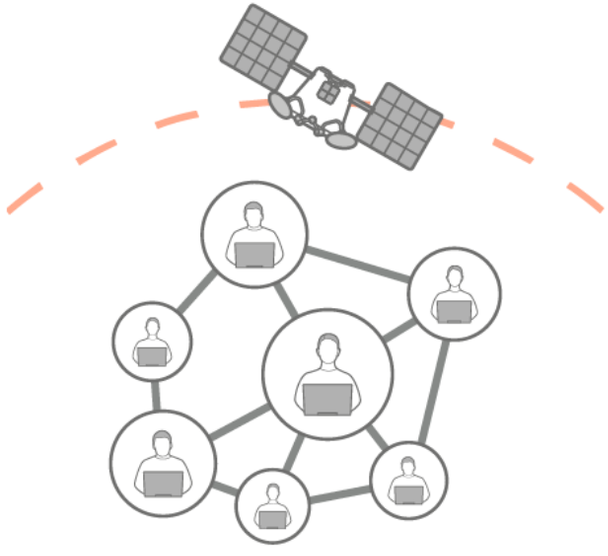
Multicast Routing

@ipeng

OneWeb

Mission

Connectivity from Space through an innovative Low-Earth-Orbit satellite network



Approach

Partnering with local stakeholders; reaching users together

OneWeb

Governments

Regulatory Agencies

Telcos

MNOs

Achieving

Broadband speeds
Supports all major applications

Low latency
At-home experience, anywhere

Global coverage
Access anywhere; multiple use cases

OneWeb

Supply

Demand

Mass satellite production	Launch costs reduction	New usage	Data consumption

Performance and coverage improves as you get closer to Earth

LEO 1,200km orbit
~ 50ms Latency

OneWeb

MEO
> 135ms Latency

GEO
> 560ms Latency

Low latency unlocks high demand applications

Gaming **<100ms**

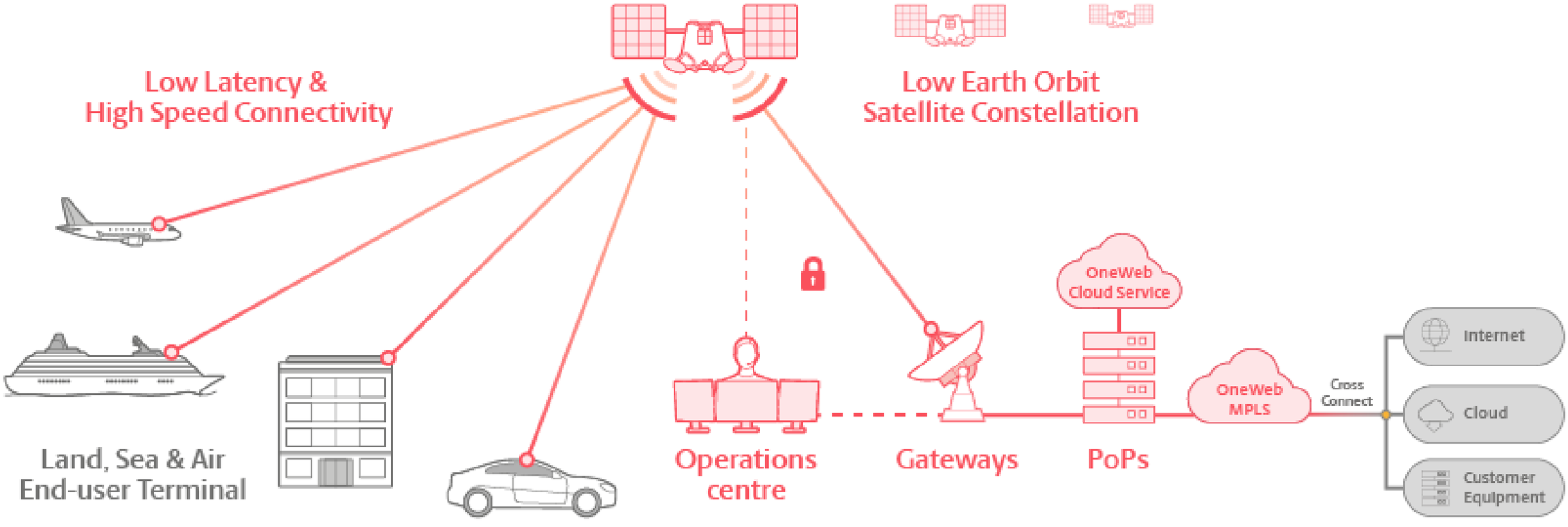
FORTNITE XBOX LIVE oculus

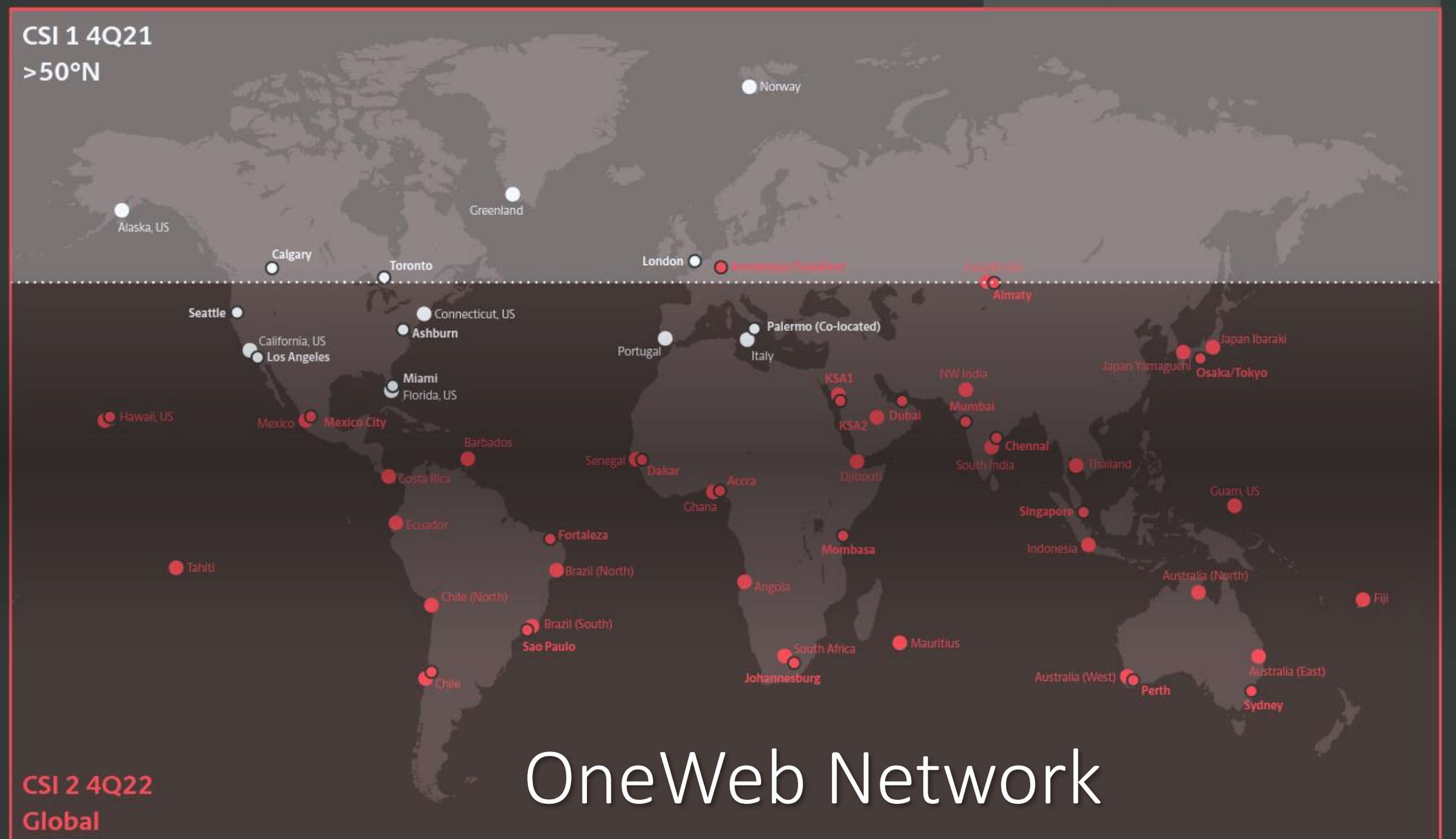
VoLTE **<150ms**

vodafone Super VoLTE AT&T HDVoice verizon VoLTE

Cloud **<200ms**

Google Cloud Azure aws ORACLE CLOUD SAP


Low Latency &
High Speed Connectivity
Low Earth Orbit
Satellite Constellation
OneWeb
Cloud Service
OneWeb
MPLS
Cross
Connect
Internet
Cloud
Customer
Equipment
Land, Sea & Air
End-user Terminal
Operations
centre
Gateways
PoPs



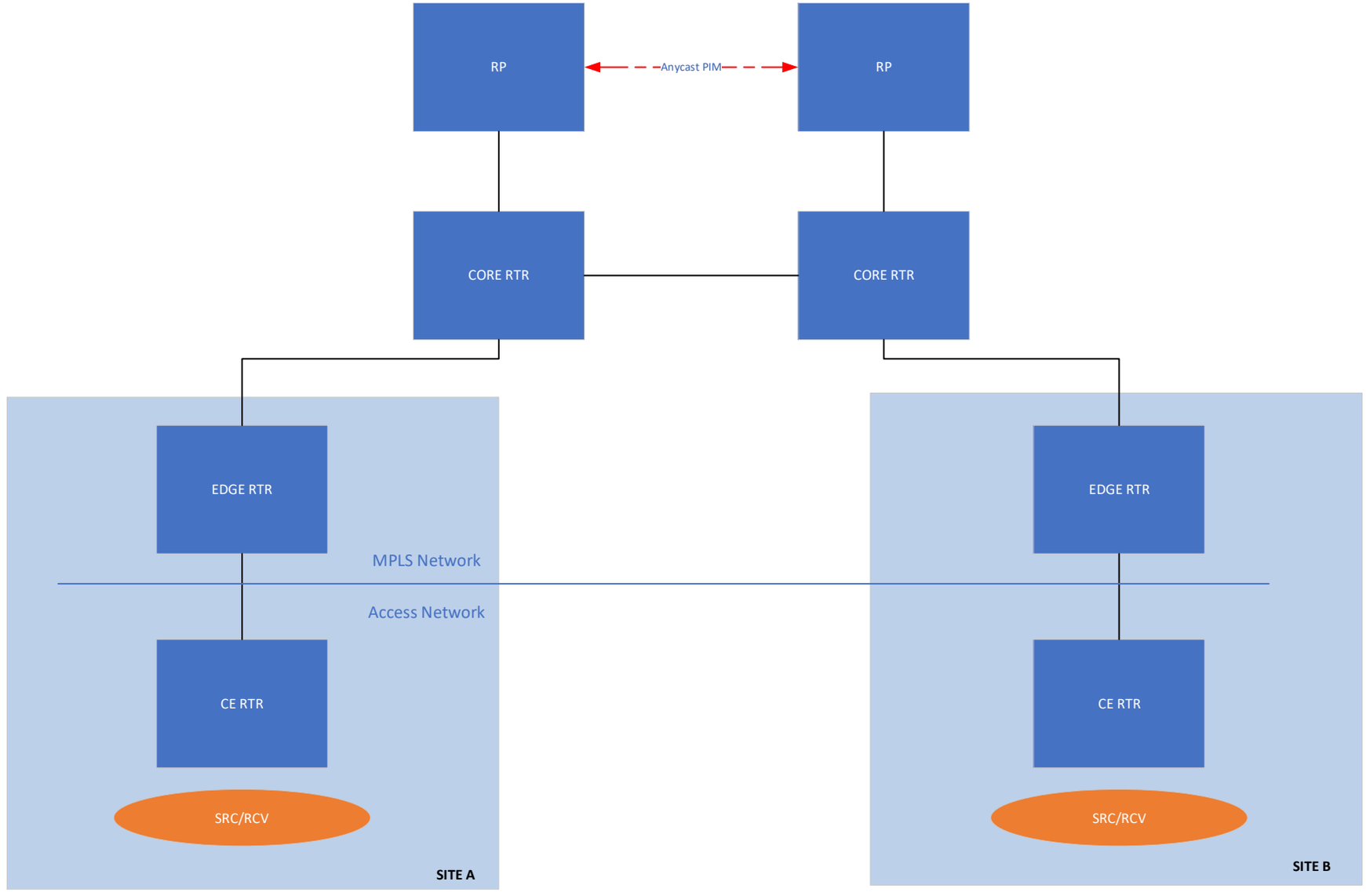
OneWeb Network

OneWeb Network Topology

Access:
- CE with services in VRF
- BGP PE-CE routing

EDGE/CORE:
- ISIS
- LDP(legacy) and SR(pref)
- MP-BGP
- PIM ASM

RP:
- Anycast PIM



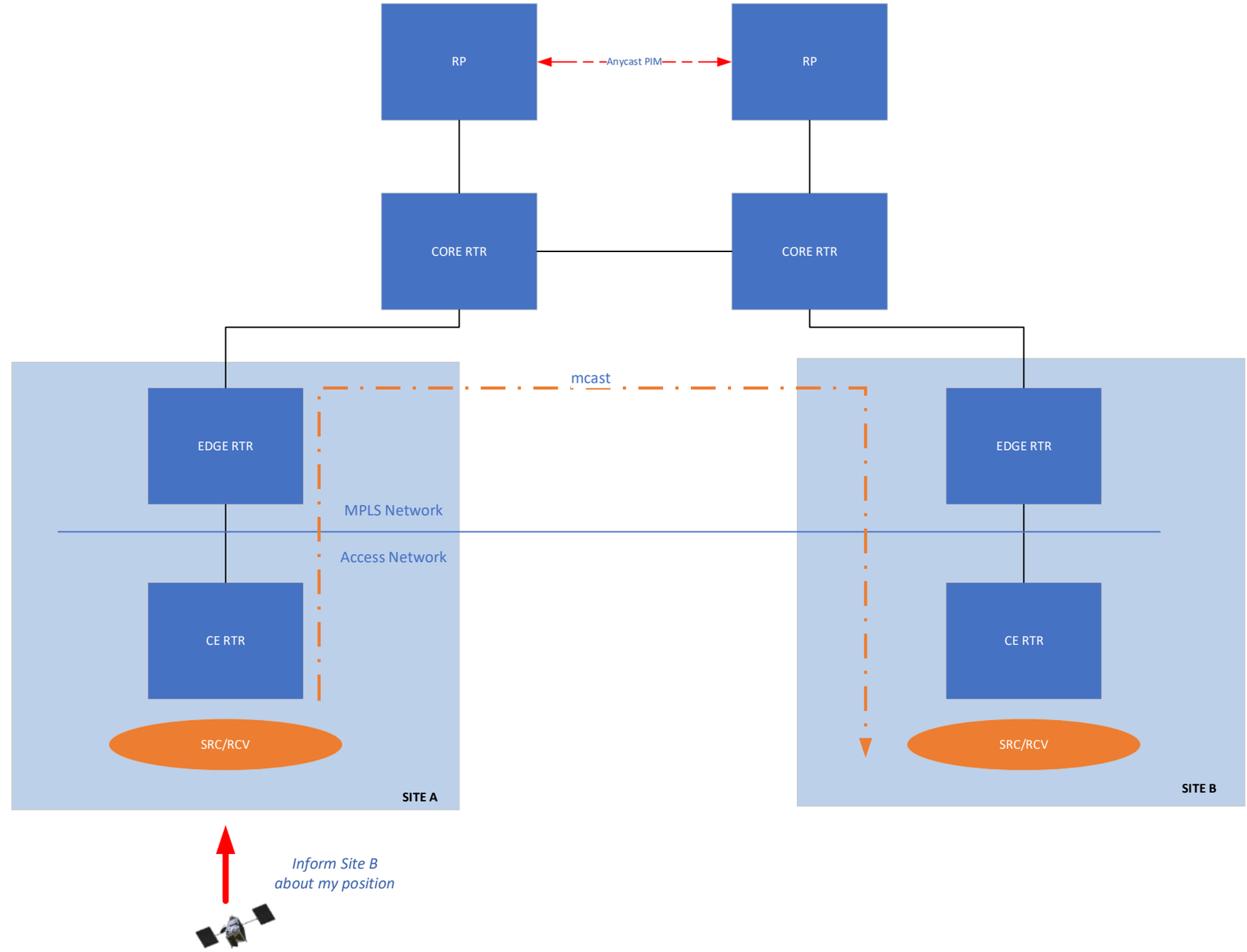
OneWeb Multicast

Messages:

- Satellite telemetry
- Site components telemetry(Antennas , and auxiliary systems)
- Data handover
- 150+ sources per Site

Today discussion:

- Access Site A needs to inform Access SITE B a satellite handover is about to happen.



OneWeb Multicast

Typical site traffic:

- gRPC info

A:lonXXXprrb# show router bgp neighbor 10.XX.XX.1 advertised-routes mvpn-ipv4 brief | match "Routes :"
Routes : **8840**
A:longXXXprrb# show router bgp neighbor 10.XXX.XXX.1 advertised-routes mvpn-ipv4 type source-ad | match "Routes :"
Routes : **7696**

OneWeb Multicast

challenges:

1. MoFRR (multicast fast reroute) not available in VRF context.
   We need sub-second convergence.
2. P nodes need to keep PIM state
3. P nodes need BGP table (RPF checks)

Solution:

mVPN with MLDP

- LDP already active in OW network. Was used for unicast pre SR migration.
- mLDP fast reroute (ECMP / LFA dependent)
- Move PIM state to BGP
- MPLS switched multicast -> BGP free core

Multicast mVPN – refresher

OneWeb

10k feet - Multicast VRF Control Plane

RFC 6514 new BGP family

Taken from RFC 6513

“A PMSI is a conceptual "overlay" on the P-network with the following property: a PE in a given MVPN can give a packet to the PMSI, and the packet will be delivered to some or all of the other PEs in the MVPN, such that any PE receiving the packet will be able to determine the MVPN to which the packet belongs.”

Type 1 routes: I-PMSI to signal all PE’s in an mVPN. (analogue to default MDT)
Type 3 routes: S-PMSI to signal a subset of PE’s in mVPN (analogue to data MDT)
Type 5 routes: S-ACTIVE to signal that a PE has a active sending source.
Type 6 routes: Shared tree join to join a (*,G)
Type 7 routes: Source tree join to join a (S,G)

Multicast mVPN – refresher

mLDP - RFC6388

Used in I/S-PMSI as an attribute.

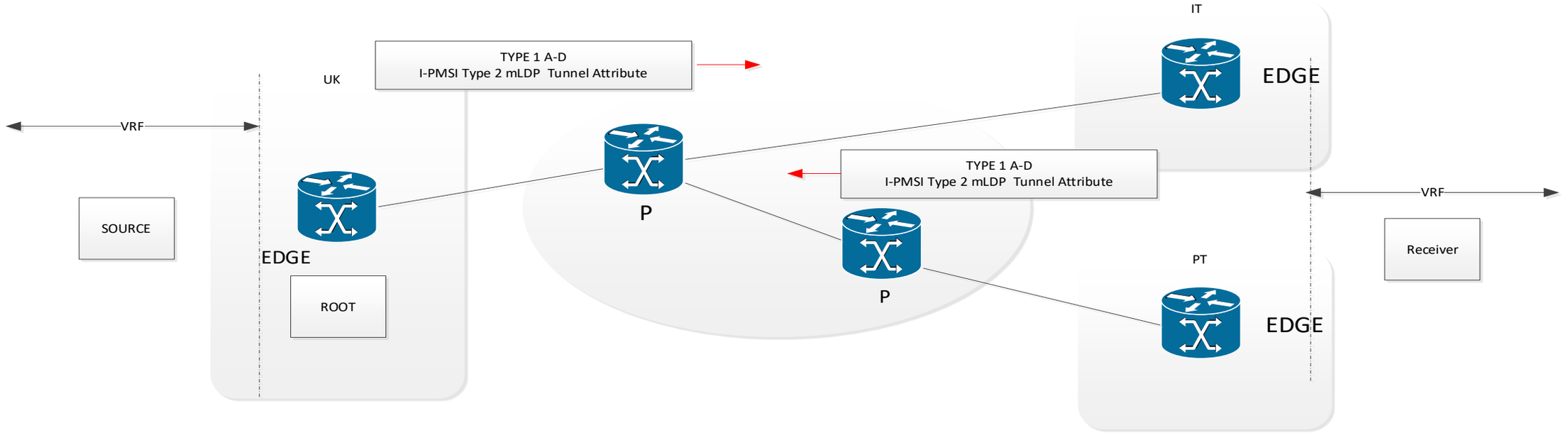
- Type 1 routes: I-PMSI to signal all PE's in an mVPN. (analogue to default MDT)
- Type 3 routes: S-PMSI to signal a subset of PE's in mVPN (analogue to data MDT)

```
Border Gateway Protocol - UPDATE Message
  Marker: ffffffffffffffffffffffffffffffff
  Length: 108
  Type: UPDATE Message (2)
  Withdrawn Routes Length: 0
  Total Path Attribute Length: 85
Path attributes
  Path Attribut - MP_REACH_NLRI
    Flags: 0x90: Optional, Non-transitive, Complete, Extended Length
    Type Code: MP_REACH_NLRI (14)
    Length: 23
    Address family: IPv4 (1)
    Subsequent address family identifier: MCAST-VPN (5)
    Next hop network address (4 bytes)
    Subnetwork points of attachment: 0
    Network layer reachability information (14 bytes)
      Route Type: Intra-AS I-PMSI A-D route (1)
      Length: 12
  Path Attribut - ORIGIN: IGP
  Path Attribut - AS_PATH: empty
  Path Attribut - LOCAL_PREF: 100
  Path Attribut - COMMUNITIES: NO_EXPORT
  Path Attribut - EXTENDED_COMMUNITIES
  Path Attribut - PMSI_TUNNEL_ATTRIBUTE
    Flags: 0xc0: Optional, Transitive, Complete
    Type Code: PMSI_TUNNEL_ATTRIBUTE (22)
    Length: 22
    Flags: 0
    Tunnel Type: mLDP P2MP LSP (2)
    MPLS Label Stack (withdrawn)
      MPLS Label: 0 (withdrawn)
    Tunnel ID: Type: P2MP FEC root node: 10.1.1.1 Id: 1
      mLDP P2MP FEC element type: P2MP FEC (6)
      mLDP P2MP FEC element address family: IPv4 (1)
      mLDP P2MP FEC element address length: 4
      mLDP P2MP FEC element root node address: 10.1.1.1 (10.1.1.1)
      mLDP P2MP FEC element opaque length: 7
      mLDP P2MP FEC element opaque value type: Generic LSP Identifier (1)
      mLDP P2MP FEC element opaque value length: 4
      mLDP P2MP FEC element opaque value unique Id: 1
Border Gateway Protocol - UPDATE Message
```

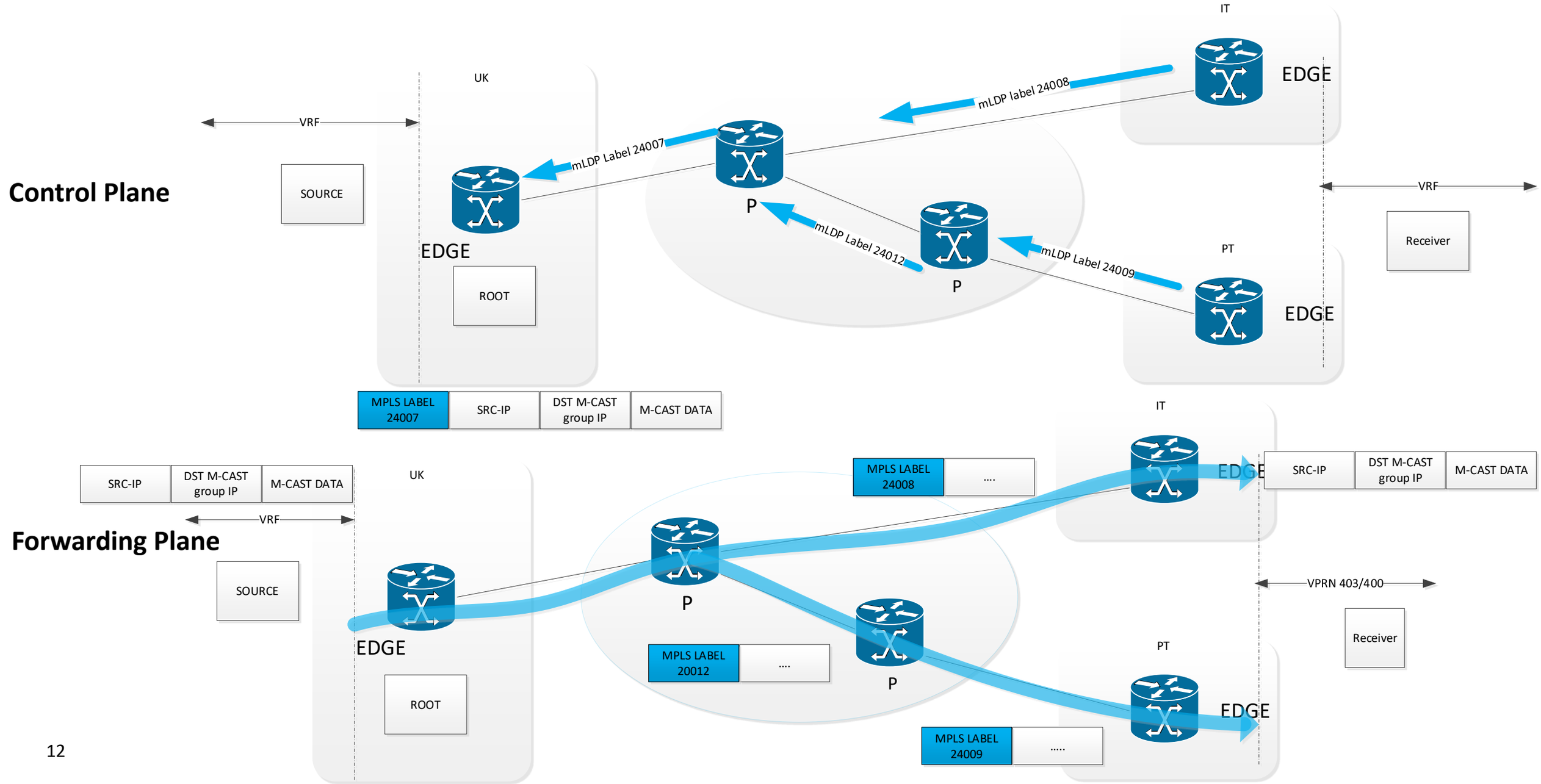
Type 1 – AD Route

Tunnel type 2

mLDP tunnel attributes



Multicast mVPN – refresher



OW Multicast mVPN - Migration

No big bang migration – Just stating the obvious

Phase1:
- Deploy new VRF with multicast support
- Deploy RP inside the VRF at **central** sites
- Link GRT with VRF using a physical loop. <--yes physical loop , you read it right. Next Gen
- Deploy MSDP between GRT RP and VRF RP

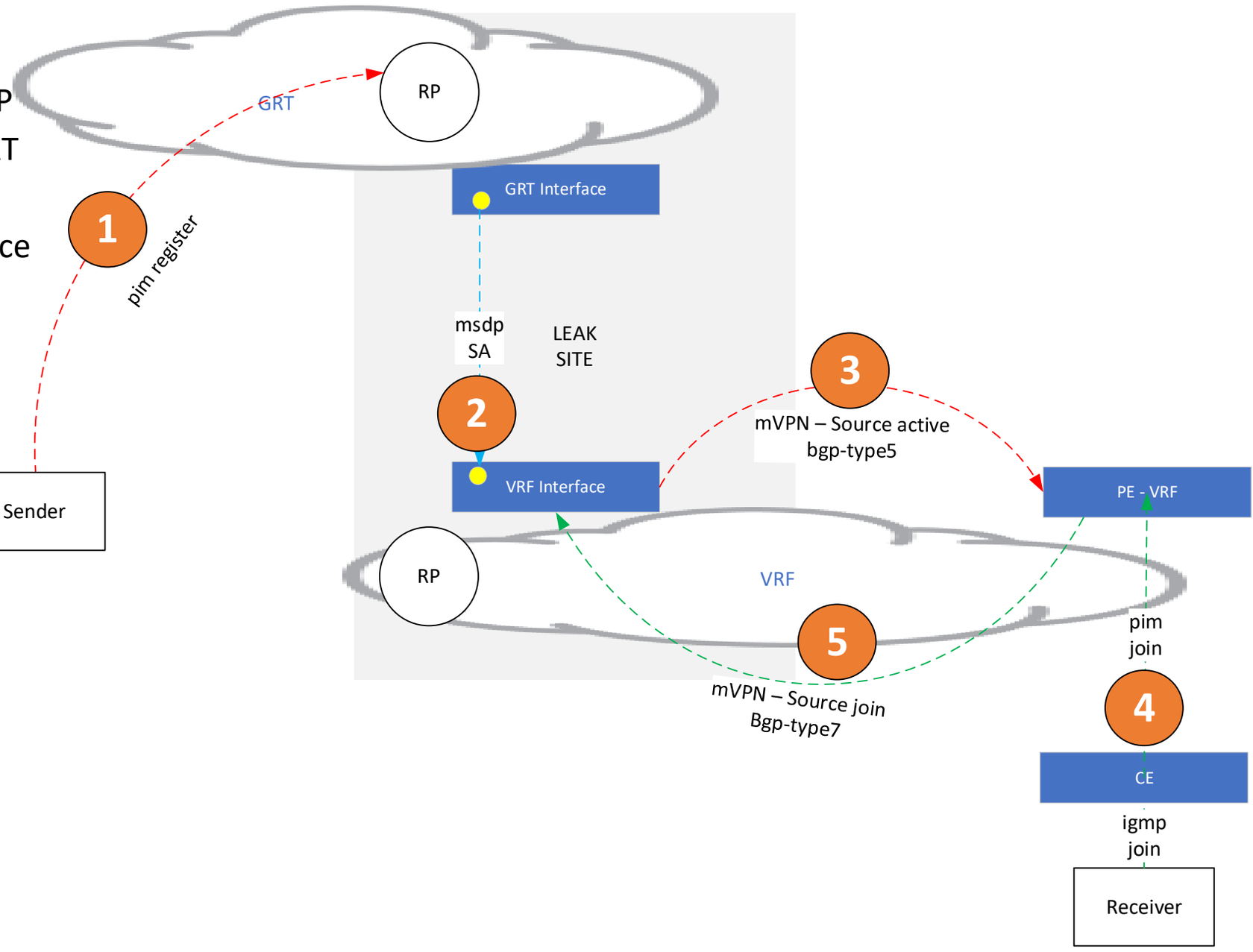
Phase2:
- On each access site
- Move all interfaces from GRT to VPRN (one router at the time)

Phase2- challenges :
If **Central** site is far away :
- Migrate at once 2-3 sites that needs to talk to each other directly and not via central site. Number can be higher depending on handover scenarios
- Accept latency increase for a short time

OW Multicast mVPN – Migration

Sender in GRT , receiver in mVPN

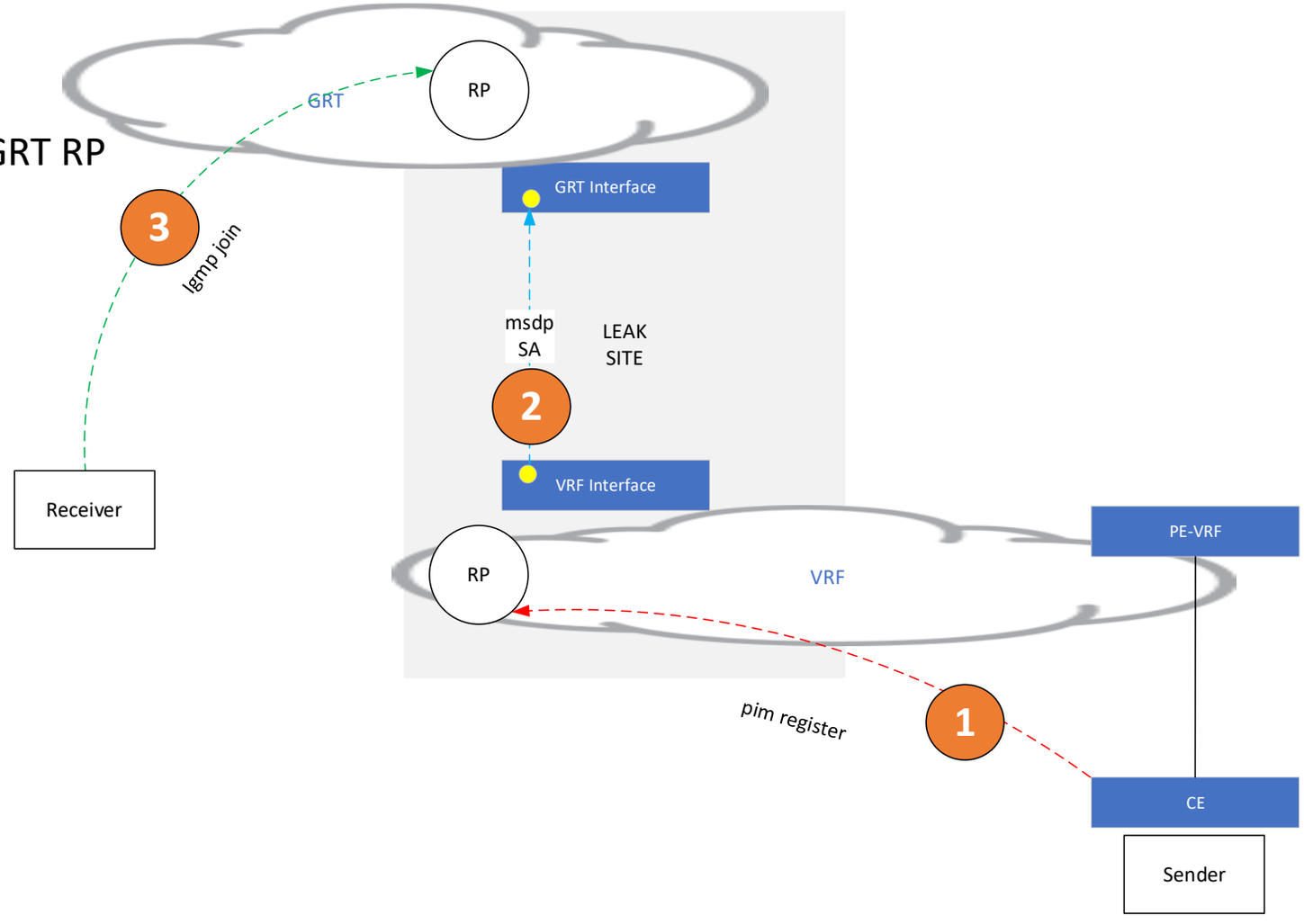
(1) Sender sends a unicast register to RP
(2) MSDP source active is send from GRT RP to VPRN RP
(3) VPRN RP will generate a type 5 source active route into mVPN
(4) Receiver in VPRN 403 will join the group.
(5) ERT in mVPN will generate a type 7 source join towards the PE which the next hop of the Sender.



OW Multicast mVPN - Migration

Sender in mVPN receiver in GRT

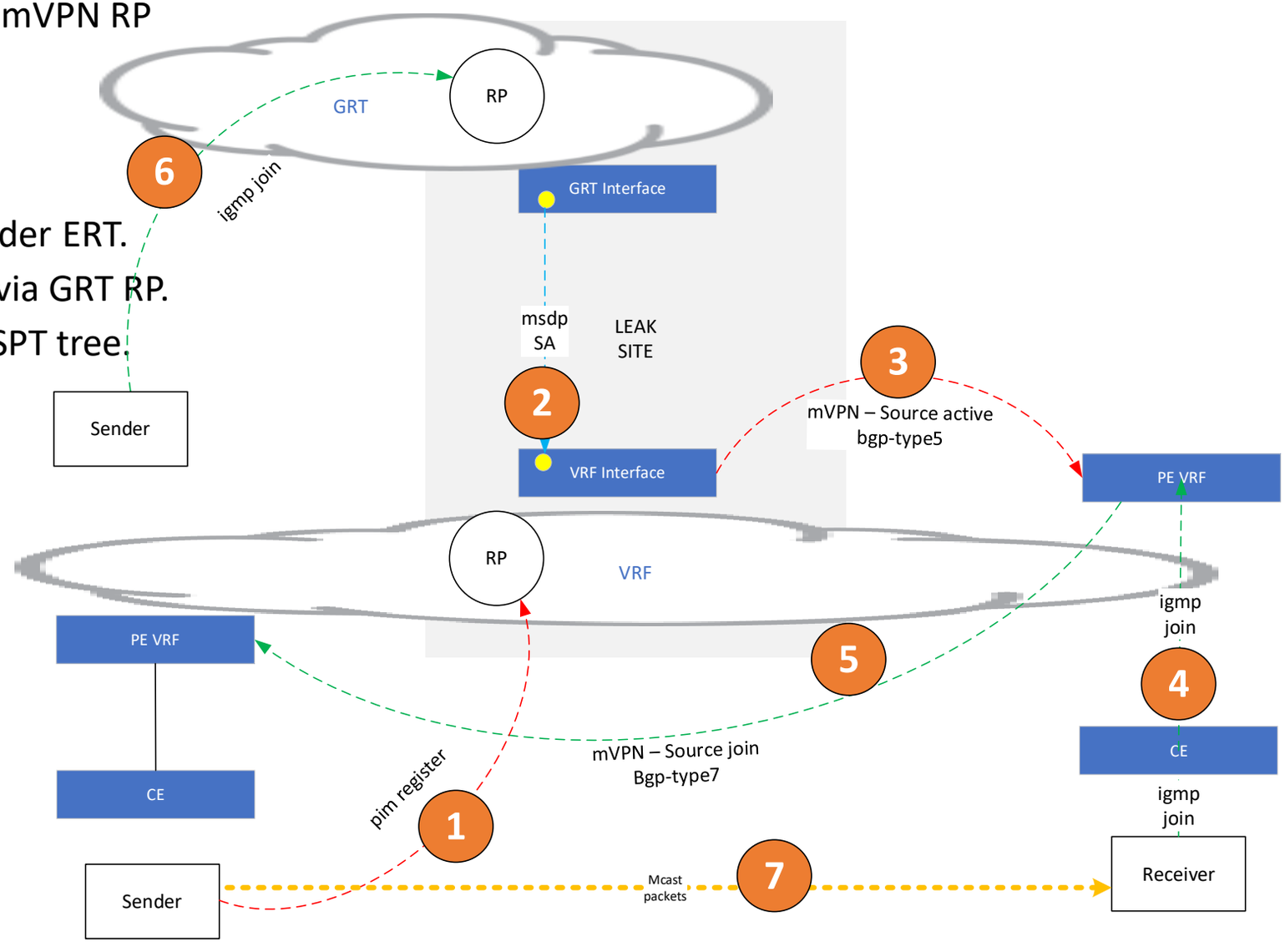
(1) Sender sends a unicast register to mVPN RP
(2) MSDP source active is send from mVPN RP to GRT RP
(3) Receiver in GRT will join the group



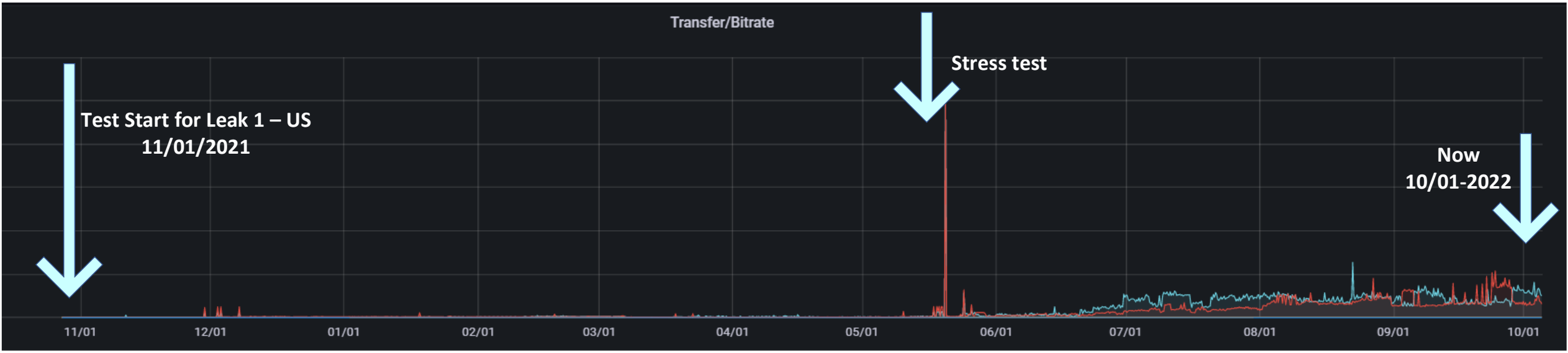
OW Multicast mVPN - Migration

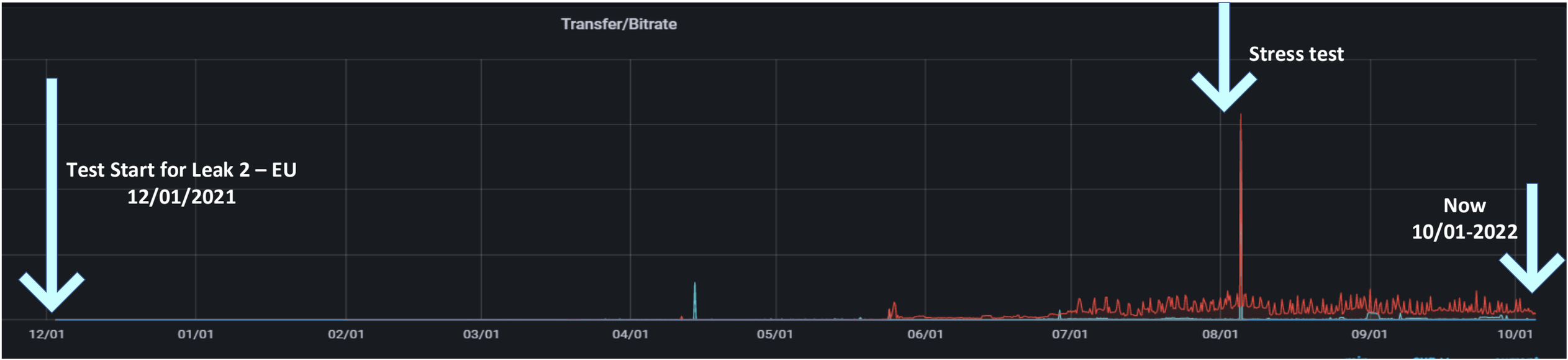
All sites migrated

(1) Sender in mVPN sends a unicast register to mVPN RP
(2) VPRN RP sends a SA to GRT RP
(3) VPRN RP sends a type 5 into mVPN
(4) Receiver joins group
(5) mVPN ERTA sends a type 7 towards the sender ERT.
(6) Optional a GRT receiver can join the group via GRT RP.
(7) Traffic flows from SENDER to RECEIVER via SPT tree.



OW Multicast mVPN – Migration Timeline





OW Multicast mVPN - Summary

OneWeb

Lesson learned

- BGP convergence is key, make sure you know the convergence time for vpnv4 and mvpn (next slide)
- Dedicated RR's help.
- LFA is your friend…not R-LFA or TI-LFA. (next slide)

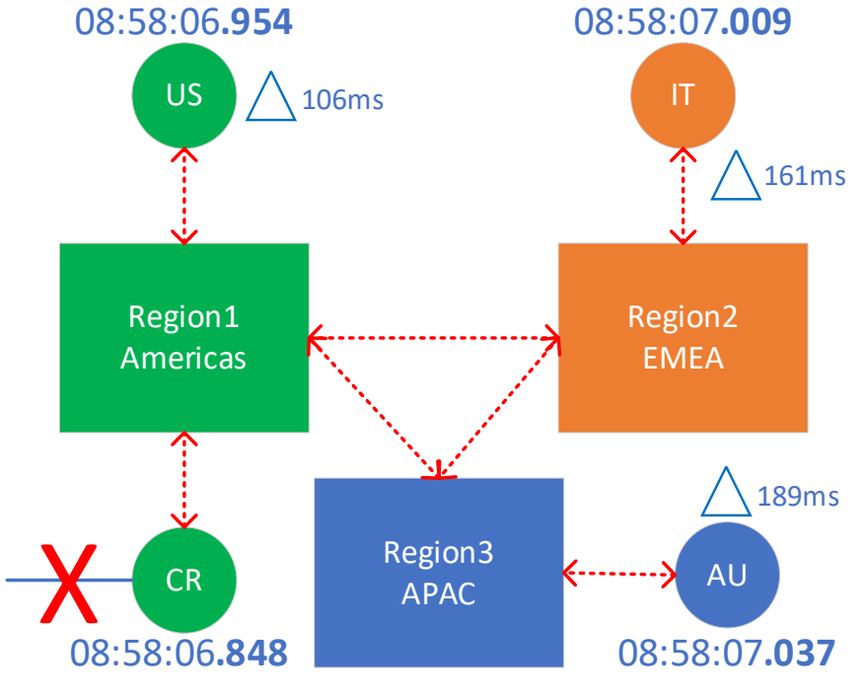
Prepare to hit ~~bugs~~ sw defects …(some below / some **have** a fix / some **not** documented in vendor Release notes)

- Fragmented PIM JP PDUs are discarded.
- MSDP fails to send Source Active if packet >1500B.
- PIM register-stop not process if received on non-PIM interface
- High CPU due to VRF vs MP-BGP best path selection.
- Underlying routing changes in a core network may cause MVPN tunnels to flap
- Platform scale , number of mVPN neigh -> increase in convergence

Prepare to revisit some design decisions

- BFD on multicast routes
- BGP backup path (VRF / VRF-Lite)
- LDP BFD
- IGMP timers
- Repeat - E2E convergence test couple of times

OW – BGP Convergence

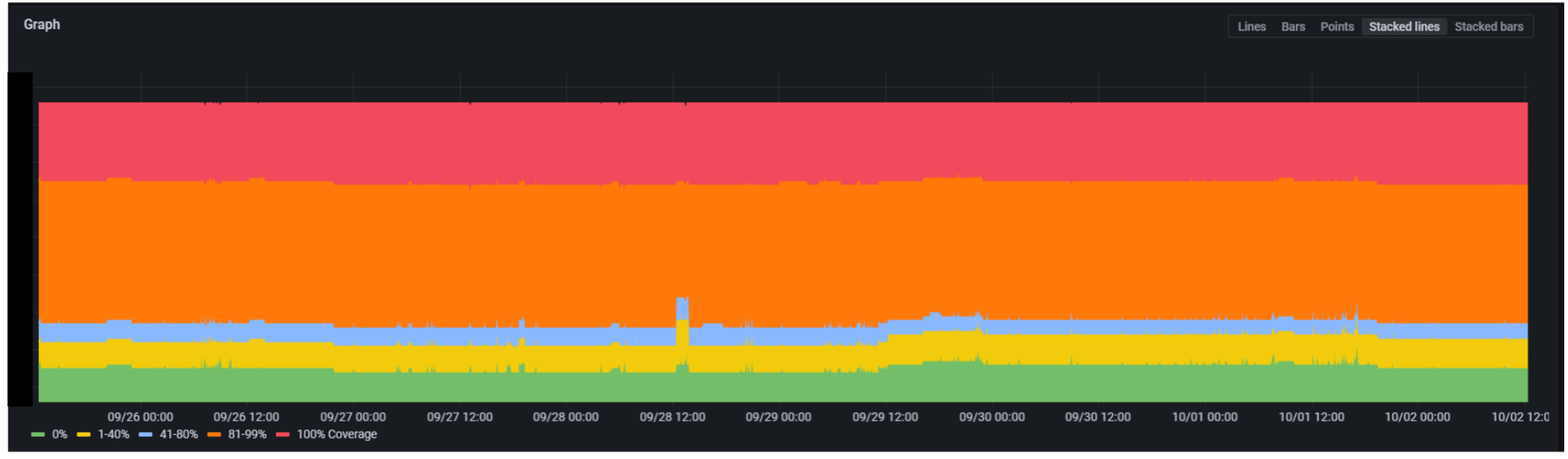


BGP Update

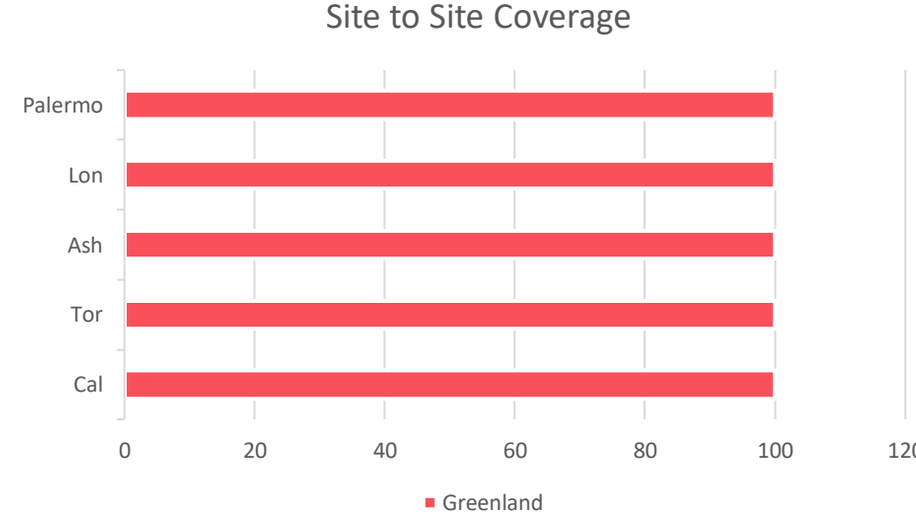
device	ip_prefix	log_type	peer_ip_name	timestamp ↑
cascr[redacted]	[redacted].29.192/31	withdraw	cascr[redacted]	2022-10-01 08:58:06.848485
miaus[redacted]	[redacted].29.192/31	withdraw	qasus[redacted]	2022-10-01 08:58:06.909069
miaus[redacted]	[redacted].29.192/31	withdraw	laxus2[redacted]	2022-10-01 08:58:06.940322
miaus[redacted]	[redacted].29.192/31	withdraw	qasus[redacted]	2022-10-01 08:58:06.954571
miaus[redacted]	[redacted].29.192/31	withdraw	laxus2[redacted]	2022-10-01 08:58:06.954571
cascr[redacted]	[redacted].29.192/31	withdraw	qasus[redacted]	2022-10-01 08:58:06.983383
cascr[redacted]	[redacted].29.192/31	withdraw	laxus2[redacted]	2022-10-01 08:58:06.985885
pmoit[redacted]	[redacted].29.192/31	withdraw	longb[redacted]	2022-10-01 08:58:07.001516
perau[redacted]	[redacted].29.192/31	withdraw	sydau[redacted]	2022-10-01 08:58:07.004743
pmoit[redacted]	[redacted].29.192/31	withdraw	amsnl[redacted]	2022-10-01 08:58:07.009682
perau[redacted]	[redacted].29.192/31	withdraw	sydau[redacted]	2022-10-01 08:58:07.019856
perau[redacted]	[redacted].29.192/31	withdraw	sinsg3[redacted]	2022-10-01 08:58:07.023535
perau[redacted]	[redacted].29.192/31	withdraw	sinsg3[redacted]	2022-10-01 08:58:07.037793

*OW – ISIS LFA Coverage

- ❑ Full - 26%
- ❑ 81-99 – 43%
- ❑ 41-80 – 5%
- ❑ 1-40 – 9%
- ❑ 0 – 17%



* We care about specific site to site coverage:
  - Greenland talks with London/Toronto/Calgary and Ashburn
  - Greenland to Palermo with 100% **doesn't** help.
- It's a balance between latency/capacity and LFA coverage.
- We do design for **low latency(1)** and **high coverage(2)** ; still room to improve.



OW Multicast mVPN - Summary

Looking forward

- Move away from mLDP
- BIER
- **SR-P2MP policies**
- Wait for our vendor to support R-LFA and TI-LFA for mLDP
  - https://www.rfc-editor.org/rfc/rfc7715
  - https://datatracker.ietf.org/meeting/105/materials/slides-105-pim-mofrr-based-on-tilfa-00
- Scaling optimisation
  - Redesign applications ?
  - Limit Source-Active to specific sites only.
  - S-PMSI threshold decrease (now at 300kbs)

Come join us shape the future of space

https://oneweb.net/work-us/latest-vacancies

Department -> Core Network

#OneWebOneWorld