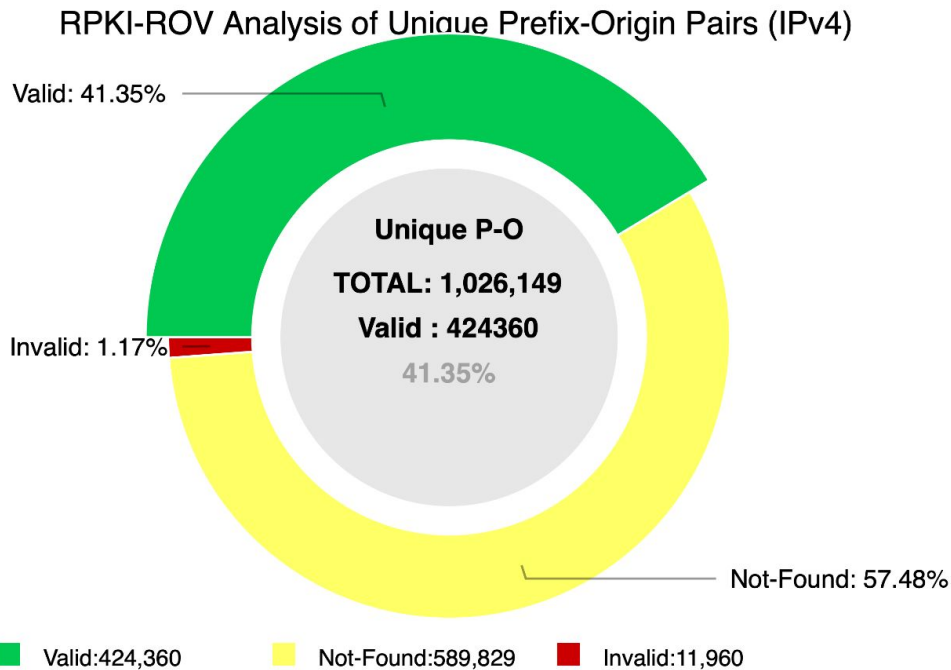


BGP security today

- Resource Public Key Infrastructure (RPKI)
 - A framework to support improved BGP security:
 - A secure way to map AS numbers to IP prefixes.
 - A distributed repository system for storing and disseminating the mappings.
- RPKI operations
 - RPKI relies on cryptographic certificates (X.509)
 - The certificate infrastructure mimics the way IP prefixes are distributed: from IANA, to Regional Internet Registries (RIR), to end-customers.
 - A Route Origination Authorization (ROA) states which AS is authorized to originate certain IP prefixes.

BGP security today - long way to go



NIST RPKI Monitor: RPKI-ROV Analysis

Protocol: IPv4

RIR: All

Date: 2023-03-02 06:00

BGP is extremely vulnerable

- Several high-profile outages
- Many smaller examples
 - Blackholing a single destination prefix
 - Hijacking unallocated addresses to send spam
- Why isn't it an even bigger deal?
 - Really, most big outages are configuration errors
 - Most bad guys want the Internet to stay up
 - ... so they can send unwanted traffic (e.g., spam, identity theft, denial-of-service attacks, port scans, ...)

BGP is hard to fix

- Complex system
 - Large, with around 65,000 ASes
 - Decentralized control among competitive ASes
 - Core infrastructure that forms the Internet
- Hard to reach agreement on the right solution
 - S-BGP with public key infrastructure, registries, crypto?
 - Who should be in charge of running PKI and registries?
 - Worry about data-plane attacks or just control plane?
- Hard to deploy the solution once you pick it
 - Hard enough to get ASes to apply route filters
 - Now you want them to upgrade to a new protocol
 - ... all at the exact same moment? A “flag day”

Because BGP is based on policy - it is not guaranteed to converge

- ASes are free to choose and advertise any paths they want
 - network stability argues against this
- Guaranteeing the absence of oscillations is hard
 - even when you know all the policies

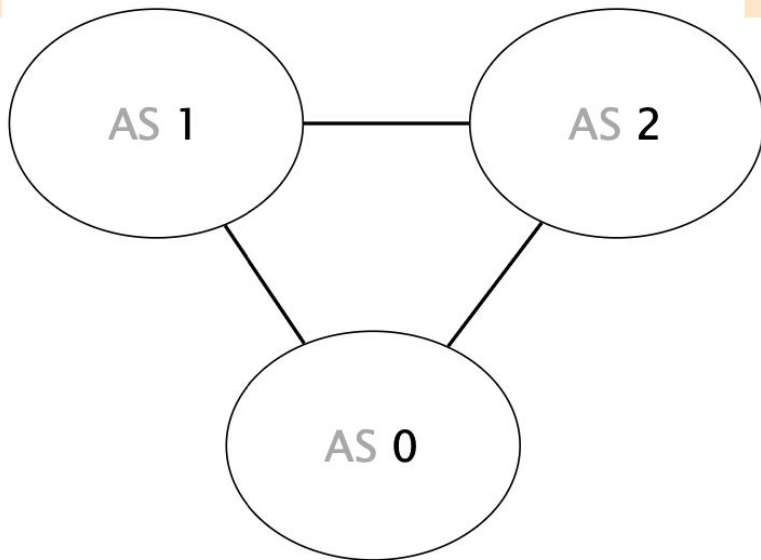
Because of policy, BGP can have multiple stable states

preference list

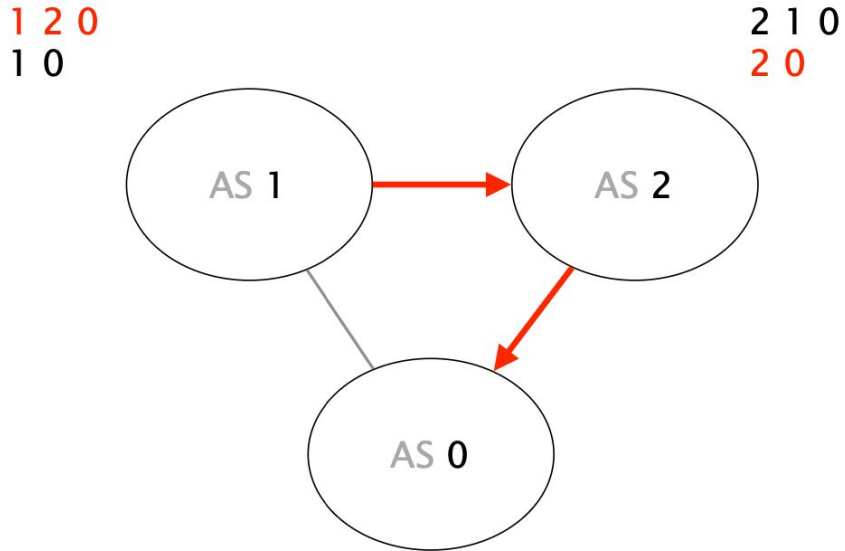
1 prefers to reach 0
via 2 rather than directly

1	2	0
1	0	

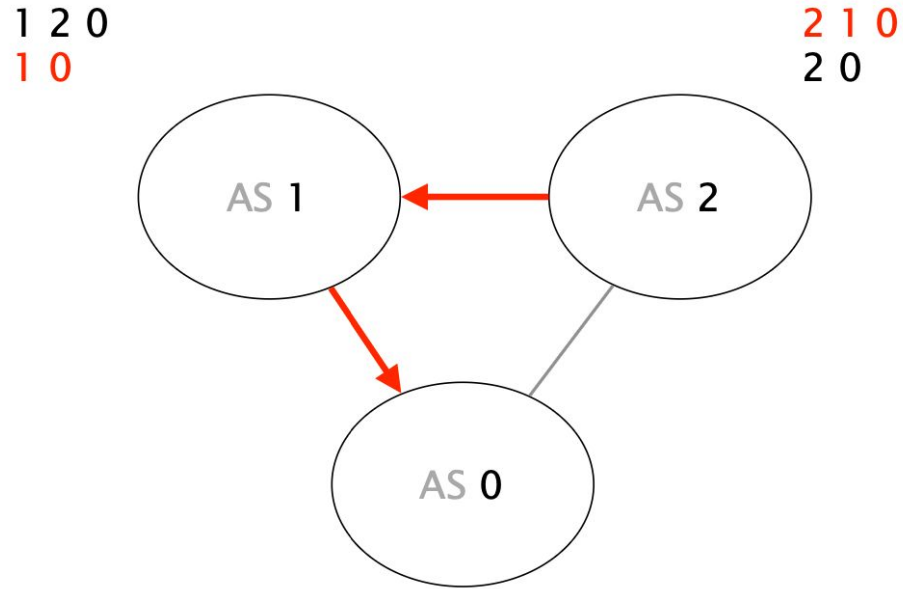
2	1	0
2	0	



If AS2 is the first to advertise [2 0],
the system stabilizes in a state where AS 1 is happy

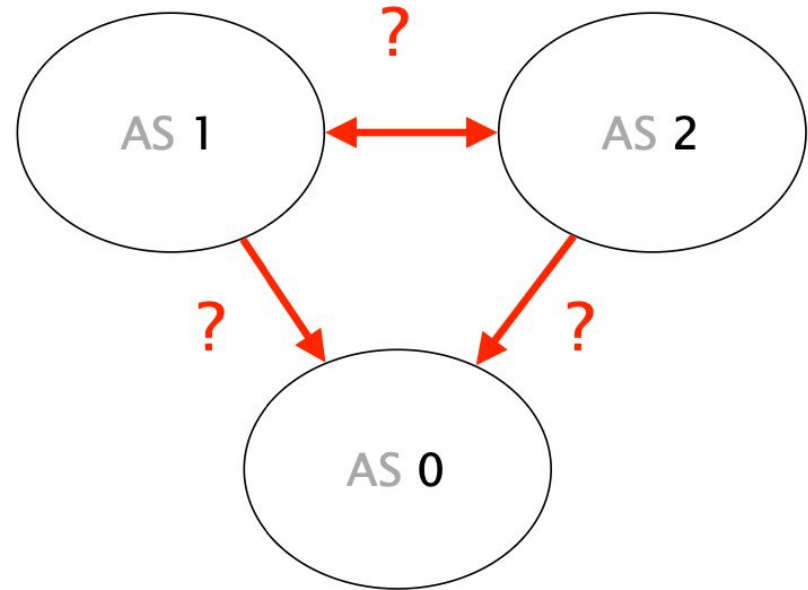


If AS1 is the first one to advertise [1 0],
the system stabilizes in a state where AS 2 is happy

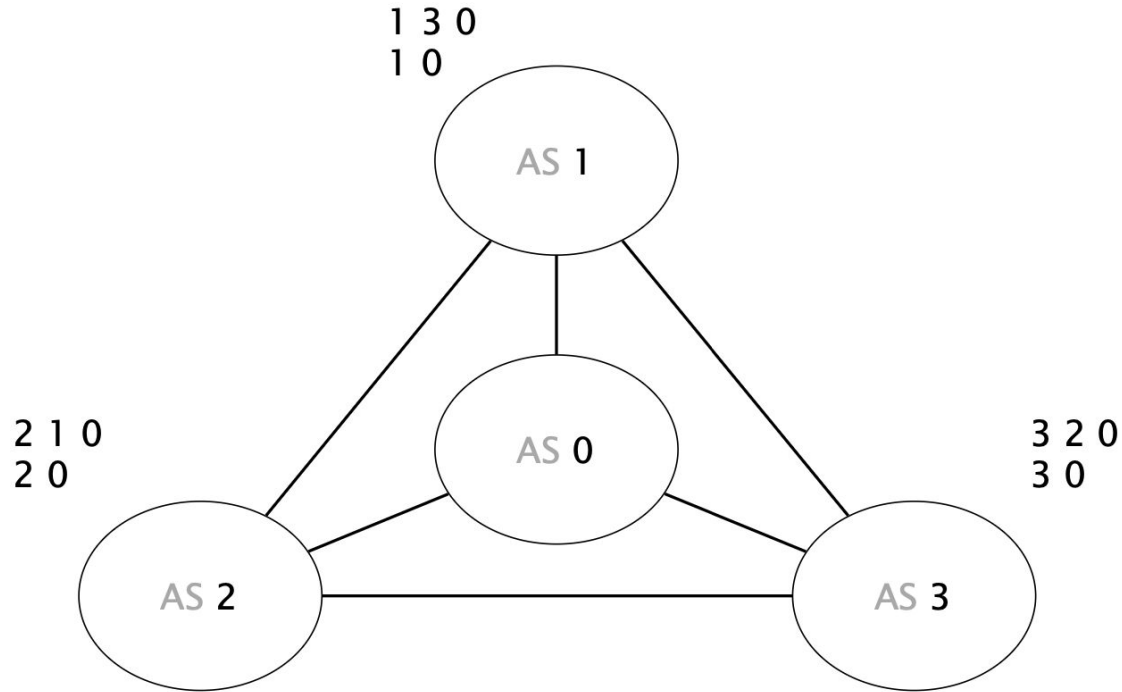


The actual assignment depends on the ordering between the messages

Note that AS1/AS2 could change the outcome by manual intervention



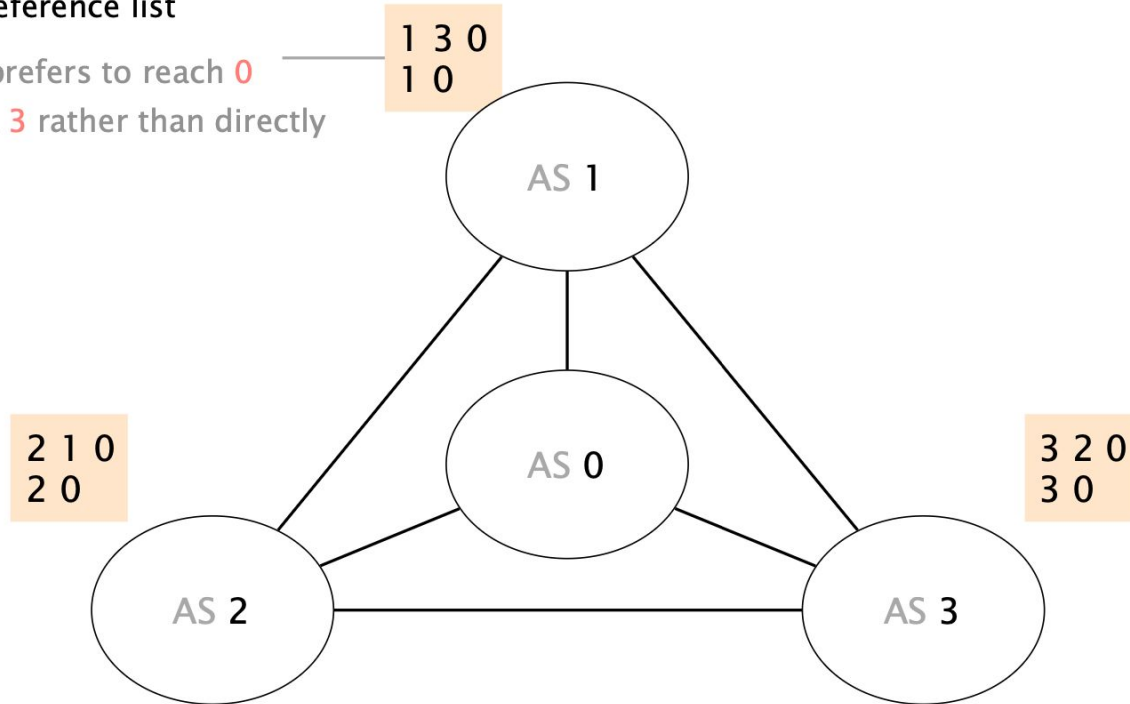
With arbitrary policies, BGP may fail to converge



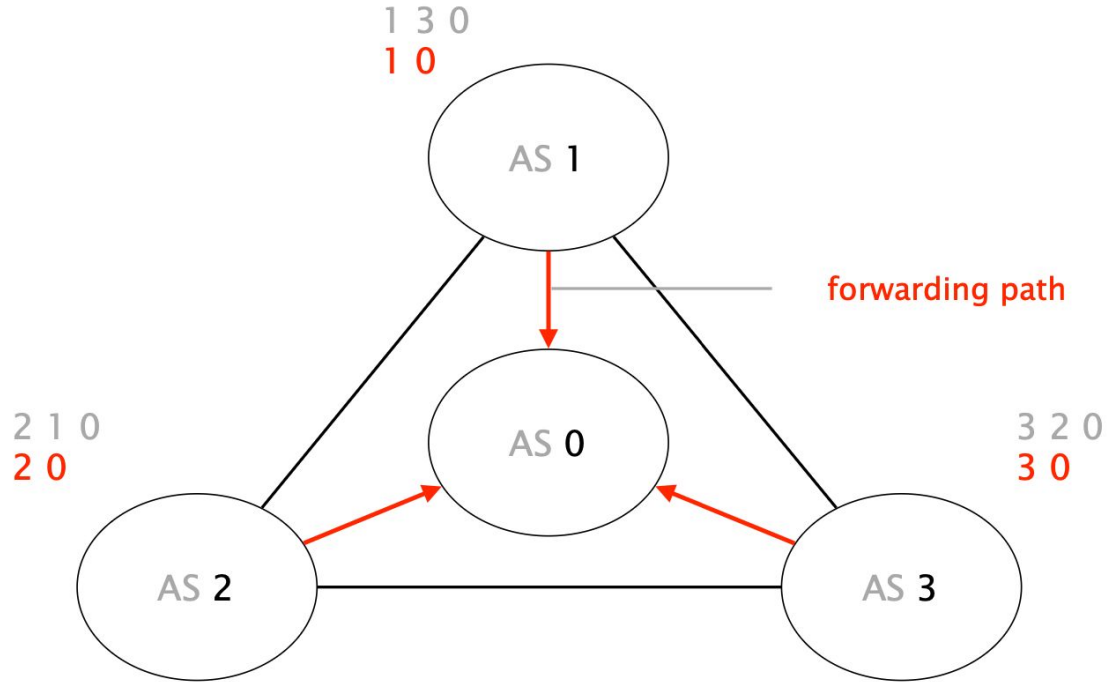
With arbitrary policies, BGP may fail to converge

preference list

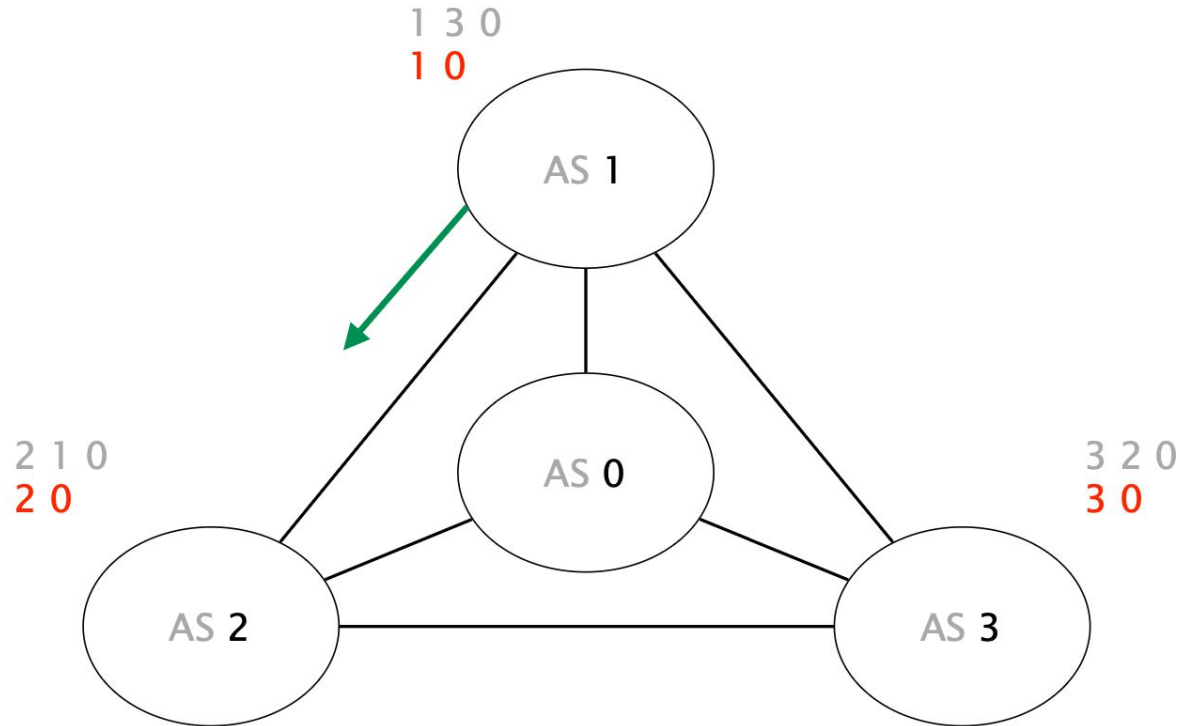
1 prefers to reach 0
via 3 rather than directly



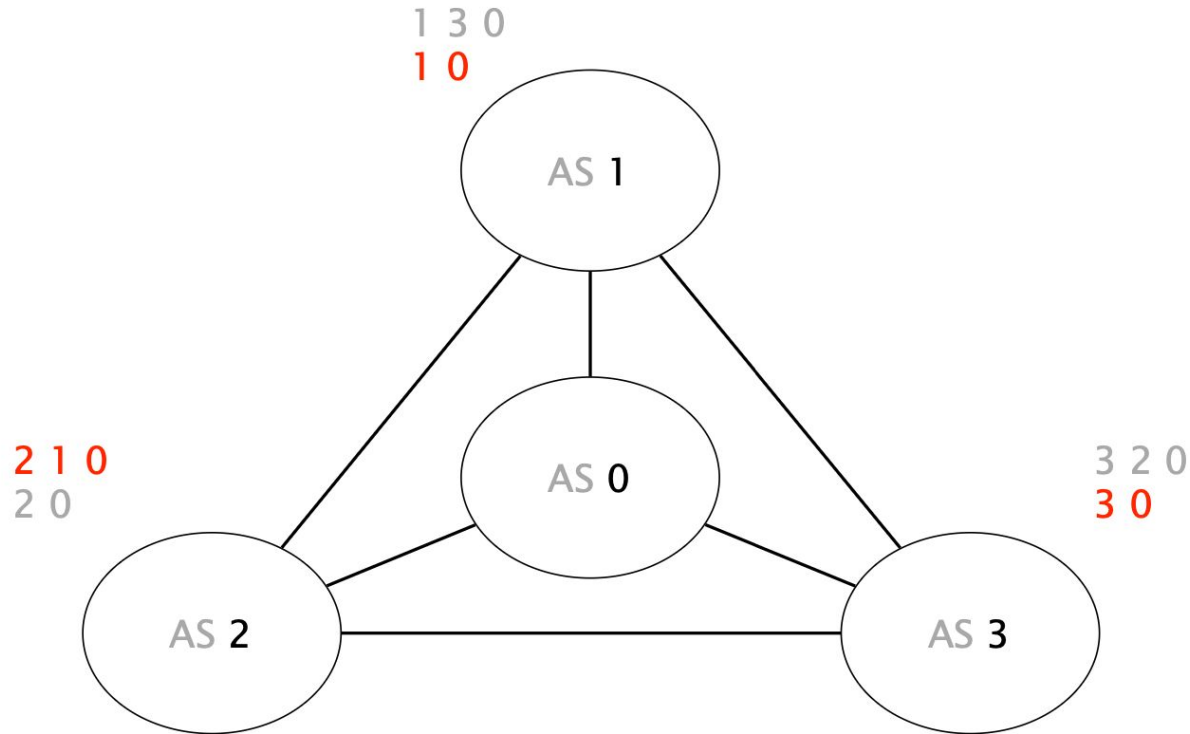
Initially, all ASes only know the direct route to 0



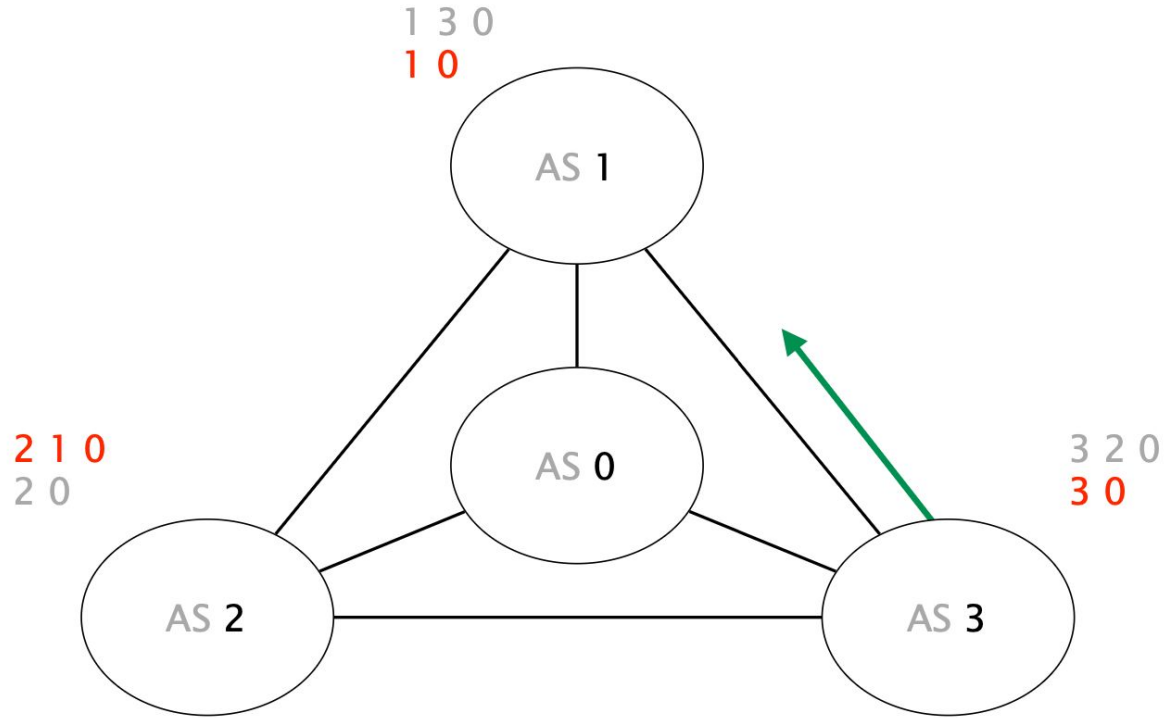
AS 1 advertises its path to AS 2



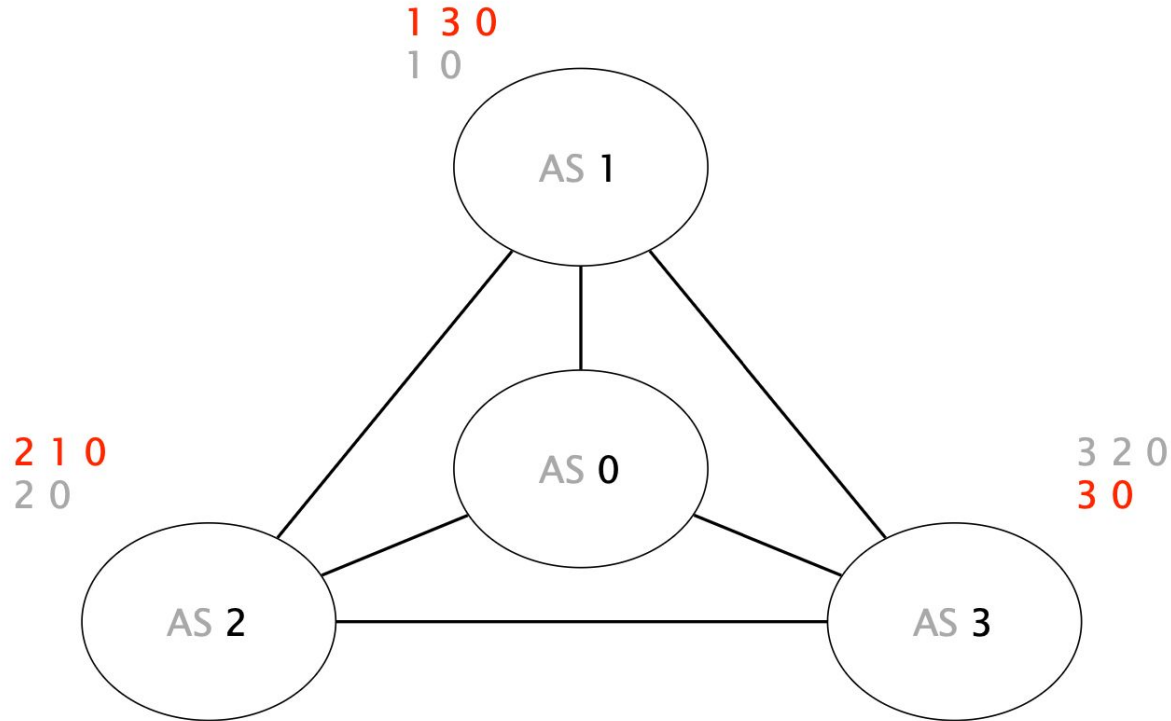
Upon reception, AS 2 switches to 2 1 0 (preferred)



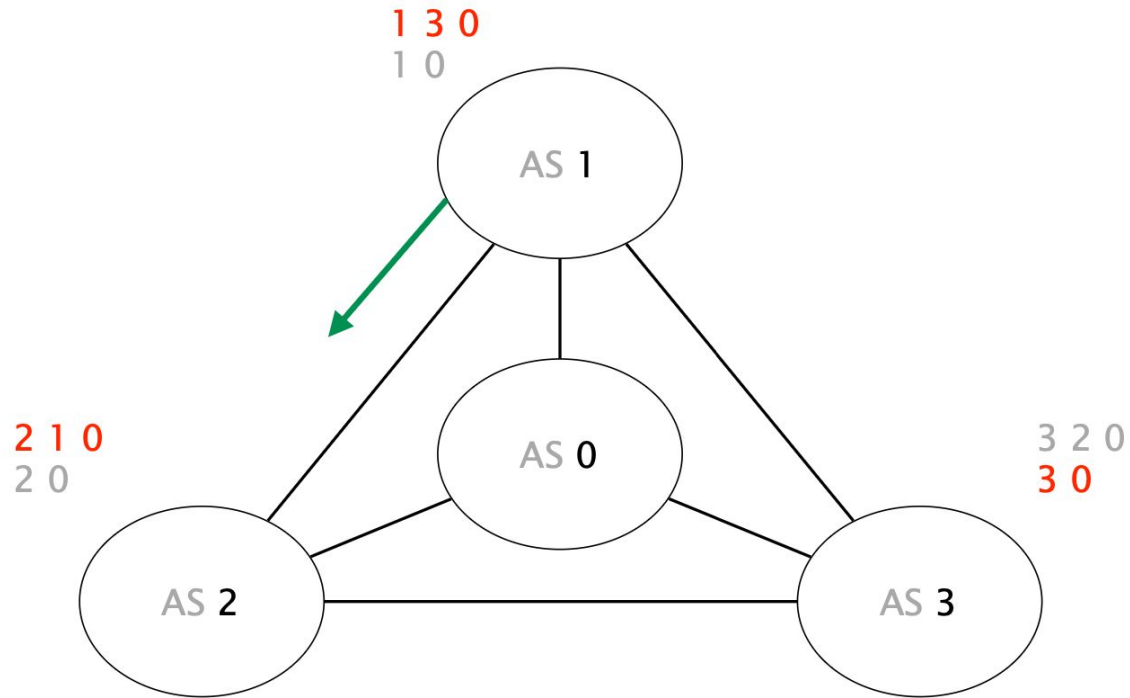
AS 3 advertises its path to AS 1



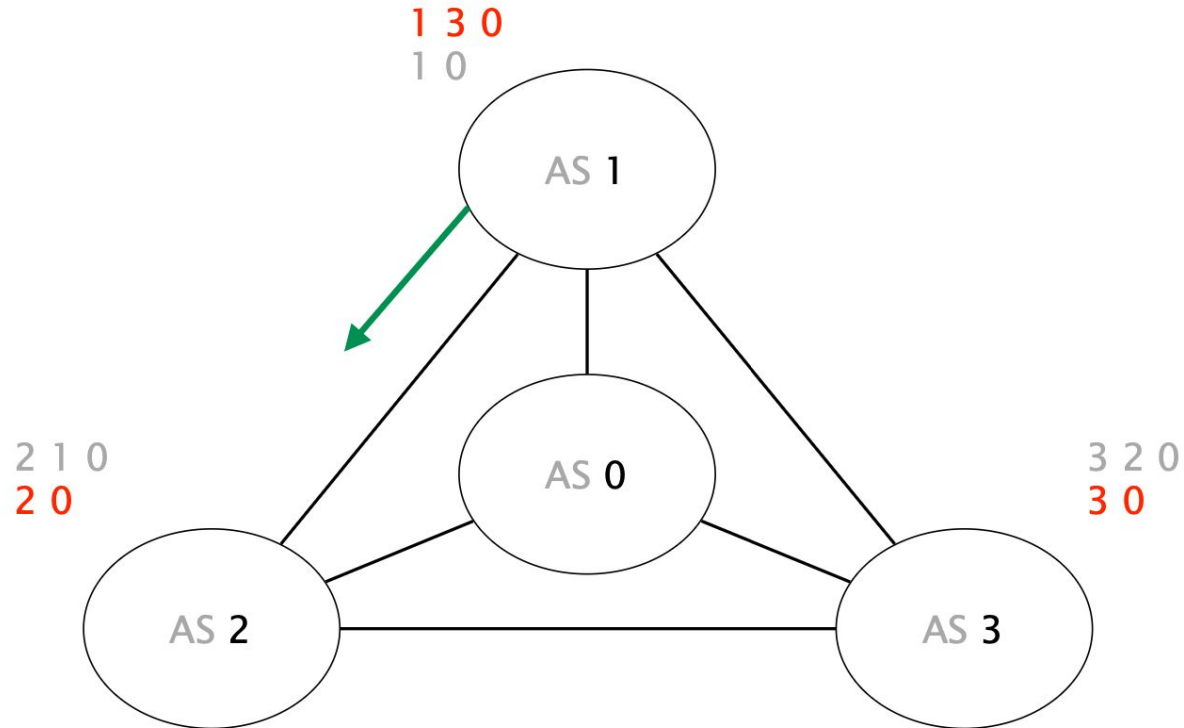
Upon reception, AS 1 switches to 1 3 0 (preferred)



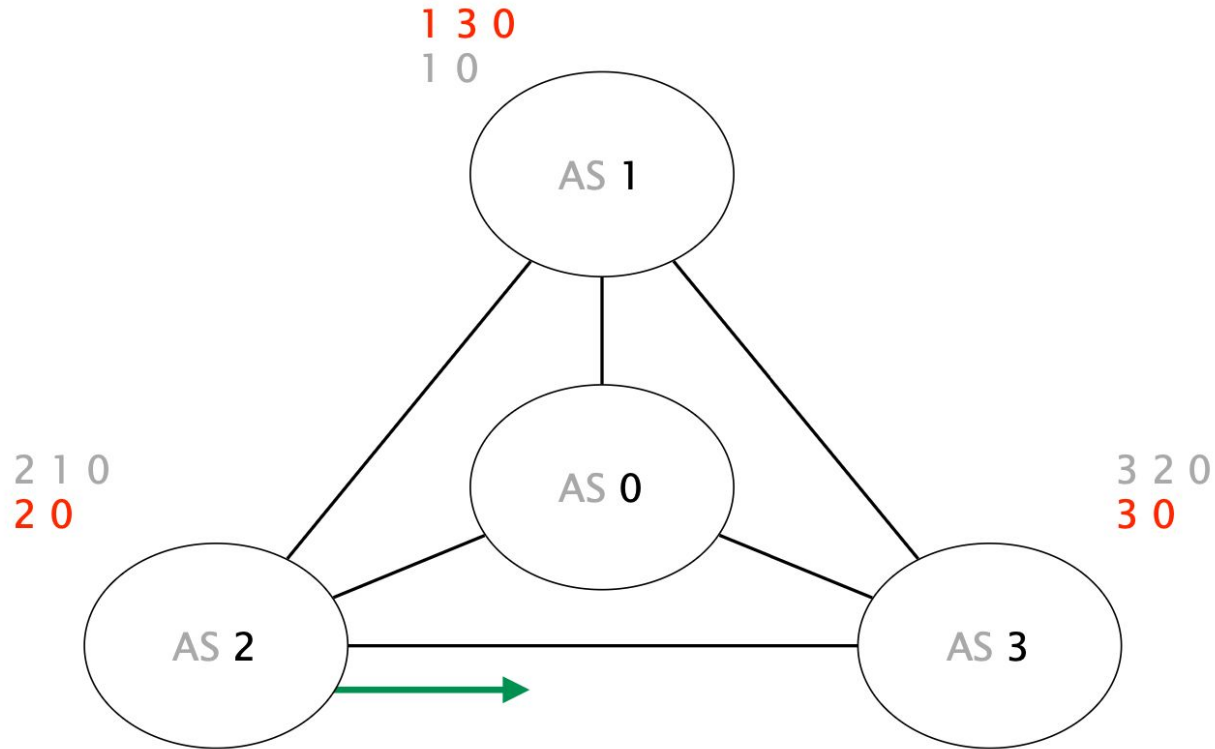
AS 1 advertises its new path 1 3 0 to AS 2



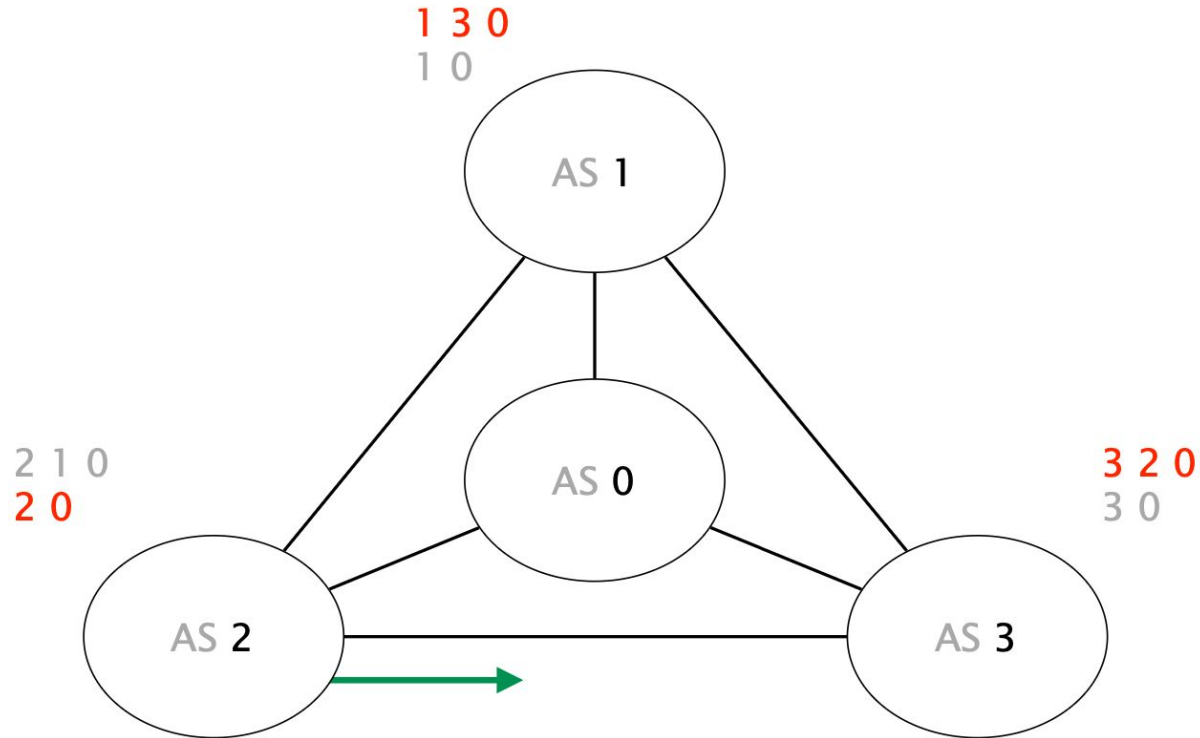
Upon reception, AS 2 reverts back to its initial path 2 0



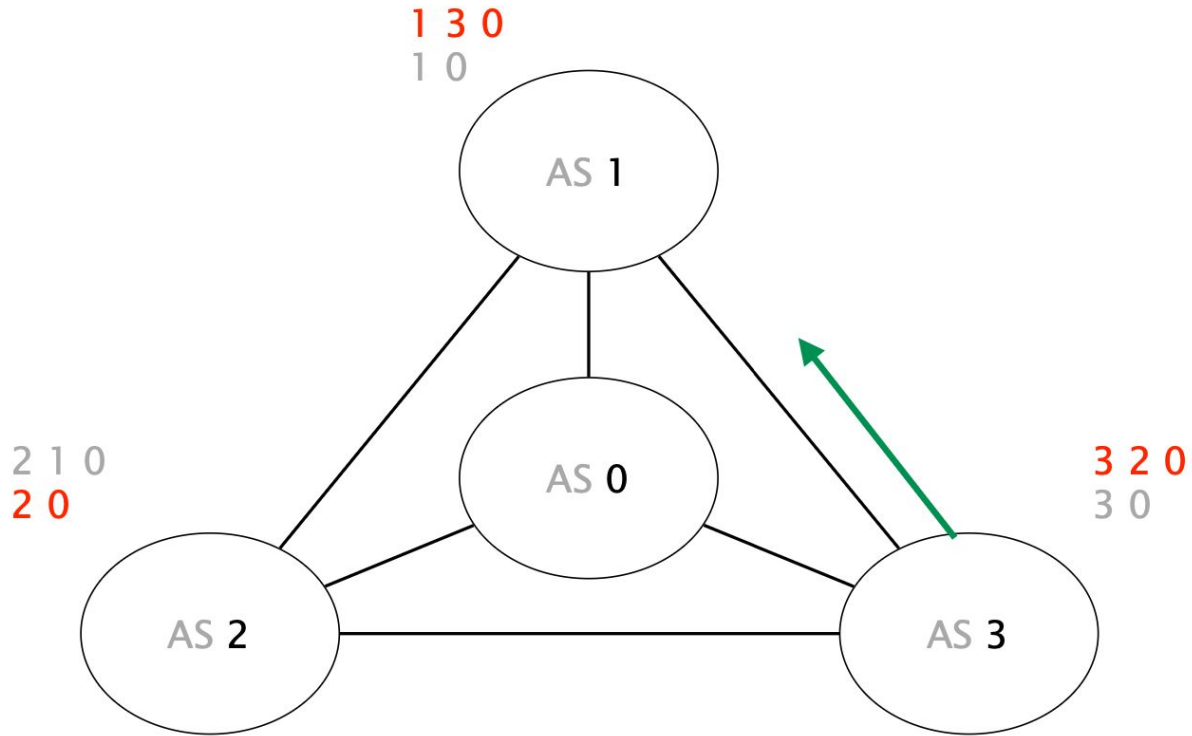
AS 2 advertises its path 2 0 to AS 3



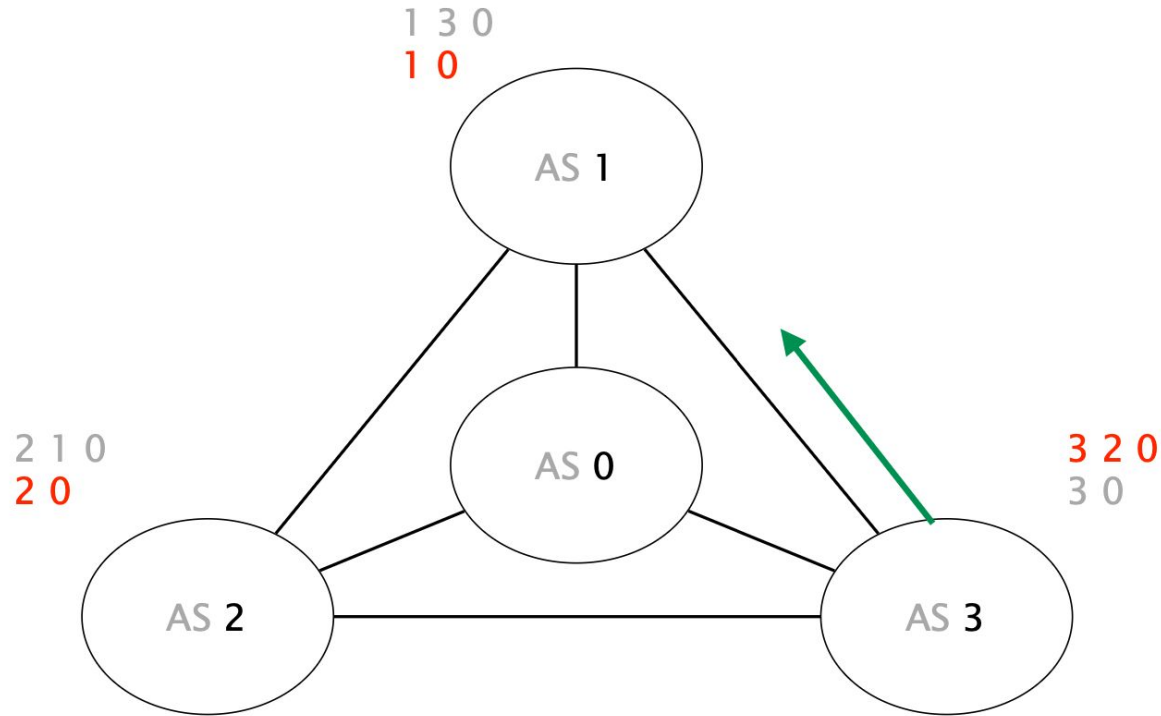
Upon reception, AS 3 switches to 3 2 0 (preferred)



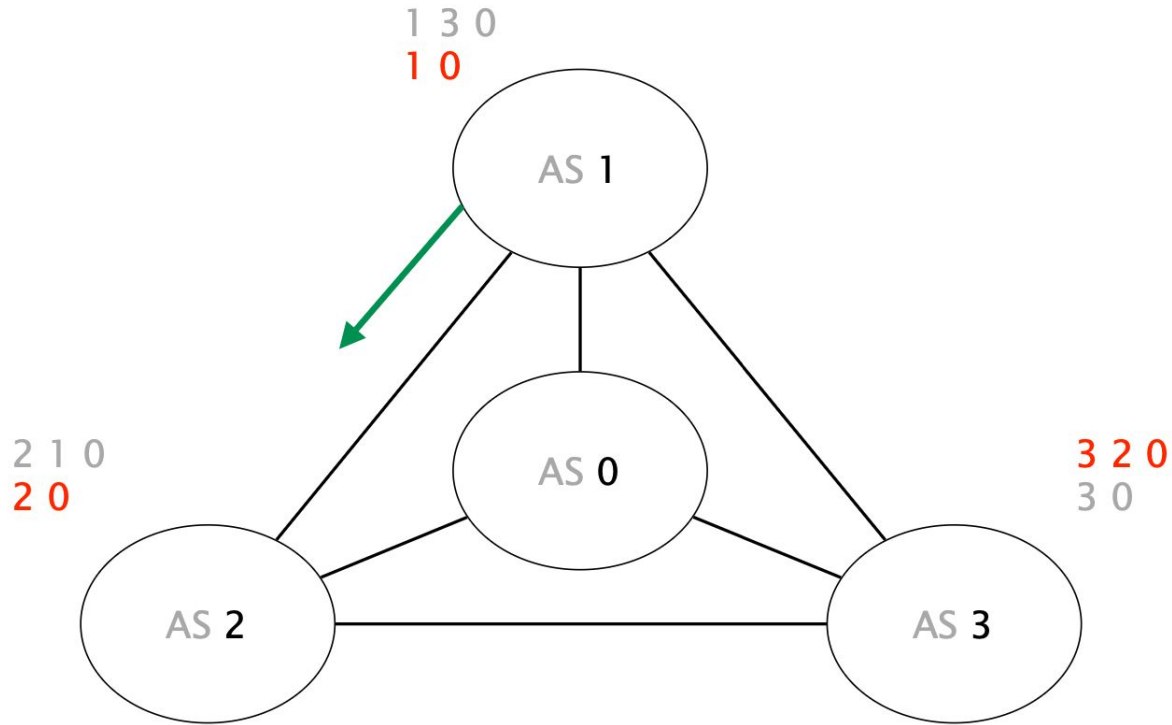
AS 3 advertises its new path 3 2 0 to AS 1



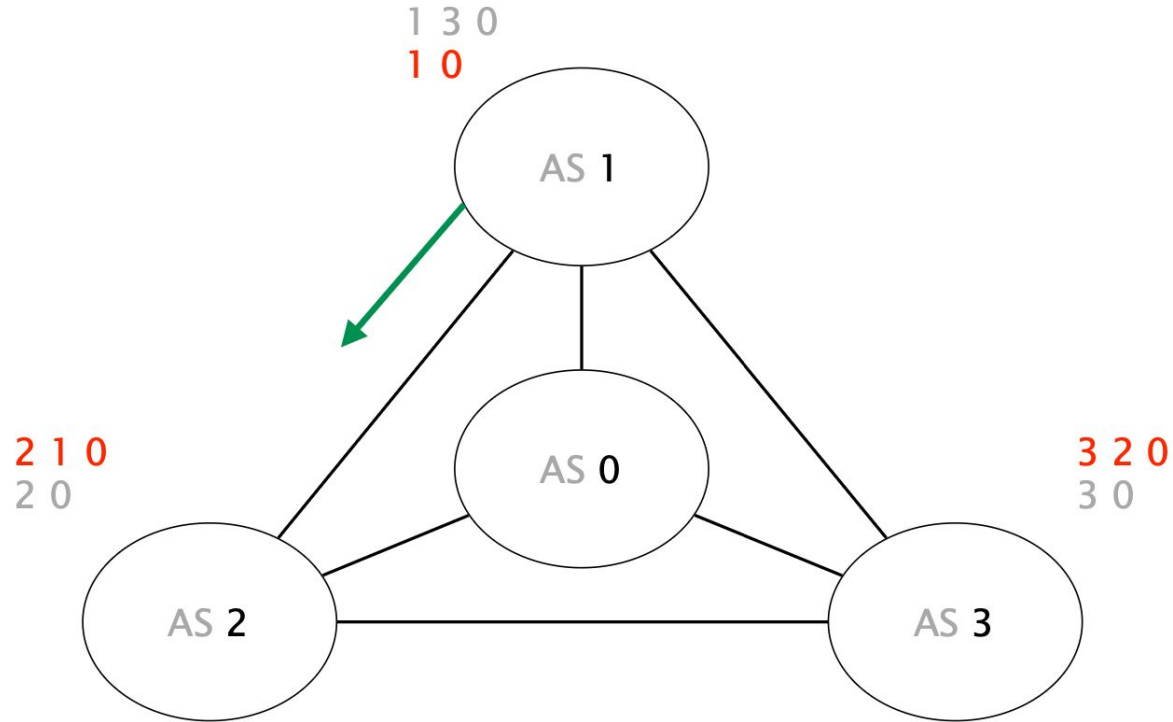
Upon reception, AS 1 reverts back to 1 0 (initial path)



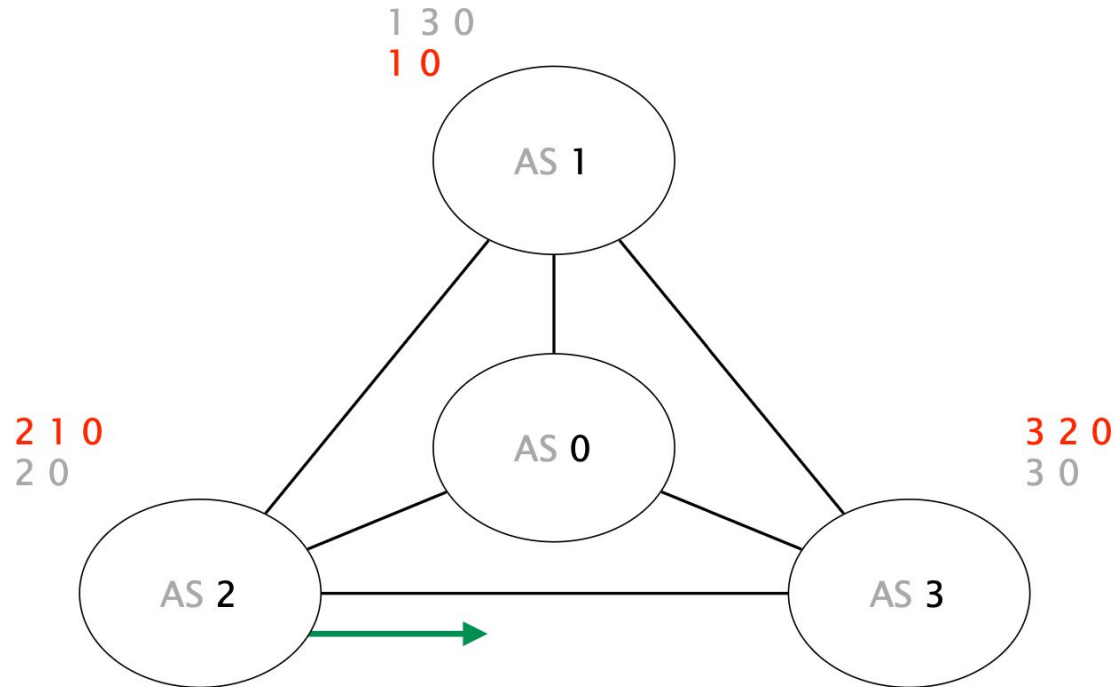
AS 1 advertises its new path 1 0 to AS 2



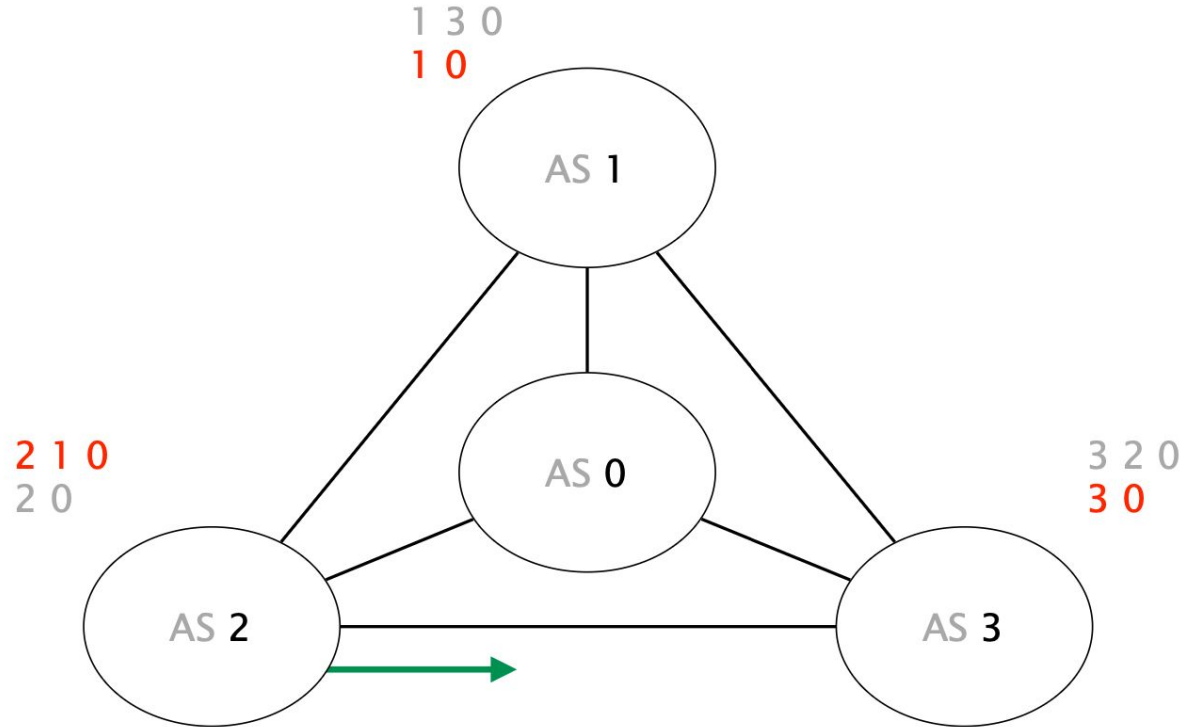
Upon reception, AS 2 switches to 2 1 0 (preferred)



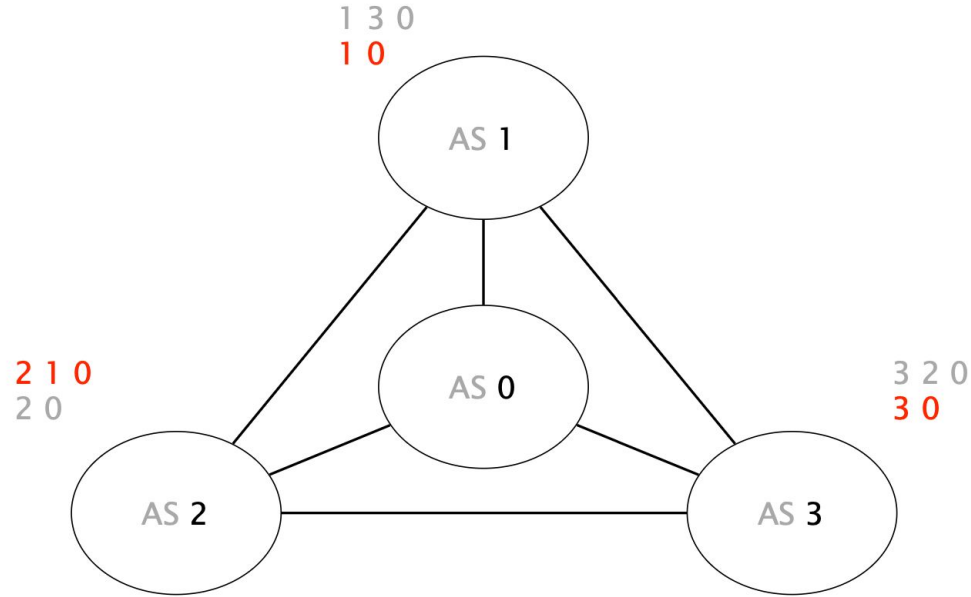
AS 2 advertises its new path 2 1 0 to AS 3



Upon reception, AS 3 switches to its initial path 3 0



We are back where we started, from there on, the oscillation will continue forever



Policy oscillations are a direct consequence of policy autonomy

ASes are free to choose and advertise any paths they want
network stability argues against this

Guaranteeing the absence of oscillations is hard
even when you know all the policies!

In practice, BGP does not oscillate “that” often

known as “Gao-Rexford” rules

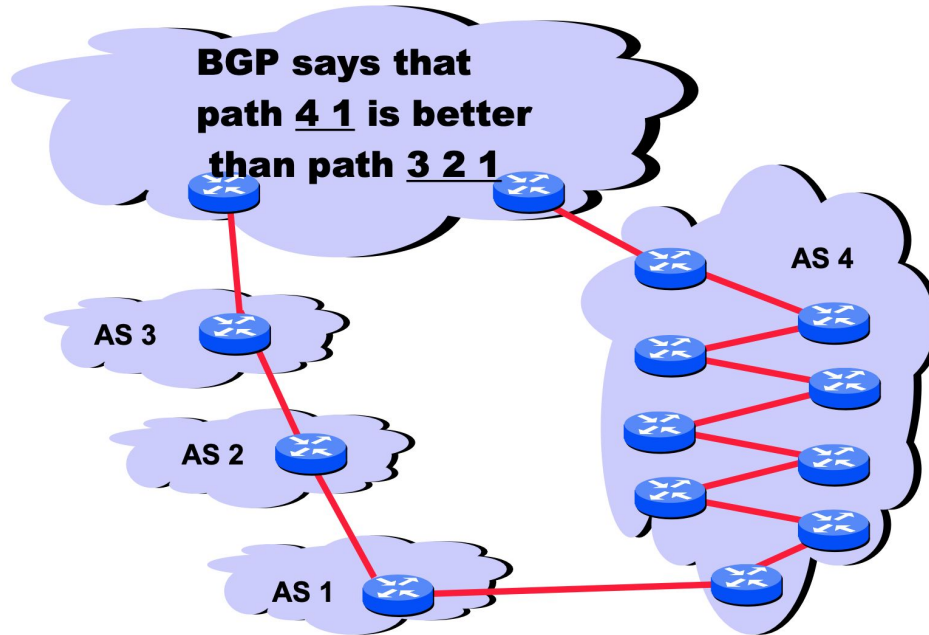
Theorem

If all AS policies follow the cust/peer/provider rules,
BGP is **guaranteed** to converge

Intuition

Oscillations require “preferences cycles”
which make no economical sense

BGP path selection is mostly economical, not based on accurate performance criteria



BGP is fragile

BGP is both “bloated” and underspecified

lots of knobs and (sometimes, conflicting) interpretations

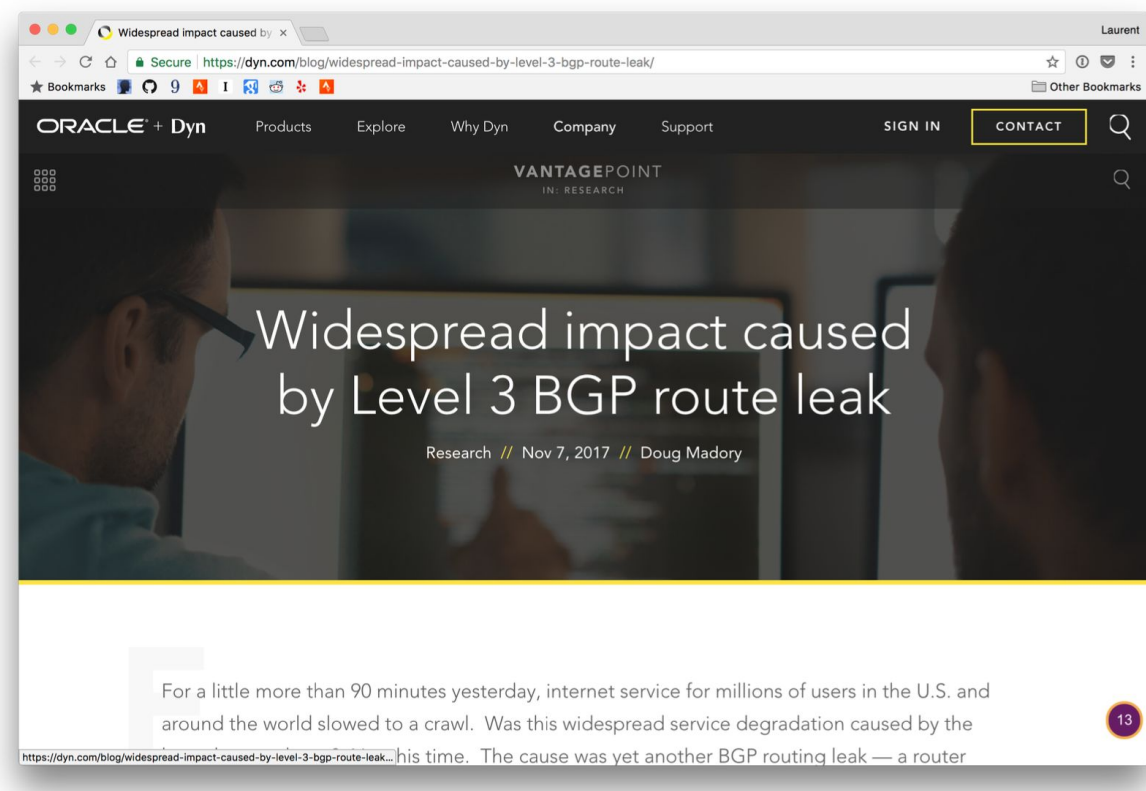
BGP is often manually configured

humans make mistakes, often

BGP abstraction is fundamentally flawed

disjoint, router-based configuration to effect AS-wide policy

Many outages



The image shows a screenshot of a web browser displaying a blog post on the Dyn website. The browser's address bar shows the URL: <https://dyn.com/blog/widespread-impact-caused-by-level-3-bgp-route-leak/>. The page header includes the Oracle + Dyn logo and navigation links for Products, Explore, Why Dyn, Company, and Support. There are also links for SIGN IN and CONTACT. The main content area features a large title: "Widespread impact caused by Level 3 BGP route leak" and a subtitle: "Research // Nov 7, 2017 // Doug Madory". Below the title, the text begins: "For a little more than 90 minutes yesterday, internet service for millions of users in the U.S. and around the world slowed to a crawl. Was this widespread service degradation caused by the". A small purple circle with the number 13 is visible in the bottom right corner of the page.

Widespread impact caused by Level 3 BGP route leak

Research // Nov 7, 2017 // Doug Madory

For a little more than 90 minutes yesterday, internet service for millions of users in the U.S. and around the world slowed to a crawl. Was this widespread service degradation caused by the

[https://dyn.com/blog/widespread-impact-caused-by-level-3-bgp-route-leak...](https://dyn.com/blog/widespread-impact-caused-by-level-3-bgp-route-leak/) his time. The cause was yet another BGP routing leak — a router

Many outages

For a little more than 90 minutes [...],

Internet service for millions of users in the U.S. and around the world slowed to a crawl.

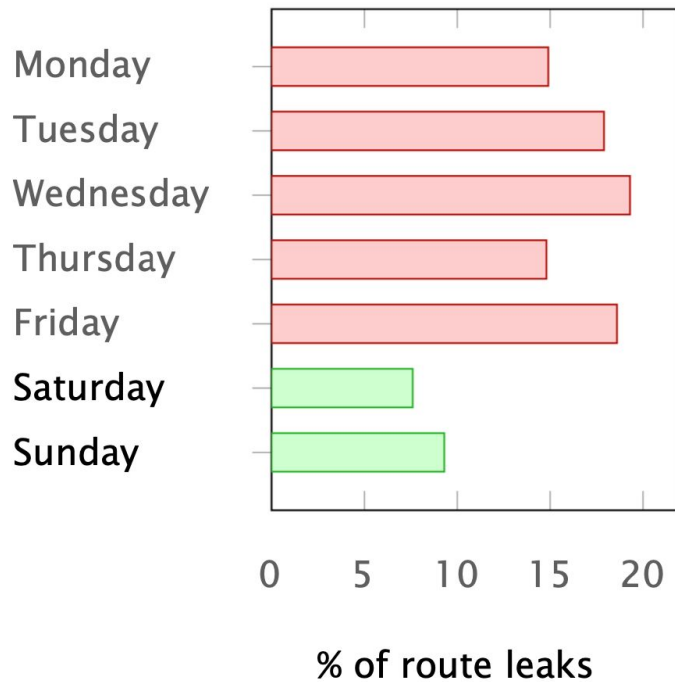
The cause was yet another BGP routing leak, a **router misconfiguration** directing Internet traffic from its intended path to somewhere else.

Many outages

“Human factors are responsible
for 50% to 80% of network outages”

Juniper Networks, *What's Behind Network Downtime?*, 2008

Ironically, this means the Internet works better on the weekends



source: Job Snijders (NTT)

BGP continues to have many problems

- Instability
 - Route flapping (network x.y/z goes down... tell everyone)
 - Not guaranteed to converge, NP-hard to tell if it does
- Scalability still a problem
 - ~485,000 network prefixes in default-free table today
 - Tension: Want to manage traffic to very specific networks (eg. multihomed content providers) but also want to aggregate information.
- Performance
 - Non-optimal, doesn't balance load across paths

The world of BGP is changing

- “Flattening” of the Internet
- ISPs are now eyeballs talking to content networks
 - e.g., Spectrum and Netflix/Spotify/YouTube
- Transit becomes less important and less profitable
 - traffic move more and more to interconnection points
- No systematic practices, yet
 - details of peering arrangements are private anyway