BGP

Border Gateway Protocol

BGP – Border Gateway Protocol Basics

Why BGP ?

- If the requirement is to use a routing protocol on the Public Internet then only choice is Border Gateway Protocol aka BGP

- BGP is the most scalable routing protocol and considered as very robust as it runs over TCP and TCP is inherently reliable

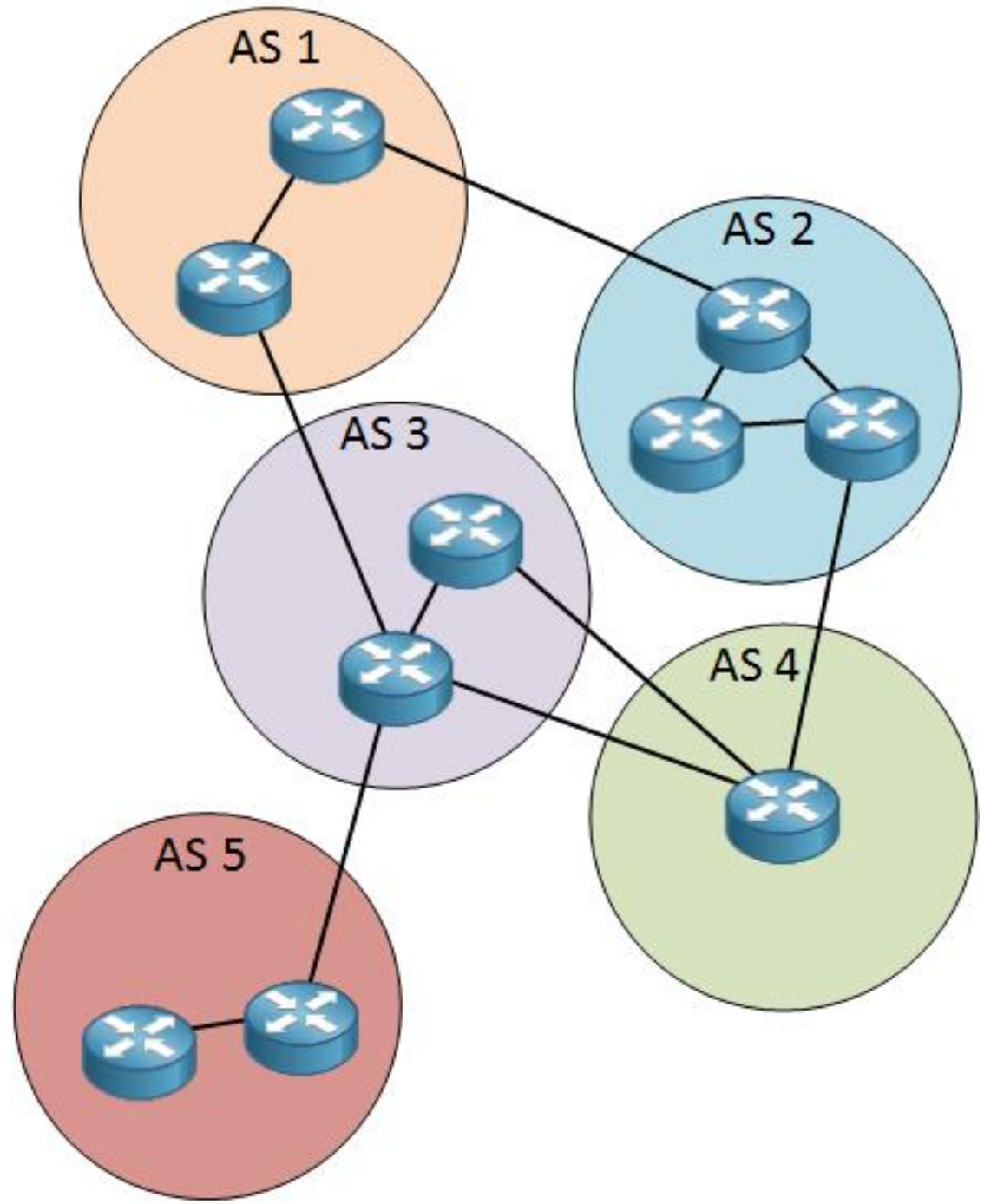
- BGP is a multi protocol , with the new NLRI it can carry many address families. Today almost a 20 different NLRI is carried over BGP. New AFI, SAFI is defined for the new address families (https://orhanergun.net/tag/multi-protocol-bgp/)

BGP – Border Gateway Protocol Basics

Autonomous System

An Autonomous System (AS) is a collection of routers whose prefixes and routing policies are under common administrative control. This could be a network service provider, a large company, a university, a division of a company, or a group of companies

An Exterior Gateway Protocol (EGP) is a routing protocol that handles routing between Autonomous Systems (inter-AS routing). BGP version 4, the Border Gateway Protocol, is the standard EGP for inter-AS routing



BGP Add-Path

- With Shadow RR or Shadows sessions, there are secondary IBGP sessions between RR and PEs. But same behavior can be achieved with BGP ADD-Path without extra IBGP session

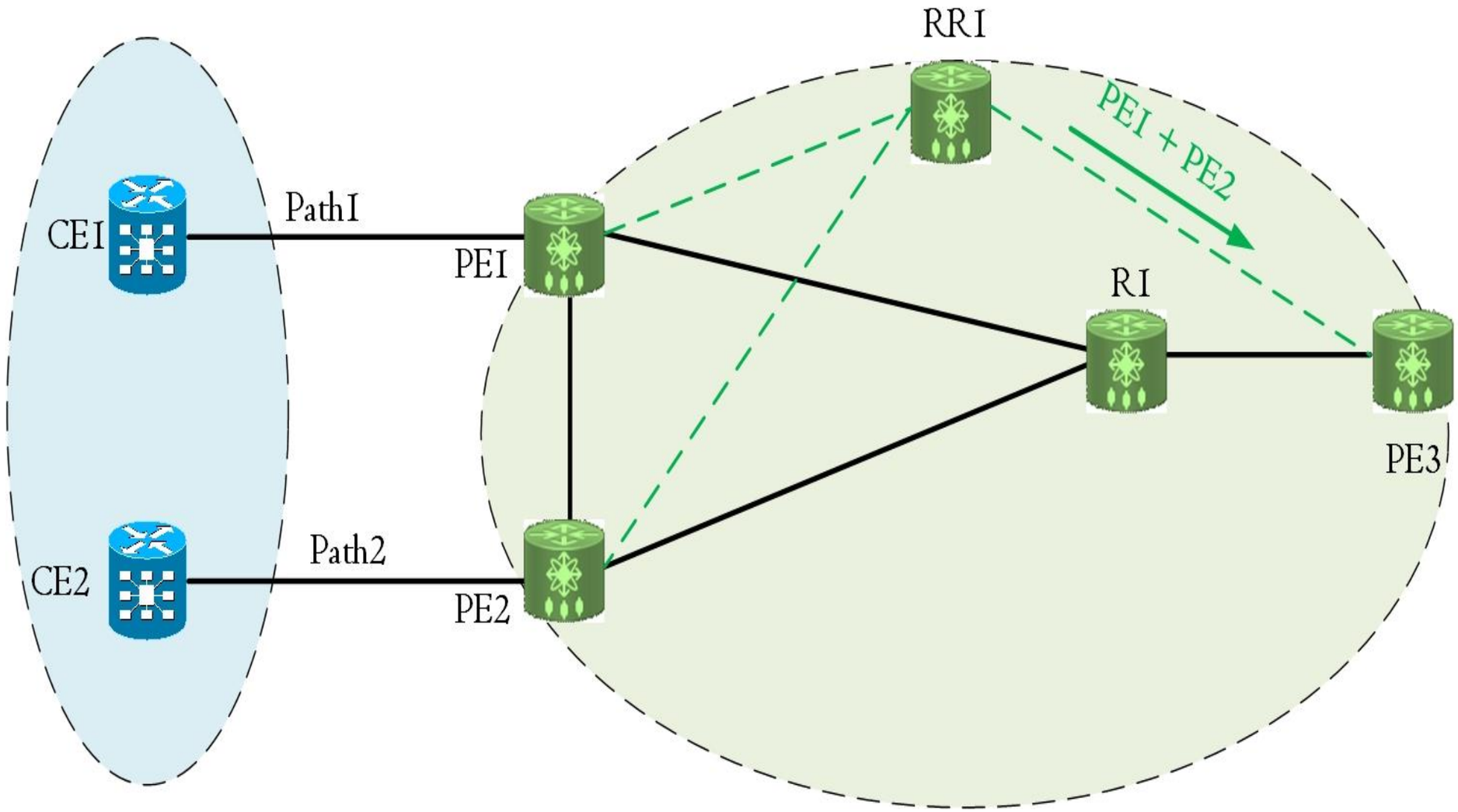
- Add-path uses path-identifier to distinguish the different next hops over one IBGP session

BGP Add-Path

- In IBGP, if multiple paths are sent over the same BGP session, last one is kept by the receiving BGP speaker, because for the first one implicit withdrawn is sent, if route is completely gone then BGP Explicit withdrawn is sent

- With Add-Path, withdrawn is not sent thus receiving BGP router keeps all the paths and can make a best path selection based on its own view

BGP Add-Path

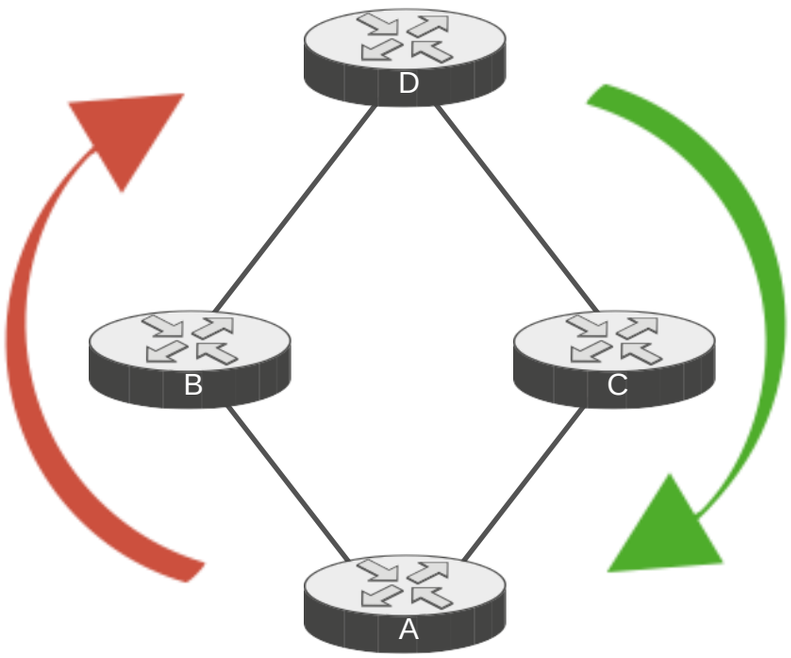


BGP – IGP Interactions

- When a router needs to be restarted or maybe having a resource issue, operator may want to remove it from the network path

- In order to do this without losing packet, a router signal it’s IGP neighbor that they shouldn’t send the traffic towards it anymore

BGP – IGP Interactions



Node B or C removal can create packet loss
if nodes don’t signal their neighbors with increased IGP metric

IS-IS Overload Bit
OSPF Max-Metric Router LSA
If BGP as an IGP , BGP Graceful Shutdown Community
EIGRP Stub Feature

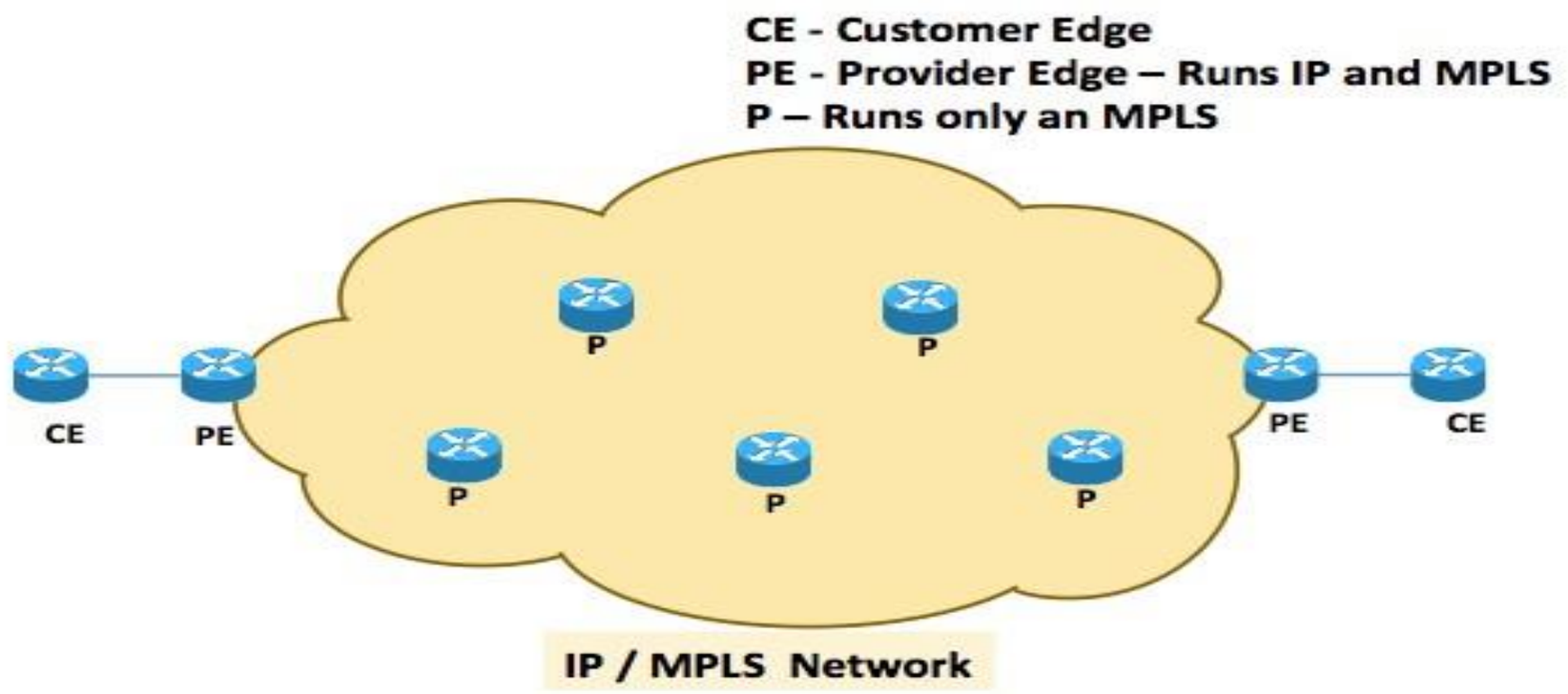
BGP – MPLS Interactions

- When BGP and MPLS is used together, generally it is used for VPN services

- MPLS removes the requirement of having BGP in the Mid/Core routers

- This phenomena is known in design as ‘ BGP Free Core ‘ design

BGP – MPLS Interactions



BGP Only needs to run at the PE routers in the above topology

BGP Route Leak Types

Based on RFC 7908, several Route Leak Type is defined

1. Hairpin Turn with Full Prefix
2. Lateral ISP-ISP-ISP Leak
3. Leak of Transit Provider Prefixes to Peer
4. Leak of Peer Prefixes to Transit Provider
5. Prefix Re-origination with Data Path to Legitimate Origin
6. Accidental Leak of Internal Prefixes and More-Specific Prefixes

BGP Route Leak - Hairpin Turn with Full Prefix

- A multihomed AS learns a route from one upstream ISP and simply propagates it to another upstream ISP

- It should be noted that leaks of this type are often accidental (not malicious)

- The leak often succeeds (the leaked update is accepted and propagated) because the second ISP prefers customer announcement over peer announcement of the same prefix

BGP Path Hunting

- When BGP is used in the DC, based on the AS allocation, it might suffer from the BGP Path Hunting behavior

- BGP Path Hunting will slow down the convergence based on the topology and ASN allocation schema when there is a failure in the network, link or node failure

- Let’s have a look at the details of BGP Path Hunting before start talking about BGP ASN allocation

BGP Path Hunting

- Without topology information a Router does not know the physical link state of every other node in the network, it doesn’t know whether the route is truly gone (because the node at the end went down itself) or is reachable via some other path

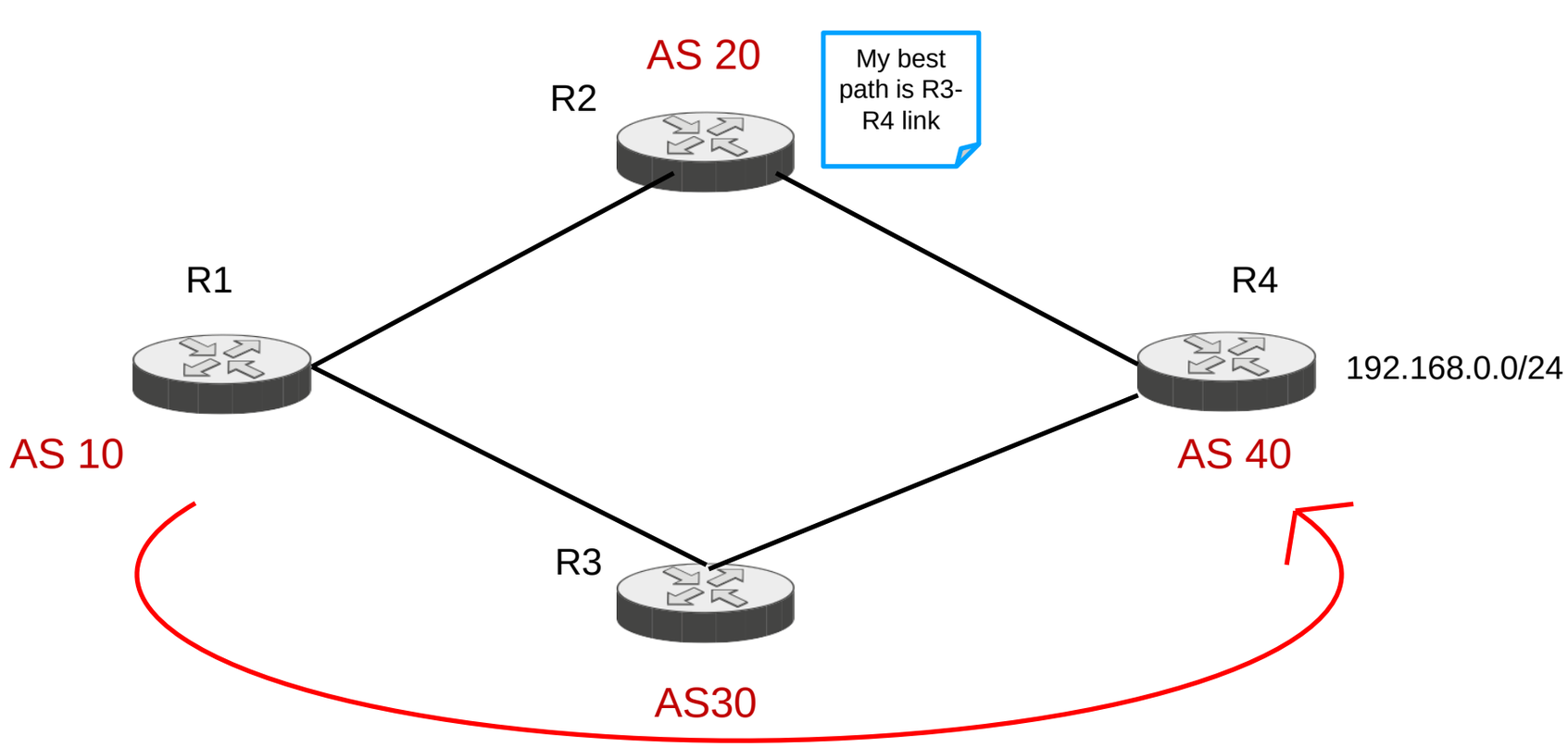
- That’s why a Router proceeds to hunt down reachability to the destination via all its other available paths. This is called path hunting

BGP Path Hunting

Let’s assume R1 selected best path to 192.168.0.0/24 as R3

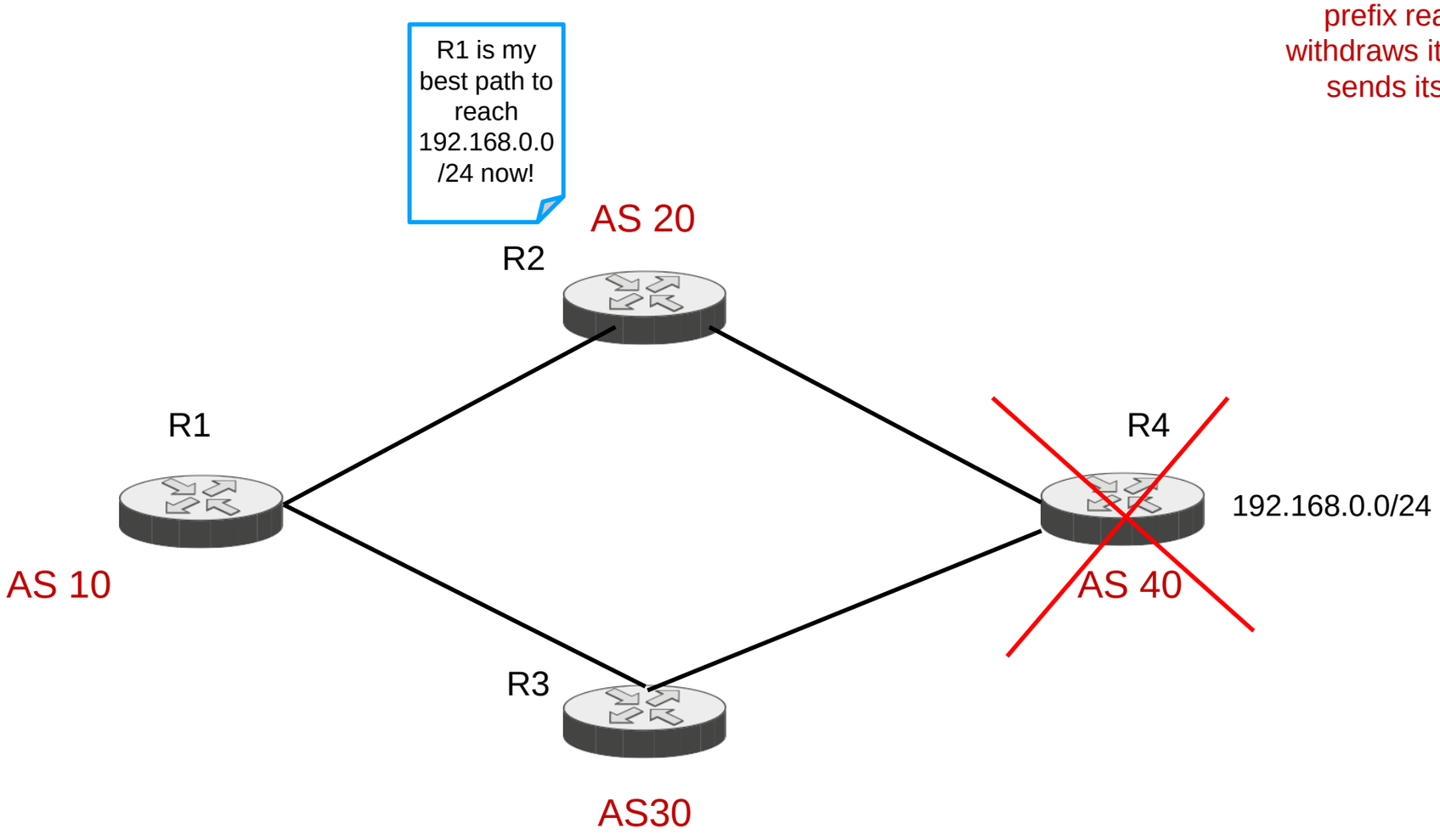
R1 advertises [R1 R3 R4] As-Path to R2

R2 accepts the advertisement but doesn’t use it , as R2 has shorter path to 192.168.0.0/24



BGP Path Hunting

Now, when the Router R4 fails, R2 loses its best path to 192.168.0.0/24, and so it re-computes its best path via R1, AS_PATH [R1, R3, R4] and sends this message to R1.
R2 also sends a route withdrawal message for the prefix to R1
When R3's withdrawal for the prefix reaches R1, R1 also withdraws its route to prefix, and sends its withdrawal to R2



BGP Path Hunting

- EBGP AS number allocation will trigger path hunting when there is a failure to the destination

- Path Hunting will slow down the convergence which is not good for the Datacenter BGP

BGP Path Hunting

- Path Hunting in BGP is a normal process for convergence, you cannot say I don’t want Path Hunting, it is how protocol works (Similar to EIGRP)

- We will look at next how ASN allocation should happen to reduce convergence impact of BGP Path Hunting behavior when EBGP is used inside the Datacenter