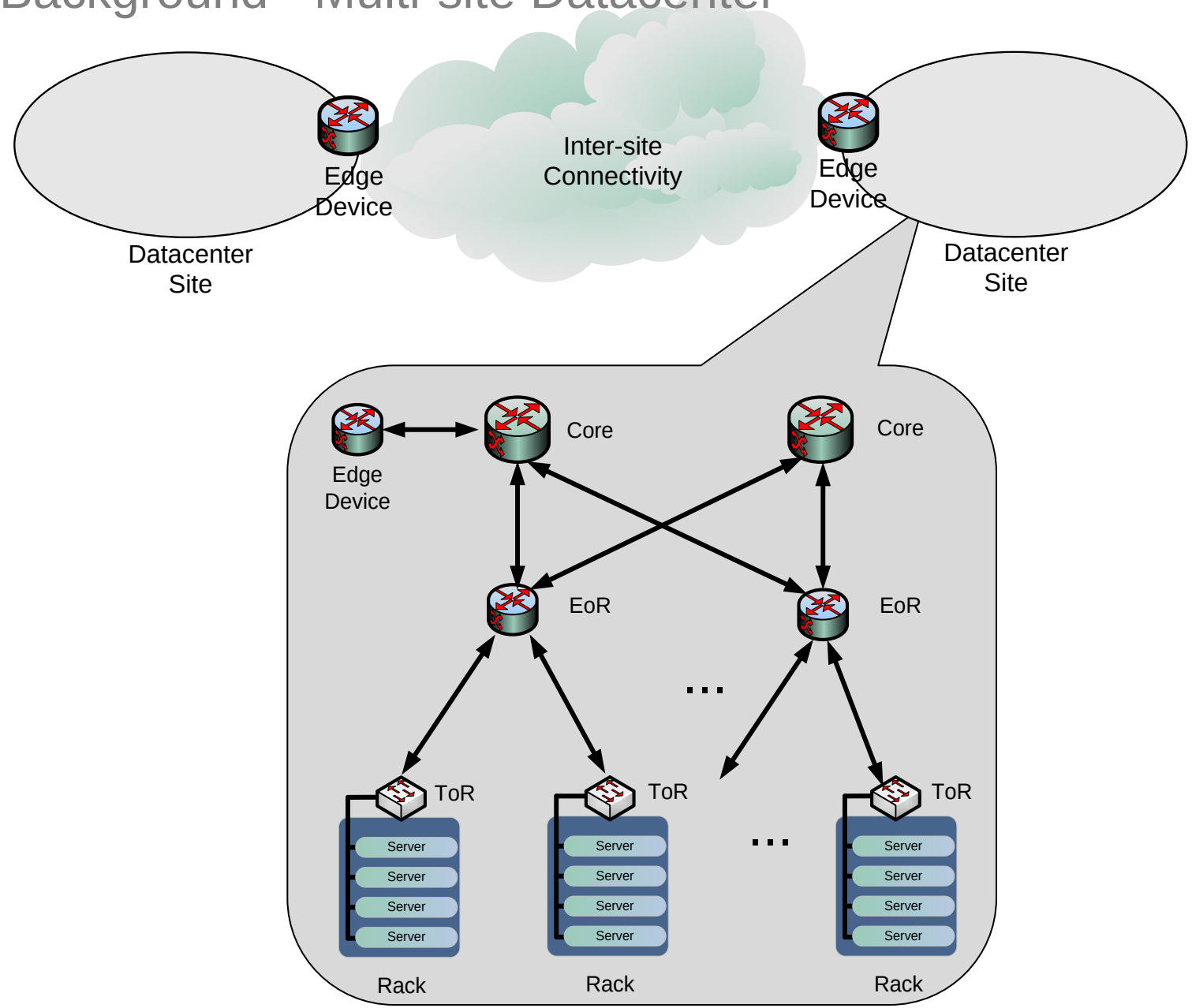
Scaling the Address Resolution Protocol for Large Data Centers (SARP)

draft-nachum-sarp

Youval Nachum	**Marvell**
Linda Dunbar	**Huawei**
Ilan Yerushalmi	**Marvell**
Tal Mizrahi	**Marvell**

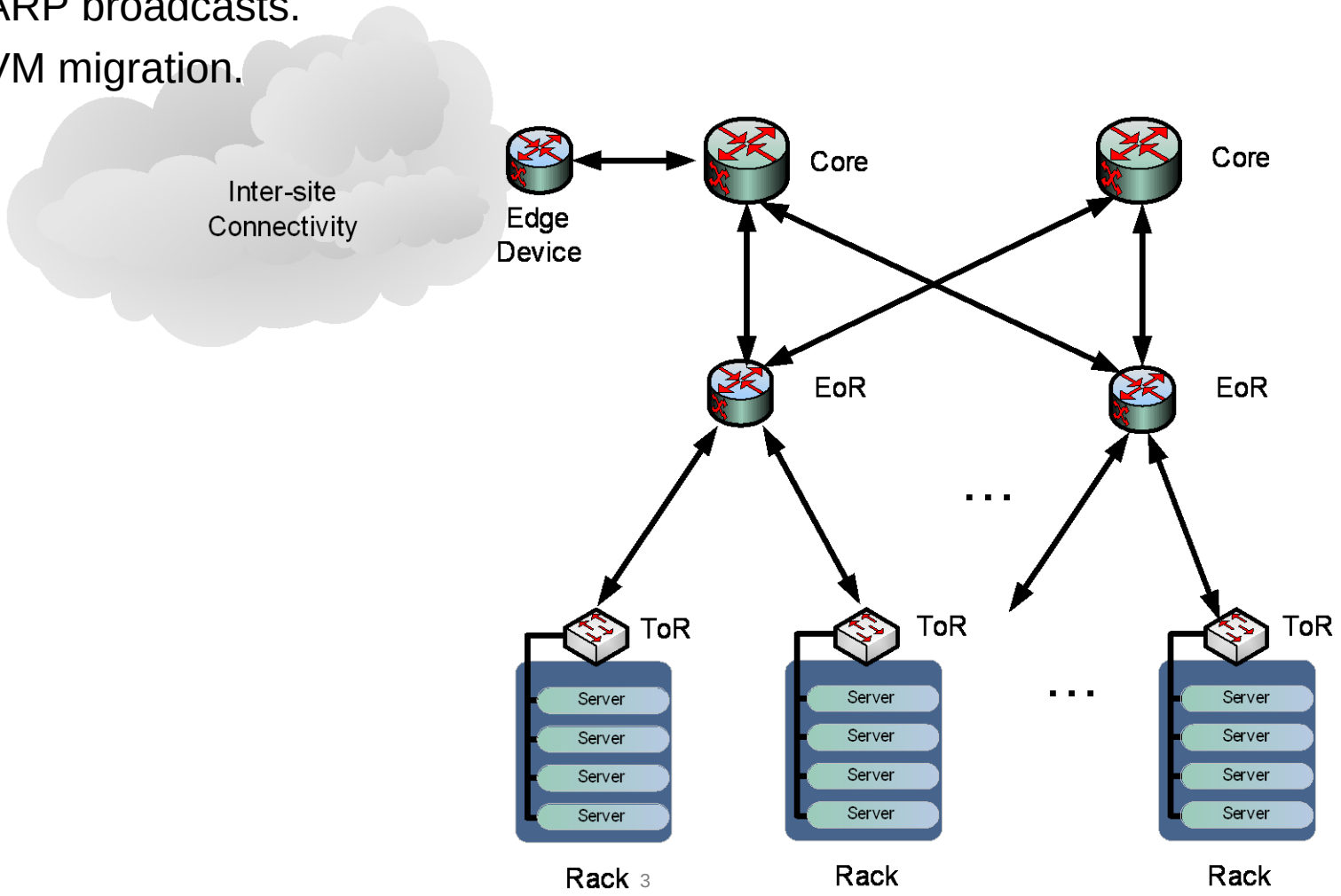
IETF Meeting 84, July 2012

Background - Multi-site Datacenter



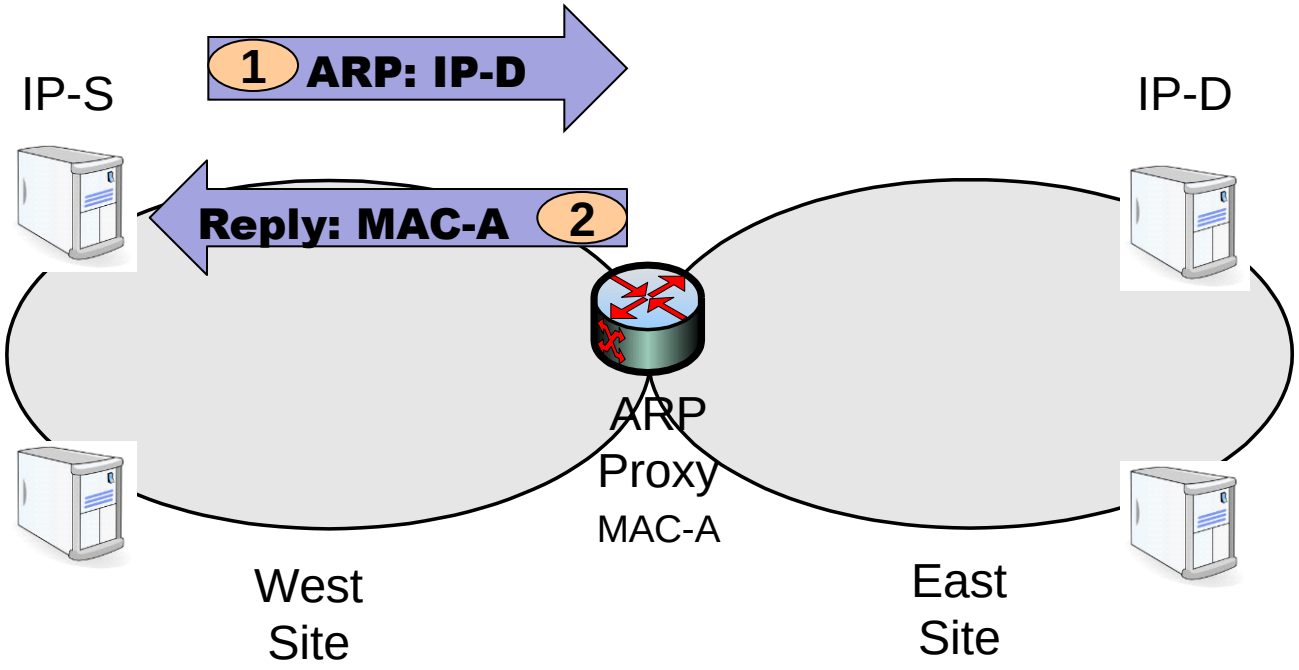
Background

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



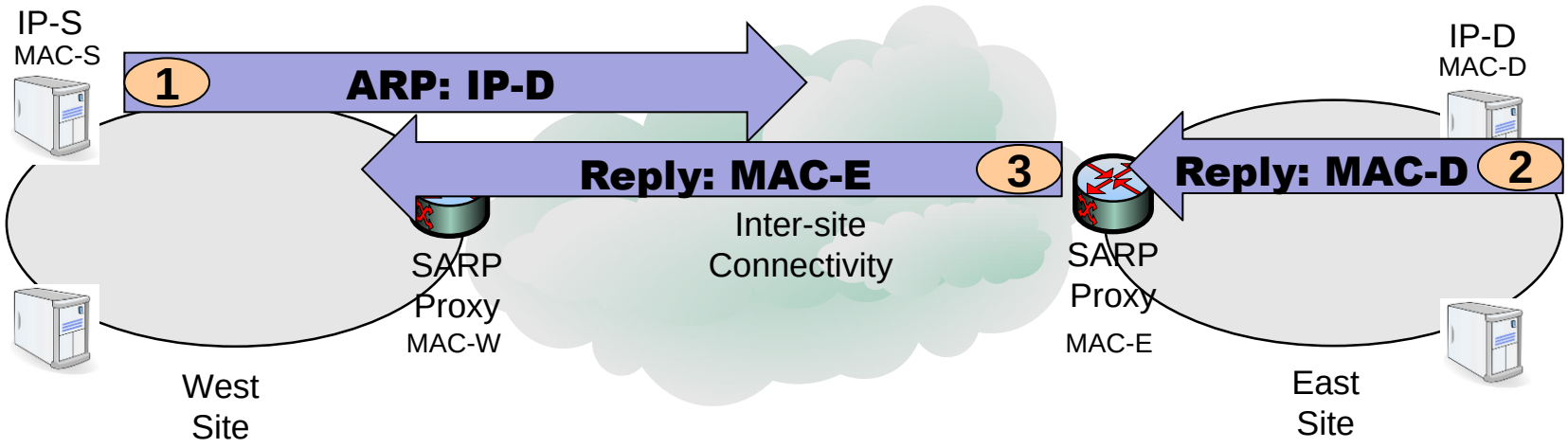
Background – Proxy ARP

- **Proxy ARP (RFC 1027, RFC 1009, RFC 925).**
- **Proxy ARP responds based on IP subnet.**
  - Assumption: IP subnet implies location.

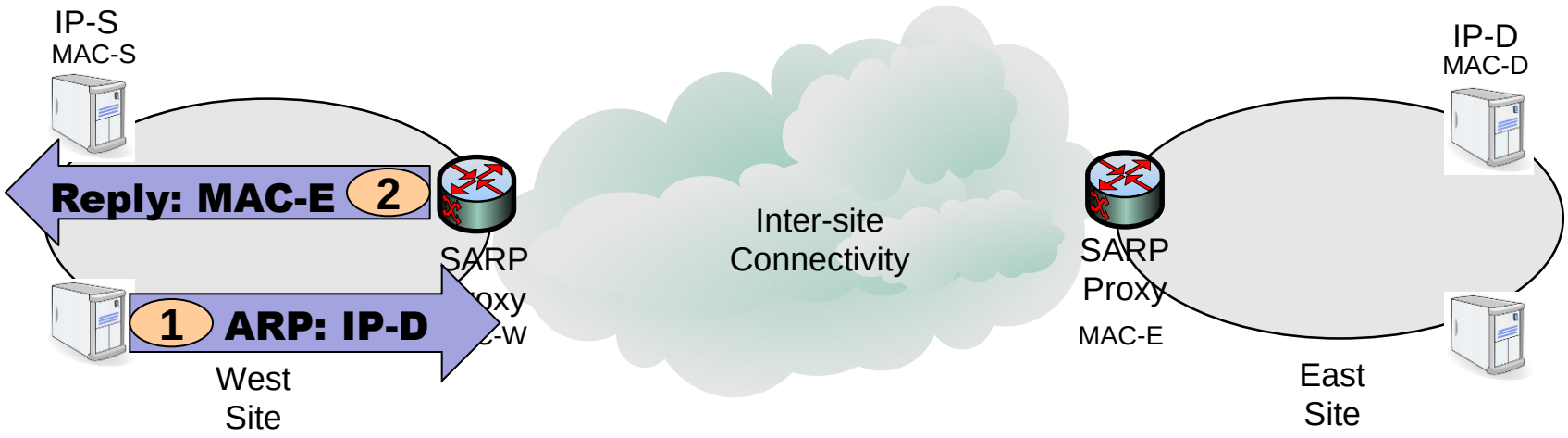


SARP

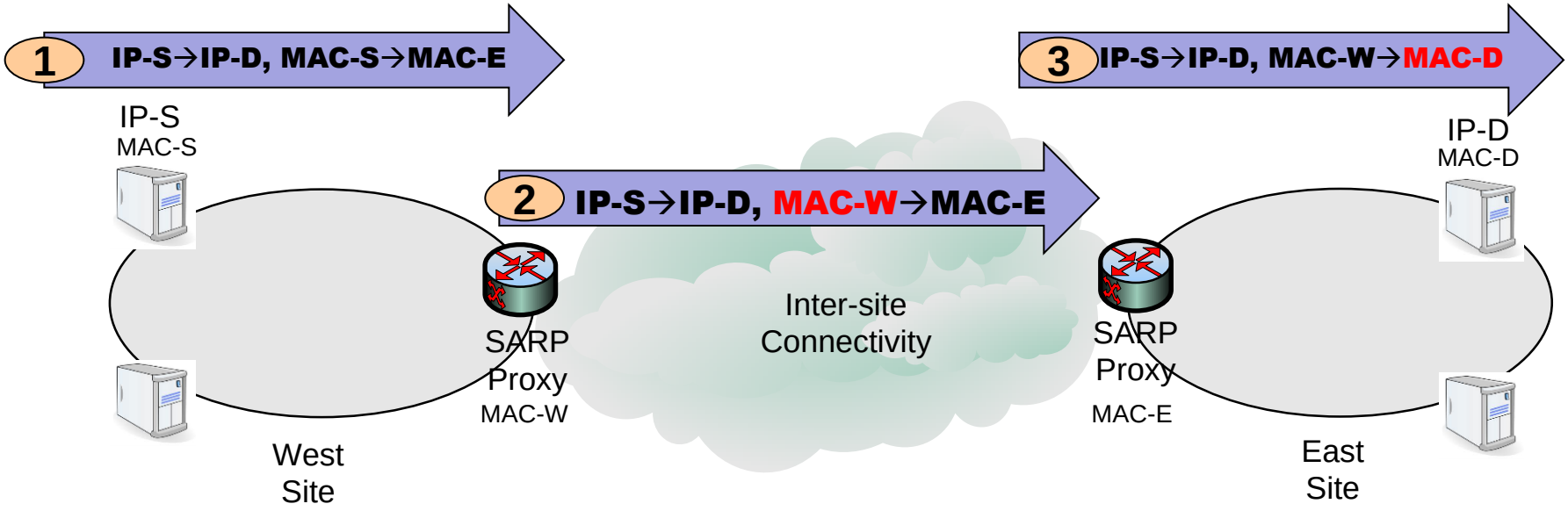
- **Edge devices: proxy SARP.**
- **IP subnet does not imply location.**
- **MAC-W / MAC-E imply location.**



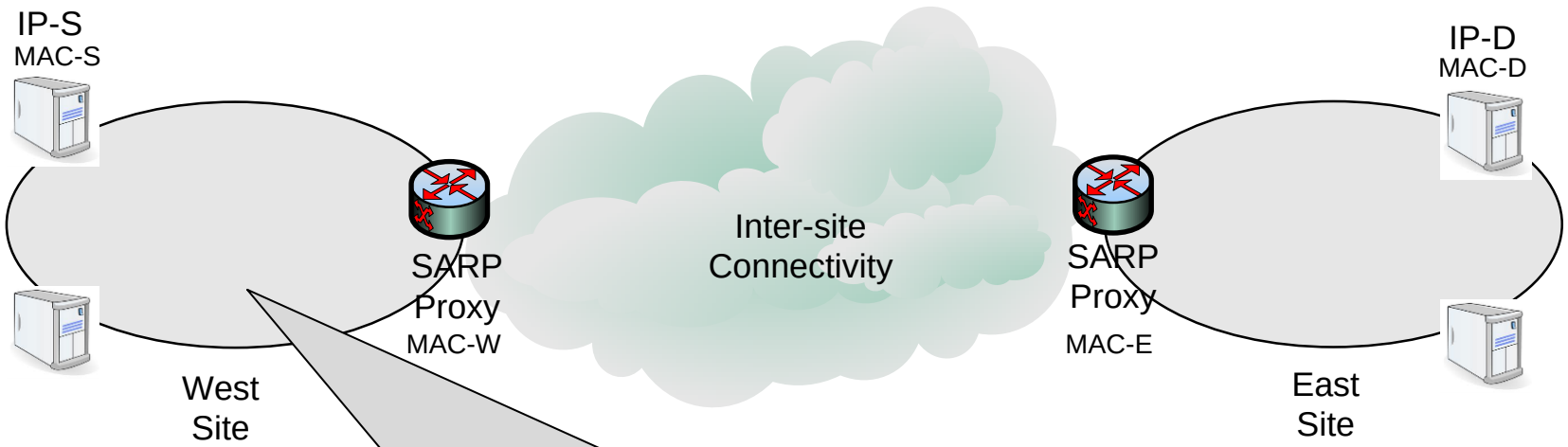
SARP Cache


IP-S
MAC-S
Reply: MAC-E
2
1
ARP: IP-D
West
Site
Inter-site
Connectivity
SARP
Proxy
MAC-E
IP-D
MAC-D
East
Site

SARP – Data Plane



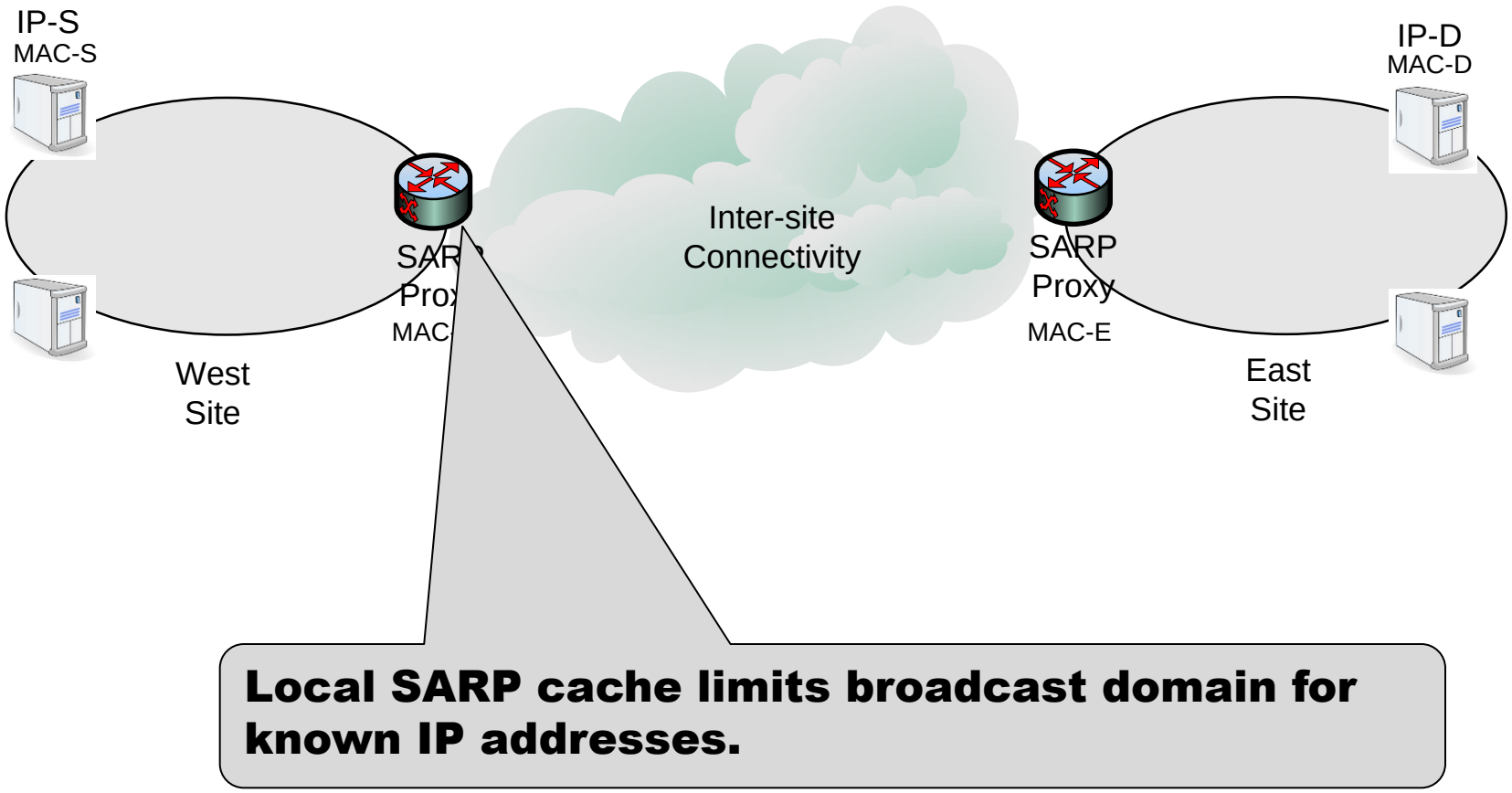
SARP – MAC Address Tables



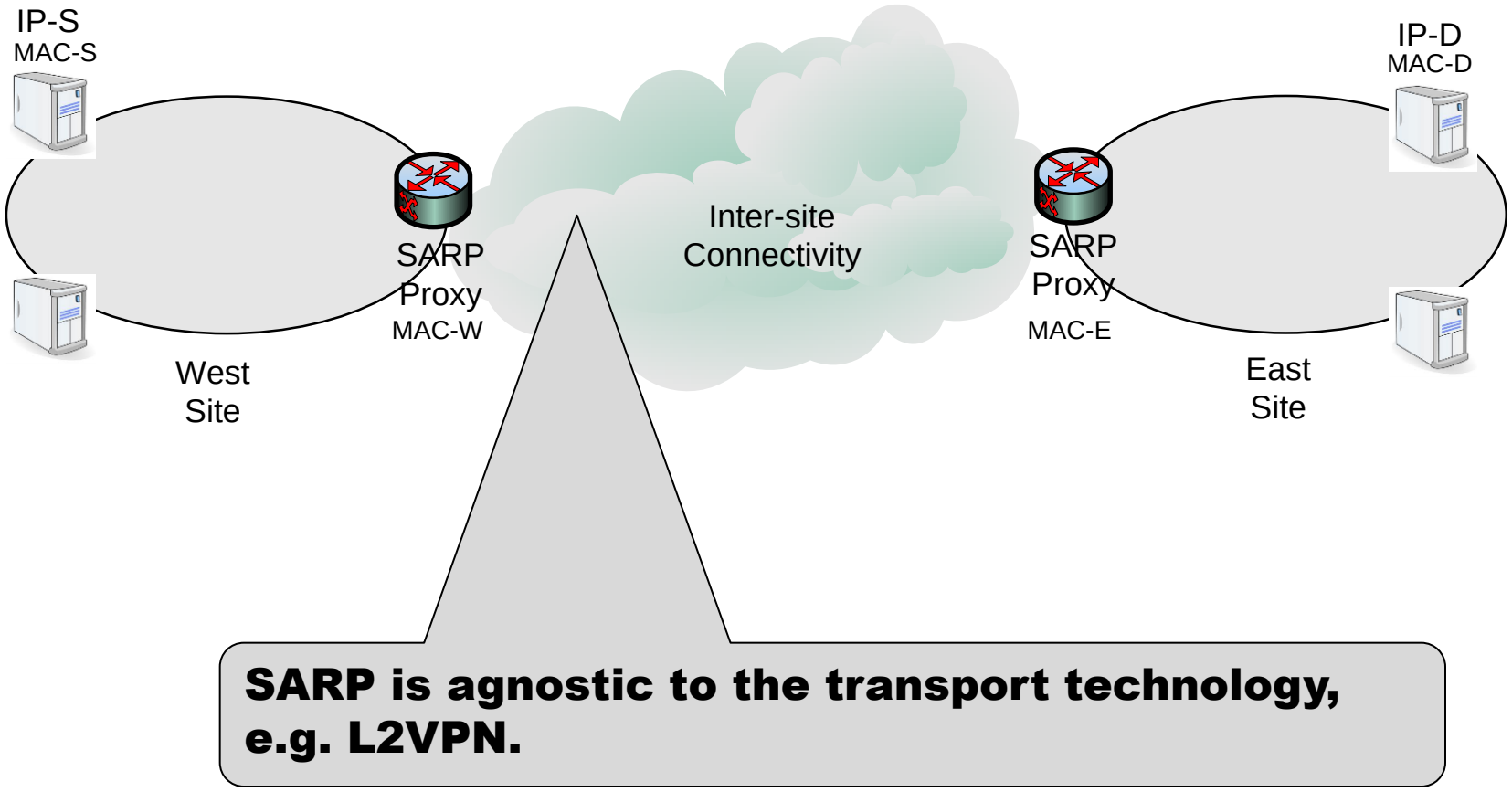
MAC address table of bridges in the west site:

- **Local site addresses, e.g., MAC-S.**
- **Edge devices, e.g., MAC-E.**
- **No need for addresses of remote sites.**

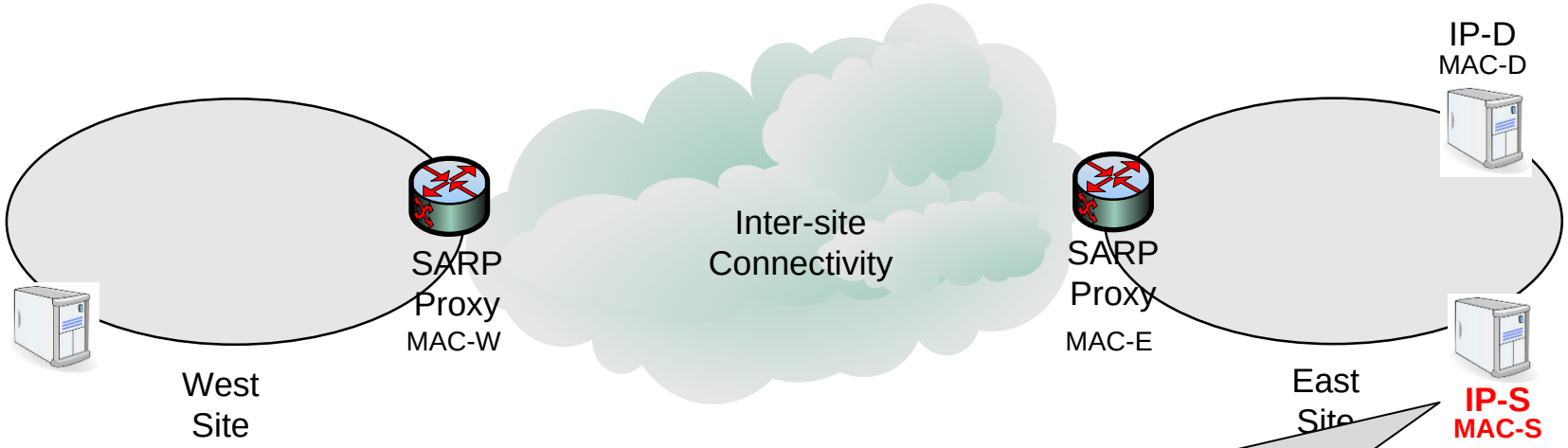
SARP – ARP Broadcast Domains



SARP over Overlay Network



SARP with VM Migration



- **Gratuitous ARP is used to notify network about migration.**
- **No need for additional control protocols.**
- **Transparent to inter-site network and protocols.**

Next Steps

History:

- **March 2012 – draft 00.**
- **June 2012 – draft 02.**
- **Discussion in ARMD mailing list.**

Next steps:

- **Receive feedbacks from WG.**
- **WG adoption.**

Thanks