Supermicro Enterprise Advanced SONiC v4.5.0 Release Notes

Revision 1.0

Updated June 2025

Revisions

Version	Date	Notes
1.0	June 2, 2025	Initial version

Supported Hardware

SSE-C4632

SSE-T8032

SSE-T8164

SSE-T8196*

*Newly introduced hardware in 4.5.0

New Features and Enhancements

Layer 2

IGMP Snooping scale up to 8K

LLDP for DCBx (for ROCEv2 use case)

Q-in-Q Decoupling of CV-ID and SV-ID

Q-in-Q scaling increased to 8000 VLAN translations

Q-in-Q VLAN stacking C-VID/S-VID decoupling

Bridged Layer-2 Protocol Tunneling

Layer 3

Partition ECMP groups (3k for ECMP overlay, 1k for underlay)

Resilient Hashing

UDF ECMP Hashing - Additional option to choose the offset and number of bytes

Weighted ECMP (UCMP)

BGP Max Prefix with route drop

Fast Link Failover (FLF) support

OSPFv3/IPv6 OSPF

Proxy ARP

Manageability

sFlow support up to 128+128+128 bytes depth

Capture mirrored packets on the CPU without forwarding them

MCLAG

MCLAG peer liveliness and split-brain management

QoS

ETS Configuration support

Extending non default ROCE support

SONiC RoCE buffer config based on cable length per interface

WRED Profile threshold range based on the platform capability

Security

RADIUS over TLS

Support special characters # (pound) and , (comma) for TACACS+ Server Key and RADIUS

Federal Certs - Common Criteria

ISG Zero Trust Requirements: MFA (RSA SecurID)

PKI Enhancements for Certification Validation and Certificate Maintenance

Port MAC Security Enhancements: Sticky MAC, MLL

SSH CAC-PIV Support

System & Platform Infrastructure

Linear Pluggable Optics support

Telemetry

gNMI Subscription for LLDP Neighbor status and Tunnel count change

VXLAN

Import directly connected routes across the VRF and install entry in the HW so that pkts can flow in HW instead in SW

Import host routes into local address family in addition to Type-2 routes

Previously Introduced Features/Enhancements

ACLs

ACL Consistency Checker

Control Plane ACL

Layer 2 ACLs

Layer 3 ACLs

Layer 2

IGMP

IGMP Snooping (v1, v2, v3)

IGMP Snooping (v1, v2, v3) over MCLAG

LLDP

LLDP TLV Subtype (Vlan Name, Link Aggr, MFS) Support

LACP Fast Rate and LACP Fallback

LACP Graceful Shutdown

LACP Individual for VxRail

Port Channel Min-Links configuration enhancement

Static Port-channel (LAG)

VLAN stacking (QinQ)

VLAN stacking (QinQ) enhancements

Enhancement in showing the MSTP configuration to display the default region name

Enhancement to configuring the MST region name as an empty string to interoperate with OS10/Arista/Cisco

PVST and RPVST+

PVST and RPVST+ over MCLAG

RPVST+ Scaling to 3500 VLAN Ports

DHCP snooping

Interface Hold-Down

UDLD

Uplink Tracking

VLAN Auto-state

Layer 3

DHCP Relay

DHCP Relay and DHCP Snooping support on the same VLAN

DHCP Relay Circuit-Id Format Support

DHCP Relay Circuit-Id Option

DHCP Relay configuration to support 16 DHCP servers

DHCP Relay Hop Count Configuration

DHCP Relay Option 82, Sub Option 151 VRF Name/ID Option

DHCP Relay Option 82, Sub Option 5 Link-Selection Option RFC3527

DHCP Relay Over IPv4 Unnumbered Interfaces

DHCP Relay over IPv6 Link-Local Interfaces with RFC5549 Routes

DHCP Relay over VXLAN Overlay Interfaces

DHCP Relay Source Interface Selection (e.g. Loopback)

IP Helper

Smart DHCP Relay (DHCPv4 only)

Avoid Netlink for Handling IPv6 Link-Local Address

CLI commands for RA Retransmission Interval, RADv Disable

CPU Offload for Neighbor Suppression

CPU Offload for Slowpath ARP Flooding

RFC 8106 IPv6 Router Advertisement (RA) options

ECMP

ECMP Group Partitioning

Flow-based Hashing

Hash Algorithm, Hash Offset and Ingress Port

QPN Based Hashing (UDF Hashing)

Symmetric Hashing

Versatile Hashing

IPv4 PIM-SSM Support

L3 IGMP

L3 Multicast with PIM operates on L3 interfaces only

1 Million Route Scale

Adaptive Routing (DLB) support

BFD IS-CLIs

BFD Optimizations to Support 5x100msec Aggressive Timers in SW

BFD Profile

BFD with VRF

BGP Docker Warm Restart

BGP Dynamic Neighbor

BGP Route Filtering Representation

BGP Support for ASDot and ASDot+ notation for 4 byte BGP ASNs

BGPv4 and BGPv6

Drop Neighbor Entry to protect CPU from Unknown IP Packets hitting the CPU

Higher Route scale

OSFPv2 GRRouter Advertisement Support (via KLISH/REST/gNMI only)

OSPFv2

PBR Enhancements for Service Chaining

Policy-based Routing (IPv4 and IPv6)

RIB/FIB Consistency Checker

Route Leaking across VRFs including Management VRF

Route Maps

Route Updates Performance Improvements

Routed Subinterface

Static Route pointing to unnumbered interface

Static Routing

4K L3 VLAN Interface Scale for SAG and Unique-IP Cases

IP SLA (ICMP and TCP tracker)

IPv4 Unnumbered Interfaces

L3 VLAN Scale to 4K

NAT

Next Hop Group (NHG) Support

Nexthop Resolution using Default Route

RIF Counters for L3 Interfaces

Support for 4K L3 VLAN Interfaces

VRRP

VRRPv3, VRRP/VRRPv3 over VRF

Manageability

Ability to Filter Logs based on Facility and Severity

Application bring up for Events and Alarms

Audit Logging and Syslogs

AuditD

In-memory Debug Logging

Interface Events (Notify Interface Oper Up/Down in show event)

Option to send Audit Log Messages to Syslog Server

Separate Channel for audit logs

Syslog High Threshold Notifications and clear for CPU/Temperature

Syslog over TCP

Syslog over TLS

VRF Support for Syslog

sFlow

sFlow on Management VRF

SNMP Configuration Traps and OIDs

SNMP new CLIs for enabling or disabling individual traps

SNMP Polling

SNMP Trap Enablement on Interface Instead of Global

Support to Read Service Tag via SNMP

TEC certification SNMP “set” command for sysName

Broadcom Debug Tool

Command to return Interfaces to the default configuration

Configuration Services Chef for EVPN

Host Table Resource Reservation for Local Hosts

ifindex support for physical interfaces in show interface output

Industry Standard CLI (IS-CLI)

Management VRF Hardening

NTP Prefer Option

NTP Server and NTP Authentication

Port Mirroring on Port Channel and VLAN Scalability Improvements

SCP, SFTP, TFTP, FTP

Session timeout configuration

SSH

Time Zone Command Support

VRF support for SSH.in

Wildcard support for OC-YANG interface's Ethernet counters and individual counters

Zero Touch Provisioning (ZTP)

ZTP Provisioning using a USB Drive

MCLAG

Advertise PIP for both ACT-ACT and ACT-STBY on the Same Leaf Pair

L2 LVTEP

L2 MCLAG

L3 LVTEP

L3 MCLAG

Management VRF for MCLAG

MCLAG Fallback

MCLAG Graceful Shutdown

MCLAG Peer Gateway

xSTP over MCLAG

QoS

ACL-based CoPP

CoPP (Control Plane Policing)

Per Platform CoPP

ACL-based DSCP Map/Remarking

ACL-based PCP Map/Remarking

ECN (Explicit Congestion Notification)

PFC (Priority Flow Control)

Programmable PFC priority-to-queue mapping

RoCEv2 support

RoCEv2 with Cut Through mode support

WRED (Weighted Random Early Detection)

ACL-based Rate Limiting

BUM Storm Control

L2 QoS Maps

L3 QoS Maps

Port and Priority Shaping

QoS Map Support for Remarking and SVI

Queue and Buffer Size Configuration

Traffic Priority scheduling (Strict, WFQ)

Security

LDAP Integration

LDAP multiple role mapping

LDAPS (secure LDAP over TLS)

Map LDAP groups to SONiC (RBAC) roles

AAA Authorization support with TACACS+

RADIUS and TACACS+

RADIUS/TACACS+ Password Obfuscation

Custom password complexity

Port MAC Security

RBAC and HAMd Enhancements

Role-based Access Control (RBAC)

Separate Authentication Methods for Local and Remote Access

SSHD Configurability

System Platform & Infrastructure

Auto Negotiation and Link Training

CMIS 4.0 Optics Support

Default Auto-Negotiation configuration

DOM Information Display

Dynamic Port Breakout

Forwarding Plane Drop Counters

Interface Aliasing (IS-Standard and IS-Standard-Extended Interface Naming)

Interface Beacon LED

Link Flap Error-Disable

Link Statistics Enhancements

Link-Down Reason Codes

Maintenance Mode for LACP, OSPFv2, and BGP

Media AutoFEC for FEC Type automation

Port Auto-Breakout and Auto-Detect for Port Speed support

Transceiver Parameter Tuning

Flexible DPB

Hardware Resource Allocation and Reservation

Hardware Watchdog

Kdump Support

Limiting CPU/Memory/Disk Usage for Third Party Containers

Linux Buster to Bullseye migration

Linux Kernel upgrade to 5.10.162

Memory Histogram

Multi-Instance Redis DB

Patching Support in SONiC (Patch Host/Containers)

PDDF and PDK Framework

Secure Boot Process and Reference Implementation

Support Option to Bind the Third Party Container to the Management VRF

System Locator LED Support (Beacon)

System Ready for Services and Applications

Third-Party Container Management

Transformer Infrastructure: Sonic Yang Singleton Container Support

Warm Boot

Telemetry

Bulking support in both REST(YANG patch) and gNMI

Dynamic configuration of REST server and gNMI server

gNMI Subscription support for Limited YANG Paths (OnChange, Interval, Once, Poll, Target defined)

REST and gNMI Interfaces via OpenConfig YANG (OC-YANG)

Scalar Encoding support for gNMI

Drop Monitor

Inband Flow Analyzer (IFA) 2.0

Query parameter for REST and filtering support for gNMI

Tail Stamping

Watermarks, Thresholds, and Snapshots

VXLAN

BGP for EVPN (with MLAG)

DSCP Marking Preservation for VXLAN

IP fabric over IPv6 underlay RFC5549

L2 VXLAN

L3 VXLAN

Multi Site Data Center Interconnect (DCI)

RoCEv2 over VXLAN

Standards-based ESI based EVPN Multihoming support

VxLAN creation to non /32 prefixes

VXLAN over SVI Interface

Feature Limitations

ACLs

Layer 2

STP

STP loop guard is not supported for PVST.

When STP (PVST/RPVST) is enabled in the network and configuration changes are made, use the config save CLI command followed by a device reboot using the reboot command instead of using the config reload and fast-reboot CLI commands. This recommendation helps to avoid STP loops in the network during the config reload or fast reboot process because the STP control plane will be down on the nodes undergoing a config reload.

STP PortFast – Change in the Default Configuration

When users upgrade from SONiC version 3.0.6 or earlier versions to SONiC 3.0.7 or SONiC 3.1.0 or later versions with PVST PortFast enabled on individual ports, this configuration will be overwritten with PortFast disable on these ports. Users must reconfigure the PVST PortFast settings after upgrading to SONiC 3.0.7 or later releases.

Layer 3

Using System routing_config_mode

To retain FRR feature configuration, such as BGP, BFD, OSPF, and so on, across a BGP docker restart and SONiC system reboot, users should do the following:

- Configure separated mode in the following cases:
    - When using the IS CLI or REST interface for all FRR features.
    - When using the IS CLI or REST interface for the complete configuration of selected features, such as BGP, BFD, OSPF, and so on, and when using VTYSH CLI for the complete configuration of remaining FRR features, such as PIM, static route, and so on.

  Starting with the SONiC 3.1.0 release, separated mode is the default routing docker configuration mode. Users are not required to explicitly configure any routing mode. If the routing mode is not configured to any value, it is also considered as "separated" routing mode.
- Configure split mode for the following use-cases:
    - When using FRR VTYSH for managing all FRR features, such as BGP, BFD, OSPF, and so on.
    - In split mode, the user is expected to save the configuration on VTYSH by executing the write mem command to make the configuration persistent across a reboot, and then execute a config reload. This extra step is required in addition to saving the configuration in IS-CLI/Click CLI configuration save.

  Using the IS-CLI or REST interface to perform partial configuration of an FRR feature (for example, BGP) and also using FRR VTYSH CLIs to configure other aspects of the feature is not supported. That is, an FRR feature such as BGP should be either completely configured using IS-CLI/REST or completely configured by using VTYSH CLI.

CLICK CLI Syntax is as follows:

```
sudo config routing_config_mode {unified|split|separated}
```

For example:

```
admin@sonic:~$ sudo config routing_config_mode split
```

- If VRF-VNI mapping is configured using the sudo config vrf add_vrf_vni_map <vrf> <vni> configuration command, then the FRR configuration is not saved, and the device is rebooted or reloaded, the VRF-VNI mapping configuration in FRR is lost after a reboot or reload, and L3VNI will not be functional.

  To avoid this behavior, after configuring VRF-VNI mapping using the config vrf add_vrf_vni_map command, save the FRR configuration using the vtysh -c write memory command.

Forwarding Behavior

Traffic ingressing on a particular VRF instance interface but destined to leaked connected-routes in a different VRF is software forwarded. For example, connected routes from Vrf-1 are leaked into another VRF Vrf-2. Traffic ingresses on Vrf-2 with a destination IP belonging to Vrf-1 will be software forwarded. If connected routes are leaked, ARP/ND entries learned in one VRF are not leaked into another VRF, causing traffic towards connected hosts to be software forwarded. Route leaking works well when remote routes learned in one VRF are leaked into another VRF. A workaround may require revisiting the VRF subnet and VRF leak configuration requirements and relying on remote route leaking instead of connected route leaking.

High Memory Spike on Running BGP Commands

Running the following commands with the JSON option in VTYSH CLI with a large number of BGP routes (for example, more than 20,000), results in the BGP docker’s increased memory utilization momentarily for a few seconds during CLI execution.

The impacted commands are as follows:

- show ip route
- show bgp ipv4
- show bgp ipv6
- show bgp l2vpn evpn

The BGP docker's memory utilization will return to normal after the CLI execution is complete. A momentary memory utilization spike also occurs in the clish process, but it will return to normal after the CLI execution is complete.

For example, running the following CLI commands in VTYSH CLI can lead to the issue described in this section:

- show bgp l2vn evpn json
- show bgp l2vpn evpn route type prefix json
- show bgp l2vpn evpn route detail json

Adaptive Routing (DLB)

On SSE-T8164 and SSE-T8196 platforms, the maximum configurable value for the following ARS profile tunable parameters should be half of the maximum supported value.

- The maximum configurable value for load-current-max-val should be ≤ 133169151 ÷ 2.
- The maximum configurable value for load-future-max-val should be ≤ 266338303 ÷ 2.
- The maximum configurable value for load-past-max-val should ≤ 10000 ÷ 2.
- The minimum configurable values for the same parameters should be less than or equal to the maximum value configured.

Manageability

NTP

NTP listens on certain IP addresses based on CONFIG_DB configurations.

- First preference: MGMT_INTERFACE table (management interface IP address).
- Second preference: LOOPBACK_INTERFACE table (but it only listens on the Loopback0 IP address).
- Third preference: eth0.
- If MGMT_INTERFACE table is not defined, the LOOPBACK_INTERFACE table is defined, but Loopback0 is not configured with an IP address, then no interfaces (other than 127.0.0.1) will be listened on, and the NTP will not synchronize with any configured NTP server. Configure one of the following:
    - MGMT_INTERFACE or LOOPBACK_INTERFACE | Loopback0 with a valid IP address
    - NTP source-interface

NTP Server: When the SONiC device is acting as an NTP server and fields queries from NTP clients, additional actions must be performed on the SONiC device acting as the NTP server for faster time synchronization on the NTP clients.

- The SONiC NTP server itself must be able to synchronize with an upstream NTP master and servers by configuring suitable upstream NTP servers on the SONiC device acting as the NTP server.
- Minimize the root distance of the SONiC NTP server. If the root distance is greater than approximately 1.5 seconds (default), the NTP client will fail to select this NTP server. The root distance can be reduced by:
    - Using upstream NTP master/servers that are closer to the SONiC NTP server (for example, RTT in the low or sub millisecond range).
    - Setting the clock on the SONiC NTP server as close as possible to the actual time (the time at the NTP master). For example, the following Linux bash commands can be used on the SONiC NTP server that has an upstream server configured:
        - If the management VRF is enabled and NTP is configured to use the management VRF:
          ```
          sudo ntp service stop;
          sudo  cgexec -g l3mdev:mgmt /usr/sbin/ntpd -q -g;
          sudo ntp service start
          ```
        - Otherwise, for NTP in the default VRF:
          ```
          sudo ntp service stop;
          ```

```
sudo /usr/sbin/ntpd -q -g;
sudo ntp service start
```

SNMP

The MIB object atTable(1.3.6.1.2.1.3.1) is disabled. Use ipNetToMediaTable(1.3.6.1.2.1.4.22) instead of atTable.

DCHP and VRRP CoPP traps are not installed by default. DHCP and VRRP CoPP traps will be installed when the respective feature is enabled with configuration commands. Removing the feature configuration commands will disable the respective CoPP traps.

Tech Support Collection

Before running the show techsupport command to collect the tech support data, it is recommended to run the sonic-clear logging command.

MCLAG

QoS

Port-Channel Configuration

Users must remove an existing dscp-tc or dot1p-tc or an existing tc-dscp or tc-dot1p from the physical port before adding it as a member of a port-channel. When the QoS maps are applied on the port-channel, SONiC applies the same configuration on all the member ports. Adding such maps on physical member ports is not allowed.

Security

System Platform & Infrastructure

Third-Party Container Management – New Docker Install

The SONiC Debian package has been upgraded to Bullseye starting with the SONiC 4.4.0 release. The seccomp implementation in Bullseye introduces stricter restrictions on invoking various system calls. As a result, in SONiC 4.4.0 and newer releases, it is necessary to disable the default seccomp profile for the isc_dhcp TPCM docker container.

The -security-opt seccomp=unconfined option is required for the isc_dhcp docker launch.

The full command is as follows:

```
tpcm install name SEN_TPCM pull networkboot/dhcpd args "-v /home/admin/data:/data --network=host --security-opt seccomp=unconfined"
```

Warm Reboot

- Warm Reboot is only supported on the SSE-C6432 platform
- The following multi-D scales are qualified for warm boot:
  - 24K IPv4 prefix routes

  - 20K ARP (Includes both local and remote ARP learned over a VXLAN tunnel)
  - 24K IPv6/ND
  - 24K MAC
- In the scaled configurations users should increase BGP warm-restart timer to avoid routes being prematurely removed from hardware due to WB reconciliation resulting in traffic loss. The default warm-restart timer value is 120 seconds.

  ```
  warm-restart bgp timer <Value>
  ```
- There are also cases of local MAC ageout and replacement by the remote MACs in MCLAG configuration resulting in data traffic loss after a warm boot.
- The warm boot design requires that OrchAgent does not have any entries in the pending state. Pending state entries are those entries that are waiting for some SW entities to become available before the entry can be added to the HW. For example, a route entry without a valid next-hop or corresponding neighbor entry will be in a pending state in OrchAgent until the next-hop or neighbor is resolved. The warm boot restart check will fail in such cases.
- Users should also configure the following knobs in BGP, specially in the scaled configurations. When the system takes a longer time to come up, this prevents the neighbor router from withdrawing the routes.

  ```
  graceful-restart enable
  graceful-restart preserve-fw-state
  graceful-restart restart-time <Value>
  graceful-restart stalepath-time <Value>
  ```

Upgrades/Downgrades

When a new SONiC image is installed, changes made to the Linux rootfs (for example, user-installed packages) are not automatically migrated. Users are advised to use a configuration management tool, such as Chef or Ansible, to manage changes made to Linux rootfs.

Port Groups

On the SSE-C4632, port groups have a design constraint requiring all ports on the same port group to share the same speed.

Auto-Negotiation

When a SSE-T8164 or SSE-T8196 switch is connected to another device using an 800G DAC cable, auto-negotiation should be enabled with a specific speed. That is, speed auto alone is not enough to bring up the link. Instead, the user is expected to execute the speed auto 800000 command to allow the switch to advertise only 800G to negotiate with the peer (even though auto-negotiation is ON).

Telemetry

VXLAN

Support of Any-prefix Underlay Routes to Reach VxLAN Tunnel Destination

In SONiC version 4.1.0 and earlier versions, the VxLAN feature imposes a limitation. Specifically, the VxLAN tunnel becomes operationally active only if there is a /32 underlay host route leading to the tunnel destination

IP. The VxLAN tunnel remains operationally inactive even in the presence of non /32 routes to the destination IP.

SONiC is extending the support of Any Prefix Underlay Routes to reach the tunnel destination in SONiC version 4.2.0. It uses the longest prefix match route available for the destination IP to bring up the tunnel. The best route will always be chosen to establish a tunnel, and if the route is deleted, the next available best route is used. If no best route is available for the tunnel destination IP, it will use the valid default route to establish a tunnel.

Any-prefix Underlay Route mode is enabled only with a cold reboot. Users will not see any behavior change if they perform an ISSU from older versions or with a warm reboot. For ISSU and warm boot, SONiC will maintain the existing route mode to support only /32 host underlay routes to avoid any hardware update or traffic disruption during/after upgrade. When the Any-prefix underlay route mode is enabled with a cold boot, it will persist across an ISSU, warmboot, and cold reboot.

Users will observe the following change in behavior when Any-prefix underlay route mode is enabled after cold-reboot –

- The VxLAN tunnel becomes operationally up with a valid default route if no best route is available for the destination.
- When there is a best route change for the tunnel destination IP due to a route addition or deletion, the VxLAN tunnel starts pointing to the new best route.
    - The tunnel will flap one time if the new best route has a set of underlay Next-Hops that is completely different from the underlay Next-Hops pointed by the old best route.
    - The tunnel will not flap if there is at least one common underlay Next-Hop between the old best route and the new best route.

Broadcast, Unknown Unicast, and Multicast (BUM) Behavior

Broadcast, unknown unicast, and multicast (BUM) traffic sent over VXLAN tunnels will not be load-balanced across the available ECMP paths. It will be forwarded on only one of the paths. The selection of the BUM path is per VXLAN tunnel. However, for all the VLANs that are extended over the VXLAN tunnel, the same BUM next-hop path is used for forwarding the BUM traffic.

EVPN Multihoming

The hairpin switching use case does not work on ports that have ESI configured in an EVPN Multihoming deployment.

Cut-Through Switching

Cut-through switching is not applicable for VXLAN encapsulation flows in SSE-T8164 and SSE-T8196 platforms

Known/Open Issues

Defect ID: SONIC-102017

Component: PFC

Customer Probability: High

Customer Symptom: Drop counters per queue do not get updated.

Customer Condition: PFC watchdog action = DROP configured for lossless priority.

Customer Workaround: None. The device does not support per-queue drop counters.

Defect ID: SONIC-58585

Component: L3 protocols – BFD

Customer Probability: High

Customer Symptom: BFD session between MC-LAG client and MC-LAG node may flap momentarily. It gets restored within few milliseconds after MC-LAG PortChannel is shut/no shut

Customer Condition: All uplink ports of MC-LAG node is shut and no shut and BFD session timeout is 900 milli-second or less

Customer Workaround: Use BFD session timeout > 900 milli-second

Defect ID: SONIC-58678

Component: L3 protocols – BFD

Customer Probability: Medium

Customer Symptom: BFD sessions running between MC-LAG client and MC-LAG node (Active/Standby) flap.

Customer Condition: User issues reboot (cold-reboot/config-reload) on one of the MC-LAG nodes (Active or Standby) in a scaled environment when BFD session timeout is configured as 900 milliseconds.

Customer Workaround: User may configure BFD session timeout greater than 900 milliseconds.

Defect ID: SONIC-72435

Component: Neighbor Discovery

Customer Probability: Medium

Customer Symptom: If ARP request/IPv6 neighbor solicit is received for known local neighbors on neighbor suppression-enabled VLANs, Kernel bridge driver floods these ARP requests only to local ports and suppresses flooding over VxLAN tunnel. But with HW assisted flooding enabled on the platform, flooding decisions are offloaded to hardware. Since hardware is unaware of neighbor suppression configuration on the VLAN, it will continue to flood on local VLAN members and VXLAN tunnels.

Customer Condition: Neighbor Suppression is enabled and hardware assisted flooding supported on the platform.

Customer Workaround: None

Defect ID: SONIC-70517

Component: Port Breakout

Customer Probability: Low

Customer Symptom: SAI_API_PORT errors may be seen during port breakout operation

Customer Condition: Applying port breakout configuration

Customer Workaround: None - no functional impact

Defect ID: SONIC-64785

Component: System

Customer Probability: Low

Customer Symptom: The following error message appears on console and syslog during the bootup: ERR kernel: [7.842752] ata1.00: failed to set xfermode (err_mask=0x40)

Customer Condition: system bootup

Customer Workaround: None - no functional impact

Defect ID: SONIC-89198

Component: ECMP

Customer Probability: Low

Customer Condition: JENKINS_HASH_HI hash must be configured. The issue is seen for non-RoCE traffic only.

Customer Symptom: When ip load-share hash algorithm is configured to JENKINS_HASH_HI with QPN hashing enabled (ip load-share hash roce qpn), for non-RoCE traffic, it is not distributed equally across ECMP member ports.

Customer Workaround: Use the CRC_32HI hash algorithm, which is the default configuration.

Defect ID: SONIC-89758

Component: L2Protocols - RPVST

Customer Probability: High

Customer Condition: Issue will hit only when MCLAG port channel connected down stream devices are STP unaware

Customer Symptom: Active/Standby reboot of MCLAG node can result in MCLAG port channel to staying in Discarding state.

Customer Workaround: Configure edge ports on the Firewall facing ports or on the ports where there is no active spanning tree

Defect ID: SONIC-91400

Component: ACL

Customer Probability: Very Low

Customer Condition: If customer uses ingress ACL on router port and use OUTER-VLAN-ID as a match criteria; The packet will not be filtered. This issue is applicable for SSE-T8164 and SSE-T8196 platforms only.

Customer Symptom: Ingress ACL on a router port will be unable to match the traffic based on OUTER-VLAN-ID match criteria.

Customer Workaround: Customer can use other packet fields (e.g. src-ip/dst-ip).

Defect ID: SONIC-92983

Component: MCLAG

Customer Probability: When all of below conditions are met, this issue might be observed:
In a two tier MCLAG topology, when querier configured in upper tier MCLAG peers.
Then, IGMP reports sent from client connected to lower tier MCLAG peers.
Later when lower tier MCLAG port is shutdown and no shutdown on MCLAG port issued after 210 seconds which is the multicast entry timeout.

Customer Condition: In a two tier MCLAG setup, when IGMP reports destined to local MCLAG member node started reaching to remote MCLAG node in the upper tier MCLAG.

Customer Symptom: Traffic impact will be there for the missing multicast group until the next IGMP report is received.

Customer Workaround: Add a static multicast group entry using the command ip igmp snooping static-group <group-address> interface <interface> at the upper MCLAG tier to retain deleting the entry from timeout.

Defect ID: SONIC-96778

Component: MCLAG

Customer Probability This issue might be observed from the following sequence:
a. In a MCLAG topology, learn snooping entries by sending IGMPv3 reports from MCLAG client and orphan ports.
b. Send IGMPv3 leave on MCLAG client and orphan port.

Customer Condition: In an MCLAG setup, sending IGMPv3 leave from MCLAG clients and orphan ports.

Customer Symptom: Sometimes, one of the MCLAG peers might be holding stale IGMPv3 snooping entries until cleaned up explicitly.

Customer Workaround: Disable and enable IGMP snooping on the VLAN where stale entries are present.

Defect ID: SONIC-96117

Component: MCLAG

Customer Probability: When all of the following conditions are met, this issue might be observed:– In a MCLAG topology, learn IGMP snooping entries on orphan ports from both MCLAG peers.– Now, switch the IGMP snooping version on MCLAG peers without the host sending IGMP leave on older version.– Again, move the IGMP snooping version on MCLAG peers to the originally configured version keeping the IGMP host version intact on orphan ports.

Customer Condition: In an MCLAG setup, changing the IGMP snooping version on MCLAG peer without host sending IGMP leave for the groups learned on orphan ports.

Customer Symptom: Traffic impact will be there for a multicast group that is not synced from the remote peer.

Customer Workaround: Send an IGMP leave from a host connected to the orphan port before switching the IGMP snooping version on the MCLAG peers.

Defect ID: SONIC-96357

Component: VRF

Customer Probability: Upon a config replace of SAG configuration with VRF binding on new VLAN interfaces.

Customer Condition: If the SAG and VRF binding configuration on a VLAN arrives before corresponding VLANs are configured upon config replacement with SAG configuration on the VLANs.

Customer Symptom: The SAG VLAN netdev sysctl properties arp_accept, arp_announce, and ipv6.forwarding are not configured in the Linux kernel, which prevents IPv4 and IPv6 neighbors from learning on the VLAN.

Customer Workaround: Clearing the MAC addresses on the corresponding PortChannel will force the system to recover from the issue.

Defect ID: SONIC-94646

Component: MCLAG

Customer Probability: Low.

Customer Condition: After completing all MCLAG and MSTP configurations, the peer-link port channel (PO) is unconfigured, then configured with the MCLAG configuration. This may cause incorrect STP convergence.

Customer Symptom: Traffic disruption due to the wrong port state

Customer Workaround: Clearing the MAC addresses on the corresponding PortChannel will force the system to recover from the issue.

Defect ID: SONIC-101917

Component: PFC

Customer Probability: High

Customer Condition: Port is receiving lossy Pause Streams from the partner in Asymmetric PFC Disabled mode.

Customer Symptom: PFC on this port should honor only lossless 3 & 4 Pauses streams, but it honors the lossy streams as well. PFC Rx gets enabled on all the priorities even though it is enabled for a single priority.

Customer Workaround: None. Device supports perPort control instead of perPriority control for PFC Rx.

Defect ID: SONIC-102017

Component: PFC

Customer Probability: High

Customer Condition: PFC watchdog action=DROP configured for lossless priority.

Customer Symptom: Drop counters per queue do not get updated.

Customer Workaround: None. Device does not support perQueue Drop Counters.

Defect ID: SONIC-103097

Component: PFC

Customer Probability: Medium

Customer Condition: PFC watchdog is enabled for all the ports with lower detection-time (< 1sec).

Customer Symptom: PFC watchdog when enabled on all ports restore time takes double the configured time. As the number of ports with watchdog increases, restoration period increases due to high CPU usage.

Customer Workaround: When PFC watchdog is enabled for all the ports, detection-time & restoration-time should be configured with higher value.

Defect ID: SONIC-99653

Component: L3 - IPv6

Customer Probability: Low

Customer Symptom: IPv6 hosts are not reachable over the static default route with the interface as the next hop.

Customer Condition: An IPv6 static default route with only interface as the next hop does not work. In this case, destination IPv6 neighbors must be resolved on the next hop interface, which may be out of the subnet of the next hop interface that is not supported.

Customer Workaround: Configure an IPv6 default static route with a global or link-local IPv6 address as the next hop.

Upgrade/Downgrade Guidelines

Supported Upgrade/Downgrade Paths

From Version	To SONiC 3.1.x	To SONiC 3.2.x	To SONiC 3.4.x	To SONiC 3.5.x	To SONiC 4.x.x
SONiC 3.1.x	Yes	Yes	Yes	Yes	Yes
SONiC 3.2.x	N/A	Yes	Yes	Yes	Yes
SONiC 3.4.x	N/A	N/A	Yes	Yes	Yes
SONiC 3.5.x	N/A	N/A	N/A	Yes	Yes
SONiC 4.x.x	N/A	N/A	N/A	N/A	Yes

Note: Downgrading to a release prior to SONiC 4.1.0 is not supported if port breakout is performed in a system running with SONiC 4.1.0 or a later release image.

Refer to the preceding table to determine if the saved startup configuration file, /etc/sonic/config_db.json, can be migrated to a different image version of Enterprise SONiC. In scenarios where configuration migration is not supported, users are expected to install the SONiC image from the ONIE prompt or by using the sonic_installer install --skip_migration command or proceed with a supported multi-hop upgrade path.

After upgrading from version A to version B, and if the user intends to safely go back to version A (which is available in the secondary partition), use the image set-default command, as shown in the following example:

```
sonic# show image list

 Current: SONiC-OS-4.0.0_Enterprise_Advanced

 Next: SONiC-OS-4.0.0_Enterprise_Advanced

 Available:

 SONiC-OS-4.1.0_Enterprise_Advanced

 sonic# image set-default ?

          String  Image name

 sonic# image set-default SONiC-OS-4.1.0_Enterprise_Advanced
```

When upgrading devices from SONiC versions earlier than SONiC 3.1.0 to SONiC 3.1.0 or higher versions with FRR features configured using VTYSH CLIs, make sure that routing mode is configured to split mode before the upgrade. In this case, if split mode is not configured in releases prior to 3.1.0, the FRR VTYSH configuration will be lost after upgrade.

When upgrading from to a newer version of Enterprise SONiC using the sonic_installer command, the startup configuration file located at /etc/sonic/config_db.json is migrated to the new version. Because the configuration schema can change in the newer version of Enterprise SONiC, the contents of the migrated config_db.json file are transformed to be conforming to the newer version. To preserve these configuration transformations, the user is expected to save the configuration after the new image boots. If the configuration is not saved, the configuration loaded from the config_db.json file is transformed every time. To avoid this overhead, save the configuration.