Cristina Nita-Rotaru

CS670: Network security

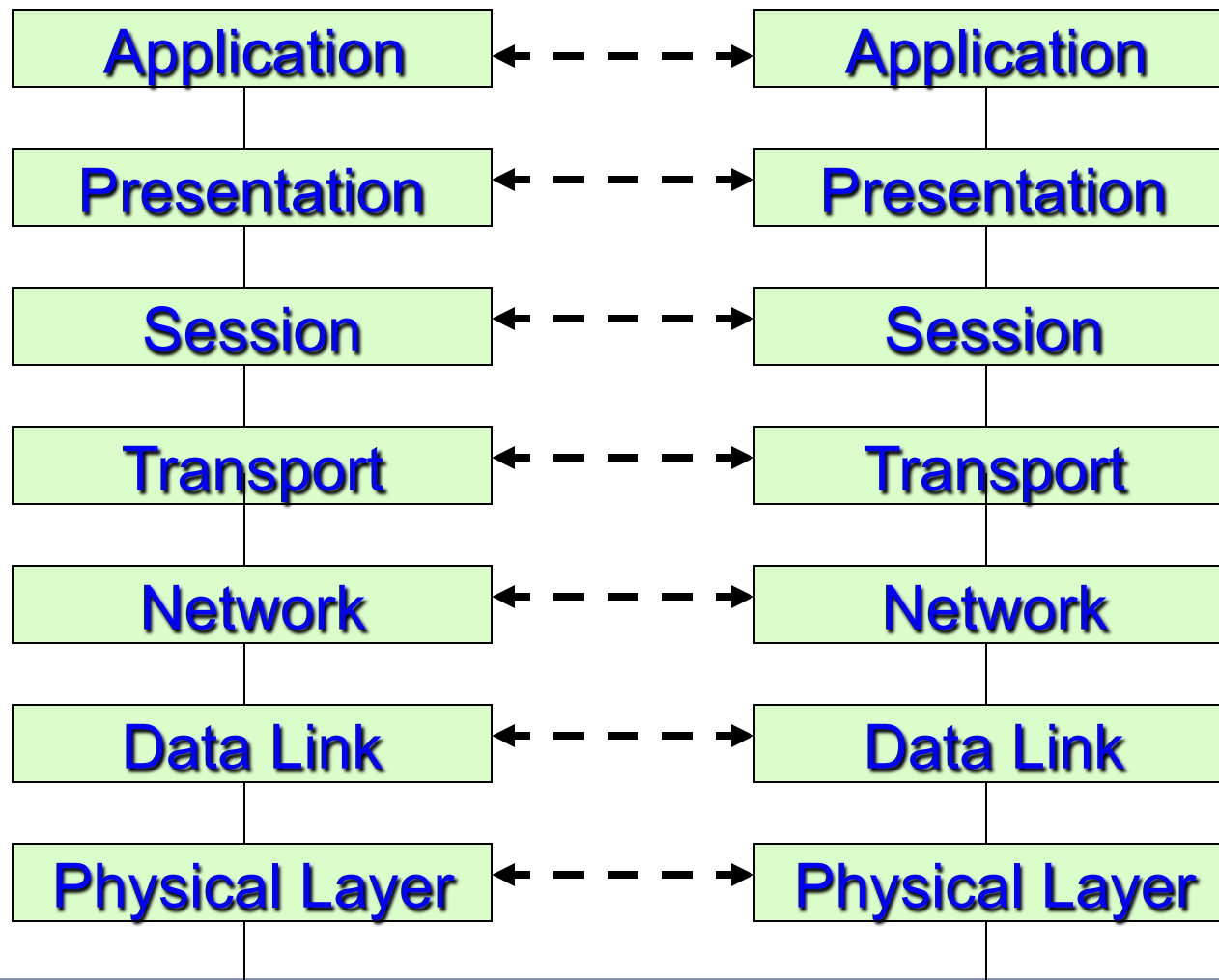## BGP. S-BGP. RPKI

Sources

1. Many slides courtesy of Christo Wilson and Wil Robertson, Northeastern

2. Secure-BGP: http://www.net-tech.bbn.com/sbgp/IEEE-JSAC-April2000/IEEE-JSAC-S-BGP.html

3. RPKI and ROA courtesy of Sharon Goldberg, BU http://queue.acm.org/detail.cfm?id=2668966

4. Slides by Jen Rexford, Princeton

1: BGP Details

OSI/ISO Model

A network of networks

▸ Internet is a "network of networks"

▸ Autonomous System (AS): a network, a single administrative domain, can span more organizations

▸ How do those networks connect
  ▸ Internet Exchange (IX)
  ▸ Network Access Points (NAP)
  ▸ Metropolitan Area Exchange (MAE)

The economy of traffic

The relationships between different ASes can be:

▶ **Transit** (or pay) – The network operator pays money (or settlement) to another network for Internet access (or transit).

▶ **Peer** (or swap) – Two networks exchange traffic between their users freely, and for mutual benefit. (often referred as peering).

▶ **Customer** (or sell) – A network pays another network money to be provided with Internet access.

Routing service

▶ Function:
  ▶ Set up routes between networks

▶ Key challenges:
  ▶ Implementing provider policies
  ▶ Creating stable paths

BGP

- Border Gateway Protocol
  - De facto inter-domain protocol of the Internet
  - Policy based routing protocol
  - Uses a Bellman-Ford path vector protocol

- Relatively simple protocol, but…
  - Complex, manual configuration
  - Policies driven by economics
    - How much **$$$** does it cost to route along a given path?
    - Not by performance (e.g. shortest paths)
  - Entire world sees advertisements
    - Errors can screw up traffic globally
  - No authentication of announcements :(

Components of BGP

- **BGP protocol**
  - Definition of how two BGP neighbors communicate
  - Message formats, state machine, route attributes, etc.
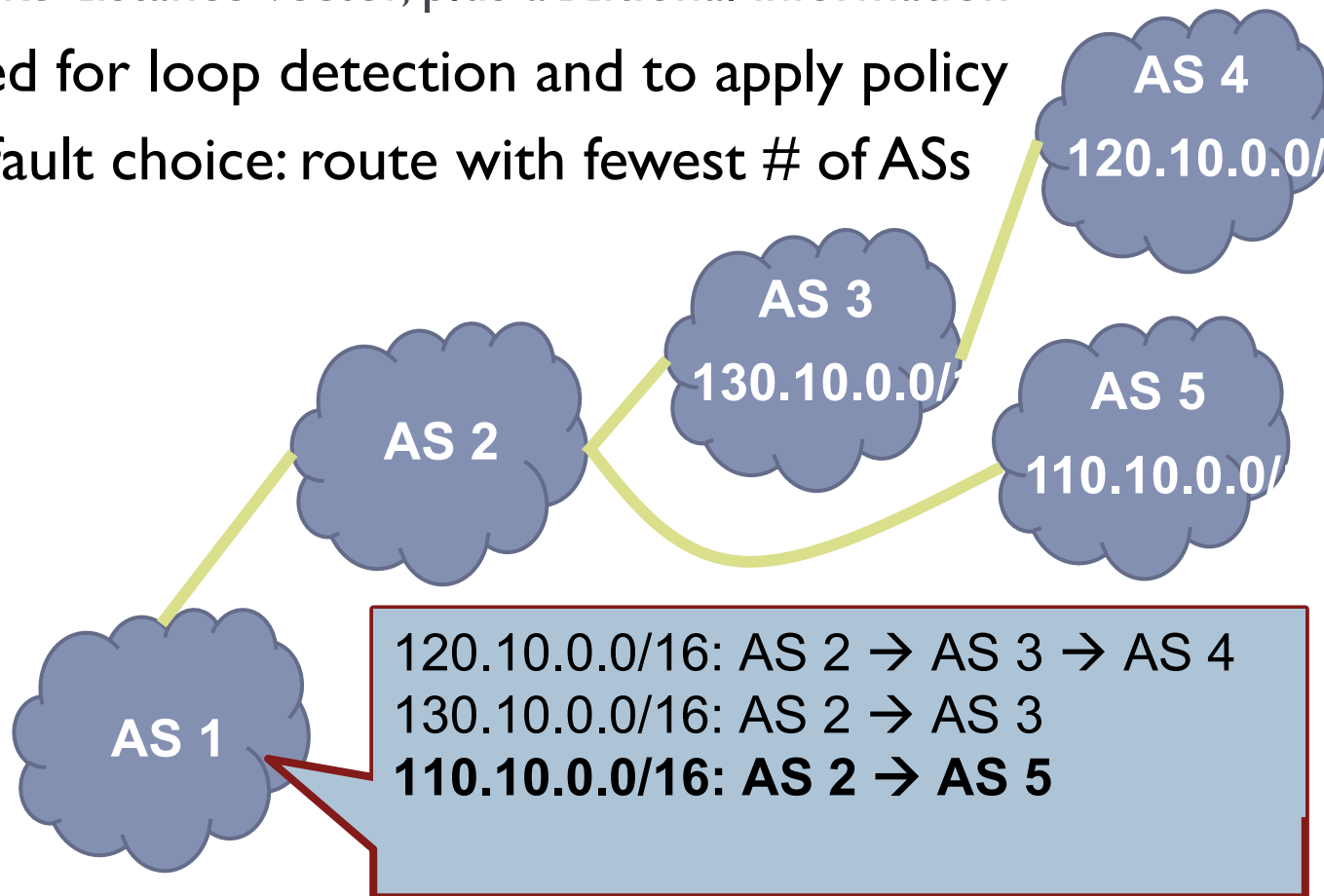 - Standardized by the IETF

- **Policy specification**
  - Flexible language for filtering and manipulating routes
  - Indirectly affects the selection of the best route
  - Varies across vendors, though constructs are similar

- **BGP decision process**
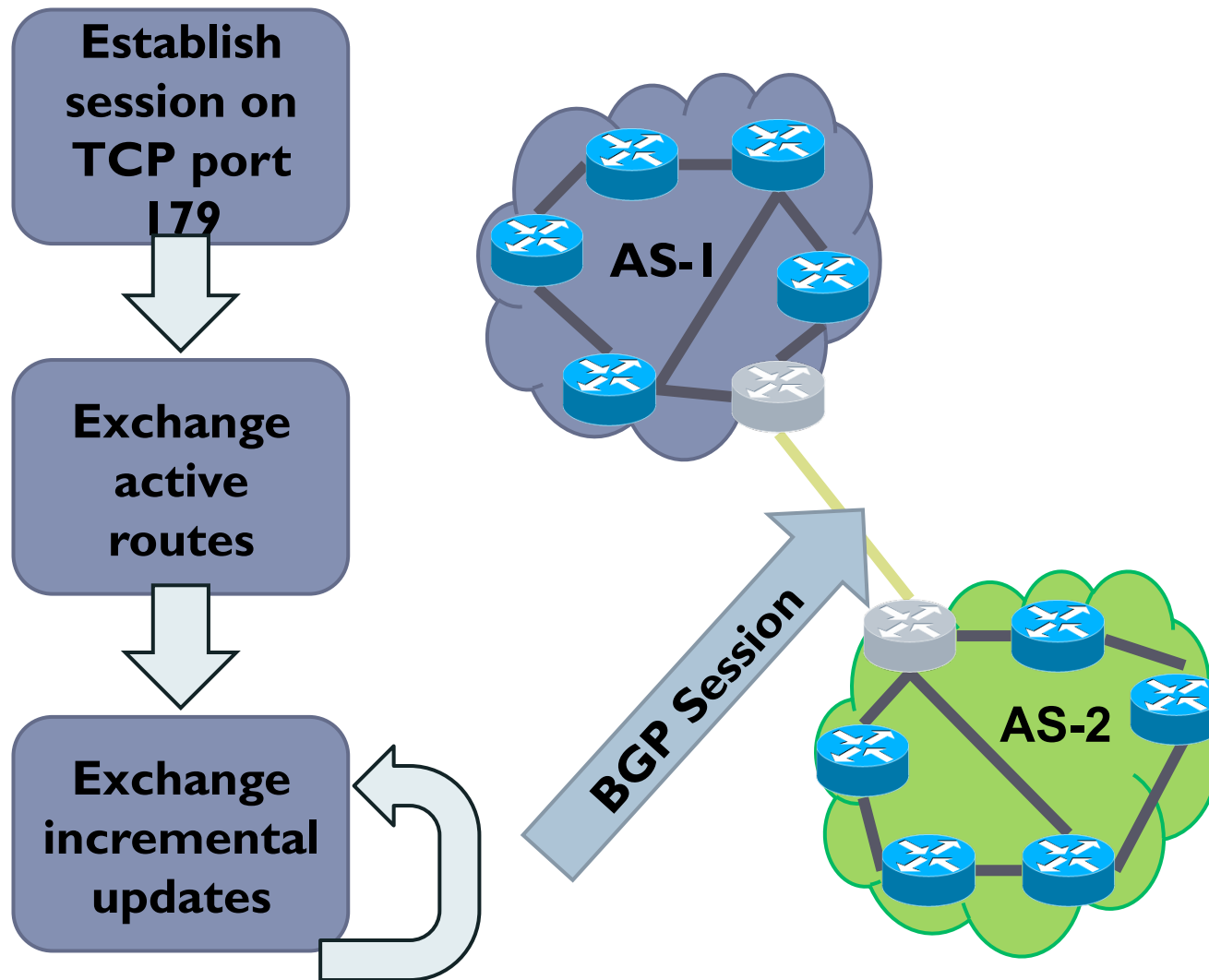 - Complex sequence of rules for selecting the best route
  - De facto standard applied by router vendors
  - Being codified in a new RFC for BGP coming soon

Path vector protocol

▸ **AS-path: sequence of ASs a route traverses**

  ▸ Like distance vector, plus additional information

▸ **Used for loop detection and to apply policy**

▸ **Default choice: route with fewest # of ASs**

AS 4
120.10.0.0/

AS 3
130.10.0.0/

AS 5
110.10.0.0/

AS 2

AS 1

120.10.0.0/16: AS 2 → AS 3 → AS 4
130.10.0.0/16: AS 2 → AS 3
110.10.0.0/16: AS 2 → AS 5

BGP

BGP operations (simplified)

Establish session on TCP port 179

↓

Exchange active routes

↓

Exchange incremental updates

AS-1

AS-2

BGP Session

Four types of BGP messages

- **Open**: Establish a peering session.
- **Keep Alive**: Handshake at regular intervals.
- **Notification:** Shuts down a peering session.
- **Update:** Announce new routes or withdraw previously announced routes.
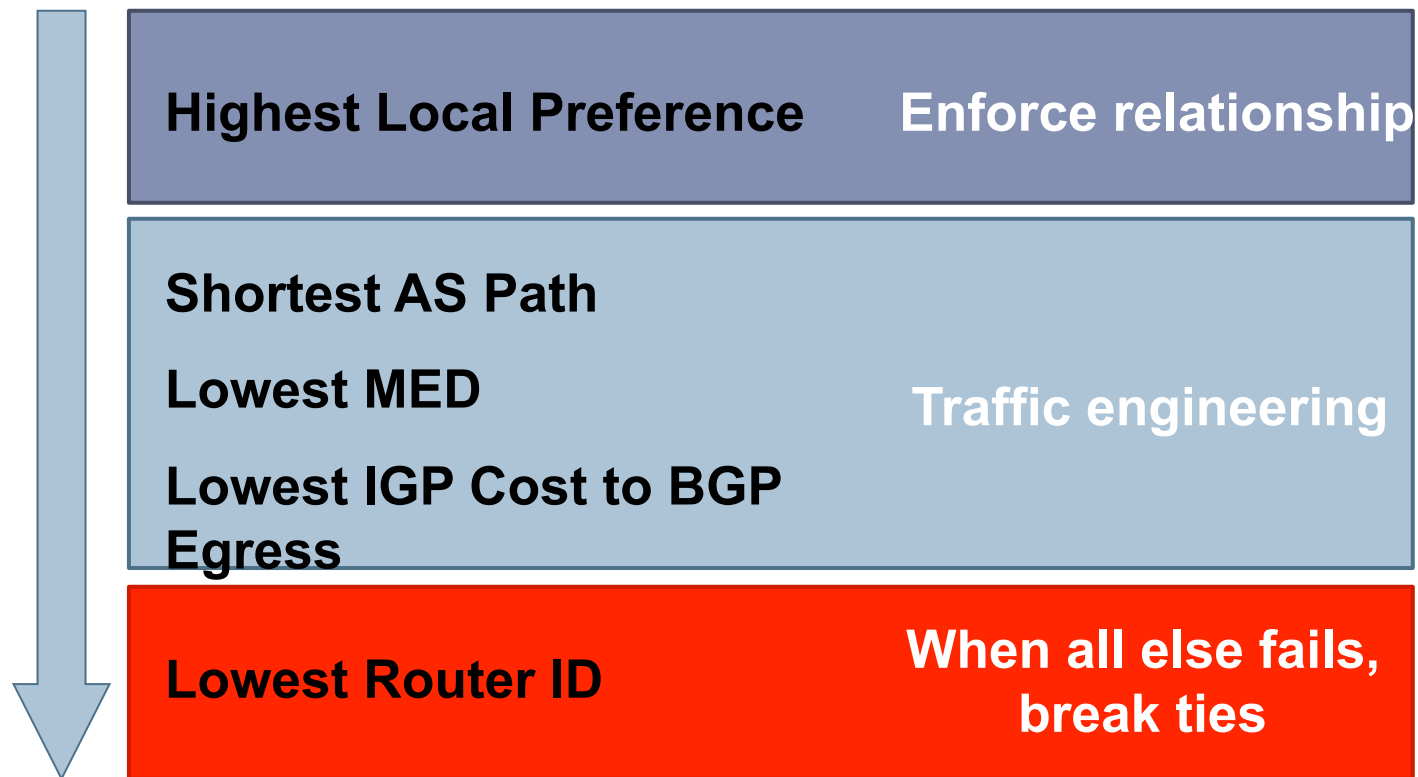
announcement = IP prefix + <u>attributes values</u>

BGP attributes summary

▸ **Well-known**: must be supported by every BGP implementation

▸ **Mandatory** attributes : must be included with every route entry. If one attribute is missing, it will result in an error message

▸ – Ex: ORIGIN, AS_PATH, NEXT_HOP, LOCAL_PREF

▸ **Discretionary** : every BGP router must recognize, but they don't have to be present with every route entry

▸ – Ex. ATOMIC_AGGREGATE

▸ **Optional**: not necessarily supported by all BGP implementations. It can be either transitive or non-transitive.
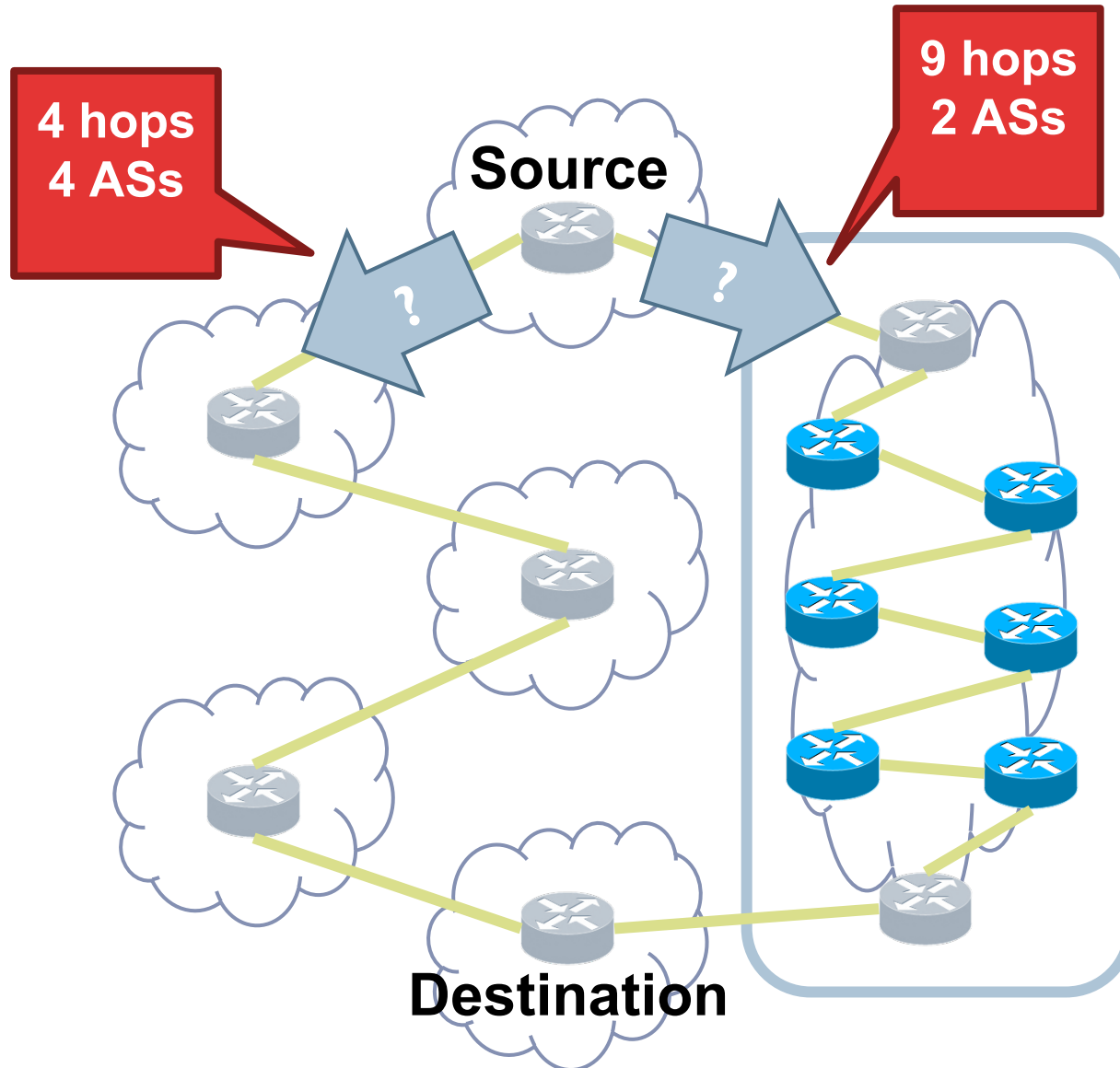
▸ – AGGREGATOR, COMMUNITY, MULTI_EXIT_DISC

BGP

Attributes used to select "best" path

- ▸ LocalPREF
  - ▸ Local preference policy to choose most preferred route
  - ▸ Overrides default fewest AS behavior
- ▸ Multi-exit Discriminator (MED)
  - ▸ Specifies path for external traffic destined for an internal network
  - ▸ Chooses peering point for your network
- ▸ Import Rules
  - ▸ What route advertisements do I accept?
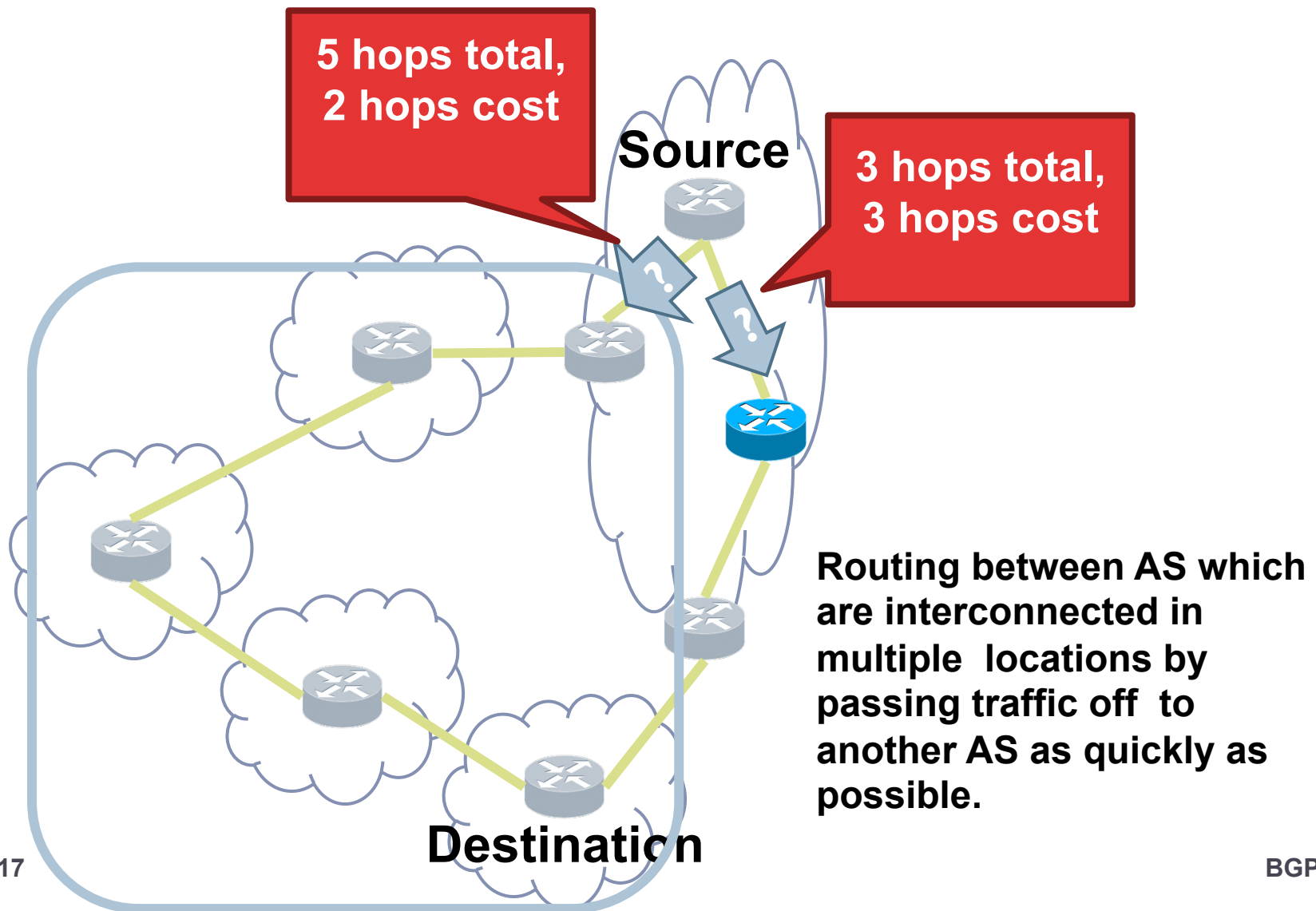- ▸ Export Rules
  - ▸ Which routes do I forward to whom?

Route selection summary

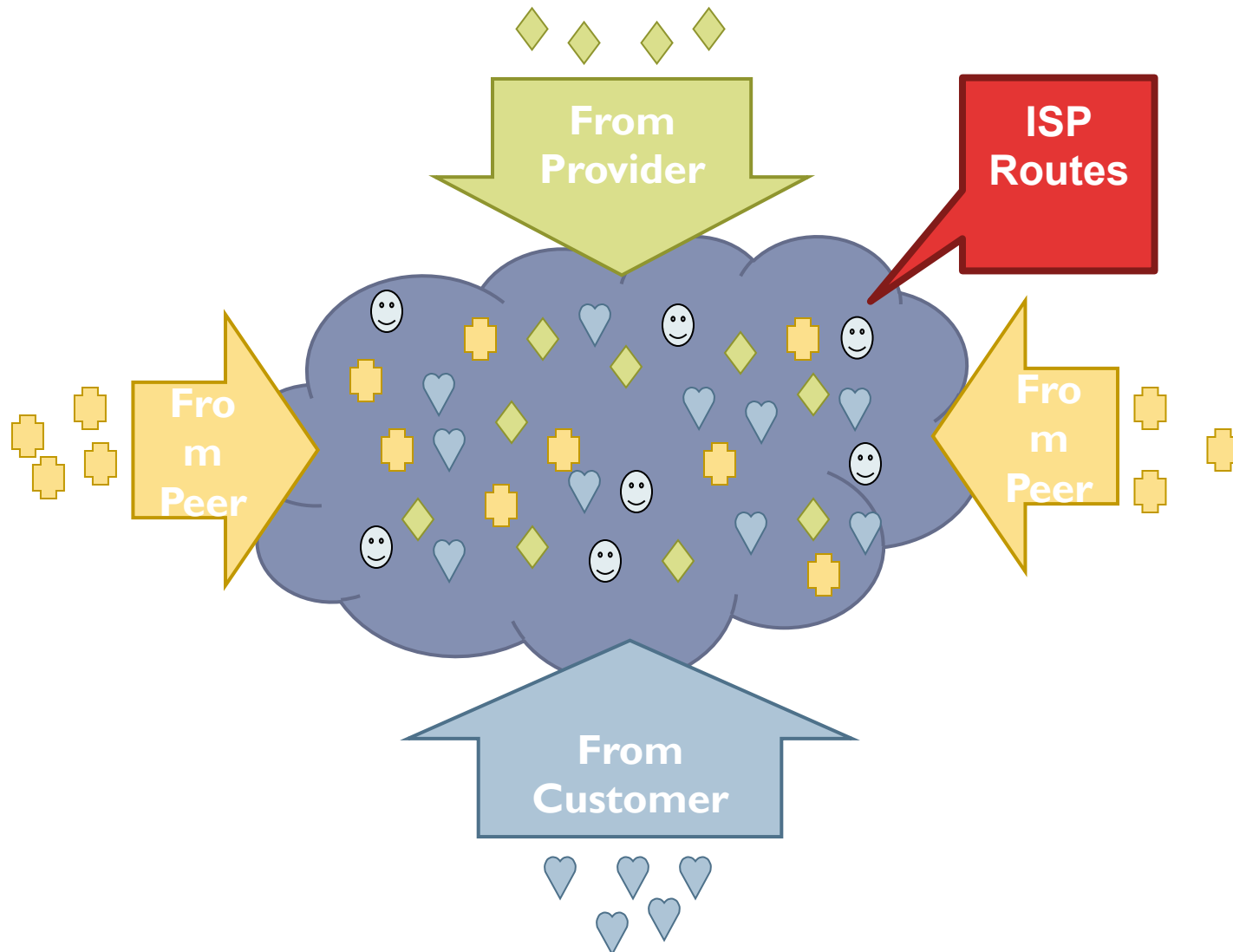**Highest Local Preference**   **Enforce relationship**

Shortest AS Path

Lowest MED   **Traffic engineering**

Lowest IGP Cost to BGP Egress

Lowest Router ID   **When all else fails, break ties**

Shortest AS path != Shortest path

4 hops
4 ASs

9 hops
2 ASs

Source

Destination

Hot potato routing

5 hops total,
2 hops cost

Source

3 hops total,
3 hops cost

Routing between AS which
are interconnected in
multiple locations by
passing traffic off to
another AS as quickly as
possible.

Destination
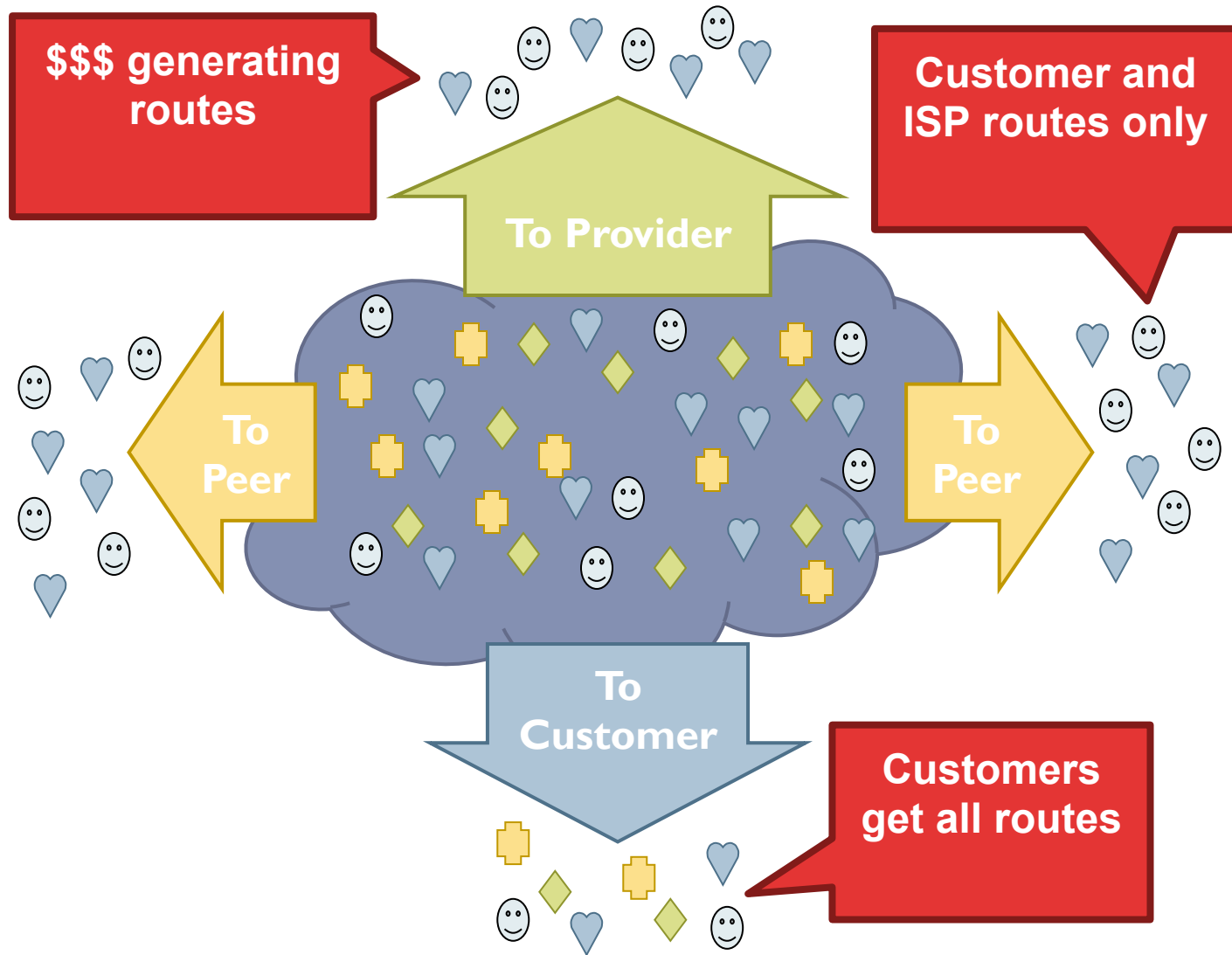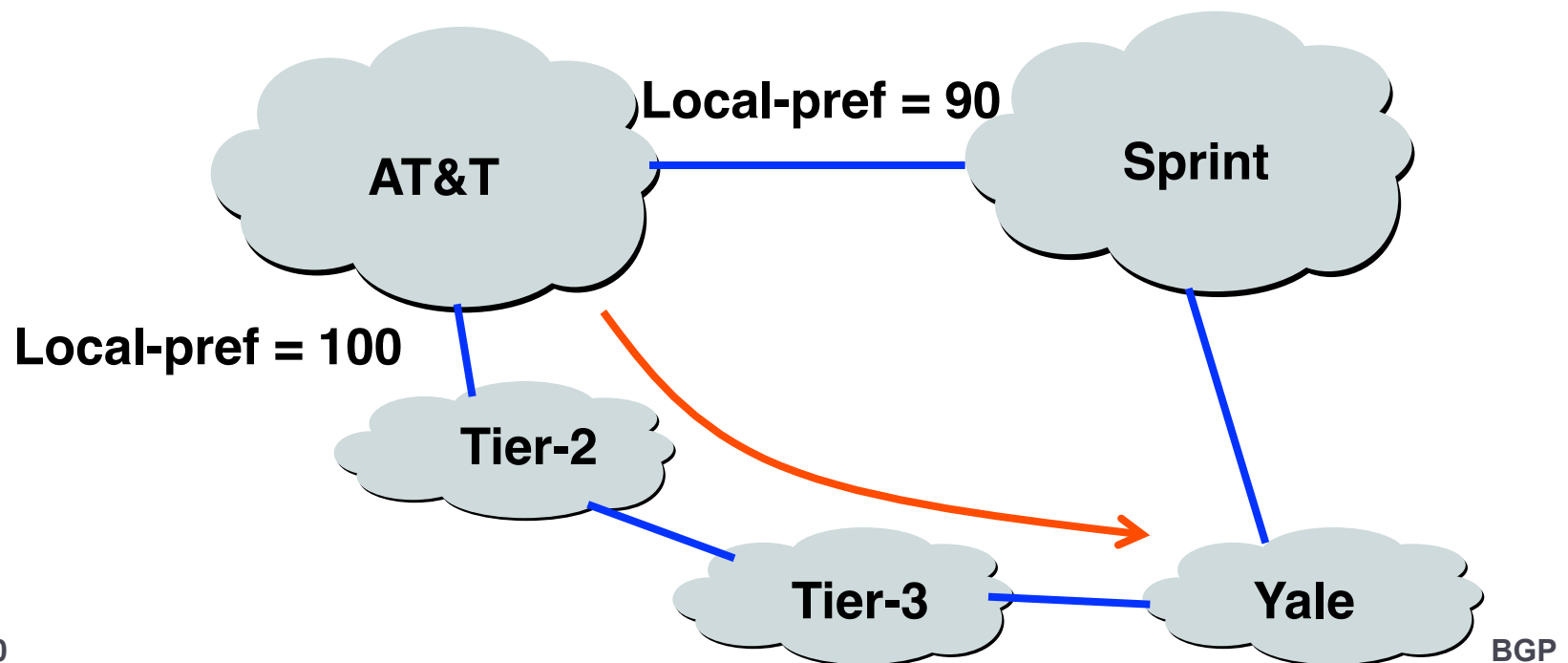
Importing routes

Exporting routes

Import policy: Local preference

▸ Favor one path over another
  ▸ Override the influence of AS path length
 ▸ Apply local policies to prefer a path

▸ Example: prefer customer over peer

Local-pref = 90

AT&T

Sprint

Local-pref = 100

Tier-2

Tier-3

Yale

BGP

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

AT&T

USLEC

Princeton

128.112.0.0/16
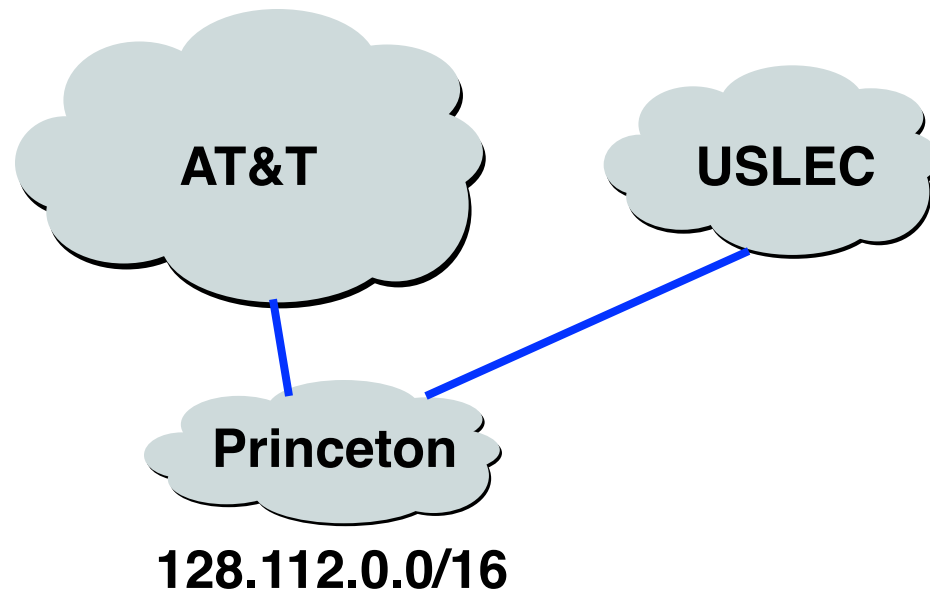
Export policy: Filtering

▸ **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

128.112.0.0/16
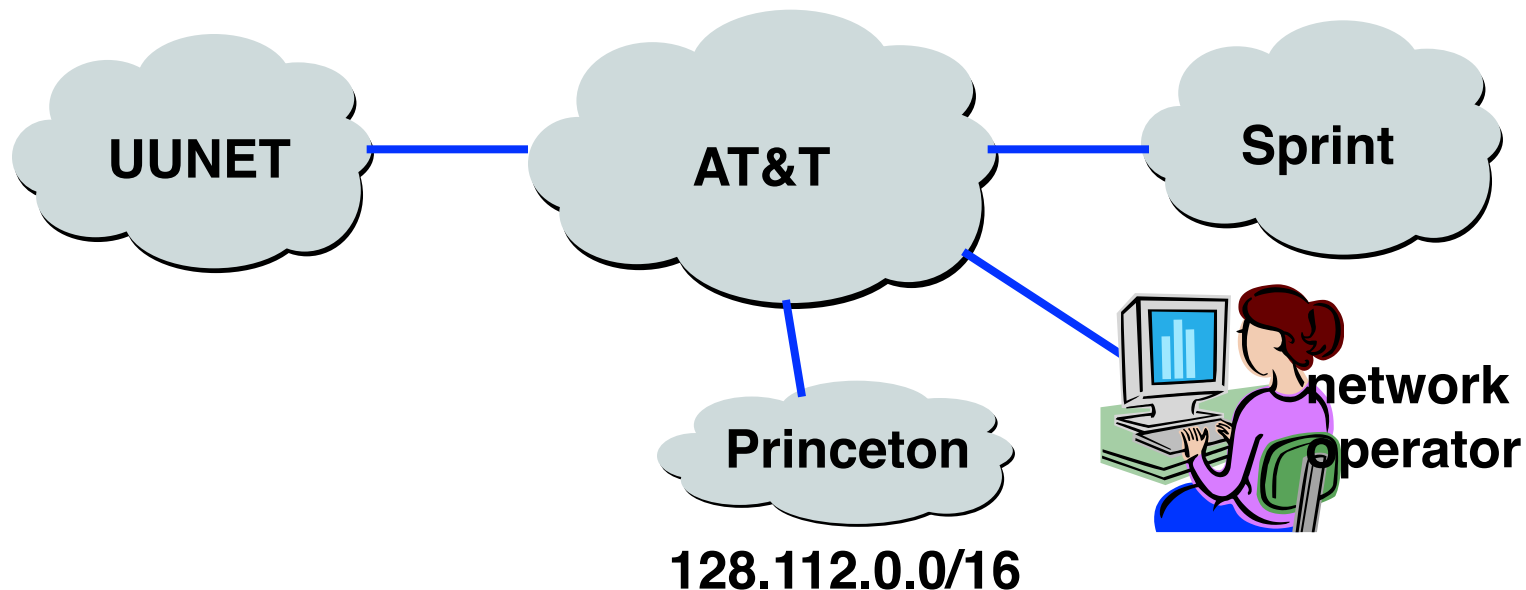
Export policy: Attribute manipulation

▸ **Modify attributes of the active route**

  ▸ To influence the way other ASes behave

▸ **Example: AS prepending**

  ▸ Artificially inflate AS path length seen by others

  ▸ Convince some ASes to send traffic another way

AT&T   Sprint   USLEC

**88 88**   **Princeton**   **88**

128.112.0.0/16
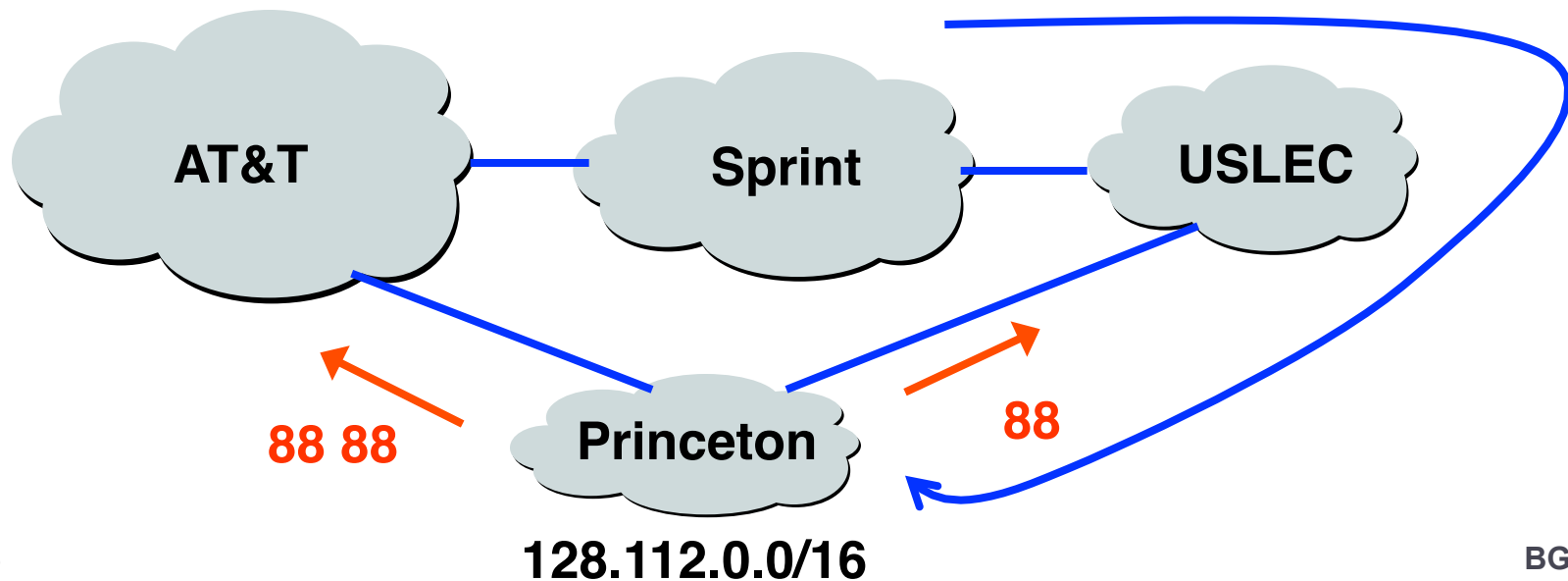
BGP Policy Configuration

▸ **Routing policy languages are vendor-specific**

  ▸ Not part of the BGP protocol specification

  ▸ Different languages for Cisco, Juniper, etc.

▸ **Still, all languages have some key features**

  ▸ Policy as a list of clauses

  ▸ Each clause matches on route attributes

  ▸ … and discards or modifies the matching routes

▸ **Configuration done by human operators**

  ▸ Implementing the policies of their AS

  ▸ Business relationships, traffic engineering, security

BGP authentication

- How are BGP sessions authenticated?
  - Shared secrets
- BGP relies on transitive trust
  - You trust your neighbor's routers...
  - Your neighbor trusts some other routers...
  - Etc.
- Are there any guarantees that:
  - An advertised route is "real"?
  - Advertised routes aren't tampered with when forwarded?

AS-1

Manually configured shared secrets between routers

BGP Session

AS-2

BGP
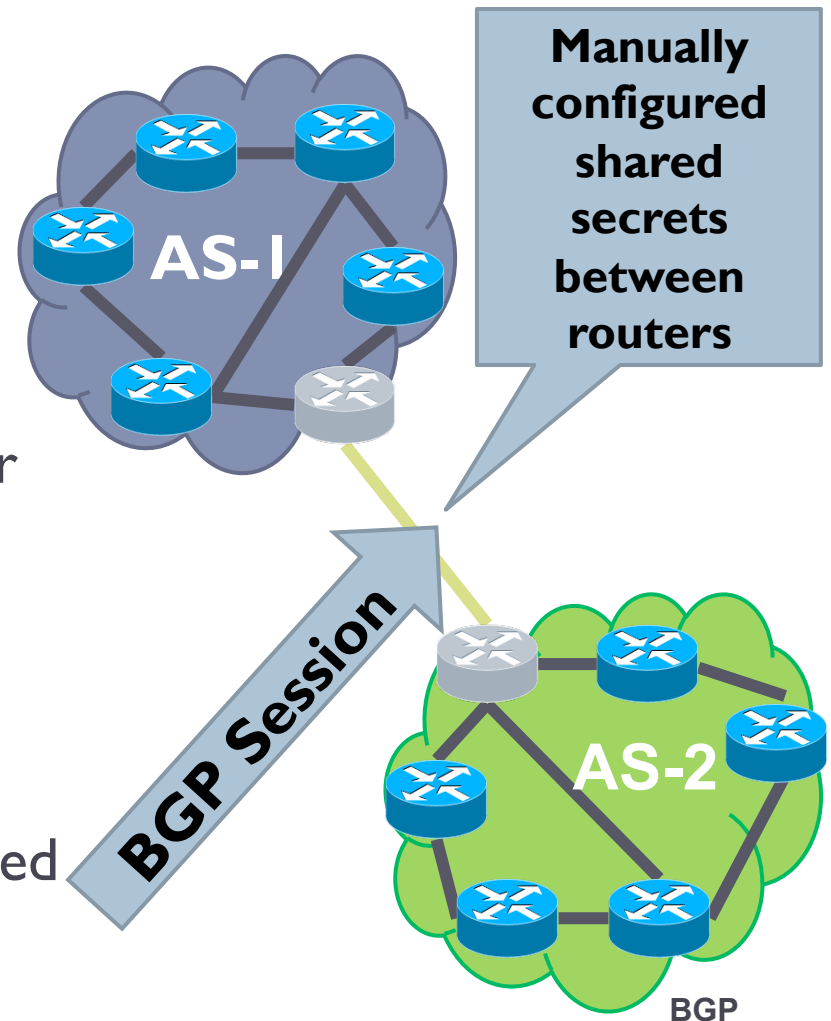
2: Attacks against BGP

AS 7007 Incident

- **Famous incident in 1997 where AS 7007 announced its internal routing table to the world**
- **Very specific (/24) routes caused many ASes to route traffic through AS 7007, creating routing black holes**

North American Network Operators Group

Date Prev | Date Next | Date Index | Thread Index | Author Index | Historical

7007 Explanation and Apology

- *From:* Vincent J. Bono
- *Date:* Sat Apr 26 19:42:16 1997

Dear All,

I would like to sincerely apologize to everyone everwhere who experienced problems yesterday due to the 7007 AS announcements.

If anyone cares to know, here is what happened:

At 11:30AM, EST, on 25 Apr 1997, our border router, stamped with AS 7007, recieved a full routing view from a downstream ISP (well, a view contacing 23,000 routes anyway).

YouTube Hijacking: A RIPE NCC RIS case study

> ℹ This content has been archived and is no longer actively maintained.

Introduction

On Sunday, 24 February 2008, Pakistan Telecom (AS17557) started an unauthorised announcement of the prefix 208.65.153.0/24. One of Pakistan Telecom's upstream providers, PCCW Global (AS3491) forwarded this announcement to the rest of the Internet, which resulted in the hijacking of YouTube traffic on a global scale.

In this report we show how the events were seen by RIPE NCC's Routing Information Service (RIS) and how, in general, one can use the RIS tools to obtain hard data on network events.

Event Timeline

- **Before, during and after Sunday, 24 February 2008:** AS36561 (YouTube) announces 208.65.152.0/22. Note that AS36561 also announces other prefixes, but they are not involved in the event.

- **Sunday, 24 February 2008, 18:47 (UTC):** AS17557 (Pakistan Telecom) starts announcing 208.65.153.0/24. AS3491 (PCCW Global) propagates the announcement. Routers around the world receive the announcement, and YouTube traffic is redirected to Pakistan.

- **Sunday, 24 February 2008, 20:07 (UTC):** AS36561 (YouTube) starts announcing 208.65.153.0/24. With two identical prefixes in the routing system, BGP policy rules, such as preferring the shortest AS path, determine which route is chosen. This means that AS17557 (Pakistan Telecom) continues to attract some of YouTube's traffic.

- **Sunday, 24 February 2008, 20:18 (UTC):** AS36561 (YouTube) starts announcing 208.65.153.128/25 and 208.65.153.0/25. Because of the longest prefix match rule, every router that receives these announcements will send the traffic to YouTube.

28   BGP

Why hijack?

▸ Human or software errors

- ▸ Routers leak internal routes to the world
- ▸ People fat finger routing entries

▸ Censorship

- ▸ Many ASs are obliged block access to specific IP ranges (e.g. Facebook, YouTube)
- ▸ Sometimes these black hole routes leak to world

▸ Spying Easy to monitor or MiTM traffic once it's routed through your network ☺

▸ Cybercrime

- ▸ Recent incident where a prefix hijack was used to steal Bitcoins from a large mining operation

BGP

Hijacking techniques

1. ## Prefix hijack
   - Most basic attack
   - Announce a prefix that the attacker doesn't actually own
   - Neighbors may route traffic for the prefix to the attacker, depending on preferences and AS topology
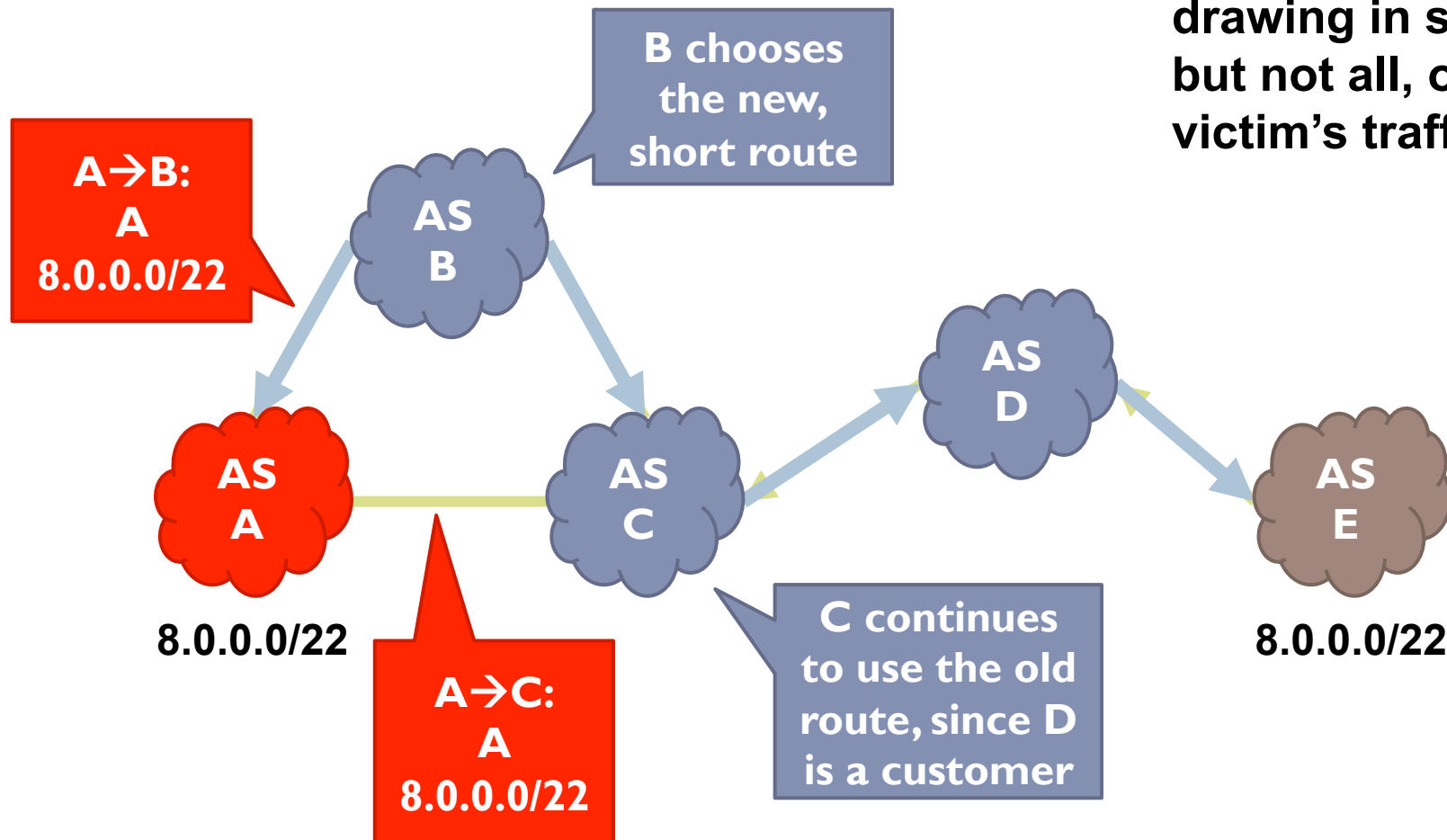
2. ## Subprefix hijack
   - Most devastating type of hijack
   - Announce a very specific prefix (e.g. a /24)
   - Routing is based on longest prefix matching, so the attackers route is likely to be selected by all ASes globally

3. ## Path shortening
   - Announce a bogus update with very few ASes on the path
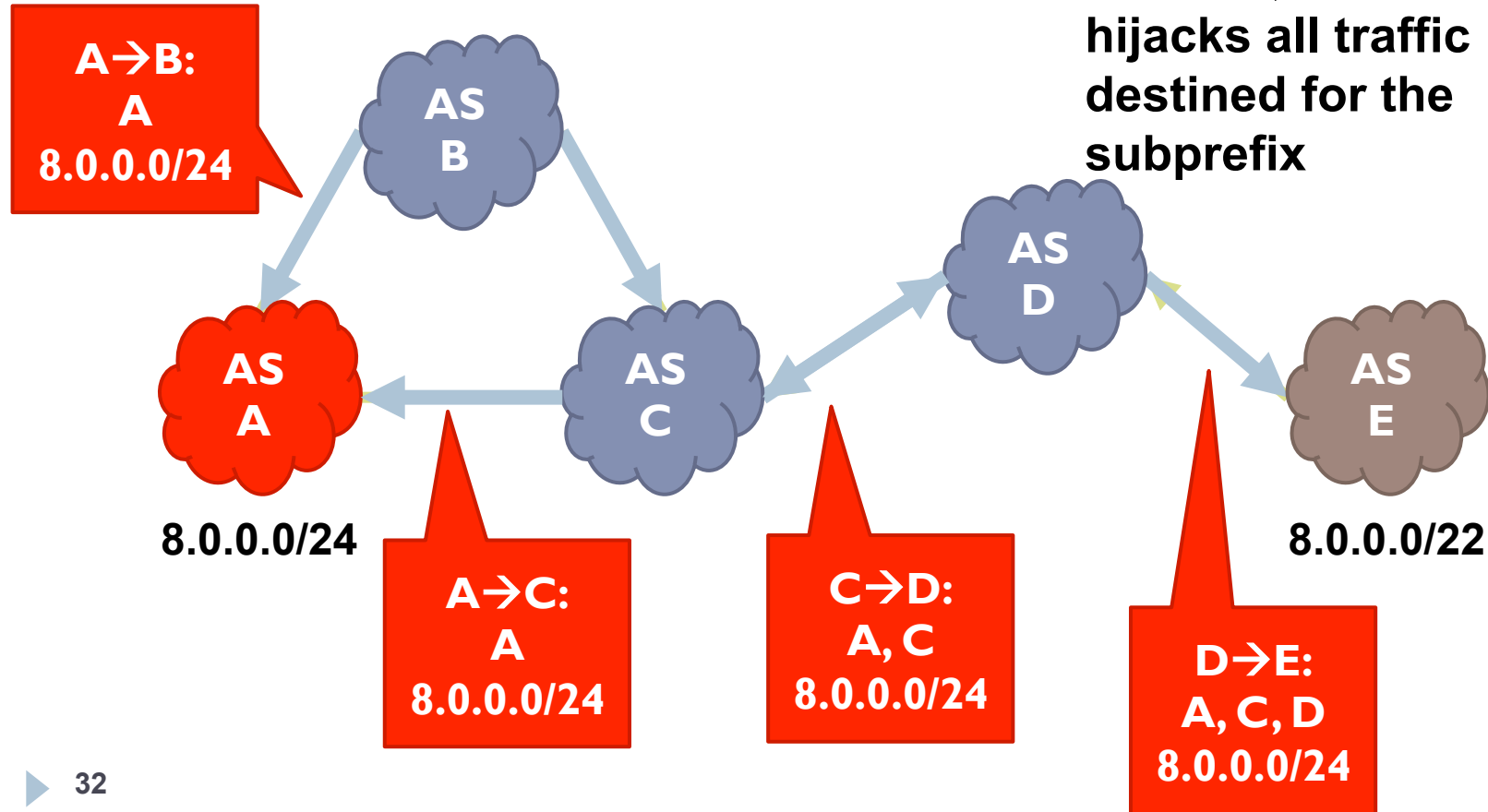   - Neighbors are likely to select the bogus path because it has a short AS path

Prefix hijack example

- **Prefix hijacking is successful at drawing in some, but not all, of the victim's traffic**
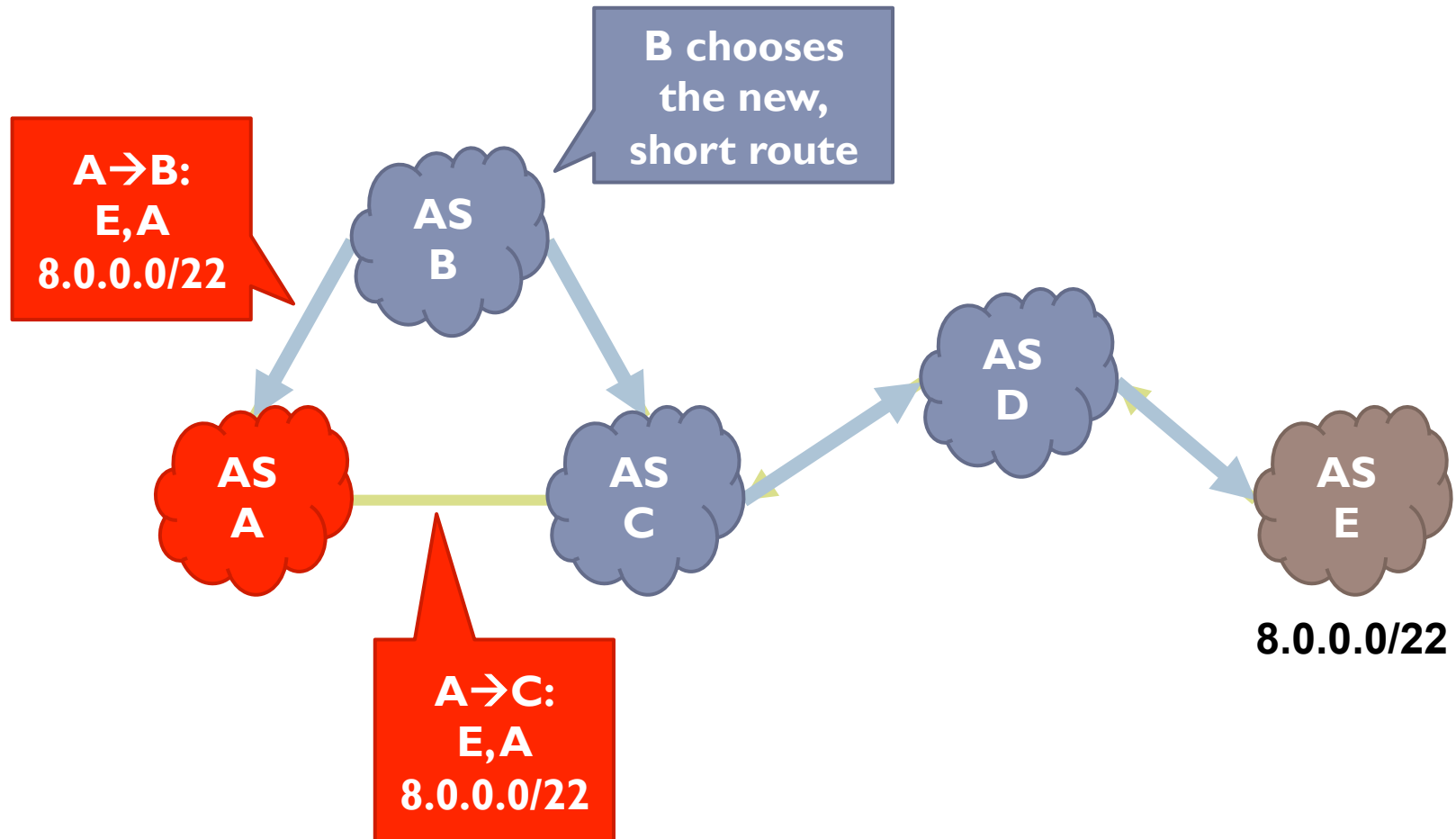
B chooses the new, short route

A→B:
A
8.0.0.0/22

AS B

AS A

8.0.0.0/22

A→C:
A
8.0.0.0/22

AS C

C continues to use the old route, since D is a customer

AS D

AS E

8.0.0.0/22

Subprefix hijack example

- **Announcement for a novel subprefix is likely to propagate**
- **/24 is more specific than /22, successfully hijacks all traffic destined for the subprefix**

A→B:
A
8.0.0.0/24

AS B

AS D

AS A

AS C

AS E

8.0.0.0/24

8.0.0.0/22

A→C:
A
8.0.0.0/24

C→D:
A, C
8.0.0.0/24

D→E:
A, C, D
8.0.0.0/24

BGP

Short path hijack

Mechanisms to secure BGP

▸ Many mechanisms have been proposed over the years

▸ We'll discuss three

  ▸ Secure BGP (S-BGP)

  ▸ RPKI and ROAs

  ▸ Anomaly Detection

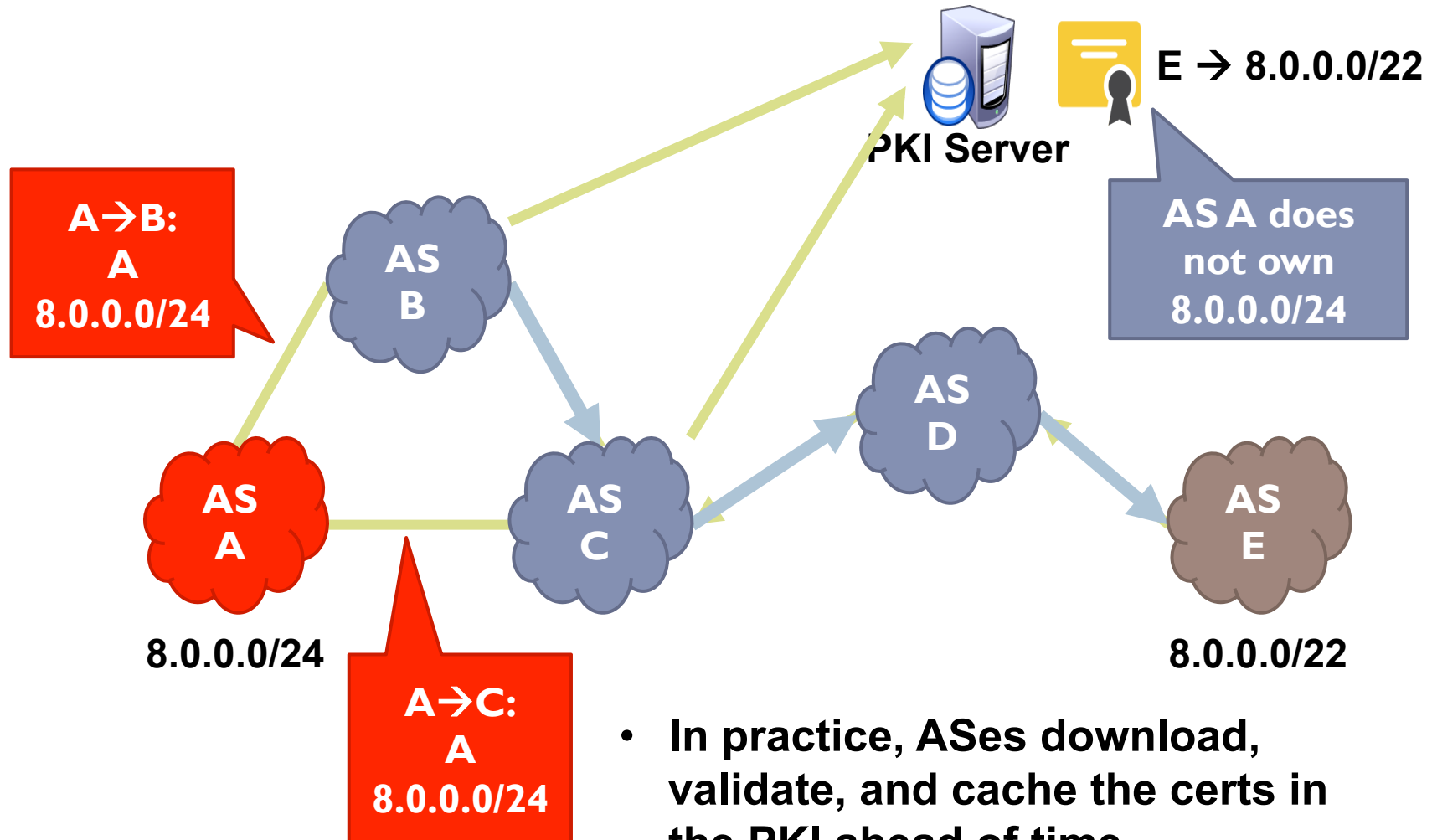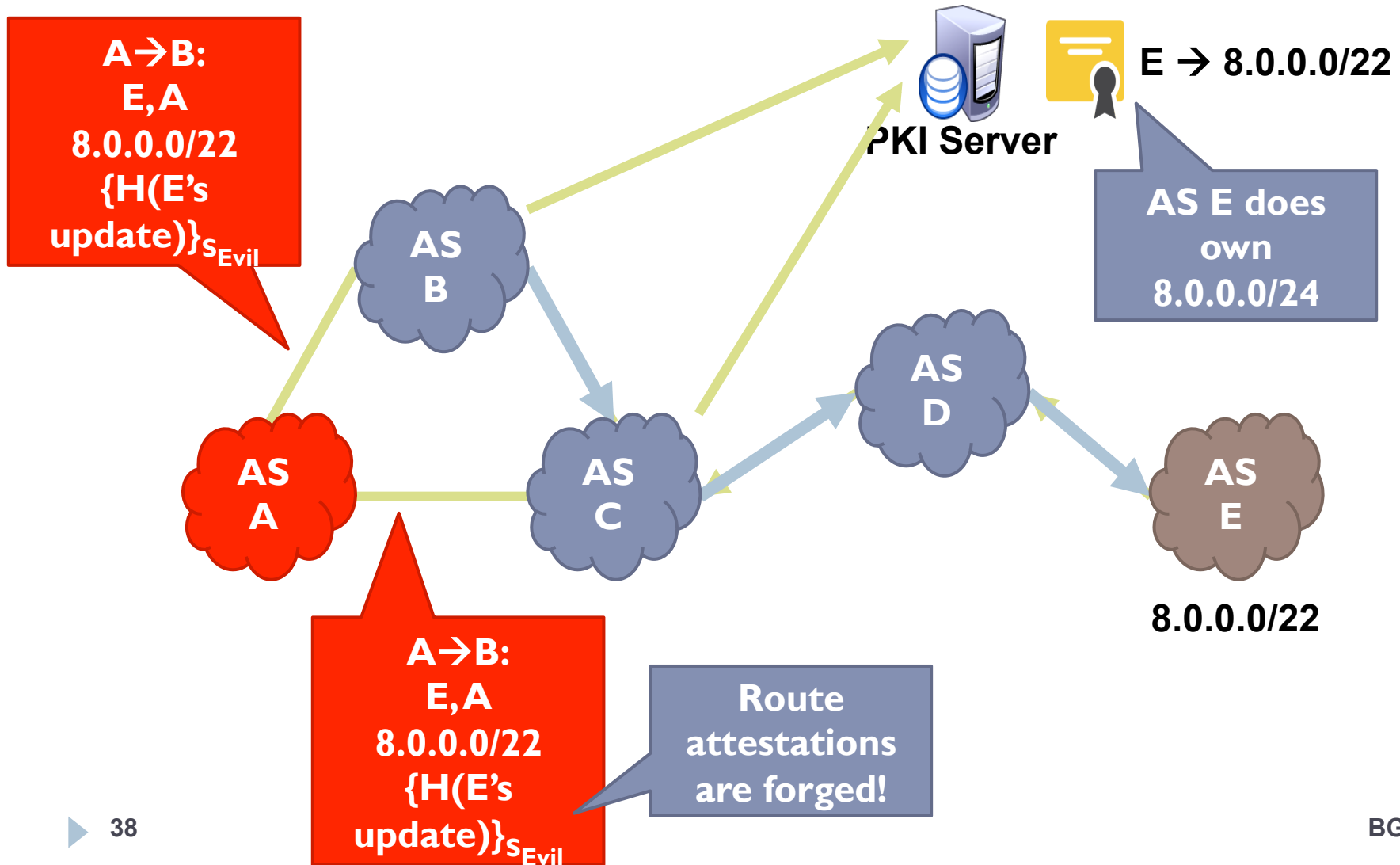# 2: S-BGP

Secure BGP

1. ## Use PKI to authenticate BGP

   ▸ Dual hierarchies of certificates bind prefix ownership to ASes and routers to ASes

   ▸ Certificate hierarchy distributed and validated out-of-band

   ▸ Routers only accept updates that are covered by valid certificates

2. ## Route attestations using "onion" signatures

   ▸ Each BGP update is signed by the announcer

   ▸ These signatures accumulate as the update propagates

   ▸ Any AS receiving the announcement can verify the signature added by each AS back to the source

S-BGP vs. subprefix hijack example

A→B:
A
8.0.0.0/24

AS B

AS A

8.0.0.0/24

A→C:
A
8.0.0.0/24

AS C

AS D

AS E

8.0.0.0/22

PKI Server

E → 8.0.0.0/22

AS A does not own 8.0.0.0/24

- **In practice, ASes download, validate, and cache the certs in the PKI ahead of time**

S-BGP vs. short path hijack

(Lack of) S-BGP deployment

▸ **S-BGP was proposed at least a decade ago, and implementations were available soon afterwards**

▸ **But, it was never deployed. Why?**

  ▸ Trust rooted in ICANN, a US organization

    ▸ Other countries are wary of centralizing power in the US

  ▸ Verification of signed attestations is costly in terms of CPU

    ▸ Routers are expensive and resource constrained

    ▸ Entire chain of attestations must be cryptographically validated for each received update

    ▸ In contrast, PKI validation can be done out of band and applied using simple filters

3: RPKI

Resource PKI

▸ Resource Public Key Infrastructure (RPKI) achieves some of what S-BGP does – i.e., origin validation

  ▸ RPKI prevents prefix and subprefix hijacking

  ▸ But, security is optional and can be incrementally deployed

▸ Regional Internet Registries (RIRs) publish signed attestations of prefix ownership and how those prefixes can be announced

  ▸ Five RIRs: ARIN (North America), LACNIC (Latin America), APNIC (Asia and Australia), RIPE (Europe, Russia, Middle East), AfriNIC (Africa)

  ▸ Attestations called Route Origin Authorizations (ROAs)

▸ By default, RPKI does not include path attestations

  ▸ Thus, RPKI is vulnerable to short path hijacks

  ▸ BGPSEC is an RPKI extension that adds cryptographic path attestation back in

Route origin attestations (ROA)

▸ Route origin authorizations (ROA) bind ownership of network prefixes to ASes

▸ ROAs also define the minimum specificity of a route announcement

▸ e.g., 192.168.0.0/16 (min=22) ⟶ AS 7007

  ▸ AS 7007 "owns" 192.168.0.0/16

  ▸ Route announcements within this prefix cannot be more specific than /22

RPKI Hierarchy Example
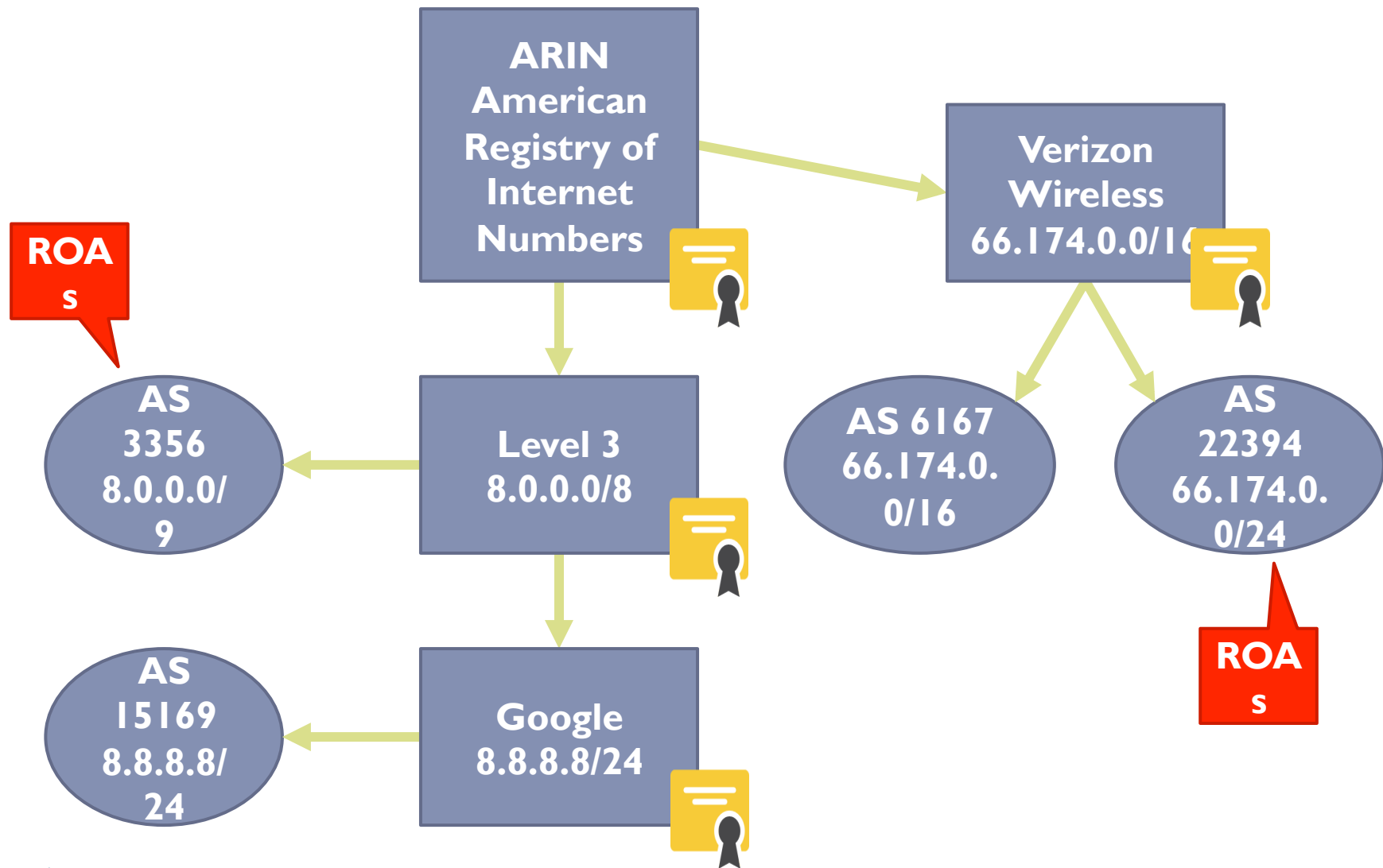
RPKI deployment

▸ The 5 RIRs have finished the deployment of RPKI, and are now offering RPKI services to their members.

▸ A number of countries (Ecuador, Japan, Bangladesh, etc.) have also started to test and deploy RPKI interiorly.

▸ However, RPKI is still in its early stages of global deployment.

  ▸ current routing table holds about 595817 IP prefixes in total, and the RPKI validation state has been determined for 38398 IP prefixes, which means that only 6.44% of the prefixes in the routing table can be validated.

▸ After years of efforts we are still a long way away from cryptographically secured BGP

3: Anomaly Detection

Routing anomalies

▸ Cryptographic attestation and verification is one approach to securing BGP (i.e. S-BGP and RPKI)

▸ Out-of-band monitoring is another – e.g., detection of routing anomalies

    ▸ Route announcements collected at many different Internet vantage points

    ▸ Use heuristics to filter updates that seem suspicious

▸ Do route announcements make sense with respect to...

    ▸ Geography?

    ▸ Internet topology?

    ▸ AS classifications?

Anomalous features

- **Geography**
  - Routes should not advertise paths that "jump" across large geographical distances
  - e.g., a route from CA to NY that transits Russia

- **Topology / AS classifications**
  - Routes should not enter and exit the Internet core (i.e., transit ASes) multiple times

- **Origin authenticity**
  - Multiple ASes should not announce ownership of the same prefix

Geographic anomaly example

AS 17557
Pakistan
Telecom
208.65.152.0/24

AS 36561
YouTube
208.65.152.0/22

7700 miles

BGP
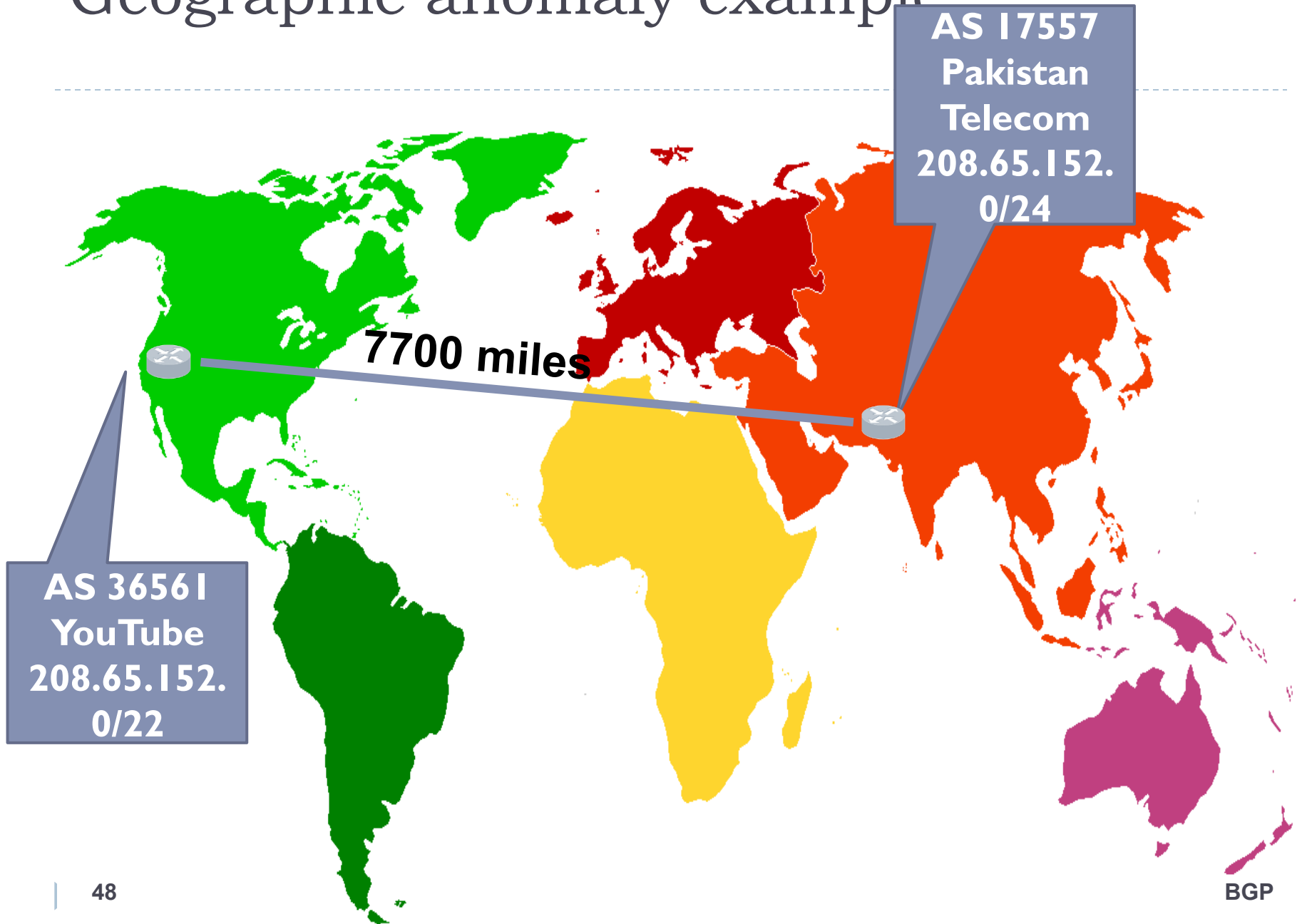
Topology anomaly example

- ▸ The A, E route doesn't make sense
- ▸ It violates the typical customer, provider relationship
- ▸ It also enters the core multiple times
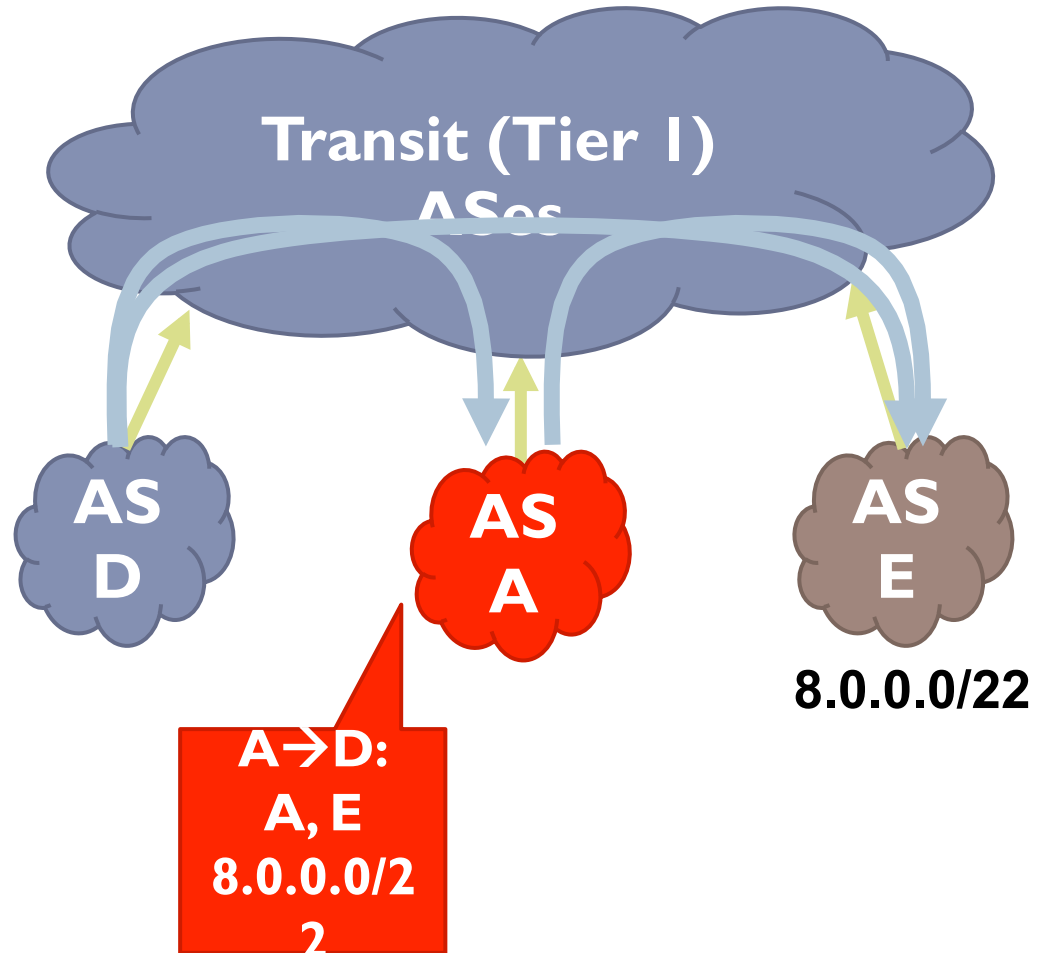
Summary

▸ DNS and BGP are crucial to the security of the Internet

▸ Both are fundamentally insecure

  ▸ Protocols lack strong (or any, in the case of BGP) authentication

▸ Solutions exist for both, but they are not yet fully deployed

  ▸ DNS is doing better than BGP, but not by much