ECS

Version D- and U-series (Gen 1 and Gen 2)

Networks Guide for D- and U- Series (Gen 1 and Gen 2) Hardware

302-005-209

01

DELL EMC

Published August 2018

Dell EMC
Hopkinton, Massachusetts 01748-9103
1-508-435-1000 In North America 1-866-464-7381
www.DellEMC.com

CONTENTS

Chapter 1 **ECS Network Documentation** **5**
Introduction....6

Chapter 2 **Public Network** **7**
Public network.... 8
Types of traffic on the public network....8
Public switches....8
Connecting the public network to the customer networks.... 9
Single uplink connection....10
Multiport uplink connection to a single customer switch....10
Multiport uplink connection to multiple customer switches using MLAG or vPC 11

Chapter 3 **Private and OBM Networks** **13**
Private and OBM networks.... 14
Private switch.... 14
Private network....14
Private.4 network 15
Overlaying private.4 network and customer uplink....15
Private.4 topology....16
OBM (Out of Band Management) network....20
Customer uplink for OBMs.... 20

Chapter 4 **Access to Switches and Configuration Files** **21**
Access to the switches....22
ECS switch configuration files 22

Appendix A **ECS Networking Glossary of Terms** **23**

CHAPTER 1

ECS Network Documentation

- Introduction 6

Introduction

This document describes the ECS public, private, and OBM networks, and how they are used to connect to the customer networks, and connect within the ECS system.

This document is intended for use with systems running on D- and U- Series (Gen 1 and Gen 2) hardware.

This document assumes you have read, and are familiar with the ECS networking concepts and operations provided in the following documents.

For information about ECS:	See the ECS	Which is available from:
Network ports, and encryption	*Security Configuration Guide*	support.emc.com
Network cabling	*Hardware Guide*	ECS Product Documentation page
Switch configuration, and the steps to configure the static IP addresses	*ECS Software Installation Guide for D- and U- Series (Gen 1 and Gen 2) Hardware*	SolVE

Other documentation resources

For information about:	See the ECS	Which is available from:
General ECS networking principles, guidelines, and best practices	*Networking and Best Practices*	emc.com For Mac users copy and paste URL into your web browser: https://www.emc.com/collateral/white-paper/h15718-ecs-networking-bp-wp.pdf
Steps and examples for configuring the private switch for ECS Gen 1 and Gen 2 systems.	*Configuring the RMM Interface on the Private Arista Switches*	SolVE
Troubleshooting, or monitoring the Arista switch.	*Arista EOS Command Line Interface documentation*	Arista EOS Command Line Interface documentation

CHAPTER 2

Public Network

- Public network 8
- Types of traffic on the public network 8
- Public switches 8
- Connecting the public network to the customer networks 9

Public network

The default network of the appliance with connections to the Public Switch. By default, all types of public traffic will use the public network unless explicitly defined.

Interface name: public (slave-0, slave-1)

The public network:

- Provides connectivity between the ECS traffic and the customer networks.
- Enables replication of data to other sites in an ECS federation.

Types of traffic on the public network

The following types of traffic run on the public network.

- Data — customer data and I/O requests
- Management — hosts the ECS web portal, all common infrastructure services such as NTP, DNS, DHCP as well as Secure remote services (SRS).
- Replication — data replicated to other nodes within a replication group.

Note

The gateway is always configured on the public network.

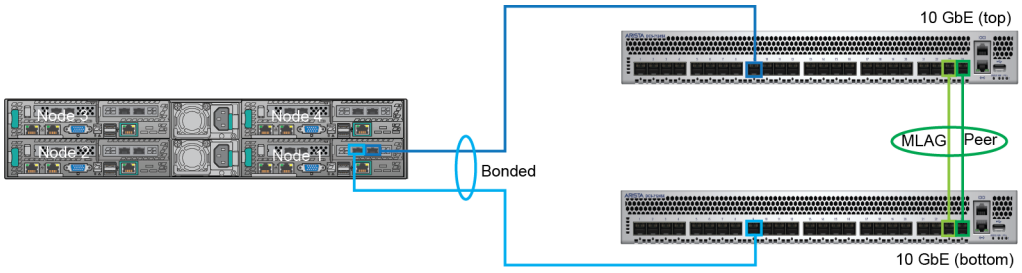
In a standard ECS deployment the data, management, and replication traffic runs through the public network. This chapter provides examples of a standard ECS deployment. Optionally, the traffic can be separated to run on dedicated networks as described in *ECS Network Separation Guide*, which is available in SolVe. Be sure to refer to the version of the guide that corresponds to the ECS software version you are running.

Public switches

The switches that connect to the customer's network. This includes the default public network and any defined separated networks such as management, replication or data. The switches are stacked on top of each other, and are located on the top of the ECS rack, and are sometimes referred to as the "top-of-rack" switches. The terms rabbit (top) switch, and hare (bottom) switch are further used to differentiate between the two switches.

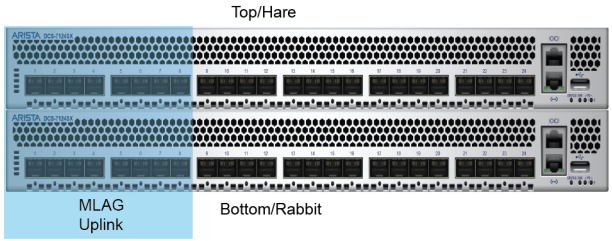
The two switches work in tandem using LACP/MLAG to create High Availability network for the nodes in the rack. By leveraging the ability of Link aggregations sub-second recovery time, the ECS has a near zero fault tolerant recovery time. Each node in the rack is connected to both rabbit and hare through two 10Gb NICs which are aggregated together using a Linux bonding driver. The node is configured to bond the two 10Gb NIC into a single LACP bonding interface also known as a "mode 4" bond. This bonding interface connects one port to rabbit and another port to the hare. The two switches are configured with LACP/MLAG to create a single LAG interface as demonstrated in the following image.

Figure 1 Physical Topology



Connecting the public network to the customer networks

Both rabbit (top) and hare (bottom) have eight 10Gb ports available to connect to the customer network. The configurations of the rabbit and hare have the first 8 ports on each of the switches configured as a single LACP/MLAG interface.



Fallback mode

The LAG interfaces also provide a fallback mode for connections to customer networks that cannot provide LACP to the ECS rack. The LACP fallback mode will not provide redundancy and is only activated if no LACP advertisements are detected by the protocol. Once the fallback criteria have been met then the lowest active port will be activated. All other linked ports in the LAG will be disabled. If any of the uplink ports receive advertisement from a neighbor then fallback mode is disabled and the LACP will establish the state of the port.

Types of connectivity to the customer networks

The public switches can be connected to the customer network in anyone of the following ways:

- Single uplink connection (not recommended for a production system)
- Multiport uplink connection to a single customer switch
- Multiport uplink connection to multiple customer switches using MLAG or vPC

Spanning Tree in a customer environment

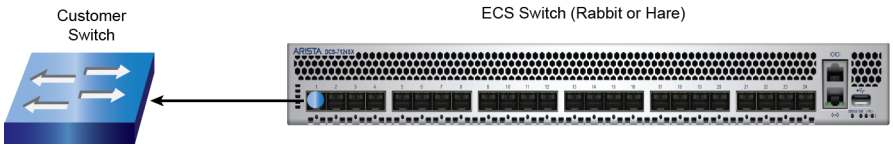
The switches in the ECS rack have been configured to not participate in the customers spanning tree topology. Both switches in the rack will be presented as an edge device or host device. Both rabbit and hare have been configured to filter all spanning tree BPDUs from the uplink ports. This will ensure that none of the ECS racks will disrupt the customer network due to link or switch failure in the rack. This will simplify the configuration of the ECS switches in the customer network. The ECS will only support single link or single/multi switch LAG configuration. An RPQ is required to make any changes to the default configurations.

Single uplink connection

Connecting a single ECS switch to the customer network is strongly discouraged.

However, this setup can be used pre-production installation or for customer's that cannot provide a LAG connection to the ECS rack. This configuration does not provide high availability redundancy.

Figure 2 Single uplink connection



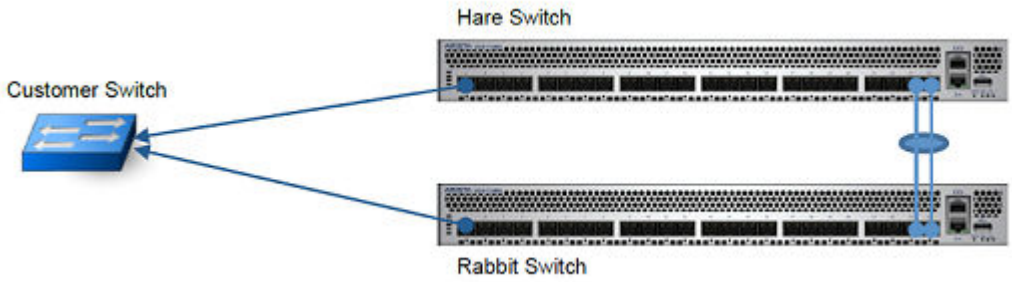
Multiport uplink connection to a single customer switch

When a single customer switch is connected to both of the ECS public switches, the customer switch will need to be configured to allow the connection to the rabbit and hare MLAG port channels. This is done by creating the matching port channel on the customer's switch using LACP in active or passive mode. No additional work should have to be done on the ECS public switches.

Example of a two port LAG port configuration

The following example is demonstrating with a Cisco switch.

Figure 3 Two port LAG port configuration



Example configurations:

- Arista preconfigured on the ECS switches

```
interface Ethernet 1-2
channel-group 100 mode active
```

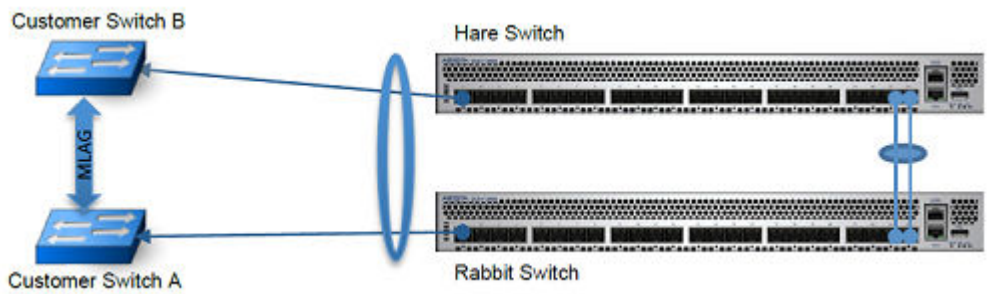
- Cisco configuration on the customer switch

```
interface Ethernet1/1
channel-group 100 mode active
interface Ethernet1/2
channel-group 100 mode active
```

Multiport uplink connection to multiple customer switches using MLAG or vPC

To connect the ECS rack to a multi-switch LAG, the customer will need to configure their switches with a port channel as described in Multiport uplink connection to a single customer switch and both port channels will need to be connected together using a multi-switch LAG protocol like Arista MLAG or Cisco vPC to connect to the rabbit/hare uplink MLAG port channel. All port channels that are participating in the multi-switch LAG must have LACP in either active or passive mode.

Figure 4 Example of a four port multi-switch LAG



Example configurations:

- ECS Rabbit preconfigured Arista switch

```
interface Ethernet 1-2
channel-group 100 mode active
interface port-channel 100
mlag 100
```

- ECS Hare preconfigured Arista switch

```
interface Ethernet 1-2
channel-group 100 mode active
interface port-channel 100
mlag 100
```

- Customer Cisco Switch A configuration

```
interface Ethernet1/1
channel-group 100 mode active
interface Ethernet1/2
channel-group 100 mode active
interface port-channel 100
vpc 100
```

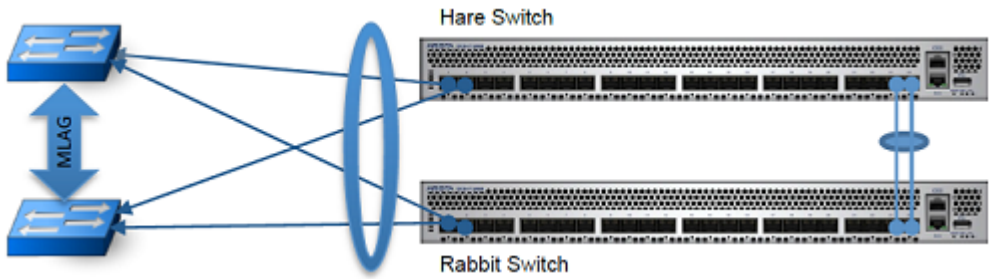
- Customer Cisco Switch B configuration

```
interface Ethernet1/1
channel-group 100 mode active
interface Ethernet1/2
channel-group 100 mode active
interface port-channel 100
vpc 100
```

Example of a multi-switch LACP with four ports

If more than two links are connected to each of the rabbit and hare switches then the links from each of the switches should spread to both of the customer switches in a bowtie fashion to provide optimal redundancy and performance during failures or scheduled downtime.

Figure 5 Multi-switch LACP with four ports



CHAPTER 3

Private and OBM Networks

- Private and OBM networks 14
- Private switch 14
- Private network 14
- Private.4 network 15
- OBM (Out of Band Management) network 20

Private and OBM networks

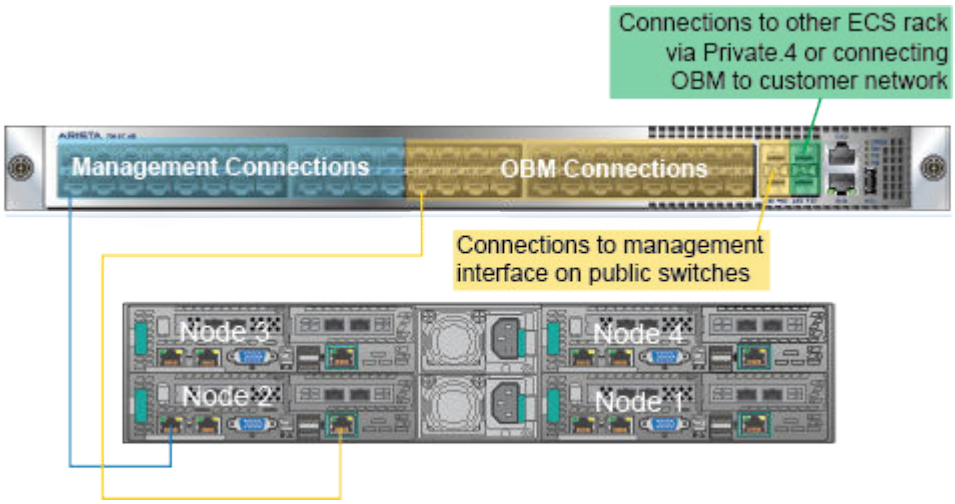
Two private (private, and private.4) networks, and the Out of Band Management network are configured on the same private switch.

Private switch

The switch used for internal maintenance including the ECS private network, private.4 network and Out of Band Management (OBM) connectivity. This is sometimes referred to as the Turtle switch.

All server nodes in an ECS intra-rack have two connections going to the private switch. The first is the 1Gbe port from the server node, which connects the node to the private.4 network. The other is the connection to the Out of Band Management (OBM) of the node through a virtual KVM network application as demonstrated below.

Figure 6 Private (Turtle) switch



Spanning tree is enabled on all ports. This is to ensure that we can create a loop free environment in the private.4 topology or customer management network.

Private network

The private network is an intra-rack only network used for service operations such as; Install, reinstall, and expansion. The private network connects all nodes and switches to a local area network to limit traffic to members of the intra-rack. Limiting the traffic to the intra-rack allows for easy management of all the nodes in the intra-rack and to reduce scalability issues.

All the nodes in an intra-rack are connected to the private switch through a 1Gb Ethernet port. IP address 192.168.219.0/24 is reserved for the private network. the physical Ethernet port on the nodes is configured with the private IP address and the node ID, for example; 192.168.219.*port_number*.

Figure 7 Private network

The subnet is not routable and is reused by all intra-racks in an ECS inter-rack network.

Private.4 network

The private.4 network interconnects multiple, co-located ECS intra-rack networks into a single inter-rack network through VLAN 4.

From the 1Gb management interface this LAN also uses the first 24 ports, ports 25-48 are used to connect the OBM (optional) and ports 51 and 52, to connect segments together using a VLAN tag.

All participants in the inter-rack LAN will tag their IP traffic with VLAN ID 4, and communicate using the IPv4 link local subnet 169.254.0.0/16.

All nodes in the rack will be assigned an address in this network once the intra-rack has been given color designation. This color is mapped to the rack ID. The rack ID along with the node ID will make up the new inter-rack IP address for every node belonging to the inter-rack for example, (169.254.{RackID}.{NodeID})

Figure 8 Private.4 network

Inter-rack Interconnect
The last two ports on the private switch are used to connect to other intra-rack LANs. Refer to Private.4 (NAN) topology for details.

Overlaying private.4 network and customer uplink

Ports 51 and 52 on the private switch are setup in a hybrid mode. A hybrid mode allows the switches to behave like both an access port and a trunk port at the same time. The hybrid mode is useful when reusing the same ports for multiple purposes.

If you are using an external switch to connect ECS private switches together then a similar configuration must exist on the external switch to ensure proper connectivity to all entities in the system and to prevent leaking of unwanted traffic to other networks. The external switch should be configured to forward VLAN 4 as tagged traffic to only the ports that are connected to ECS private switches. This will guarantee that private.4 is a closed network.

Example: Arista switch configuration for hybrid port from turtle

```
interface Ethernet51
description Nile Area Network Uplink
mtu 9212
switchport trunk native vlan 6
switchport trunk allowed vlan 4,6
switchport mode trunk
```

where

- `switchport trunk native vlan 6` — Assigns untagged packets to VLAN 6 on the way in, and untags packets from VLAN 6 on the way out. VLAN 6 is only required when the OBM option is configured. For details refer to Customer uplink for OBMs.
- `switchport trunk allowed vlan 4,6` — Only traffic from the listed VLAN(s) are allow to forward traffic.
- `switchport mode trunk` — Enables tagged and untagged traffic on the port.

Private.4 topology

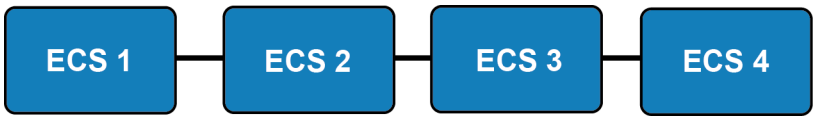
The ECS intra-rack backend management networks are connected together to create the inter-rack topology.

By connecting either port 51 or 52 to another private switch from another ECS intra-rack, the inter-rack network is created. Through these connections nodes from any intra-rack can communicate to any other node on the inter-rack network as described below:

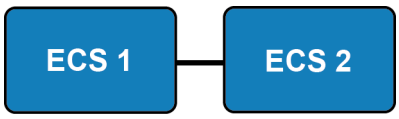
- Daisy chain or line topology
- Ring topology
- Star topology

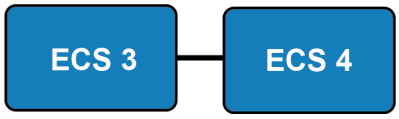
Daisy chain or line topology

The simplest topology to connect the intra-racks together does not require any extra equipment. All the private switches can be connected together a linear or daisy chain fashion as demonstrated below.

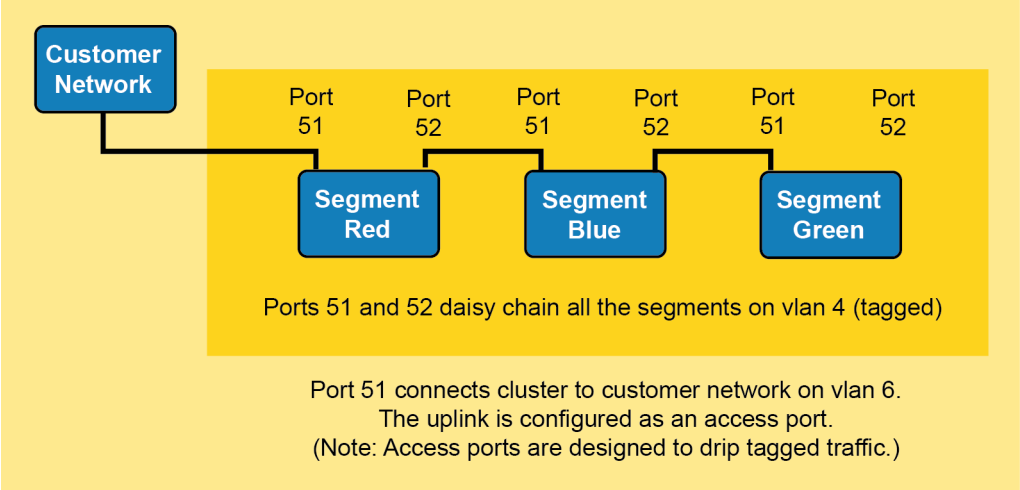


This linear or daisy-chain topology is the least dependable setup and is easily susceptible to split-brain topologies as demonstrated below.





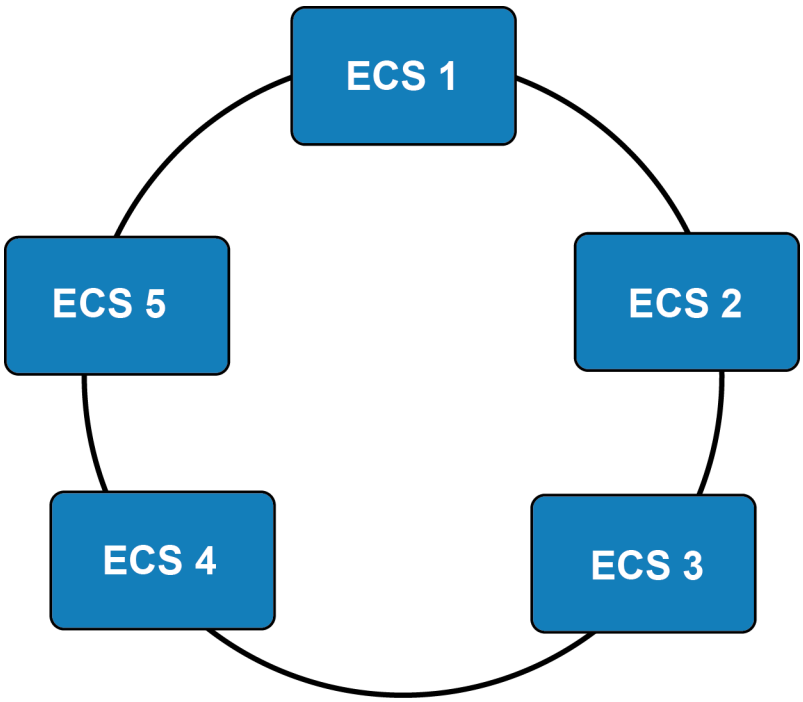
With a daisy chain, or linear topology, the intra-racks are connected together end-to-end using ports 51 and 52 as demonstrated below.



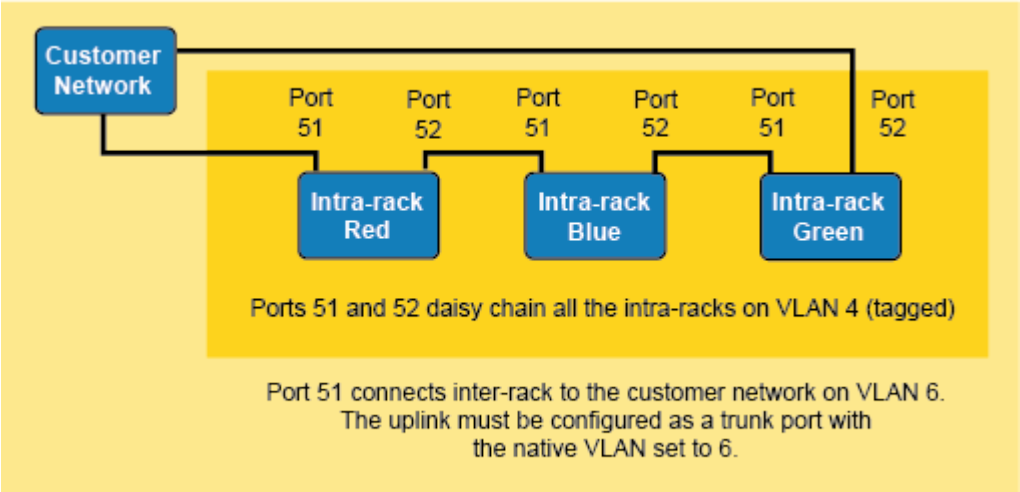
The uplink to the customer network can be attached from either end of the daisy chain. In the above example, the customer uplink is attached at the red intra-rack, however, the uplink to the customer network could also be attached at from the green intra-rack. This will allow connectivity to the cluster LAN using VLAN 4, and provide connectivity to customer facing OBM interfaces from all the rack segments.

Ring topology

For a more reliable network, the ends of the daisy chain topology can be connected together to create a ring network as demonstrated below. The ring topology would require two physical link breaks in the topology to create split-brain issue in the private.4 network.



The ring topology is very similar to the daisy chain/line topology, except that it is more robust since it requires two points of failure to break the topology which would cause a split-brain issue.



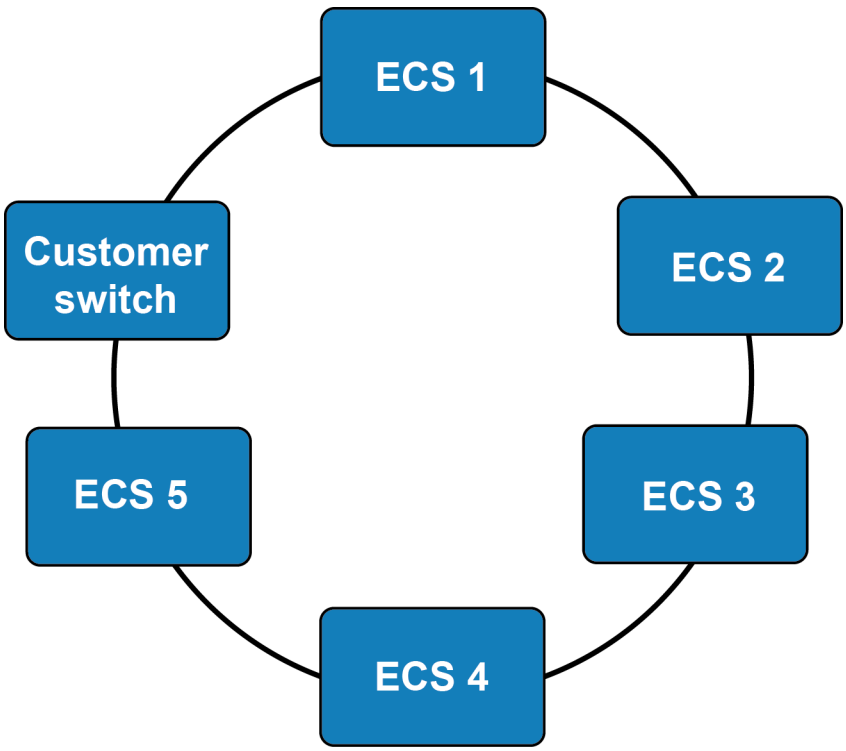
The customer uplink ports need to be configured to , and to also , and to

- Isolate VLAN 4 traffic to only the uplinks
- Allow tagged and untagged traffic to pass between the two ports for VLAN 4 traffic
- Allow the untagged traffic to be forwarded up the customer network.

For details about configuring the customer ports see: Overlaying private.4 network and customer uplink.

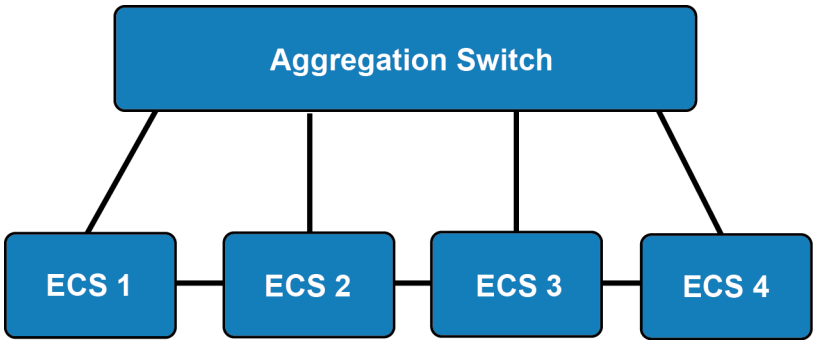
The big drawback to the ring topology is the OBM ports cannot be connected to the customer network unless the external customer or aggregation switch is added to the ring as demonstrated below.

For details about the OBM network refer to OBM (Out of Band Management) network.



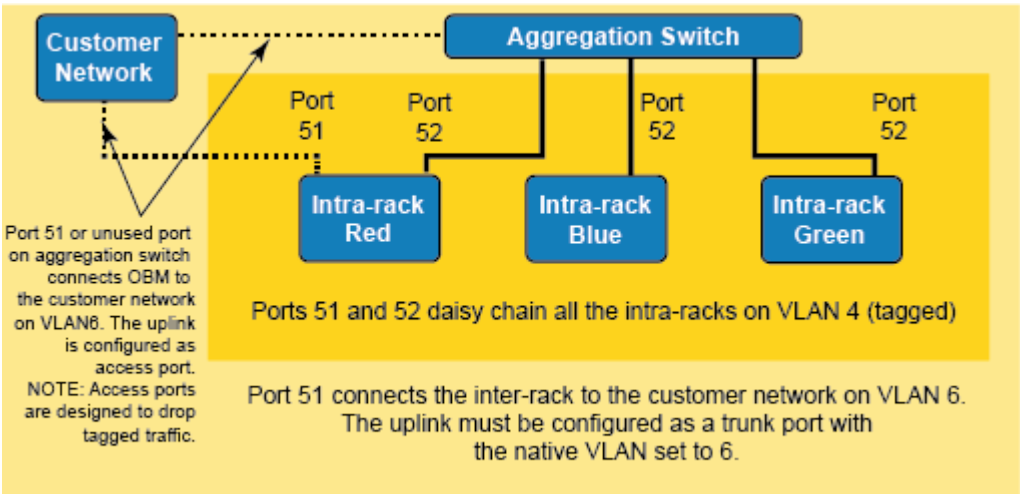
Star topology

The limitation with the daisy chain or ring topologies is that they do not scale well for large installations. For ten or more ECS racks an aggregation switch should be added support the large installation.



By using an aggregation switch to connect to all intra-racks, the star topology provides better protection against the split-brain issue than both the daisy chain/line, or ring topologies. With an aggregation switch, link failures are isolated to a single intra-rack in the private.4 network.

The aggregation switch connecting to the intra-racks must be setup as a trunk and allow VLAN traffic to flow between all ports in the inter-rack.



You have the option of connecting out-of-band management (OBM) to the customer network using one of the unused ports on one of the private switches. However if you'd like the same fault tolerance for the customer OBM as configured for the intra-racks, the customer uplink can be plugged into the aggregation switch. The configuration of the aggregation switch will be similar to the configuration of ports 51 and 52 on the private switch for the connections to the ECS intra-racks. The customer uplink on the aggregation switch must be configured as an access port to prevent the leaking of VLAN 4 traffic to the customer network.

For details about the OBM network refer to OBM (Out of Band Management) network.

Example 1 Sample aggregation switch configuration for Arista 7048t/7010t

Example 1 Sample aggregation switch configuration for Arista 7048t/7010t (continued)

```
vlan 4,6
interface Ethernet1-51
description Nile Area Network Uplink
mtu 9212
switchport trunk native vlan 6
switchport trunk allowed vlan 4,6
switchport mode trunk
exit

interface Ethernet52
description Customer Uplink for RMM
mtu 9212
switchport access vlan 4
exit
```

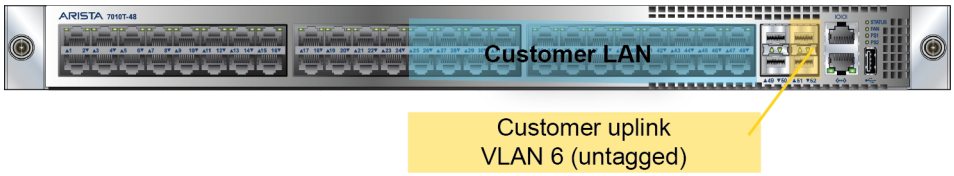
OBM (Out of Band Management) network

The Out of Band Management network, is an optional network, which allows for remote maintenance of the physical servers. There is a vendor specific software connection on-top of the OBM which goes by many different names such as RMM, iDrac, iLo, etc.

By default, the OBM interfaces are configured to be connected to the customer network (See: Customer uplink for OBMs).

Customer uplink for OBMs

By default, the OBM ports, which are connected to the private switch from ports 25-48, are configured to connect to the customer network through the last two ports on the private switch.



To allow ports 51 and 52 to be used as customer uplink, the peer connection from the customer network must setup as an access port to allow untagged traffic to and from the private switch. By configuring the customer switch as an access port and connecting to a private switch, the customer will have access to all the OBM interfaces in the entire the private, inter-rack network. It is unnecessary to have a connection from every intra-rack. It may, however, require some configuration on the customer switch or aggregation switch to allow both cluster and customer traffic to travel on the same physical link.

VLAN 6 requirement

For the OBM traffic, ports should be setup to assign untagged traffic to VLAN 6 and strip the tag off of the VLAN 6 traffic.

ECS is configured for VLAN 6 to be used for RMM connectivity. Modifying this VLAN requires an RPQ, and must be done with an ECS customer support representative.

Once the VLAN that will be used in place of VLAN 6, the VLAN 6 configuration will still exist with some modifications, however VLAN 6 will not be used for ECS OBM traffic.

CHAPTER 4

Access to Switches and Configuration Files

- Access to the switches 22
- ECS switch configuration files 22

Access to the switches

The ECS switches are used for pure layer 2 purposes and therefore cannot be reached directly.

To access the switches you must log into Node 1 of the ECS as the admin user. Once logged in, the admin user can lock down ssh access. After logging into the node, you can validate the switch names as follows:

```
nslookup hare.rack
Server:         192.168.219.254
Address:        192.168.219.254#53

Name:   hare.rack
Address: 192.168.219.253

admin@ecsparis1:~> nslookup rabbit
Server:         192.168.219.254
Address:        192.168.219.254#53

Name:   rabbit.rack
Address: 192.168.219.252

nslookup turtle.rack
Server:         192.168.219.254
Address:        192.168.219.254#53

Name:   turtle.rack
Address: 192.168.219.251
```

After collecting the information you can use another ssh session, for example `ssh admin@hare`, to perform more commands. You can also change to enable mode to access more of the commands.

Refer to the Arista EOS Command Line Interface documentation for more information.

ECS switch configuration files

The switch configuration files are located on the appliance in the following directory.

`/usr/share/emc-arista-firmware/config/ecs/`, and include the following

- nile-hare-arista-24ports.cfg
- nile-hare-arista-52+ports.cfg
- nile-rabbit-arista-24ports.cfg
- nile-rabbit-arista-52+ports.cfg
- nile-turtle-arista-48ports.cfg

APPENDIX A

ECS Networking Glossary of Terms

This appendix lists and describes the common terminology used for ECS networking.

data network
An optional separated network dedicated to the transfer and operations of all customer data. Additional data networks maybe be defined.

Note

For ECS 3.1, only one additional data network can be defined for CAS only.

The initial interface name is public.data, subsequent data networks are named as follows public.dataN, where "N" is the number used to differentiate the data network, when there are multiple data networks. For example when two data networks are configured for data traffic the names are:

- public.data
- public.data2

hare
See public switches.

management network
An optional separated VLAN network dedicated to hosting the ECS web portal, all common infrastructure services such as NTP, DNS, DHCP as well as EMC's secure remote services (ESRS).

Interface name: public.mgmt

nile area network (NAN)
See private.4 network.

OBM network
The Out of Band Management (OBM) network allows for remote maintenance of the physical servers. There is a vendor specific software connection on-top of the OBM which goes by many different names such as RMM, iDrac, iLo, etc.

private network
A rack only network used for service operations such as install, reinstall, and expansion.

private.4 network
A network which interconnects all co-located ECS racks through their private switches onto a single VLAN, which is VLAN 4 by default. Also referred to as the Nile area network (NAN).

private switch
Switch used for internal maintenance including the ECS private network, the private.4 network (also known as Nile area network (NAN)), and Out of Band Management (OBM) connectivity. The customary name for this physical switch is Turtle.

public network
The default network of the appliance that consists of two bonded interfaces with connections to the Public Switch. By default, all types of public traffic will use the public network unless explicitly defined.

Interface name: public (slave-0, slave-1)

public switches
Switches that connect to the customer's network. This includes the default public network and any defined separated networks such as management, replication or data. The customary name for the two physical components in this switch complex are Rabbit and Hare.

rabbit
See public switches.

replication network
An optional separated network dedicated to replicating objects between Virtual Data Centers.

Interface name: public.repl

turtle
See private switch.