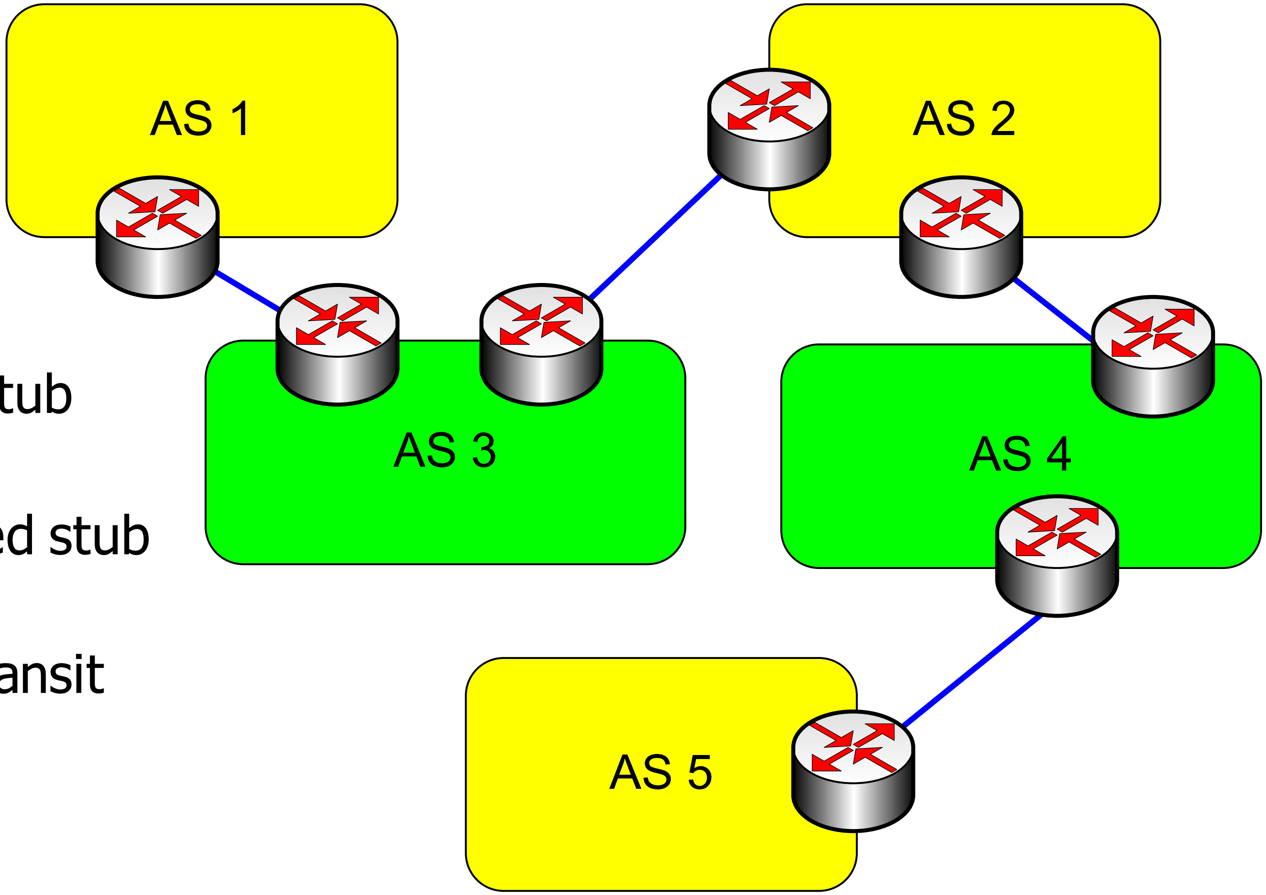




Autonomous Systems Terminology

- □ **local traffic** = traffic with source or destination in AS
- □ **transit traffic** = traffic that passes through the AS
- □ **Stub AS** = has connection to only one AS, only carry local traffic
- □ **Multihomed AS** = has connection to >1 AS, but does not carry transit traffic
- □ **Transit AS** = has connection to >1 AS and carries transit traffic

Stub and Transit Networks

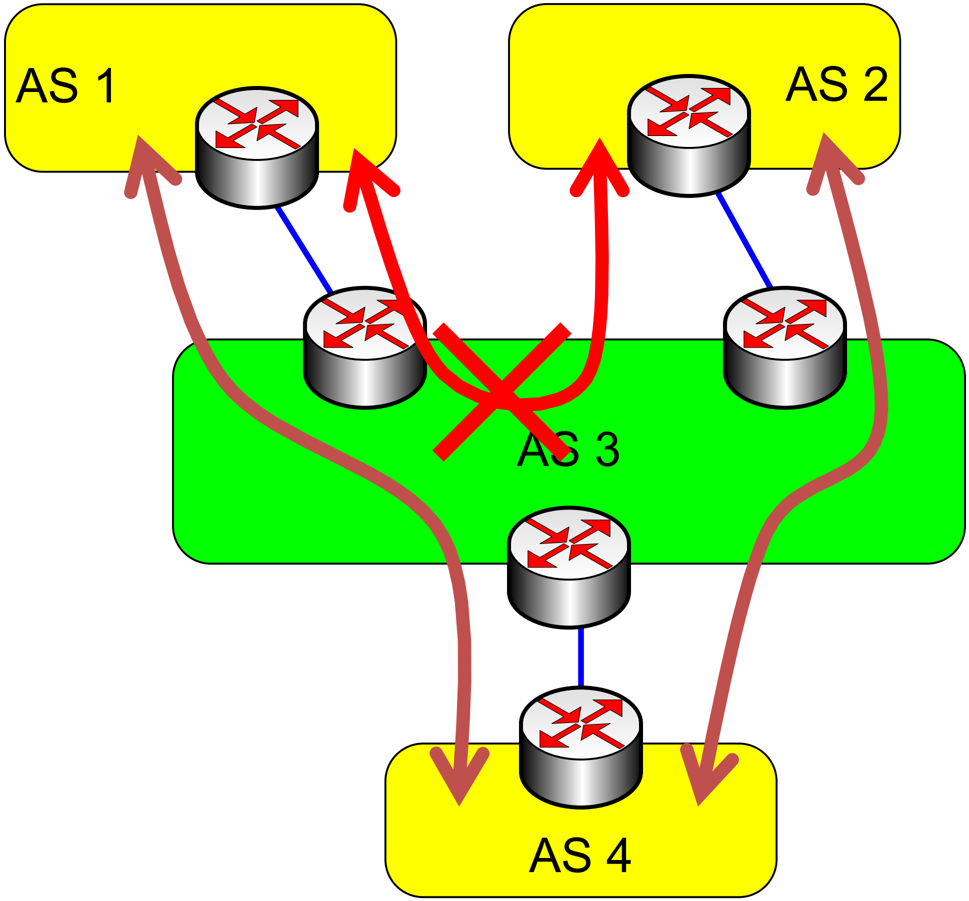


- ❑ AS 1, and AS 5 are stub networks
- ❑ AS 2 is a multi-homed stub network
- ❑ AS 3 and AS 4 are transit networks

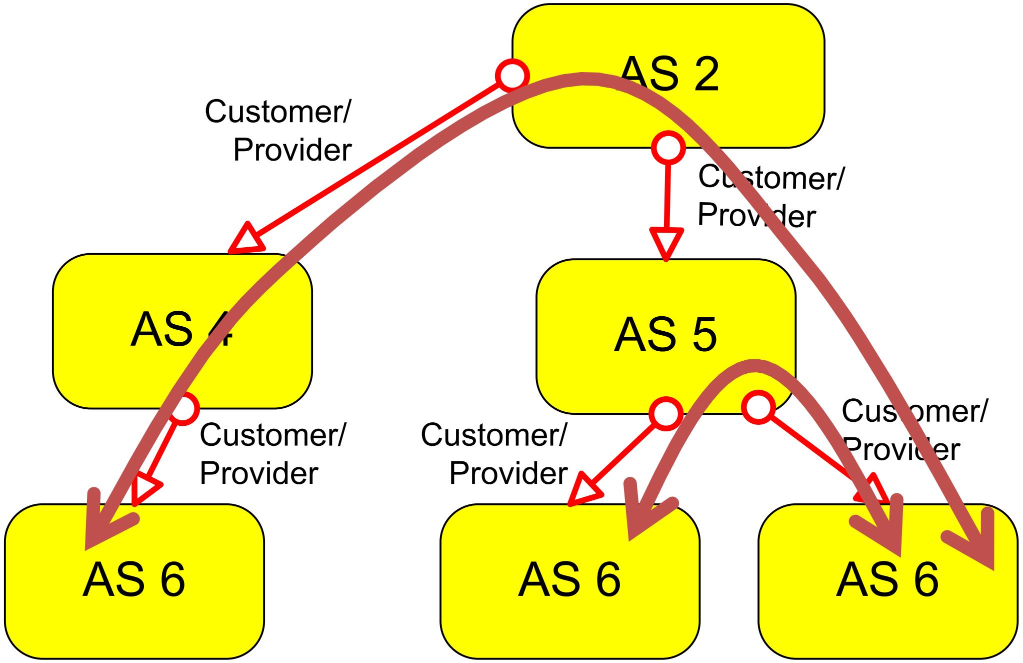
Selective Transit

Example:

- Transit AS 3 carries traffic between AS 1 and AS 4 and between AS 2 and AS 4
- But AS 3 does **not** carry traffic between AS 1 and AS 2
- The example shows a routing policy.



Customer/Provider

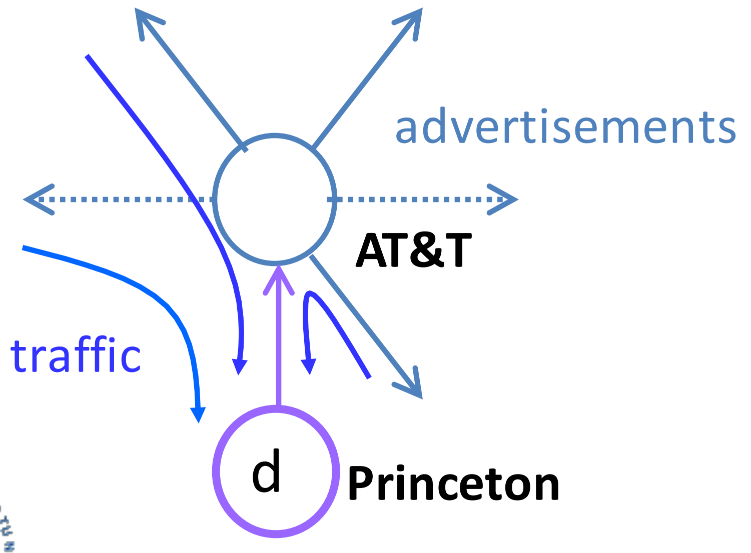


- ❑ A stub network typically obtains access to the Internet through a transit network.
- ❑ Transit network that is a provider may be a customer for another network
- ❑ Customer pays provider for service

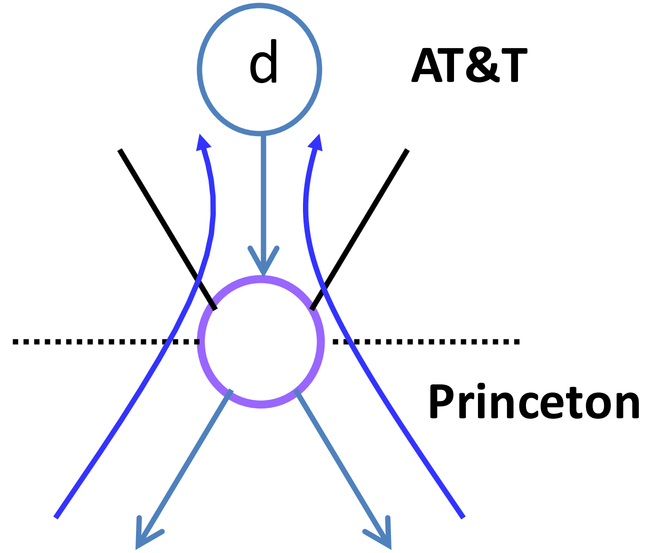
Customer-Provider Relationship

- ❑ Customer pays provider for access to Internet
  - Provider exports customer's routes to everybody
  - Customer exports provider's routes to customers

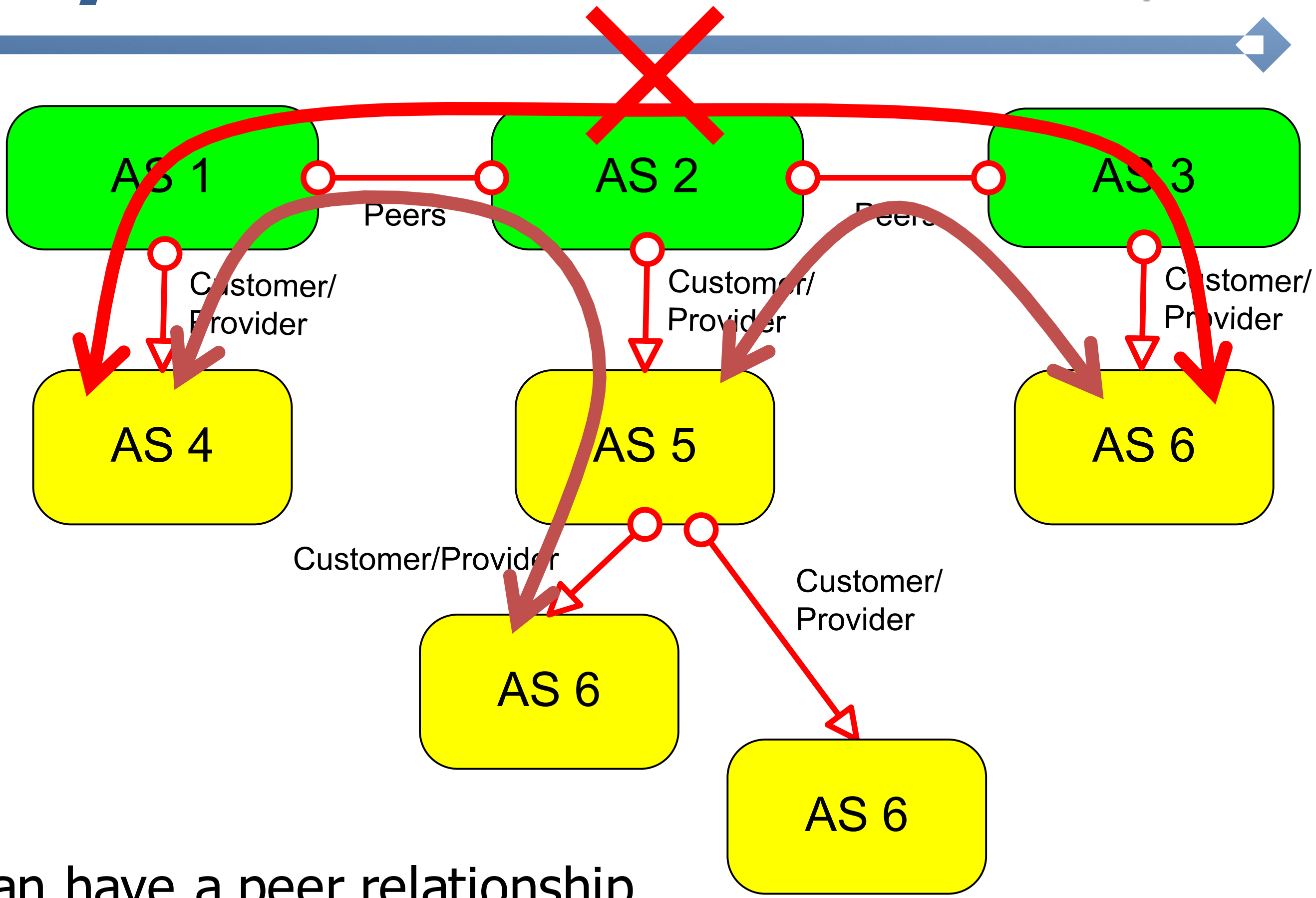
Traffic **to** the customer



Traffic **from** the customer

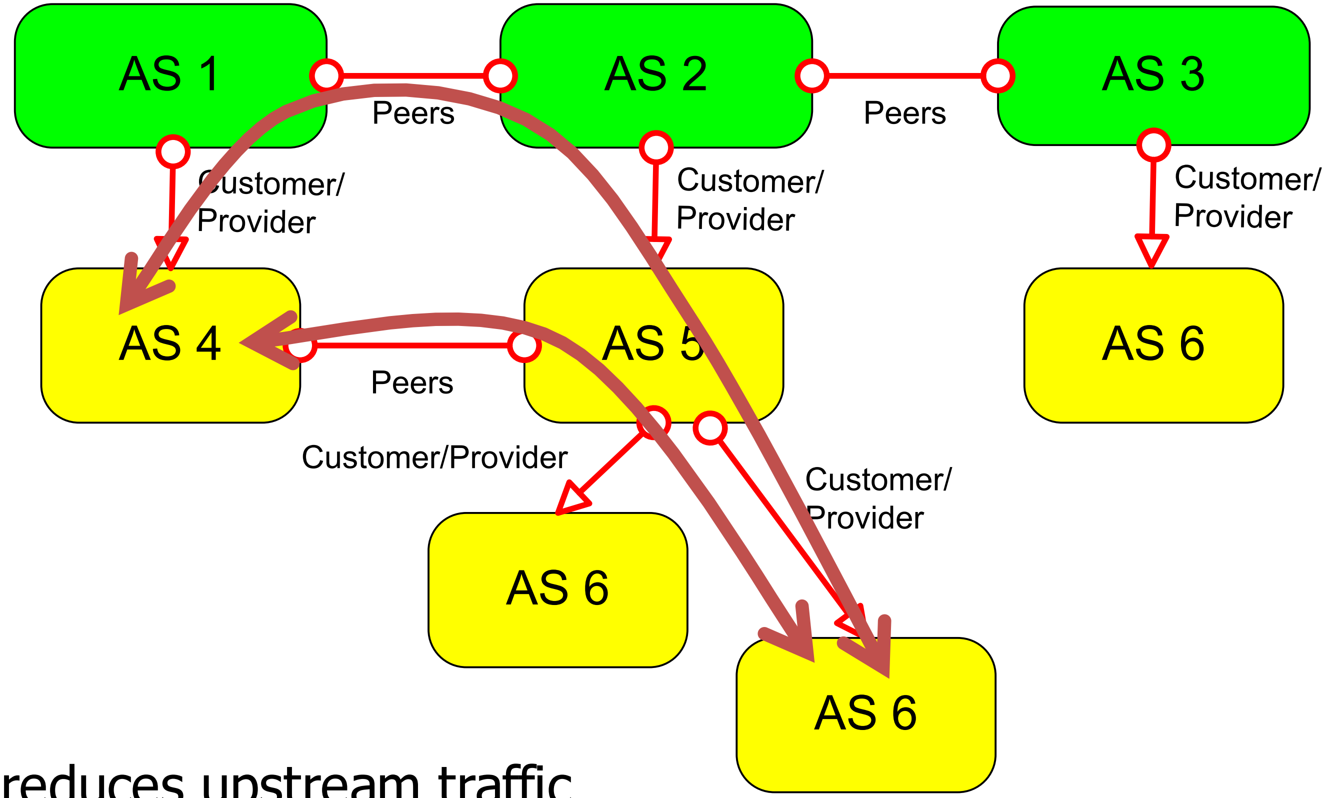


Customer/Provider and Peers



- ❑ Transit networks can have a peer relationship
- ❑ Peers provide transit between their respective customers
- ❑ Peers do not provide transit between peers
- ❑ Peers normally do not pay each other for service

Shortcuts through peering

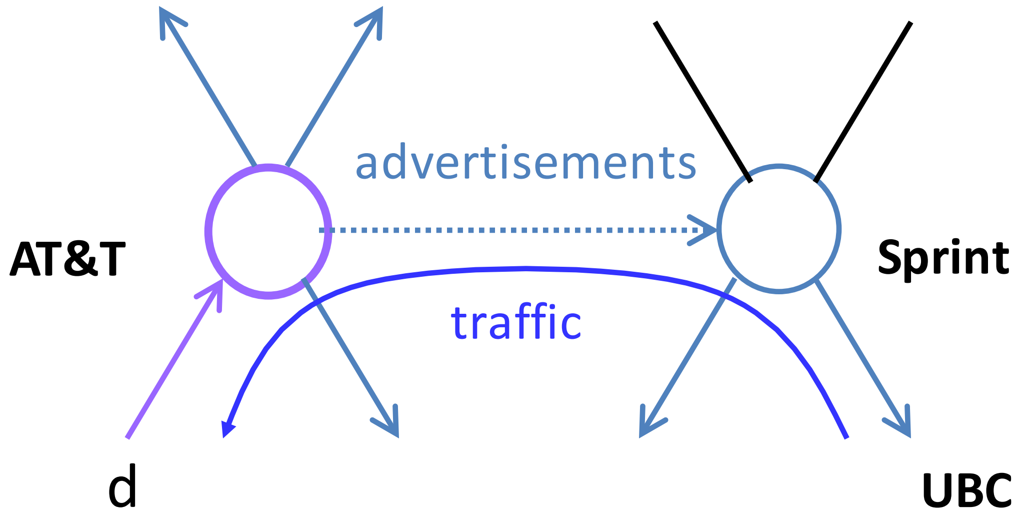


- ❑ Note that peering reduces upstream traffic
- ❑ Delays can be reduced through peering
- ❑ But: Peering may not generate revenue

Peer-Peer Relationship

- ❑ Peers exchange traffic between customers
  - AS exports *only* customer routes to a peer
  - AS exports a peer's routes *only* to its customers

Traffic to/from the peer and its customers



How Peering Decisions are Made?

Peer

❑ Reduces upstream transit costs

❑ Can increase end-to-end performance

❑ May be the only way to connect your customers to some part of the Internet (“Tier 1”)

Don’t Peer

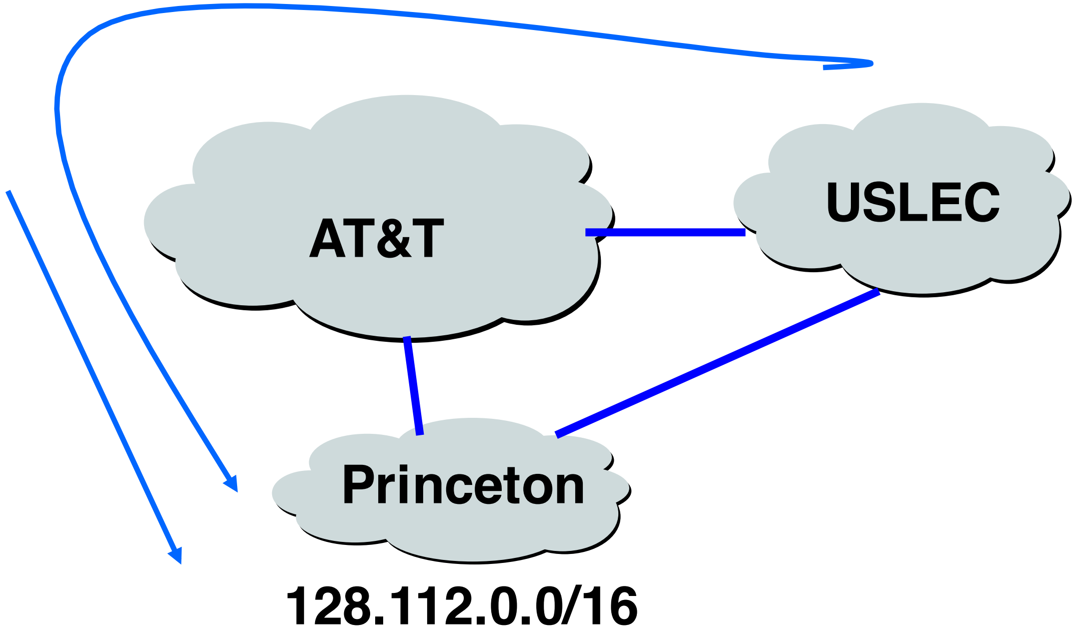
❑ You would rather have customers

❑ Peers are usually your competition

❑ Peering relationships may require periodic renegotiation

Backup Relationship

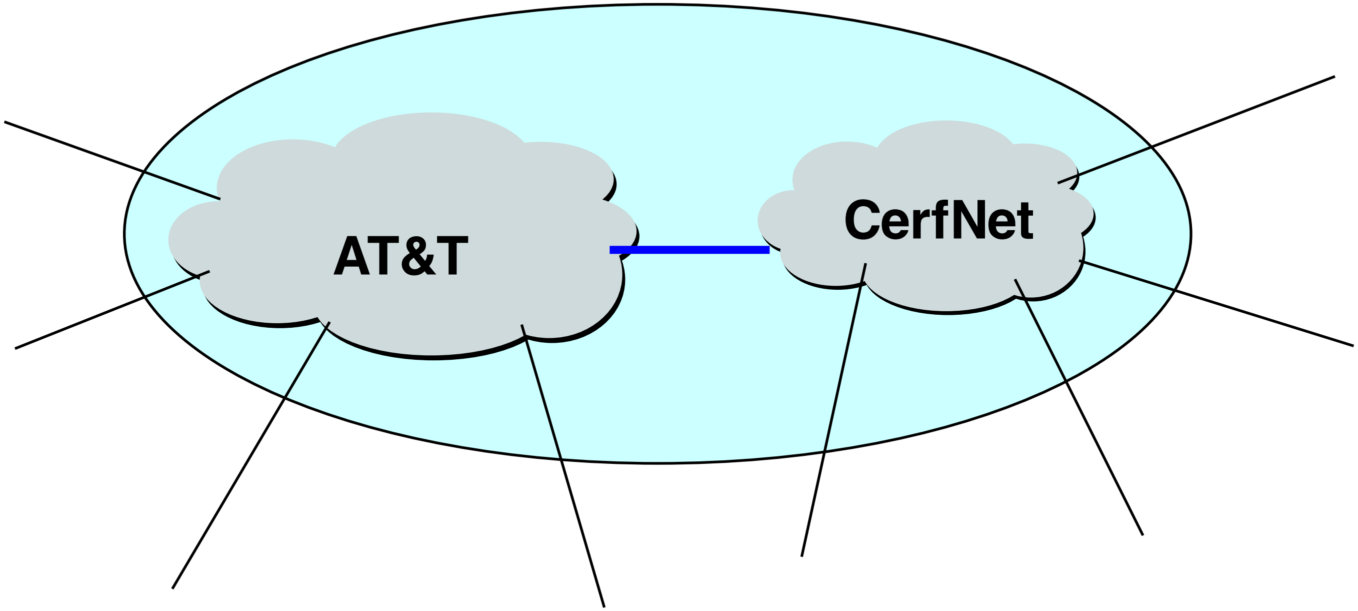
- ❑ Backup provider
  - Only used if the primary link fails
  - Routes through other paths



Sibling Relationship

❑ Two ASes owned by the same institution

- E.g., two ASes that have merged
- E.g., two ASes simply for scaling reasons
- Essentially act as a single AS



Border Gateway Protocol (BGP)

- ❑ Border Gateway Protocol is the interdomain routing protocol for the Internet for routing between autonomous systems
- ❑ Currently in version 4 (1995)
  - Network administrators can specify routing policies
  - BGP is a path vector protocol
- ❑ Uses TCP to transmit routing messages

Border Gateway Protocol (BGP)

❑ An autonomous system uses BGP to advertise its network address(es) to other AS's

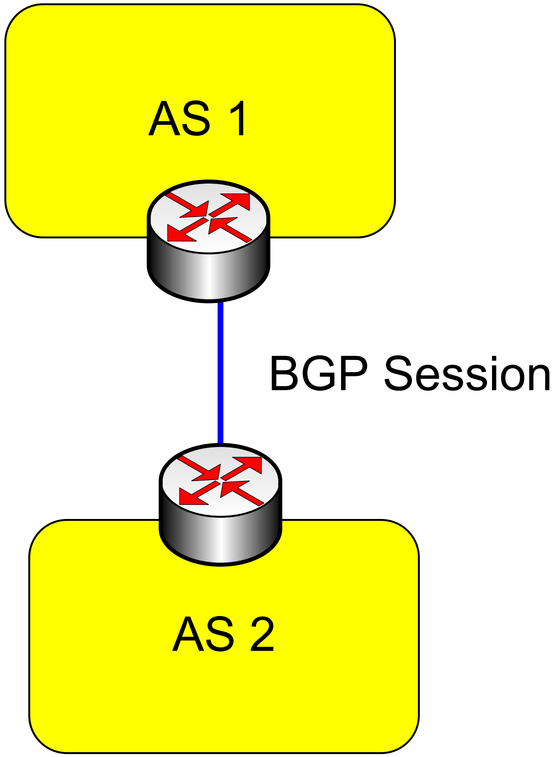
❑ BGP helps an autonomous system with the following:

1. Collect information about reachable networks from neighboring AS's
2. Disseminate the information about reachable networks to routers inside the AS and to neighboring AS's
3. Picks routes if there are multiple routes available

BGP interactions

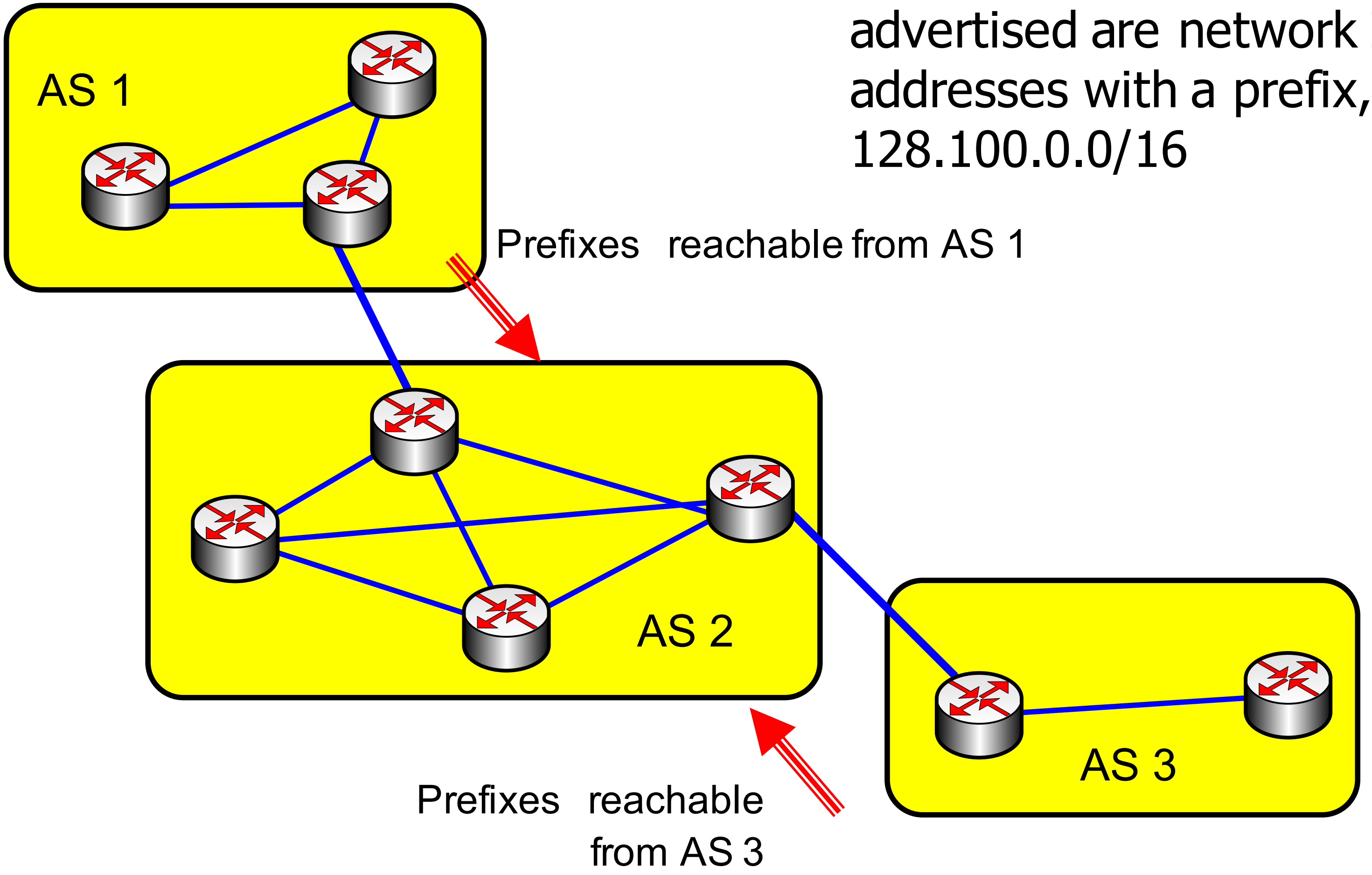
- ❑ Router establishes a TCP connection (TCP port 175)
- ❑ Routers exchange BGP routes
- ❑ Periodically send updates
- ❑ BGP is executed between two routers
  - BGP session
  - BGP peers or BGP speakers

- ❑ **Note:** Not all autonomous systems need to run BGP. On many stub networks, the route to the provider can be statically configured



BGP interactions

- The networks that are advertised are network IP addresses with a prefix, E.g., 128.100.0.0/16

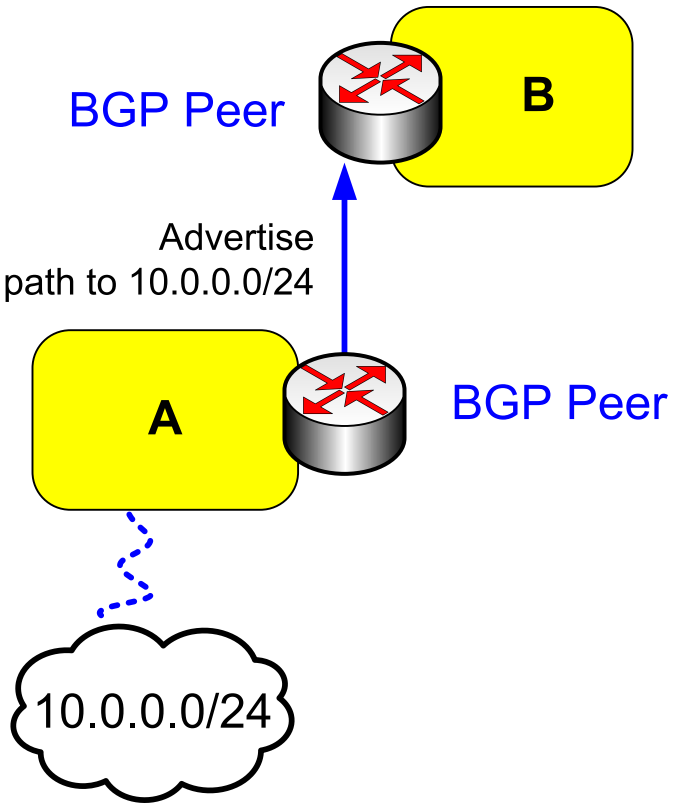


BGP interactions

- ❑ BGP peers advertise reachability of IP networks

- ❑ A advertises a path to a network (e.g., 10.0.0.0/8) to B only if it is willing to forward traffic going to that network

- ❑ **Path-Vector:**
  - ▪ A advertises the complete path to the advertised network
  - ▪ Path is sent as a list of AS's

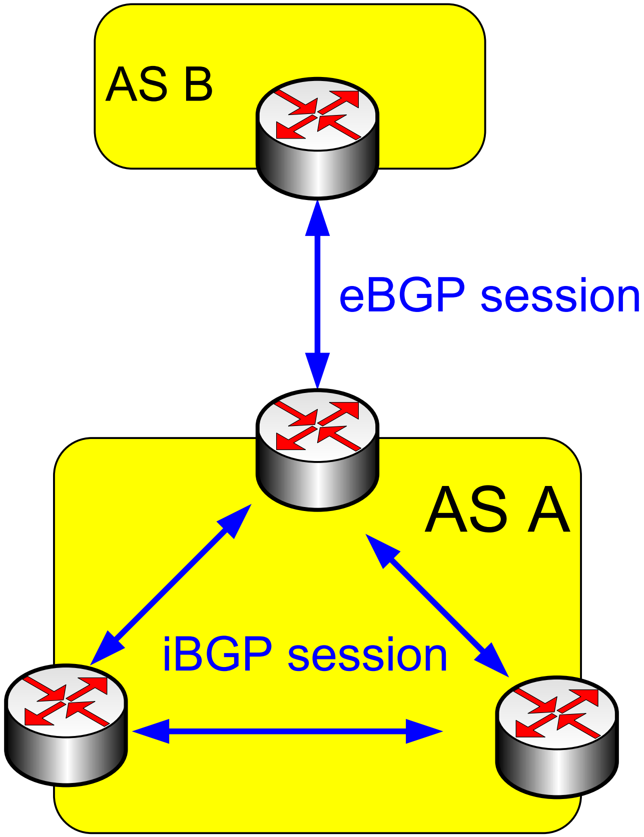


BGP Sessions

- □ **External BGP session (eBGP):**
  Peers are in different AS'es

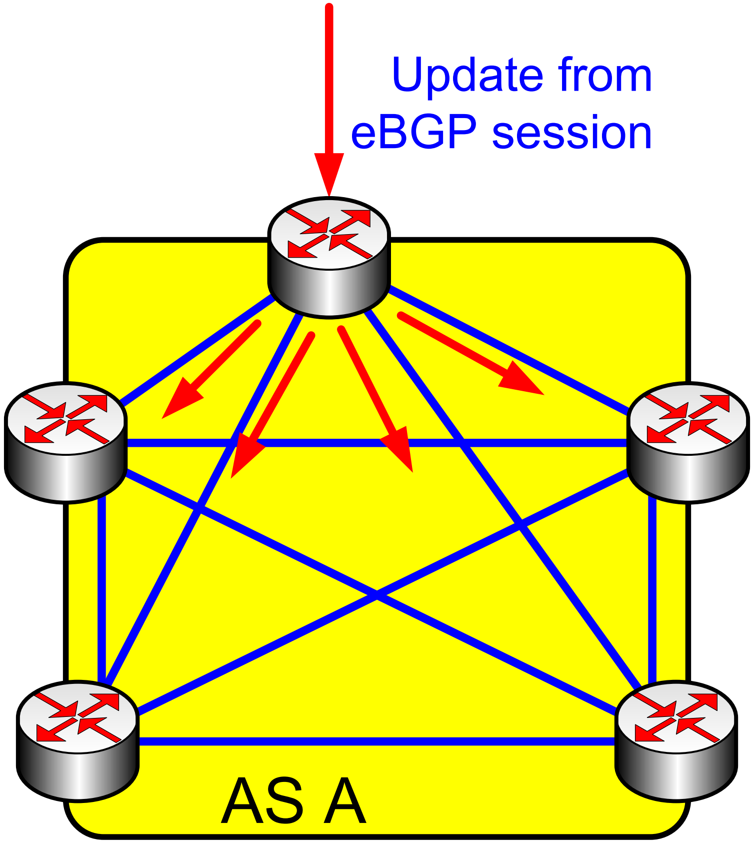
- □ **Internal BGP session (iBGP)**
  Peers are in same AS

- □ Note that iBGP sessions are going over routes that are set up by an intradomain routing protocol!



iBGP sessions

- ❑ All iBGP peers in the same autonomous system are fully meshed
- ❑ Peer announces routes received via eBGP to iBGP peers
- ❑ **But:** iBGP peers do not announce routes received via iBGP to other iBGP peers



BGP Message Types

□ **Open**: Establishes a peering session

□ **Keep Alive**: Handshake at regular intervals to maintain peering session

□ **Notification**: Closes a peering session

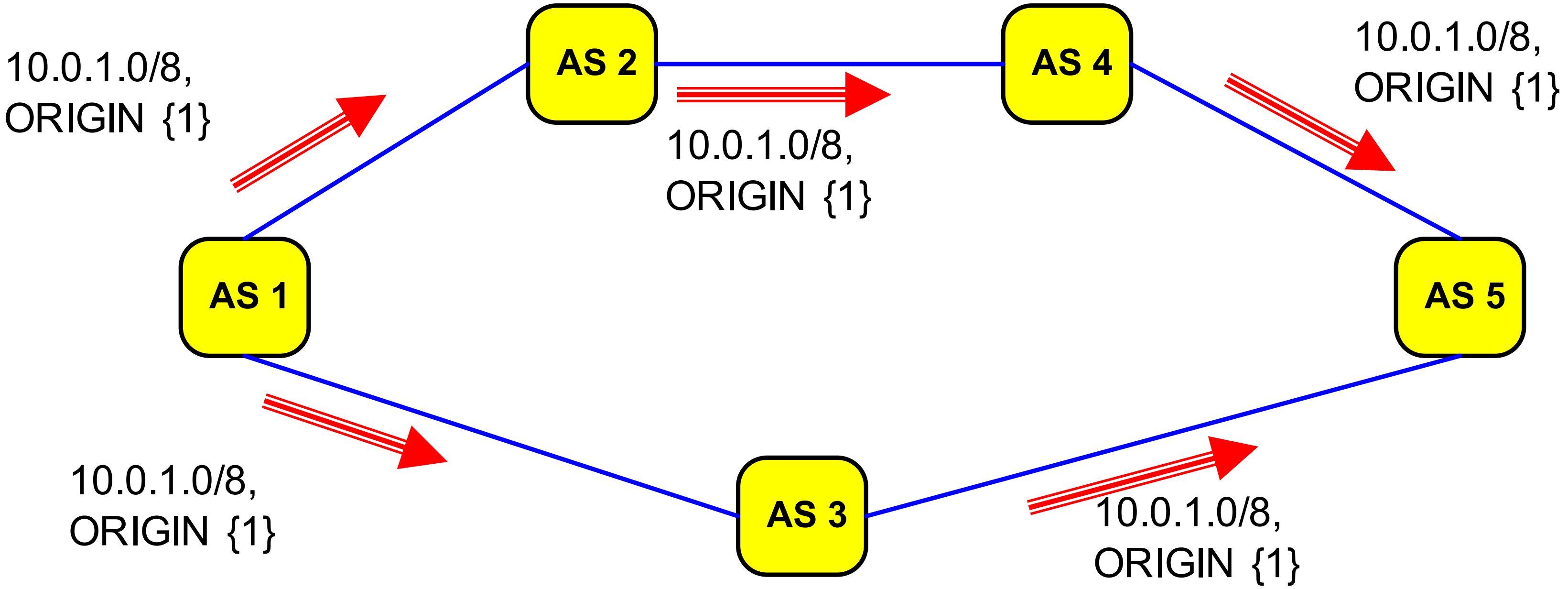
□ **Update**: Advertises new routes or withdraws previously announced routes. Each announced route is specified as a network prefix with attribute values

Content of Advertisements

- ❑ BGP routers advertise **routes**
- ❑ Each route consists of a network prefix and a list of **attributes** that specify information about a route
- ❑ Mandatory attributes:
    - **ORIGIN**
    - **AS_PATH**
    - **NEXT_HOP**
- ❑ Many other attributes
    - **LOCAL_PREF**
    - **MULTI_EXIT_DISC**
    - **.....**

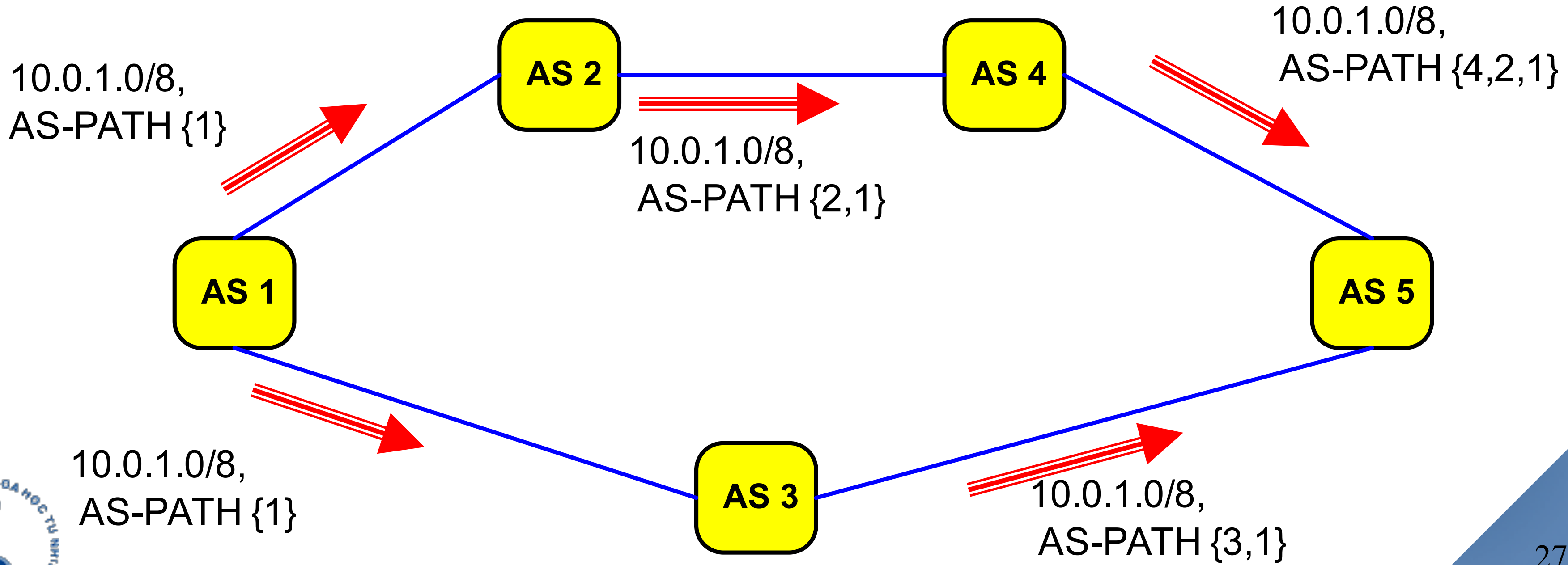
ORIGIN attribute

- ❑ Originating domain sends a route with ORIGIN attribute
- ❑ Three values: igp, egp, incomplete



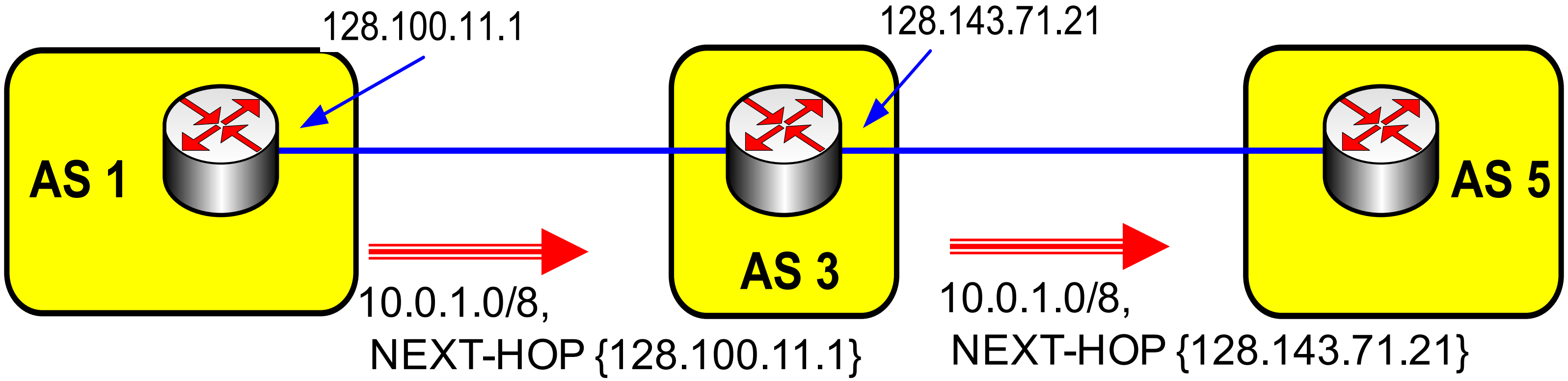
AS-PATH attributes

- Each AS that propagates a route prepends its own AS number
  - AS-PATH collects a path to reach the network prefix
- Path information prevents routing loops from occurring
- Path information also provides information on the length of a path (By default, a shorter route is preferred)

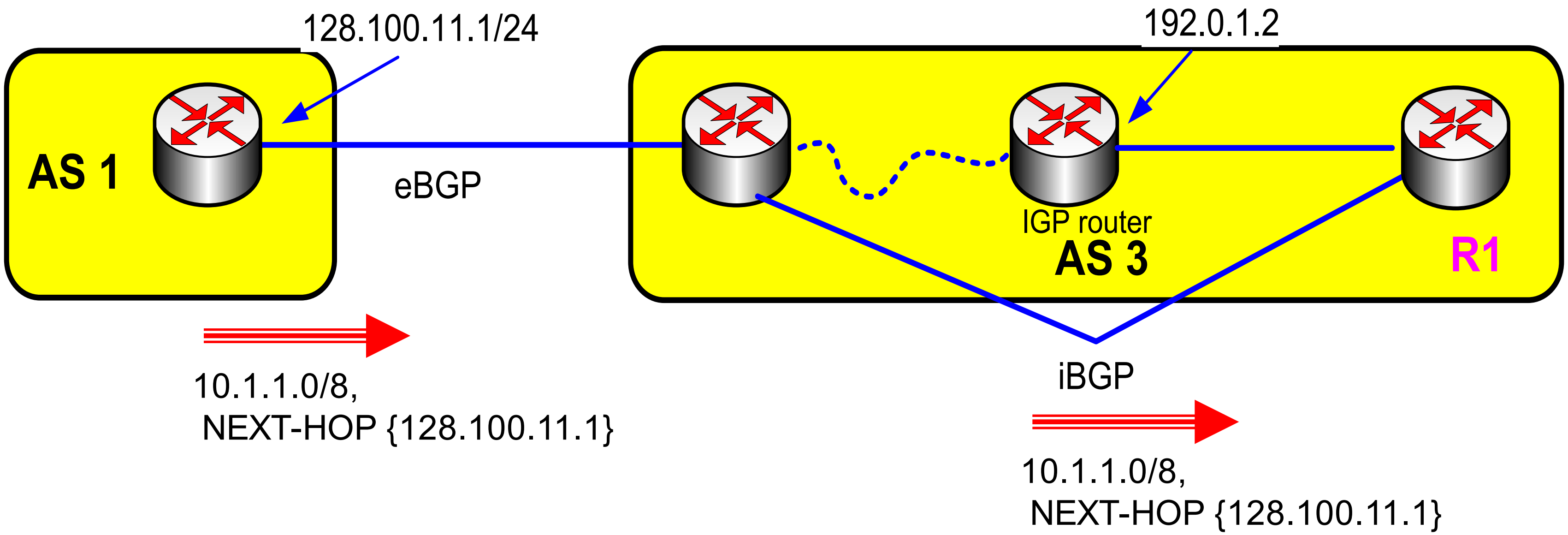


NEXT-HOP attributes

- Each router that sends a route advertisement it includes its own IP address in a NEXT-HOP attribute
- The attribute provides information for the routing table of the receiving router.



Connecting NEXT-HOP with IGP information



10.1.1.0/8,
NEXT-HOP {128.100.11.1}

10.1.1.0/8,
NEXT-HOP {128.100.11.1}

At R1:

Routing table

Dest.	Next hop
128.100.11.0/24	192.0.1.2

BGP info

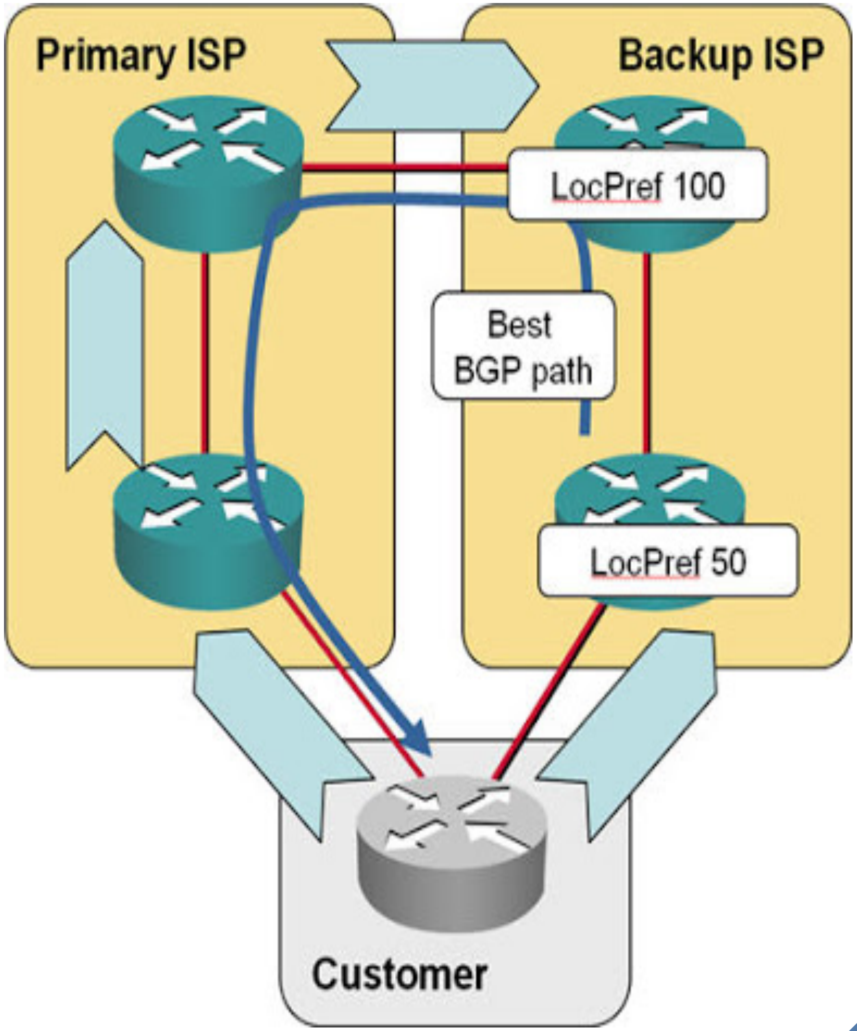
Dest.	Next hop
10.1.1.0/8	128.100.11.1

Routing table

Dest.	Next hop
128.100.11.0/24	192.0.1.2
10.1.1.0/8	192.0.1.2

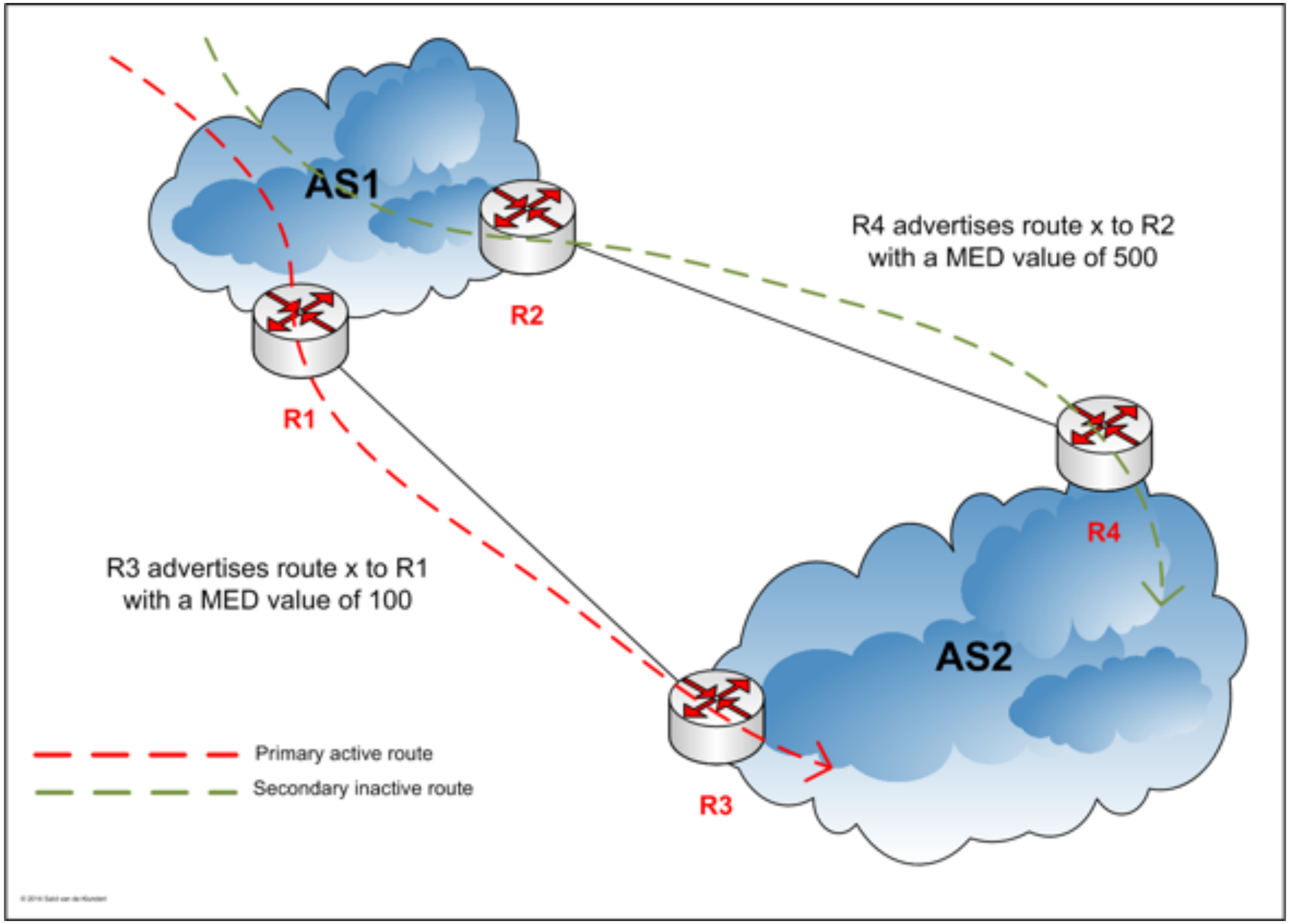
LOCAL_PREF attribute

- ❑ Local Preference



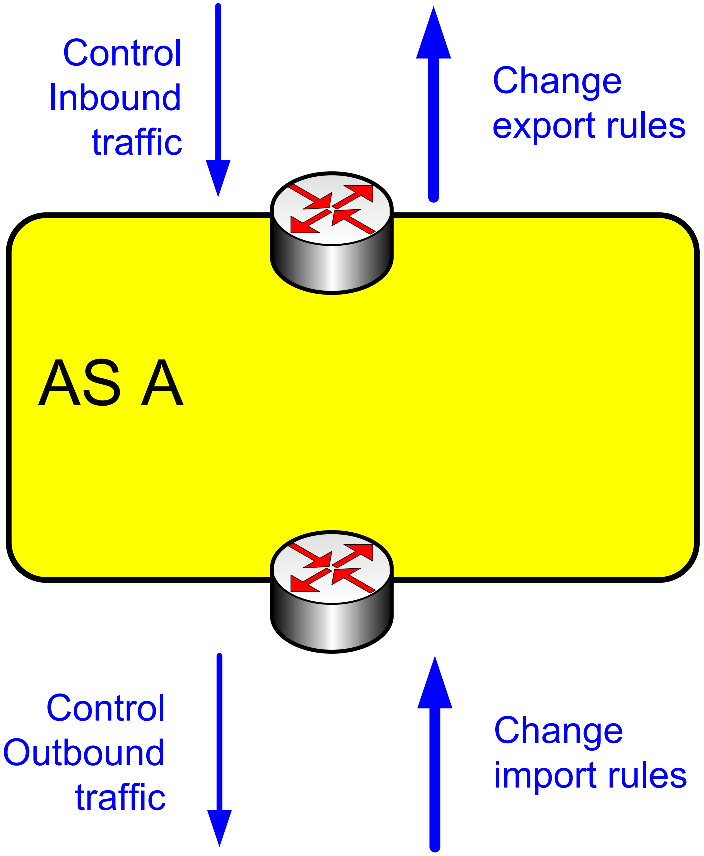
MED Attribute

- ❑ Multiple exit discriminator (MED)



Importing and Exporting Routes

- ❑ An AS may not accept all routes that are advertised
- ❑ An AS may not advertise certain routes
- ❑ Route policies determines which routes are filtered
- ❑ If an AS wants to have less inbound traffic it should adapt its export rules
- ❑ If an AS wants to control its inbound traffic, it adapts its import rules



Path Selection on a Router

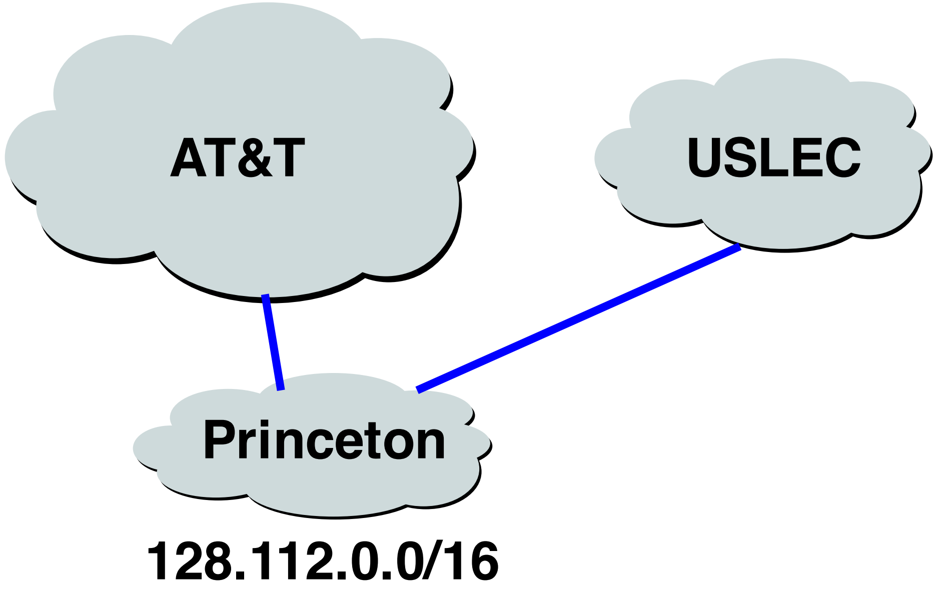
- ❑ Routing Information Base
  - ▪ Store all BGP routes for each destination prefix
  - ▪ Withdrawal message: remove the route entry
  - ▪ Announcement message: update the route entry
- ❑ Selecting the best route
  - ▪ Consider all BGP routes for the prefix
  - ▪ Apply rules for comparing the routes
  - ▪ Select the one best route
    - Use this route in the forwarding table
    - Send this route to neighbors

BGP Decision Process: Multiple Steps

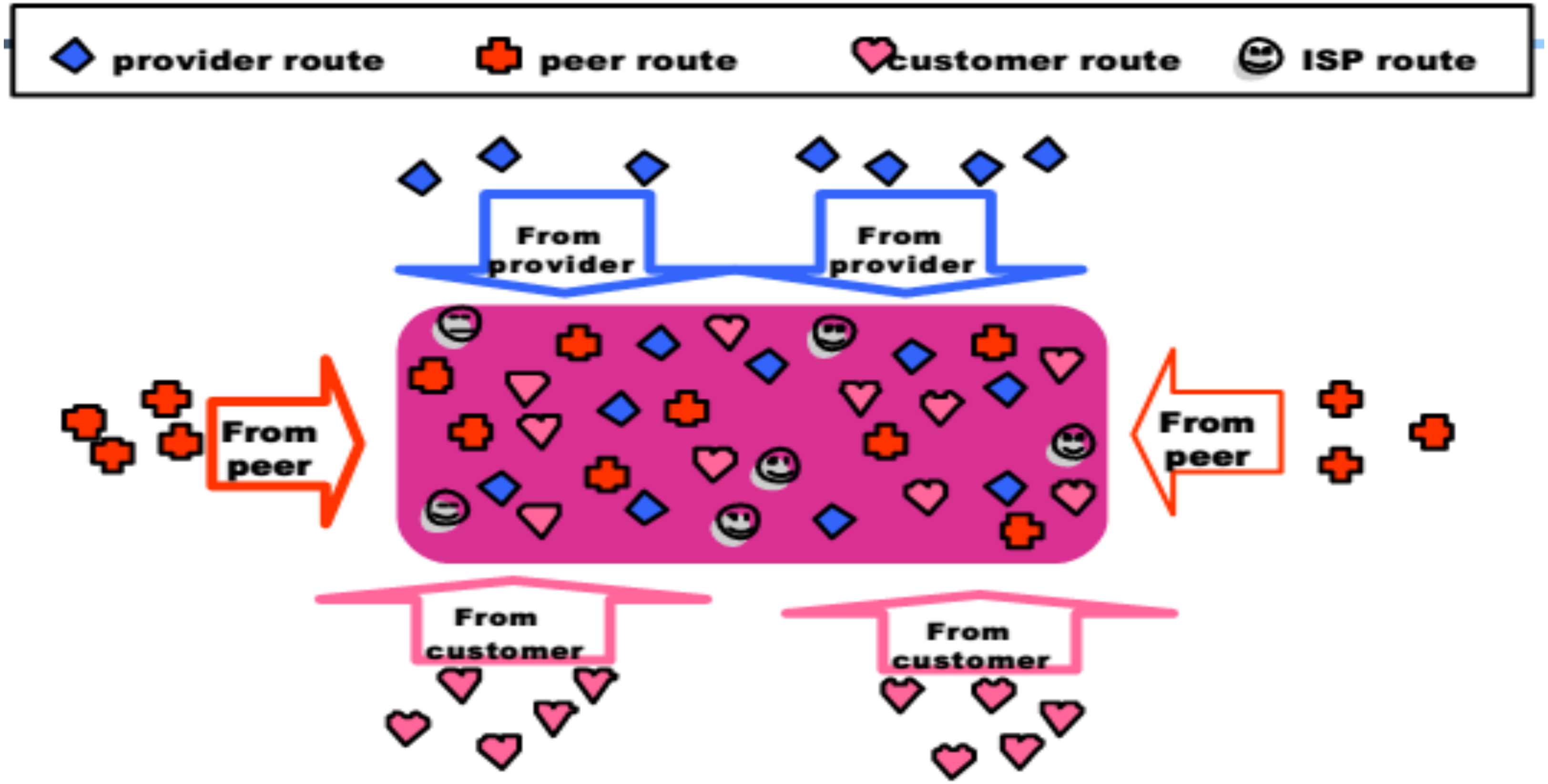
- ❑ Highest local preference
  - *Set by import policies* upon receiving advertisement
- ❑ Shortest AS path
  - Included in the route advertisement
- ❑ Lowest origin type
  - IGP < EGP < incomplete
- ❑ Smallest multiple exit discriminator (MED)
  - Included in the advertisement or reset by import policy
- ❑ Smallest internal path cost to the next hop
  - Based on intradomain routing protocol (e.g., OSPF)
- ❑ Smallest next-hop router id
  - Final tie-break

Import Policy: Filtering

- ❑ Discard some route announcements
  - Detect configuration mistakes and attacks
- ❑ Examples on session to a customer
  - Discard route if prefix not owned by the customer
  - Discard route with other large ISP in the AS path

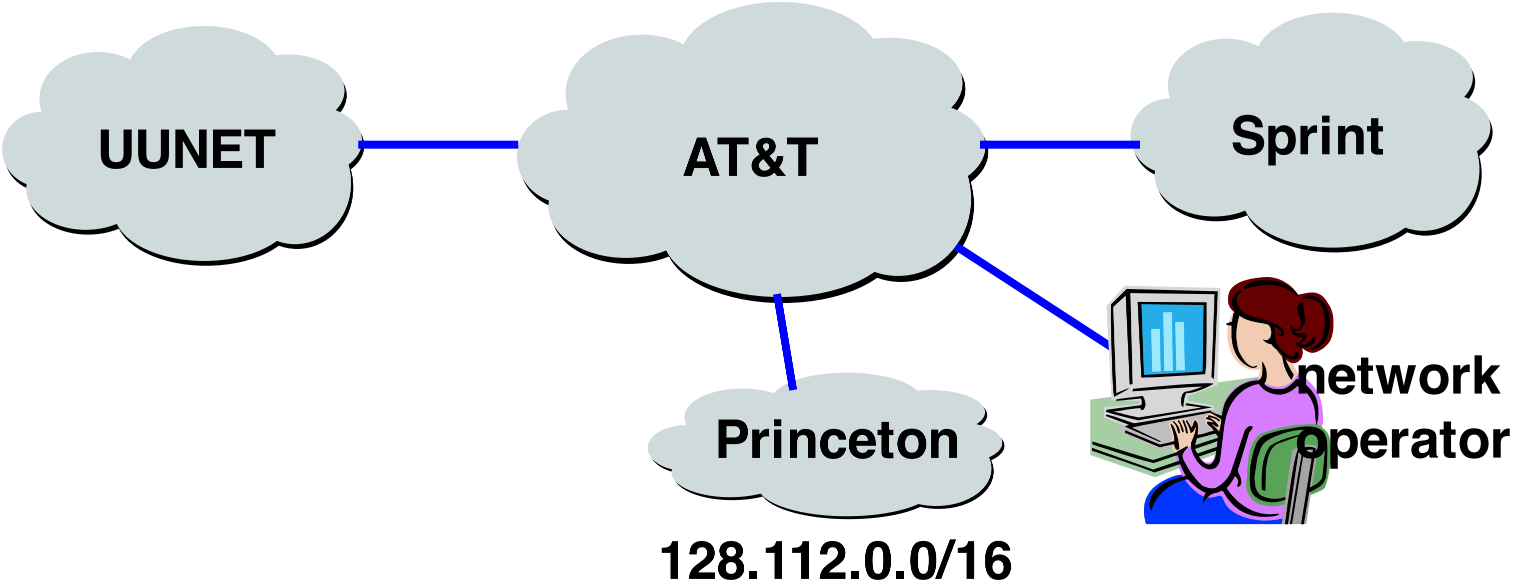


Import Rules

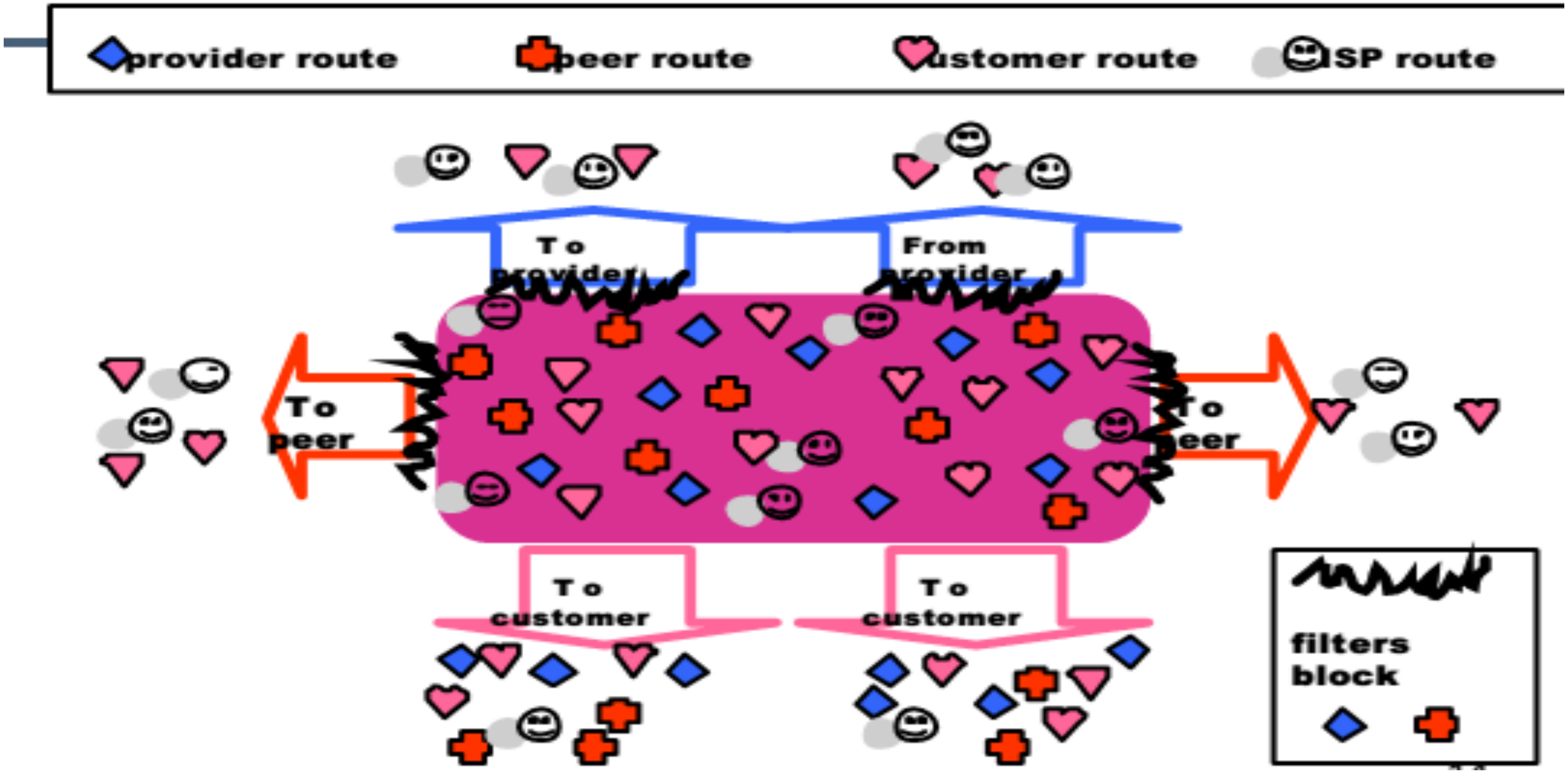


Export Policy: Filtering

- ❑ Discard some route announcements
  - Limit propagation of routing information
- ❑ Examples
  - Don’t announce routes from one peer to another
  - Don’t announce routes for management hosts



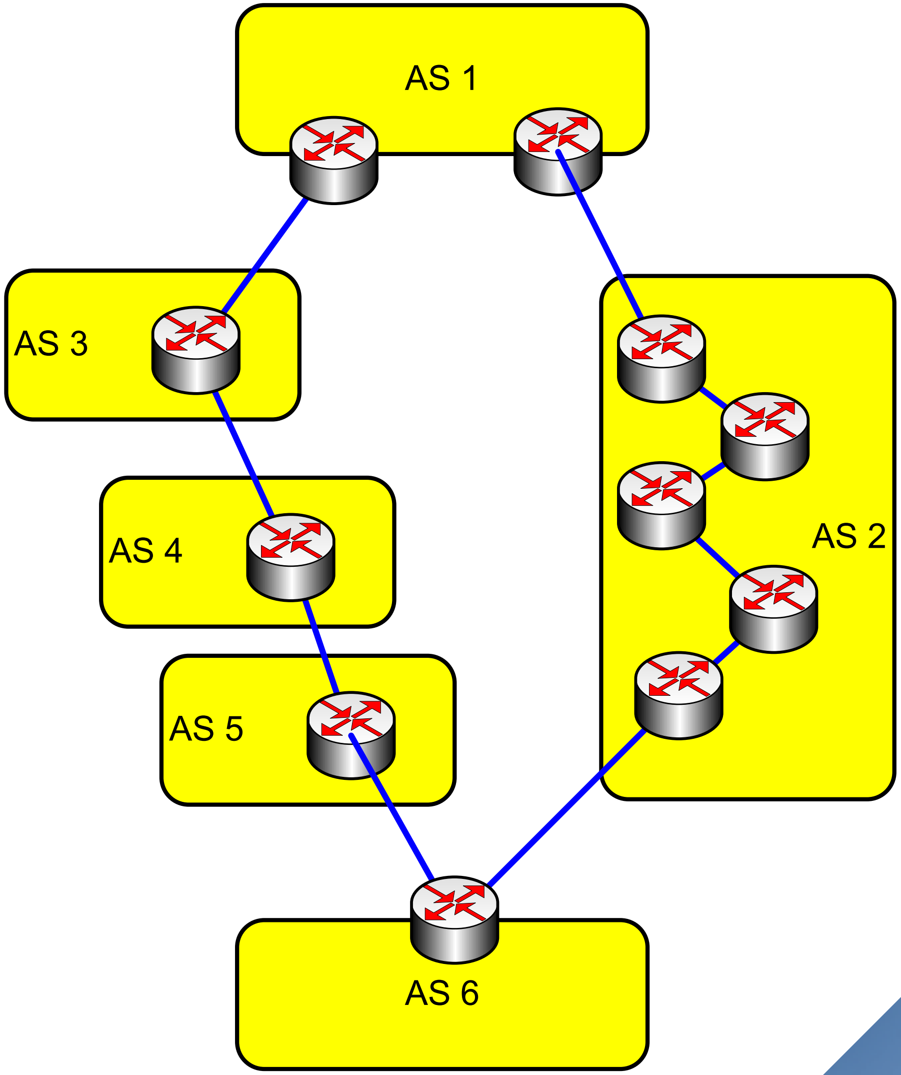
Export Rules



Short AS-PATH does not mean that route is short

- ❑ From AS 6's perspective
  - Path {AS2, AS1} is short
  - Path {AS5, AS4, AS3, AS1} is long

- ❑ But the number of traversed routers is larger when using the shorter AS-PATH

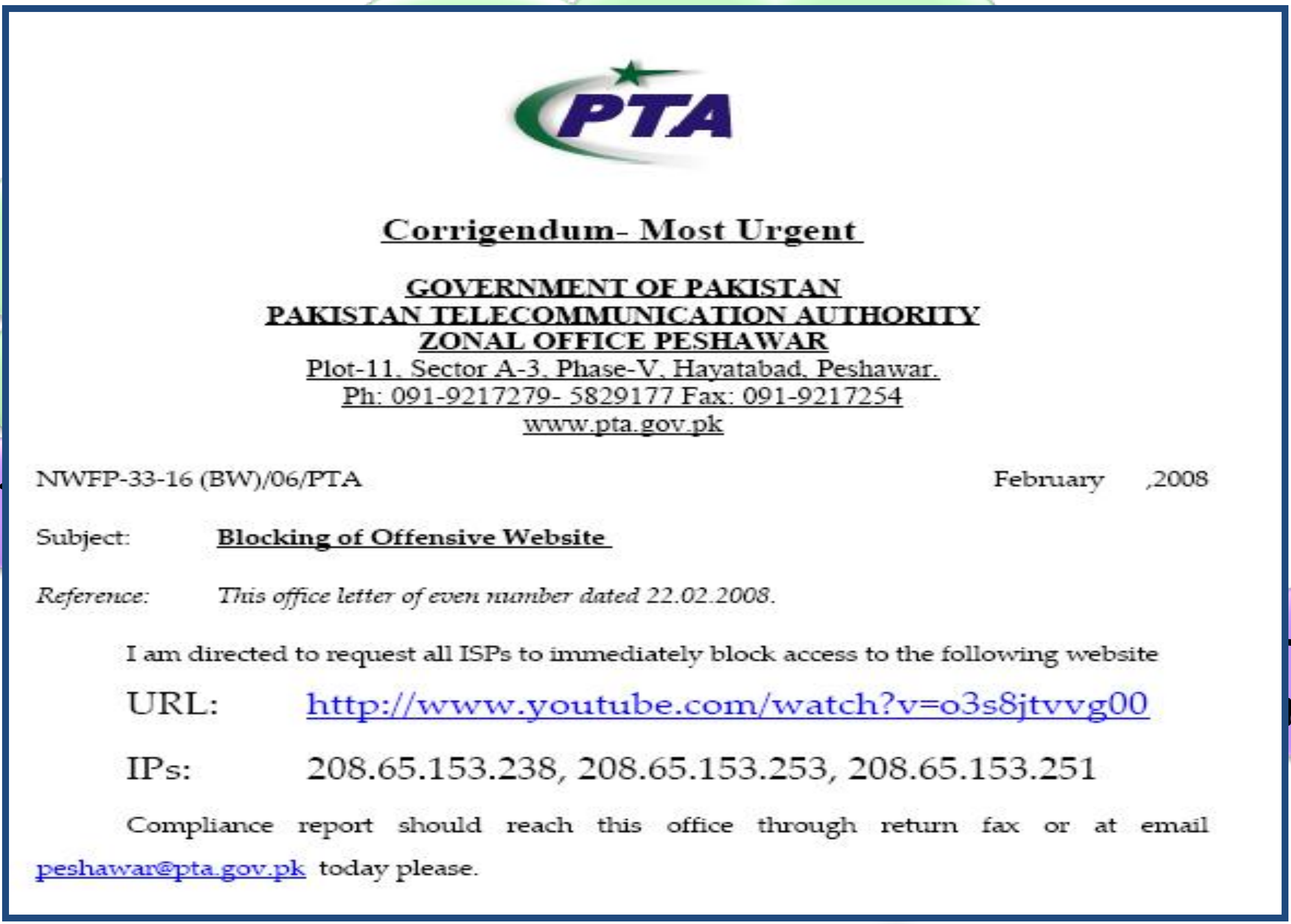


BGP Issues

- ❑ BGP is a simple protocol but it is very difficult to configure
- ❑ BGP has severe stability issue due to policies → BGP is known to not converge
- ❑ As of July 2005, 39,000 AS numbers (of available 64,510) are consumed

How Secure is Today's Internet Routing?

February 2008: Pakistan Telecom hijacks YouTube!



Corrigendum- Most Urgent

GOVERNMENT OF PAKISTAN
PAKISTAN TELECOMMUNICATION AUTHORITY
ZONAL OFFICE PESHAWAR
Plot-11, Sector A-3, Phase-V, Hayatabad, Peshawar.
Ph: 091-9217279- 5829177 Fax: 091-9217254
www.pta.gov.pk

NWFP-33-16 (BW)/06/PTA February ,2008

Subject: **Blocking of Offensive Website**

Reference: This office letter of even number dated 22.02.2008.

I am directed to request all ISPs to immediately block access to the following website

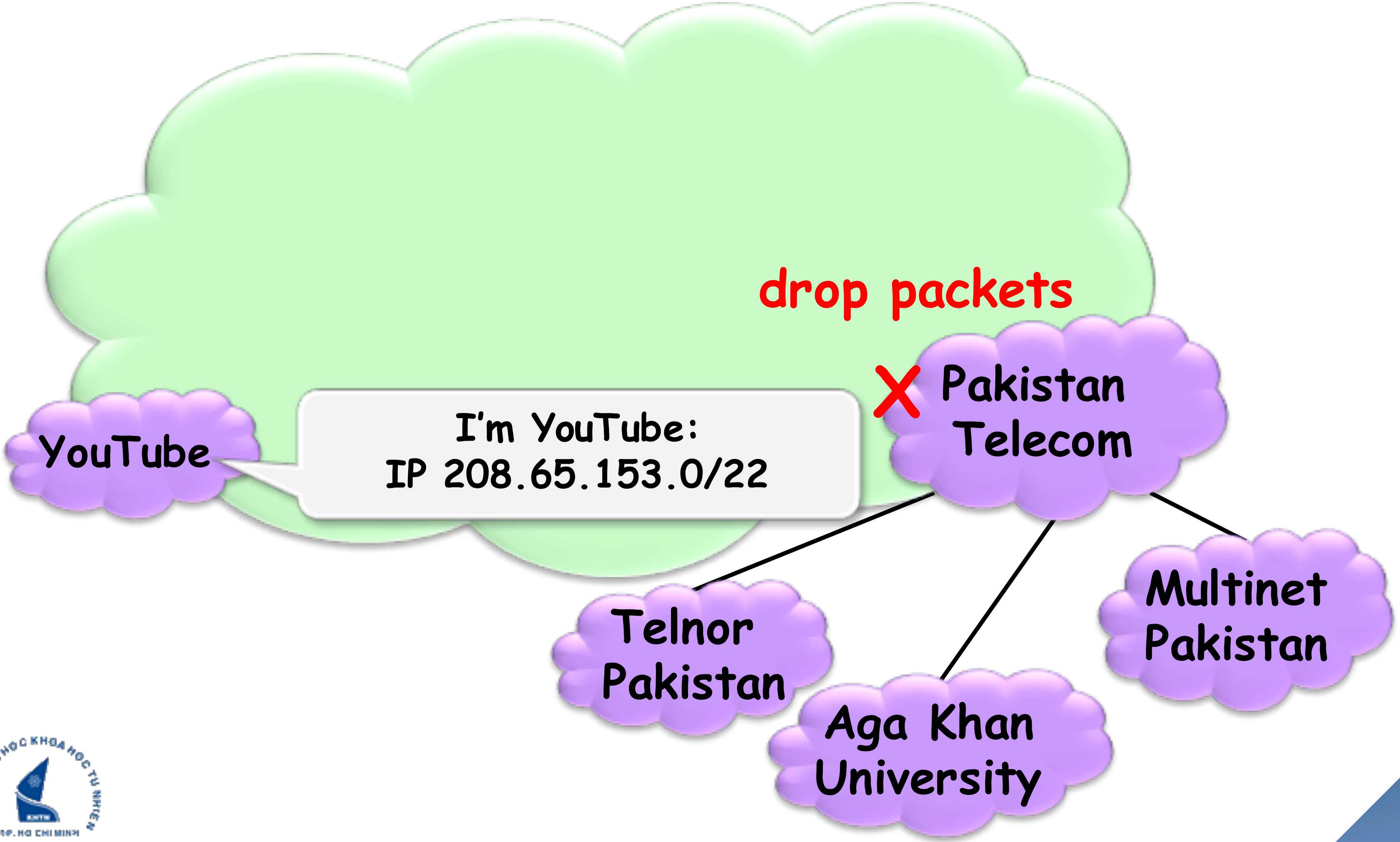
URL: http://www.youtube.com/watch?v=o3s8jtvvg00

IPs: 208.65.153.238, 208.65.153.253, 208.65.153.251

Compliance report should reach this office through return fax or at email peshawar@pta.gov.pk today please.

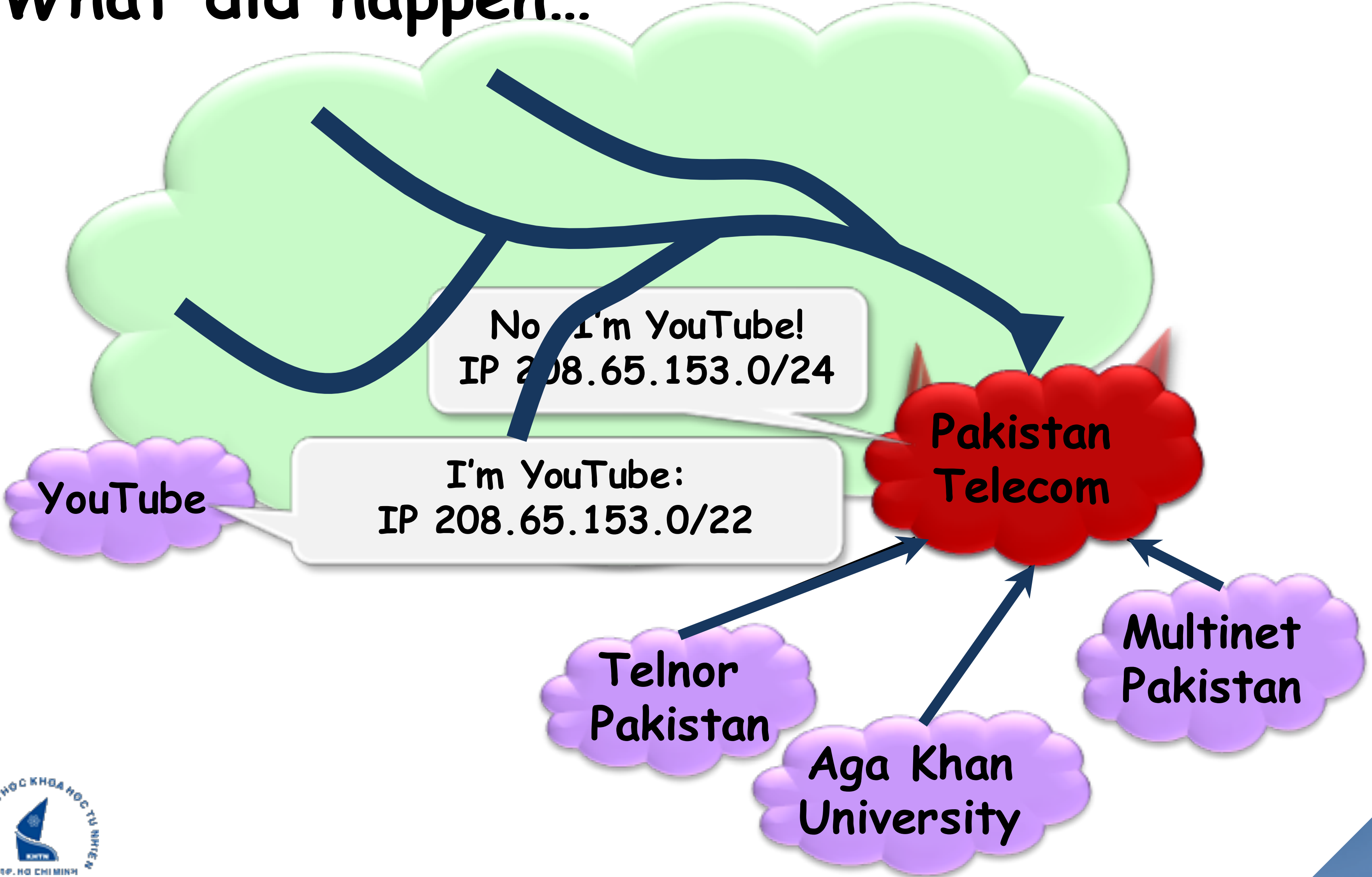
How Secure is Today's Internet Routing?

What should have happened…

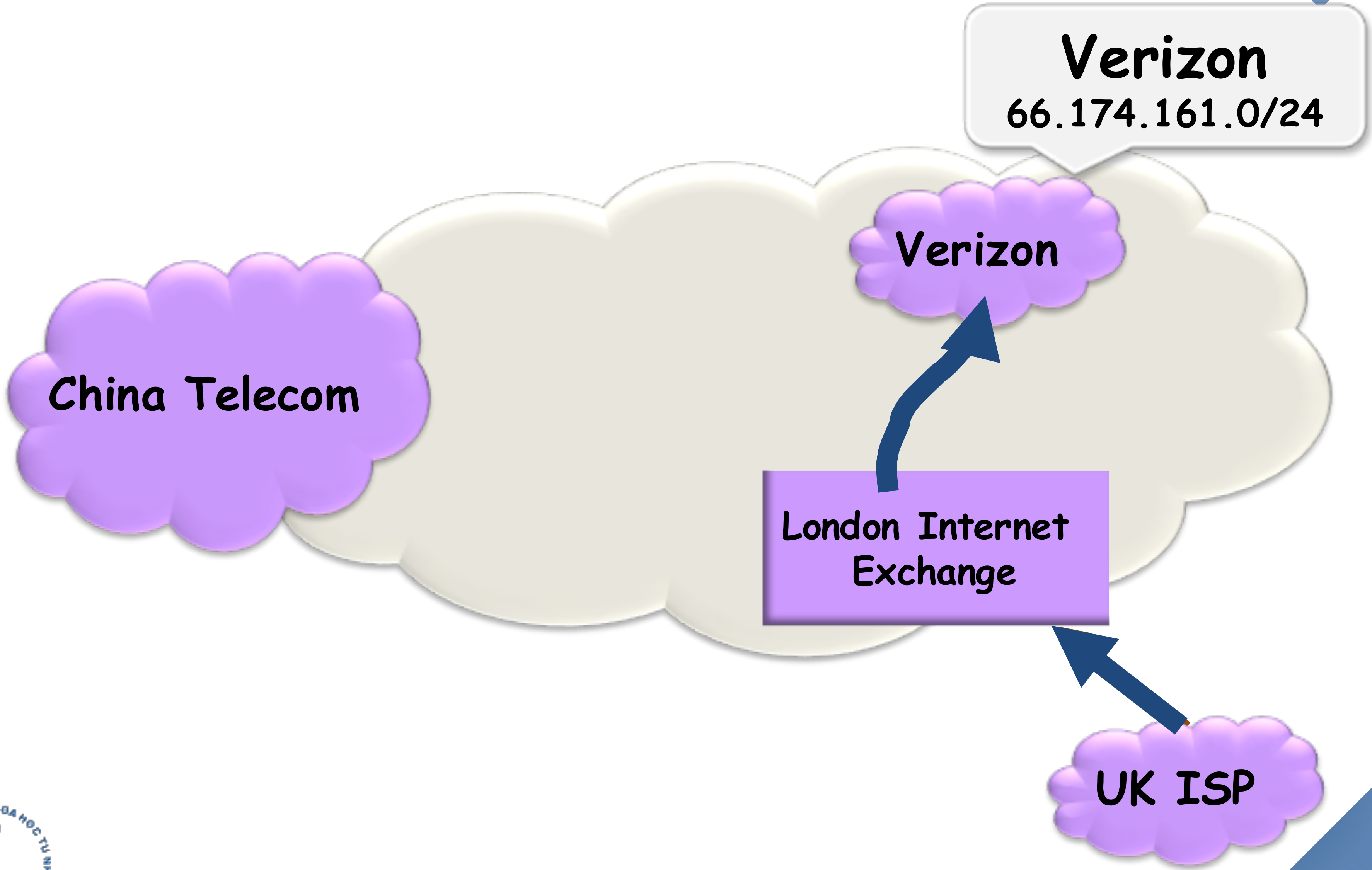


How Secure is Today's Internet Routing?

What did happen…

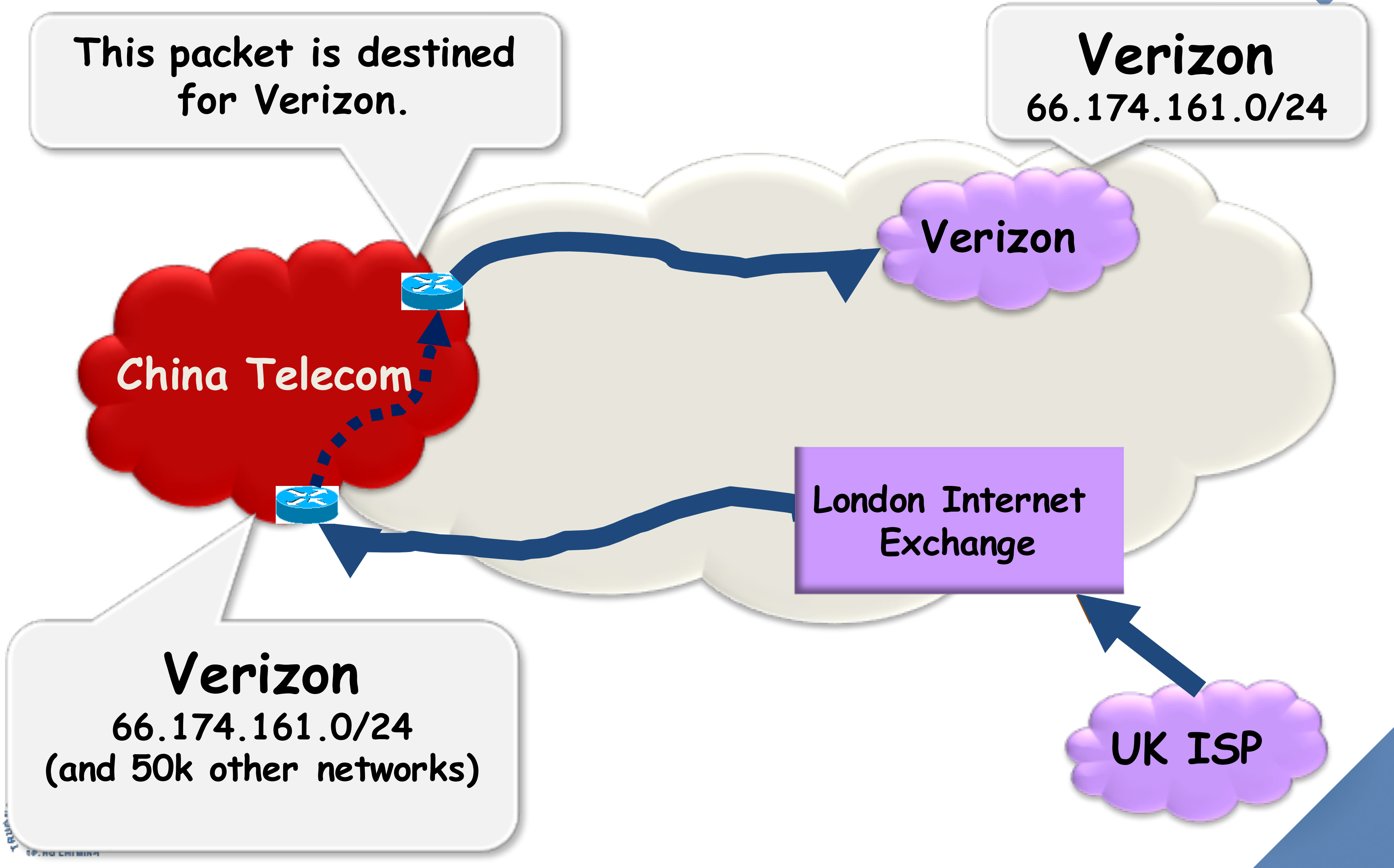


How Secure is Routing on the Internet Today? (2)



How Secure is Routing on the Internet Today? (2)

April 2010 : China Telecom intercepts traffic



This packet is destined for Verizon.

Verizon
66.174.161.0/24

Verizon

China Telecom

London Internet Exchange

Verizon
66.174.161.0/24
(and 50k other networks)

UK ISP

BGP Security Today

- Applying best common practices (BCPs)
  - Filtering routes by prefix and AS path, *etc.*

- This is not good enough!
  - Depends on vigilant application of BCPs … and not making configuration mistakes!
  - Doesn't address fundamental problems, *e.g.*, prefix hijacking!

Securing Internet Routing

- ❑ How to secure Internet routing?
  - ▪ Long standing agenda in the standards and research communities.

- ❑ Over the past 15 years, several secure Internet routing protocols have been proposed.

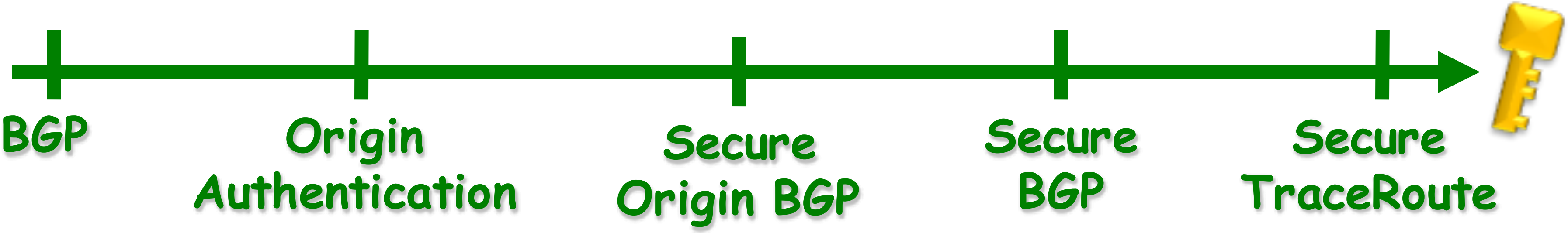
Securing Internet Routing

- ❑ The U.S. federal government is accelerating its efforts to secure the Internet's routing system … The effort … will secure the Internet's core routing protocol known as the **Border Gateway Protocol (BGP)**.

- ❑ "**BGP is one of the largest threats on the Internet**. It's incredible, the insecurity of the routing system."

(Danny McPherson, CSO at Arbor Networks, Jan 2009)

Secure Routing Protocols



BGP

Origin Authentication

Secure Origin BGP

Secure BGP

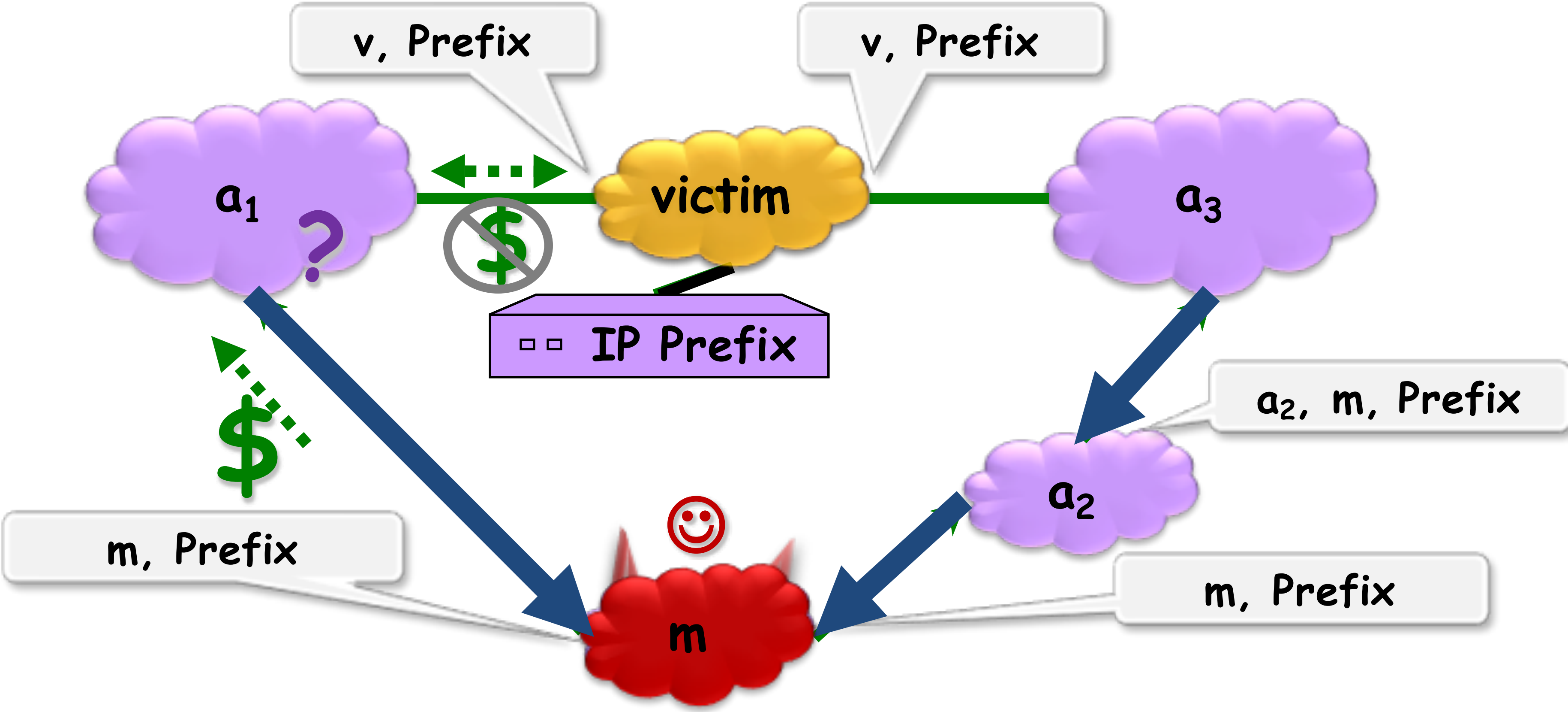
Secure TraceRoute

Prefix Hijacking and Origin Authentication

IP Address Ownership and 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



Hijacking is Hard to Debug

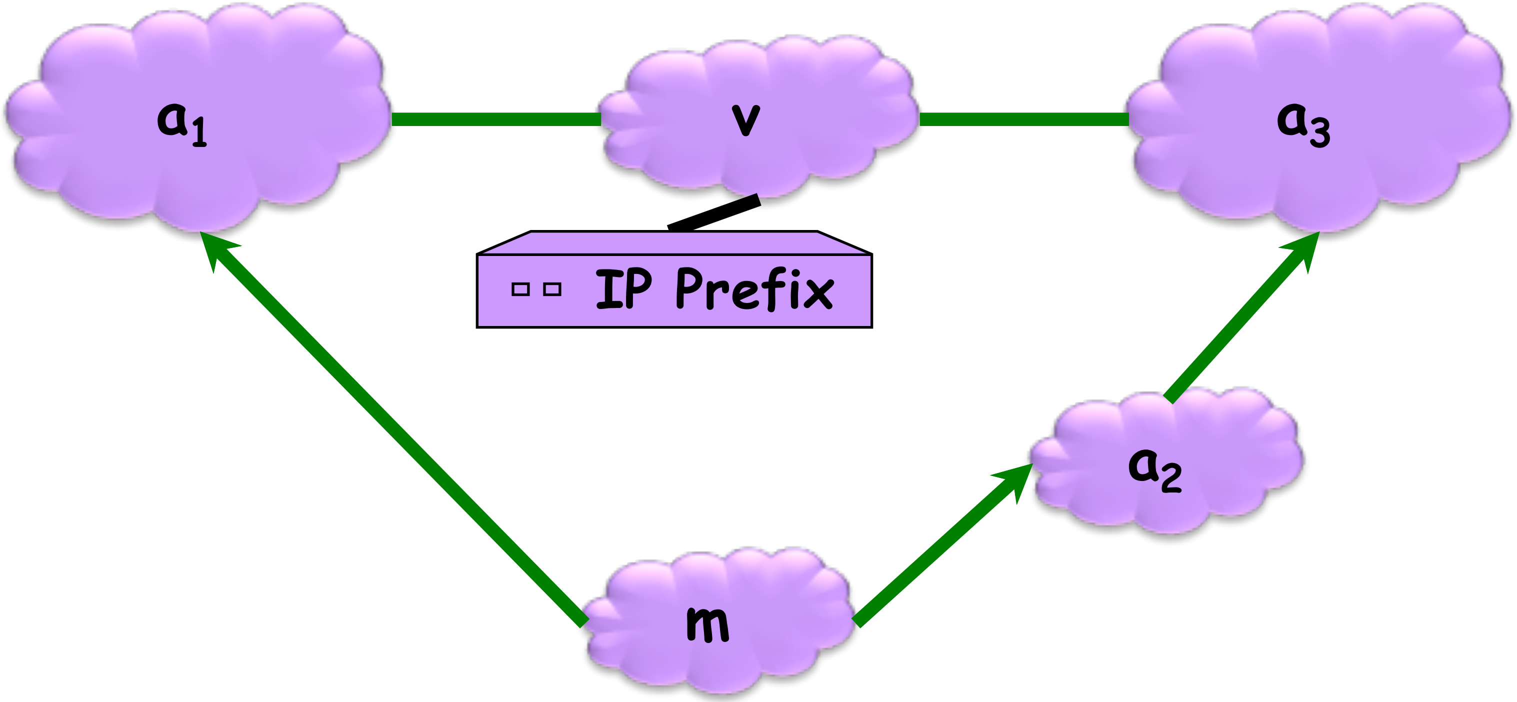
- The victim AS doesn't see the problem
  - Picks its own route
  - Might not even learn the bogus route
- May not cause loss of connectivity
  - *E.g.*, if the bogus AS snoops and redirects
  - … may only cause performance degradation
- Or, loss of connectivity is isolated
  - E.g., only for sources in parts of the Internet
- Diagnosing prefix hijacking
  - Analyzing updates from many vantage points
  - Launching traceroute from many vantage points

How to Hijack a Prefix

- ❑ **The hijacking AS has**
  - Router with BGP session(s)
  - Configured to originate the prefix
- ❑ **Getting access to the router**
  - Network operator makes configuration mistake
  - Disgruntled operator launches an attack
  - Outsider breaks in to the router and reconfigures
- ❑ **Getting other ASes to believe bogus route**
  - Neighbor ASes do not discard the bogus route
  - E.g., not doing protective filtering

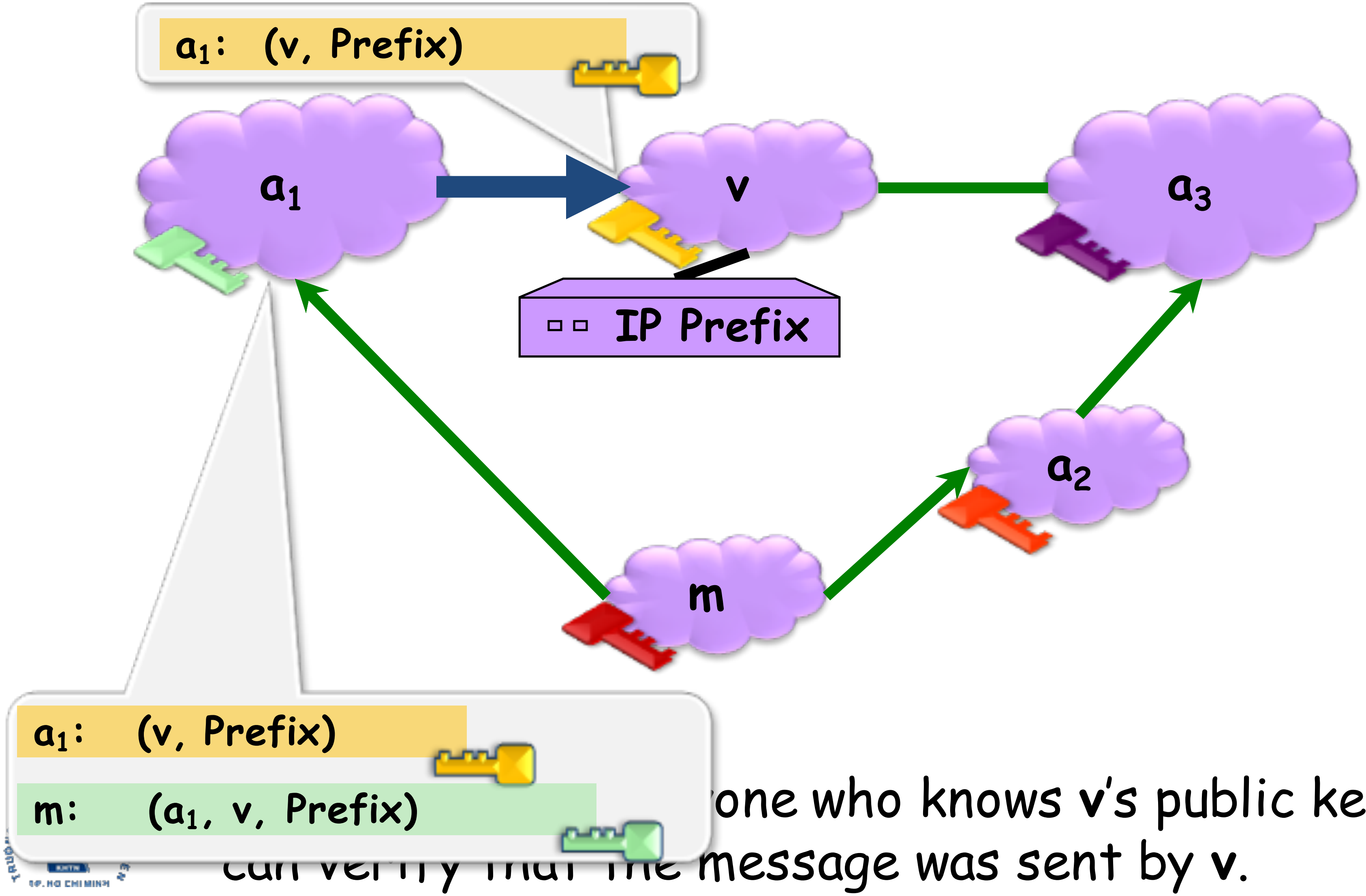
Origin Authentication

A secure database maps IP prefixes to owner ASes.



Secure BGP

Origin Authentication + cryptographic signatures



a_1: (v, Prefix)

a_1: (v, Prefix)

m: (a_1, v, Prefix)

one who knows **v**'s public key

can verify that the message was sent by **v**.

Secure BGP

❑ S-BGP can validate the order in which ASes were traversed.

❑ S-BGP can validate that no intermediate ASes were added or removed.

❑ S-BGP can validate that the route is recent.

Are We There Yet?

S-BGP Deployment Challenges

❑ **Complete, accurate registries**
- E.g., of prefix ownership

❑ **Public Key Infrastructure**
- To know the public key for any given AS

❑ **Cryptographic operations**
- *E.g.*, digital signatures on BGP messages

❑ **Need to perform operations quickly**
- To avoid delaying response to routing changes

❑ **Difficulty of incremental deployment**
- Hard to have a "flag day" to deploy S-BGP

Incremental Deployment?

- ❑ There is a necessary transition period.

- ❑ S-BGP must be backwards compatible with BGP

- ❑ Who upgrades first? Why?

Pessimistic View

ISPs would be the ones forced to ***upgrade all of their equipment*** *to support this initiative, but* ***how would it benefit them****? As commercial companies, if there is little to no benefit (potential to increase profit), why would they implement a potentially costly solution? The answer is* ***they won't****.*

[http://www.omninerd.com/articles/Did_China_Hijack_15_of_the_Internet_Routers_BGP_and_Ignorance]

unless everyone else does?

S-BGP = IPv6?

Conclusions

- ❑ **Internet protocols designed based on trust**
  - The insiders are good guys
  - All bad guys are outside the network
- ❑ **Border Gateway Protocol is very vulnerable**
  - Glue that holds the Internet together
  - Hard for an AS to locally identify bogus routes
  - Attacks can have very serious global consequences
- ❑ **Proposed solutions/approaches**
  - Secure variants of the Border Gateway Protocol

One last thing…

Harming Internet Routing Without Attacking BGP

Attacks on TCP and Data-Plane Attacks

- **Attack TCP!**
  - A BGP session runs over TCP.
- **Do not forward traffic as advertised!**
  - Drop packets!
  - Route packets along unannounced routes!

Questions