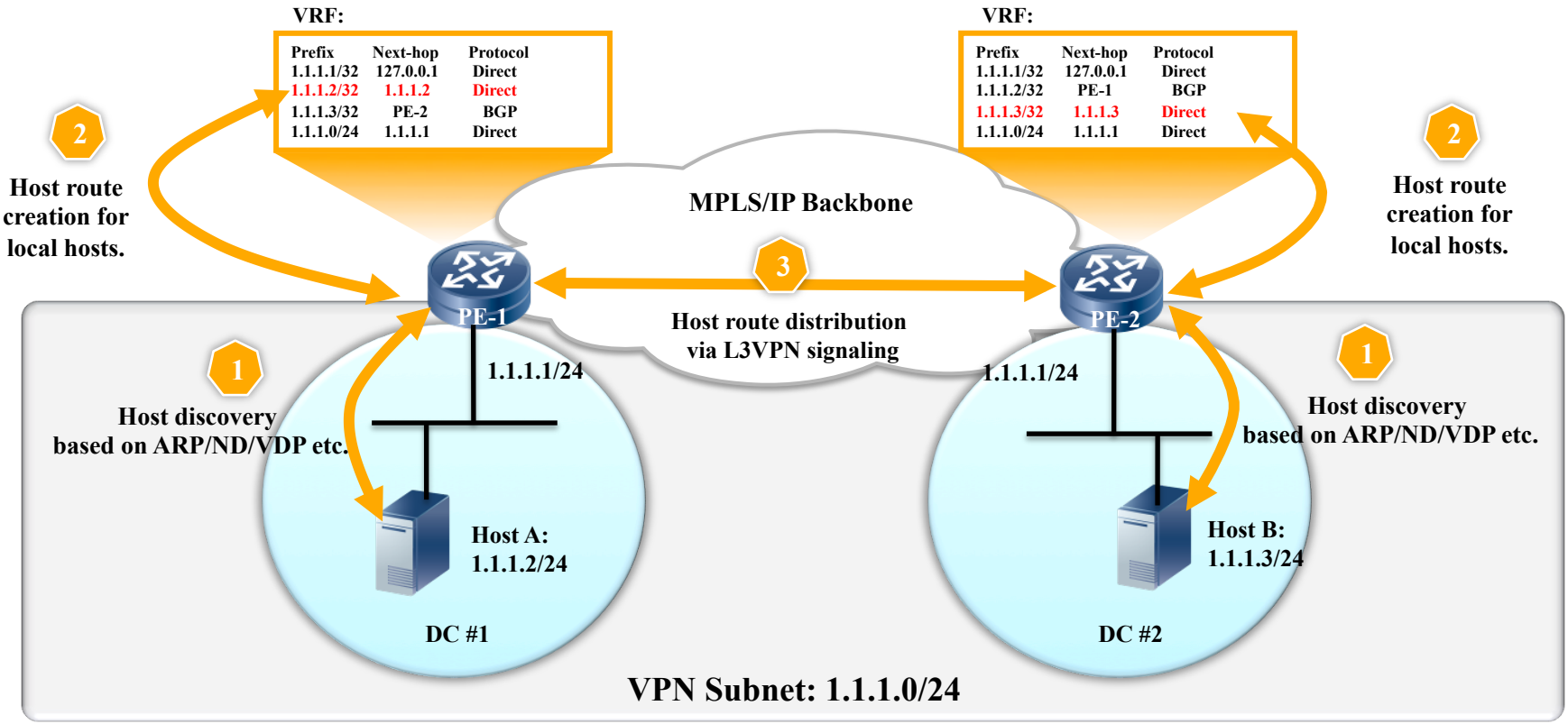
Virtual Subnet : A L3VPN-based Subnet Extension Solution

draft-xu-l3vpn-virtual-subnet-01

Xiaohu Xu (Huawei)
Susan Hares (Adara Networks)
Yongbing Fan (China Telecom)
Christian Jacquenet (Orange)
Truman Boyes (Bloomberg LP)
Brendan Fee (Extreme Networks)

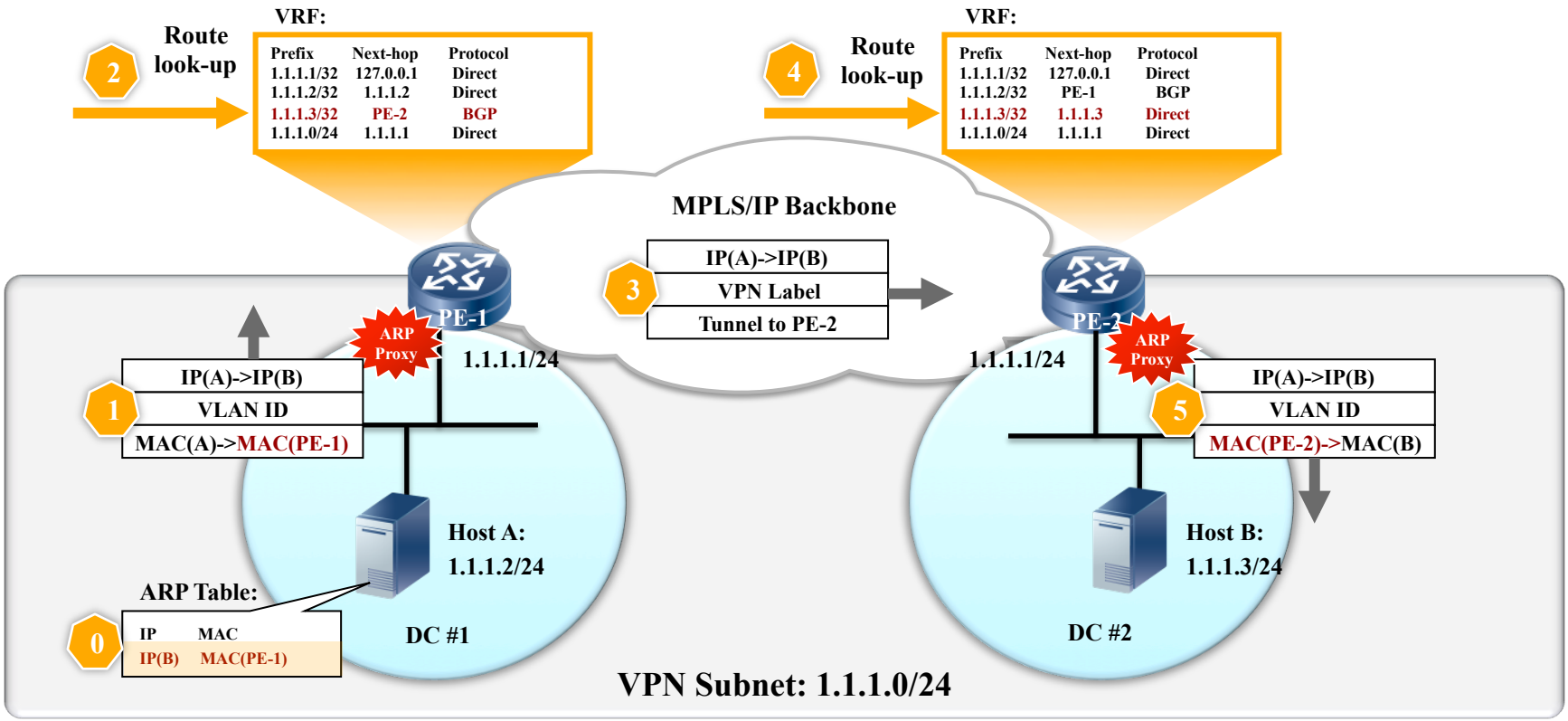
IETF88, Vancouver

Virtual Subnet at a Glance: Control Plane



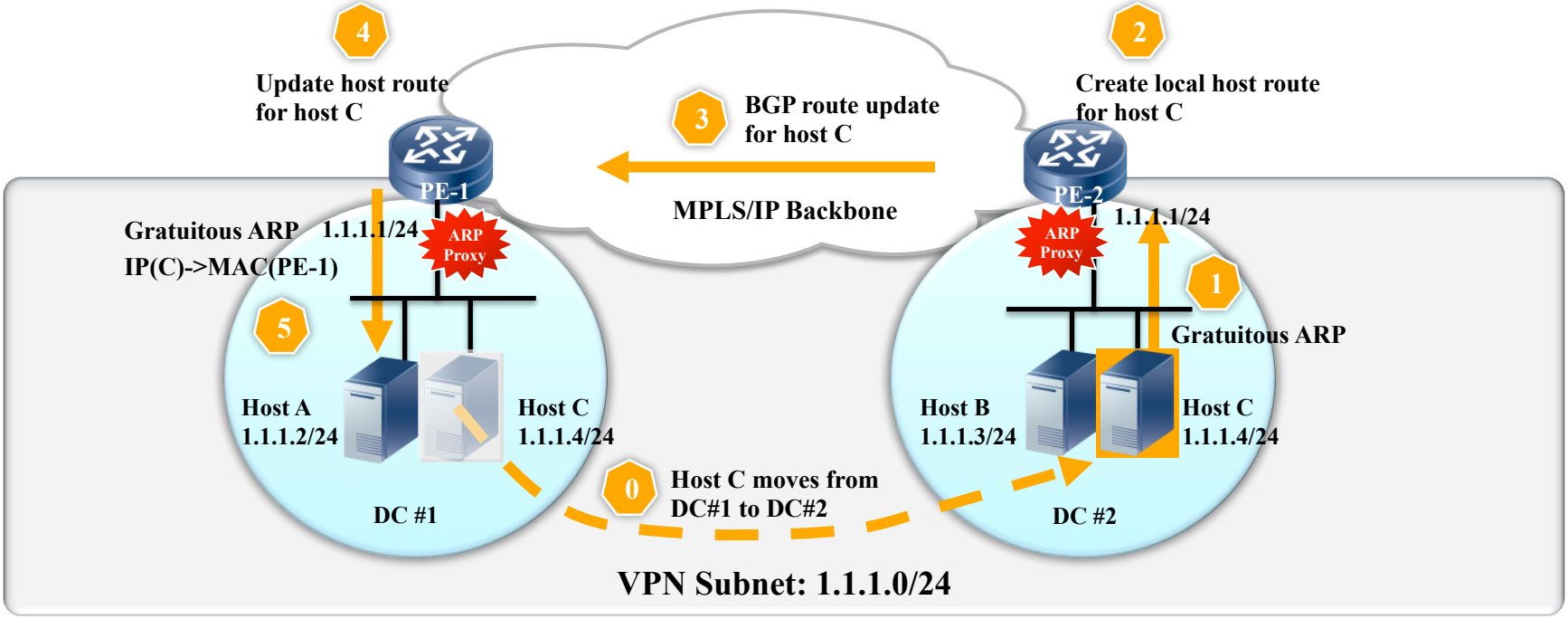
- Local CE hosts are discovered based on ARP/ND, VDP or even the interaction with the data center orchestration system.
- Host routes for local CE hosts are created automatically on PE routers and then propagated to remote PE routers via L3VPN signaling.

Virtual Subnet at a Glance: Data Plane



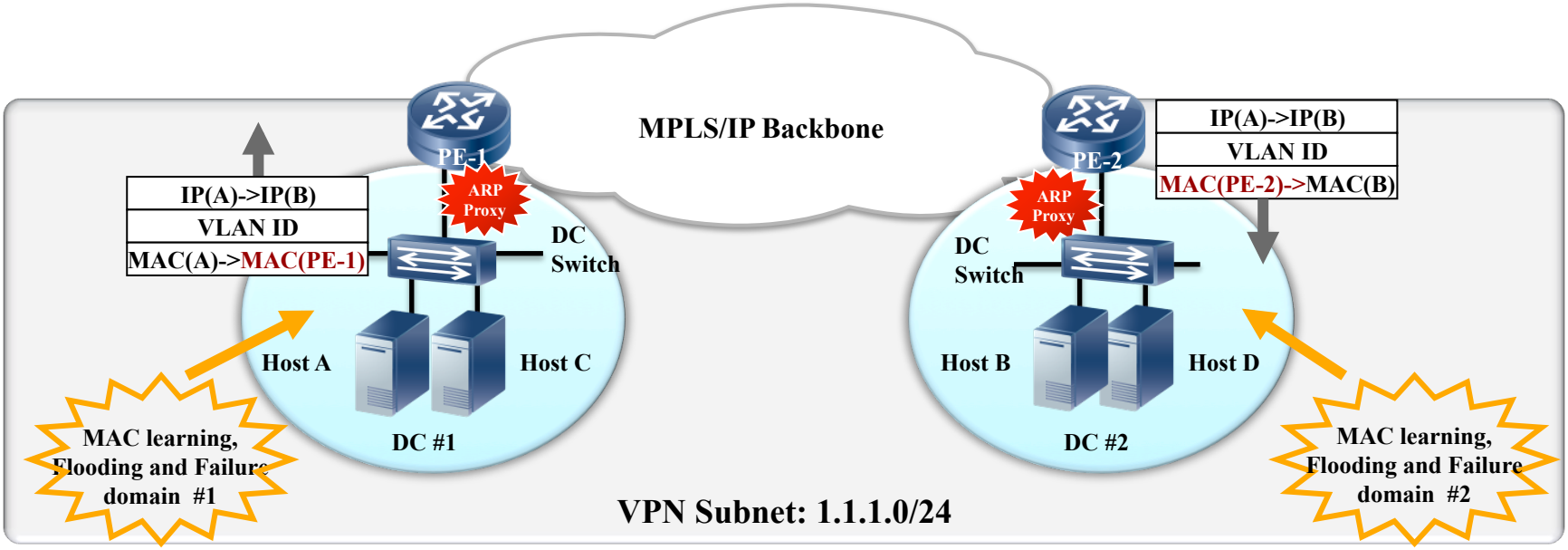
- PE routers acting as ARP proxies respond with their own MAC addresses to the ARP requests messages for remote CE hosts from local CE hosts.
- Intra-subnet traffic across data centers is forwarded according to the L3VPN forwarding process.

VM (CE Host) Mobility



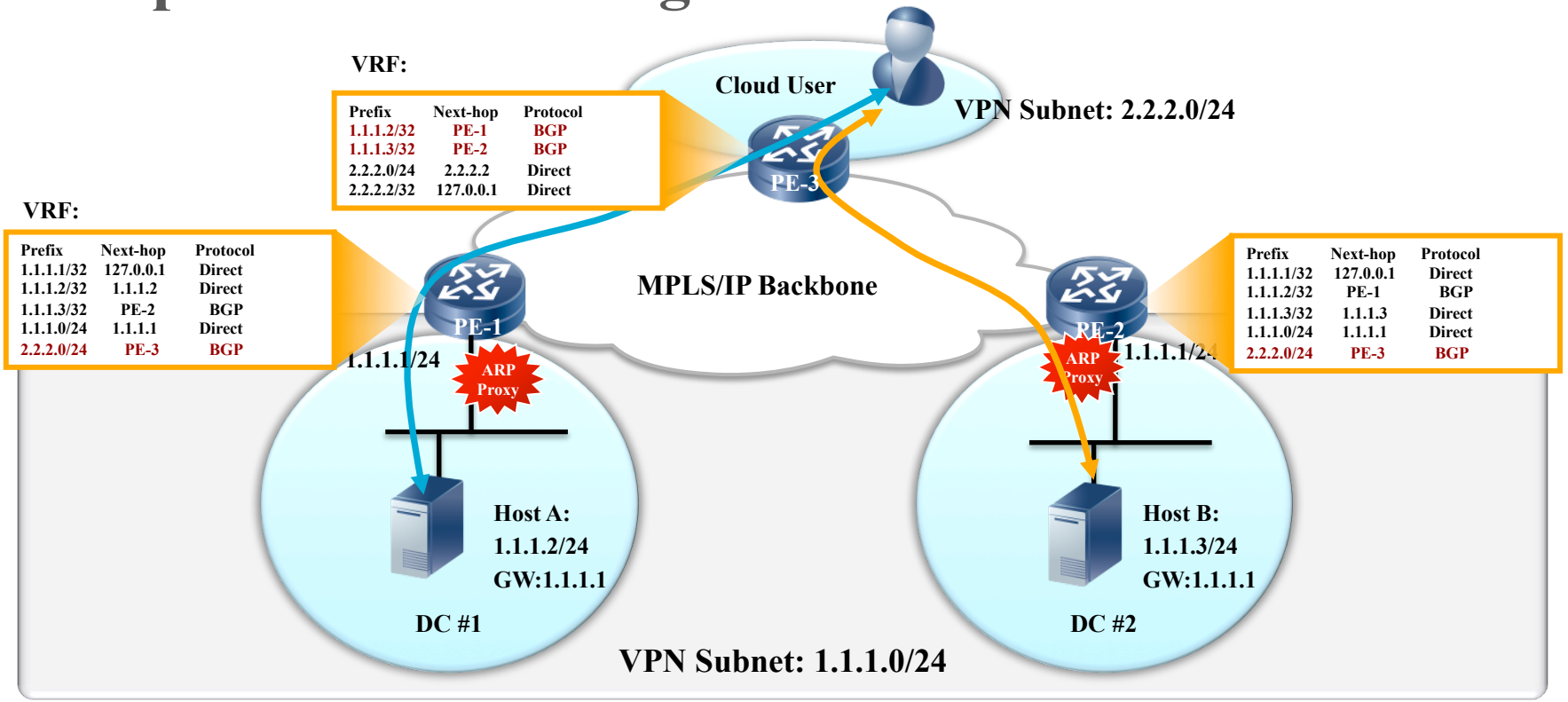
- The new PE router (i.e., PE-2) advertises a host route for the arriving VM upon receiving a notification of VM attachment (e.g., a gratuitous ARP).
- The old PE router (i.e., PE-1) withdraws the host route for the moved VM after noticing the leave of that VM. Meanwhile, it would broadcast a gratuitous ARP on behalf of that CE host with source MAC address being one of its own.

Confine MAC Learning, Flooding and Failure Domains



- MAC learning domain is confined within data centers. Therefore, switches within data centers only need to learn MAC addresses of local CE hosts.
- Flooding and failure domains are confined within data centers. As such, multicast/broadcast protocol messages (e.g., ARP/DHCP/IGMP/STP/ VRRP) from customer networks are terminated on PE routers. In addition, no flood of unknown unicast across data centers.

Optimal Forwarding for North-South Traffic



- Host routes for CE hosts within data centers are propagated to remote PE routers to which cloud users are connected. Therefore, north-to-south traffic would be delivered to the right data center without traffic tromboning.
- PE routers of each data center are default GWs. Therefore, south-to-north traffic would be forwarded to cloud users without traffic tromboning as well.

Considerations for Non-IP traffic

- Virtual Subnet is a Layer3 overlay in which IP traffic including both intra-subnet and inter-subnet would be forwarded at Layer3.
- To support non-IP traffic further, the unified L2/L3 overlay approach following the idea of “route all IP traffic, bridge non-IP traffic” could be considered (e.g., IP traffic is forwarded by using the Virtual Subnet while non-IP traffic is forwarded across Layer2 overlays (e.g., VPLS) .

Next Steps

- Implementations are already available.
- We co-authors believe this draft is ready for WG adoption.