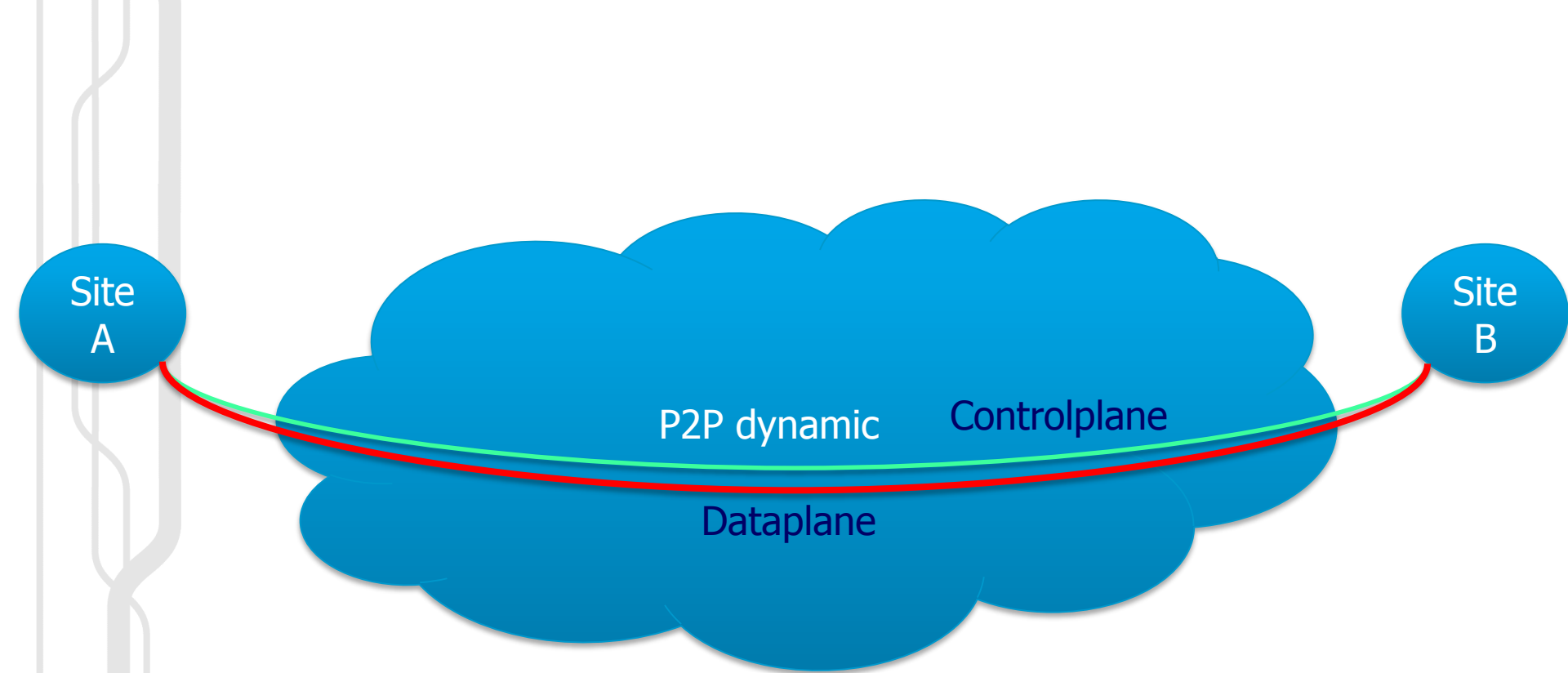


BGP Route Server Proof of Concept

Magnus Bergroth
NORDUnet

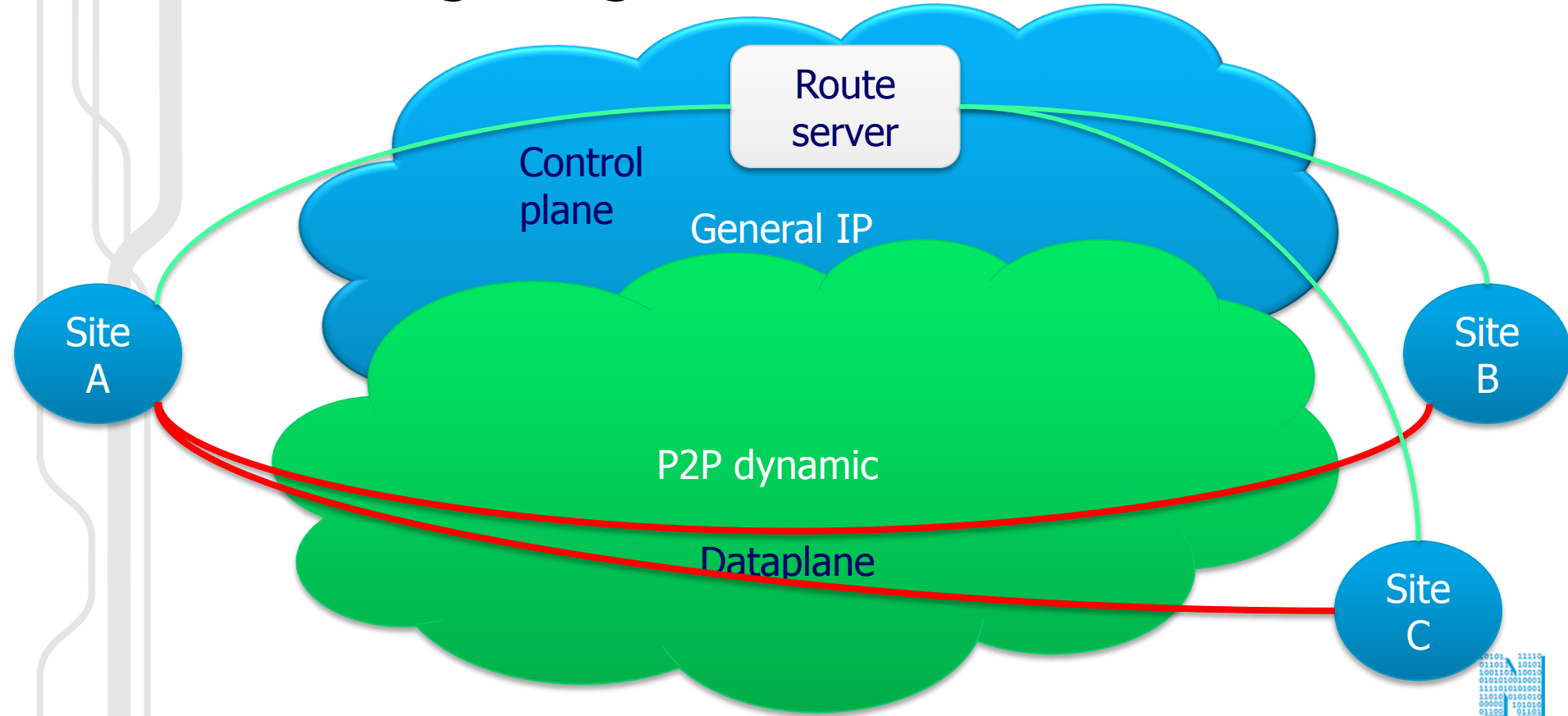
- Dynamic P2P links has two end points that normally terminates in a aggregation router at each site.
- On logical interface per destination site.
- eBGP are configured over the logical interface to each site.
- Reachability is advertised after the P2P link is up and BGP is established.

- 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.



- Controlplane shared with dataplane
- Dataplane reachability detected when controlplane goes down

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

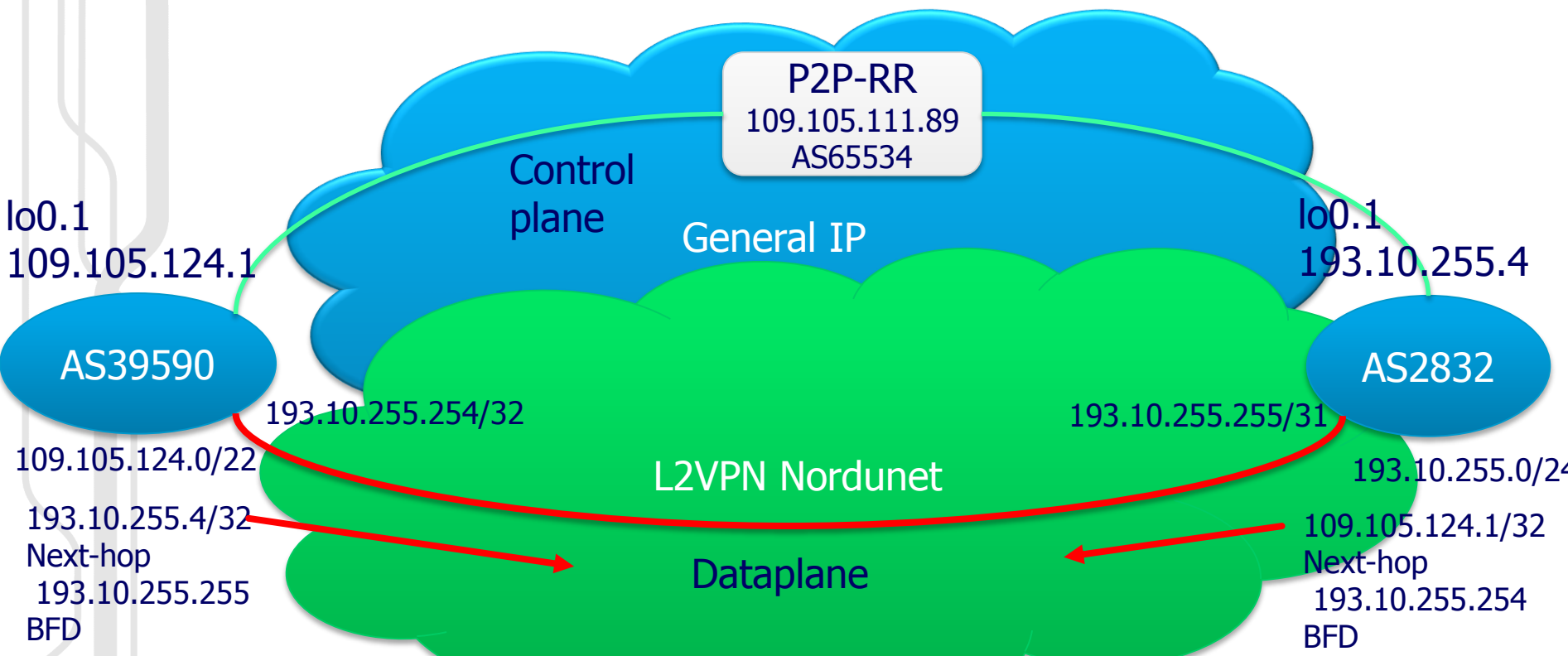


Edoardo found

- Unfortunately been quiet, and what I know no implementation ready for use.

L2VPN over NORDUnet

- NDGF(39590), juniper MX480 in Copenhagen
- AS2832, juniper T320 in Stockholm



```
define myas = 65534;

function bgp_out(int peeras){
  if (myas,peeras) ~ bgp_community then return true;
  return false;
}

protocol bgp R39590x1 {
  description "NDGF - peer 1";
  neighbor 109.105.124.1 as 39590;
  export where bgp_out(39590);
  multihop 127;
  next hop keep;
  table T39590;
  local as myas;
  rs client;
}
```


BGP Conf to Route Server (Done once)

```
protocols {
  bgp {
    group p2p-RS {
      multihop {
        ttl 127;
      }
      neighbor 109.105.111.89 {
        import [ accept-P2P reject-rest ];
        export export_on_p2p_bfd;
        peer-as 65534;
      }
    }
  }
}
```

VLAN per peer

```
interfaces {  
  ae1 {  
    unit 3001 {  
      description "p2p test";  
      vlan-id 3001;  
      family inet {  
        address 193.10.255.254/31;  
      }  
    }  
  }  
}
```

Static route

```
routing-options {  
  static {  
    route 193.10.255.4/32 { ← loopback of other side  
      next-hop 193.10.255.255;  
      bfd-liveness-detection {  
        minimum-interval 100;  
        multiplier 3;  
      }  
    }  
  }  
} Condition  
policy-options {  
  condition {  
    p2p_bfd_lo_AS2832 {  
      if-route-exists {  
        193.10.255.4/32;  
        table inet.0;  
      }  
    }  
  }  
}
```

Export policy

```
policy-options {  
  policy-statement {  
    export_on_p2p_bfd {  
      term to_AS2832 {  
        from {  
          protocol aggregate;  
          route-filter 109.105.124.0/22 exact;  
          condition p2p_bfd_lo_AS2832;  
        }  
        then {  
          community add p2p_RS_AS2832;  
          accept;  
        }  
      }  
    }  
    term reject-rest {  
      then reject;  
    }  
  }  
}
```

Peer peer conf

```
set interfaces ae1 unit 3001 description "p2p test"  
set interfaces ae1 unit 3001 vlan-id 3001  
set interfaces ae1 unit 3001 family inet address 193.10.255.254/31  
  
set routing-options static route 193.10.255.4/32 next-hop 193.10.255.255  
set routing-options static route 193.10.255.4/32 bfd-liveness-detection minimum-interval 100  
set routing-options static route 193.10.255.4/32 bfd-liveness-detection multiplier 3  
  
set policy-options condition p2p_bfd_lo_AS2832 if-route-exists 193.10.255.4/32  
set policy-options condition p2p_bfd_lo_AS2832 if-route-exists table inet.0  
  
set policy-options policy-statement export_on_p2p_bfd term to_AS2832 from protocol aggregate  
set policy-options policy-statement export_on_p2p_bfd term to_AS2832 from route-filter 109.105.124.0/22 exact  
set policy-options policy-statement export_on_p2p_bfd term to_AS2832 from condition p2p_bfd_lo_AS2832  
set policy-options policy-statement export_on_p2p_bfd term to_AS2832 then community add p2p_RS_AS2832  
set policy-options policy-statement export_on_p2p_bfd term to_AS2832 then accept  
set policy-options policy-statement export_on_p2p_bfd term reject-rest then reject  
  
set policy-options community p2p_RS_AS2832 members 65534:2832
```

Link down:

```
bergroth@dk-ndgf-re0# run show bfd session
```

Address	State	Interface	Detect Time	Transmit Interval	Multiplier
193.10.255.255	Down	ae1.3001	0.300	2.000	3

```
-----  
PING 193.10.255.255 (193.10.255.255):
```

```
56 data bytes
```

```
^C
```

```
--- 193.10.255.255 ping statistics ---
```

```
20 packets transmitted, 0 packets received, 100% packet loss
```

```
-----  
bergroth@dk-ndgf-re0# run show bgp summary | grep 109.105.111.89
```

```
109.105.111.89      65534      81111      77261      0      27 3w3d 21:10:24
```

```
Establ
```

```
-----  
bergroth@dk-ndgf-re0# run show bgp neighbor 109.105.111.89 | grep prefix
```

```
Active prefixes:      0
```

```
Received prefixes:   0
```

```
Accepted prefixes:   0
```

```
Advertised prefixes: 0
```

Link Up:

```
bergroth@dk-ndgf-re0# run show bfd session
```

Address	State	Interface	Detect Time	Transmit Interval	Multiplier
193.10.255.255	Up	ae1.3001	0.300	0.100	3

```
bergroth@dk-ndgf-re0# run show route 193.10.255.4 table inet.0
```

```
193.10.255.4/32  *[Static/5] 00:01:53  > to 193.10.255.255 via ae1.3001
```

```
bergroth@dk-ndgf-re0# run show route receive-protocol bgp 109.105.111.89
```

Prefix	Nexthop	MED	Lclpref	AS path
• 193.10.255.0/24	193.10.255.4			2832 I

```
bergroth@dk-ndgf-re0# run show route 193.10.255.0/24
```

```
193.10.255.0/24  *[BGP/170] 00:01:59, localpref 100, from 109.105.111.89
```

```
AS path: 2832 I, validation-state: unverified
```

```
> to 193.10.255.255 via ae1.3001
```

```
[BGP/170] 4w1d 14:04:59, MED 15, localpref 100
```

```
AS path: 2603 2832 I, validation-state: unverified
```

```
> to 109.105.102.49 via ae1.1
```

```
193.10.255.4/32  *[Static/5] 00:01:59
```

```
> to 193.10.255.255 via ae1.3001
```



Ping a host inside AS2832, disable P2P link.

```
64 bytes from 193.10.255.9: icmp_seq=56 ttl=62 time=9.288 ms P2P  
64 bytes from 193.10.255.9: icmp_seq=57 ttl=62 time=9.236 ms  
64 bytes from 193.10.255.9: icmp_seq=58 ttl=62 time=9.205 ms  
64 bytes from 193.10.255.9: icmp_seq=59 ttl=62 time=9.333 ms  
64 bytes from 193.10.255.9: icmp_seq=60 ttl=62 time=9.246 ms  
64 bytes from 193.10.255.9: icmp_seq=61 ttl=62 time=23.055 ms ← Switch  
64 bytes from 193.10.255.9: icmp_seq=62 ttl=59 time=44.793 ms GeneralIP  
64 bytes from 193.10.255.9: icmp_seq=63 ttl=59 time=12.311 ms  
64 bytes from 193.10.255.9: icmp_seq=64 ttl=59 time=12.273 ms  
64 bytes from 193.10.255.9: icmp_seq=65 ttl=59 time=12.221 ms  
64 bytes from 193.10.255.9: icmp_seq=66 ttl=59 time=12.472 ms  
64 bytes from 193.10.255.9: icmp_seq=67 ttl=59 time=12.360 ms  
64 bytes from 193.10.255.9: icmp_seq=68 ttl=59 time=12.221 ms
```

--- 193.10.255.9 ping statistics ---

82 packets transmitted, 82 packets received, 0% packet loss

round-trip min/avg/max/stddev = 9.169/13.165/46.212/8.627 ms



- Questions?