

Multi-Chassis Link Aggregation

Two physical switches (same platform, same EOS version for simplicity and predictability) connected via a peer-link to form one logical switch for redundancy, higher resiliency, and allowing active/active use of all interconnects.

<https://www.arista.com/en/products/multi-chassis-link-aggregation-mlag>

Configuration

Peer A

Turn off spanning tree for the MLAG VLAN:

```
no spanning-tree vlan 4094
```

Configure the VLAN for MLAG control plane traffic; can be any VLAN but recommended to use 4094:

```
vlan 4094
```

Put this VLAN into a **trunk group** which removes this VLAN out of the default **switchport mode trunk** command:

```
trunk group MLAG-Peer
```

Create the L3 interface for the MLAG VLAN which carries the control plane traffic across the peer-link and serves as the source interface for the MLAG tunnel:

```
interface Vlan 4094
```

```
ip address 169.1.1.1/30
```

Set the interface to always be "up":

```
no autostate
```

Set MTU to allow jumbo frames:

```
mtu 9214
```

Create the L2 peer link between the two peers. Use LACP (mode active) and use at least two interfaces for redundancy:

```
interface Ethernet 51/1
```

```
description MLAG Peer Link Member
```

```
channel-group 2000 mode active
```

```
interface Ethernet 52/1
```

```
description MLAG Peer Link Member
```

```
channel-group 2000 mode active
```

```
interface Port-Channel 2000
```

```
description MLAG Peer Link
```

```
switchport mode trunk
```

Add the trunk group created earlier for the MLAG VLAN to allow this traffic to traverse this link:

```
switchport trunk group MLAG-Peer
```

Configuration (cont)

> Ensure native vlan packets are tagged. In later codes (>4.21.1F), this is internal code and doesn't need to be explicitly configured:

```
switchport trunk native vlan tag
```

Configure the MLAG domain.

```
mlag configuration
```

The Domain ID is case sensitive and has to match identically to its peer:

```
domain-id AristaMLAG1
```

Set the source interface:

```
local-interface Vlan4094
```

Set the destination interface:

```
peer-address 169.1.1.2
```

Set the peer link:

```
peer-link Port-Channel2000
```

```
interface Port-Channel X
```

```
switchport mode trunk
```

Configure port-channels to be MLAG member ports:

```
mlag <mlag_ID>
```

```
interface Ethernet Y
```

```
description Example MLAG member link
```

Use LACP (mode active) whenever possible.

```
channel-group X mode active
```

Peer B

Configure identically to Peer A with differences noted below:

```
interface Vlan 4094
```

```
ip address 169.1.1.2/30
```

```
no autostate
```

```
mtu 9214
```

```
!
```

```
mlag configuration
```

```
domain-id AristaMLAG1
```

```
local-interface Vlan4094
```

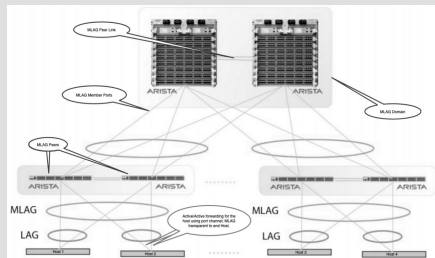
```
peer-address 169.1.1.1
```

```
peer-link Port-Channel2000
```

```
!
```

<https://www.arista.com/en/um-eos/eos-multi-chassis-link-aggregation>

Sample Topology and Terminology



*Only two switches per MLAG domain.

*The peer link is mainly for control plane traffic, but L3 traffic (these are still two "routers" even though they are simulating one "switch") and Layer 2 data plane traffic could still traverse the peer link (single-homed devices aka orphan ports or active/down MLAG interfaces.)

Confirmation

show mlag config-sanity

Run on both peers to confirm no inconsistencies or issues

show mlag detail

Confirm MLAG is active, which peer is Primary/Secondary, timers, number of active-full/active-partial interfaces, etc.

show mlag interfaces [detail|members|states]

Confirm information on MLAG interfaces.

show mlag issu warnings

Displays a warning message regarding the backward-compatibility of this feature before upgrading.

tcpdump

`#bash tcpdump -nei vlan4094 port 4432 ==> to confirm control plane traffic for mlag is being sent and received.`

Things to Note

MLAG System ID

Derived after Primary Peer is elected (lowest MAC address); persistent across reboots; LACP and STP control packets uses this to emulate one "logical" switch.

STP

Active only on the Primary Peer; configuration needs to be consistent across both peers for seamless failover; Secondary will forward BPDUs to Primary across the Peer Link.

Things to Note (cont)

TCP and UDP Port 4432

Must be permitted in control plane ACL on both peers if non-default control plane ACL is used.

MAC address table

Sync'ed between MLAG peers for active/active member ports.

IGMP Snooping Tables

Sync'ed between MLAG peers for active/active member ports; configuration should be consistent across both peers.

ARP table sync?

No ARP table sync with MLAG.

L3 Sync?

MLAG is L2 active/active technology. To get an active/active L3 Gateway use VARP, but keep in mind that though MLAG forms one logical switch between two switches, there are two separate Layer 3 control planes on each peer.

For any additional questions, please check EOS Central/Arista Forum or contact Arista TAC. <https://www.arista.com/en/support/customer-support>

Configuration Notes

Turn off STP on the MLAG peering VLAN to prevent the peer link from going into the discarding state.

Recommended to use VLAN4094 but any VLAN can be used. Use the same VLAN throughout all MLAG domains for consistency.

Put the MLAG peering VLAN into a trunk group to ensure this VLAN isn't used by any other ports and avoid any possible loop conditions being created.

Use "no autostate" on the MLAG SVI to ensure this interface remains UP.

For the Peer Link port-channel, use a minimum of two interfaces across multiple ASICs or line cards for optimal redundancy.

The Peer Link capacity should be equal to the Leaf-to-Spine total capacity to avoid losing capacity if the uplinks fail.

The MLAG Domain ID needs to be identical (case-sensitive) across Peers as well as be unique in regards to other Leaf MLAG pairs.

Configuration Notes (cont)

The MLAG IDs need to match across Peers, but it is recommended to use the same port-channel ID across both Peers if possible for simplicity in operations and troubleshooting.

MLAG timers should be kept to the default values. Reload Delay is the interval that MLAG interfaces are disabled after an MLAG peer reboots. Non MLAG Reload Delay is the interval that non-MLAG links are disabled after an MLAG peer reboots.

After 4.21.1F, "switchport mode trunk native vlan tag" is no longer needed to be explicitly configured on the MLAG Peer Link.

C

By **sh-arista**
cheatography.com/sh-arista/

Not published yet.
Last updated 25th November, 2019.
Page 3 of 3.

Sponsored by **Readable.com**
Measure your website readability!
<https://readable.com>