Virtual Subnet :
A L3VPN-based Subnet Extension Solution for Cloud Data Center Interconnect

draft-xu-virtual-subnet-09

Xiaohu Xu (Huawei)

Susan Hares (Huawei)

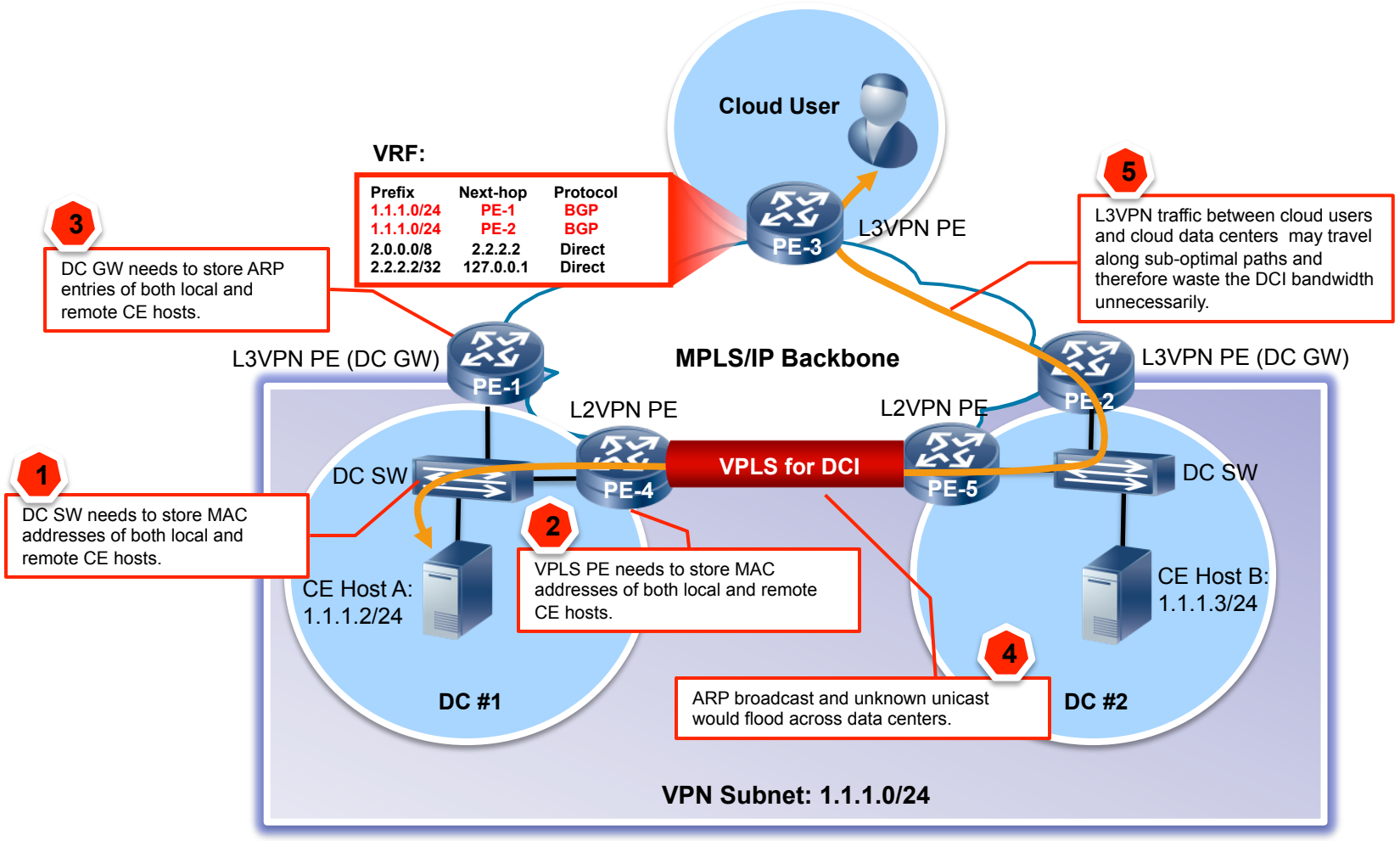
Yongbing Fan (China Telecom)
Christian Jacquenet (France Telecom)

IETF85, Atlanta

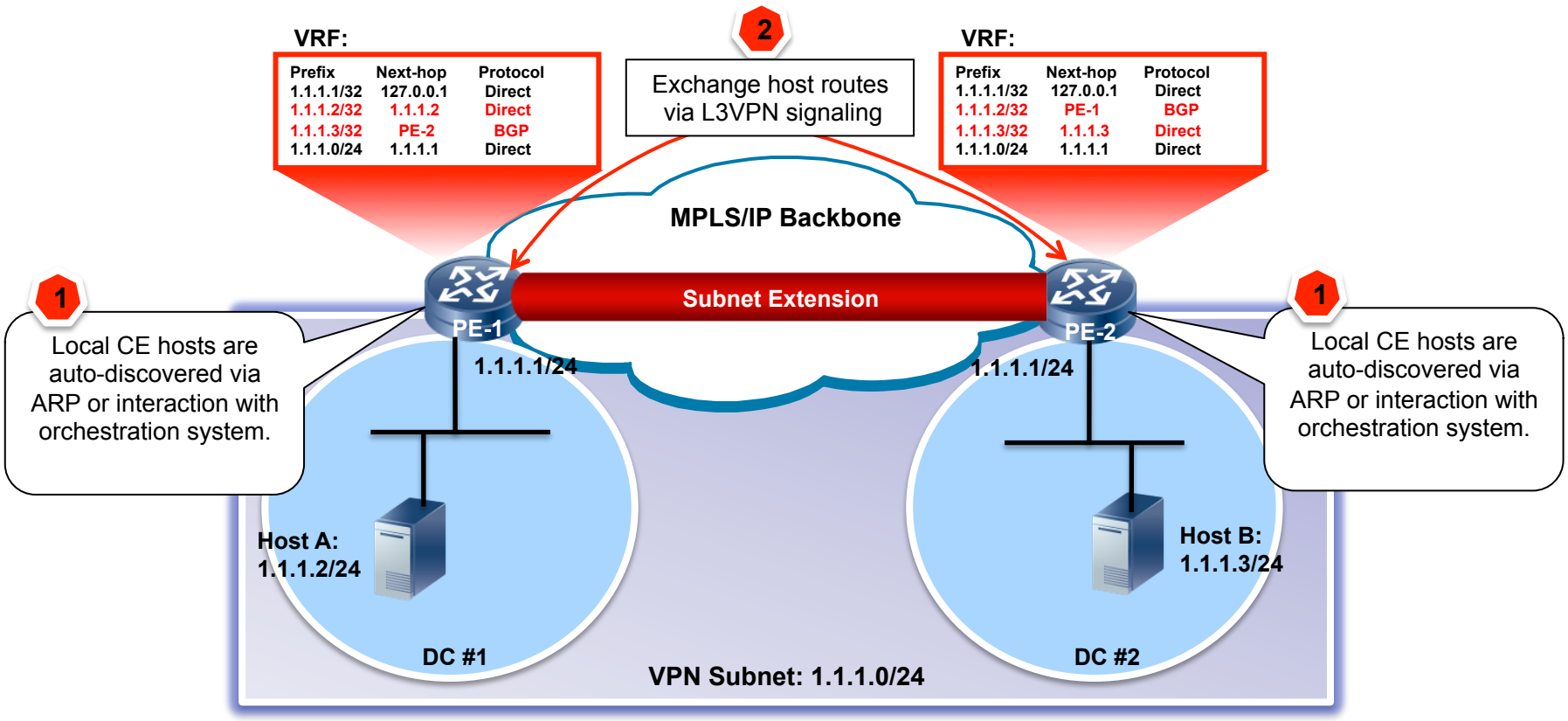
Cloud Data Center Interconnect Requirements

- **Subnet extension**
  - To allow VM migration across data centers without renumbering.
- **Forwarding table scalability**
  - Prevent the forwarding table size of DC switches and PEs from growing as the number of data centers to be connected increases.
- **ARP/ND cache table scalability**
  - Avoid ARP/ND cache table size of DC gateways from growing as the number of data centers to be connected increases.
- **Eliminate ARP/ND and unknown unicast impact**
  - Confine the flooding of ARP/ND broadcast/multicast and unknown unicast traffic within data centers.
- **Path optimization**
  - Avoid the traffic between cloud users and cloud data centers from travelling along suboptimal paths.

Why VPLS not a Good Choice for DCI

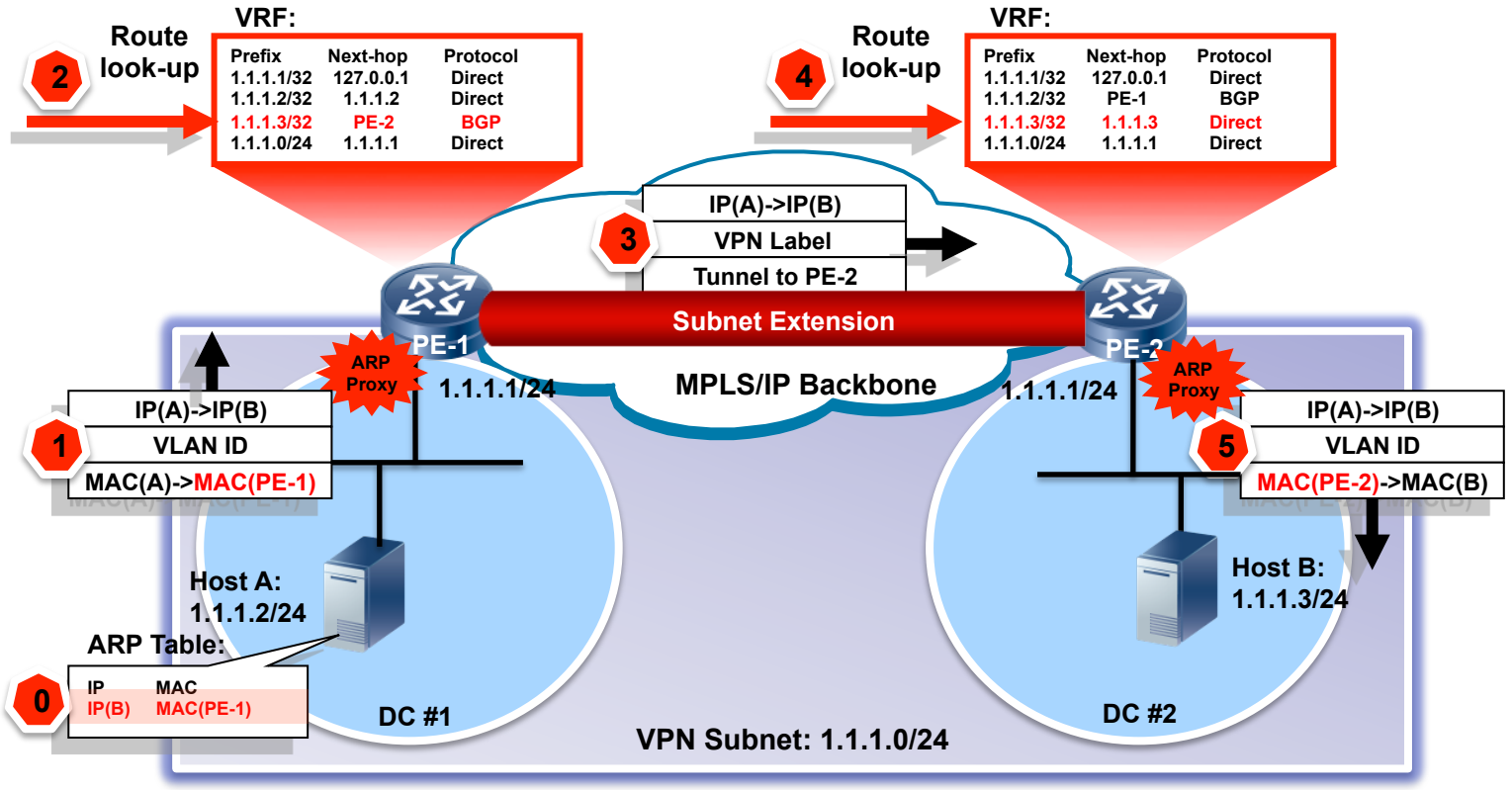


Virtual Subnet at a Glance: Control Plane



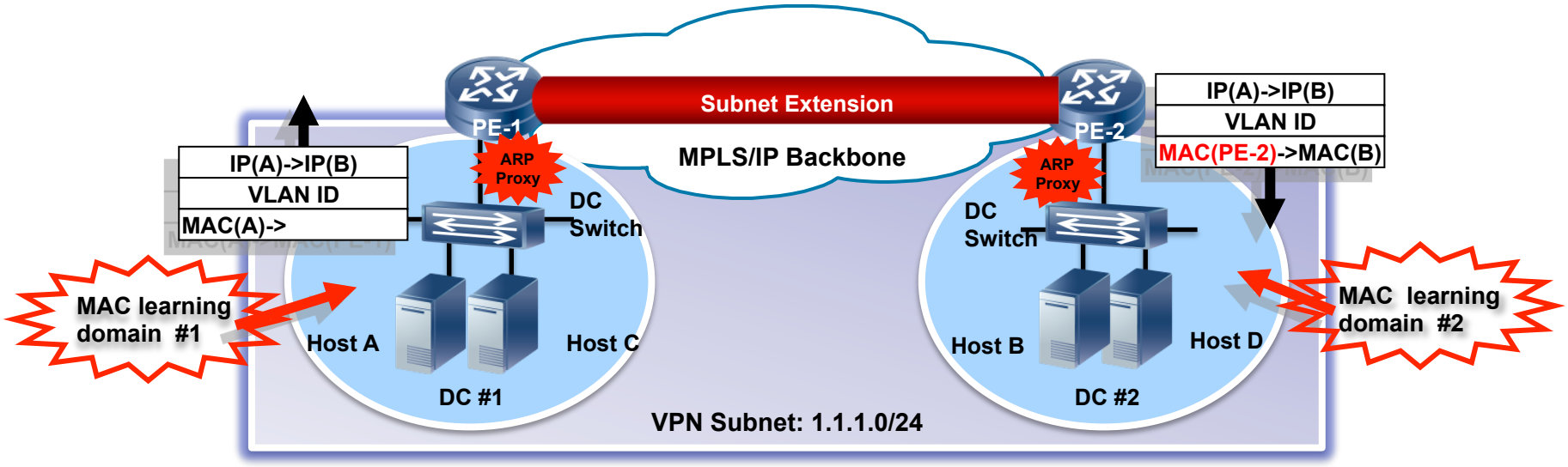
- Host routes for local CE hosts are created dynamically by PE routers and then propagated to remote PE routers via the existing L3VPN signaling.

Virtual Subnet at a Glance: Data Plane



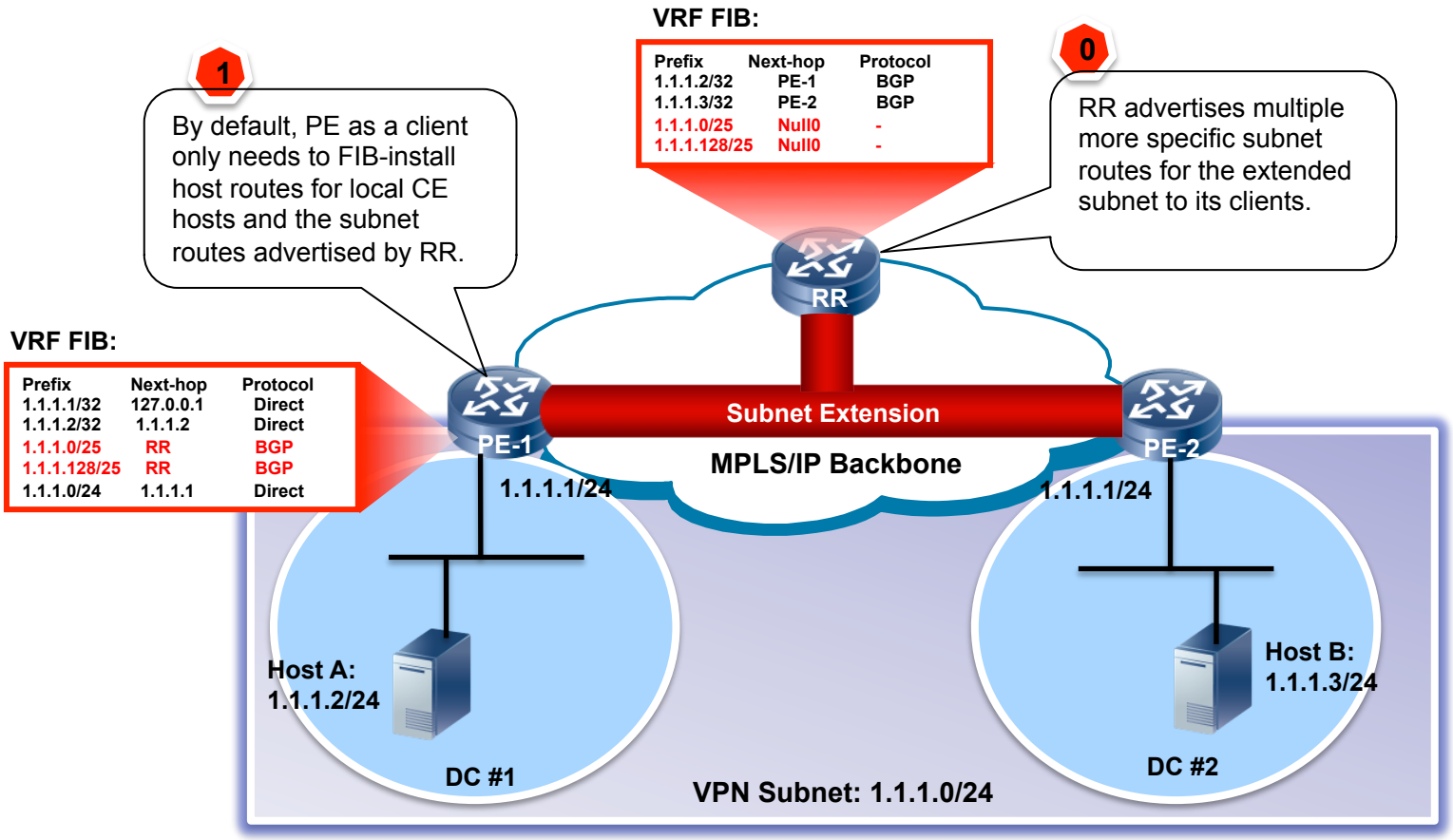
- Intra-subnet customer traffic across data centers is forwarded by PE routers in accordance with the current L3VPN forwarding procedures.

Scale MAC Tables of DC Switches



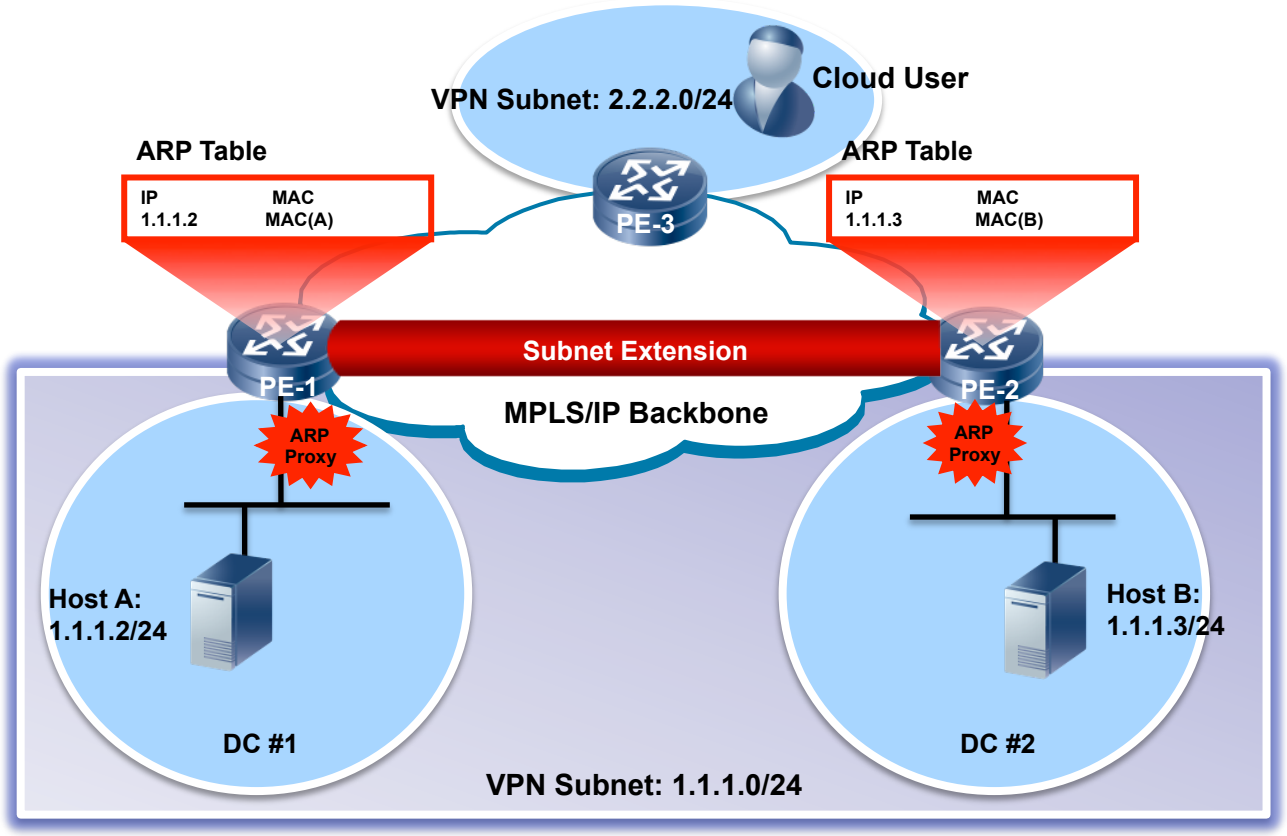
learn MAC addresses of local CE hosts, but not those of remote CE hosts.

Scale FIB Tables of PE Routers



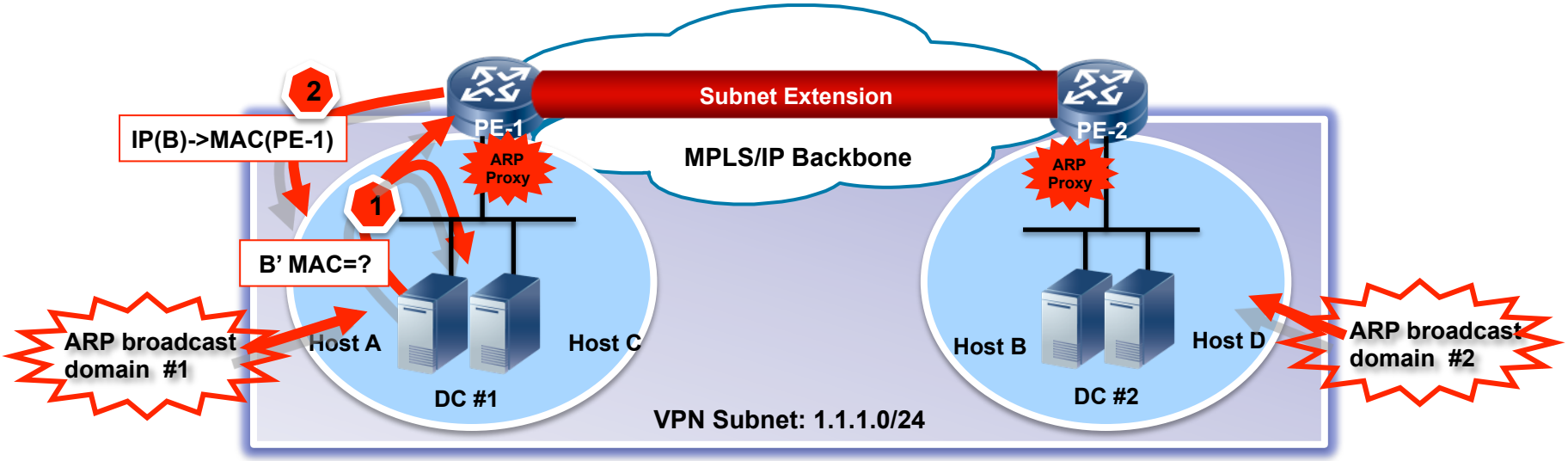
- PE routers only FIB-install host routes for local CE hosts and the subnet routes advertised by RR while remaining remote host routes still in RIB.
- ARP requests received from local CE hosts could trigger PE routers to FIB-install the corresponding host routes for the requested CE hosts from RIB.

Scale ARP/ND Cache Tables of DC Gateways



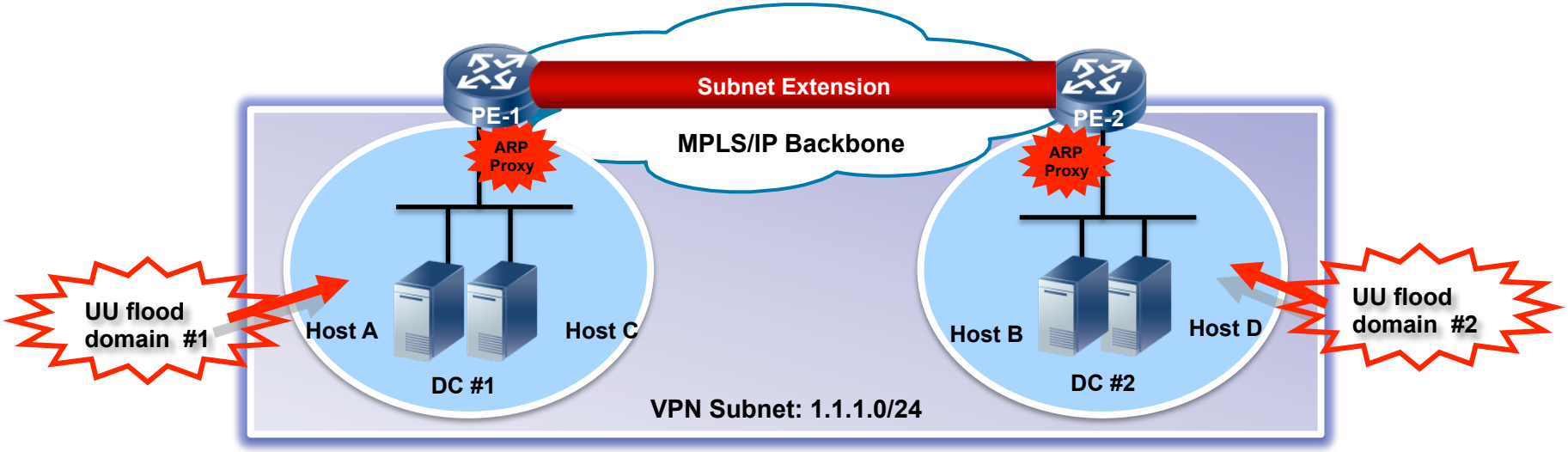
- PE routers acting as DC gateways as well only need to learn ARP entries of local CE hosts.

Avoid ARP/ND Broadcast/Multicast Flooding across Data Centers



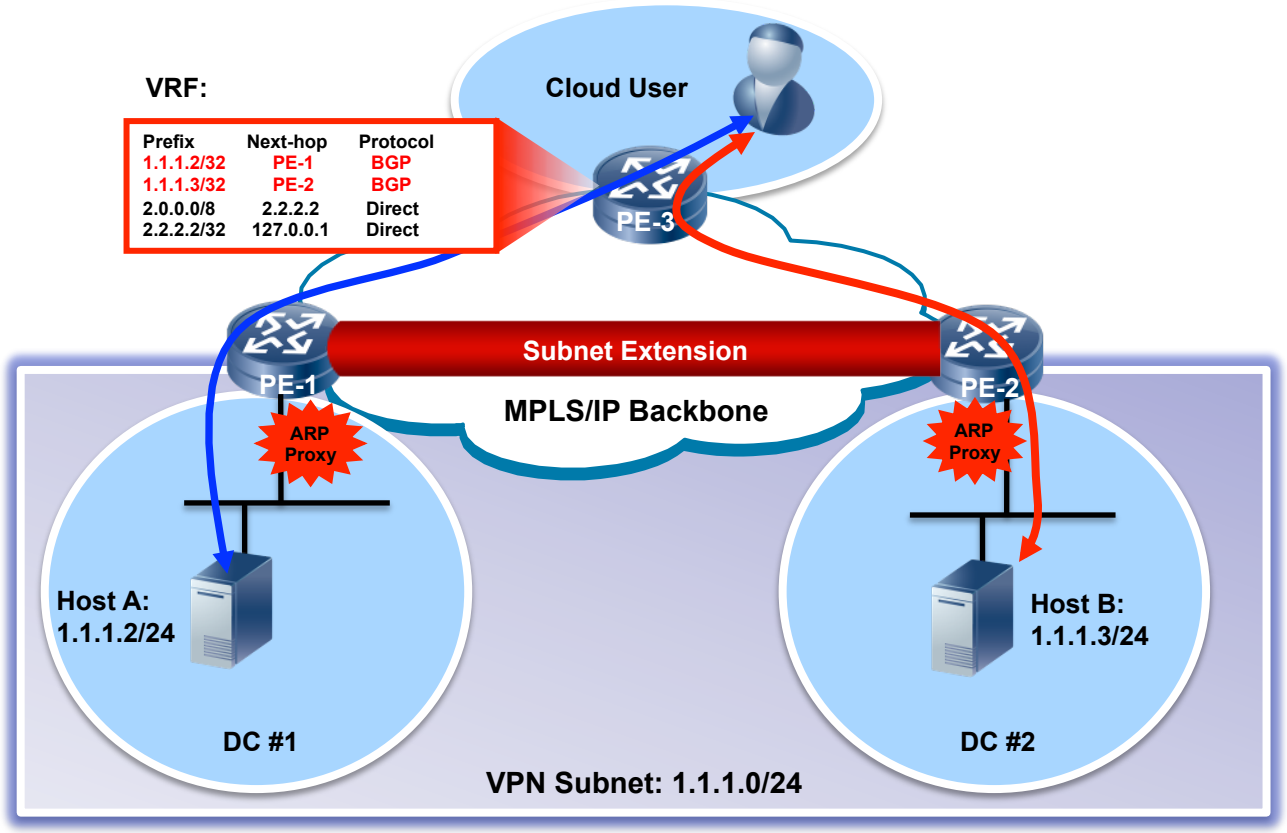
- Due to the use of ARP/ND proxy, ARP/ND broadcast/multicast traffic would not be flooded across data centers.

Avoid Unknown Unicast (UU) Flooding across Data Centers



- Since PE routers forward customer packets according to the associated VRF, there is no need for unknown unicast flood across data centers anymore.

Path Optimization for L3VPN Traffic between Cloud Users and Cloud DC



- Host routes for local CE hosts within data centers are propagated to remote PE routers to which cloud users are attached. Therefore, the inbound traffic would be forwarded according to host routes, rather than subnet route.

Next Step

- **WG adoption as an informational draft?**