



Enterprise Advanced SONiC v4.5.0

QUICK START GUIDE

Revision 1.0

The information in this User's Manual has been carefully reviewed and is believed to be accurate. The vendor assumes no responsibility for any inaccuracies that may be contained in this document, and makes no commitment to update or to keep current the information in this manual, or to notify any person or organization of the updates. Note: For the most up-to-date version of this manual, see our website at <https://www.supermicro.com>.

Super Micro Computer, Inc. ("Supermicro") reserves the right to make changes to the product described in this manual at any time and without notice. This product, including software and documentation, is the property of Supermicro and/or its licensors, and is supplied only under a license. Any use or reproduction of this product is not allowed, except as expressly permitted by the terms of said license.

IN NO EVENT WILL Super Micro Computer, Inc. BE LIABLE FOR DIRECT, INDIRECT, SPECIAL, INCIDENTAL, SPECULATIVE OR CONSEQUENTIAL DAMAGES ARISING FROM THE USE OR INABILITY TO USE THIS PRODUCT OR DOCUMENTATION, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. IN PARTICULAR, SUPER MICRO COMPUTER, INC. SHALL NOT HAVE LIABILITY FOR ANY HARDWARE, SOFTWARE, OR DATA STORED OR USED WITH THE PRODUCT, INCLUDING THE COSTS OF REPAIRING, REPLACING, INTEGRATING, INSTALLING OR RECOVERING SUCH HARDWARE, SOFTWARE, OR DATA.

Any disputes arising between manufacturer and customer shall be governed by the laws of Santa Clara County in the State of California, USA. The State of California, County of Santa Clara shall be the exclusive venue for the resolution of any such disputes. Supermicro's total liability for all claims will not exceed the price paid for the hardware product.

FCC Statement: This equipment has been tested and found to comply with the limits for a Class A or Class B digital device pursuant to Part 15 of the FCC Rules. These limits are designed to provide reasonable protection against harmful interference when the equipment is operated in industrial environment for Class A device or in residential environment for Class B device. This equipment generates, uses, and can radiate radio frequency energy and, if not installed and used in accordance with the manufacturer's instruction manual, may cause harmful interference with radio communications. Operation of this equipment in a residential area is likely to cause harmful interference, in which case you will be required to correct the interference at your own expense.

California Best Management Practices Regulations for Perchlorate Materials: This Perchlorate warning applies only to products containing CR (Manganese Dioxide) Lithium coin cells. "Perchlorate Material-special handling may apply. See <https://www.dtsc.ca.gov/hazardouswaste/perchlorate>".



WARNING: This product can expose you to chemicals including lead, known to the State of California to cause cancer and birth defects or other reproductive harm. For more information, go to <https://www.P65Warnings.ca.gov>.



AVERTISSEMENT : Ce produit peut vous exposer à des agents chimiques, y compris le plomb, identifié par l'État de Californie comme pouvant causer le cancer, des malformations congénitales ou d'autres troubles de la reproduction. Pour de plus amples informations, prière de consulter <https://www.P65Warnings.ca.gov>.

The products sold by Supermicro are not intended for and will not be used in life support systems, medical equipment, nuclear facilities or systems, aircraft, aircraft devices, aircraft/emergency communication devices or other critical systems whose failure to perform be reasonably expected to result in significant injury or loss of life or catastrophic property damage. Accordingly, Supermicro disclaims any and all liability, and should buyer use or sell such products for use in such ultra-hazardous applications, it does so entirely at its own risk. Furthermore, buyer agrees to fully indemnify, defend and hold Supermicro harmless for and against any and all claims, demands, actions, litigation, and proceedings of any kind arising out of or related to such ultra-hazardous use or sale.

Manual Revision 1.0

Release Date: January 14, 2026

Unless you request and receive written permission from Super Micro Computer, Inc., you may not copy any part of this document. Information in this document is subject to change without notice. Other products and companies referred to herein are trademarks or registered trademarks of their respective companies or mark holders.

Copyright © 2026 by Super Micro Computer, Inc.
All rights reserved.

Published in the United States of America

Document Revision History

Date	Revision	Description
1/14/2026	1.0	Initial document.

Contents

Chapter 1: Introduction	6
1.1 Overview of Supermicro Enterprise Advanced SONiC	7
1.2 Applicable Use Cases	8
Layer 3 Clos	8
VXLAN EVPN Overlay	8
AI/ML Fabrics (Backend GPU Cluster)	9
1.3 Supermicro-Validated Designs	11
Chapter 2: Hardware Setup	12
Chapter 3: Login and Configuration	13
3.1 Initial Access and Login	14
3.2 Management Interface and Hostname Configuration	17
3.3 User Account Configuration	19
3.4 NTP Configuration	20
3.5 DNS Configuration	22
Chapter 4: Layer 2 and Layer 3 Enablement	23
4.1 Configuring Breakout Ports	24
4.2 Configuring Layer 2 Interfaces and VLAN	28
4.3 Configuring Layer 3 Interfaces	30
Chapter 5: Use Case: VXLAN/EVPN Fabric	32
5.1 Loopback and VTEP Configuration	33
5.2 BGP EVPN Configuration	34
5.3 VXLAN VTEP VNI Mapping	37
5.4 Sample VXLAN EVPN Configuration	38
Topology	38
VXLAN EVPN Verification	46
Chapter 6: Use Case: AI Fabric	49
6.1 Adaptive Routing Configuration	50
Sample Adaptive Routing Configuration	51
6.2 Adaptive Routing Verification	52
6.3 RoCEv2 Configuration and Verification	55
6.4 Recommended RoCEv2 Best Practices	61

Adjust WRED Policy to Recommended Best Practices for ECN	61
Map DSCP 16 to Support RoCE ACK Packets (Only for Intel Gaudi-Based Deployments) ...	61
Chapter 7: Additional Resources	63
7.1 Password Reset	64
7.2 System RNIC Configuration Workflow Example	67
7.3 Contacting Supermicro	69

Chapter 1:

Introduction

This chapter provides a brief outline of the functions and features of Enterprise Advanced SONiC with Supermicro switches.

1.1 Overview of Supermicro Enterprise Advanced SONiC	7
1.2 Applicable Use Cases	8
Layer 3 Clos	8
VXLAN EVPN Overlay	8
AI/ML Fabrics (Backend GPU Cluster)	9
1.3 Supermicro-Validated Designs	11

1.1 Overview of Supermicro Enterprise Advanced SONiC

Supermicro Enterprise Advanced SONiC is a commercial-grade distribution of the open-source SONiC (Software for Open Networking in the Cloud) network operating system. Designed for high-performance data center deployments, it offers enhanced reliability, security, and operational features over community SONiC. Built on a modular, container-based Linux foundation and optimized for merchant silicon, Enterprise Advanced SONiC enables scalable leaf/spine and super-spine topologies.

More information on SONiC's open-source architecture can be found on the following page:

<https://github.com/Azure/SONiC/wiki/Architecture>

1.2 Applicable Use Cases

Layer 3 Clos

Enterprise Advanced SONiC is designed for modern Layer 3 data center fabric architectures, including public, private, and edge computing deployments. It supports scalable Clos topologies using Supermicro's SSE series in roles such as Top-of-Rack (ToR), Leaf, Spine, Super Spine, and Border Leaf. This makes it an excellent choice for scalable underlay networks in web-scale and enterprise data centers, supporting a wide range of workloads including Big Data platforms like Hadoop and virtualized environments such as VMware and Nutanix.

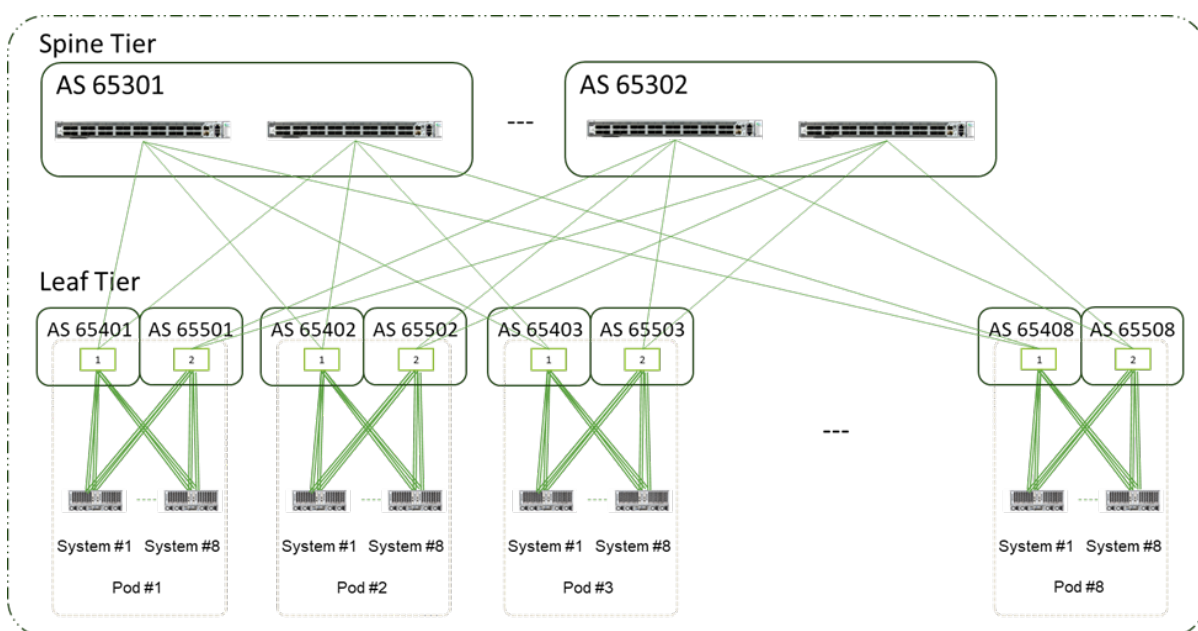


Figure 1-1. Layer 3 Clos

VXLAN EVPN Overlay

For use cases requiring multi-tenancy and endpoint mobility, Enterprise Advanced SONiC supports an overlay network using VXLAN with BGP EVPN as the control plane. This architecture enables logical Layer 2 and Layer 3 topologies that are decoupled from the physical network. This allows administrators to provision tenants programmatically while enabling seamless Layer 2/Layer 3 extension across racks or sites. VXLAN encapsulation forms the data plane, while BGP EVPN facilitates distributed MAC and IP learning and control.

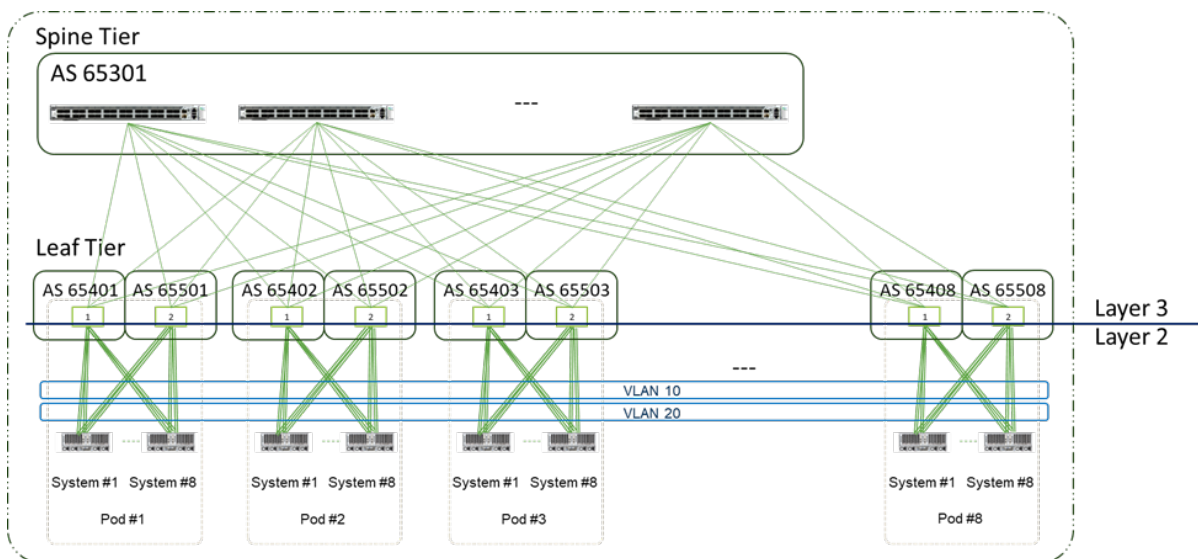


Figure 1-2. VXLAN EVPN Overlay

AI/ML Fabrics (Backend GPU Cluster)

Enterprise Advanced SONiC is purpose-built to address the demanding requirements of high-performance AI/ML infrastructure, particularly in environments with backend GPU clusters. These fabrics demand extreme bandwidth, low-latency transport, and lossless networking capabilities to support parallelized GPU workloads that are highly sensitive to jitter and congestion.

The Supermicro SSE-T8032, SSE-T8164, and SSE-T8196 platforms are specifically optimized for such deployments, supporting advanced features required in AI/ML data pipelines. Enterprise Advanced SONiC leverages these platforms with enhancements such as Adaptive Routing using Dynamic Load Balancing (DLB) to ensure optimal traffic distribution and congestion avoidance across multiple paths. RoCEv2 (RDMA over Converged Ethernet v2) is natively supported with Type of Service (ToS) marking over VXLAN tunnels to preserve traffic class consistency and latency control even in encapsulated environments.

Cut-through switching is enabled to minimize forwarding delay and accelerate inter-GPU communication. Enhanced Equal Cost Multipath (ECMP) with selectable hashing algorithms (including support for UDF-based and QPN-aware hashing) allows for deterministic load balancing across large-scale AI fabrics. Additionally, PFC lossless queues, QoS scheduler policies, and WRED/ECN configurations ensure fine-grained traffic shaping and congestion control tailored to GPU workload characteristics.

Together, these features create a robust, reliable, and high-throughput foundation for AI/ML applications that require tightly coupled GPU interconnection across a distributed fabric.

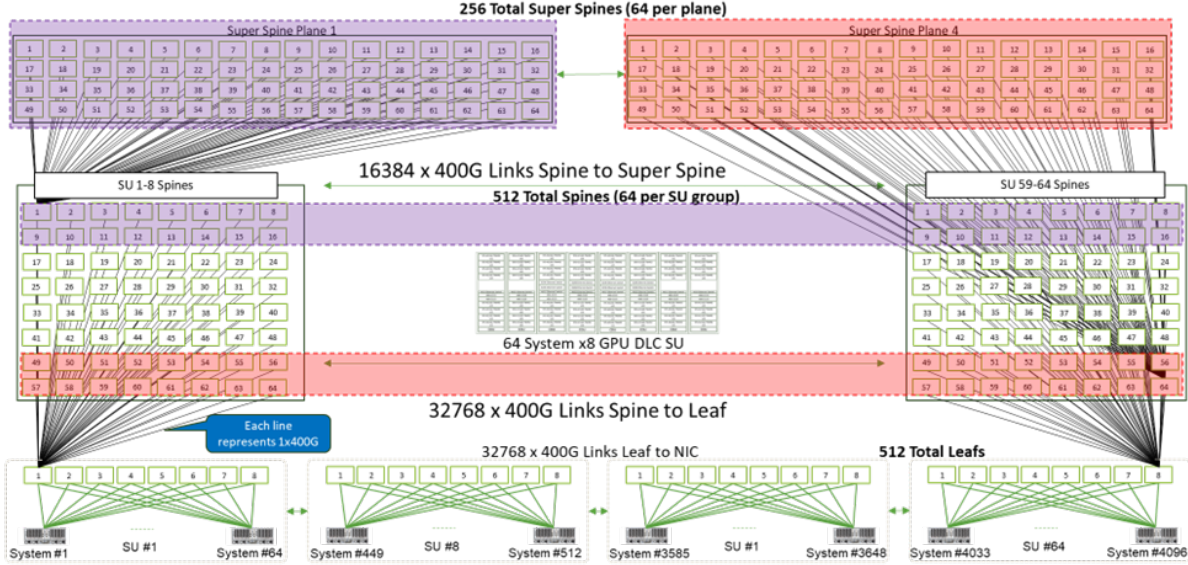


Figure 1-3. AI/ML Fabrics

1.3 Supermicro-Validated Designs

The following Enterprise Advanced SONiC designs have been validated by Supermicro:

- Building a tenant-aware AMD MI325X Accelerator Cluster with an integrated Supermicro solution: <https://www.supermicro.com/solutions/validated-design/AMD-Instinct-MI325X-Accelerator-Cluster.pdf>
- Building a tenant-aware AMD Instinct™ MI325X GPU Cluster with an integrated Supermicro solution: <https://www.supermicro.com/solutions/validated-design/AMD-Instinct-MI325X-Pensando-Pollara-GPU-Cluster.pdf>

Chapter 2:

Hardware Setup

For Supermicro switch rack mounting, interface connections, and more, refer to the hardware installation guide for your selected platform below:

- SSE-C4632S/SR: <https://www.supermicro.com/manuals/network/sse-c4632s-sr-users-guide-pdf.pdf>
- SSE-C4632SB/SRB: <https://www.supermicro.com/manuals/network/sse-c4632sb-srb-users-guide-pdf.pdf>
- SSE-T8032S: https://www.supermicro.com/manuals/network/SSE-T8032S_Users_Guide.pdf
- SSE-T8164: <https://www.supermicro.com/manuals/network/SSE-T8164.pdf>
- SSE-T8196: <https://www.supermicro.com/manuals/network/MNL-2828.pdf>

Chapter 3:

Login and Configuration

This chapter describes the initial access, login, and configuration for Supermicro Enterprise Advanced SONiC.

3.1 Initial Access and Login	14
3.2 Management Interface and Hostname Configuration	17
3.3 User Account Configuration	19
3.4 NTP Configuration	20
3.5 DNS Configuration	22

3.1 Initial Access and Login

You can log into the Supermicro Enterprise Advanced SONiC device using a serial console-based login or an SSH-based login. For SSH-based login, you can log in to the management interface (Management0) IP address.

Notes:

- Enterprise Advanced SONiC uses SSHv2 as its connection protocol for secure connectivity over a network. In this document, all mentions of SSH refer specifically to SSHv2.
 - For console-based connections, an RJ45 to DB9 serial cable is provided with the switch hardware. If you do not have a DB9 port on your client computer, you must use an RJ45 to USB serial cable (not included) to provide console-based connectivity for initial access.
1. Ensure the SSH client's connection type is set to Serial. The serial settings are 115200 baud, 8 data bits, and no parity.
 2. Log in to the Linux shell from the console or SSH connection using the default username `admin` and password `YourPaSsWoRd`.
 3. In the console, enter the following. Next to `admin@sonic's password`, enter your previously configured admin password.

```
Debian GNU/Linux 11 sonic ttyS1
```

```
SONiC login: admin
```

```
Password: YourPaSsWoRd
```

SSH from any remote server to SONiC can be done by connecting to the IP address of the Management interface.

```
user@supermicro-host: ~$ ssh admin@sonic_ip_address(or SONIC DNS Name)
```

```
admin@sonic's password:
```

4. If you are logging in for the first time, follow the information below to change your password. Otherwise, move on to Step 5.

```
You are required to change your password immediately
(administrator-enforced)
```

```
Changing password for admin.
```

```
(current) UNIX password: YourPaSsWoRd
Enter new UNIX password: *****
Retype new UNIX password: *****
Linux sonic 5.10.0-21-amd64 #1 SMP Debian 5.10.162-1 (2023-01-
21) x86_64
You are on
SONiC
-- Software for Open Networking in the Cloud --
Unauthorized access and/or use are prohibited.
All access and/or use are subject to monitoring.
Help https://azure.github.io/SONiC/
admin@sonic:~$
```

5. If ZTP is enabled but you are not using it, it can be disabled via the following command to prevent repeated log messages from being output to the console screen:

```
admin@sonic:~$ config ztp disable -y
```

6. Confirm via the following command that the switch is ready and all processes have finished starting up:

```
admin@sonic:~$ show system status
```

7. At the default Click CLI prompt, enter `sonic-cli` to access the Management Framework CLI in EXEC mode. You can run `show` commands in EXEC mode to monitor various functions on the switch. You can also use `debug` commands to troubleshoot switch operation. Use `?` to display the available commands.

```
admin@sonic:~$ sonic-cli
Leaf1#
alarm                Alarm related commands
clear                Clear commands
configure            Enter configuration mode
consistency-check    Performs consistency check
copy                Perform file operations
crypto              Configure PKI
debug                Enter debugsh mode
delete              Delete the file from local filesystem
dir                 Show folder contents
```

<code>exit</code>	Exit from the CLI
<code><Snipped Output></code>	
<code>traceroute6</code>	Print the route packets take to the IPv6 host
<code>usb</code>	Mount or un-mount usb partitions
<code>write</code>	Save config

3.2 Management Interface and Hostname Configuration

If you use the `gwaddr` option to configure the management interface, all traffic originating from the management IP address on the switch and all locally switched traffic on the management interface is forwarded to the device with the configured `gwaddr` IP address.

If this is not the intended behavior, configure the management VRF (via `ip vrf mgmt`) when assigning the `gwaddr` IP address to the management interface. The recommendation is to configure the management VRF so that the management interface is automatically added and that management traffic is fully isolated from other configured VRFs.

1. Configure an Out-of-Band management interface by entering its IP address, subnet mask, and default gateway that can then be used to manage the switch via SSH.

```
sonic(config)# interface Management0
sonic(conf-if-Management0)# description Management0
sonic(conf-if-Management0)# ip address 100.94.189.13/24 gwaddr
100.94.189.254
sonic(conf-if-Management0)# ipv6 address 2000::13/64
sonic(conf-if-Management0)#
```

2. Verify the management interface configuration.

```
sonic# show interface Management0
Management0 is up, line protocol is up
Hardware is MGMT, address is 0c:29:ef:e1:f8:00
Description: Management0
IPV4 address is 100.94.189.13/24
Mode of IPV4 address assignment: MANUAL
IPV6 address is 2000::13/64,fe80::e29:efff:fee1:f800/64
Mode of IPV6 address assignment: DHCP
IP MTU 1500 bytes
LineSpeed 1GB, Auto-negotiation True
Input statistics:
75106 packets, 15083074 octets
28980 Multicasts, 0 error, 1553 discarded
Output statistics:
```

```
3291 packets, 765626 octets
```

```
0 error, 0 discarded
```

```
Time since last interface status change: 12:56:27
```

3. If you plan to place the Management0 interface in the management VRF, you must create the management (mgmt) VRF. By default, the Management0 interface is in the default VRF. However, after the management VRF is created, the Management0 interface is automatically moved from the default VRF to this management VRF.

```
sonic(config)# do show ip interfaces
Flags: U-Unnumbered interface, A-Anycast IP
- - -          - - -          - - -          - - -          - - -
Interface      IP address/mask      VRF      Admin/      Flags
              IP address/mask      VRF      Oper
- - -          - - -          - - -          - - -          - - -
Management0    100.94.189.13/24      up/down
sonic(config)# exit
sonic(config)# ip vrf mgmt
sonic(config)# do show ip interfaces
Flags: U-Unnumbered interface, A-Anycast IP
- - -          - - -          - - -          - - -          - - -
Interface      IP address/mask      VRF      Admin/      Flags
              IP address/mask      VRF      Oper
- - -          - - -          - - -          - - -          - - -
Management0    100.94.189.13/24    mgmt      up/down
```

4. Configure the hostname the switch should be identified with.

```
sonic(config)# hostname <hostname-val>
```

5. After changing the hostname, you must exit to the Linux shell and reenter `sonic-cli` as prompted for the change to take full effect.

3.3 User Account Configuration

While Supermicro Enterprise Advanced SONiC also supports remote authentication using RADIUS, TACACS+, and LDAP, this section focuses on local user account creation.

- When logging in as an administrator, you are placed in the Linux shell using the `admin@Leaf1:~$` prompt. To access the Management Framework CLI, enter the `sonic-cli` command.
- When logging in as an operator, you are placed in the Management Framework CLI using the `Leaf1#` prompt by default.

Note: An operator cannot access the Linux shell.

1. Configure a new user or update an existing user's password.

```
Leaf1(config)# username <user-name> { { [ password <passwd> { [
role <rl> ] } ] } | { [ role <rl> { [ password <passwd> ] } ] }
| { [ sshkey <key> ] } }
```

```
Leaf1(config)# username admin password Supermicro123!
```

2. Verify that the user account is created and associated with the appropriate role.

```
Leaf1# show users all
- - -      - - -
User       Role(s)
- - -      - - -
admin      admin
Leaf1#
```

3.4 NTP Configuration

Accurate time synchronization across network infrastructure is essential for event correlation, security protocols, and operational troubleshooting.

Supernano Enterprise Advanced SONiC supports Network Time Protocol (NTP) to maintain consistent timestamps across all switches in your fabric, ensuring reliable log analysis, certificate validation, and coordinated network operations.

1. Configure an NTP server by entering its IP address or domain name to synchronize time on the switch. To configure multiple NTP servers, use the `ntp server` command.

```
Leaf1(config)# ntp server {ipv4-address | ipv6-address | ntp-
server-name} [key keyid] [prefer true | false] [maxpoll
interval] [minpoll interval]
Leaf1(config)# ntp server 100.94.121.15 key 1
Leaf1(config)# ntp source-interface Management0
Leaf1(config)# ntp authentication-key 1 md5 ntpclient2
Leaf1(config)# ntp trusted-key 1
Leaf1(config)# ntp authenticate
```

2. Verify the NTP configuration.

```
Leaf1# show ntp server
NTP Servers          minpoll    maxpoll    Prefer Authentication
                    key ID
10.89.192.129        6          10         False
134.214.100.6        5          9          True
144.126.242.176     6          10         True
Leaf1# show ntp global
- - -                - - -      - - -      - - -
NTP Global Configuration
- - -                - - -      - - -      - - -
NTP source-interface: Management0
NTP vrf: mgmt
Leaf1# show ntp associations
remote refid st t when poll reach delay offset jitter
===                ===        ===        ===
+10.11.0.1 10.11.8.1 4 u 9 64 1 0.232 -2.570 0.062
```

```
*10.11.0.2 10.11.8.1 4 u 8 64 1 0.262 -0.747 0.033
* master (synced), # master (unsynced), + selected, -
candidate
```

3.5 DNS Configuration

The Domain Name System (DNS) enables switches to resolve hostnames to IP addresses, simplifying network management and configuration.

Supernano Enterprise Advanced SONiC supports DNS client functionality, allowing administrators to reference servers and services by name rather than IP address for operations such as NTP synchronization, syslog forwarding, and remote authentication.

1. Configure the source interface on the switch that is used for the source IP address in DNS queries.

```
Leaf1(config)# ip name-server source-interface {Eth slot/port  
[/breakout-port] | Loopback number | Management0 | PortChannel  
number | Vlan vlan-id}
```

2. Configure a DNS server. Enter an IPv4 address in A.B.C.D format. Enter an IPv6 address in A::B format. Enter a VRF name to specify the VRF to use for reaching the name server.

```
Leaf1(config)# ip name-server {ip-address | ipv6-address} [vrf  
vrf-name]
```

3. To configure additional DNS servers, reenter the `ip name-server` command.
4. Verify the DNS server configuration.

```
Leaf1# show hosts
```

A sample DNS server configuration appears as follows:

```
Leaf1(config)# ip name-server source-interface Loopback0  
Leaf1(config)# ip name-server 1.2.3.4 vrf mgmt  
Leaf1(config)# ip name-server 2001:4860:4860::8888 vrf mgmt  
Leaf1(config)# ip name-server 8.8.8.8  
Leaf1# show hosts  
Source Interface: Loopback0  
Name servers are: 1.2.3.4(vrf: mgmt), 2001:4860:4860::8888(vrf:  
mgmt), 8.8.8.8
```

Chapter 4:

Layer 2 and Layer 3 Enablement

This chapter describes Layer 2 and Layer 3 enablement with Supermicro Enterprise Advanced SONiC.

4.1 Configuring Breakout Ports	24
4.2 Configuring Layer 2 Interfaces and VLAN	28
4.3 Configuring Layer 3 Interfaces	30

4.1 Configuring Breakout Ports

Breakout ports enable a single high-speed interface to be split into multiple lower-speed interfaces, providing flexibility when connecting devices that operate at different speeds.

Supermicro Enterprise Advanced SONiC supports multiple breakout options depending on the platform, allowing administrators to divide 800G or 400G ports into combinations such as 2x400G, 4x200G, 4x100G, or 8x100G depending on platform capabilities and transceiver type.

1. Before configuring the breakout interface, verify which breakout modes are supported on the SSE switch you are configuring.

```
Leaf1# show interface breakout modes
-----
Port      Pipe      Interface  Supported Modes          Default
-----
Mode
-----
1/1       1         Eth1/1     1x800G, 1x400G(8),      1x800G
          1x100G(4), 1x50G,
          2x400G(8), 2x200G(8),
          2x100G(8), 2x50G(4),
          4x200G, 4x100G(8),
          8x100G, 8x50G
1/2       2         Eth1/2     1x800G, 1x400G(8),      1x800G
          1x100G(4), 1x50G,
          2x400G(8), 2x200G(8),
          2x100G(8), 2x50G(4),
          4x200G, 4x100G(8),
          8x100G, 8x50G
1/3       3         Eth1/3     1x800G, 1x400G(8),      1x800G
          1x100G(4), 1x50G,
          2x400G(8), 2x200G(8),
          2x100G(8), 2x50G(4),
          4x200G, 4x100G(8),
          8x100G, 8x50G
```

```

1/4      4      Eth1/4      1x800G, 1x400G(8),      1x800G
                               1x100G(4), 1x50G,
                               2x400G(8), 2x200G(8),
                               2x100G(8), 2x50G(4),
                               4x200G, 4x100G(8),
                               8x100G, 8x50G

1/5      5      Eth1/5      1x800G, 1x400G(8),      1x800G
                               1x100G(4), 1x50G,
                               2x400G(8), 2x200G(8),
                               2x100G(8), 2x50G(4),
                               4x200G, 4x100G(8),
                               8x100G, 8x50G

```

<Snipped>

Leaf1#

2. Before configuring the breakout interface, verify if any existing interface configuration will be removed.

```
Leaf1# show interface breakout dependencies port 1/5
```

```
-----
```

```
Dependent Configurations
```

```
-----
```

```
QUEUE|Ethernet32|4
```

```
QUEUE|Ethernet32|0
```

```
QUEUE|Ethernet32|6
```

```
QUEUE|Ethernet32|3
```

```
PORT_QOS_MAP|Ethernet32
```

<Snipped>

3. Before configuring the breakout interface, verify that the switch has resources available to support this new breakout cable.

```
Leaf1# show interface breakout resources
```

```
Maximum ports supported in the system: 320
```

```
Current ports in the system: 66
```

```
- - - - -
```

```
Pipeline      Ports      Max-Ports  Front-panel-ports
```

```
- - - - -
```

```

1           3           10           1/1, 1/34
2           2           10           1/2, 1/33
3           3           10           1/3, 1/36
4           2           10           1/4, 1/35
5           2           10           1/5, 1/38

```

<Snipped>

4. Configure the breakout port using the command below within configuration mode.

```

Leaf1(config)# interface breakout port (slot/port) mode {1x10G |
1x25G | 1x40G | 1x50G | 1x100G | 1x200G | 1x400G | 1x800G |
2x10G | 2x25G | 2x40G | 2x50G | 2x100G | 2x200G | 2x400G | 4x10G |
4x25G | 4x50G | 4x100G | 4x200G | 2x200G | 8x10G | 8x25G | 8x50G
| 8x100G}

```

- **Example:** Leaf1(config)# interface breakout port 1/1 mode 2x400G

5. If the breakout port configuration is not supported, you will receive an error similar to the following:

```

Leaf1(config)# interface breakout port 1/1 mode 4x50G
%Error: Invalid or unsupported breakout mode 4x50G

```

6. To revert the configuration of the breakout port back to default, use the following command within configuration mode.

```
Leaf1(config)# no interface breakout port slot/port
```

A sample breakout port configuration appears as follows:

```
Leaf1(config)# no interface breakout port 1/1
```

7. Verify that the breakout configuration was completed as expected.

```

Leaf1(config)# show interface breakout
- - -      - - -      - - -      - - -
Port        Breakout Mode  Status      Interfaces
- - -      - - -      - - -      - - -
1/1         2x400G        Completed   Eth1/1/1
                                           Eth1/1/2

Leaf1(config)#

```

8. After the breakout is successfully completed, the ports will be administratively shut down and ready for configuration with any Layer 3 IP addressing, Layer 2 VLANs, and/or QoS configurations necessary.

```
Leaf1# show interface status
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Name      Desc.   Oper   Reason   Auto   Speed   MTU   Alternate
          - - -   - -   - - -   Neg   - - -   - -   Name
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Eth1/    -      down  no-tran  off   400000  9100  Ethernet0
1/1      -              sceiver
Eth1/    -      down  no-tran  off   400000  9100  Ethernet4
1/2      -              sceiver
```

4.2 Configuring Layer 2 Interfaces and VLAN

VLANs segment network traffic into isolated broadcast domains, enabling logical separation of workloads, tenants, or traffic types across shared physical infrastructure.

An access port sends and receives untagged frames from connected devices. By default, no Layer 2 VLAN or Layer 3 IP is assigned to a port interface.

Supernano Enterprise Advanced SONiC supports standard 802.1Q VLAN tagging, allowing interfaces to operate as access ports for single-VLAN connectivity or trunk ports for carrying multiple VLANs between switches and endpoints.

- To assign an interface to an access VLAN, use the `switchport access Vlan vlan-id` command.

```
Leaf1(conf-if-Eth1/1)# switchport access Vlan [vlan-id]
```

- To remove an interface from the access VLAN, use the `no switchport access Vlan` command.

1. Configure an interface as an access port by associating the untagged VLAN with the interface.

```
Leaf1(config)# interface Eth1/1
Leaf1(conf-if-Eth1/1)# switchport access Vlan 100
Leaf1(conf-if-Eth1/1)# end
```

2. Verify that the interface is now associated with the appropriate VLAN.

```
Leaf1# show Vlan
Q: A - Access (Untagged), T - Tagged
NUM      Status      Q      Ports      Autostate      Dynamic
100      Active      A      Eth1/1      Enable          No
```

- A trunk port sends and receives tagged frames from multiple VLANs and untagged frames from the access VLAN. By default, no trunk VLANs are assigned to a port interface. You can remove VLAN IDs from the allowed VLANs, and later add VLAN IDs to the list to allow their traffic to pass on the trunk.
- To configure the allowed VLAN traffic on a trunk interface, use the `switchport trunk allowed Vlan` command.

```
Leaf1(conf-if-Eth1/3)# switchport trunk allowed Vlan {vlan-
list | {add | remove | except} vlan-list | none | all}
```

3. Configure an interface as a trunk port by associating the native as well as tagged VLANs on the interface.

```
Leaf1(config)# interface Eth1/3
Leaf1(conf-if-Eth1/3)# switchport trunk allowed Vlan add 100
Leaf1(conf-if-Eth1/3)# end
```

4. Verify that the interface is now associated with the correct tagged VLANs and untagged/native VLAN.

```
Leaf1# show Vlan
Q: A - Access (Untagged), T - Tagged
NUM      Status      Q      Ports      Autostate      Dynamic
100      Active      T      Eth1/3      Enable         No
```

4.3 Configuring Layer 3 Interfaces

Layer 3 interfaces enable IP routing on switch ports, allowing direct routing between network segments without requiring an external router.

Supernano Enterprise Advanced SONiC supports routed interfaces on physical ports, VLAN interfaces (SVIs) for inter-VLAN routing, and loopback interfaces for stable router identifiers and VTEP source addresses in overlay deployments.

1. Configure a Layer 3 interface by entering its IP address and the subnet mask associated with the switch.

```
Leaf1(config)# interface Eth1/3

! similarly you can change Eth interface to vlan [1-4094] to
configure VLAN instead Ethernet interface

Leaf1(conf-if-Ethernet1/3)#

Leaf1(conf-if-Ethernet1/3)# description "Test Config"
Leaf1(conf-if-Ethernet1/3)# ip address 10.210.9.3/31
Leaf1(conf-if-Ethernet1/3)# ipv6 address 10:21:10:9::3/64
Leaf1(conf-if-Ethernet1/3)# ip access-group ipv4_ingress_acl in
Leaf1(conf-if-Ethernet1/3)# ipv6 access-group ipv6_ingress_acl
in

Leaf1(conf-if-Ethernet1/3)# mtu 9100

Leaf1(conf-if-Ethernet1/3)# speed 10000

Leaf1(conf-if-Ethernet68)# no shutdown

Leaf1(conf-if-Ethernet68)#
```

2. Verify that the interface is up and operational, as well as whether it is configured correctly and transmitting/receiving traffic.

```
Leaf1(config)# do show interface Eth1/3

Eth1/3 is up, line protocol is up, reason oper-up
Hardware is Eth, address is 0c:29:ef:e1:f8:02
Description: Test Config
IPV4 address is 10.210.9.3/31
Mode of IPV4 address assignment: MANUAL
IPV6 address is 10:21:10:9::3/64, fe80::e29:efff:fee1:f802/64
```

```
Mode of IPV6 address assignment: MANUAL
IP MTU 9100 bytes
LineSpeed 10GB, Auto-negotiation off Link-training: off
Unreliable-LOS: off
FEC: DISABLED
Events:
initialized at 2025-06-20T05:46:43.421001Z
admin-up at 2025-06-20T05:46:53.361035Z
xcvr-status-up at 2025-06-20T05:48:07.486835Z
port-enabled at 2025-06-20T05:48:07.487593Z
phy-link-up at 2025-06-20T05:48:07.729433Z
Last clearing of "show interface" counters: never
10 seconds input rate 70000 packets/sec, 769602688 bits/sec,
96200336 Bytes/sec
10 seconds output rate 19999 packets/sec, 242878752 bits/sec,
30359844 Bytes/sec
Input statistics:
2606931547 packets, 3582259145792 octets
20 Multicasts, 466699 Broadcasts, 2606464835 Unicasts
0 error, 209297229 discarded, 0 Oversize
6 Packets (128 to 255 Octects)
Output statistics:
682419292 packets, 1035901416402 octets
1877 Multicasts, 1 Broadcasts, 682417414 Unicasts
```

3. If the interface is up and operational, and an IP is configured on the endpoint in the subnet, you can also verify whether you can reach the IP configured on the switch from the endpoint or vice versa.

Chapter 5:

Use Case: VXLAN/EVPN Fabric

Virtual Extensible LAN (VXLAN) enables Layer 2 connectivity across Layer 3 network boundaries, allowing servers in different physical locations to communicate as if they were on the same local network. This capability is essential for modern data centers that require flexible workload placement and multi-tenant isolation.

Supernano Enterprise Advanced SONiC implements VXLAN as defined in RFC 7348, with support for BGP EVPN as the control plane for scalable, automated tunnel management.

5.1 Loopback and VTEP Configuration	33
5.2 BGP EVPN Configuration	34
5.3 VXLAN VTEP VNI Mapping	37
5.4 Sample VXLAN EVPN Configuration	38
Topology	38
VXLAN EVPN Verification	46

5.1 Loopback and VTEP Configuration

In a data center, the same static anycast MAC address is used in all pods. This MAC address acts as the gateway MAC that is set as the source MAC address in routed packets in VLANs, which have a configured static anycast address. If the network fabric uses filters to drop packets which have the static anycast MAC address as the source MAC address, configure the router HW MAC address to be set as the source MAC address in routed packets:

```
Leaf1(config)# [no] ip anycast-mac-address router-mac-for-forwarding
```

1. Configure a Loopback interface (0 to 16383) which will be used as the VTEP source IP address. The interface is automatically enabled and up.

```
Leaf1(config)# interface Loopback [1-16383]
Leaf1(conf-if-lo1)# ip address [ip-address/mask]
Leaf1(conf-if-lo1)# exit
```

2. Configure a source IP address using the Loopback IP address from Step 1. This VTEP source IP address is used in VXLAN-encapsulated packet headers. By default, no source IP interface is configured.
 - The VXLAN interface name is a text string up to ten characters in length. Use a VTEP name in `vtep###` format, such as `vtep1`.
 - Enter the source IP address in dotted decimal A.B.C.D format. An IPv6 address is not supported as the source VXLAN address. Only one source IP address is supported on a VTEP. On a standalone VTEP, the source IP address is used both as the source VXLAN tunnel endpoint in the overlay, and as the IP source address for routing in the underlay.

```
Leaf1(config)# interface vxlan vtep[name]
Leaf1(config-if-vxlan-vtep1)# source-ip [ipv4-address |
Loopback-interface]
Leaf1(config-if-vxlan-vtep1)# exit
```

3. Configure the anycast MAC address of the gateway used on all VTEPs to forward VXLAN traffic in the overlay. Configure the same anycast MAC address on each VTEP in the VXLAN network.

```
Leaf1(config)# ip anycast-address enable
Leaf1(config)# ip anycast-mac-address [mac-address]
Leaf1(config)# ipv6 anycast-address enable
```

5.2 BGP EVPN Configuration

BGP EVPN provides the control plane for VXLAN deployments, automating the distribution of MAC addresses, IP addresses, and VNI-to-VTEP mappings across the fabric. This eliminates the need for flood-and-learn behavior and enables efficient, scalable overlay networks with support for multi-tenancy and workload mobility.

Supermicro Enterprise Advanced SONiC can optionally implement BGP EVPN with IPv6 unnumbered peering for simplified leaf-spine fabric addressing.

If you are using BGP unnumbered, the command `capability-extended-nexthop` is required under the peer-group to support RFC5549.

1. Enable IPv6 on all fabric-facing interfaces between leaf and spine to support BGP unnumbered.

```
Leaf1(config)# interface Eth{slot/port}[/breakout-port]
Leaf1(conf-if-Eth1) ipv6 enable
Leaf1(conf-if-Eth1) no shutdown
```

2. Configure a Loopback interface (0 to 16383) which will be used as the BGP Router-ID. The interface is automatically enabled and up.

```
Leaf1(config)# interface Loopback 0
Leaf1(conf-if-lo0)# ip address [ip-address/mask]
Leaf1(conf-if-lo0)# exit
```

3. Configure the BGP IPv4 Unicast and L2VPN EVPN adjacencies.

```
Leaf1(config)# router bgp [1- 4294967295]
Leaf1(config-router-bgp)# router-id [router-id-ip]
Leaf1(config-router-bgp)# address-family ipv4 unicast
Leaf1(config-router-bgp-af)# redistribute connected
Leaf1(config-router-bgp-af)# maximum-paths [1-16]
Leaf1(config-router-bgp-af)# exit
!
Leaf1(config-router-bgp)# address-family l2vpn evpn
Leaf1(config-router-bgp-af)# advertise-all-vni
Leaf1(config-router-bgp-af)# exit
!
```

```
Leaf1(config-router-bgp)# peer-group SPINE
Leaf1(config-router-bgp-pg)# remote-as external
Leaf1(config-router-bgp-pg)# capability extended-nexthop
!
Leaf1(config-router-bgp-pg)# address-family ipv4 unicast
Leaf1(config-router-bgp-pg-af)# activate
Leaf1(config-router-bgp-pg-af)# exit
!
Leaf1(config-router-bgp-pg)# address-family l2vpn evpn
Leaf1(config-router-bgp-pg-af)# activate
Leaf1(config-router-bgp-pg-af)# exit
Leaf1(config-router-bgp-pg)# exit
!
Leaf1(config-router-bgp)# neighbor interface Eth{slot/port}
[/breakout-port]
Leaf1(config-router-bgp-neighbor)# peer-group [name]
Leaf1(config-router-bgp-neighbor)# exit
!
Leaf1(config-router-bgp)# neighbor interface Eth{slot/port}
[/breakout-port]
Leaf1(config-router-bgp-neighbor)# peer-group [name]
Leaf1(config-router-bgp-neighbor)# exit
Leaf1(config-router-bgp)# exit
Leaf1(config-router)# exit
!
Leaf1(config)# router bgp [1- 4294967295] vrf Vrf[name]
!
Leaf1(config-router-bgp)# address-family ipv4 unicast
Leaf1(config-router-bgp-af)# redistribute connected
Leaf1(config-router-bgp-af)# maximum-paths [1-16]
Leaf1(config-router-bgp-af)# exit
```

```
!  
Leaf1(config-router-bgp)# address-family l2vpn evpn  
Leaf1(config-router-bgp-af)# advertise ipv4 unicast
```

5.3 VXLAN VTEP VNI Mapping

VNI mapping associates VXLAN Network Identifiers with VLANs and VRFs, defining how traffic is encapsulated and forwarded across the overlay network. This configuration links local Layer 2 segments or Layer 3 routing instances to their corresponding VXLAN tunnels, enabling seamless communication between distributed endpoints.

As of Enterprise Advanced SONiC releases 4.4.3 and 4.5.0, the SSE-T8164, SSE-T8196, and SSE-T8032 switches support only Layer 3 VXLAN (VRF extensibility). The SSE-C4632 and SSE-F3648 switches support both Layer 2 and Layer 3 VXLAN for deployments requiring Layer 2 extension across sites.

1. Create the Layer 3 VLAN to be used for the Layer 3 VXLAN VNID.

```
Leaf1(config)# interface Vlan [1-4094]
Leaf1(config-if-Vlan-99)# ip vrf forwarding Vrf[name]
```

2. Under the existing VXLAN VTEP, map the VRF VLAN and VRF itself to the same VNI.

```
Leaf1(config)# interface vxlan vtep[name]
Leaf1(config-if-vxlan-vtep1)# map vni [1-16777215] vlan [1-4094]
Leaf1(config-if-vxlan-vtep1)# map vni [1-16777215] vrf Vrf[name]
```

5.4 Sample VXLAN EVPN Configuration

The following are sample VXLAN EVPN configurations for both the leaf and spine roles with Supermicro Enterprise Advanced SONiC, using Leaf1, Leaf2, Spine1, and Spine2 in the topology below. After configuration is complete, verify that the VXLAN tunnel is up.

Topology

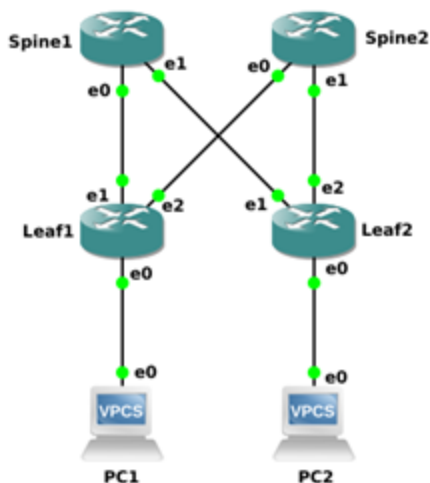


Figure 5-1. Topology Diagram

- **Leaf1 Configuration**

```
ip anycast-mac-address 00:00:de:ad:be:ef
ip anycast-address enable
!
ip vrf VrfCust-A
!
interface Lo0
description Router-ID
ip address 1.1.1.1/32
!
interface Lo1
description VTEP
ip address 10.255.0.1/32
```

```
!  
interface Vlan 10  
ip vrf forwarding VrfCust-A  
neigh-suppress  
ip anycast-address 10.0.0.1/24  
!  
interface Vlan 99  
ip vrf forwarding VrfCust-A  
!  
interface vxlan vtep1  
source-ip 10.255.0.1  
map vni 99 vlan 99  
map vni 99 vrf VrfCust-A  
!  
interface Eth0  
switchport access Vlan 10  
no shutdown  
exit  
!  
interface Eth1  
description Spine1  
ipv6 enable  
no shutdown  
!  
interface Eth2  
description Spine1  
ipv6 enable  
no shutdown  
!  
router bgp 65001  
router-id 1.1.1.1
```

```
address-family ipv4 unicast
redistribute connected
maximum-paths 2
!
address-family l2vpn evpn
advertise-all-vni
!
peer-group SPINE
remote-as external
capability extended-next-hop
!
address-family ipv4 unicast
activate
!
address-family l2vpn evpn
activate
neighbor interface Ethernet1
description Spine1
peer-group SPINE
!
neighbor interface Ethernet2
description Spine2
peer-group SPINE
!
router bgp 65001 vrf VrfCust-A
!
address-family ipv4 unicast
redistribute connected
maximum-paths 2
!
address-family l2vpn evpn
```

```
advertise ipv4 unicast
```

- **Leaf2 Configuration**

```
ip anycast-mac-address 00:00:de:ad:be:ef
```

```
ip anycast-address enable
```

```
!
```

```
ip vrf VrfCust-A
```

```
!
```

```
interface Lo0
```

```
description Router-ID
```

```
ip address 2.2.2.2/32
```

```
!
```

```
interface Lol
```

```
description VTEP
```

```
ip address 10.255.0.2/32
```

```
!
```

```
interface Vlan 20
```

```
ip vrf forwarding VrfCust-A
```

```
neigh-suppress
```

```
ip anycast-address 20.0.0.1/24
```

```
!
```

```
interface Vlan 99
```

```
ip vrf forwarding VrfCust-A
```

```
!
```

```
interface vxlan vtep1
```

```
source-ip 10.255.0.2
```

```
map vni 99 vlan 99
```

```
map vni 99 vrf VrfCust-A
```

```
!
```

```
interface Eth0
```

```
switchport access Vlan 20
```

```
no shutdown
```

```
exit
!
interface Eth1
description Spine1
ipv6 enable
no shutdown
!
interface Eth2
description Spine1
ipv6 enable
no shutdown
!
router bgp 65002
router-id 2.2.2.2
address-family ipv4 unicast
redistribute connected
maximum-paths 2
!
address-family l2vpn evpn
advertise-all-vni
!
peer-group SPINE
remote-as external
capability extended-nexthop
!
address-family ipv4 unicast
activate
!
address-family l2vpn evpn
activate
neighbor interface Ethernet1
```

```
description Spine1
peer-group SPINE
!
neighbor interface Ethernet2
description Spine2
peer-group SPINE
!
router bgp 65002 vrf VrfCust-A
!
address-family ipv4 unicast
redistribute connected
maximum-paths 2
!
address-family l2vpn evpn
advertise ipv4 unicast
```

- **Spine1 Configuration**

```
interface Loopback 0
description Router-ID
ip address 3.3.3.3/32
!
interface Ethernet0
description Leaf1
ipv6 enable
no shutdown
!
interface Ethernet1
description Leaf2
ipv6 enable
no shutdown
!
router bgp 65000
```

```
router-id 3.3.3.3
!
address-family ipv4 unicast
redistribute connected
maximum-paths 2

peer-group LEAF
remote-as external
capability extended-nextthop
!
address-family ipv4 unicast
activate
!
address-family l2vpn evpn
activate
!
neighbor interface Ethernet0
description Leaf1
peer-group LEAF
!
neighbor interface Ethernet1
description Leaf2
peer-group LEAF
```

- **Spine2 Configuration**

```
interface Loopback 0
description Router-ID
ip address 4.4.4.4/32
!
interface Ethernet0
description Leaf1
ipv6 enable
```

```
no shutdown
!
interface Ethernet1
description Leaf2
ipv6 enable
no shutdown
!
router bgp 65000
router-id 4.4.4.4
!
address-family ipv4 unicast
redistribute connected
maximum-paths 2

peer-group LEAF
remote-as external
capability extended-nexthop
!
address-family ipv4 unicast
activate
!
address-family l2vpn evpn
activate
!
neighbor interface Ethernet0
description Leaf1
peer-group LEAF
!
neighbor interface Ethernet1
description Leaf2
peer-group LEAF
```

VXLAN EVPN Verification

1. Validate that the VXLAN tunnel is up between leaves.

```
Leaf1# show vxlan tunnel
Name          SIP          DIP          source      Group        D-VNI  oper
              = = =      = = =      = = =      = = =      = = =  status
              = = =      = = =      = = =      = = =      = = =  = = =
EVPN_        10.         10.         EVPN       internal    no     oper_up
10.255       255.        255.
.0.2         0.1         0.2
```

2. Validate that the VXLAN to VRF and VXLAN to VLAN mapping is configured as expected.

```
Leaf1# show vxlan vrfvni
VRF          VNI
= = =      = = =
VrfCust-A    99
Total count: 1
Leaf1#
```

```
Leaf1# show vxlan vlanvni
VLAN        VNI
= = =      = = =
Vlan99      99
Total count: 1
```



```
10.255 0 65000 RT:65002:99
.0.2 65002 ET:8
? Rmac:0c:a1:
26:e5:00:0a

*> [5]:[0]:[24]:[20.0.0.0]
10.255 0 65000 RT:65002:99
.0.2 65002 ET:8
? Rmac:0c:a1:
26:e5:00:0a
```

Displayed 2 prefixes (3 paths)

Chapter 6:

Use Case: AI Fabric

Supermicro Enterprise Advanced SONiC delivers the high-bandwidth, low-latency, and lossless networking capabilities required for AI/ML infrastructure, including Adaptive Routing and RoCE QoS configuration. The SSE-T8164, SSE-T8196, and SSE-T8032 platforms provide purpose-built foundations for backend GPU clusters, where parallelized workloads demand predictable performance and zero packet loss.

6.1 Adaptive Routing Configuration	50
Sample Adaptive Routing Configuration	51
6.2 Adaptive Routing Verification	52
6.3 RoCEv2 Configuration and Verification	55
6.4 Recommended RoCEv2 Best Practices	61
Adjust WRED Policy to Recommended Best Practices for ECN	61
Map DSCP 16 to Support RoCE ACK Packets (Only for Intel Gaudi-Based Deployments)	61

6.1 Adaptive Routing Configuration

Adaptive Routing and Switching (ARS) dynamically distributes traffic across equal-cost paths based on real-time port utilization metrics, including past, current, and projected load. Unlike static ECMP hashing, ARS continuously monitors congestion and redirects flows to less-utilized paths, preventing hotspots and maximizing fabric throughput. This capability is essential for AI/ML workloads where GPU-to-GPU traffic patterns create unpredictable bursts that can overwhelm individual links.

1. Configure the switch-level binding.

```
Leaf1(config)# ars profile <profile-name>
Leaf1(config-ars-profile-default)# exit
!
Leaf1(config)# ars bind profile <profile-name>
!
Leaf1(config)# ars profile <profile-name>
Leaf1(config-ars-profile-default)# port-load-past
Leaf1(config-ars-profile-default)# port-load-future
Leaf1(config-ars-profile-default)# port-load-current
```

2. Configure the port-level binding.

```
Leaf1(config)# ars port-profile <port-profile-name>
Leaf1(config-ars-port-profile)# enable
Leaf1(config)# interface Eth{slot/port}[/breakout-port]
Leaf1(config-if)# ars bind <port-profile-name>
```

3. Configure the next-hop-group-level binding.

```
Leaf1(config)# ars object <ars-object-name>
Leaf1(config)# route-map <map-name> permit <sequence-number>
Leaf1(config-route-map)# set ars-object <ars-object-name>
Leaf1(config-route-map)# exit
!
Leaf1(config)# ip protocol any route-map <map-name> [vrf {vrf-name}]
```

```
Leaf1(config)# ipv6 protocol any route-map <map-name> [vrf {vrf-  
name}]
```

Sample Adaptive Routing Configuration

```
ars profile default  
port-load-past  
port-load-future  
port-load-current  
exit  
!  
ars bind profile default  
!  
ars port-profile default  
enable  
exit  
!  
interface range Eth 1-2  
ars bind default  
exit  
!  
ars object default  
exit  
!  
route-map ars-map permit 10  
set ars-object default  
exit  
!  
ip protocol any route-map ars-map  
ipv6 protocol any route-map ars-map
```

6.2 Adaptive Routing Verification

After configuring adaptive routing with Supermicro Enterprise Advanced SONiC, be sure to confirm that the routing functions as intended.

1. Review the ARS switch profile.

```
Leaf1# show ars profile default
Switch Attributes Info:
    Profile Name: default
    ARS algorithm used for quality computation: EWMA
    Sampling interval of data for quality measure
    computation: 16 microsecond(s)
    Random seed: 10
    Past port load as the quality parameter: Enabled
    Past port load weight: 2
    Future port load as the quality parameter: Enabled
    Future port load weight: 2
    Current port load calculation: Disabled
    Exponential port load value: 2
    Minimum past load threshold value: 3000
    Maximum past load threshold value: 6000
    Minimum future load threshold value: 2097152
    Maximum future load threshold value: 12582912
    Minimum current load threshold value: 1048576
    Maximum current load threshold value: 6291456
```

2. Review the ARS port profile.

```
Leaf1# show ars port-profile default
Port Attributes Info:
    Port Profile Name: default
    ARS for a port is: Enabled
    Port load scaling factor: 0
    Past port load quality weight in percent: 80
    Future port load quality weight in percent:
    10
```

3. Review the ARS object details.

```
Leaf1# show ars object default
ARS Object Attributes Info:
    ARS Object Name: default
    Idle duration to classify a flow-let in a macro
    flow: 80 microsecond(s)
    Maximum flow size: 256
    ARS path (re)assignment mode: Per flow-let quality
```

4. Review the ARS next-hop-group statistics.

```
Leaf1#show ars nhg-statistics
ARS NHG Statistics Info:
    NHG ID: 200
    Number of packet drops: 50
    Number of next hop reassignments: 40
    Number of port reassignments: 30
```

5. Verify the ARS binding configuration.

```
Leaf1# show ars bind
ARS Profile binding Info (Switch Name : Profile Name):
    SWITCH: default
Port Profile binding Info (Port Profile Name : Interfaces):
    default : Eth1, Eth2
ARS Object binding Info (Routemap, Seq.No : Object Name):
    ars-map, 1 : default
```

6. Verify the routing table and next-hop-group details.

```
Leaf1# show ip route nexthop-group
Codes: K - kernel route, C - connected, S - static, B - BGP, O
- OSPF
> - selected route, * - FIB route, q - queued route, r -
rejected route
Destination      Gateway          NHG      Dist/      Last
                  via              ID        Metric     Update
- - -            - - -           - - -    - - -     - - -
B>*              via              Ethernet1 68         20/0       00:00:05
20.0.0.0/24      10.255.0.2                      ago
```

7. Verify the port-quality calculations based on the ARS configuration and congestion.

```
Leaf1# show ars port-quality
```

```
DLB Quality on egress ports:
```

```
PortName      Quality [0-7]
```

```
Eth1          6
```

```
Eth2          7
```

6.3 RoCEv2 Configuration and Verification

RDMA over Converged Ethernet version 2 (RoCEv2) enables direct memory access between servers without CPU overhead, delivering the ultra-low latency and high throughput required for GPU cluster interconnects and high-performance storage fabrics.

Supermicro Enterprise Advanced SONiC provides a streamlined RoCEv2 enablement command that automatically configures lossless buffer allocation, Priority Flow Control (PFC), WRED/ECN marking, and QoS scheduling policies optimized for RDMA traffic. For advanced configurations, refer to the Supermicro Enterprise Advanced SONiC User Guide.

Important: Enabling and Disabling RoCE requires a switch restart. You will be prompted with a warning; after you input `Y`, the configuration will be saved and the switch will be reloaded.

1. Configure the switch to enable RoCEv2 with the default RoCEv2/iSCSI lossless buffer settings, as well as the default WRED/ECN, scheduling, and QoS map configurations that are defined for the switch.

```
Leaf1(config)# roce enable {force-defaults}
```

2. (VXLAN only) If you are configuring RoCEv2 in combination with VXLAN, you must set the QoS Mode of the VXLAN VTEP interface to uniform mode to copy the DSCP value from the inner header to the outer VXLAN header.

```
Leaf1(config)# interface vxlan vtep[name]
```

```
Leaf1(config-if-vtep1)# qos-mode uniform
```

3. (VXLAN only) Similarly, to ensure ECN trims and marks the packet as expected in a VXLAN topology, you must configure a WRED policy and associate it to the VXLAN VTEP interface.

```
Leaf1(config)# qos wred-policy <wred-policy-name>
```

```
Leaf1(config-wred-wred-green)# green minimum-threshold <minimum-
threshold-value> maximum-threshold <maximum-threshold-value>
drop-probability <drop-probability-value>
```

```
Leaf1(config-wred-wred-green)# ecn green
```

```
!
```

```
Leaf1(config)# interface vxlan vxlan-interface-name
```

```
Leaf1(config-if-vtep1)# queue [0-7] wred-policy <wred-policy-
name>
```

4. After the switch reboots and the system status is ready, you can review and verify the default RoCE QoS policy and behavior.

```
Leaf1# show qos map dscp-tc
```

```
DSCP-TC-MAP: ROCE
```

```
  - - -   - - -  
DSCP     TC  
  - - -   - - -
```

```
0         0
```

```
1         0
```

```
2         0
```

```
3         0
```

```
4         4
```

```
5         0
```

```
<Snipped>
```

```
23        0
```

```
24        3
```

```
25        0
```

```
26        3
```

```
27        0
```

```
28        0
```

```
<Snipped>
```

```
47        0
```

```
48        6
```

```
49        0
```

```
<Snipped>
```

```
62        0
```

```
63        0
```

```
!
```

```
Leaf1# show qos map dot1p-tc
```

```
DOT1P-TC-MAP: ROCE
```

```
  - - -   - - -  
DOT1P     TC  
  - - -   - - -
```

```
0         0
```

```

1          0
2          0
3          3
4          4
5          0
6          0
7          0

!

! These ingress traffic classes are assigned to RoCEv2 Priority
Groups

!

Leaf1# show qos map tc-pg
Traffic-Class-Priority-Group-MAP: ROCE

- - - - -
TC      PG
- - - - -
0        7
1        7
2        7
3        3
4        4
5        7
6        7
7        7
- - - - -

! Ingress traffic classes are assigned to egress queues 0, 3, 4,
and 6. No front-panel ports are mapped to egress queues 1, 2, 5,
and 7

! Traffic generated by switch CPU is sent using queue 7

!

Leaf1# show qos map tc-queue
Traffic-Class-Queue-MAP: ROCE

- - - - -

```

TC	Queue
- - -	- - -
0	0
1	0
2	0
3	3
4	4
5	0
6	6
7	0

!

! The PFC priority traffic is assigned to RoCEv2 PFC priority queues

!

Leaf1# show qos map pfc-priority-queue

PFC-Priority-Queue-MAP: ROCE

- - -	- - -
PFC Priority	Queue
- - -	- - -
0	0
1	1
2	2
3	3
4	4
5	5
6	6
7	7

!

! The default scheduler policy for WRED/ECN configures the PFC priority queues for RoCEv2 traffic

!

Leaf1# show qos scheduler-policy

Scheduler Policy: ROCE

```
Queue: 0
    type: dwrr
    weight: 50
Queue: 3
    type: dwrr
    weight: 50
Queue: 4
    type: dwrr
    weight: 50
Queue: 6
    type: strict
```

!

! The default WRED policy has a minimum and maximum threshold value, drop rate, and ECN traffic filter configured

Note: The output will vary depending on your switch platform.

```
Leaf1# show buffer pool
egress_lossless_pool:
size                : 31617024
type                : egress
mode                : static
egress_lossy_pool:
size                : 24320512
type                : egress
mode                : dynamic
ingress_lossless_pool:
size                : 32157184
type                : ingress
shared-headroom-size : 2621440
mode                : dynamic
```

!

! Various buffer profiles are created and associated with an ingress or egress buffer pool

! This specifies reserved memory, static/dynamic thresholds, and optional pause/resume thresholds

!

! By default, all switch interfaces are assigned to PFC priority groups with ingress buffer profiles

!

```
Leaf1# show buffer interface Ethernet all priority-group
Interface          priority-group  Profile
Ethernet0          3-4            pg_lossless_25000_40m_
                  profile
Ethernet0          7              ingress_lossy_profile
Ethernet1          3-4            pg_lossless_25000_40m_
                  profile
Ethernet1          7              ingress_lossy_profile
Ethernet2          3-4            pg_lossless_25000_40m_
                  profile
Ethernet2          7              ingress_lossy_profile
<Snipped>
```

!

! By default, all switch interfaces are assigned to egress queues with egress buffer profiles

!

```
Leaf1# show buffer interface Ethernet all queue
Interface          queue          Profile
CPU                0-47          egress_lossy_cpu_profile
Ethernet0          0-2,5-19     egress_lossy_profile
Ethernet0          3-4          egress_lossless_profile
Ethernet1          0-2,5-19     egress_lossy_profile
Ethernet1          3-4          egress_lossless_profile
Ethernet2          0-2,5-19     egress_lossy_profile
Ethernet2          3-4          egress_lossless_profile
<Snipped>
```

6.4 Recommended RoCEv2 Best Practices

Refer to the following best practices depending on your Supermicro Enterprise Advanced SONiC setup.

Adjust WRED Policy to Recommended Best Practices for ECN

Depending on the lowest common denominator interface speed of the GPU RNIC, it is recommended to adjust the WRED Policy to configure different minimum and maximum thresholds, as well as drop probability.

- **If the lowest interface speed is 200G and below:**

```
Leaf1(config)# qos wred-policy ROCE
Leaf1(config-wred-ROCE)# green minimum-threshold 1000 maximum-
threshold 3000 drop-probability 20
```

- **If the lowest interface speed is 400G and above:**

```
Leaf1(config)# qos wred-policy ROCE
Leaf1(config-wred-ROCE)# green minimum-threshold 500 maximum-
threshold 1500 drop-probability 20
```

Map DSCP 16 to Support RoCE ACK Packets (Only for Intel Gaudi-Based Deployments)

In addition to mapping RoCEv2 traffic to DSCP 26, Intel Gaudi maps RoCE ACK packets separately to DSCP 16. Because of this, take additional steps to map DSCP 16 to Traffic Class 2 using the configuration below.

```
Leaf1(config)# qos map dscp-tc ROCE
Leaf1(config-dscp-tc-map-ROCE)# dscp 16 traffic-class 2
Leaf1(config-dscp-tc-map-ROCE)# exit
!
Leaf1(config)# qos map tc-pg ROCE
Leaf1(config-tc-pg-map-ROCE)# traffic-class 2 priority-group 2
Leaf1(config-tc-pg-map-ROCE)# exit
!
Leaf1(config)# qos scheduler-policy ROCE
Leaf1(config-sched-policy-ROCE) queue 2
```

```
Leaf1(config-scheduler-ROCE-queue-2)# type dwrr
Leaf1(config-scheduler-ROCE-queue-2)# weight 50
Leaf1(config-scheduler-ROCE-queue-2)# exit
!
Leaf1(config)# interface Ethx/y
Leaf(config-if-Ethx/y)# queue 2 wred-policy ROCE
Leaf(config-if-Ethx/y)# priority-flow-control priority 2
```

Chapter 7:

Additional Resources

The following additional resources are available for Supermicro Enterprise Advanced SONiC switches.

In addition, Supermicro Enterprise Advanced SONiC user guides are under construction for external publishing. In the meantime, reach out to your Sales team or open a Support case to obtain the latest user guides available.

7.1 Password Reset	64
7.2 System RNIC Configuration Workflow Example	67
7.3 Contacting Supermicro	69

7.1 Password Reset

If you lose or forget your Supermicro Enterprise Advanced SONiC account credentials, including the admin password, you can reset the password by completing the following steps.

1. Connect to the serial or USB console port. The serial settings are 115200 baud, 8 data bits, and no parity.
2. Reboot or power on the switch. If you reboot, be sure to save any unsaved configuration changes by entering the write memory command in Exec mode.
3. Wait for the GRUB menu to appear.



Figure 7-1. GRUB Menu

4. In the GRUB menu, select the Enterprise Advanced SONiC version. Then press the **E** key to open the GRUB editor. Use the arrow keys to navigate to the line that starts with `linux /image-`. This line becomes highlighted.

Note: If the arrow keys do not work in the GRUB editor, use the following key combinations to navigate to the line:

- To go up: **Ctrl+P**
- To go down: **Ctrl+N**
- To go left: **Ctrl+B**
- To go right: **Ctrl+F**

```

GNU GRUB version 2.02

+-----+
|setparams 'SONiC-0S-4.4.0-Enterprise_Base'
|
|   search --no-floppy --label --set=root SONiC-0S
|   echo  'Loading SONiC-0S OS kernel ...'
|   insmod gzio
|   if [ x = xxen ]; then insmod xzio; insmod lzopio; fi
|   insmod part_msdos
|   insmod ext2
|   linux /image-4.4.0-Enterprise_Base/boot/vmlinuz-5.10.0-21-amd64 r\
|oot=UUID=91719e4d-2a0e-4cc0-9571-ade38f570486 rw console=tty0 console=ttyS0\
|,115200n8 quiet intel_idle.max_cstate=0 net.ifnames=0 bios\
|devname=0          loop=image-4.4.0-Enterprise_Base/fs.squashfs loop\|v
+-----+

Minimum Emacs-like screen editing is supported. TAB lists
completions. Press Ctrl-x or F10 to boot, Ctrl-c or F2 for
a command-line or ESC to discard edits and return to the GRUB menu.

```

Figure 7-2. GRUB Editor

- In the line that starts with `linux /image-`, type `init=/bin/bash` between `rw` and `console=tty0` as shown in the output below. To undo any text you have typed and return to the GRUB menu without saving, press the **Esc** key.

Note: This text insertion does not persist across reboots; you do not have to re-edit GRUB after you recover your password and use it to log in again.

```

GNU GRUB version 2.02

+-----+
|   insmod ext2
|   linux /image-4.4.0-Enterprise_Base/boot/vmlinuz-5.10.0-21-amd64 r\
|oot=UUID=91719e4d-2a0e-4cc0-9571-ade38f570486 rw init=/bin/bash console=tty\
|0 console=ttyS0,115200n8 quiet intel_idle.max_cstate=0 net\
|.ifnames=0 biosdevname=0          loop=image-4.4.0-Enterprise_Base/f\
|s.squashfs loopfstype=squashfs          crashk\
|ernel=0M-2G:256M,2G-4G:256M,4G-8G:384M,8G-:448M          apparmor=1 \
|security=apparmor varlog_size=4096 usbcore.autosuspend=-1 cgroup_enable=me\
|memory swapaccount=1 cma=0          systemd.unified_cgroup_hierarchy=f\
|alse systemd.legacy_systemd_cgroup_controller=false memmap=64M\16G softlock\
|up_all_cpu_backtrace=1
|   echo  'Loading SONiC-0S OS initial ramdisk ...'
+-----+

Minimum Emacs-like screen editing is supported. TAB lists
completions. Press Ctrl-x or F10 to boot, Ctrl-c or F2 for
a command-line or ESC to discard edits and return to the GRUB menu.

```

Figure 7-3. init=/bin/bash Location

- Reboot by pressing the **F10** key or **Ctrl+X**.

```
Booting a command list

Loading SONiC-OS OS kernel ...
Loading SONiC-OS OS initial ramdisk ...
bash: cannot set terminal process group (-1): Inappropriate ioctl for device
bash: no job control in this shell
root@(none):/# passwd admin
New password:
Retype new password:
passwd: password updated successfully
root@(none):/# echo "b" > /proc/sysrq-trigger
```

Figure 7-4. Sample Boot Sequence

7. To reset the password, use the Linux command `passwd username`. For example:

```
root@(none):/# passwd admin

Enter new UNIX password: xxxxxx

Retype new UNIX password: xxxxxx

passwd: password updated successfully
```

8. Reboot the switch by entering the Linux command `echo "b" > /proc/sysrq-trigger` or `reboot -f`. For example:

```
root@(none):/# echo "b" > /proc/sysrq-trigger
```

Or

```
root@(none):/# reboot -f
```

9. After the switch reboots, you can log in with your username and the new password.

7.2 System RNIC Configuration Workflow Example

The following Supermicro Enterprise Advanced SONiC configuration provides an example of the necessary steps to configure the breakout ports, configure the interfaces with an IP address, enable QoS for RoCEv2, and configure Adaptive Routing and Switching.

1. When RDMA-enabled NIC (RNIC) is configured for RoCE, it's important to configure the proper DCQCN settings on the system RNIC as follows:
 - RoCEv2 packets are assigned a DSCP value of 24 and 26, and internally use Traffic Class 3.
 - Intel Gaudi assigns RoCEv2 ACK packets a DSCP value of 16, with all other RoCEv2 retaining the default DSCP value of 26. Because of this, you must use the configuration described in ["Map DSCP 16 to Support RoCE ACK Packets \(Only for Intel Gaudi-Based Deployments\)"](#) on page 61.
 - ECN CNP packets are assigned a DSCP value of 48 and internally use Traffic Class 6.
 - PFC is enabled specifically for Traffic Class 3.
 - Three or four traffic classes are established depending on whether your deployment is Intel Gaudi-based:
 - TC0 for non-RoCE traffic
 - TC2 for RoCE ACK traffic (only for Intel Gaudi)
 - TC3 for RoCE traffic
 - TC6 for ECN CNP traffic
2. Depending on the RNIC capability and the default port speed of the switch, you might need to configure breakout ports. For example, SSE-T8164 has 64x800G ports, potentially requiring you to break an 800G port into 2x400G ports.
For details, see ["Configuring Breakout Ports"](#) on page 24.
3. To enable IP-based connectivity between 2 or more RNICs you will need to configure an IP address between the RNIC and switch interface. While not required, it is recommended to configure a /31 subnet between the RNIC and switch interface to maximize IPv4 addressing.
For details, see ["Configuring Layer 3 Interfaces"](#) on page 30.

4. After configuring the breakout port (if applicable) and IP address on the interface, it is recommended to validate the connectivity between the leaf switch and the RNIC.

From the leaf switch, enter the command `sonic-cli` if not already entered, and then validate whether you can ping the RNIC.

```
Leaf# ping <RNIC-IP> -I <Leaf-IP>
```

5. To ensure DSCP 0, 24, 26, and 48 are mapped to the appropriate queues and schedulers, enable the RoCE default configuration, which configures the underlying QoS functionality.

For details, see ["RoCEv2 Configuration and Verification" on page 55](#) and ["Recommended RoCEv2 Best Practices" on page 61](#).

6. (Optional) If you are using multiple leaf switches with leaf-leaf or leaf-spine fabric connections, it is recommended to configure Adaptive Routing and Switching (ARS) which redistributes flows based on past/current/future load.

For details, see ["Adaptive Routing Configuration" on page 50](#) and ["Adaptive Routing Verification" on page 52](#).

7.3 Contacting Supermicro

Headquarters

Address: Super Micro Computer, Inc.
980 Rock Ave.
San Jose, CA 95131 U.S.A.

Tel: +1 (408) 503-8000

Fax: +1 (408) 503-8008

Email: Marketing@supermicro.com (General Information)
Sales-USA@supermicro.com (Sales Inquiries)
Government_Sales-USA@supermicro.com (Gov. Sales Inquiries)
Support@supermicro.com (Technical Support)
RMA@Supermicro.com (RMA Support)
Webmaster@supermicro.com (Webmaster)

Website: <https://www.supermicro.com>

Europe

Address: Super Micro Computer B.V.
Het Sterrenbeeld 28, 5215 ML
's-Hertogenbosch, The Netherlands

Tel: +31 (0) 73-6400390

Fax: +31 (0) 73-6416525

Email: Sales_Europe@supermicro.com (Sales Inquiries)
Support_Europe@supermicro.com (Technical Support)
RMA_Europe@supermicro.com (RMA Support)

Website: <https://www.supermicro.nl>

Asia-Pacific

Address: Super Micro Computer, Inc.
3F, No. 150, Jian 1st Rd.
Zhonghe Dist., New Taipei City 235 Taiwan (R.O.C)

Tel: +886 (2) 8226-3990

Fax: +886 (2) 8226-3992

Email: Sales-Asia@supermicro.com.tw (Sales Inquiries)
Support@supermicro.com.tw (Technical Support)
RMA@supermicro.com.tw (RMA Support)

Website: <https://www.supermicro.com.tw>