Master Course
Computer Networks
IN2097

Prof. Dr.-Ing. Georg Carle
Christian Grothoff, Ph.D.
Stephan Günther

Chair for Network Architectures and Services
Department of Computer Science
Technische Universität München
http://www.net.in.tum.de

Routing Security

BGP “security” today – a sad topic…

- BGP sessions use TCP
  - No encryption – interceptors can read everything
  - “Authentication”: accept or decline AS number in OPEN message
  - Further authentication (recommended, but optional): TCP-MD5, TCP-AO
    - TCP header option contains cryptographic signature of packet
    - TCP connections only accepted from peers with accepted signature
    - No protection against replay attacks, against eavesdropping, …
  - Only accept BGP sessions from specific IP addresses?
- Defensive filtering
  - Provider knows prefixes of its (stub) AS customers:
    - Don’t accept updates for other prefixes from them
    - Don’t accept updates with other ASNs from them

BGP Routing security case study 1: How Pakistan Telecom inadvertently hijacked Youtube

-
-
  access to a certain YouTube video
- Only feasible choice was to block all YouTube traffic (208.65.152.0/22)
- They created an internal “black hole route” for their network:
  - Manual insertion of a new route for 208.65.152.0/24 into IGP
  - Packets sent via that route get discarded at the endpoint
  - Longest prefix match ⇨ This route absorbs ¼ of the /22 traffic (in this case: the part containing the servers)
- Unfortunately, this black hole route slipped into eBGP…
  - … so BGP routers world-wide saw the new route and used it
- Quick remedy by Google/YouTube?
  - Announcement of even longer prefixes 208.65.152.0/25 and 208.65.152.128/25

Youtube hijacking: Assessment

- Which security mechanisms could have worked here?
- Authentication?
  - No!
  - Pakistan Telecom is a legit BGP speaker
  - Not known for malicious behaviour
- Defensive filtering?
  - Probably not!
  - Pakistan Telecom ist not just some tiny stub AS with only one or two prefixes

BGP Routing security case study 2: How a small Czech provider terrorized the world's BGP routers

- On 2009-02-16, there was a world-wide surge in BGP updates
- Small Czech provider SuproNet (AS 47868) wanted to announce their prefix with AS path prepending
- Cisco syntax: `[…] as-path prepend 47868 47868 47868`
- …but they used MikroTik routers. Syntax: `bgp-prepend 3`
- 47868 cast into 8 bits: 47868 mod 256 = 252
- Result: AS path of length 252 (=unusually long)
- Path became longer as the announcement travelled through the world… and approached length 256 (=maximum)
- Many Cisco routers could not handle the long AS path and sent out invalid BGP messages
- Result = BGP session resets at their BGP neighbours
  - Remove all BGP routes learned from the crashed router
  - Accordingly, send BGP updates to neighbours

AS path terror: Assessment (1)

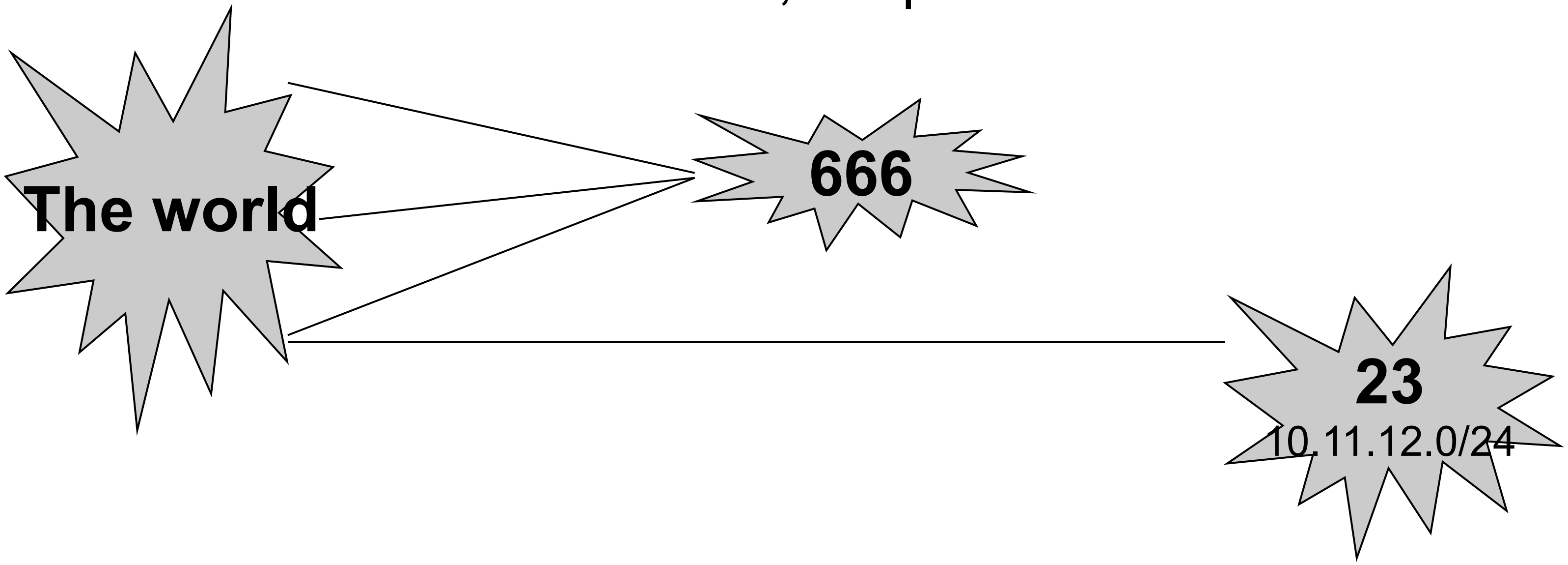
- So… who is to blame?
- SuproNet
  - Network administrator principle:
    Thou shalt read the documentation of your router…
  - …especially if it is about BGP
- MikroTik
  - Number was way too large
  - UI design principle:
    Thou shalt do error checking on user input!
    (If a user can enter garbage, he will do it.)
- Cisco
  - Strange input (long AS path) resulted in malformed output
  - Network software design principle:
    - Thou shalt do error checking on network input
    - Error checking on network output also is a good idea

AS path terror: Assessment (2)

- Which security mechanisms could have worked here?
- Authentication?
  - No!
  - SuproNet is a legitimate BGP speaker
  - Not known for malicious behaviour
- Defensive filtering?
  - SuproNet just announced their very own prefix
- Intercepting malformed BGP updates?
  - That’s exactly what crashed those BGP sessions…

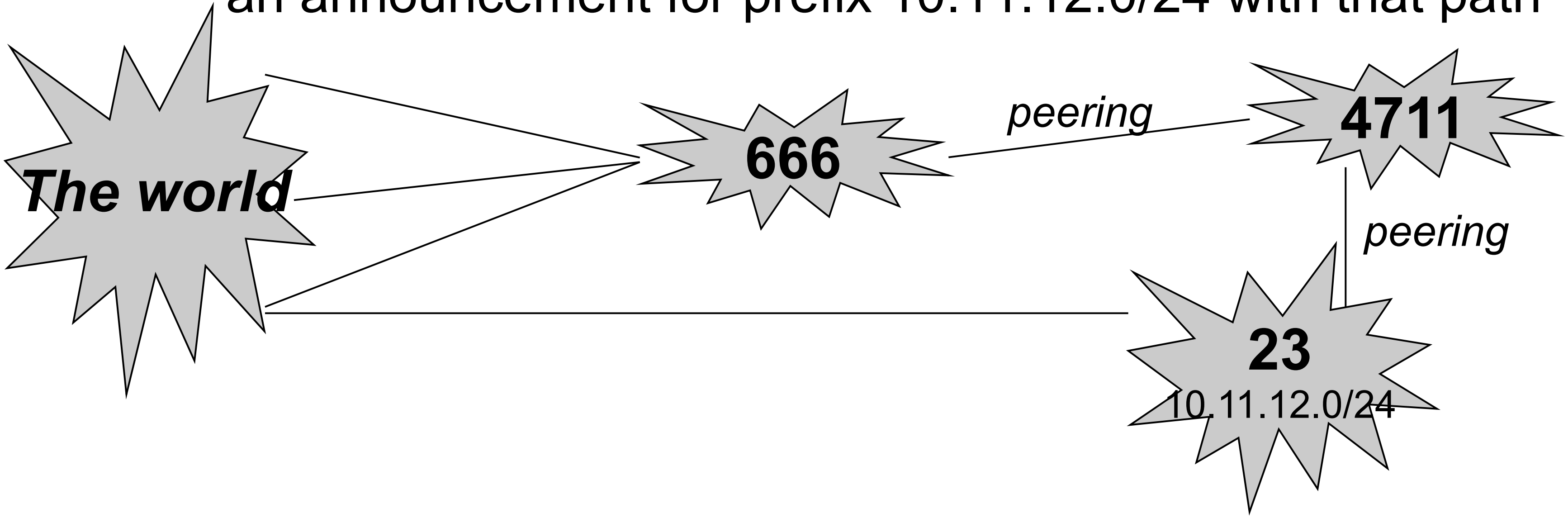
BGP security: Suggested mechanisms (1)

- **Origin authentication:** Only ASes that “own” a prefix can announce it
  - Can secure this cryptographically (PKI)
  - Can we outsmart this?
    - Let 10.11.12.0/24, owned by AS23, be the prefix to be hijacked
    - Rogue AS 666 can lie by announcing non-existent paths: Prefix: 10.11.12.0/24, AS path: 666 23



BGP security: Suggested mechanisms (2)

- **Secure origin authentication:** Only paths that physically exist can announce it
  - Cryptographically secured path database
  - Can we outsmart this?
    - Can announce paths that we should not see
    - Rogue AS666 knows paths 23–4711 and 4711–666 exist
    - Can announce 66 4711 23, even though it never received an announcement for prefix 10.11.12.0/24 with that path



S-BGP

- Secure BGP (S-BGP)
  - Discussed in Interdomain Routing (IDR) Working Group
  - draft-clynn-s-bgp-protocol-01.txt, June 2003
  - c.f. http://www.ir.bbn.com/sbgp/
- Three security mechanisms
  - Secure origin authentication using a Public Key Infrastructure (PKI)
  - Additional attribute ("attestations") allows to carry signatures of routing information in a BGP UPDATE
  - IPsec protects updates, providing data and sequence integrity and router authentication
- Can we outsmart this?
  - Rogue AS666 can still announce a “good” route but then actually use a “bad” route – or even drop the traffic

BGP security: Further reading

- Renesys blog:
  - Posts with ‘security’ tag: www.renesys.com/blog/security/
  - Entry “Reckless driving on the Internet” http://www.renesys.com/blog/2009/02/the-flap-heard-around-the-worl.shtml
  - Entry “Longer is not always better” http://www.renesys.com/blog/2009/02/Longer is not always better.shtml
  - Entry “Pakistan hijacks YouTube” http://www.renesys.com/blog/2008/02/pakistan-hijacks-youtube-1.shtml
  - Entries that match “Syria”
- Butler, Farley, McDaniel, Rexford:
  A survey of BGP security issues and solutions
  Proceedings of the IEEE, January 2010
  http://ix.cs.uoregon.edu/~butler/pubs/bgpsurvey.pdf
- Goldberg, Schapira, Hummon, Rexford:
  How secure are secure interdomain routing protocols?
  Proceedings of ACM SIGCOMM, August 2010
  http://dl.acm.org/citation.cfm?id=1851195

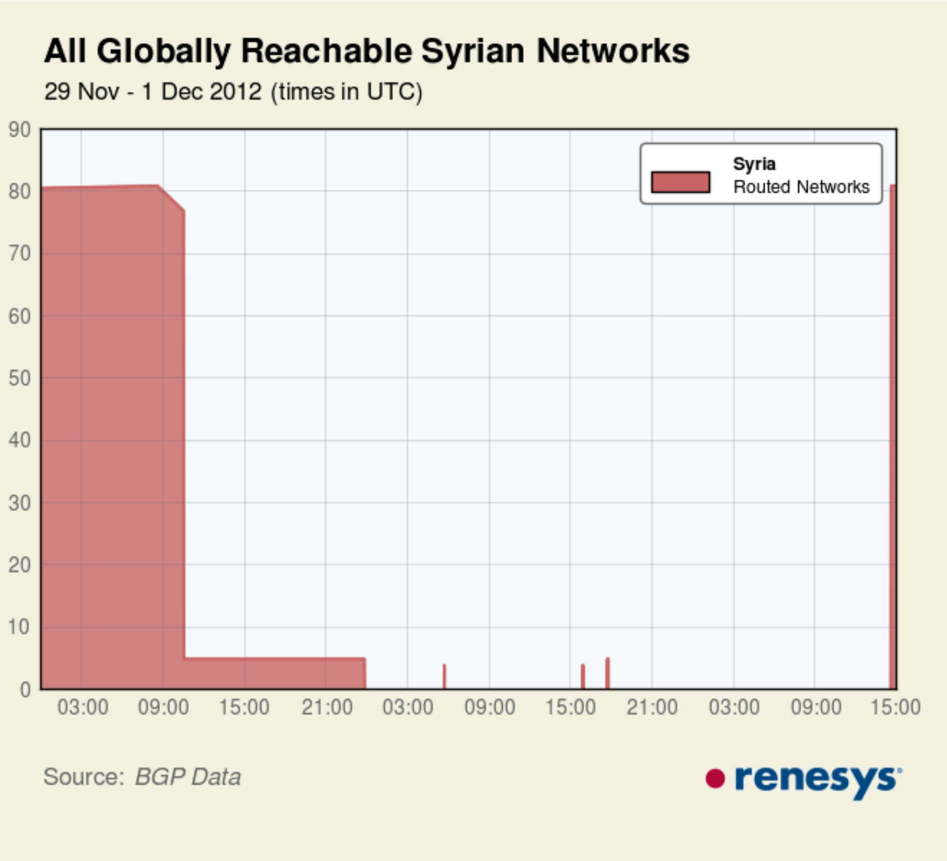
Syrian Internet Connectivity

- Observations
  - On Thursday November 29, 2012 (10:26 UTC), Syria's international Internet connectivity was disrupted: all 84 of Syria's IP address blocks (Syrian Telecommunications Establishment AS with its customer networks) became unreachable
  - 5 networks of Syrian-registered IP space stayed reachable via Tata Communications AS routes until November 30, 01:00 UTC, then became unreachable
  - Restoration of Syrian Internet on December 1 (14:32 UTC)
  - Transit providers: Telecom Italia, Tata Communications, Turk Telecom, and PCCW
- Renesys blog:

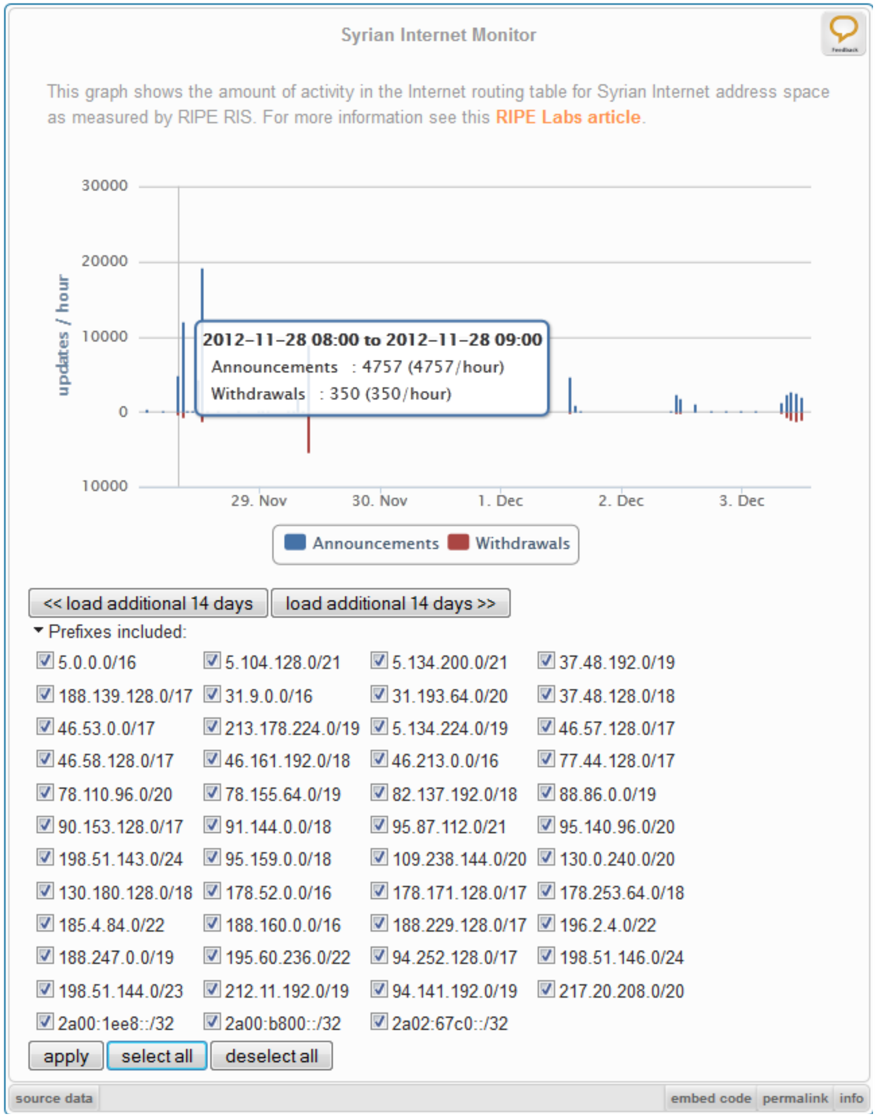
  http://www.renesys.com/blog/2012/11/syria-off-the-air.shtml

  http://www.renesys.com/blog/2012/12/restoration-in-syria-1.shtml
- https://labs.ripe.net/Members/emileaben/monitor-syrian-blackout-with-ripestat

Syrian Internet Connectivity



Source: *BGP Data*



Syrian Internet Monitor

This graph shows the amount of activity in the Internet routing table for Syrian Internet address space as measured by RIPE RIS. For more information see this RIPE Labs article.

2012-11-28 08:00 to 2012-11-28 09:00
Announcements : 4757 (4757/hour)
Withdrawals : 350 (350/hour)

Prefixes included:

☑ 5.0.0.0/16 ☑ 5.104.128.0/21 ☑ 5.134.200.0/21 ☑ 37.48.192.0/19
☑ 188.139.128.0/17 ☑ 31.9.0.0/16 ☑ 31.193.64.0/20 ☑ 37.48.128.0/18
☑ 46.53.0.0/17 ☑ 213.178.224.0/19 ☑ 5.134.224.0/19 ☑ 46.57.128.0/17
☑ 46.58.128.0/17 ☑ 46.161.192.0/18 ☑ 46.213.0.0/16 ☑ 77.44.128.0/17
☑ 78.110.96.0/20 ☑ 78.155.64.0/19 ☑ 82.137.192.0/18 ☑ 88.86.0.0/19
☑ 90.153.128.0/17 ☑ 91.144.0.0/18 ☑ 95.87.112.0/21 ☑ 95.140.96.0/20
☑ 198.51.143.0/24 ☑ 95.159.0.0/18 ☑ 109.238.144.0/20 ☑ 130.0.240.0/20
☑ 130.180.128.0/18 ☑ 178.52.0.0/16 ☑ 178.171.128.0/17 ☑ 178.253.64.0/18
☑ 185.4.84.0/22 ☑ 188.160.0.0/16 ☑ 188.229.128.0/17 ☑ 196.2.4.0/22
☑ 188.247.0.0/19 ☑ 195.60.236.0/22 ☑ 94.252.128.0/17 ☑ 198.51.146.0/24
☑ 198.51.144.0/23 ☑ 212.11.192.0/19 ☑ 94.141.192.0/19 ☑ 217.20.208.0/20
☑ 2a00:1ee8::/32 ☑ 2a00:b800::/32 ☑ 2a02:67c0::/32

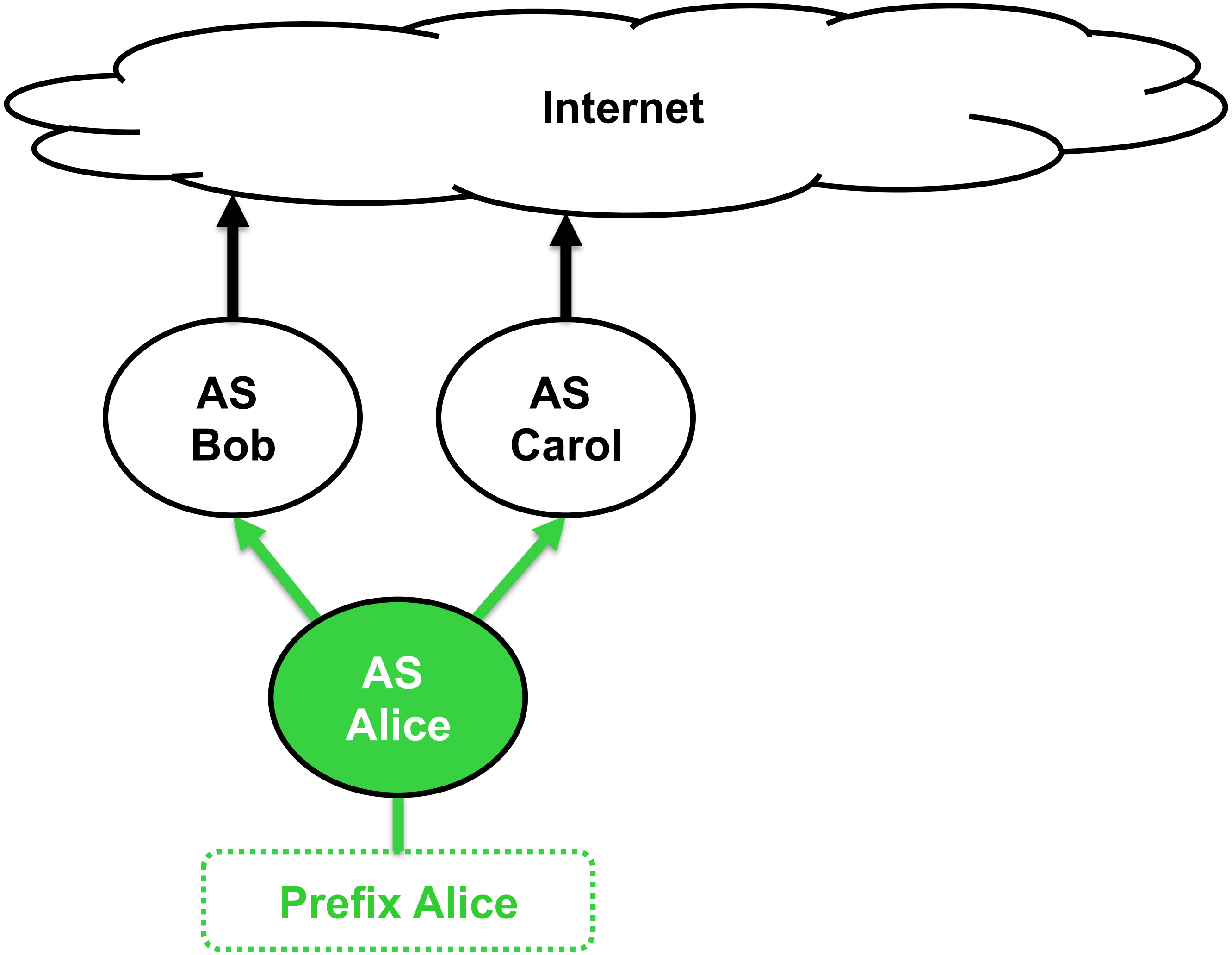
http://www.renesys.com/

https://labs.ripe.net

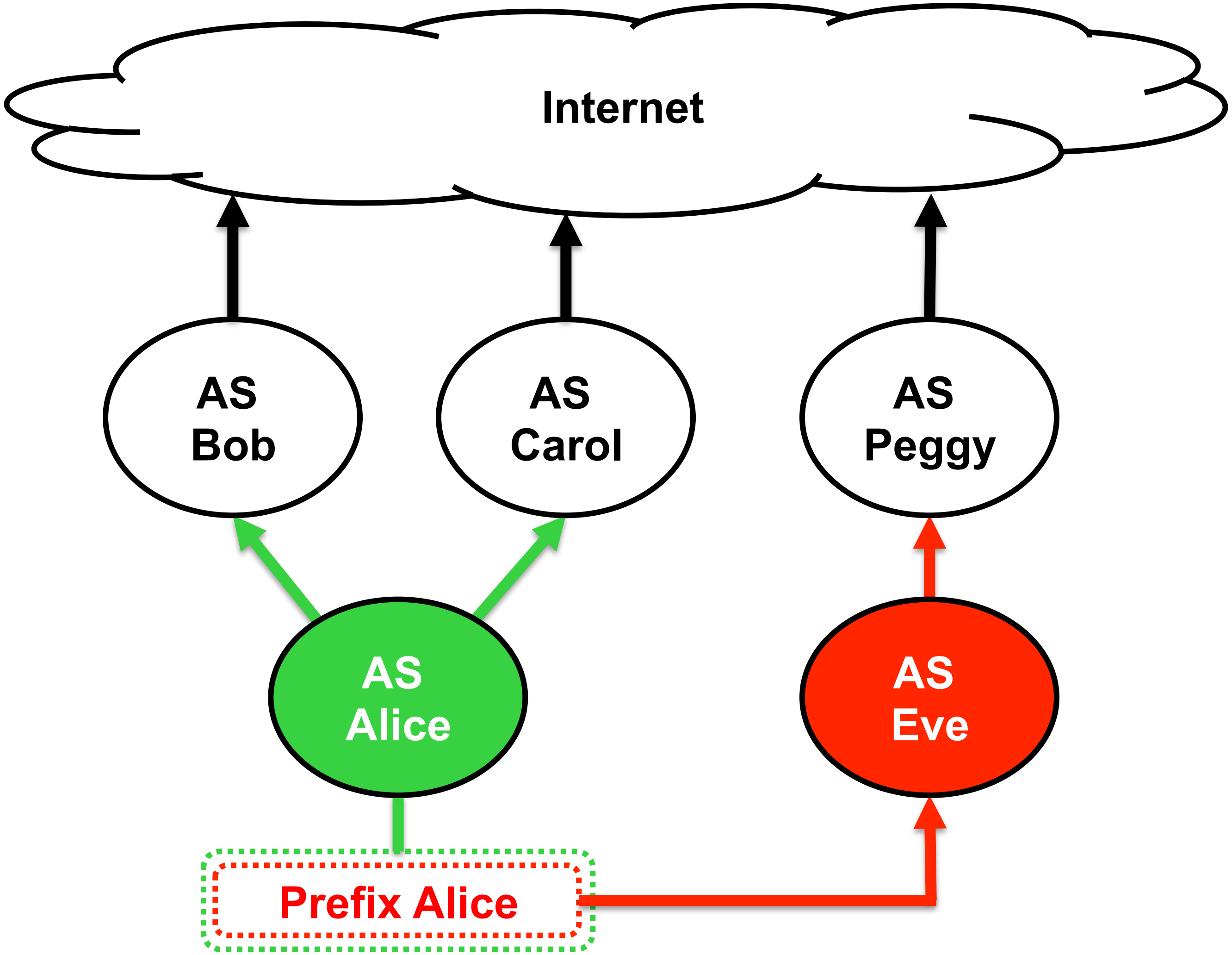
Threats to Robust Routing

- Prefix hijacking
  - Malicious AS announces prefix it does not own
  - Symptoms
    - Depend on position in global Internet Topology
  - Prevention
    - BGP Security (S-BGP, soBGP, psBGP, BGPSec)
    - Cryptographic means for Route Origin Authorisation (ROA)
    - BGPSec
      - c.f. Secure Interdomain Routing Working Group
      - RPKI ROA infrastructure
      - Resource Public Key Infrastructure (RPKI)
        RFC 6480: M. Lepinski and S. Kent,
        An Infrastructure to Support Secure Internet Routing

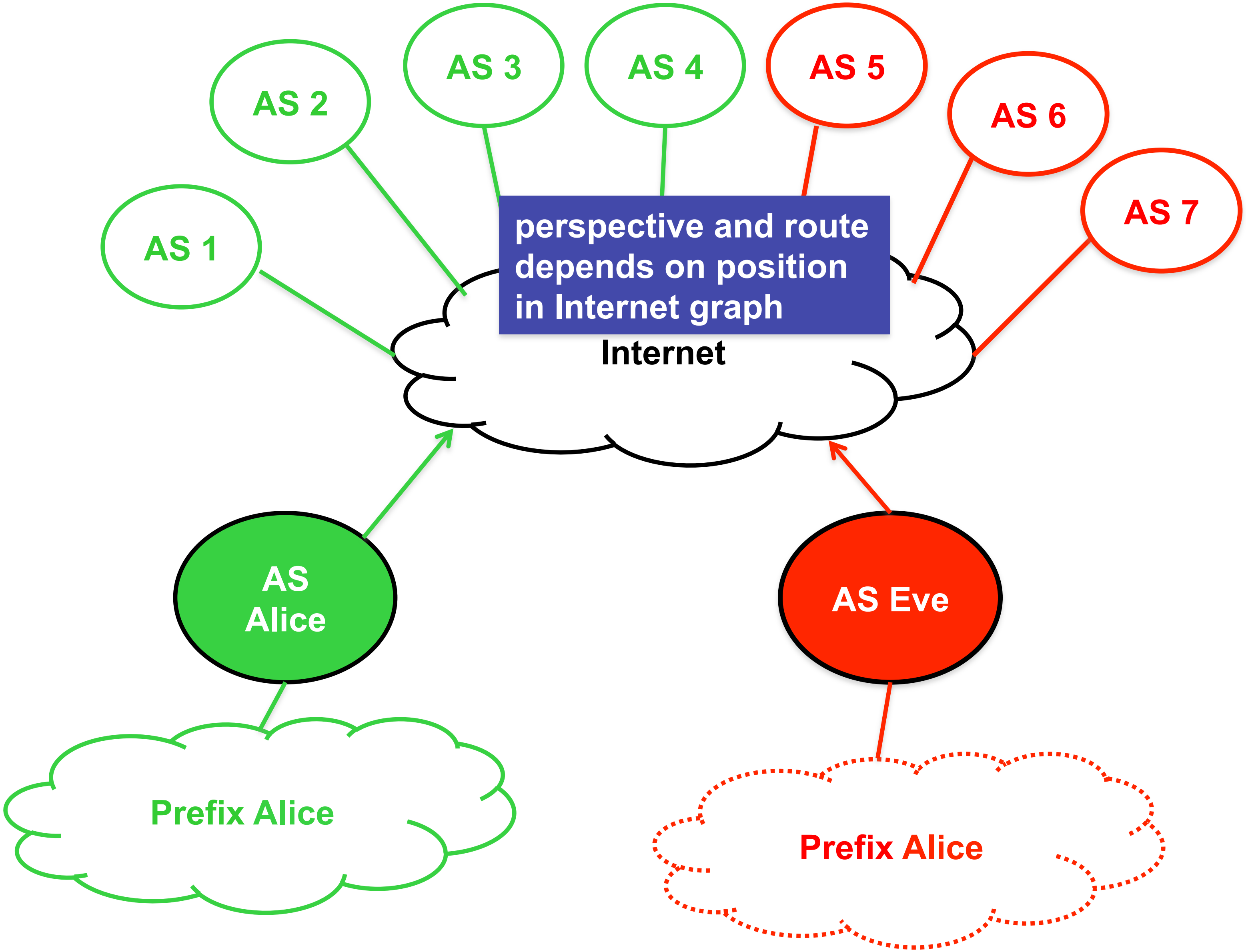
Prefix Announcement



Prefix Hijacking

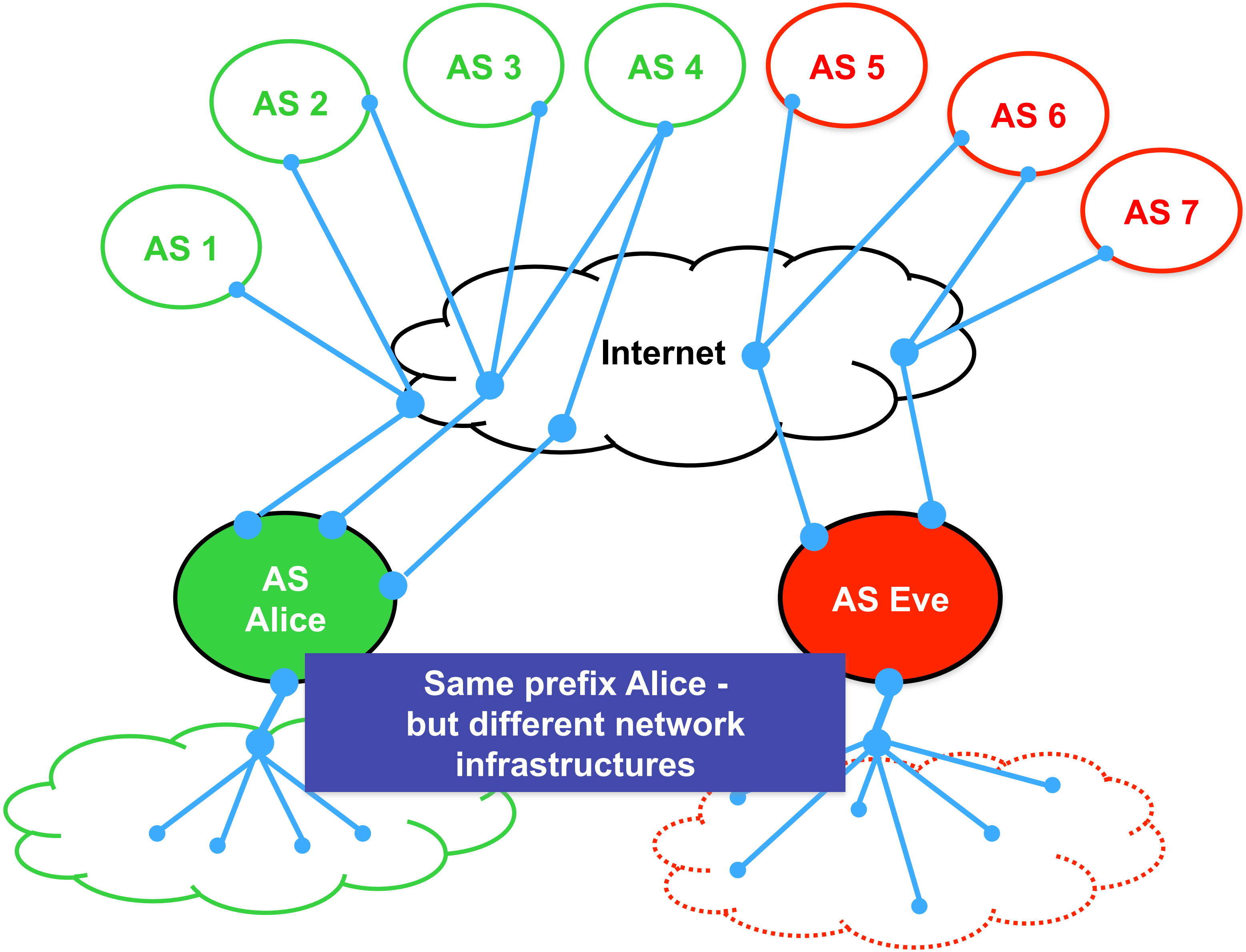


Prefix Hijacking



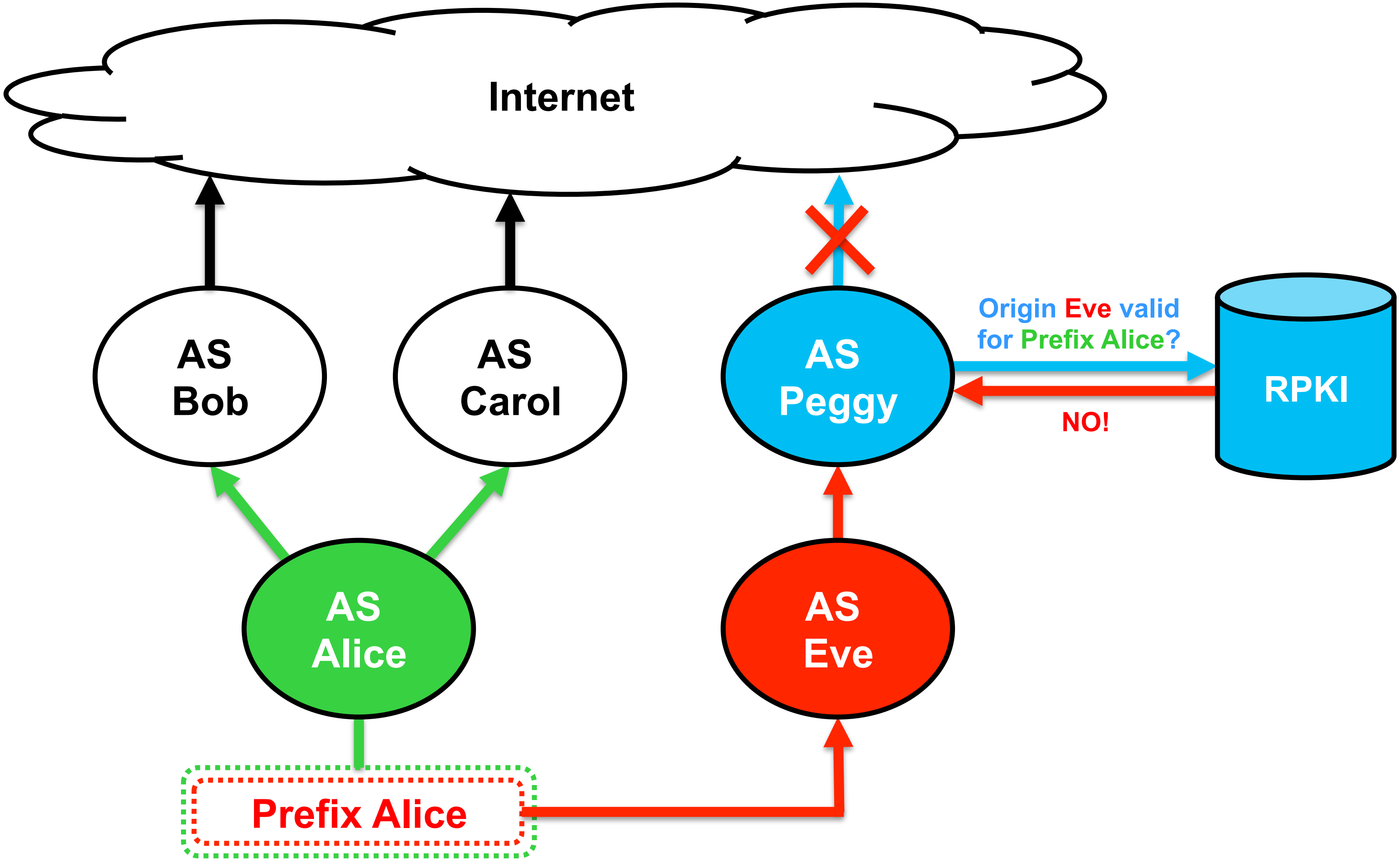
perspective and route depends on position in Internet graph

Detection of Prefix Hijacking



Same prefix Alice -
but different network
infrastructures

RPKI - Resource Public Key Infrastructure



Exkursus: KLIK Team

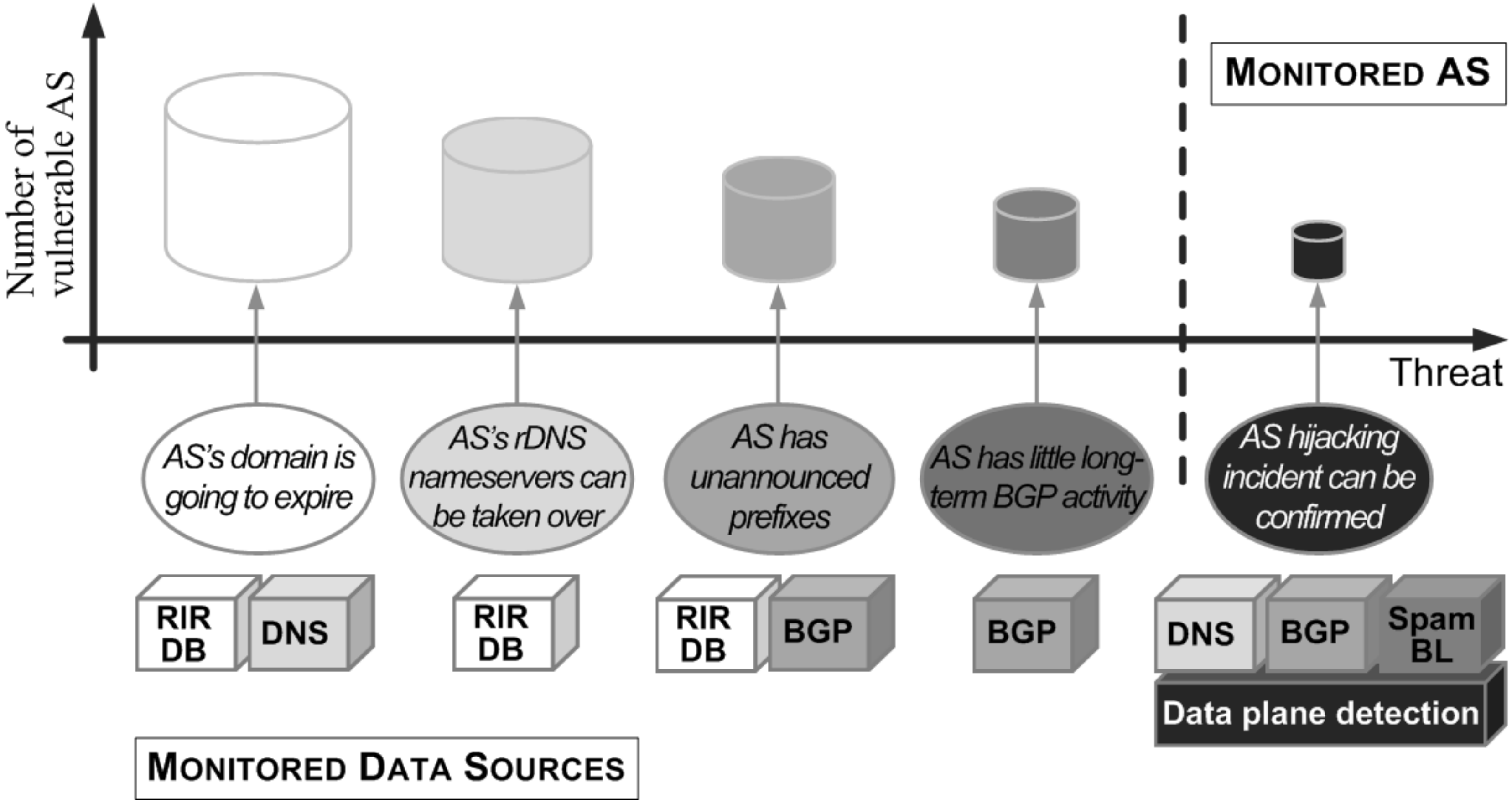
- **KLIK Team**
  annual party
- Gifts
  - Macbooks
  - Briefcase with money
  - Car

Autonomous System Hijacking

- AS hijacking
  - Attacker claims ownership of whole autonomous system and its prefixes
  - Best current praxis
    - Transit providers install prefix filters to protect against wrong routes received from BGP-speaking customers
    - Transit provides install prefix filters towards peers
    - Transit providers request Letter Of Authorisation (LOA) from ISPs who want to propagate their customers' routes
    - LOA comes from customer, and confirms that ISP is authorised to announce routes on their behalf
  - AS hijacking attack
    - Establishing fraudulent business relationship with upstream provider
      - Forged „Letter of authorisation“(LOA)
      - Electronic payment

Autonomous System Hijacking Early Warning

- Observation of DNS Expiry and new domain registrations
- Analysis of reverse DNS und BGP announcements


Number of vulnerable AS
Threat
MONITORED AS
AS's domain is going to expire
AS's rDNS nameservers can be taken over
AS has unannounced prefixes
AS has little long-term BGP activity
AS hijacking incident can be confirmed
RIR DB
DNS
RIR DB
RIR DB
BGP
BGP
DNS
BGP
Spam BL
Data plane detection
MONITORED DATA SOURCES

Traffic Engineering

Routing: Optimization purposes

- Inter-AS routing
  - Optimality = select route with highest revenue/least loss
  - Mainly policy driven (as we have seen)

- Intra-AS routing
  - Optimality = configure routing such that network can host as much traffic as possible
  - Traffic engineering methods

Traffic Engineering

- Collect traffic statistics: Traffic Matrix
  - How much traffic is flowing from A to B?
  - Often difficult to measure!
    - Drains router performance
    - Therefore often estimated – active research area
    - Alternative: Build lots of MPLS tunnels, measure each tunnel
- Optimize routing
  - E.g., calculate good choice of OSPF weights
  - Typical goal: minimize maximum link load in entire network; keep average link load below 50% or 70%
    - (Why? Fractal TCP traffic leads to spikes.)
- Deploy new routing
  - Performance may deteriorate during update
  - E.g., routing loops during OSPF convergence

Dynamic Traffic Engineering

Why static? Why don’t we do it dynamically?

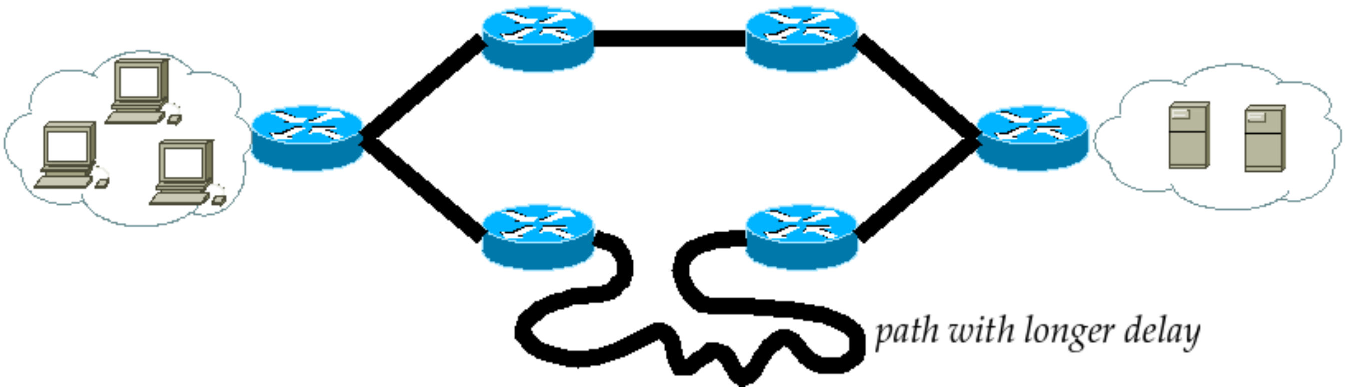
- ❑ Prone to oscillations and chaotic behaviour
  - Bad experiences in the ARPANET
  - Ex.: Route A congested, route B free
    → Everyone switches from A to B
    → Route A free, route B congested → …
- ❑ Routing loops during convergence → packet losses
- ❑ Packet reordering:
  - Packet P1 arrives later than Packet P2
  - TCP will think that P1 got lost! ⇒ congestion control!
- ❑ Actually, a difficult problem
  - Stale information
  - Interaction with TCP congestion control
  - Interaction with dynamic TE mechanisms in other ASes
- ❑ Thus: Congestion control in end hosts (TCP), usually not in network

Multipath routing

- Routing = finding best-cost route
- But: What if more than one best route exists?
- Some routing protocols allow Equal-Cost Multipath (ECMP) routing, e.g., OSPF
  - ≥ 2 routes of same cost exist to destination prefix?
    → Evenly distribute traffic across these routes

Multipath Routing: TCP Problem

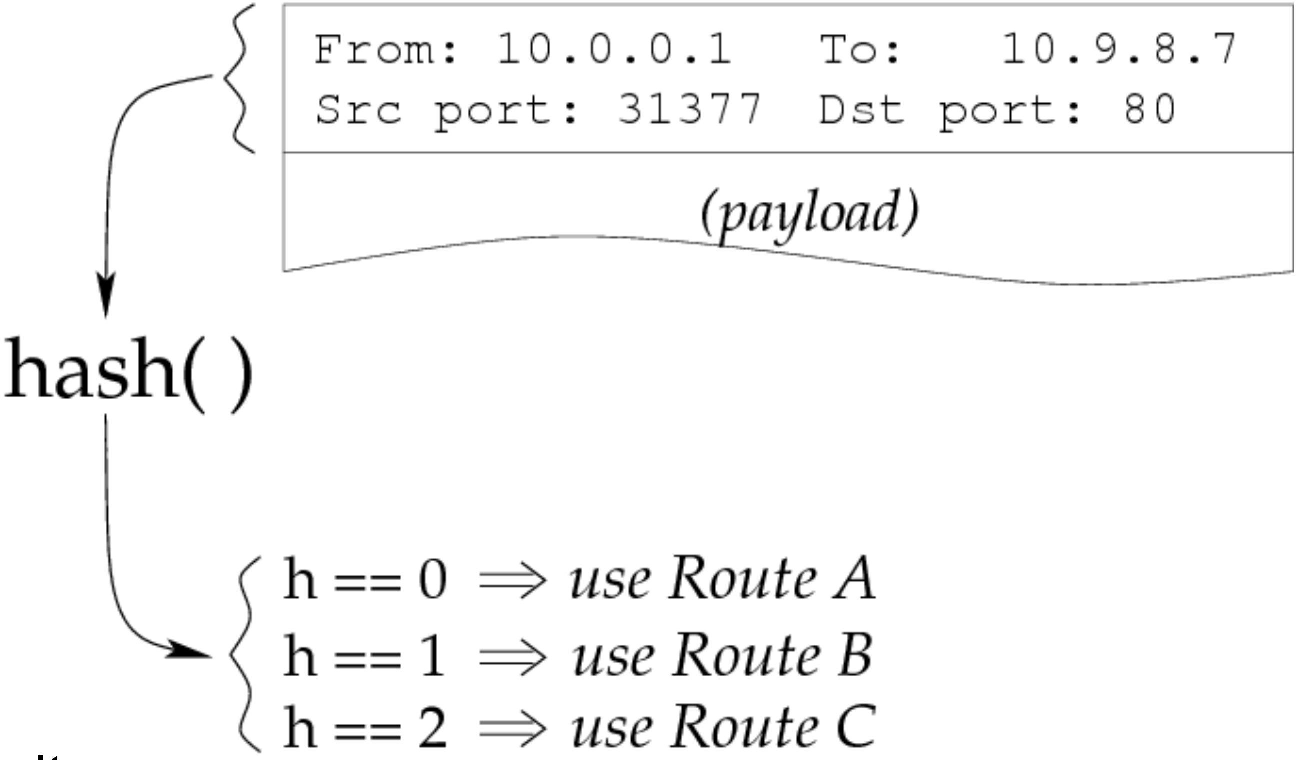
- How to distribute traffic? Naïve approaches:
  - Round-robin
  - Distribute randomly
- Equal cost does not mean equal latency:



- Problem with TCP = Packet reordering!
  - Packets sent: P1, P2
  - Packets received: P2, P1
  - Receiver receives P2 → believes P1 to be lost → triggers congestion control mechanisms → performance degrades

Multipath routing: Solution

- Hash "randomly"…
- …but use packet headers as "random" values:



```
From: 10.0.0.1    To:    10.9.8.7
Src port: 31377   Dst port: 80
```

(payload)

hash( )

h == 0 ⇒ *use Route A*

h == 1 ⇒ *use Route B*

h == 2 ⇒ *use Route C*

- Result:
  - Packets from same TCP connection yield same hash value
  - No reordering within one TCP connection possible