modulus

VXLAN & EVPN For a VoIP Provider

Kostas Lotsos

About me

- Kostas Lotsos - Chief Technical Officer
- I have been working at modulus for the past 12 years!
- Background: Networking
- I would like to live in Chania

modulus®

modulus

Contents

- Intro
- Infrastructure
- How BGP is being used
- Our VXLAN journey
- Configuration
- Results


modulus
Intro

Our History

2007

Our journey began in 2007 in the IT sector, with a vision to offer something different in telecommunications

2013

In **2013** modulus was the first VoIP provider connected to GRIX

2017

In **2017** we redesign our Core SIP infrastructure (SBCs, Media Gateways etc.)

2019

In **2019** we started to provide Wholesale VoIP Services

2020

In **2020** we refreshed our whole network infrastructure. We connected to DECIX, SEECIX & later LINX

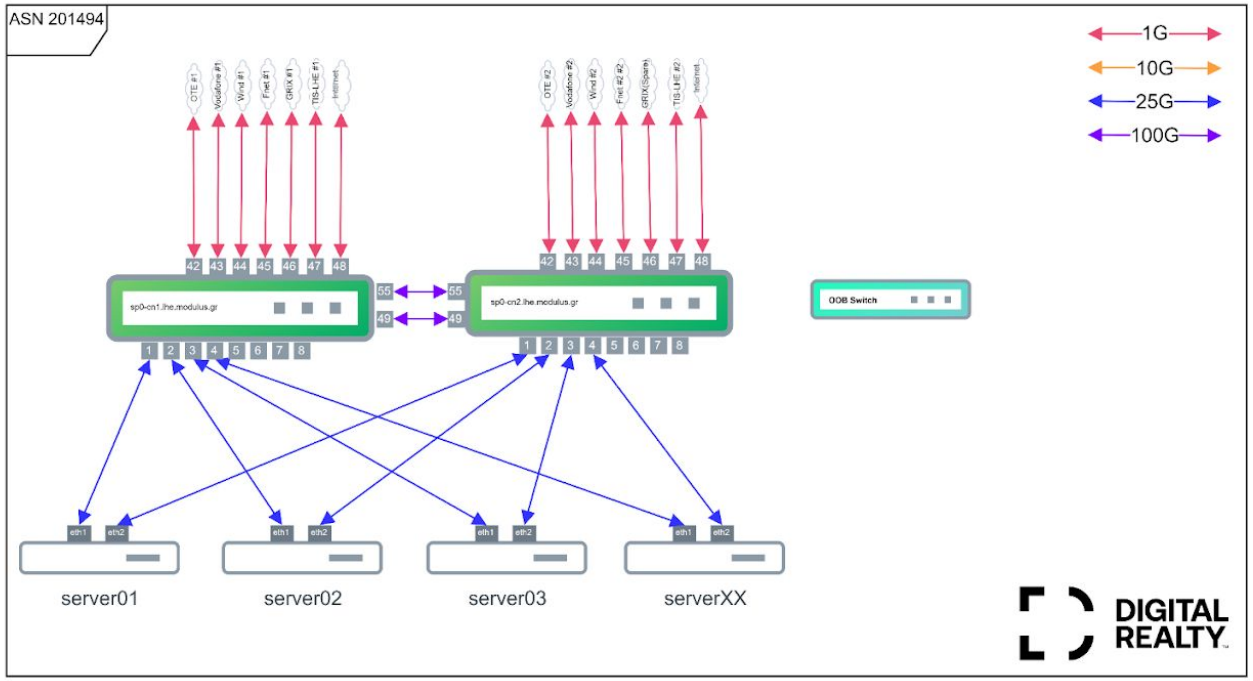
2022

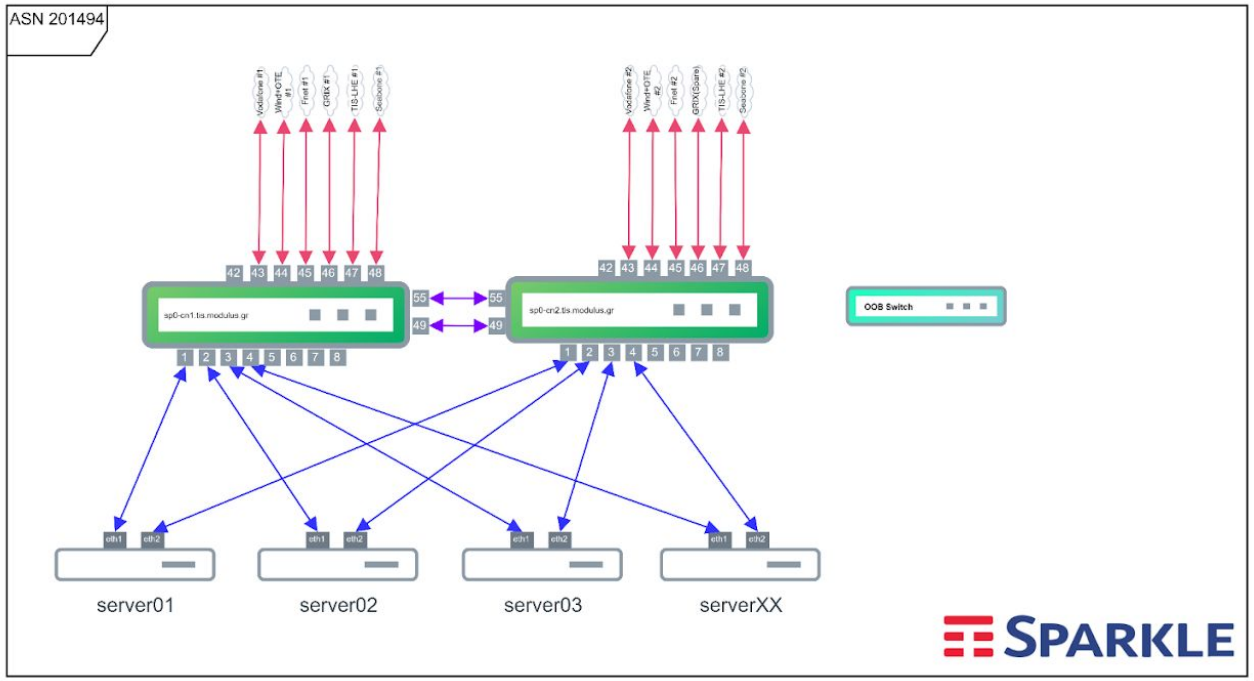
In **2022** we completed interconnections with all Greek Tier 1 providers for Fixed & Mobile Networks

modulus®


modulus
Infrastructure

High Level Design





modulus®

Some Details

- We have presence in two DCs (Ti Sparkle & Digital Realty)
- We are using ARISTA Networks for our network infrastructure
- We have two boxes per DC and we are using “Spline” topology (Leaf & Spine in the same box)
- We have MLAG between boxes with 2x100G per DC
- We have two L2 P2P redundant links between two DCs. We have decided to configure them as L3 links to separate the STP domains
- All of our peers, upstreams and servers are connected via two FO pairs in LACP
- We use isolated VRFs for every national provider that we are peering
- As for internal routing, we are using OSPF

modulus®

Internet Connectivity

Internet Upstreams

Sparkle | Seabone

DIGITAL REALTY™

Internet Exchanges

GRIX powered by grnet

linx

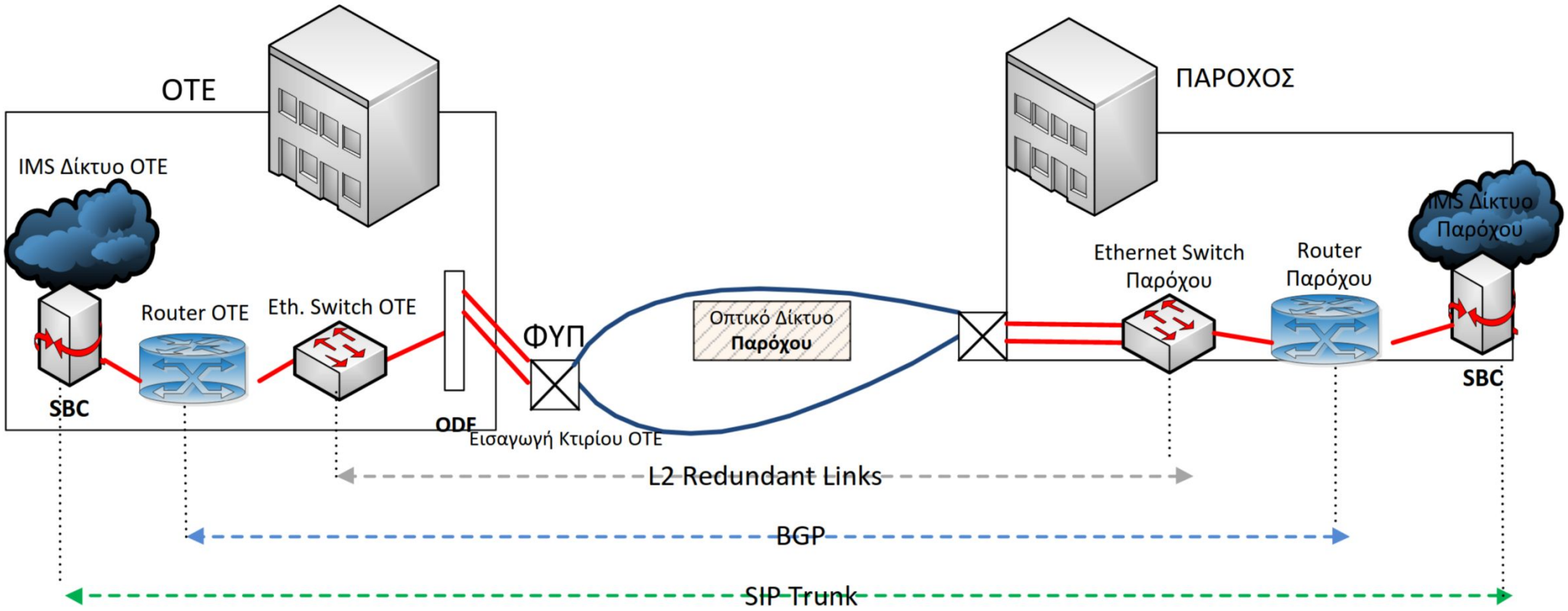
DE CIX

modulus®

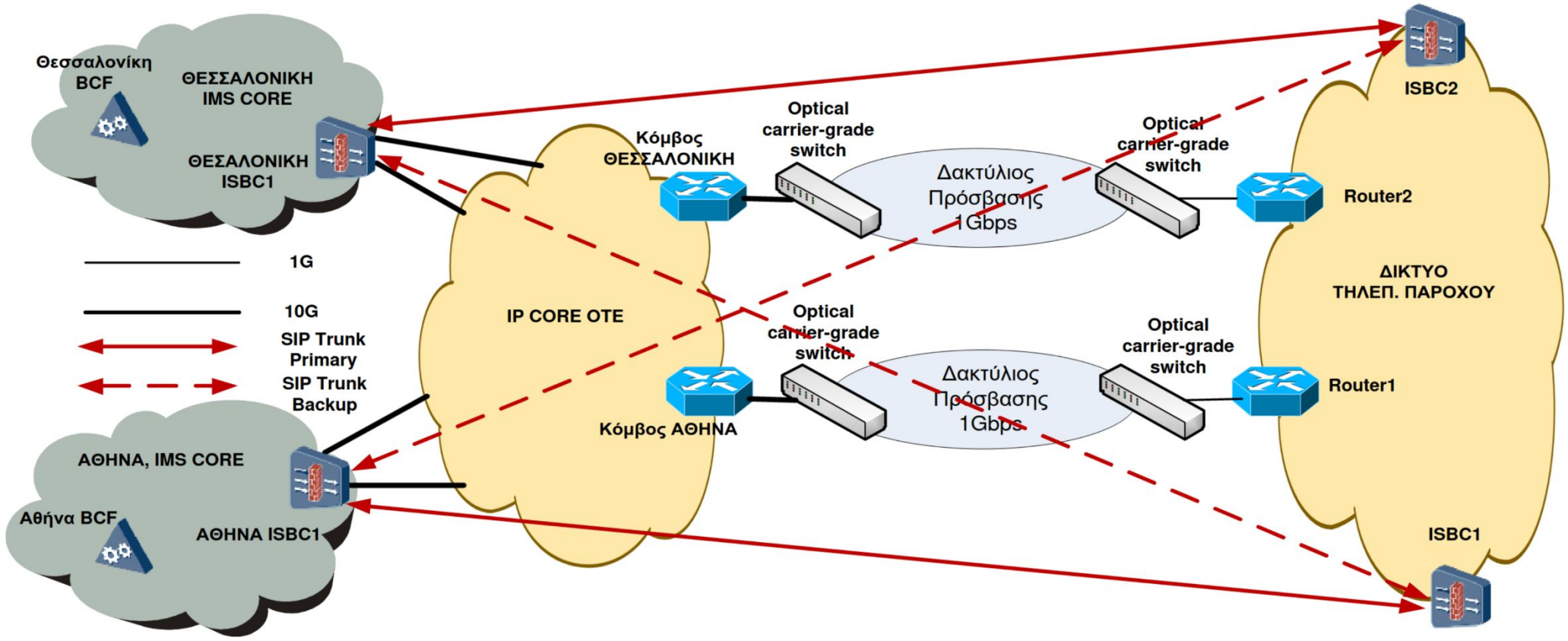
Private Network Interconnections (PNIs)

modulus®

PNIs Network Topology (High-Level)



IMS Topology (High-Level)



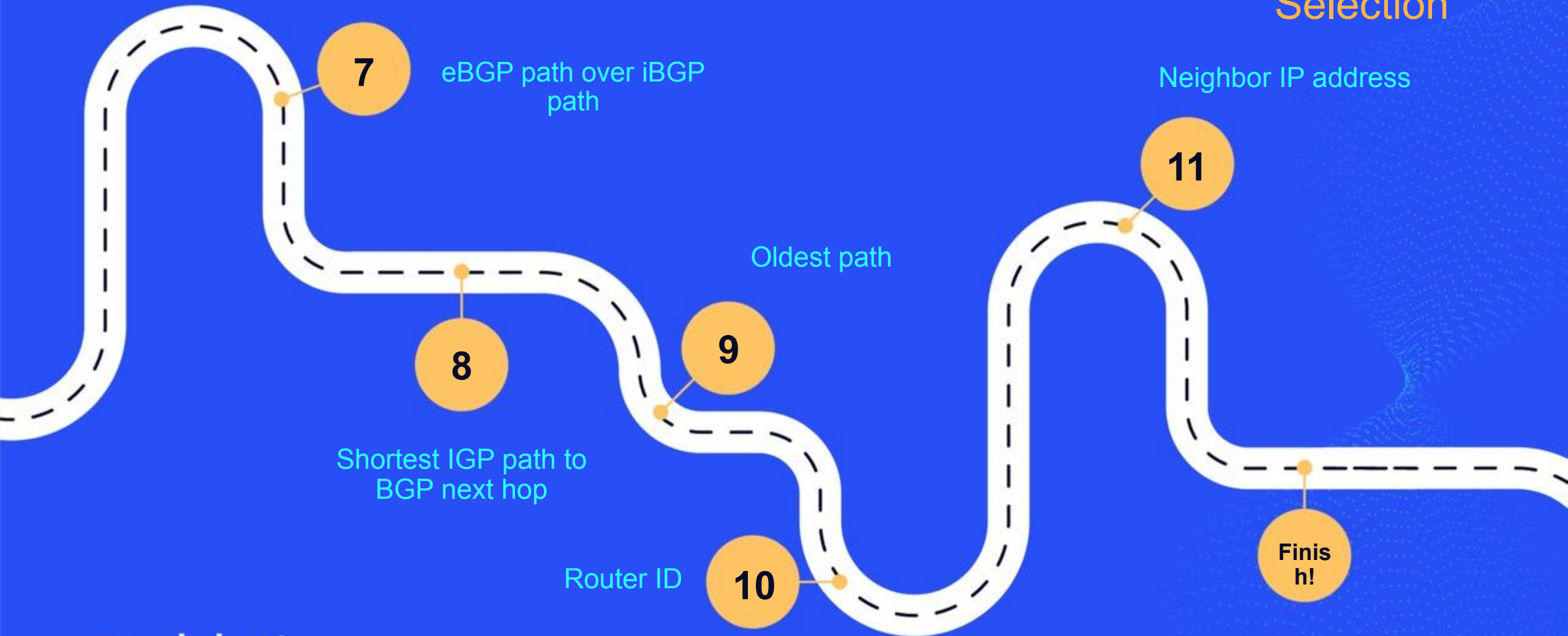
modulus®

How BGP is being used

BGP Attributes and Path Selection

modulus®

BGP Attributes and Path Selection



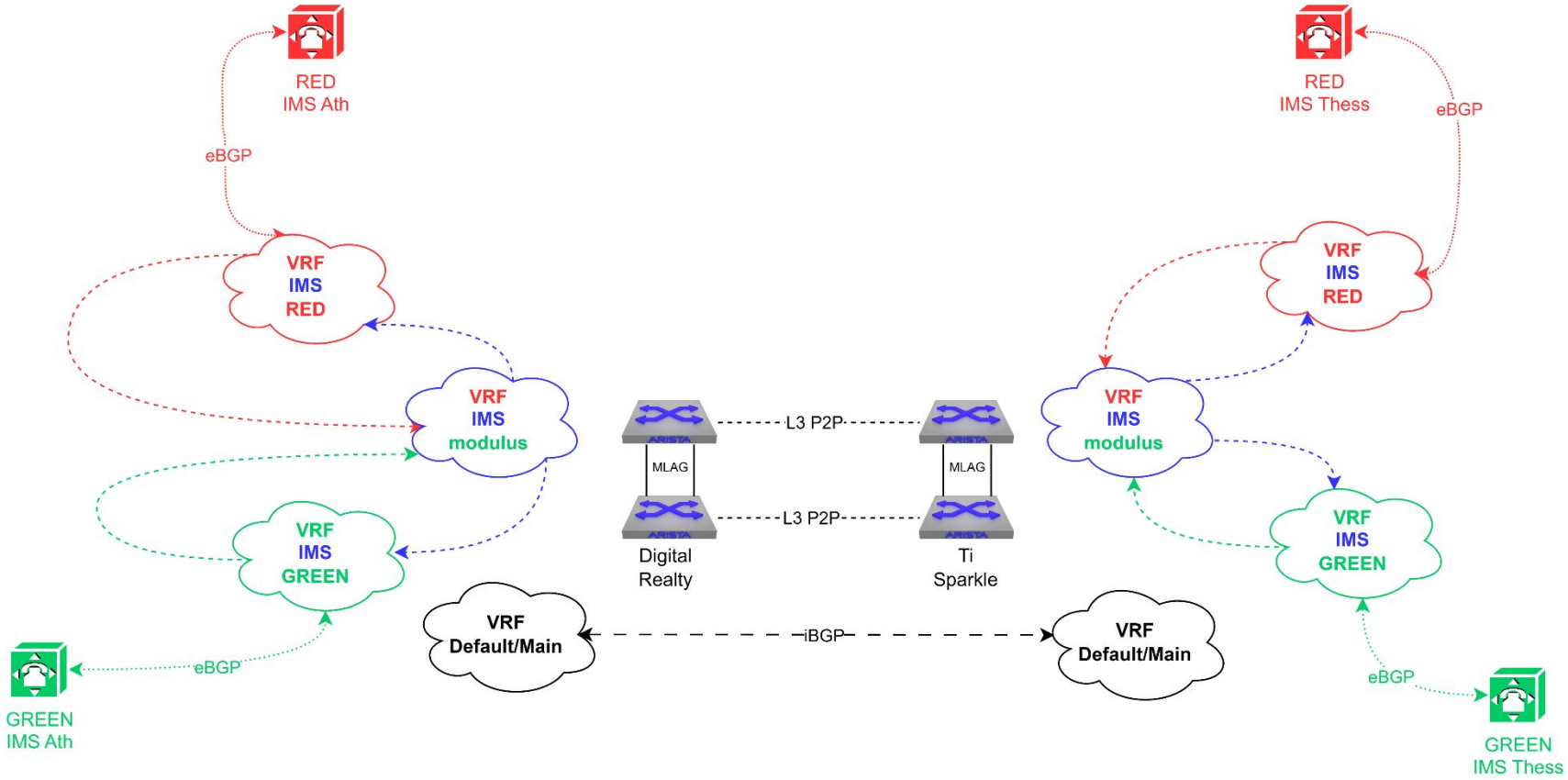
modulus®

We Love MED Because..

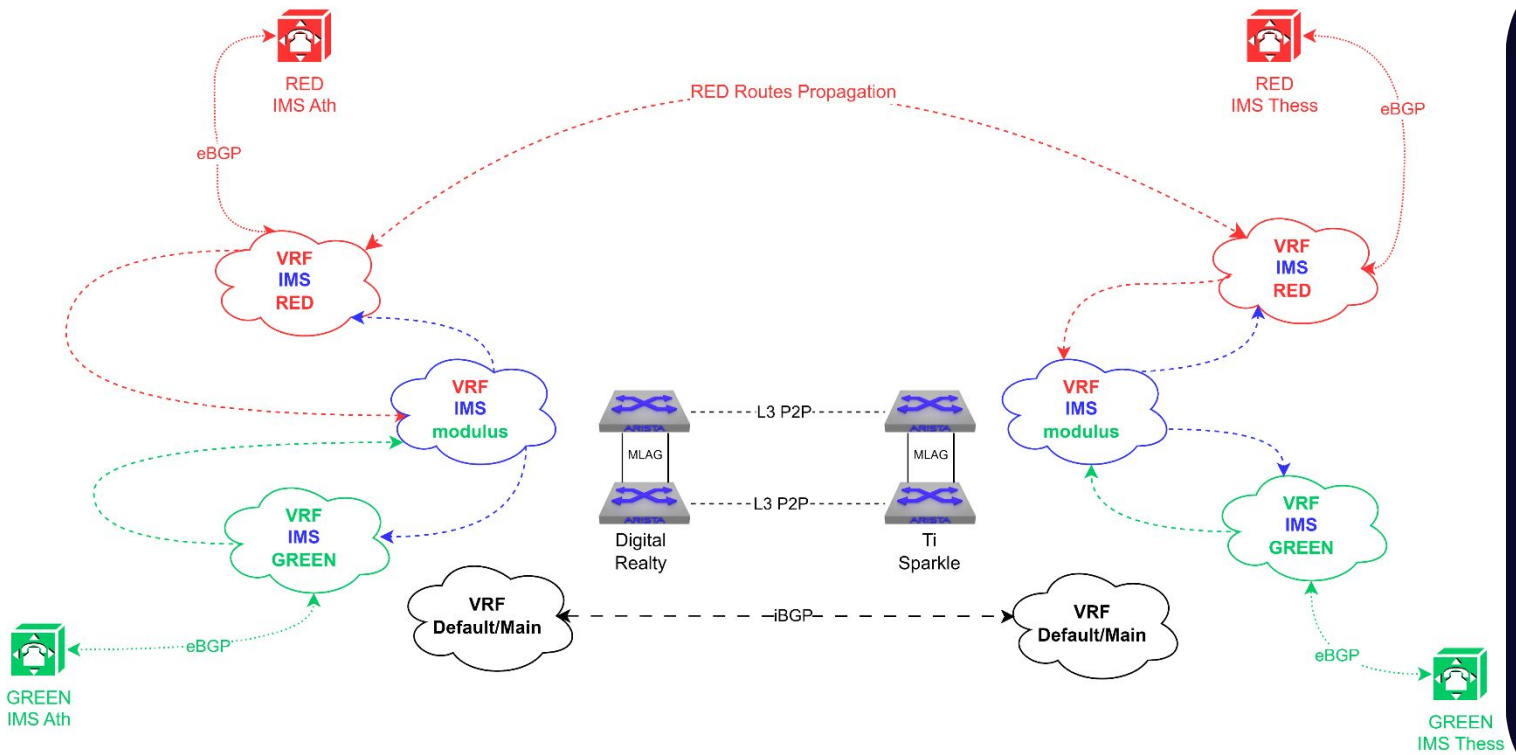
- It is not propagated across the entire Internet but is limited to adjacent autonomous systems (ASs).
- MED is capable to influence how other ASs enter your AS to reach a certain prefix.
- The AS path rather than the number of hops (4th) between hosts is the primary criterion for BGP route selection; an AS with multiple connections to a peer AS can have multiple equivalent AS paths. When the routing table contains two routes to the same host in a neighboring AS, a MED (6th) metric assigned to each route can determine which to include in the forwarding table.
- MED and BGP Maintenance (RFC6198) help us to make hitless upgrades.
- The smallest MED Wins!

modulus®

BGP Route Advertisements



BGP Route Advertisements



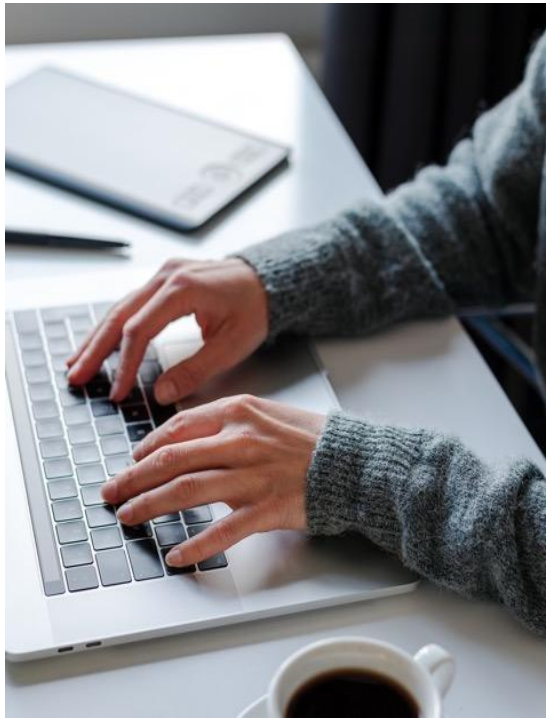
We aim to achieve the following:

- Propagate all received prefixes while preserving the original MED values.
- Leak the prefixes between VRFs.
- Route traffic for each prefix to the peer with the smallest MED value.
- Simplify the configuration as much as possible.

The ideal scenario proved to be impractical to implement

- To implement route leaking across all routers, we needed to build a full mesh for every VRF.
- With 4 routers, each router would require (N-1) iBGP neighbors and (N*(N-1))/2 = 6BGP sessions. Multiplying this by 6 VRFs results in 36iBGP sessions in total. Adding another provider would further increase the session count.
- Managing such a large number of sessions became very challenging, so we began exploring alternatives.
- The most common alternative, was MPLS.
- Unfortunately, we received the following mail from our vendor...

modulus®



Hi Kostas,

So basically you have two sites each one with few VRFs and you need to share routes between them?
This can be achieved using MP-BGP protocol.
The data plane can be MPLS (incase of vpnv4 routes) or VxLAN (incase of evpn routes) your platform does not support MPLS so we left with VxLAN, which is anyway the best practice and the preferred option.

Configure BGP session with AF EVPN between sites (loopback interfaces)
Example:

```
router bgp 100
   router-id 172.16.0.1
   neighbor 172.16.0.2 remote-as 200
   neighbor 172.16.0.2 update-source Loopback0
   neighbor 172.16.0.2 ebgp-multihop
   neighbor 172.16.0.2 send-community
   neighbor 172.16.0.2 maximum-routes 12000
```

Than activate the AF EVPN

```
   !
   address-family evpn
      neighbor 172.16.0.2 activate
   !
```

Configure the VRF inside the BGP protocol, and use the import/export command to share routes as you wish

```
  vrf VRF1
      rd 172.16.0.1:1
      route-target import evpn 1:1
      route-target export evpn 1:1
      route-target export evpn 2:2
      redistribute connected
   !
   vrf VRF2
      rd 172.16.0.1:2
      route-target import evpn 1:1
      route-target import evpn 2:2
```

Dont forget to create Vxlan Tunnel and add the VRFs

```
interface Vxlan1
   vxlan source-interface Loopback0
   vxlan udp-port 4789
   vxlan vrf VRF1 vni 100000
   vxlan vrf VRF2 vni 200000
```

More detailed example can be found on this link:
https://eos.arista.com/l3-evpn-vxlan-configuration-guide/

Our VXLAN journey just started

What is VXLAN (Virtual eXtensible Local-Area Network)?

- VXLAN is a network encapsulation technology that enables the creation of virtualized Layer 2 networks over a Layer 3 infrastructure..
- The VXLAN tunneling protocol encapsulates Layer 2 Ethernet frames in Layer 4 UDP packets.
- Each subnet is uniquely identified by a VXLAN Network Identifier (VNI).
- The device responsible for encapsulating and decapsulating packets is known as the VXLAN Tunnel Endpoint (VTEP). A VTEP can be a physical network device, such as a router or switch, or a virtual switch deployed on a server.
- It's a data plane technology/protocol

modulus®

What is EVPN

- EVPN is an extension to Border Gateway Protocol (BGP) that allows the network to carry endpoint reachability information such as Layer 2 MAC addresses or IP routes
- It is a control plane technology that uses MP-BGP for distribution, adding an additional address family.
- Since MAC learning is handled in the control plane, EVPN avoids the flooding typical of Layer 2 networks. It is a loop-free technology that eliminates the need for Spanning Tree Protocol (STP)
- In EVPN-VXLAN architectures, VXLAN provides the overlay data-plane encapsulation
- Active-active redundant links fully utilizing network bandwidth

modulus®

VXLAN/EVPN Pros...

VXLAN/EVPN Cons...

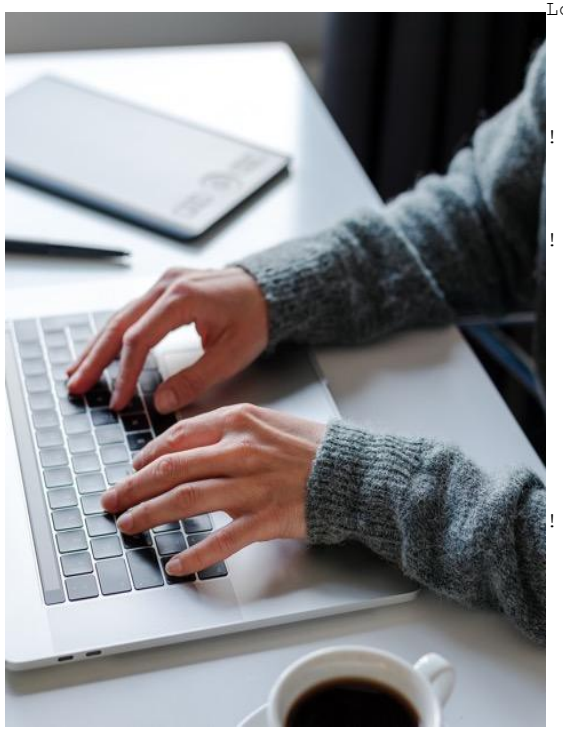
- We don’t know anything so far…

modulus®

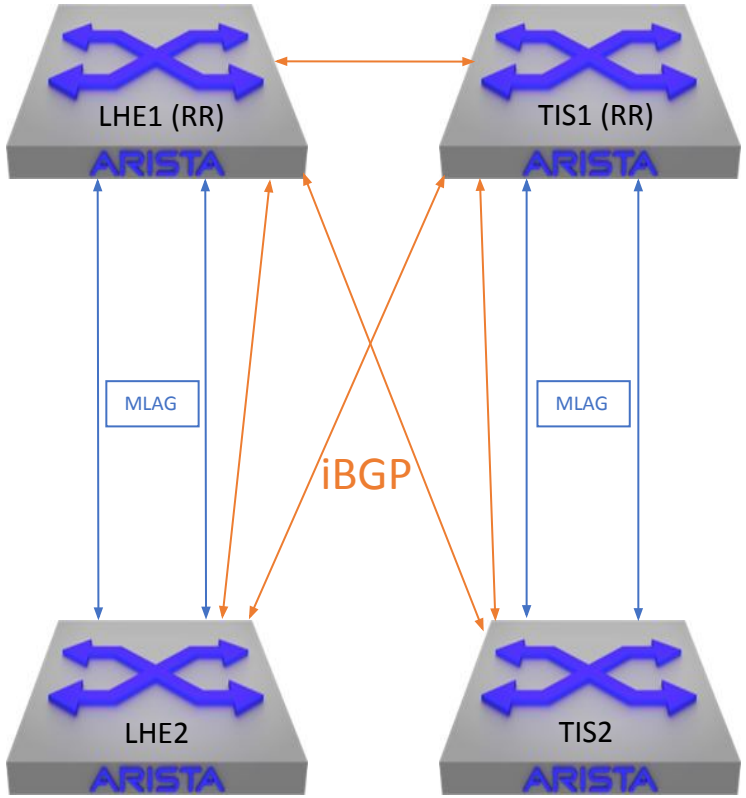
modulus

Configuration

iBGP Configuration



```
!iBGP Config for LHE1
  neighbor REFLECTORS peer group
  neighbor REFLECTORS remote-as 201494
  neighbor REFLECTORS next-hop-self
  neighbor REFLECTORS update-source
Loopback0
  neighbor REFLECTORS bfd
  neighbor REFLECTORS idle-restart-timer 60
  neighbor REFLECTORS send-community
  neighbor $TIS1 peer group REFLECTORS
!
  address-family evpn
    neighbor REFLECTORS activate

!           =========
  neighbor CLIENTS peer group
  neighbor CLIENTS remote-as 201494
  neighbor CLIENTS next-hop-self
  neighbor CLIENTS update-source Loopback0
  neighbor CLIENTS bfd
  neighbor CLIENTS route-reflector-client
  neighbor CLIENTS idle-restart-timer 60
  neighbor CLIENTS send-community
  neighbor $TIS2 peer group CLIENTS
  neighbor $LHE2 peer group CLIENTS
!
  address-family evpn
    neighbor CLIENTS activate
```

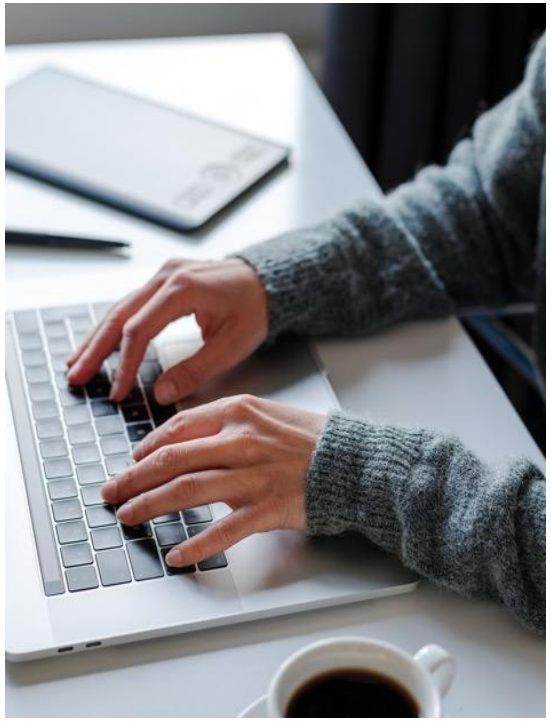



```
!iBGP Config for TIS1
  neighbor REFLECTORS peer group
  neighbor REFLECTORS remote-as 201494
  neighbor REFLECTORS next-hop-self
  neighbor REFLECTORS update-source
Loopback0
  neighbor REFLECTORS bfd
  neighbor REFLECTORS idle-restart-timer 60
  neighbor REFLECTORS send-community
  neighbor $LHE1 peer group REFLECTORS
!
  address-family evpn
    neighbor REFLECTORS activate

!           =========
  neighbor CLIENTS peer group
  neighbor CLIENTS remote-as 201494
  neighbor CLIENTS next-hop-self
  neighbor CLIENTS update-source Loopback0
  neighbor CLIENTS bfd
  neighbor CLIENTS route-reflector-client
  neighbor CLIENTS idle-restart-timer 60
  neighbor CLIENTS send-community
  neighbor $TIS2 peer group CLIENTS
  neighbor $LHE2 peer group CLIENTS
!
  address-family evpn
    neighbor CLIENTS activate
```

```
!iBGP Config for LHE2
  neighbor REFLECTORS peer group
  neighbor REFLECTORS remote-as 201494
  neighbor REFLECTORS next-hop-self
  neighbor REFLECTORS update-source
Loopback0
  neighbor REFLECTORS bfd
  neighbor REFLECTORS idle-restart-timer 60
  neighbor REFLECTORS send-community
  neighbor $LHE1 peer group REFLECTORS
  neighbor $TIS1 peer group REFLECTORS
!
  address-family evpn
    neighbor REFLECTORS activate
```

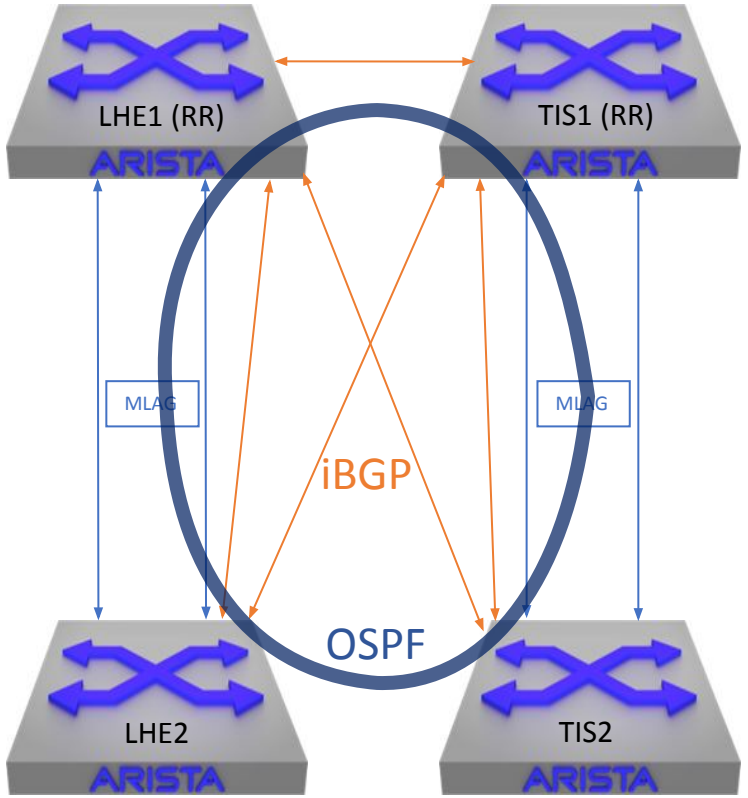
```
!iBGP Config for TIS2
  neighbor REFLECTORS peer group
  neighbor REFLECTORS remote-as 201494
  neighbor REFLECTORS next-hop-self
  neighbor REFLECTORS update-source
Loopback0
  neighbor REFLECTORS bfd
  neighbor REFLECTORS idle-restart-timer 60
  neighbor REFLECTORS send-community
  neighbor $LHE1 peer group REFLECTORS
  neighbor $TIS1 peer group REFLECTORS
!
  address-family evpn
    neighbor REFLECTORS activate
```

VXLAN Configuration



```
LHE1#sh run int Lo1
interface Loopback1
  ip address 198.51.100.105/32
  ip ospf area 0.0.0.0

LHE1#sh run sec vxlan1
interface vxlan1
   vxlan source-interface
Loopback1
   vxlan udp-port 4789
   vxlan vrf ims-modulus vni 1111
   vxlan vrf ims-red vni 2222
   vxlan vrf ims-green vni 3333
   vxlan flood vtep $Lo1 (TIS)
```

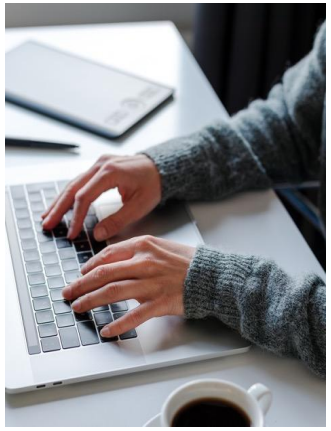
```
LHE1#sh run int Lo1
interface Loopback1
  ip address 198.51.100.105/32
  ip ospf area 0.0.0.0

LHE1#sh run sec vxlan1
interface vxlan1
   vxlan source-interface
Loopback1
   vxlan udp-port 4789
   vxlan vrf ims-modulus vni 1111
   vxlan vrf ims-red vni 2222
   vxlan vrf ims-green vni 3333
   vxlan flood vtep $Lo1 (TIS)
```

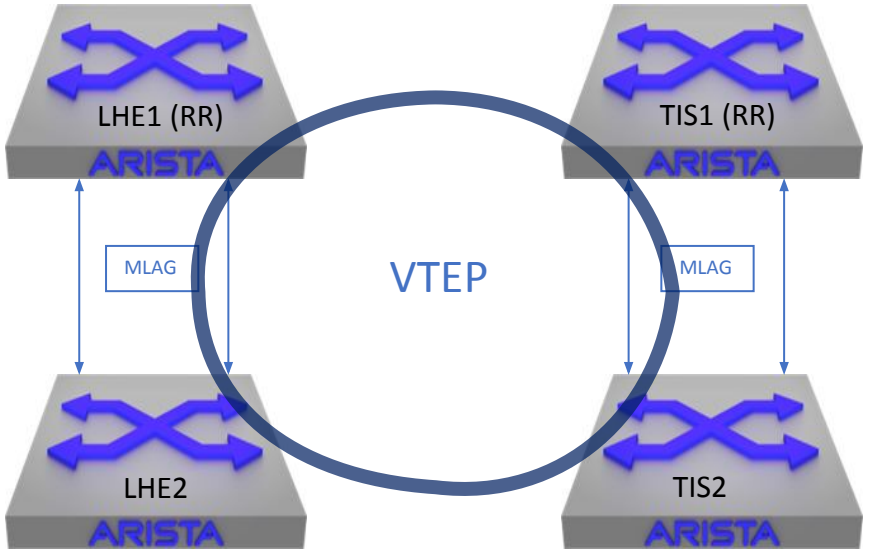



```
LHE1#sh run int Lo1
interface Loopback1
  ip address 198.51.100.106/32
  ip ospf area 0.0.0.0

LHE1#sh run sec vxlan1
interface vxlan1
   vxlan source-interface
Loopback1
   vxlan udp-port 4789
   vxlan vrf ims-modulus vni 1111
   vxlan vrf ims-red vni 2222
   vxlan vrf ims-green vni 3333
   vxlan flood vtep $Lo1 (LHE)
```

```
LHE1#sh run int Lo1
interface Loopback1
  ip address 198.51.100.106/32
  ip ospf area 0.0.0.0

LHE1#sh run sec vxlan1
interface vxlan1
   vxlan source-interface
Loopback1
   vxlan udp-port 4789
   vxlan vrf ims-modulus vni 1111
   vxlan vrf ims-red vni 2222
   vxlan vrf ims-green vni 3333
   vxlan flood vtep $Lo1 (LHE)
```

VXLAN / EVPN






```
!VXLAN Config (LHE1 & LHE2)
  interface Vxlan1
  vxlan flood vtep $TIS1(Loopback)

! VRF Instances
  vrf instance ims-modulus
    rd 1111:10
  vrf instance ims-red
    rd 2222:10
  vrf instance ims-green
    rd 3333:10

! Enable vrf routing
  ip routing vrf ims-modulus
  ip routing vrf ims-red
  ip routing vrf ims-green

! EVPN Config
  vxlan vrf ims-modulus vni 1111
  vxlan vrf ims-red vni 2222
  vxlan vrf ims-green vni 3333

! Route Leaking between VRFs

! modulus
vrf ims-modulus
   rd 1111:10
   route-target export evpn 1111:10
   route-target import evpn 1111:10
   route-target import evpn 2222:10
   route-target import evpn 2222:10
   route-target import evpn 3333:10
   route-target import evpn 3333:10
   redistribute static
   redistribute connected route-map
ims-route-leaking

! red
vrf ims-red
  rd 2222:10
    route-target export evpn 2222:10
    route-target import evpn 2222:10
    route-target import evpn 1111:10

! green
vrf ims-green
  rd 3333:10
    route-target export evpn 3333:10
    route-target import evpn 3333:10
    route-target import evpn 1111:10
```

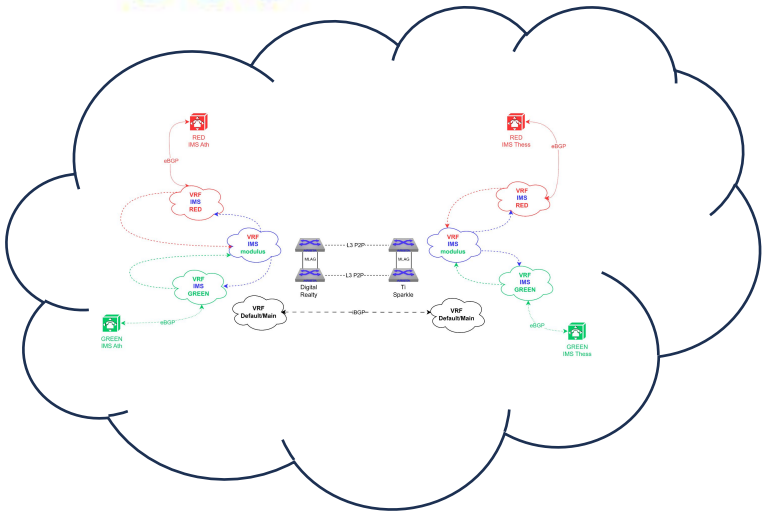
```
!VXLAN Config (TIS1 & TIS2)
  interface Vxlan1
  vxlan flood vtep $LHE1(Loopback)

! VRF Instances
  vrf instance ims-modulus
    rd 1111:10
  vrf instance ims-red
    rd 2222:10
  vrf instance ims-green
    rd 3333:10

! Enable vrf routing
  ip routing vrf ims-modulus
  ip routing vrf ims-red
  ip routing vrf ims-green

! EVPN Config
  vxlan vrf ims-modulus vni 1111
  vxlan vrf ims-red vni 2222
  vxlan vrf ims-green vni 3333

! Route Lacking between VRFs

! modulus
vrf ims-modulus
   rd 1111:10
   route-target export evpn 1111:10
   route-target import evpn 1111:10
   route-target import evpn 2222:10
   route-target import evpn 2222:10
   route-target import evpn 3333:10
   route-target import evpn 3333:10
   redistribute static
   redistribute connected route-map ims-route-leaking

! red
vrf ims-red
  rd 2222:10
    route-target export evpn 2222:10
    route-target import evpn 2222:10
    route-target import evpn 1111:10

! green
vrf ims-green
  rd 3333:10
    route-target export evpn 3333:10
    route-target import evpn 3333:10
    route-target import evpn 1111:10
```

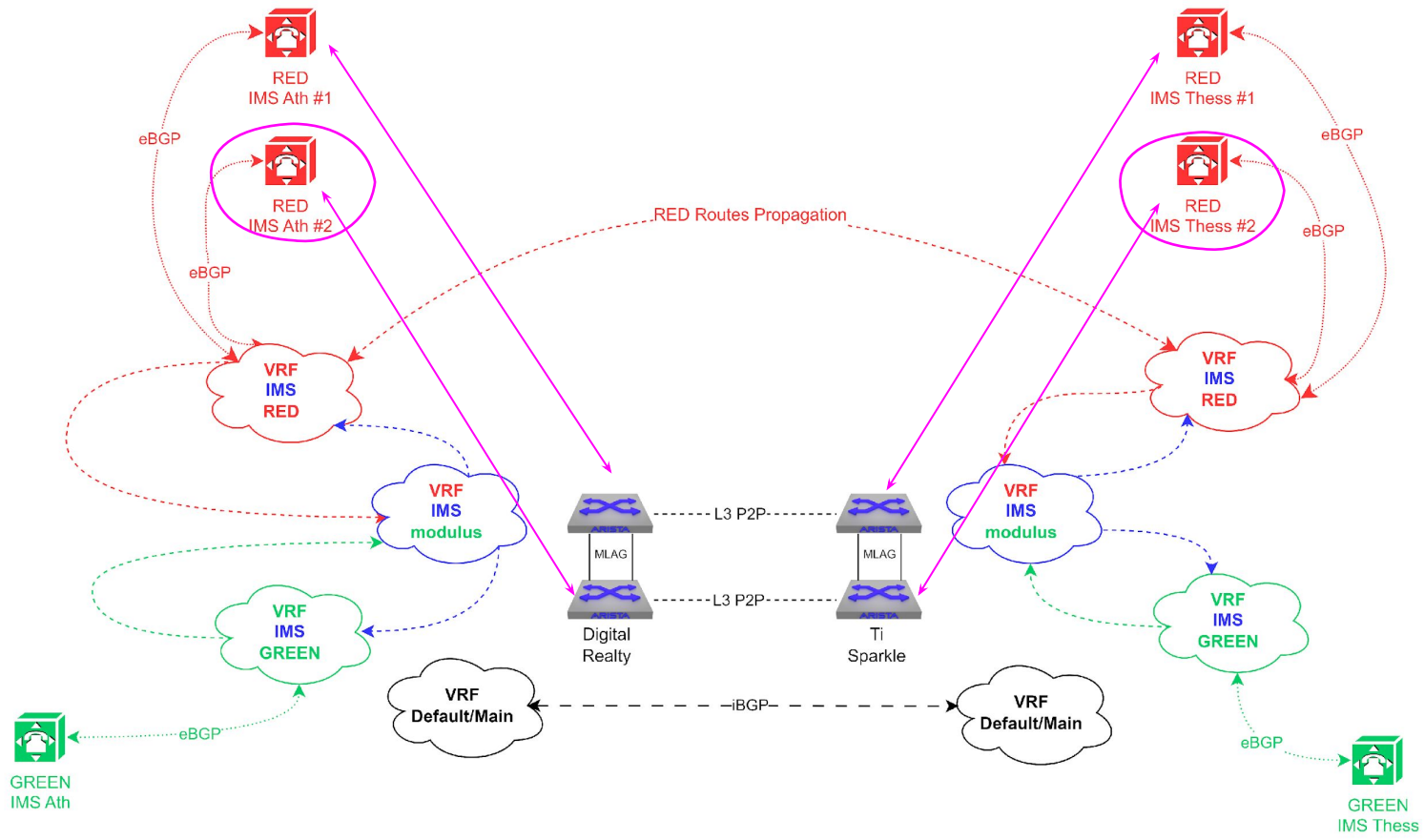
modulus®

modulus

Results

BGP Route Advertisements



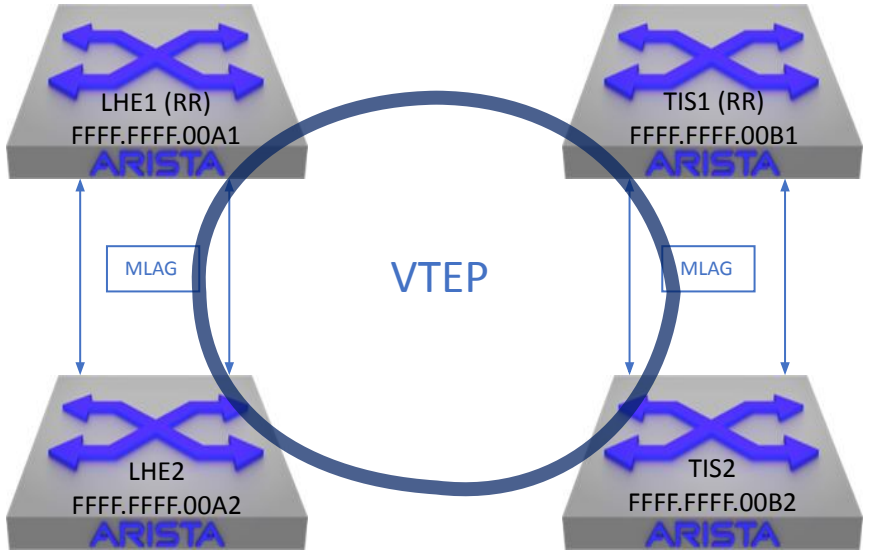


modulus®

VXLAN / EVPN

```
show ip bgp neighbors $RED_LHE1 received-routes vrf ims-red
          Network          Next Hop        Metric  AIGP    LocPref Weight  Path
*      I 192.0.2.0/28      198.51.100.4    100     -       -       -       2222

show ip route vrf ims-red | i 192.0.2.0/28
B I      192.0.2.0/28 [200/5] via VTEP $TIS1 VNI 2222 router-mac FFFF.FFFF.00B1
```

```
show ip bgp neighbors $RED_TIS1 received-routes vrf ims-red
          Network          Next Hop        Metric  AIGP    LocPref Weight  Path
*      I 192.0.2.0/28      198.51.100.0    5       -       -       -       2222

show ip route vrf ims-red | i 192.0.2.0/28
B E 192.0.2.0/28 [20/5] via 198.51.100.0, Vlan611
```




```
show ip bgp neighbors $RED_LHE2 received-routes vrf ims-red
          Network          Next Hop        Metric  AIGP    LocPref Weight  Path
*      I 192.0.2.0/28      198.51.100.6    110     -       -       -       2222

show ip route vrf ims-red | i 192.0.2.0/28
B I      192.0.2.0/28 [200/5] via VTEP $TIS1 VNI 2222 router-mac FFFF.FFFF.00B1
```

```
show ip bgp neighbors $RED_TIS2 received-routes vrf ims-red
          Network          Next Hop        Metric  AIGP    LocPref Weight  Path
*      I 192.0.2.0/28      198.51.100.2    10      -       -       -       2222

show ip route vrf ims-red | i 192.0.2.0/28
B I      192.0.2.0/28 [200/5] via VTEP $TIS1 VNI 2222 router-mac FFFF.FFFF.00B1
```

modulus®

Our goals!

Your Voice, Our Connection!

- An absolutely trouble-free solution
- Easy setup & maintenance
- Scalable and flexible solution

modulus

www.modulus.gr

Thank you!