Scaling the Address Resolution Protocol for Large Data Centers (SARP)
draft-nachum-sarp

Youval Nachum   **Marvell**

Linda Dunbar   **Huawei**

Ilan Yerushalmi   **Marvell**

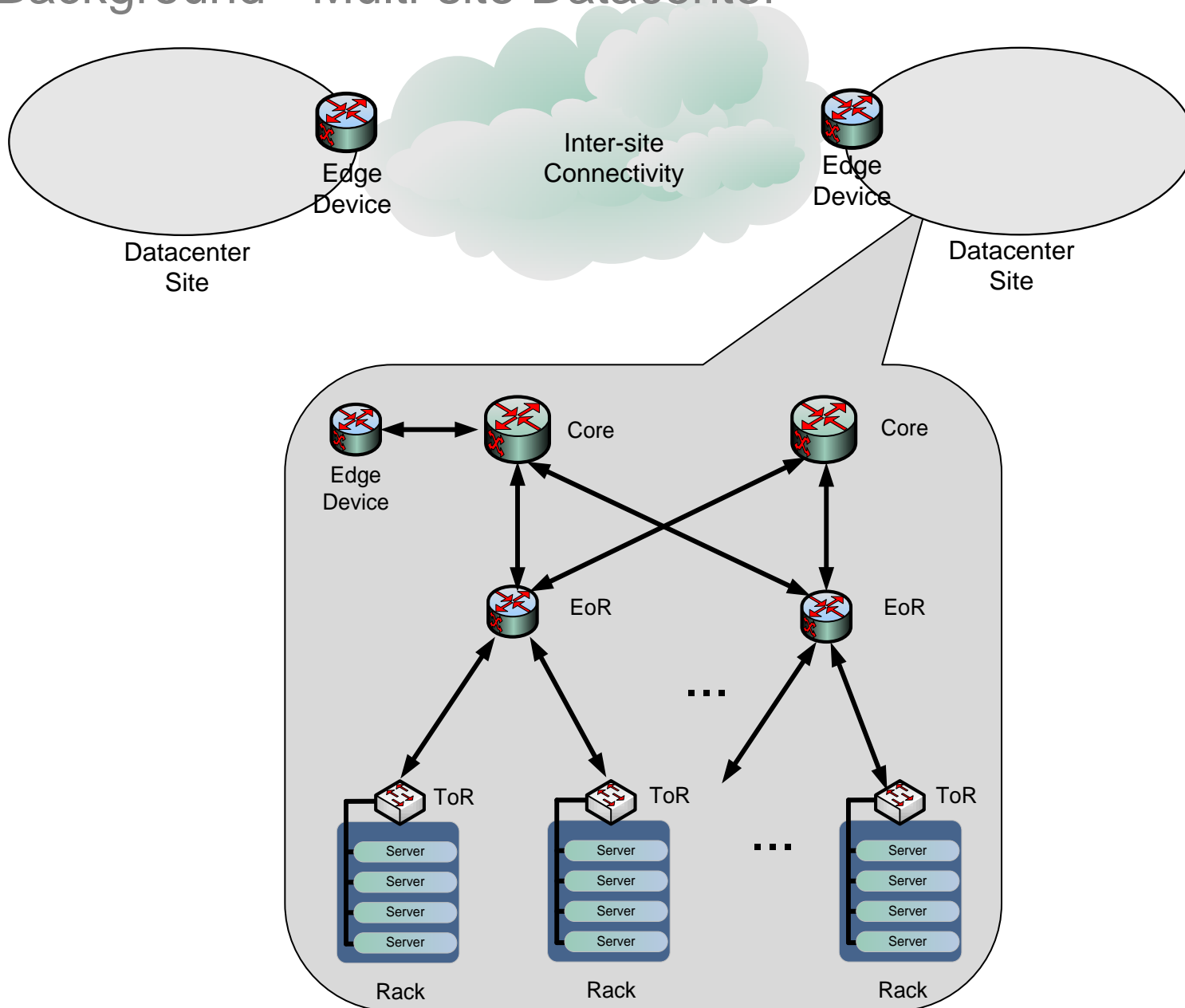**Tal Mizrahi**   **Marvell**

IETF Meeting 85, November 2012

History of this Draft

▶ **March 2012 – draft 00.**

▶ **Discussion in ARMD mailing list.**

▶ **July 2012 – IETF 84 – presented in INTAREA WG.**
- Main feedback: need to equally address IPv4 and IPv6.

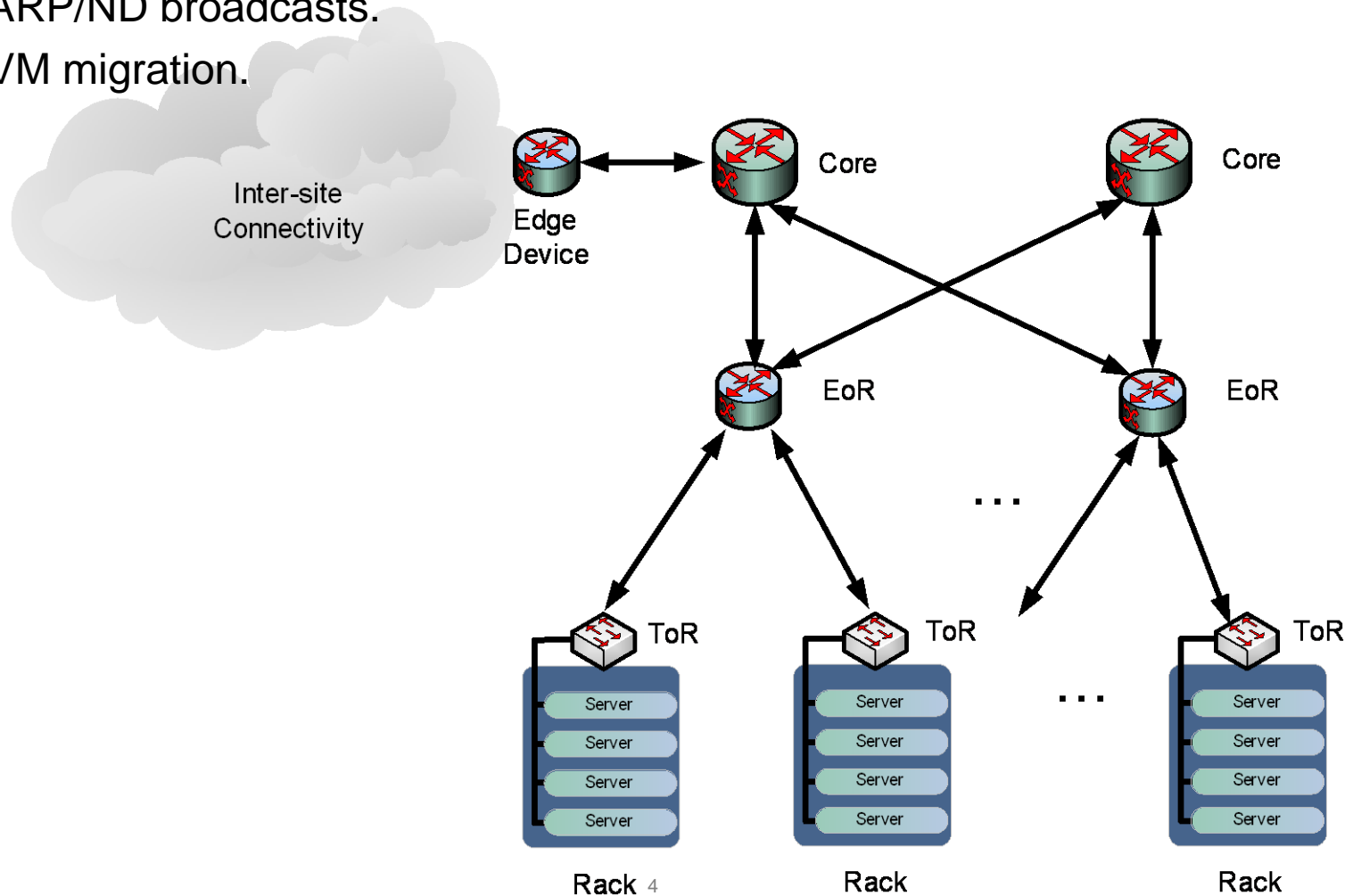▶ **October 2012 – draft 03.**
- More details about SARP with IPv6.
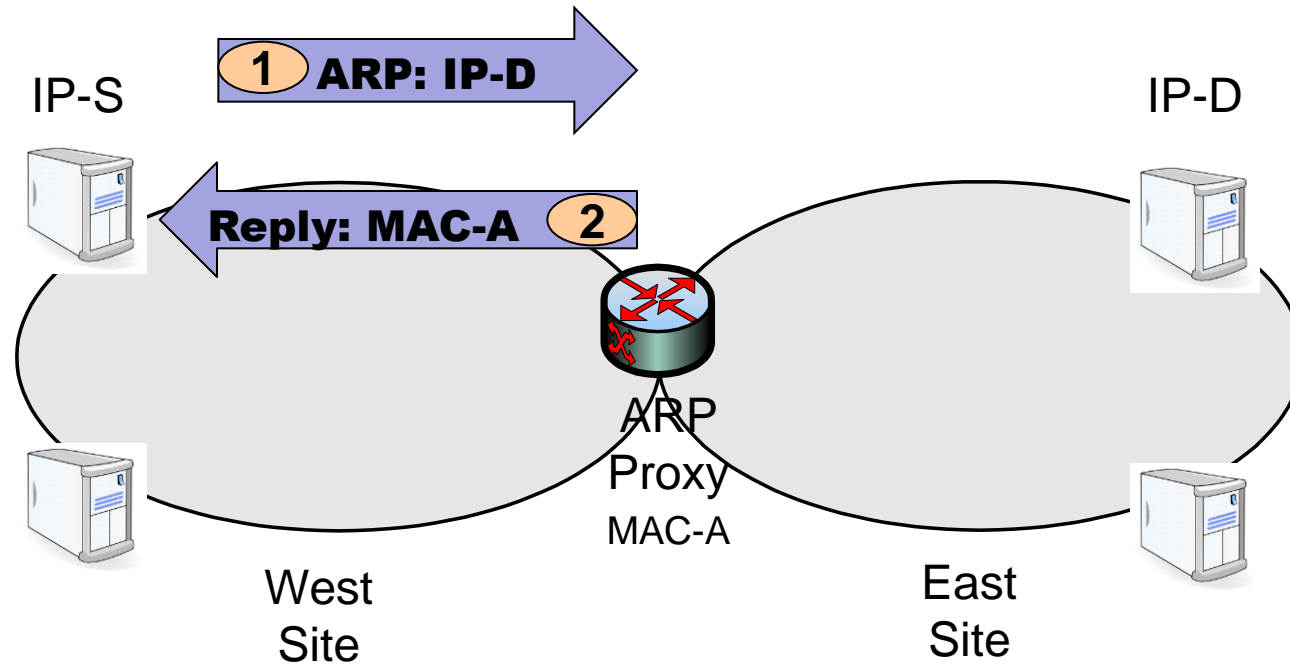
Background - Multi-site Datacenter

Background

▶ **Challenges in datacenter network scaling:**
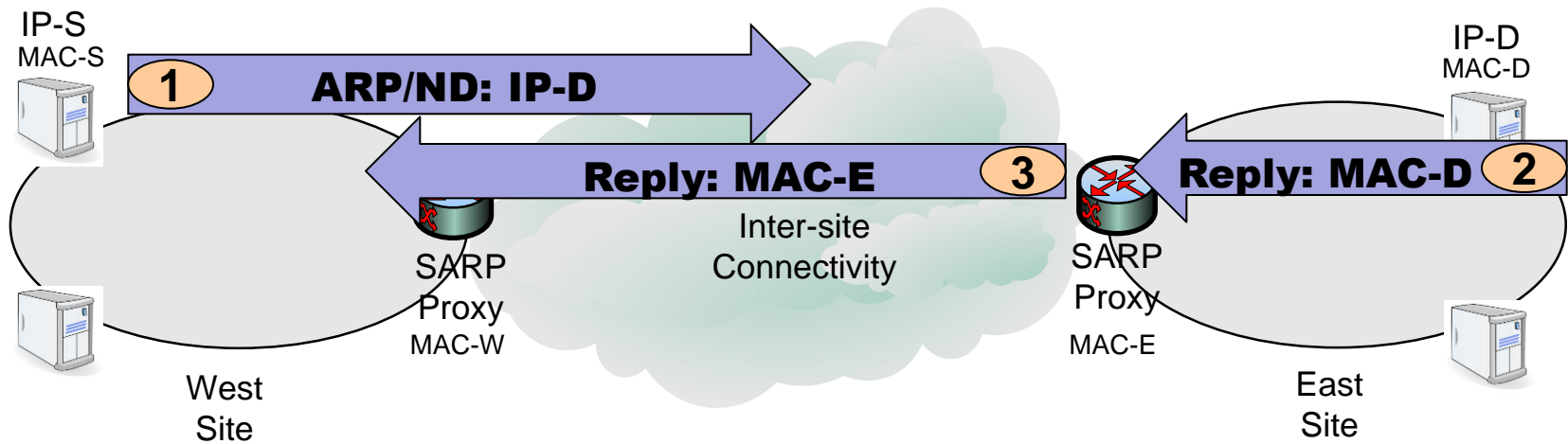- Large MAC address tables.
- ARP/ND broadcasts.
- VM migration.

Background – Proxy ARP

- **Proxy ARP (RFC 1027, RFC 1009, RFC 925).**

- **Proxy ARP responds based on IP subnet.**
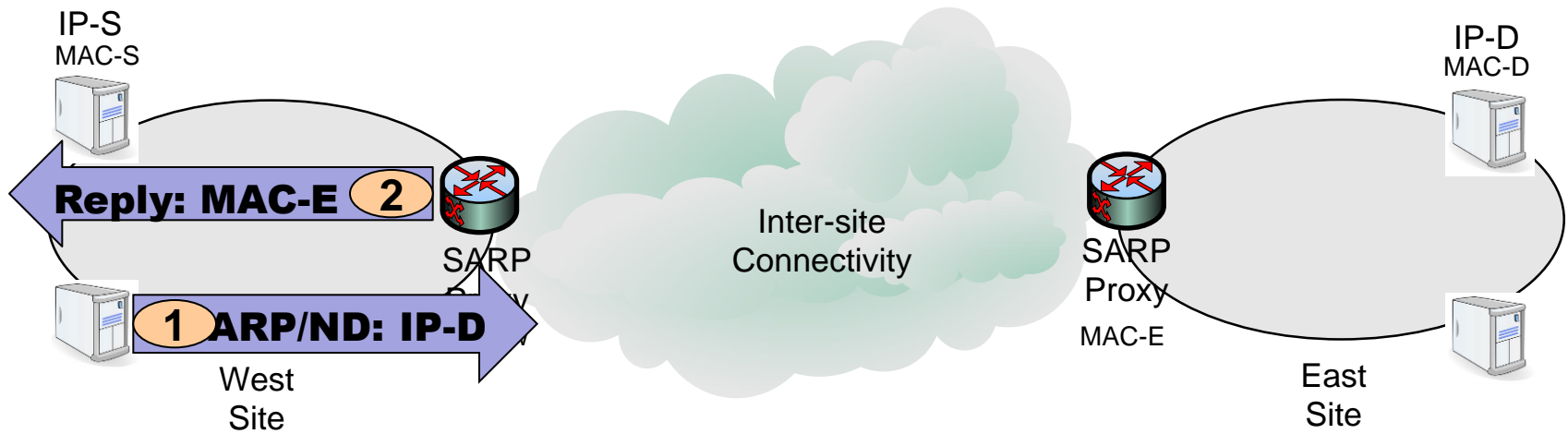 - Assumption: IP subnet implies location.

SARP
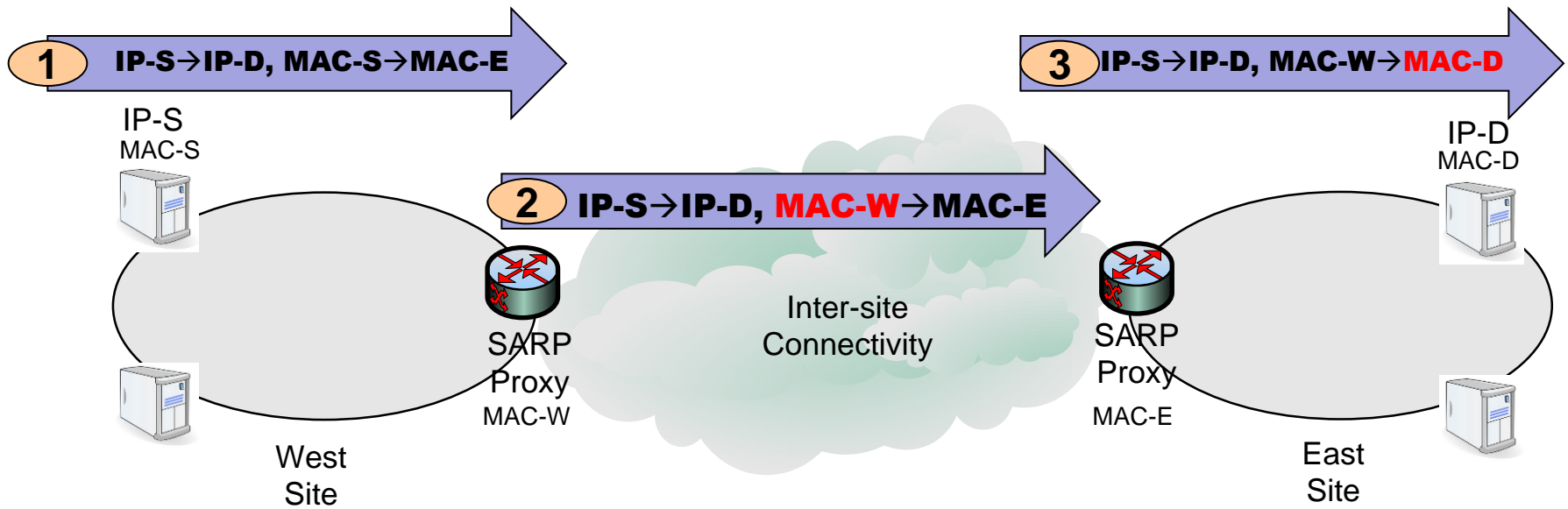
▶ **Edge devices: proxy SARP.**

▶ **IP subnet does not imply location.**

▶ **MAC-W / MAC-E imply location.**
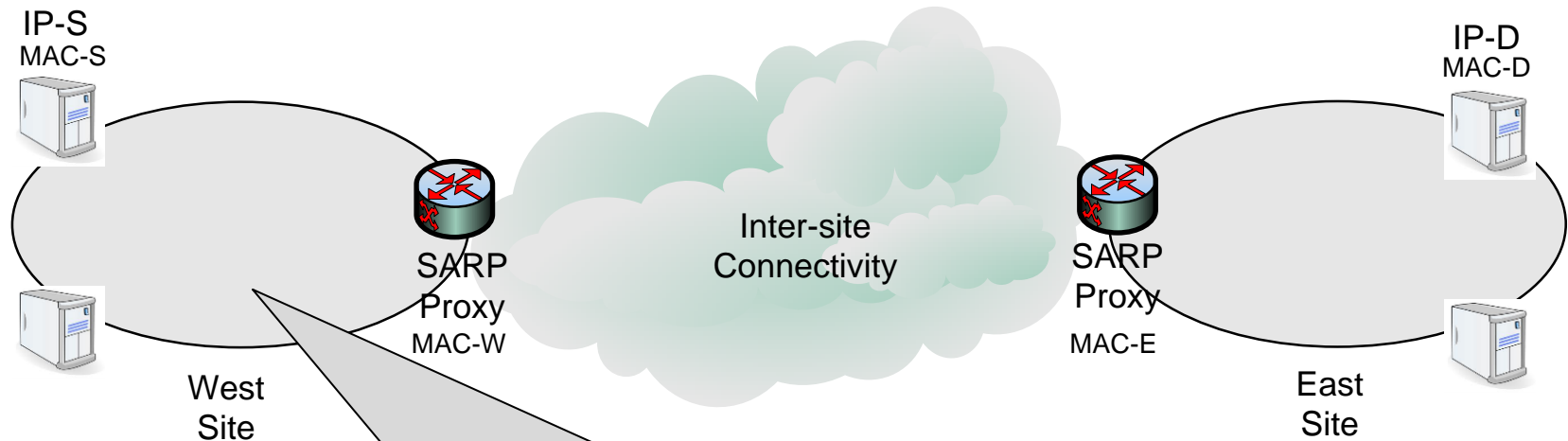
SARP Cache

SARP – Data Plane

1 IP-S→IP-D, MAC-S→MAC-E

3 IP-S→IP-D, MAC-W→MAC-D

2 IP-S→IP-D, MAC-W→MAC-E

IP-S
MAC-S

IP-D
MAC-D

SARP
Proxy
MAC-W

Inter-site
Connectivity

SARP
Proxy

MAC-E

West
Site

East
Site

SARP – MAC Address Tables

IP-S
MAC-S

IP-D
MAC-D

SARP
Proxy
MAC-W

Inter-site
Connectivity

SARP
Proxy
MAC-E

West
Site

East
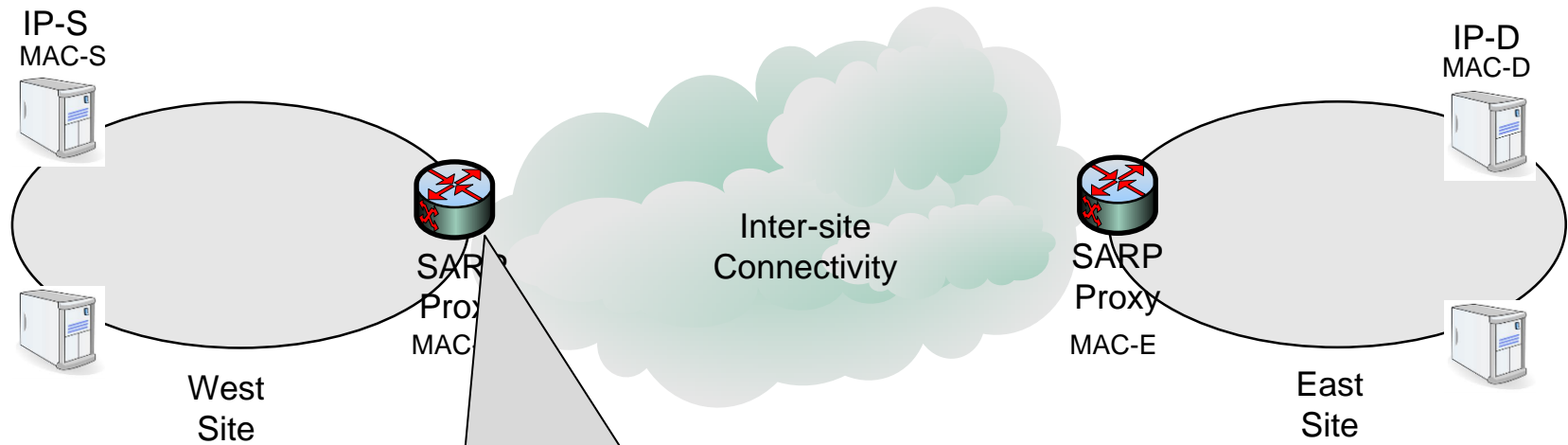Site

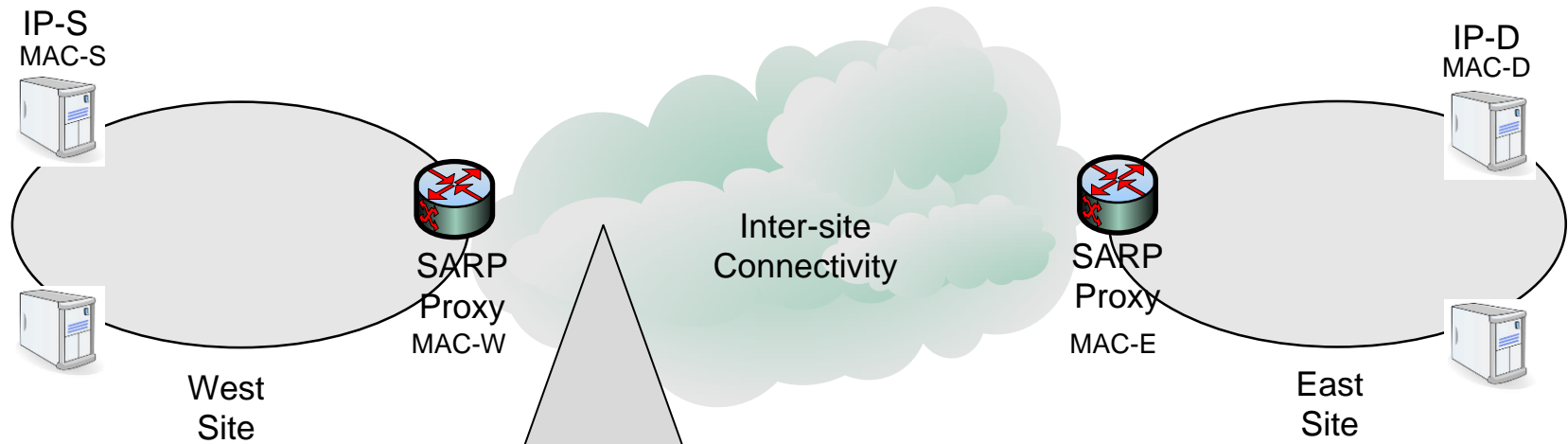**MAC address table of bridges in the west site:**
- **Local site addresses, e.g., MAC-S.**
- **Edge devices, e.g., MAC-E.**
- **No need for addresses of remote sites.**
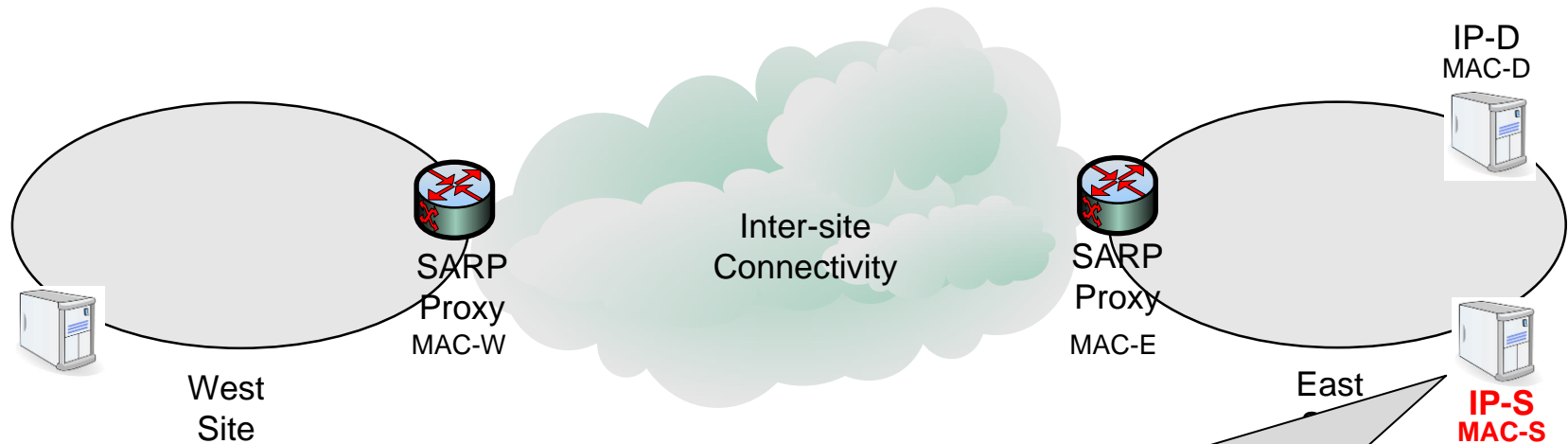
SARP – ARP Broadcast Domains

IP-S
MAC-S

IP-D
MAC-D

Inter-site
Connectivity

SARP
Proxy
MAC-

SARP
Proxy
MAC-E

West
Site

East
Site

Local SARP cache limits broadcast domain for known IP addresses.

SARP over Overlay Network

IP-S
MAC-S

IP-D
MAC-D

SARP
Proxy
MAC-W

Inter-site
Connectivity

SARP
Proxy
MAC-E

West
Site

East
Site

SARP is agnostic to the transport technology, e.g. L2VPN.

SARP with VM Migration

IP-D
MAC-D

SARP
Proxy
MAC-W

Inter-site
Connectivity

SARP
Proxy
MAC-E

West
Site

East

IP-S
MAC-S

- **IPv4: Gratuitous ARP is used to notify network about migration.**
- **IPv6: unsolicited neighbor advertisement is used.**

- **No need for additional control protocols.**
- **Transparent to inter-site network and protocols.**

Next Steps

▶ **Receive feedbacks from WG.**

▶ **WG adoption.**

Thanks