fastly

Selectively Sharing Multipath Routes in BGP

Trisha Biswas
Fastly, Inc.

LISA21

June 3, 2021

Agenda

1. **BGP Overview**

   Overview of Border Gateway Protocol.

2. **Best Path Selection**

   Best path algorithm and route sharing in BGP.

3. **Additional Paths**

   Sharing additional routes over an add path session.

4. **Selective Add Paths**

   Improving path diversity with selective add paths.

5. **Policy Based filtering of Add Paths**

6. **Demo**

   Demonstration of selective add-path.

7. **Conclusions**

BGP Overview

- Border Gateway Protocol is an external routing protocol that is used between Internet Service Providers (ISPs)
- Routers running BGP are called BGP speakers
- The protocol is best suited for a network of networks, or a network of Autonomous Systems (ASes)

BGP Overview

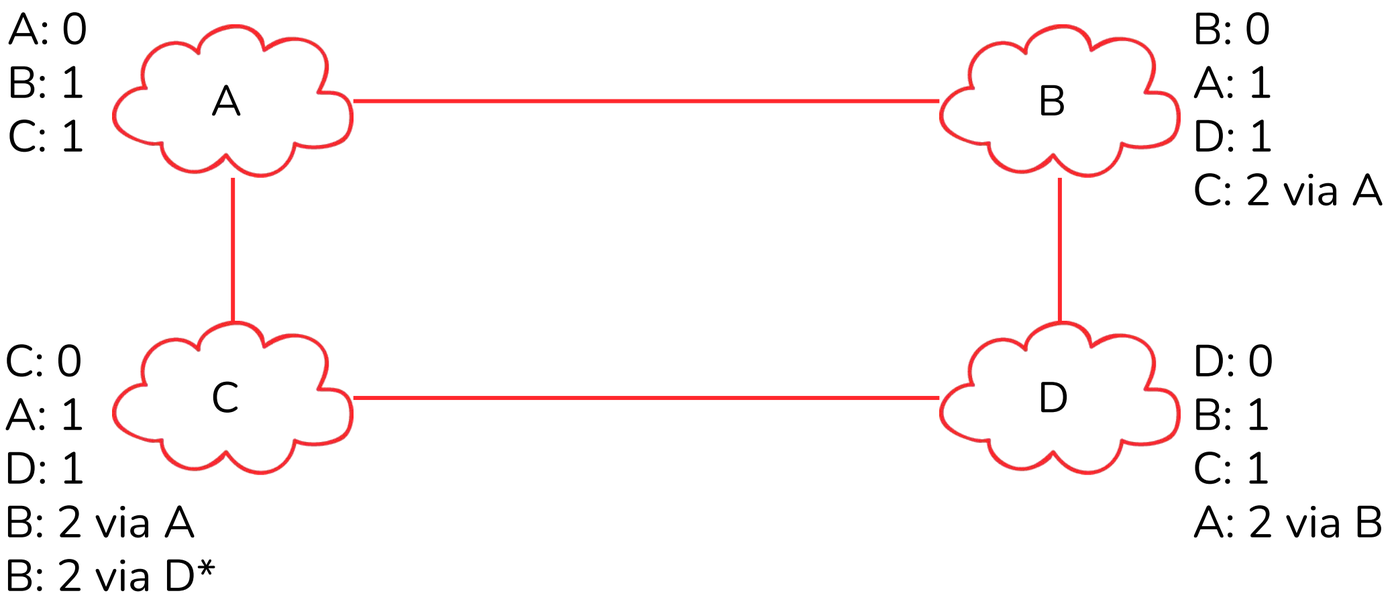
- ASes run Interior Gateway Protocol (IGP) within the system, and Exterior Gateway Protocol (EGP) between them
- Routing between autonomous systems is called interdomain routing
- BGP neighbors are called peers and must be statically configured
- Peers in different ASes use external BGP or eBGP sessions for communication, whereas peers within the AS form internal or iBGP sessions

BGP Overview

- A route to a network consists of a list of autonomous systems as well as other attributes on the path to that network
- BGP is a path-vector protocol, where speakers advertise reachability to other networks (represented as prefixes) with their peers
- The advertisement of a prefix replaces the previous announcement of that prefix. This behavior is known as an implicit withdrawal

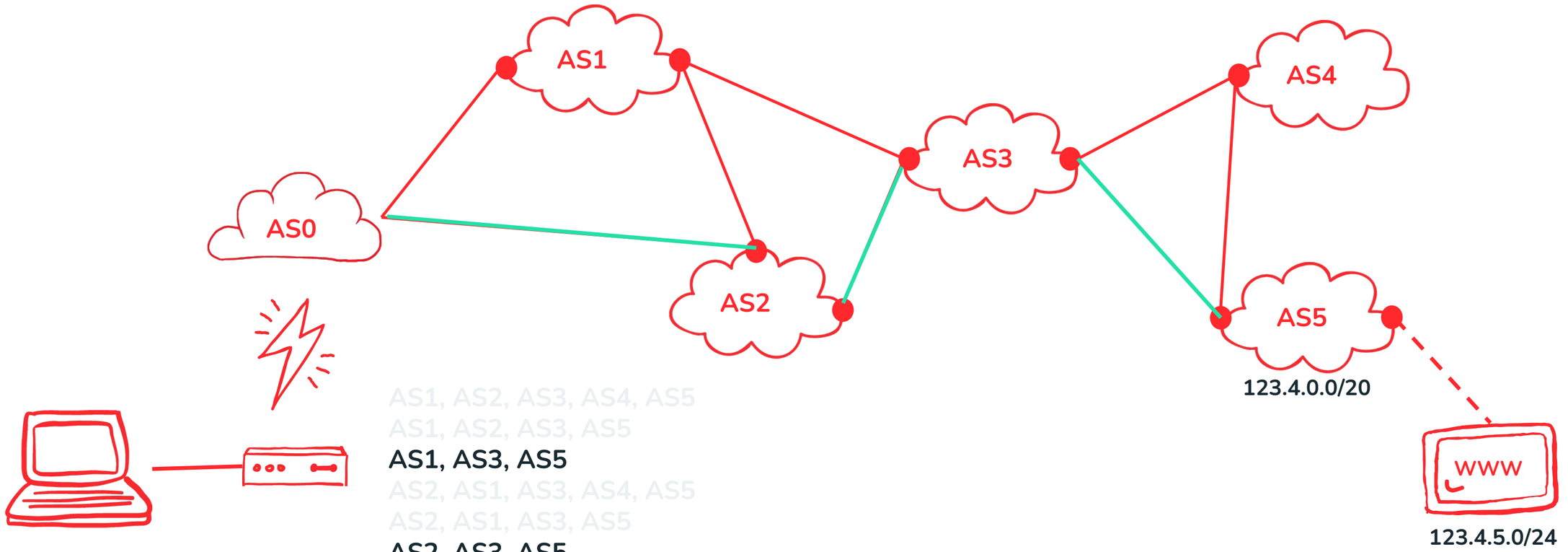
Best Path Selection

Example 1



Best Path Selection

Example 2



Additional Paths

- BGP routers propagate only their best path over their sessions
- This achieves better scaling, but at the cost of path diversity
- BGP Additional Paths (RFC7911*) feature allows sharing of multiple paths for the same prefix without the new paths implicitly replacing previous paths
- Add paths helps achieve faster re-convergence after network failures

A's routing table entry:

```
157.52.89.0  proto bird  metric 10
   nexthop via 192.168.100.1  dev et14 weight 1 *
   nexthop via 192.168.100.5  dev et15 weight 1
```

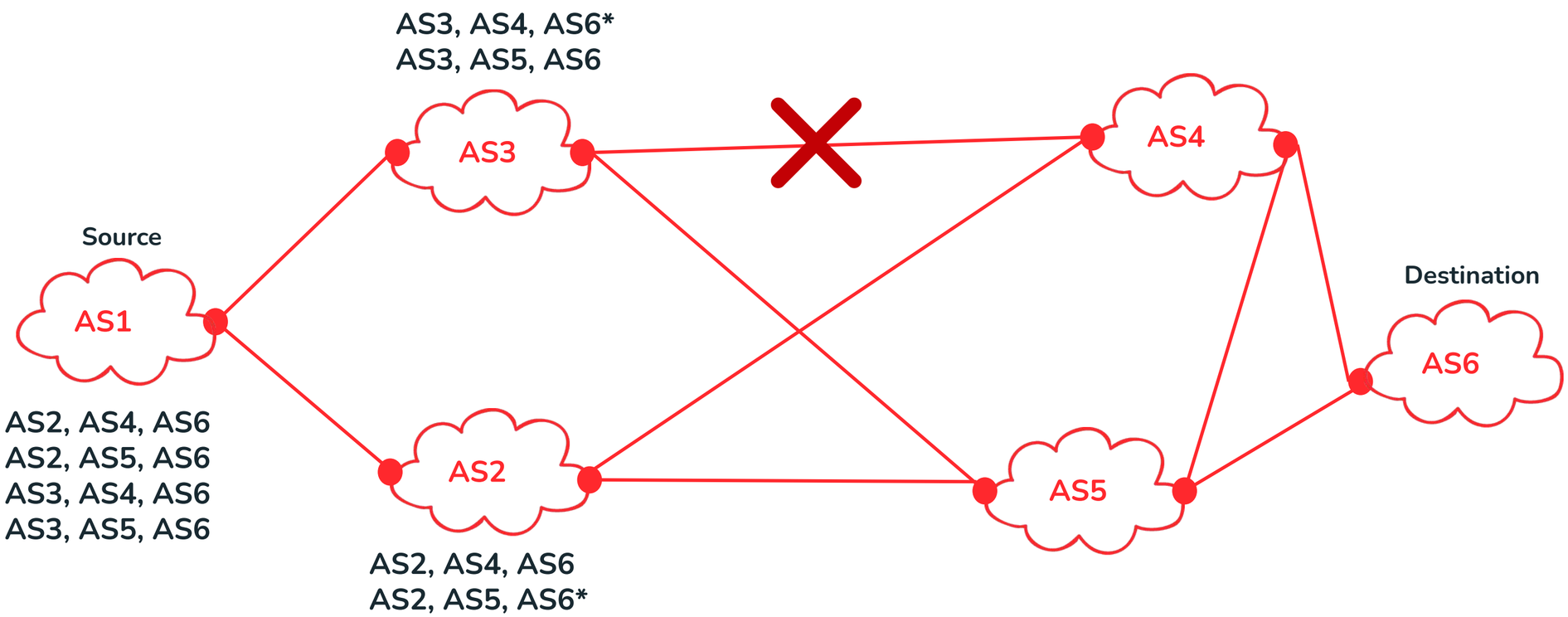
What A advertises to B:

```
157.52.89.0  proto bird  metric 10
   nexthop via 192.168.100.1  dev et14 weight 1
```

* RFC7911: https://tools.ietf.org/html/rfc7911

Additional Paths

Example



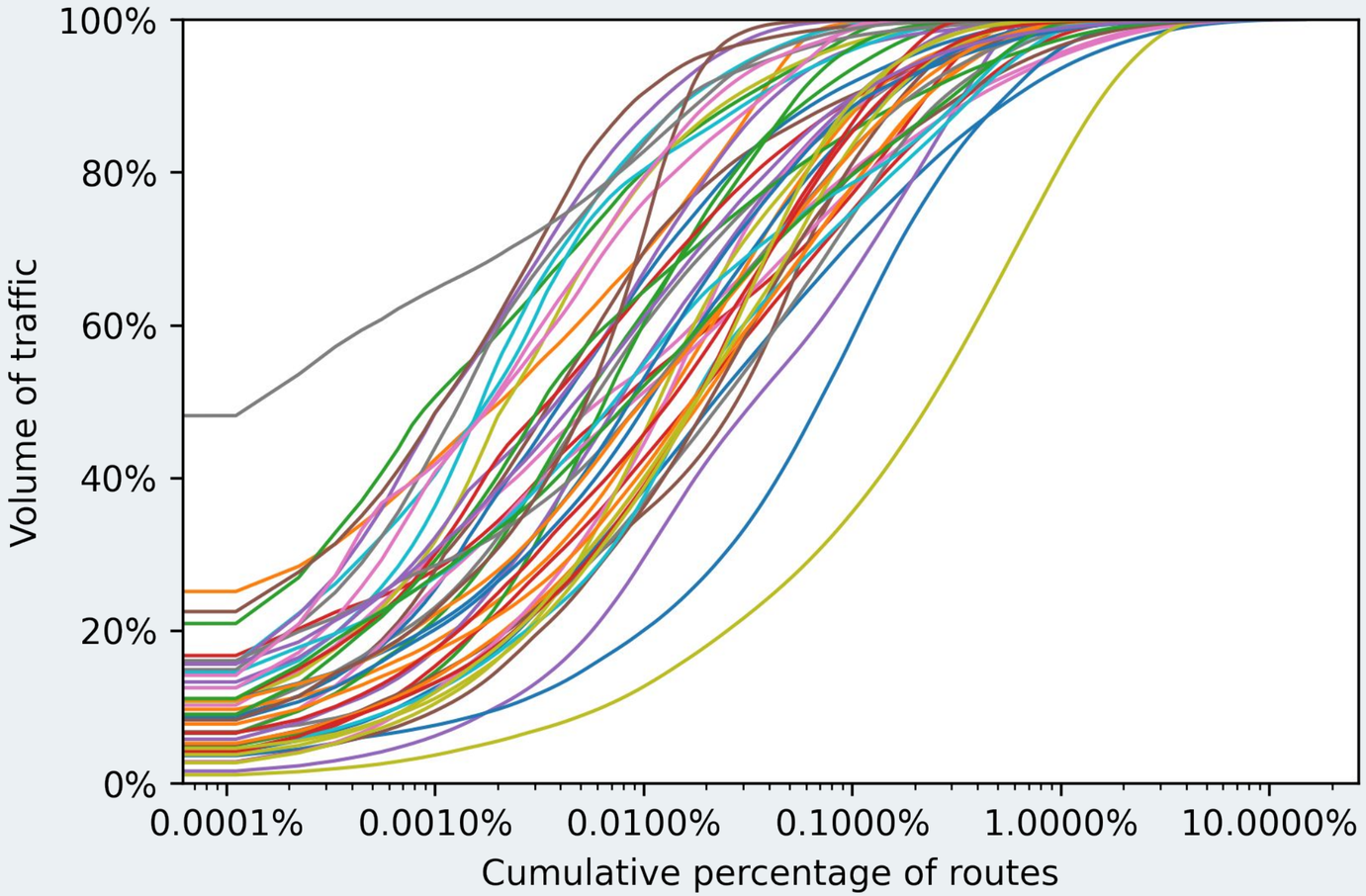
Selective Add Paths

- BGP Add paths causes all additional paths to be sent to the peer
- In today's networks, ASes have thousands of prefixes, and there could be thousands of paths to each prefix if every AS used full add paths
- Enabling add-paths over a BGP session could result in sharing millions of routes with a peer, potentially overwhelming it
- Selective add paths* feature extension allows sending of specific routes for a given prefix based on additional selection criteria

* IETF draft: https://tools.ietf.org/html/draft-keyupate-idr-bgp-selective-add-paths-00

Selective Add Paths

Experimental data



In a dense network, although there may be thousands of prefixes, only a handful serve most of the traffic. Sharing multiple paths for only those prefixes unlocks the potential of multipath BGP, without compromising peer performance.

Policy Based Filtering of Add Paths

- Routes can be filtered based on any of the BGP route attributes
- Best or preferred path should always be advertised

```
bird > show route 192.0.2.0/24 all
192.0.2.0/24
unicast [bgp_1 14:22:16.301] * (100) [AS64512i]
   via 203.0.113.1 on po491
   BGP.as_path: 64512
   BGP.next_hop: 203.0.113.1
   BGP.med: 96
   BGP.local_pref: 55
   BGP.ext_community: (ro, 2, 2) (ro, 3, 1)
unicast [bgp_2 14:22:16.279] (100) [AS64512i]
   via 203.0.113.2 on po491
   BGP.as_path: 64512
   BGP.next_hop: 203.0.113.2
   BGP.med: 98
   BGP.local_pref: 55
   BGP.ext_community: (ro, 2, 2)
```

Demo

Conclusions

- ➔ BGP add paths helps achieve faster re-convergence but can affect the overall performance of the peer due to the large number of routes advertised
- ➔ Although ASes have hundreds of thousands of routes, only a few hundred serve most of the traffic and are likely worth sharing
- ➔ Selective add paths helps achieve the best of both worlds by leveraging BGP multipath without overloading the peers

Acknowledgements: Lorenzo Saino & Jeremiah Millay

Thank You!