

Dynamic P2P with BGP Route Servers

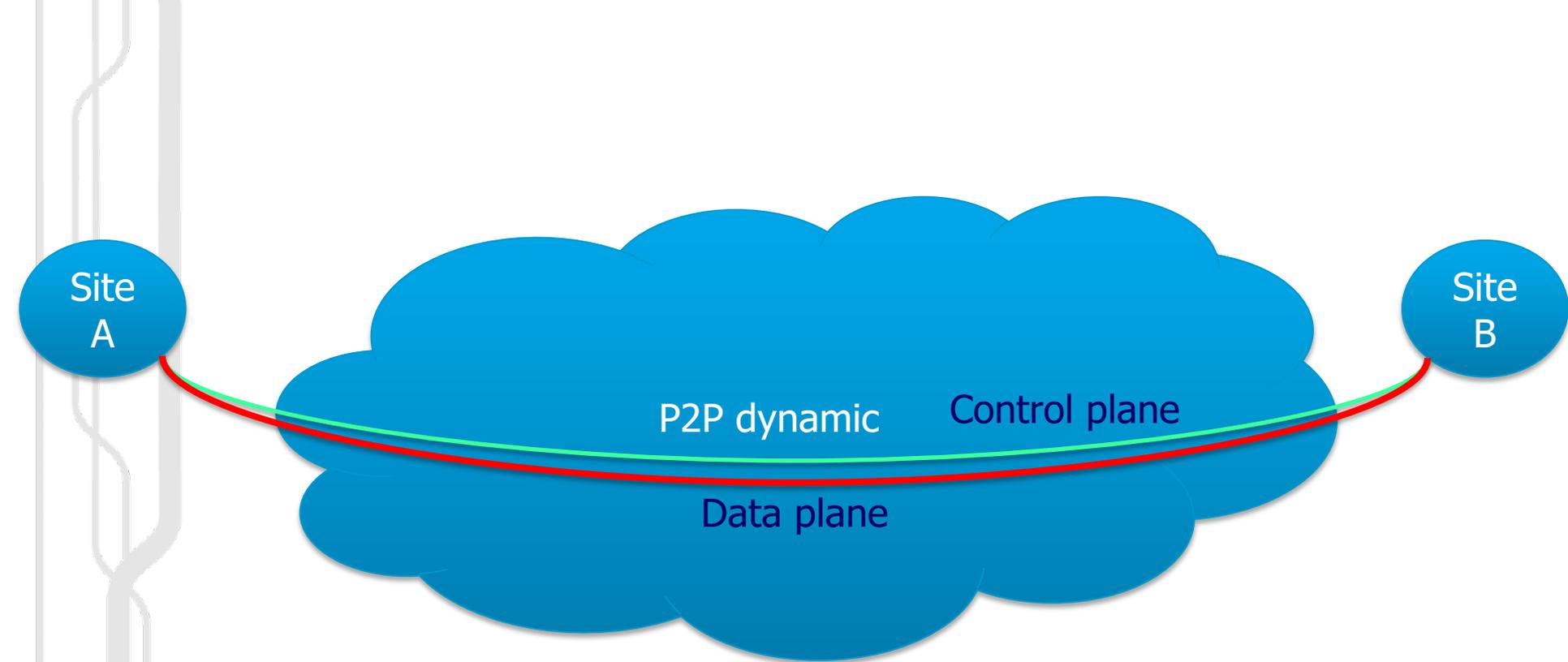
BFD for data-plane verification

Magnus Bergroth

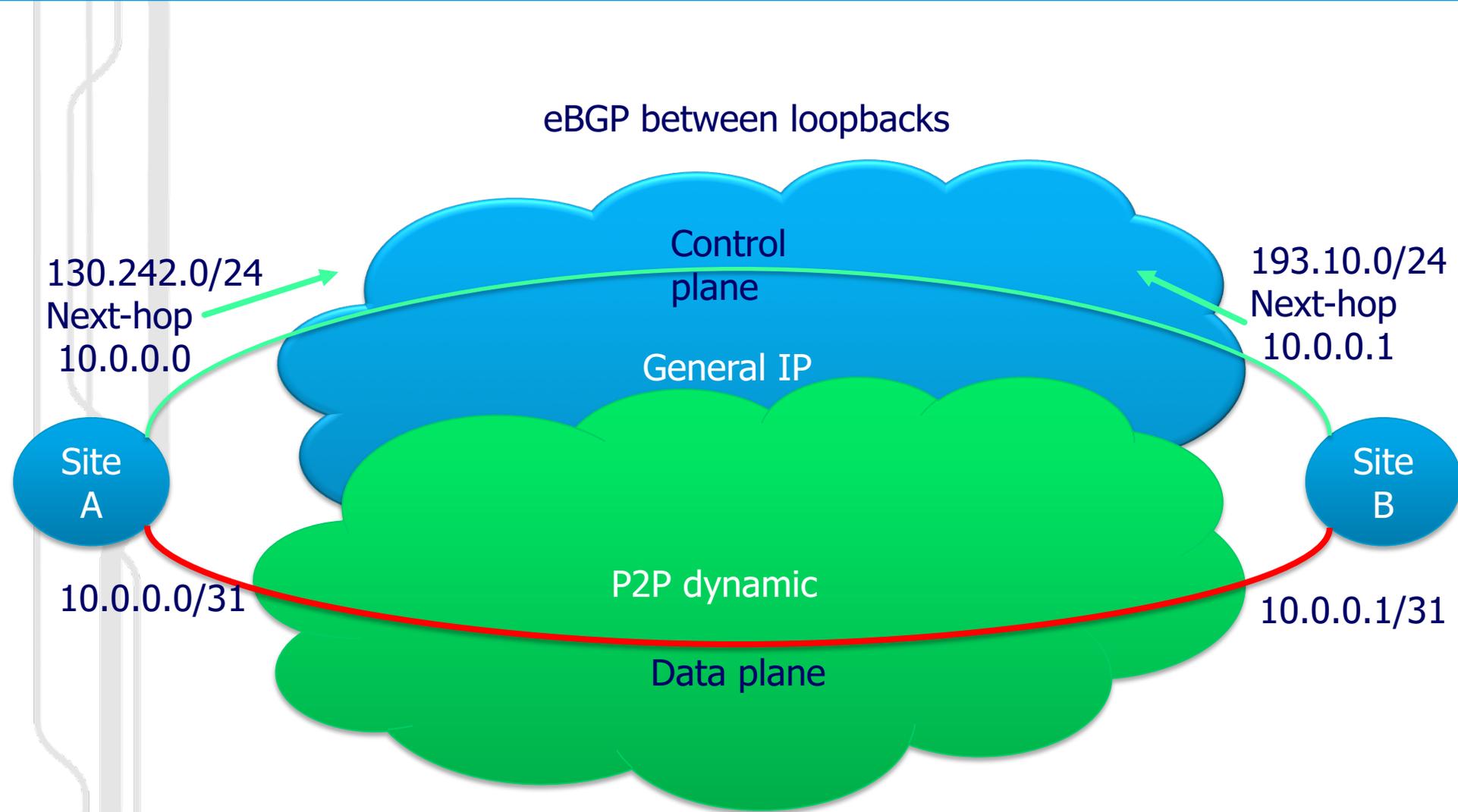
NORDUnet

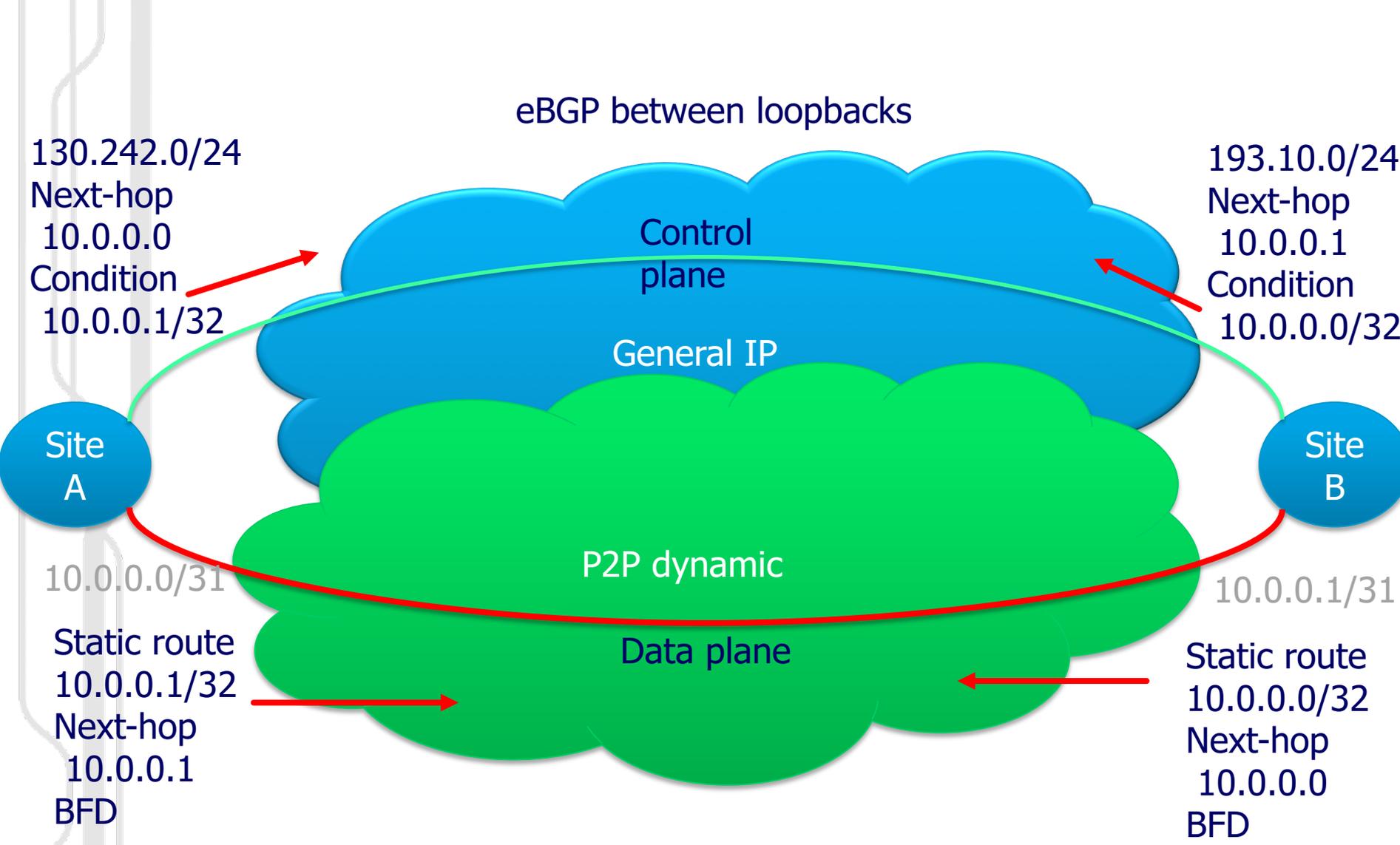
- Dynamic P2P links has two end points that normally terminates in a aggregation router at each site.
- On logical interface per destination site.
- eBGP are configured over the logical interface to each site.
- Reachability is advertised after the P2P link is up and BGP is established.

- Full mesh of BGP sessions.
- Extensive amount of configuration.
- BGP sessions over short lived P2P links are most of the time down and causes alarms.

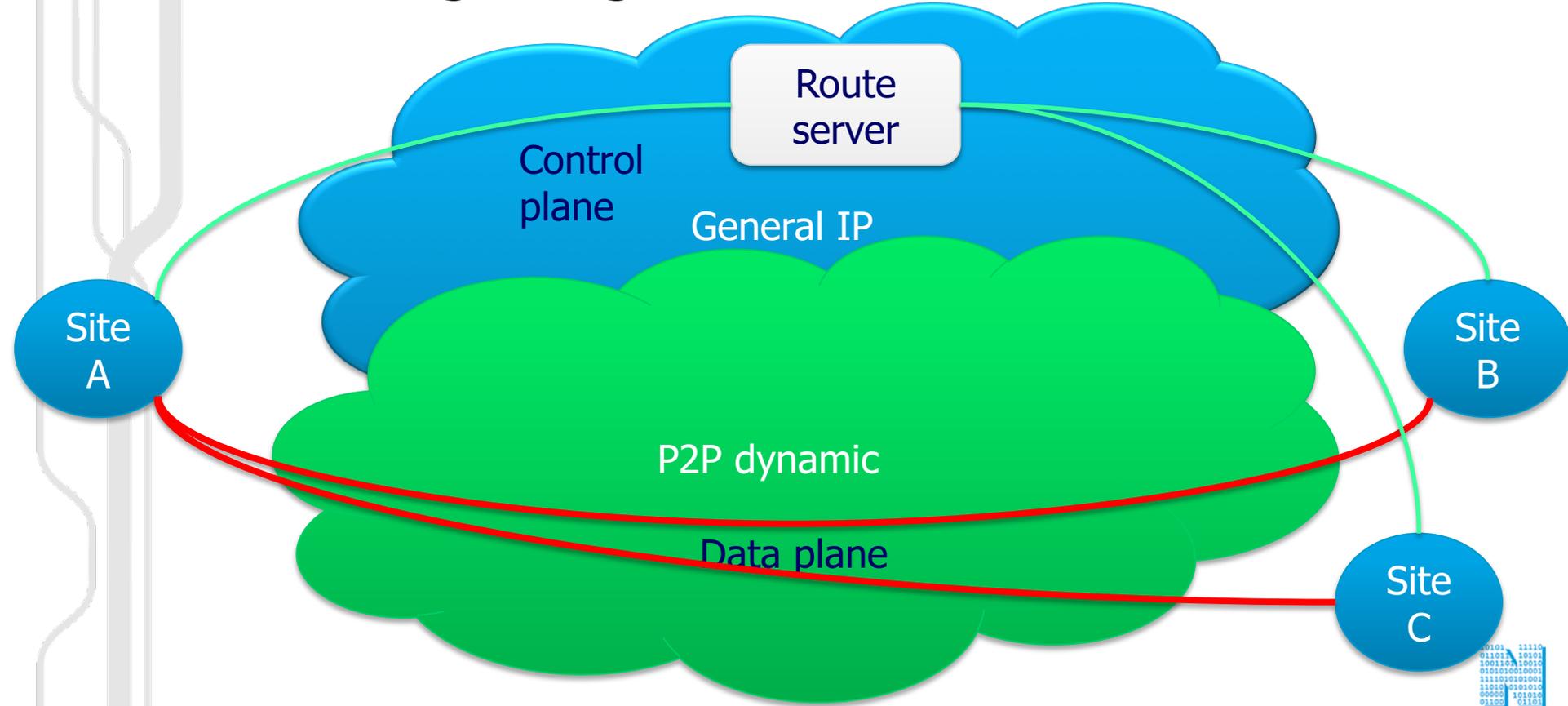


- Control plane shared with data plane
- Data plane reachability detected when control plane goes up





- Simplify the BGP setup
- Only one BGP session per site
- Route server with one outgoing RIB per site, steering using communities



- RFC5549
 - Advertising IPv4 Network Layer Reachability Information with an IPv6 Next Hop
- Logical interface per site, link local v6
- BFD per site
- Community per site

130.242.0/24

Next-hop

2001:6b0:ffff:1

Condition

2001:6b0:ffff::2/128

add community site_B

Condition

2001:6b0:ffff::3/128

add community site_C

lo0.1
2001:6b0:ffff::1

Site A

fe80::1
fe80::1

Static route

2001:6b0:ffff:2

Next-hop Static route

fe80::2 2001:6b0:ffff:3

BFD Next-hop

fe80::3

BFD

Control plane

General IP

P2P dynamic

Data plane

Route server

lo0.1
2001:6b0:ffff::2

Site B

fe80::2
fe80::2

fe80::3

Site C

fe80::3

lo0.1
2001:6b0:ffff:13



```
root@chimay-re0# show interfaces xe-4/0/2
description "to duvel xe-4/0/2";
flexible-vlan-tagging;
encapsulation flexible-ethernet-services;
unit 10 {
    vlan-id 10;
    family inet;
    family inet6 {
        address fe80::1/64;
    }
}
```

```
root@chimay-re0# show routing-options
rib inet6.0 {
    static {
        route 2001:6b0:1e:50:1234::2/128 {
            qualified-next-hop fe80::2 {
                interface xe-4/0/2.10;
            }
            bfd-liveness-detection {
                minimum-interval 100;
            }
            no-readvertise;
        }
    }
}
```

```
root@chimay-re0# show protocols bgp group COOKIE
multihop {
    ttl 10;
}
local-address 2001:6b0:25:1::11;
family inet {
    unicast {
        extended-nexthop;
    }
}
family inet6 {
    unicast;
}
export [ adv-v4 reject ];
neighbor 2001:6b0:1e:50::2 {
    peer-as 2832;
}

root@orval-re0# show policy-options policy-statement adv-v4
term 1 {
    from {
        protocol static;
        route-filter 172.18.0.0/16 exact;
    }
    then {
        next-hop 2001:6b0:1e:50:1234::2; (secondary loopback)
        accept;    }
}
```



```
root@chimay-re0# run show route 172.18.0.0/16
inet.0: 22 destinations, 22 routes (17 active, 0 holddown, 5 hidden)
+ = Active Route, - = Last Active, * = Both
172.18.0.0/16      *[BGP/170] 03:56:22, localpref 100, from 2001:6b0:1e:50::2
                   AS path: 2832 I, validation-state: unverified
                   > to fe80::2 via xe-4/0/2.10
```

```
root@chimay-re0# run show bfd session
```

Detect Address	Transmit	State	Interface	Time	Interval	Multiplier
fe80::2		Up	xe-4/0/2.10	0.300	0.100	3

```
root@chimay-re0# run show bfd session
```

Address	State	Interface	Detect Time	Transmit Interval	Multiplier
fe80::2	Down	xe-4/0/2.10	0.300	2.000	3

```
root@chimay-re0# run show route 172.18.0.0/16 hidden
```

```
172.18.0.0/16      [BGP/170] 00:00:08, localpref 100, from 2001:6b0:1e:50::2
                   AS path: 2832 I, validation-state: unverified
                   Unusable
```

- Created for route-servers at IXP
- Client routers verifies connectivity of dataplane with BFD Bidirectional Forwarding Detection, RFC5880.
- Advertising NH-Reach state in BGP
A new BGP SAFI, the NH-Reach SAFI

BFD defaults $1000000 \text{ us} \times 3 = 3\text{sec}$ up down detection

- Only one BGP session with the Router-server. Minimize the router configuration.
- The BGP session will always be up.
- Connectivity to new sites are added by adding communities to advertised prefixes.
- Prefixes are learned via BGP only when the Data-plane(dynamic P2P) is up.
- Fast detection when the dynamic P2P are teared down. No risk of black holing

- Still a draft
- Quagga and Bird are working on implementation.
- ~~The draft suggest a dataplane to be consider permanent down after 24hours. This needs to be configurable to be tested for ever.~~
- ~~The Routeserver init state is that the dataplane is working. For dynamic P2P is this not the case.~~

- Arista DC feature, no routeserver
 - BGP neighbor with ipv6 link-local addresses. Router advertisement finds neighbor.
 - Neighbor is set as connecting interface
 - P2P links only.

- GoBGP

```
[global.config]
```

```
as = 64512
```

```
router-id = "192.168.255.1"
```

```
[[neighbors]]
```

```
[neighbors.config]
```

```
neighbor-interface = "eth0"
```

Quagga **Cumulus Linux**
BGP Unnumbered

Link local RA

```
interface swp1
```

```
    ipv6 nd ra-interval 5
```

```
    no ipv6 nd suppress-ra
```

```
!
```

```
router bgp 65535
```

```
    neighbor swp1 interface
```

```
    neighbor swp1 remote-as 65534
```

```
    neighbor swp1 capability extended-nexthop
```

- For Troubleshooting and statistics
 - Route server for lookingglas
- All NSPs peers using multihop BGP
- Add-path, 5 extra paths