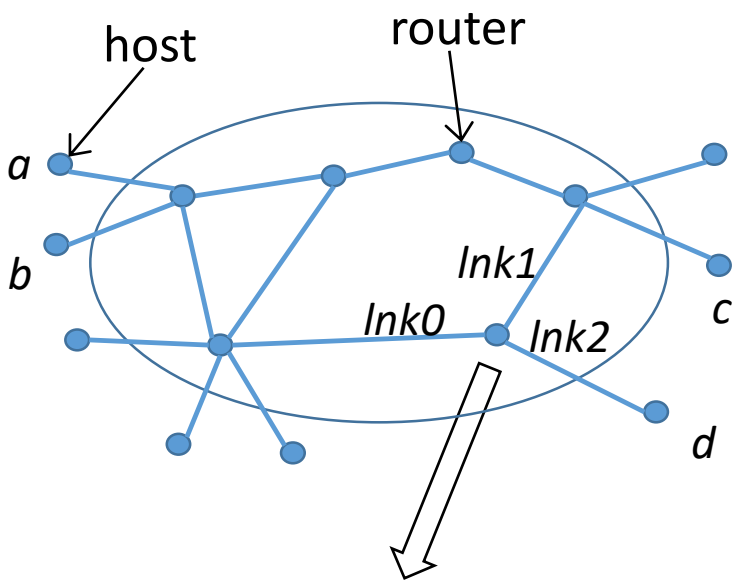
http://wanttoworkintelevision.com/how-to-use-this-site/signpost/

Routing, etc.

With material from:
Dave Levin

Internetwork: oversimplified

- Internet: nodes and links
  - nodes: routers + hosts
  - each node has an address
- Packet forwarding
  - each node has a forwarding table
  - forwards each incoming packet to the outlink of the packet's destination
- Route selection
  - maintains the forwarding table
  - initially, table has neighbors only
  - populated by exchanging routing msgs, eg, [sender id, reachable hosts]
  - adapts to changes in nodes and links

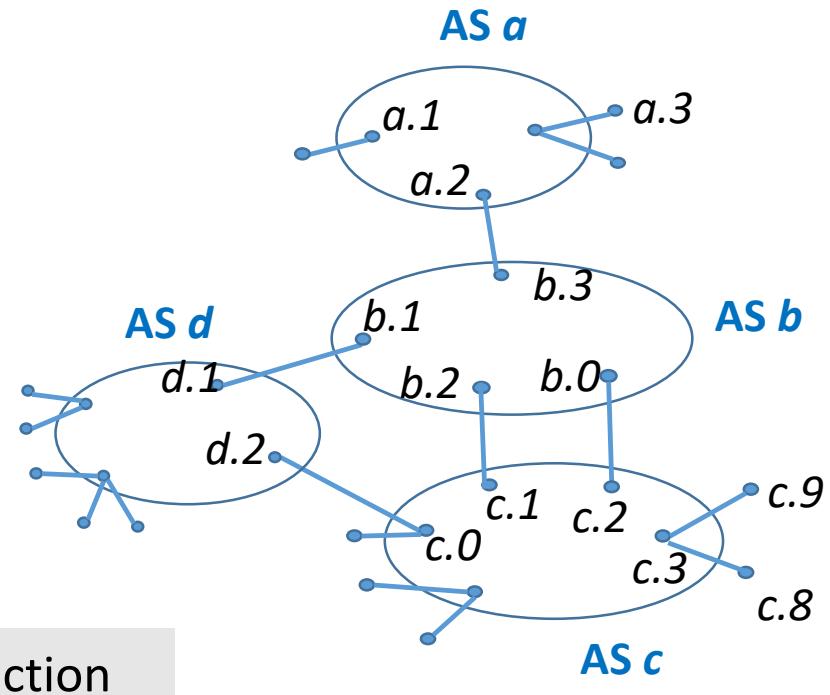


forwarding table

destination	outlink
a	*lnk0*
b	*lnk0*
c	*lnk1*
d	*lnk2*
. . .	. . .

Internetwork: bit closer to reality

- Internet is a network of autonomous systems (ASes)
- AS: "independent" network
  - Node addr: [AS #, node #]



forwarding table at *c.1*

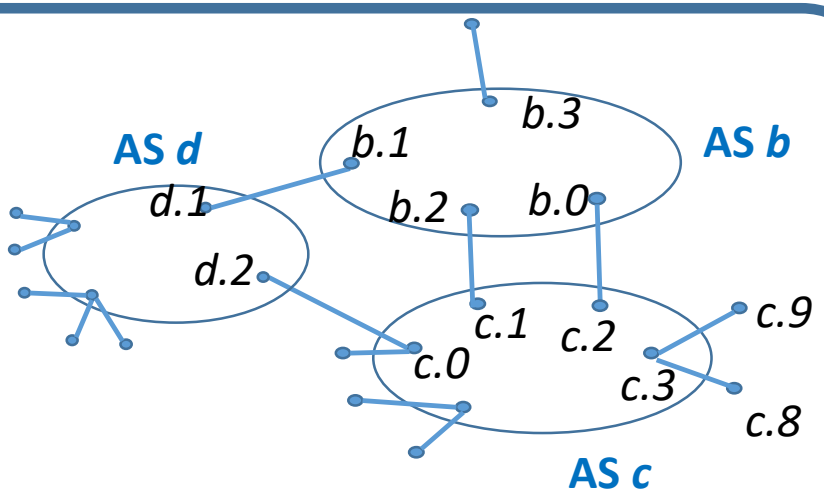
dest	outlink
c.0	. . .
c.2	. . .
c.8	*c.3*
c.9	*c.0*
. . .	
a, b	*b.2*
d	*c.0*
. . .	. . .

intra-AS route selection
(based on performance)

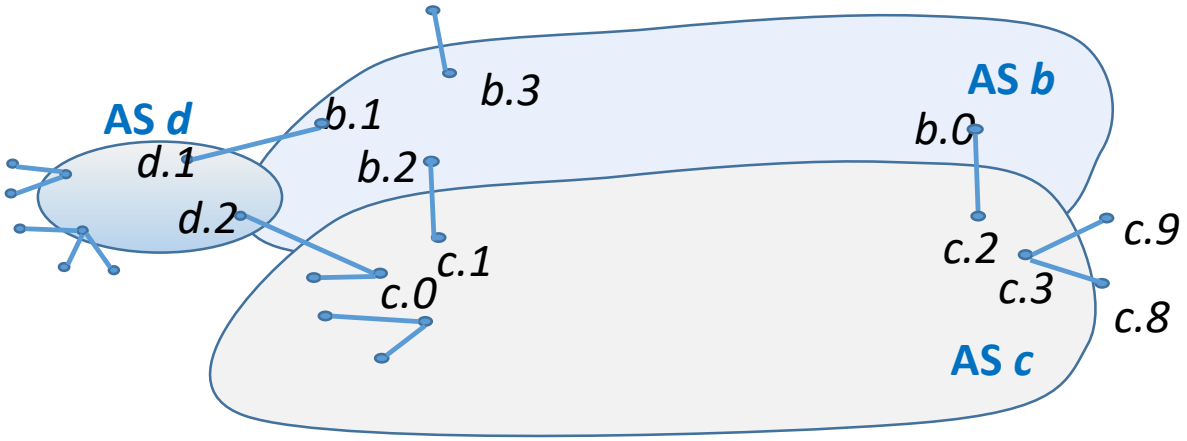
inter-AS route selection
(based on agreements with peer Ases)

ASes can overlap in geography

Layout by addresses



Layout by geography



Internetwork: closer to reality

- AS's address set defined by a list of arbitrary-length prefixes
- Forwarding table
  - Incoming packet is forwarded according to longest-prefix match
- Intra-AS route selection
  - various protocols: RIP, OSPF, . . .
- Inter-AS route selection
  - BGP: Border Gateway Protocol
  - BGP update: [AS path, set of prefixes]
  - When a router receives a BGP update, it may:
    - modify its forwarding table
    - propagate update to routers in its AS and to neighboring ASes

Internet: name-to-address mapping

- Humans prefer to identify nodes by name instead of address
  - www.cs.umd.edu instead of to 128.8.127.30
- So need to map names to addresses at run time
- DNS (domain naming system) provides this service
  - Domain name space divided into zones
    - eg: edu, umd.edu, cs.umd.edu
  - Each zone has one or more nameservers
  - Root nameservers point to top-level nameservers, who point to lower-level nameservers, and so on
  - Eg: to get the address of www.cs.umd.edu
    - query a root nameserver to get addr of an edu nameserver
    - query that to get the addr of umd.edu nameserver
    - and so on
- Cacheing for efficiency

Quick Intro to BGP

Interdomain routing

- Network belonging to one organization is an ***autonomous system (AS)***
  - Continuous block of IP addresses (prefix)
- Interdomain routing: routing between ASes
  - Target address is in a specific block

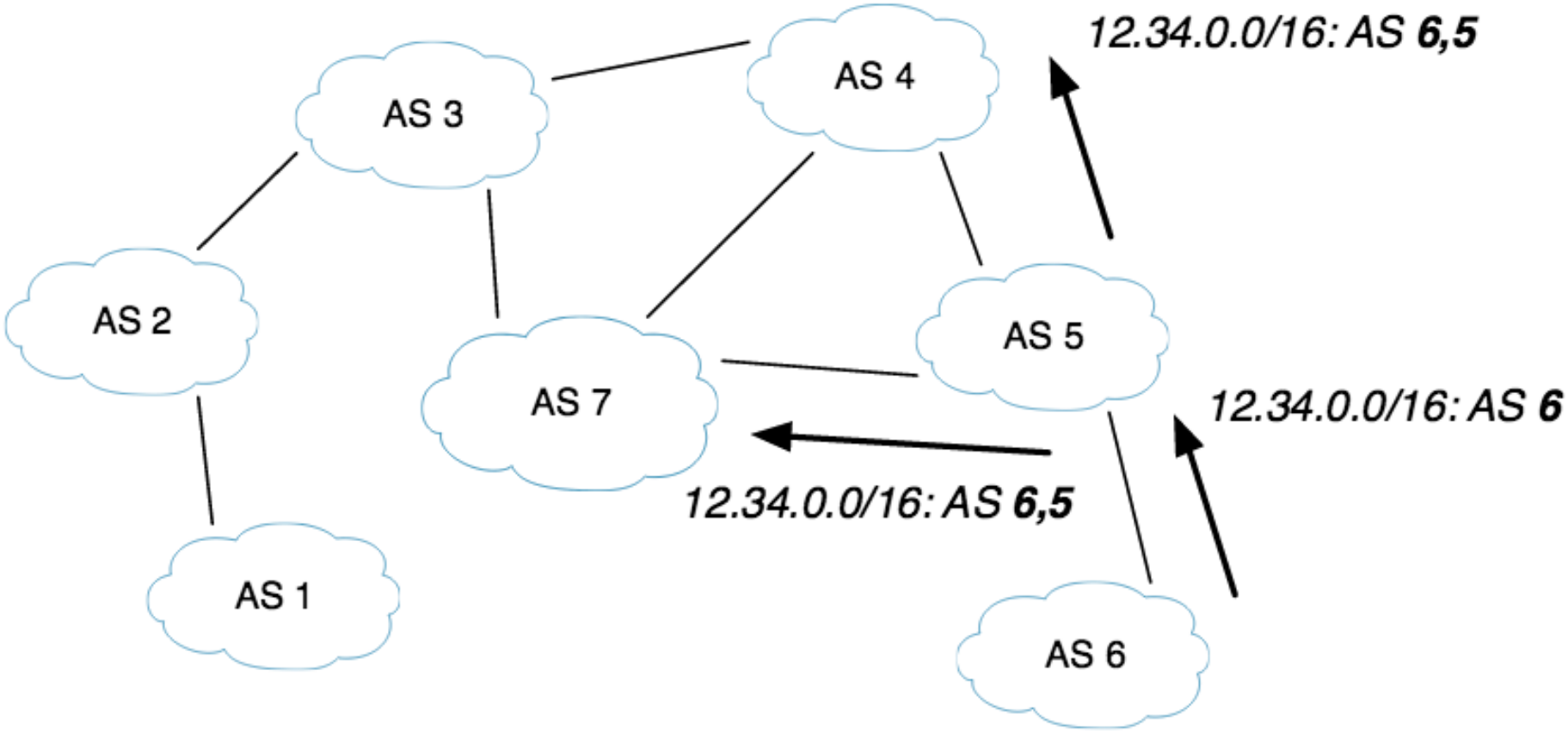
Reading an IP Prefix

- **Address Range / Bit count**
- Start from the leftmost part of the address range and count bits.
  - Included bits point at the network
  - Excluded bits are *wild cards*
- Example: 192.168.0.0 / 16
  - Really means: 192.168.(any).(any)
  - Bit count not divisible by 8 is more confusing

Introducing an AS

- ASes talk to each other to establish routes
- Advertise your own prefix to others: AS#X controls the IP prefix A.B.C.D
- Advertise prefixes you know to your neighbors:
  - Get to AS#3 via me
- Longest prefix rule: Most specific block wins

Suppose AS#6 controls 12.34.X.X



http://www.cs.princeton.edu/~jrex/papers/bgp-security08.pdf

Who can spot the problem here?

BGP Insecurity

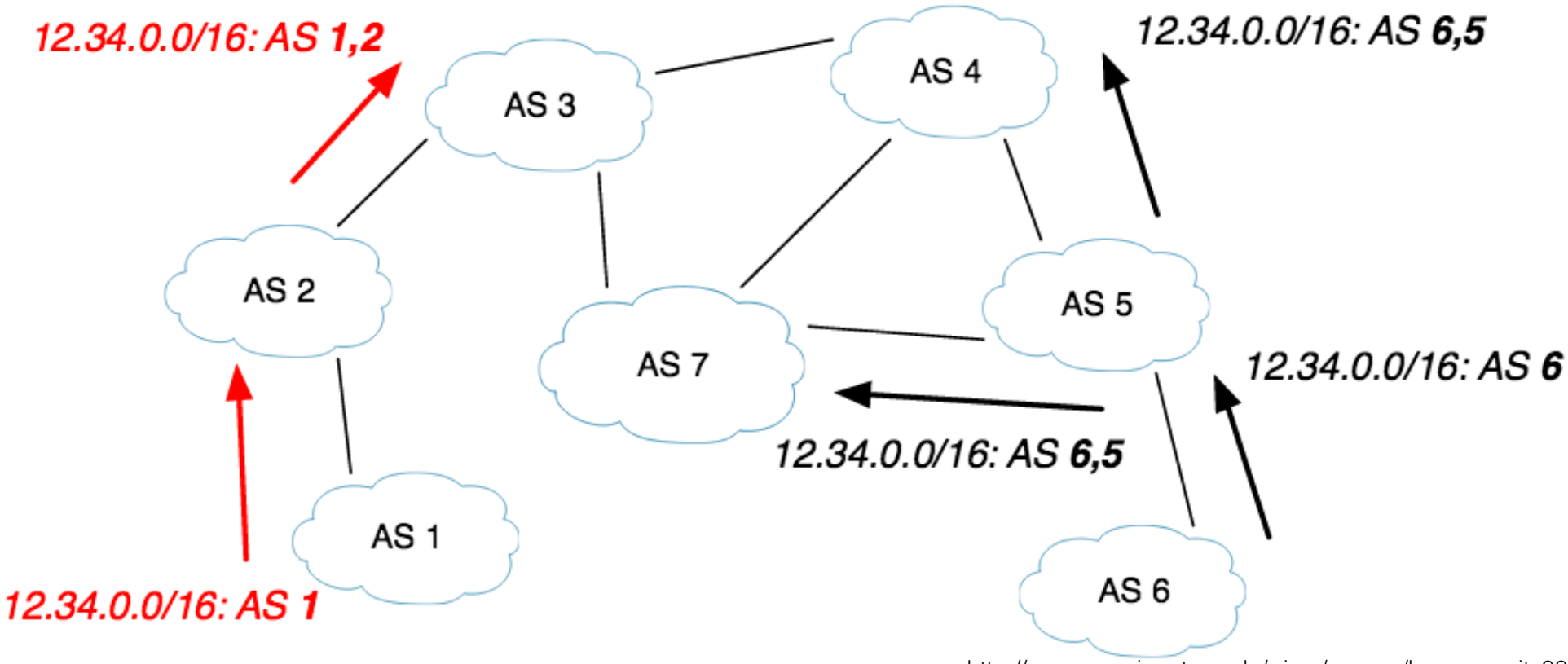
“In the early days of the Internet, getting stuff to work was the primary goal. There was no concept that people would use this to do malicious things. . . . Security was not a big issue.”

–Kirk Lougheed, co-inventor of BGP

Prefix hijacking

- Announce a prefix you don't actually own
  - Neighbors may route that traffic to you
  - Neighbors may pass wrong info along to others
  - Exploit longest-prefix rule
- Black hole: Simply drop all traffic to target
- Impersonation, interception
  - Analyze traffic, possibly forward to real dest.

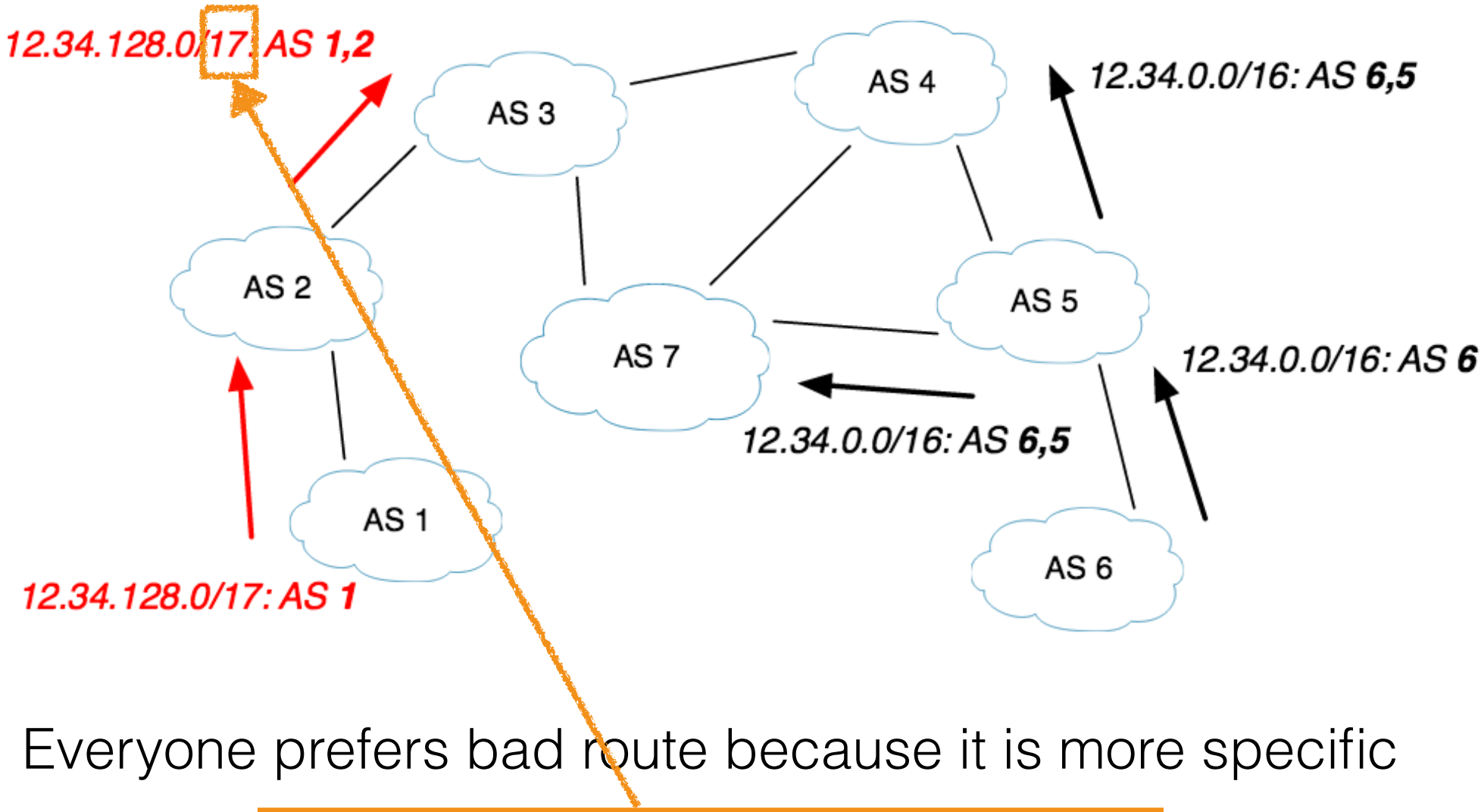
Suppose AS#6 controls 12.34.X.X



http://www.cs.princeton.edu/~jrex/papers/bgp-security08.pdf

AS2 and AS3 prefer bad route because it is shorter (fewer AS hops)

Suppose AS#6 controls 12.34.X.X



Everyone prefers bad route because it is more specific

17 bits is more specific than 16 bits!

http://www.cs.princeton.edu/~jrex/papers/bgp-security08.pdf

Examples (not malicious per se)

- 1997: Small ISP in FL claimed optimal route to all destinations
  - Most traffic rerouted there; crashed
  - Took down Internet for ~ 2 hours
- 2008: Pakistani govt orders YouTube blocked
  - Pakistan Telecom claims route to YouTube
  - More specific than real youtube prefix
  - Took down YouTube for ~2/3 of internet for ~2 hours

Examples, cont.

- 2010: China telecom advertises best routes for thousands of networks (16k in U.S.)
- Lots of traffic funneled to Beijing for 18 min
  - Including U.S. military data
  - “Most likely an accident”

TCP attacks/DOS

- Force AS router offline; neighbors withdraw routes
- Comes back up, reestablish (“route flapping”)
  - Flapping routes are disfavored
- Force traffic through your AS (eavesdropping)

Route attribute attacks

- ASes set policy (generally financial)
  - QoS guarantees, payment for transit, etc.
- BGP UPDATE messages set attribute values
  - Path length, congestion, paying customer, etc.
- Bogus announcements
  - Make path look shorter or longer
  - Add victim's AS# to suggest a loop

BGP Defenses

“TTL security hack”

- Set TTL in announcement to max value (255)
- If received at less than 254 (one away), ignore
- Prevents remote attacks (multi hops away)
  - No defense against malicious/subverted insiders
  - No defense against tunneling

Defensive filtering

- AS filters routes advertised by its customers
  - Don’t allow prefixes customer does not own
  - Would have prevented Pakistan/YouTube
- Logistical challenges (customers are complex)
- Works best if everyone does it
- Also, rewrite attributes to your preferred values

Authenticated registry

- Public registry of accurate routing data
  - Filter BGP updates accordingly
  - Can also include public keys (we'll get to why)
- Registry must be complete, accurate, secure
- Routing policy / topology within org is proprietary?

Digest for integrity

- Hash-MAC of TCP + BGP data per packet
  - Early attempt at including crypto

✓ Cannot be spoofed / fake routes are ignored

✓ Fits in existing TCP extension

– Requires shared secret
  - Public-key cert extension / IPSec (coming up)

– No confidentiality

BGPSec (was S-BGP)

- Built on certificates (public key)
- Address attestation: Claim right to an IP prefix
  - Hierarchical delegation up to ICANN
  - Distributed out-of-band
- Route attestation
  - Distributed within modified BGP UPDATE
  - Signed by each AS **in transit** (nested)

BGPSec (was S-BGP)

- Full authentication of origins and paths
- Expensive in route convergence time, storage
- Incremental deployment?
  - Need unbroken path from origin to you
- On the way to becoming a standard
  - Global adoption so far: 5%
  - North America: <1% (June 2015)

http://www.washingtonpost.com/sf/business/2015/05/31/net-of-insecurity-part-2/