BROADCOM®

SONiC

Release Notes

Enterprise SONiC Distribution by Broadcom, Version 4.1.1

Broadcom reserves the right to make changes without further notice to any products or data herein to improve reliability, function, or design. Information furnished by Broadcom is believed to be accurate and reliable. However, Broadcom does not assume any liability arising out of the application or use of this information, nor the application or use of any product or circuit described herein, neither does it convey any license under its patent rights nor the rights of others.

Table of Contents

1 Overview 4
1.1 Packages 5
1.2 List of Features not Supported in all Packages 6
1.3 Date of Delivery 6
1.4 Items Included in the Delivery....7
2 Customer Use Cases 8
3 Supported Platforms....12
4 Newly Introduced and Qualified Software Features 14
5 Supported Software Features from Previous Releases 14
6 Limitations, Restrictions, Workarounds, and Guidelines 19
7 SONiC 4.1.1 Image Upgrade and Downgrade Considerations....22
8 Silicon Support and Feature Matrix....24
9 List of Supported Optics and Media....26
10 Scalability Matrix....29
10.1 Guidelines for MAC Aging Timer Settings for Scaled ARP Hosts with Unidirectional Traffic 35
10.2 Guidelines for Scale Limitation and CoPP Setting for Trident4 35
10.3 Guidelines for BFD Interval Configuration 36
10.4 “show CRM resources” Enhancement for IPv4 and IPv6 Routes 36
10.5 Guideline for Configuring “radv enable” for IPv6 Router Advertisement....37
11 Deprecation of CLI Commands....37
12 Open Known Defects 38
Revision History....43
SONiC-RN411; July 24, 2023 43

1 Overview

SONiC is an open source network operating system based on Linux that runs on merchant silicon based ODM platforms. The open source SONiC project is at https://github.com/Azure/SONiC/wiki.

SONiC is in production today at multiple web-scale companies for data center fabric deployments and has a healthy developer community and vendor ecosystem around it. The underlying architecture of SONiC is here: https://github.com/Azure/SONiC/wiki/Architecture.

The Enterprise SONiC distribution by Broadcom is a commercial offering based on open source SONiC with feature enrichment and hardening targeted at Data Center leaf and spine/super-spine use cases. The Enterprise SONiC distribution by Broadcom supports ODM and OEM platforms based on the StrataXGS® family of silicon from Broadcom.

The Enterprise SONiC distribution by Broadcom offers benefits such as high performance and simplicity based on industry leading merchant silicon and standards based IP CLOS architecture. It also provides agility based on a flexible management framework with programmatic APIs and an extensible, container-based architecture. Finally, an open source foundation and standardized ecosystem provide strong economic benefits for a data center fabric solution.

Starting with release 4.0.0, Enterprise SONiC can also be used for campus use cases where customers have a two-layer fabric (access and aggregation) and edge devices such as POS terminals, Thin Clients, Security Cameras, can be attached to the access layer, and the aggregation layer can connect to existing Data Center deployment already running Enterprise SONiC.

The Enterprise SONiC distribution by Broadcom version 4.1.1 GA uses the software components listed in the following table.

Component	Version
Broadcom SAI Adapter	8.5.0.0
Broadcom SDK	6.5.27
Base Operating System	Debian GNU/Linux 10.13 (Buster)
Linux Kernel	5.10.0-8-2-amd64
Config DB	version_4_1_1
FRR	8.2.2
Community SONiC last rebase	202012

1.1 Packages

The Enterprise SONiC distribution by Broadcom, version 4.1.1 packages are as follows:

- Cloud_Base Package
  - This includes features/functionality (for example, eBGP, ZTP, programmatic API, QoS, ACL, and security features such as TACACS+) required for DC fabric underlay leaf, spine, super-spine use cases
  - Base telemetry features: Thresholds and Snapshots (BST)
- Cloud_Advanced Package
  - This includes all features in Cloud_Base
  - Inband Flow Analyzer (IFA, version 2.0), Tail Stamping and Drop Monitor
  - Linux PTP (KNETSync)
- Enterprise_Base Package
  - This includes underlay features (for example, eBGP, ZTP, programmatic API, QoS, ACL, and so on) required for DC fabric underlay
  - Overlay features (BGP EVPN, VXLAN, and so on) for DC fabric overlay use cases (for PINs leaf, spine, super-spine)
  - Enterprise Features (such as RPVST+, IP Multicast, and so on)
  - Base telemetry features: Thresholds and Snapshots (BST)
- Enterprise_Advanced Package
  - This includes all features in Enterprise_Base
  - Inband Flow Analyzer (IFA, version 2.0), Tail Stamping and Drop Monitor
  - Linux PTP (KNETSync)
  - Broadcom Debug Tool
- Campus Package
  - This includes the following features in addition to L2/L3 features:
  - POE, POE+ and POE-bt
  - Features related to Port Access Control (IEEE 802.1X, MAB, dynamic ACLs, RADIUS-assigned VLANs, RADIUS support)
  - LLDP-MED
  - Port Security
  - Digital Optical Monitoring and Time Domain Reflectometry
  - EVPN VXLAN
  - Campus package is supported on campus-specific platforms only

The packages are supported on various platforms according to the mapping in Section 3, Supported Platforms.

The Enterprise SONiC distribution by Broadcom, version 4.1.1 packages are available as a subscription license.

1.2 List of Features not Supported in all Packages

The following table lists the SONiC features that are supported only in certain packages, but not all packages. Any feature not listed in this table is supported by all Enterprise SONiC Distribution by Broadcom, Version 4.1.1 packages.

Feature	Cloud Base	Cloud-Adv	Enterprise-Base	Enterprise-Adv	Campus
VXLAN, LVTEP	Not Supported	Not Supported	Supported	Supported	Supported
NAT	Supported	Supported	Supported	Supported	Not Supported
gNMI	Supported	Supported	Supported	Supported	Not Supported[a]
PVST, RPVST+	Not Supported	Not Supported	Supported	Supported	Supported
Threshold/BST	Supported	Supported	Supported	Supported	Not Supported
EVPN control plane	Not Supported	Not Supported	Supported	Supported	Supported
L3 IGMP and IGMP snooping	Not Supported	Not Supported	Supported	Supported	Supported
IPv4 PIM SSM	Not Supported	Not Supported	Supported	Supported	Supported
Multisite DCi	Not Supported	Not Supported	Supported	Supported	Not Supported
MSTP	Not Supported	Not Supported	Supported	Supported	Supported
ACL-based replication	Not Supported	Not Supported	Supported	Supported	Not Supported
Broadcom debug tool	Not Supported	Not Supported	Not Supported	Supported	Not Supported
Inband Flow Analyzer tail stamping and drop monitor	Not Supported	Supported	Not Supported	Supported	Not Supported
PTP	Not Supported	Supported	Not Supported	Supported	Not Supported
PAC (802.1x, MAB) - support for both Data/Voice clients	Not Supported	Not Supported	Not Supported	Not Supported	Supported

[a] GNMI is supported on select campus platforms (E3248-PXE, E3248-P, and AS4630-54PE)

1.3 Date of Delivery

The Enterprise SONiC distribution by Broadcom, version 4.1.1 GA release is delivered in July, 2023.

1.4 Items Included in the Delivery

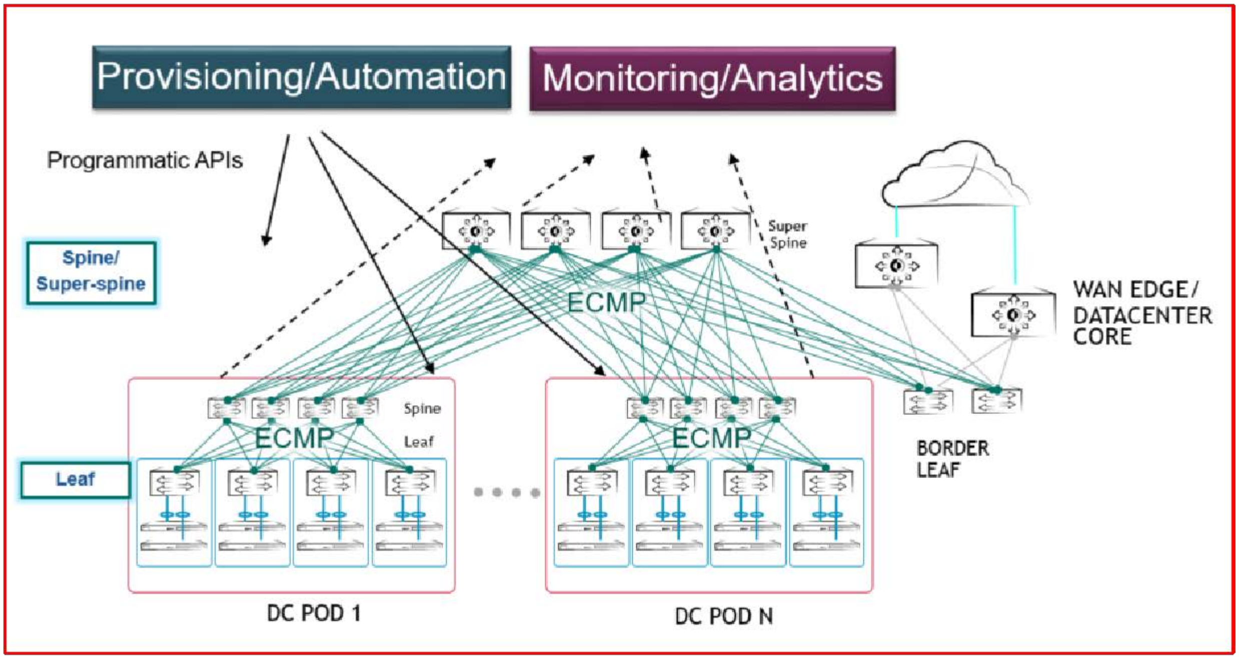
- Binaries:
  - Enterprise SONiC distribution by Broadcom, version 4.1.1 Cloud Base executable and md5sum
  - Enterprise SONiC distribution by Broadcom, version 4.1.1 Cloud Advanced executable and md5sum
  - Enterprise SONiC distribution by Broadcom, version 4.1.1 Enterprise Base executable and md5sum
  - EnterpriseSONiC distribution by Broadcom, version 4.1.1 Enterprise Advanced executable and md5sum
  - Enterprise SONiC distribution by Broadcom, version 4.1.1 Campus executable and md5sum
- Documentation
  - Release Notes (this document)
  - The following Enterprise SONiC 4.1.0 distribution by Broadcom documents are applicable to the SONiC 4.1.1 release as well:
    - User Guide
    - Enterprise SONiC Validated Design
    - gNMI Reference
    - REST Reference
    - IS-CLI Command Reference
    - MIB and OID Guide
    - ConfigDB Manual
    - OSS Attribution Reference Manual
  - The following Enterprise SONiC 4.0.0 distribution by Broadcom documents are applicable to the SONiC 4.1.1 release as well:
    - Product Brief
    - Chef User Guide
    - Ansible User Guide
    - *Inband Flow Analyzer* application note (56870-AN5xx)

2 Customer Use Cases

Data Center L3 CLOS (underlay) use case

The Enterprise SONiC distribution by Broadcom Cloud editions are targeted for Data Center Fabric deployments (Public, Private, and Edge compute). The Enterprise SONiC distribution by Broadcom can be deployed at various Places-In-Network (PIN) - ToR, Leaf, Spine, Super spine, Border Leaf PINs.

The Enterprise SONiC distribution by Broadcom-based data center fabrics can be deployed in an underlay use case for web-scale data center architectures, or for data center PODs in Enterprises or Service Providers for select workloads such as Hadoop which require an underlay network. The Enterprise SONiC distribution by Broadcom can also be used in Enterprises and Service Providers as an underlay data center fabric for VMware based PODs deploying VMware ESX, NSX, vSAN and other VMware solutions.

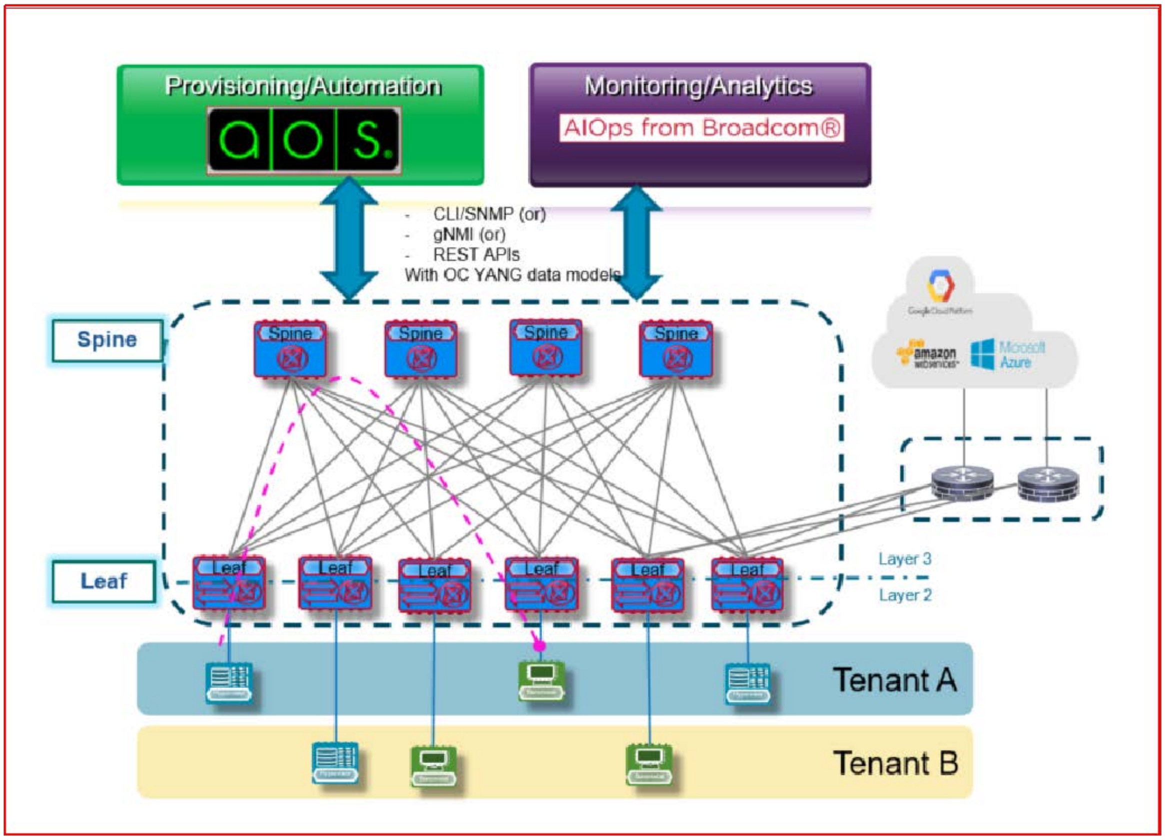


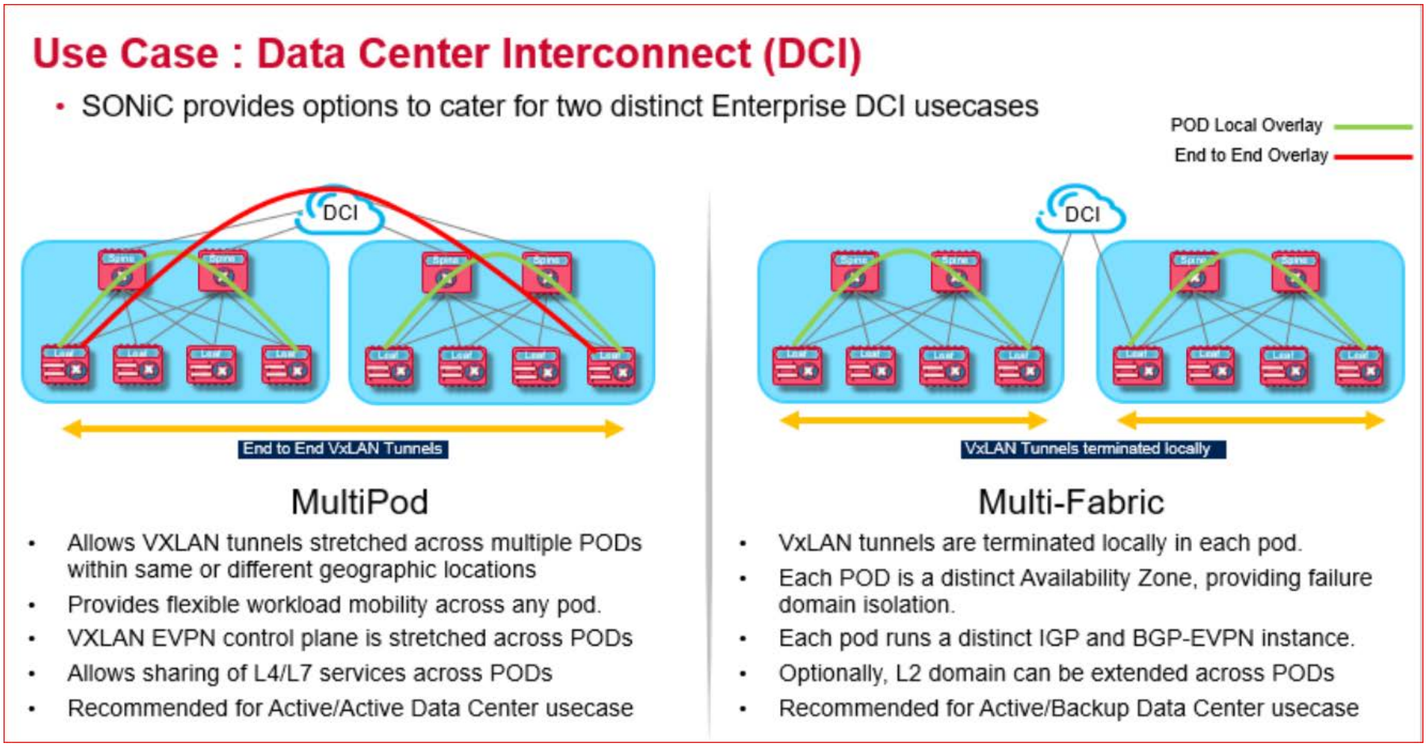
Data Center L3 CLOS Overlay Use case (with VXLAN and BGP-EVPN) and DCI

The Enterprise SONiC distribution by Broadcom can also be deployed in Enterprises or Service Providers for select workloads such as Hadoop, which require an overlay in order to support multi-tenancy.

Using an overlay architecture in the data center allows end users (network admins) to place endpoints (servers or virtual machines) anywhere in the network and remain connected to the same logical Layer 2 (or Layer 3) network, enabling the virtual topology to be decoupled from the physical topology. This decoupling allows the data center network to be programmatically provisioned at a per-tenant level.

Overlay networking generally supports both Layer 2 and Layer 3 transport between servers or VMs. It also supports a much larger scale. SONiC overlay networks use a control-plane protocol (BGP-EVPN) to facilitate learning and sharing of endpoint information, and use VXLAN tunneling protocol to create the data plane for the overlay layer.





Use Case : Data Center Interconnect (DCI)

- SONiC provides options to cater for two distinct Enterprise DCI usecases

MultiPod

- Allows VXLAN tunnels stretched across multiple PODs within same or different geographic locations
- Provides flexible workload mobility across any pod.
- VXLAN EVPN control plane is stretched across PODs
- Allows sharing of L4/L7 services across PODs
- Recommended for Active/Active Data Center usecase

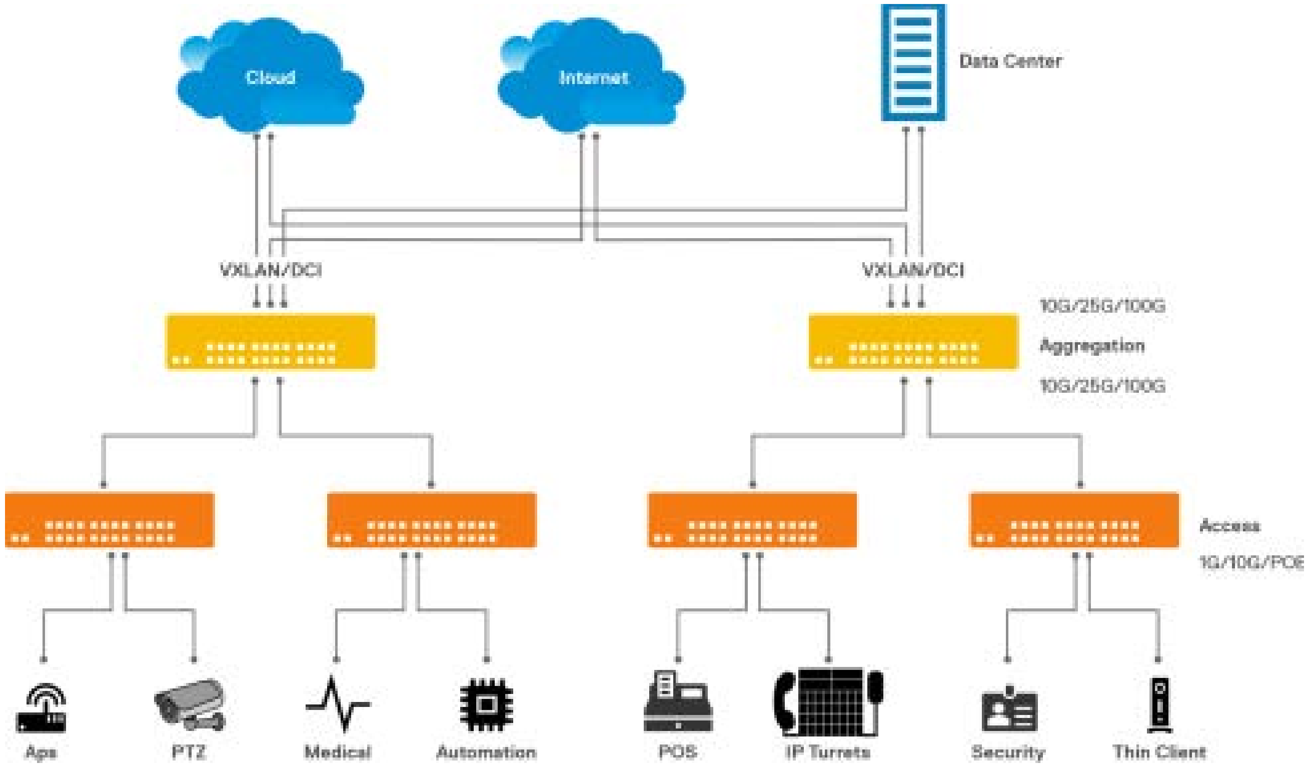
Multi-Fabric

- VxLAN tunnels are terminated locally in each pod.
- Each POD is a distinct Availability Zone, providing failure domain isolation.
- Each pod runs a distinct IGP and BGP-EVPN instance.
- Optionally, L2 domain can be extended across PODs
- Recommended for Active/Backup Data Center usecase

Campus Use case with Access/Aggregation layers that can be connected to existing DCs

Enterprise SONiC can serve as the unified Network Operating System (NOS) that you use to connect edge devices (Campus devices such as POS, thin Clients, Security Cameras, etc as shown in diagram below) and a data center. The Campus bundle extends a DC fabric to remote locations using the same DC NOS. Additionally the Campus package caters to the traditional 3 Tier architecture of Access-Aggregation-Core.

- A two-layer fabric is implemented in which there is an aggregation and an access layer. Customers can use VXLANs to stretch the fabric.
- The CLOS network Leaf/Spine architecture allows for future scaling. Each leaf-layer access switch is connected to each spine-layer aggregation switch in a full-mesh topology.
- In the aggregation layer, VXLAN EVPN supports multi-tenancy and multi-site data center interconnection (DCI).
- Customers can leverage the automation and management tools in the data center to configure and maintain edge switches



3 Supported Platforms

PIN	Platform Name	Port Configuration	ODM/OEM Vendor	Broadcom Silicon	Hardware Revision	BIOS Versions
Management Switch	BES2348T	48 × 10M/100M/1000 RJ45 + 4 × 25G	Alpha Networks	BCM56274 (TD3-X2)	A1 Rev. 00	HRV.05.10.12.0035
Management Switch	SCG60D0-484T	48 × 10M/100M/1000 RJ45 + 4 × 25G	Alpha Networks	BCM56274 (TD3-X2)	A1 Rev. 00	HRV.05.10.12.0035
Management Switch	Belgite (E1070)	48 × 10M/100M/1000 RJ45 + 8 × 10G	Celestica	BCM56277 (TD3-X2)	Rev: 05	COMe-Dnvt.2.02.00
Management Switch	Wistron	48 × 10M/100M/1000 RJ45 + 4 × 25G	Wistron	BCM56275 (TD3-X2)	2	OOB_1.0.8
Management Switch Campus Switch	AS4630-54TE	48 × 1G + 4 × 25G + 2 × 100G	Accton	BCM56371 (TD3-X3)	R01A	v47.01.01.00
Management Switch Campus Switch	AS4630-54PE	48 × 1G + 4 × 25G + 2 × 100G	Accton	BCM56371 (TD3-X3)	R0D	v37.0b.01.00
Management Switch Campus Switch	N3248TE	48 × 1G + 4 × 10G + 2 × 100G	DellEMC	BCM56370 (TD3-X3)	A00 or later	3.45.0.9-3
Leaf	Questone_2A	48 × 25G + 8 × 100G	Celestica	BCM56771 (TD3-X5)	4/Questone-IIA	0.00.15
Spine	Silverstone	32 × 400G	Celestica	BCM56980(TH3)	6	2.0.0
Campus Switch	N3248X	48 × 10G + 4 × 25G + 2 × 100G	DellEMC	BCM56771 (TD3-X5)	A00 or later	3.45.0.9-4
Campus Switch	N3248PXE	48 × 10G + 4 × 25G + 2 × 100G	DellEMC	BCM56771 (TD3-X5)	A00 or later	3.45.0.9-4
Campus Switch	E3248PXE	48 × 10G + 4 × 25G + 2 × 100G	DellEMC	BCM56771 (TD3-X5)	A00 or later	3.57.0.9-3
Campus Switch	E3248P	48 × 1G + 4 × 10G + 2 × 100G	DellEMC	BCM56371 (TD3-X3)	A00 or later	3.57.0.9-3
Management Switch Campus Switch	AS4630-54NPE	36 × 2.5G + 12 × 10G + 4 × 25G + 2 × 100G	Accton	BCM56370	R02A	v47.01.02.00
Leaf	AS7326-56X	48 × 25G + 8 × 100G	Accton	BCM56873 (TD3-X7 2.0T)	R01B, R01G	v36.01.00.01
Leaf	AS7726-32X	32 × 100G	Accton	BCM56870 (TD3-X7 3.2T)	R01B	v36.01.00.02
Leaf	IX8	48 × 25G + 8 × 100G	Quanta	BCM56873 (TD3-X7 2.0T)	C1D	3A06
Leaf	IX8-BWDE	48 × 25G + 8 × 100G	Quanta	BCM56873 (TD3.X7 2.0T)	1/C3D	3A13
Leaf	IX7	32 × 100G	Quanta	BCM56870 (TD3-X7 3.2T)	1/B2A	3A06
Leaf	IX7-BWDE	32 × 100G	Quanta	BCM56870 (TD3-X7 3.2T)	1/B3G	3A13
Leaf	IX8A-BWDE	48 × 25G + 8 × 100G	Quanta	BCM56771 (TD3-X5)	1/B3D	3A13
Leaf	SLX9150-48Y	48 × 25G + 8 × 100G	Extreme Networks	BCM56873 (TD3-X7 2.0T)	B3F	3A10
Leaf	SLX9250-32C	32 × 100G	Extreme Networks	BCM56870 (TD3-X7 3.2T)	A3F	3A10
Leaf	S5296f	96 × 25G + 8 × 100G	DellEMC	BCM56870 (TD3-X7 3.2T)	A00 or later	v3.40.0.9-11
Leaf	S5248f	48 × 25G + 8 × 100G	DellEMC	BCM56870 (TD3-X7 3.2T)	A00 or later	v3.40.0.9-11

PIN	Platform Name	Port Configuration	ODM/OEM Vendor	Broadcom Silicon	Hardware Revision	BIOS Versions
Leaf	S5232f	32 × 100G	DellEMC	BCM56870 (TD3-X7 3.2T)	A00 or later	v3.40.0.9-11
Leaf	Seastone2	32 × 100G	Celestica	BCM56870 (TD3-X7 3.2T)	2	Version 2.19.1266
Leaf/TOR	S5224f	24 × 25G + 4 × 100G	DellEMC	BCM56771 (TD3-X5)	A00 or later	v3.40.0.9-11
Leaf/TOR	S5212f	12 × 25G + 3 × 100G	DellEMC	BCM56771 (TD3-X5)	A00 or later	v3.40.0.9-11
Leaf/TOR	S5448	48 × 100G + 4 × 400G	DellEMC	BCM56780 (TD4-X9)	A00 or later	v3.52.0.D-7
Leaf/TOR	SNJ61D0-320F	32 × 400G	Alpha Networks	BCM56881 (TD4-X11)	01A	05.10.12.0035-03/17/2022
Leaf/Spine	Z9432-O32	32 × 400G	DellEMC	BCM56881 (TD4-X11)	A00 or later	3.51.0.D-6
Spine/TOR	AS5835-54X	48 × 10G + 6 × 100G	Accton	BCM56771 (TD3-X5)	R01B	v37.0b.01.02
Spine/TOR	AS5835-54T	48 × 10G + 6 × 100G	Accton	BCM56771 (TD3-X5)	R01B	v37.0b.01.02
Spine	AS7712-32X	32 × 100G	Accton	BCM56960 (TH)	R01A	v36
Spine	Z9332	32 × 400G	DellEMC	BCM56980 (TH3)	A00 or later	v2.07
Spine	Z9664	64 × 400G	DellEMC	BCM56990 (TH4)	A00 or later	v3.54.0.D-6
Spine	AS9736-64D	64 × 400G	Accton	BCM56990(TH4)	R01A	v50.02.03.00
Spine	Z9264f	64 × 100G	DellEMC	BCM56970 (TH2)	A00 or later	v 3.42.0.9-13
Spine/Superspine	AS7816-64X	64 × 100G	Accton	BCM56970 (TH2)	R01A	v36.01.00.01 v36.01.00.03
Spine/Superspine	AS9716-32D	32 × 400G	Accton	BCM56980 (TH3)	R0BB	v36.01.00.02
Spine/Superspine	IX9	32 × 400G	Quanta	BCM56980 (TH3)	C3C	3A13
Spine/Superspine	IX4	64 × 100G	Quanta	BCM56970 (TH2)	A1A	3A13

4 Newly Introduced and Qualified Software Features

The software features in the following list have been introduced and qualified in Enterprise SONiC distribution by Broadcom, version 4.1.1:

- Support PAC clients on L3 (Routing) VLAN
- Scaling of DHCP snooping entries
- LDAP enhancements:
  - Provide configuration option for `nss_getgrent_skipmembers`. This option specifies whether or not to populate the members list in the group structure for group lookups.
  - Map LDAP groups to SONiC (RBAC) roles

5 Supported Software Features from Previous Releases

The software features in the following list were introduced and qualified in a previous Enterprise SONiC distribution by Broadcom release and are available in the 4.1.1 release:

- System and Platform Infrastructure Features
  - Dynamic Port Breakout
  - DOM Information Display
  - System Locator LED Support (Beacon)
  - CMIS 4.0 Optics Support
  - Hardware Watchdog
  - 1G/10G BASE-T copper BASE-T Support (On select platforms)
  - Multi-rate support on N3248-X, N3248-PXE and E3248-PXE platforms
  - 100 Mbps and 10 Mbps Support on N3248-TE, E3248-P Platforms
  - CoPP (Control Plane Policing)
  - Transceiver Parameter Tuning
  - Third-Party Container Management
  - PDDF and PDK Framework
  - Interface Aliasing (IS-Standard and IS-Standard-Extended Interface Naming)
  - Kdump Support
  - Maintenance Mode
    - LACP Graceful Shut
    - BGP Graceful Shut
    - OSPFv2 Maximum Metrics
  - Multi-Instance Redis DB
  - Hardware Resource Allocation and Reservation
  - Zero Touch Provisioning (ZTP)
  - Auto Negotiation and Link Training
  - Link Statistics Enhancements
  - Link-Down Reason Codes
  - Link Flap Error-Disable
  - Forwarding Plane Drop Counters
  - Time Zone Command Support
  - 2 × 50G Speed Support
  - Broadcom Debug Tool
  - Memory Histogram

  - System Ready for Services and Applications
  - Secure Boot Process and Reference Implementation
  - Syslog High Threshold notifications and clear for CPU/Temperature
  - Per Platform CoPP
  - Interface Beacon LED
  - ZTP Provisioning using a USB Drive
  - Flexible DPB
  - Support Option to Bind the Third Party Container to the Management VRF
  - Limiting CPU/Memory/Disk Usage for Third Party Containers
  - Patching Support in SONiC (Patch Host/Containers)
  - Option to Send Audit Log Messages to Syslog Server
  - Ability to Filter Logs based on Facility and Severity
  - I2C Error Statistics for Accton AS7326 and AS7816 Platforms
  - Half Duplex Support for TD3-X5 - N3248/E3248 Platforms
  - Media AutoFEC for FEC Type automation
- L2 Features:
  - VLAN Auto-state
  - Interface Hold-Down
  - LACP Graceful Shutdown
  - Uplink Tracking
  - L2 MC-LAG
  - L2 VXLAN
  - L2 LVTEP
  - LACP Fast Rate and LACP Fallback
  - Static LAG
  - LLDP
  - UDLD
  - MC-LAG Fallback
  - MC-LAG Graceful Shutdown
  - xSTP over MC-LAG
  - PVST and RPVST+
  - PVST and RPVST+ over MC-LAG
  - DHCP snooping
  - VLAN stacking (QinQ) on TD3 and TD4 platforms
  - Port Channel Min-Links configuration enhancement
- L3 Features:
  - DHCP Relay Enhancements
    - DHCP Relay over VXLAN Overlay Interfaces
    - DHCP Relay Source Interface Selection (e.g. loopback)
    - DHCP Relay over IPv6 Link-Local Interfaces with RFC5549 Routes
    - DHCP Relay Hop Count Configuration
    - DHCP Relay Over IPv4 Unnumbered Interfaces
    - DHCP Relay Option 82, Sub Option 151 VRF Name/ID Option
    - DHCP Relay Option 82, Sub Option 5 Link-Selection Option RFC3527
    - Support for Circuit-Id Formats
    - DHCP Relay Circuit-Id Option
    - DHCP Relay and DHCP snooping support on the same VLAN

- Support for 4K L3 VLAN Interfaces
- Dynamic BGP Neighbor
- IP Fabric over IPv6 underlay RFC5549
- IP Helper
- L3 MC-LAG
- L3 VXLAN
- L3 LVTEP
- Routing Subinterface (on TD3-X5, TD3-X7, TD4-X11 platforms)
- Advertise PIP for both ACT-ACT and ACT-STBY on the Same Leaf Pair
- Route Leaking across VRFs including Management VRF
- BGP Docker Warm Restart
- Avoid Netlink for Handling IPv6 Link-Local Address
- BFD Optimizations to Support 5x100msec Aggressive Timers in SW
- IP SLA (ICMP and TCP tracker)
- IPv4 Unnumbered Interfaces
- RPVST+ over MC-LAG
- RPVST+ Scaling to 3500 VLAN Ports
- Symmetric Hashing
- VXLAN over SVI Interface
- BGP for EVPN (with MLAG)
- BFD IS-CLIs
- BFD with VRF
- VRF support for syslog
- VRF support for SSH.in
- VRRPv3, VRRP/VRRPv3 over VRF
- Management VRF Hardening
- NAT
- OSPFv2
- Multi Site Data Center Interconnect (DCI)
- RIB/FIB Consistency Checker
- Next Hop Group (NHG) Support
- RIF Counters for L3 Interfaces
- BFD Profile
- 4K L3 VLAN Interface Scale for SAG and Unique-IP Cases
- Nexthop resolution using default route
- MC-LAG Peer Gateway
- 1 Million Route Scale
- Route Updates Performance Improvements
- Drop Neighbor Entry to Protect CPU from Unknown IP Packets Hitting the CPU
- CPU Offload for Neighbor Suppression
- CPU Offload for Slowpath ARP Flooding
- OSFPv2 GR
- Router Advertisement KLISH/REST/gNMI Support
  - CLI Commands for RA Retransmission Interval, RADv Disable
  - RFE-8106

- ACL and Flow-Based Services
  - PBR Enhancements for Service Chaining
  - ACL-based CoPP
  - ACL DSCP Map/Remarking
  - ACL Rate Limiting
  - Control Plane ACL
  - Policy-based Routing (IPv4 and IPv6)
  - ACL-based Packet Replication
  - ACL Consistency Checker
- Security Features:
  - RADIUS and TACACS
  - RADIUS/TACACS Password Obfuscation
  - NTP Server and NTP Authentication
  - NTP Prefer Option
  - LDAP Integration
  - AAA Authorization support with TACACS+
- Manageability Features:
  - Industry Standard CLI (IS-CLI)
  - REST and gNMI Interfaces via OpenConfig YANG (OC-YANG)
  - Role-Based Access Control (RBAC)
  - RBAC and HAMd Enhancements
  - SNMP Configuration Traps and OIDs
  - Configuration Services – Chef for EVPN
  - gNMI Subscription Support for Limited YANG Paths (OnChange, Interval, Once, Poll, Target defined)
  - Bulking support in both REST(YANG patch) and gNMI
  - Query parameter for REST and filtering support for gNMI
  - Scalar encoding support for gNMI
  - Support to Read Service Tag via SNMP
  - SNMP Trap Enablement on Interface Instead of Global
- Multicast Features / Enhancements:
  - L3 Multicast with PIM operates on L3 interfaces only
  - IGMP
  - IGMP Snooping (v1, v2, v3) (with MLAG)
  - IPv4 PIM-SSM Support
- Debuggability / Serviceability features
  - In-memory Debug Logging
  - Audit Logging and Syslogs
  - Command to Return Interfaces to the Default Configuration
  - Port Mirroring on Port Channel and VLAN
- Scalability improvements
  - L3 VLAN Scale to 4K (for TD3-X7 based Platforms)
  - Host Table Resource Reservation for Local Hosts
- QoS:
  - RoCE V2 Support
  - DSCP Marking Preservation for VXLAN

  - QoS Map Support for Remarking and SVI
  - BUM Storm Control
  - Port and Priority Shaping
- Telemetry and Instrumentation Features
  - sFlow on Management VRF
  - Inband Flow Analyzer (IFA) 2.0
  - Drop Monitor
  - Tail Stamping
  - BST – Watermarks,
  - Thresholds, and Snapshots
  - Linux PTP (KNETSync)
- SONiC readiness for Campus use case
  - Infrastructure level changes have been made extensively to make sure SONiC can run in Campus platforms (with lower memory 8G)
  - POE, POE+ and POE-bt
  - Port Access Control
    - 802.1X
    - MAC Authentication Bypass (MAB)
    - RADIUS Support
      - Multiple RADIUS Servers
      - RFC 2865 — RADIUS Client
      - RFC 2868 — RADIUS Attributes for Tunnel Protocol Support
      - RFC 3579 — RADIUS Support for EAP
    - Authentication Tiering
    - Downloadable ACLs
    - Dynamic ACL
      - Named ACLs
      - Per-session ACLs
      - Filter-Id
    - Posture Assessment (Redirect ACL/URL, COA Using DAS)
    - Guest VLAN
    - Unauth VLAN
    - Open VLAN
    - Monitor Mode
    - STP
    - Port Default to Access VLAN
    - Aruba Clearpass, Cisco ISE, FreeRadius Interoperability
  - MSTP
  - Port MAC Security
  - LLDP-MED
  - Digital Optical Monitoring (DOM)
  - Time Domain Reflectometry (TDR)
  - EVPN VXLAN (scale is different for VXLAN on Campus platforms, see Section 10, Scalability Matrix)
- Warmboot (see Warm Reboot for details)

6 Limitations, Restrictions, Workarounds, and Guidelines

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
```

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`

Warm Reboot

- 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
- Specially in the scaled configurations users should increase BGP warm-restart timer in order 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 <Values>
  graceful-restart stalepath-time <Values>
  ```

STP Loop Guard

STP loop guard is not supported for PVST.

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.

Additional Limitations and Restrictions

- 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.
- 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.
- BCM56870 (TD3-X7) port groups have a design constraint requiring all ports on the same port group to share the same speed.
- Some Accton platforms, like AS7816-64X, do not support batteries and thus do not retain RTC clock. On a power cycle, the date is reset to the default as set in BIOS. Because the RBAC feature uses date-based certificates for user authentication and authorization, the invalid date causes the REST/gNMI requests to fail when this feature is in use. To overcome this issue, the following workarounds may be used:
  - Configure NTP on the switch/router
  - After every power cycle, set the date on the router/switch using the Linux command `- timedatectl`.
  - Disable the RBAC feature.

- 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
  - For DHCP behavior please refer the NTP HLD SONiC_OC_NTP_HLD.md
- 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
        ```

- 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.
- In L2 profile mode, the VXLAN tunnel feature is supported only on TD3-X5, TD3-X7 and TD4-X9 platforms.
- 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.
- 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.
- 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.
- Speed setting of less than 100G on 400G native ports is not supported on BCM56980 (Tomahawk3) platforms.

- When a RADIUS server assigns a VLAN to a PAC client that is being authenticated, the assigned VLAN must already be created on the switch.
- 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.
- Before running the `show techsupport` command to collect the tech support data, it is recommended to run the `sonic-clear logging` command.

7 SONiC 4.1.1 Image Upgrade and Downgrade Considerations

From	To SONiC 3.1.x	To SONiC 3.2.x	To SONiC 3.4.x	To SONiC 3.5.x	To SONiC 4.0.x	To SONiC 4.1.x
SONiC 3.1.x	Yes	Yes	Yes	Yes	Yes	Yes
SONiC 3.2.x	N/A	Yes	Yes	Yes	Yes	Yes
SONiC 3.4.x	N/A	N/A	Yes	Yes	Yes	Yes
SONiC 3.5.x	N/A	N/A	N/A	Yes	Yes	Yes
SONiC 4.0.x	N/A	N/A	N/A	N/A	Yes	Yes
SONiC 4.1.x	N/A	N/A	N/A	N/A	Yes[a]	Yes

[a] 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.

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.

Migrating FRR Configuration from vtysh CLI to IS-CLI

To migrate all FRR configuration settings (which includes feature configurations such as BGP, BFD, and OSPF) form vtysh CLI to the IS-CLI or REST interface in the SONiC 3.1.0 release, use the following procedure:

1. Save and backup the FRR vtysh configuration – preferably on external storage as a reference configuration to configure equivalent IS-CLIs.
2. Configure the routing config mode as *separated* mode or delete the *split* routing config mode in config_db.json, and then perform a config save followed by a config reload or device reboot.

   For example:

   ```
   admin@sonic:~$ sudo config routing_config_mode separated
   ```
3. When the system is ready, configure all FRR vtysh equivalent CLIs with IS-CLI or with the REST interface.

These procedures are a one-time process to migrate the FRR vtysh configuration to IS-CLI or the REST interface.

Migrating DHCP relay Configuration from 3.0.x

The SONiC 3.1.x release added support for VRF-aware DHCP relay, where a DHCP client and the DHCP server can be in different L3 VRF domains. When upgrading from SONiC 3.0.x to SONiC 3.1.x or later releases, the DHCP relay configuration is migrated to derive the server VRF from the VRF associated with the DHCP client interface. For example, if the DHCP client is in VrfRed, then the DHCP server is also assumed to be reachable in VrfRed. The migrated configuration needs to be reviewed and updated depending on whether the DHCP server is reachable in the client VRF.

The following is an example of a 3.0.x to 3.1.x migrated configuration:

```
interface Vlan100
  ip vrf forwarding Vrf-untrust
  ip dhcp-relay 10.89.0.13 10.97.0.31 vrf Vrf-untrust
```

If the DHCP server is not reachable in `Vrf-untrust`, the migrated configuration must be updated to specify the correct server VRF, for example `Vrf-trust` shown in the following output. The link-select and source interface configuration must also be reviewed and updated depending on the network reachability.

```
interface Vlan100
  ip vrf forwarding Vrf-untrust
  ip dhcp-relay 10.89.0.13 10.97.0.31 vrf Vrf-trust
```

8 Silicon Support and Feature Matrix

Enterprise SONiC distribution by Broadcom, version 4.1.1 is supported on the StrataXGS family of silicon only. StrataDNX™ platforms will be supported in future releases.

Features Supported	Feature Description	TH/TH2	TH3	TD3-X2	TD3-X3	TD3-X5	TD3-X7	TD4-X9	TD4-X11	TH4
Layer2	STP – PVST/RPVST+	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	PVST/RPVST+ over MC-LAG	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No
	Link Aggregation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Link Aggregation Fallback	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	LLDP	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	UDLD	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	IGMP Snooping	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
	ACL based Layer2 Forwarding	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	DHCP snooping	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	QinQ	No	No	No	Yes	Yes	Yes	Yes	Yes	No
Layer3	BGP v4, v6	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	IPv4, v6 Static Routing	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	BFD	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	VRRP	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	ECMP	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	DHCP Relay	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	OSPFv2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	NAT	Yes	No	No	No	Yes	Yes	No	No	No
	L3 IGMP	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes
	PIM SSM	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes
	IP SLA (ICMP/TCP track)	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
	Policy Based Routing	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Service Chaining	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
	ACL based replication	Yes	Yes	No	Yes	Yes	Yes	No	No	No
	L3 VLAN scale	Yes	Yes	—	Yes	Yes	Yes	Yes	Yes	Yes
	Routing Subinterface	No	No	No	No	Yes	Yes	Yes	Yes	No
	Route Maps	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	1M Route Scaling	No	No	No	No	No	No	Yes	Yes[a]	No
ACLs	L2 ACLs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	L3 ACLs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Features Supported	Feature Description	TH/TH2	TH3	TD3-X2	TD3-X3	TD3-X5	TD3-X7	TD4-X9	TD4-X11	TH4
QoS	L2 QoS Maps	Yes	Yes	—	Yes	Yes	Yes	Yes	Yes	Yes
	L3 QoS Maps	Yes	Yes	—	Yes	Yes	Yes	Yes	Yes	Yes
	Queue and buffer size configuration	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Traffic Priority scheduling (Strict, WFQ)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	WRED	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	ECN	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Priority Flow Control	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
	BUM/Storm control	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No
	ACL-based PCP Remarking	Yes	Yes	Yes	Yes	Yes	Yes	Ingress	Ingress	Yes
	ACL-based DSCP	Yes	Yes	Yes	Yes	Yes	Yes	Ingress	Ingress	Yes
	ACL-based Rate Limiting / Policing	Yes	Ingress	Yes	Yes	Yes	Yes	Ingress	Ingress	Ingress
	RoCE V2	No	Yes	No	No	No	Yes	No	No	No
Monitoring	Telnet	No	No	No	No	No	No	No	No	No
	SSH	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	SCP	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	TFTP	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	FTP	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	NTP	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	NTP Server, Authentication	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	SNMP Monitoring	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	ZTP	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	sFlow	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MC-LAG	L2	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No
	L3	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No
	L2 LVTEP	Yes	No	No	Yes	Yes	Yes	Yes	Yes	No
	L3 LVTEP	Yes	No	No	Yes	Yes	Yes	Yes	Yes	No
	Advertise PIP	Yes	No	No	Yes	Yes	Yes	Yes	Yes	No
	IGMP Snooping	Yes	No	No	Yes	Yes	Yes	Yes	Yes	No
	MC-LAG graceful shutdown	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No
	MC-LAG fallback	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No
EVPN	L2 VXLAN	No	No	No	Yes	Yes	Yes	Yes	Yes	No
	L3 VXLAN	No	No	No	Yes	Yes	Yes	Yes	Yes	No
Security	TACACS+	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	RADIUS	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	RBAC	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Silicon Telemetry	Thresholds	Yes	Yes	No	No	Yes	Yes	No	No	No
	Snapshots	Yes	Yes	No	No	Yes	Yes	No	No	No
	Drop Monitor	No	Yes	No	No	No	Yes	No	No	No
	Inband Flow Analyzer (IFA, version 2.0)	No	No	No	No	No	Yes	No	No	No
	Tail Stamping	Yes (TH2)	Yes	No	No	No	Yes	No	No	No

Features Supported	Feature Description	TH/TH2	TH3	TD3-X2	TD3-X3	TD3-X5	TD3-X7	TD4-X9	TD4-X11	TH4
System	Dynamic Port Breakout	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes
	Interface Aliasing	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Display running configuration using IS-CLI	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Third Party Container Management	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes
	Multi Instance REDIS DB	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Debugging	Broadcom Debug Tool	Yes	Yes	No	No	Yes	Yes	No	No	No
Timing	Linux PTP	Yes (TH)	No	No	No	No	No	No	No	No

[a] On select platforms.

9 List of Supported Optics and Media

#	Transceiver/Cable	Type	Speed	Vendor	Part Number	Length	Support Level[a]
1	Transceiver	SFP	1G	Finisar	FTLF1321P1BTL	—	Full
2	Transceiver	SFP	1G	Finisar	FTLF8519P3BNL	—	Full
3	Transceiver	SFP+	10G	Finisar	FTLX8571D3BCL	—	Full
4	Transceiver	SFP+	10G	Finisar	FTLF8524P2BNL	—	Full
5	Transceiver	QSFP+	40G	Avago	AFBR-79EIPZ	—	Full
6	Transceiver	QSFP+	40G	Avago	AFBR-79E4Z	—	Full
7	Transceiver	QSFP+	40G	Finisar	FTL410QE2C	—	Full
8	Transceiver	QSFP+	40G	Finisar	FTL4C1QE1C	—	Full
9	DAC	QSFP+	40G	Amphenol	603020001	1M	Full
10	DAC	QSFP+	40G	Amphenol	603020003	3M	Full
11	DAC	QSFP+	40G	Amphenol	603020005	5M	Full
12	DAC	QSFP+	40G	Amphenol	603020007	7M	Full
13	DAC	QSFP28	100G	Yamachi	CAU120-038-D050	—	Full
14	DAC	QSFP28	100G	Amphenol	NDAAFF-0001	1M	Full
15	DAC	QSFP28	100G	Amphenol	NDAAFF-0003	3M	Full
16	DAC	QSFP28	100G	Amphenol	NDAAFJ-0004	5M	Full
17	DAC	SFP28	25G	Amphenol	NDCCGF-0005	3M	Full
18	DAC	SFP28	25G	Amphenol	NDCCGF-0001	1M	Full
19	DAC	SFP28	25G	Amphenol	NDCCGJ-0005	5M	Full
20	DAC	SFP28	25G	Amphenol	SF-NDCCGF28GB-002M	—	Full
21	DAC-QSFP28_to_Quad_ SFP28 Breakout Cable	QSFP28 → SFP28	100 → 4 × 25G	Amphenol	NDAQGF0003	3M	Full
22	DAC - QSFP28_to_Quad_ SFP28 Breakout Cable	QSFP28 → SFP28	100 → 4 × 25G	Amphenol	NDAQGF-0001	1M	Full
23	DAC - QSFP28_to_Quad_ SFP28 Breakout Cable	QSFP28 → SFP28	100 → 4 × 25G	Amphenol	NDAQGF-0005	5M	Full
24	DAC	QSFP+	40G	Finisar	FCBG410QB1C03	—	Full
25	DAC	SFP+	10G	Amphenol	586710005	10M	Full
26	DAC	SFP+	10G	Amphenol	586710007	1M	Full

#	Transceiver/Cable	Type	Speed	Vendor	Part Number	Length	Support Level[a]
27	DAC	SFP+	10G	Amphenol	586710003	5M	Full
28	DAC	SFP+	10G	Amphenol	586710004	7M	Full
29	40G_to_40G_Fiber_ Patch_Cable	QSFP+	40G	Amphenol	100164	—	Full
30	40G_to_4*10G_Fiber_ Splitter_Cable	QSFP+ → SFP+	40G → 4 × 10G	Molex	CU-3M-QSFP-4SFP10G-C	—	Full
31	40G_to_4*10G_Fiber_ Splitter_Cable	QSFP+ → SFP+	40G → 4 × 10G	Amphenol	100208	—	Full
32	Amphenol_100G_Fiber_ 15M	QSFP28	100G	Amphenol	FOQQD33P00015	—	Full
33	QSFP28_100G_QSFP_ TRANSCEIVER_FOXCONN	QSFP28	100G	Foxconn	AFBR-89BDDZ	—	Full
34	SFP+ Passive DAC	SFP+	10G	Foxconn	CUFCP14-CCF05-EF	3M	Full
35	SFP+ Passive DAC	SFP+	10G	Foxconn	CUFCP13-DCF05-EF	5M	Full
36	SFP28 Passive DAC	SFP28	25G	Foxconn	CUFCP34-CCF05-EF	3M	Full
37	SFP28 Passive DAC	SFP28	25G	Foxconn	CUFCP32-DCF05-EF	5M	Full
38	QSFP+ Fanout Passive DAC	QSFP+ → SFP+	40G → 4 × 10G	Foxconn	CURCP14-CCF05-EF	3M	Full
39	SFP 10G SR	SFP	10G	Brocade	57-0000075-01	—	Limited
40	QSFP 40G 150M	QSFP	40G	Brocade	57-1000128-01	—	Limited
41	QSFP28 SR4 150M	QSFP28	100G	Brocade	57-1000326-01	—	Limited
42	QSFP28 AOC 10M	QSFP28	100G	Brocade	57-1000347-01	—	Limited
43	QSFP 40G LR	QSFP	40G	Brocade	57-1000263-01	—	Limited
44	QSFP28-100G-DAC	QSFP28	100G	FS	QSFP28-100G-DAC	—	Limited
45	QSFP28 - 4xSFP28 DAC 3M	QSFP28 - 4xSFP28	100G	FS	Q28-PC03	—	Limited
46	QSFP28-SR4-100G	QSFP28	100G	FS	QSFP28-SR4-100G	—	Limited
47	QSFP- 4xSFP DAC	QSFP-4xSFP	40G	FS	QSFP-4SFP10G-DAC	—	Limited
48	QSFP-SR4-40G	QSFP	40G	FS	QSFP-SR4-40G	—	Limited
49	QSFP-SR4-40G	QSFP	40G	FS	QSFP-SR4	—	Limited
50	40GE BiDi QSFP+	QSFP-BiDi	40G	FS	QSFP-BD-40G	—	Limited
51	QSFP-DD DR4 400G	QSFP-DD	400G	FOIT	AFCT-91DRDDZ	—	Limited
52	QSFP28 DR1 100G	QSFP28	100G	FOIT	AFCT-89SFDZ	—	Limited
53	QSFPDD 400G DR4	QSFP-DD	400G	AVAGO	AFCT-93DRPHZ-AZ2	—	Limited
54	QSFPDD 400G eDR4	QSFP-DD	400G	Dell-EMC	6MGDY	—	Limited
55	QSFPDD 400G LR4	QSFP-DD	400G	Dell-EMC	KW5H2	—	Limited
56	Transceiver	QSFP28 SR4	100G	Extreme Networks	EQPT1H4SR4UCM100	—	Full
57	DAC	QSFP28DD	400G → 8x25	DELL-EMC	DAC-Q28DD-8S28-25G-1M	1M	Full
58	DAC	QSFP28DD	400G → 8x25	DELL-EMC	AOC-Q28DD-8S28-25G-7M	7M	Full
59	Transceiver	40G-PSM4	40G → 4x10G	DELL-EMC	QSFP28 100GBASE-PSM4-DUALRATE	—	Limited
60	Transceiver	400G-SR4.2	400G	DELL-EMC	Q56-DD SR4.2	—	Limited
61	Transceiver	SFP	1G	ProLabs	SFP-1000BASE-SX-C	—	Full
62	DAC	SFP+	10G	ProLabs	470-AAVH-C	1M	Full
63	Transceiver	SFP+	10G	Edgecore Networks	ET5402-SR4	—	Full

#	Transceiver/Cable	Type	Speed	Vendor	Part Number	Length	Support Level[a]
64	Transceiver	QSFP+	40G	Edgecore Networks	ET8401-SR4	—	Full
65	DAC	SFP28	25G	10G Tek	CAB-ZSP/ZSP-P2M	2M	Full
66	Transceiver	QSFP28	100G	Finisar	FTLC1157RGPL-1Y	—	Full
67	Transceiver	QSFPDD	400G	Cisco-pre (Precision)	QSFP56DD-DR4+-C	—	Full
68	Transceiver	QSFPDD	400G	Precision	QSFP56DD-DR4+-D3	—	Full
69	DAC	QSFP28	100G	FS	Q28-PC005	.5M	Full
70	Transceiver	QSFP28	100G	FS	QSFP28-BLR4-100G (DE)	—	Full
71	Transceiver	QSFP	40G	FS	QSFP-LR4-40G (DE)	—	Full
72	Transceiver	QSFP+	40G	DELL-EMC	XW7J0	—	Full
73	Transceiver	QSFPDD	100G	Amphenol	NDYYYF-0006 Rev H	.5M	Full
74	DAC	QSFP28	100G	Amphenol	NDAAFF-0004 Rev V	.5M	Full

[a] Full indicates Broadcom has tested these optics. Limited indicates these optics were used in Broadcom labs for SONiC testing.

10 Scalability Matrix

SONiC 4.1.1 Scaling Numbers	TH	TH2	TD3-X3	TD3-X5	TD3-X7	TH3	TD4-X11	TD4-X9	TH4
Number of port channels per system	128	128	128	128	128	128	128	128	128
Number of port channel members per system	128	128	128	128	128	128	128	128	128
Maximum number of port channel members per port channel.	32	32	32	32	32	32	32	32	32
Maximum number of PVST instances	510	510	510	510	510	62	254	254	62
Maximum number of RPVST instances	510	510	510	510	510	62	254	254	62
Maximum number of RPVST VLAN ports	5000	5000	1000	5000	5000	5000	5000	5000	5000
Number of VLAN members per system	24K	24K	24K (L2 profile) 4K (L3 profile)	24K (L2 profile) 12K (L3 profile)	24K (L2 profile) 12K (L3 profile)	24K	15K (with VLAN stacking) 24K(without VLAN stacking)	15K (with VLAN stacking) 24K(without VLAN stacking)	24K
Maximum number of VLAN members per VLAN	Up to maximum number of system ports + port channel								
Maximum number of SAG and MC LAG unique IP VLAN L3 interfaces (SVI) per system	1K	1K	4K	4K	4K	512	4K	4K	512
Maximum number of VLAN L3 interfaces (SVI) per system (**if not** SAG and MC LAG unique IP interfaces)	4K	4K	4K	4K	4K	4K	4K	4K	4K
Maximum number L3 interfaces on physical, port channel and subinterface ports supported per system.	1K	1K	1K	1K	1K	512	1K	1K	512
Maximum number of MAC per system	40K/136K	40K/256K	16K	64K (L3 profile) 224K (L2 profile)	64K (L3 profile) 224K (L2 profile)	8K/8K	32K	32K (L3 profile) 224K (L2 profile)	8K
Maximum number of VRF per system	1000	1000	128	1000	1000	512	1000	1000	1000
Maximum number of IPv4 routes per system (default mode)	65K	196K	16K	81K	81K	196K	1M	1M	1M
Maximum number of IPv6 routes per system (default mode)	24K (≤64b), 14K (>64b)	32K (≤64b), 25K (>64b)	8K (≤64b), 4K (>64b)	32K (≤64b), 25K (>64b)	32K (≤64b), 25K (>64b)	32K (≤64b), 25K (>64b)	1M 1M	1M 1M	1M 1M

SONiC 4.1.1 Scaling Numbers	**TH**	**TH2**	**TD3-X3**	**TD3-X5**	**TD3-X7**	**TH3**	**TD4-X11**	**TD4-X9**	**TH4**
Maximum number of IPv4 routes per system (route scale max mode)	N/A	280K	NA	160K	160K	N/A	N/A	NA	NA
Maximum number of IPv6 routes per system (route scale max mode)	N/A	32K (≤64b, >64b)	NA	65K (≤64b, >64b)	65K (≤64b, >64b)	N/A	N/A	NA	NA
Maximum number of static routes per system	2K	2K	1024	1024	2K	2K	2K	2K	2K
Maximum number of BGP routes per system	1M	1M	8158	256K	1M	1M	1M	1M	1M
Maximum number of BGP routes RIB IN.	1M	1M	8158	256K	1M	1M	1M	1M	1M
Maximum number of BGP neighbors	512	512	128	256	512	512	512	512	512
Maximum number of paths in an ECMP group	128	128	16	128	128	128	128	128	128
Maximum number of BFD sessions	128	128	64	64	128	128	128	128	128
IP SLA - ICMP tracker	50	50	25	50	50	50	50	50	50
IP SLA - TCP tracker	50	50	25	50	50	50	50	50	50
Maximum number of ARP supported per system	32K	32K	6144	16K	32K	16K	30,000	30,000	16K
Maximum number of ND supported per system	16K	16K	2560	8K	16K	8K	15,000	15,000	8K
Maximum number of OSPF routers per OSPF area	50	50	50	50	50	50	50	50	50
Maximum number of OSPF enabled L3 interfaces	128	128	64	128	128	128	128	128	128
Maximum number of OSPFv2 neighbors	128	128	64	128	128	128	128	128	128
Maximum number of OSPF **intra** area routes	5000	5000	2500	5000	5000	5000	5000	5000	5000
Maximum number of **inter** area summary routes	5000	5000	2500	5000	5000	5000	5000	5000	5000
Maximum number of OSPF external (type-5) routes	40,000	40,000	8158	20000	40,000	40,000	40,000	40,000	40,000
Maximum number of VRRP instances	128	128	64	128	128	128	128	128	128

SONiC 4.1.1 Scaling Numbers	TH	TH2	TD3-X3	TD3-X5	TD3-X7	TH3	TD4-X11	TD4-X9	TH4
Maximum number of VRRP enabled interfaces	128	128	64	128	128	128	128	128	128
Maximum number of VRRP instances per interface	16	16	8	16	16	16	16	16	16
Maximum number of tracked interfaces per VRRP Instance	8	8	8	8	8	8	8	8	8
Maximum IP address per VRRP instance	4	4	4	4	4	4	4	4	4
Maximum number of Ingress MAC only ACL tables/features applied to interface	3	3	4	3	3	2	2[a]	2[a]	2
Maximum number of MAC ACL rules per applied ACL table/feature at ingress	256	256	512	768	768	256[b]	1023	1023	255
Maximum number of ingress IPv4 only ACL tables/features applied to interface	3	3	4	3	3	2	2[a]	2[a]	2
Maximum number of IPv4 ACL rules per applied ACL table/feature at ingress	256	256	512	768	768	256[b]	1023	1023	255
Maximum number of Ingress IPv6 Only ACL Tables/Features Applied to Interface	1	1	4	1	1	1	2[a]	2[a]	2
Number of IPv6 ACL Rules Per Applied ACL Table/Feature at Ingress	256	256	512	768	768	256[b]	1023	1023	255
Maximum number of Egress MAC Only ACL Tables/Features Applied to Interface	1	1	1	1	1	1	0	0	1
Maximum number of MAC ACL Rules Per Applied ACL Table/Feature at Egress	256	256	512	512	512	128	0	0	127
Maximum number of Egress IPv4 Only ACL Tables/Features Applied to Interface	1	1	1	1	1	1	1	1	1
Maximum number of IPv4 ACL Rules Per Applied ACL Table/Feature at Egress	256	256	512	512	512	128	511	511	127

SONiC 4.1.1 Scaling Numbers	TH	TH2	TD3-X3	TD3-X5	TD3-X7	TH3	TD4-X11	TD4-X9	TH4
Maximum number of Egress IPv6 Only ACL Tables/Features Applied to Interface	1	1	1	1	1	1	1	1	1
Number of IPv6 ACL rules per applied ACL table/feature at egress	256	256	256	512	512	128	511	511	127
Maximum number of NAT entries (static + dynamic)	1024	1024	0	1024	1024	0	0	0	0
Maximum number of port mirroring sessions per system (both ingress and egress).	4	4	4	4	4	4	4	4	4
Maximum number of ERSPAN mirroring sessions per system (both ingress and egress).	4	4	4	4	4	4	4	4	4
Maximum number of BGP-EVPN sessions	256	256	64	128	256	256	256	256	256
Maximum number of EVPN MAC addresses	Not Supported	Not Supported	16K	40K/288K	40K/288K	Not Supported	40K/288K	40K/288K	Not supported
Maximum number of VXLAN tunnels	Not Supported	Not Supported	256 (Without PIP) 128(With PIP)	512 (Without PIP) 256 (With PIP)	512 (Without PIP) 256 (With PIP)	Not Supported	512 (Without PIP) 256 (With PIP)	512 (Without PIP) 256 (With PIP)	Not supported
Maximum number of VNIs – L2VNI	Not Supported	Not Supported	4K	4K	4K	Not Supported	4K	4K	Not supported
Maximum downstream VNI encap entries	Not Supported	Not Supported	Not Supported	4K (without PIP) 2K (with PIP)	4K (without PIP) 2K (with PIP)	Not Supported	4K (without PIP) 2K (with PIP)	4K (without PIP) 2K (with PIP)	Not supported
VRF (L3 VNI)	Not Supported	Not Supported	128	1K	1K	Not Supported	1K	1K	Not supported
Overlay routes	Not Supported	Not Supported	IPv4 – 8K, IPv6 – 4K	IPv4 – 81K, IPv6 – 25K	IPv4 – 81K, IPv6 – 25K	Not Supported	IPv4 – 81K, IPv6 – 25K	IPv4 – 81K, IPv6 – 25K	Not supported
Maximum number of IPMC (L3) forwarding entries	8K	8K	2K	8K	8K	512	8K	8K	512
Maximum number of flow groups for silicon telemetry features	Not Supported	253	Not Supported	Not Supported	253	253	Not Supported	Not Supported	Not Supported
Maximum number of IFA sessions	Not Supported	Not Supported	Not Supported	Not Supported	249	Not Supported	Not Supported	Not Supported	Not Supported

SONiC 4.1.1 Scaling Numbers	TH	TH2	TD3-X3	TD3-X5	TD3-X7	TH3	TD4-X11	TD4-X9	TH4
Maximum number of drop monitor sessions	Not Supported	Not Supported	Not Supported	Not Supported	253	253	Not Supported	Not Supported	Not Supported
Maximum number of Tail stamping sessions	Not Supported	253	Not Supported	Not Supported	253	Not Supported	Not Supported	Not Supported	Not Supported
Maximum number of IGMP snooping entries	512	512	512	512	512	512	512	512	512
Maximum number of DHCPv4 snooping entries (dynamic/static)	12K/1K	12K/1K	12K/1K	12K/1K	12K/1K	12K/1K	12K/1K	12K/1K	12K/1K
Max Number of DHCPv6 Snooping entries (dynamic/static)	12K/1K	12K/1K	12K/1K	12K/1K	12K/1K	12K/1K	12K/1K	12K/1K	12K/1K
PAC: Maximum clients (802.1x/MAB) supported on the switch	N/A	N/A	512	512	N/A	N/A	N/A	N/A	N/A
PAC: Maximum clients (802.1x/MAB) supported per-port	N/A	N/A	48	48	N/A	N/A	N/A	N/A	N/A
PAC: Maximum ACL rules supported per (802.1x/MAB) client	N/A	N/A	60	60	N/A	N/A	N/A	N/A	N/A

a. For TD4, IP Helper/DHCP relay and Ingress ACL-based QoS cannot be enabled at the same time.

b. For TH3, one ACL table can be 512 entries while other can be at 256. However, it is not deterministic which table will be size 512 when tables are added and deleted; hence, it is marked as 256.

NOTE: The actual scale numbers may be lowered for select packages.

CoPP Parameters

Table 1: CoPP Parameters and Scaling

Protocol	CPU Queue	Scheduling Weight	Rate Limit (PPS)
EAPOL	24	1	500
LACP	23	10	500
UDLD	22	10	500
STP, PVRST	21	30	16000
BFD, BFDv6	20	10	1500
PTP	19	3	5000
LLDP	18	2	500
VRRP, VRRPv6	17	2	500
ICCP	16	4	5000
OSPF	15	4	5000
BGP, BGPv6	14	6	10000
PIM	13	2	5000
IGMP Query	12	2	2000
ARP Suppress, ND Suppress	11	2	3000
ARP Req, AP Resp, Neighbor Discovery	10	2	3000
DHCP, DHCPv6	9	2	300
ICMP	8	2	1000
IP2ME	7	2	6000
Subnet	6	2	2000
Source NAT Miss, Dest NAT Miss	5	1	300
L3 MTU Error	4	1	500
Sample Packet (sflow)	3	1	8000
TTL	0	1	100
DEFAULT	0	1	100

Priority queues are serviced in a weighted round-robin manner, with packets taken from each receive queue according to their relative weight before moving to the next priority queue. In this release, the per-platform CoPP feature introduces default rate limits and scheduling weights that are tuned for the platform's CPU and switch ASIC capabilities. The platform's CoPP values can differ from the values in Table 1.

10.1 Guidelines for MAC Aging Timer Settings for Scaled ARP Hosts with Unidirectional Traffic

This section describes the Broadcom recommendations for ARP scaling with unidirectional traffic in the SONiC 4.1.1 release.

Based on your ARP scaling requirements with unidirectional traffic, configure the corresponding MAC aging timer value according to the guidelines in the following table. Using these values avoids issues with MAC age-out and traffic loss.

By default, the MAC aging timer is set to 10 minutes, and the ARP aging timer is set to 30 minutes. The MAC aging timer and ARP aging timer are both configurable

ARP Scale	MAC Aging Timer (Minimum)	ARP Aging Timer (Minimum)
2000	10 (default)	30 (default)
4000	20	30 (default)
6000	30	60

NOTE:

- To further increase ARP scaling (>10K), a step-up in MAC aging timer value is also required.
- The scale numbers described in this section are not applicable to bidirectional traffic. No specific MAC aging timer settings are necessary to avoid traffic loss or MAC age-out issues.

10.2 Guidelines for Scale Limitation and CoPP Setting for Trident4

For Trident4 based platforms, even though the HW support for MAC and ARP entries are large, in a multi-D scaling environment below is the limit

Entities	Scale in Hardware	Recommended in Mult-D Scale
ARP	30000	18000
MAC	32000	18000
ND	15000	8000

With scaled configs (ARP, MAC, BFD), it is recommended that CoPP is configured more aggressively to protect CPU control traffic on TD4-based platforms, such as the Dell Z9432f. Note that lower rates will impact ARP learning performance.

- copp-system-suppress: 1000 pps
- copp-system-arp: 1000 pps
- copp-system-ip2me: 1000 pps
- copp-system-subnet: 1000 pps
- copp-system-sflow: 1000 pps

Protocol	CPU Queue	Scheduling Weight	Rate Limit (PPS)
ARP Suppress, ND Suppress	11	2	1000
ARP Req, AP Resp, Neighbor Discovery	10	2	1000
IP2ME	7	2	1000
Subnet	6	2	1000
Sample Packet (sflow)	3	1	1000

10.3 Guidelines for BFD Interval Configuration

On devices with slow CPUs like AS4630, N3248, Dell Z9432f, Dell S5232f platforms, we recommend using default BFD timers i.e. 300ms BFD interval and 3 as multiplier. Configuring aggressive BFD intervals (less than 300ms) can cause BFD sessions to go down.

10.4 “show CRM resources” Enhancement for IPv4 and IPv6 Routes

The changes this section describes are for enhancing `show crm` commands to reflect the maximum IPv4 or IPv6 route resources available in hardware. These changes are applicable to TH, TH2, TH3, TD2, TD3, and TD3-X5 platforms. For the TD4 and TH4 platforms, these resources are already taken from statistical route projections. So, the `show crm` changes are not applicable to TD4 and TH4 platforms. The changes have been done as part of defect SONIC-36154. The commands modified are as follows:

- `show crm resources ipv4 route`
- `show crm resources ipv6 route`
- `show crm resources all`

Problem: In the current code, while displaying the available IPv4 and IPv6 route count, the hardware fetches the minimum guaranteed routes from the available LPM and ALPM resources. The LPM resources are shared between IPv4 and IPv6 routes. Also, the introduction of ALPM to increase the maximum number of L3 prefixes supported makes it difficult to indicate the actual number of routes that are supported on the platform. This is because the maximum route count now depends on various parameters. For example, the maximum number of IPv4 prefixes depends on how many IPv6 prefixes are already present (because the banks are shared). Similarly, the count also depends on whether the IPv6 prefix length is 64 or 128. In the ALPM case, it also depends on the actual routes present because that information affects the packing of the routes inside LPM and ALPM. The minGuranteed value indicates the worst-case scenario for IPv4 and IPv6 route count.

The output below is from a system with the following build and platform:

- Build: sonic_3.3.0_daily_210629_0400_227
- Platform: x86_64-accton_as7326_56x-r0 (Acton TD3)

```
sonic# show crm resources all | no-more
Resource Name           Used Count    Available Count
--------------------  ------------  -----------------
ipv4_route                       1              32751
ipv6_route                       3              10238
```

This output suggests that it is possible to program only 32K IPv4 routes into the Acton-TD3 hardware. However, it is possible to program a much higher IPv4 route count. This count can be further increased based on the route-scale profile.

Solution: To show the realistic available route count, the the maximum IPv4/IPv6 route count is now fetched from the hardware. Thereafter, the available count (after subtracting the used IPv4/IPv6 resources) is indicated as follows:

```
sonic# show crm resources ipv4 route
Resource Name           Used Count    Available Count
--------------------  ------------  -----------------
ipv4_route                       3              98254

sonic# show crm resources ipv6 route
Resource Name           Used Count    Available Count
--------------------  ------------  -----------------
ipv6_route                       3              32766
```

The values in the preceding output are best case possibilities (unidimensional). The IPv4 route resources are shared with IPv6 route resources. So, if the IPv6 routes are present and the maximum available IPv4 routes are sent to the hardware, the TABLE_FULL scenario can occur before the maximum available IPv4 route count is reached.

In CRM, the low threshold watermark is 70% of the available route count, and the high threshold watermark is 85% of the available count. It is recommended not to exceed the high threshold watermark for IPv4 and IPv6 routes in unidimensional scenarios. For example, in TD3, the available IPv4 route count is 98K. The low threshold is 68K, and high threshold is 83K for this platform.

NOTE:

- In CRM, the default polling interval is 300 seconds. When the system comes up for the first time, it is necessary to wait 300 seconds to see the available resources. Any change in the used hardware resources will be visible in the `show` command only after 300 seconds, even though the changes may have occurred much earlier. The polling-interval is configurable.

  ```
  sonic(config)# crm polling interval
  <0..9999>  Seconds
  ```

- While programming the IPv4/IPv6 routes, on exceeding the high threshold, syslog is generated as follows.

  ```
  root@sonic:~# tail -f /var/log/syslog | grep 'checkCrmThresholds\|SAI_STATUS_TABLE_FULL'

  Jul 05 12:20:29.889642 2021 sonic WARNING swss#orchagent: :- checkCrmThresholds: IPV6_ROUTE
  THRESHOLD_EXCEEDED for TH_PERCENTAGE 85% Used count 56005 free count 9795

  Jul 05 12:25:29.834589 2021 sonic WARNING swss#orchagent: :- checkCrmThresholds: IPV6_ROUTE
  THRESHOLD_EXCEEDED for TH_PERCENTAGE 85% Used count 56005 free count 9795
  ```

  Only 10 instances of the above syslog are generated. Within this duration, the user is expected to reduce the programmed route count to below the low threshold watermark.

10.5 Guideline for Configuring “radv enable” for IPv6 Router Advertisement

The `radv enable` command is provided only for backward compatibility with community SONiC (radv docker) and is disabled by default. This command is not required for the switch to send router advertisements. It is strongly recommended not to enable the `radv enable` command. If you use this command to configure the switch, you must save the switch configuration and reload the switch.

11 Deprecation of CLI Commands

1. The `buffer_pool` CLI command is deprecated in SONiC Release 4.1.0. It is advised to use the `buffer-pool` CLI command in place of the `buffer_pool` CLI command.
2. The `passive-interface Ethernet <number> non-passive` CLI command under the `router-ospf` configuration mode is deprecated in SONiC Release 4.1.0. The `no passive-interface Ethernet  <number>` is command is the equivalent CLI command that can be used.

12 Open Known Defects

1. **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

2. **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.

3. **Defect ID**: SONIC-43354

   Component: QoS-ACL

   Customer Probability: Medium

   Customer Symptom: DSCP value of payload isn't modified as per the ingress ACL configuration. DSCP value of the outer header isn't the same as the configured value of PIPE mode DSCP.

   Customer Condition: VXLAN tunnel origination device receiving payload with IP header and qos-mode configured as PIPE with a DSCP value. Ingress ACL configured on the access side interface to remark the payload DSCP value. Issue applicable to TD4 devices only.

   Customer Workaround: None

4. **Defect ID**: SONIC-61318

   Component: SNMP

   Customer Probability: High

   Customer Symptom: SNMPv3 walk may fail

   Customer Condition: The customer has configured Engine ID on a Dell EMC S52xx series platform, and running SONiC 3.5.1 or earlier versions and migrating to 4.x.x.

   Customer Workaround: Configure the Engine ID before upgrading to the SONiC 4.x.x release using the following steps, or remove and reapply the SNMPv3 configuration after the upgrade.

   Follow these steps to configure the Engine ID before upgrade to SONiC 4.x.x:

   a. Take Engine ID from the snmpd.conf file:

      ```
      admin@sonic:~$ show run snmp | grep exactEngineID exactEngineID
      0x80000137030c29efd7db21 admin@sonic:~$
      ```

   b. In KLISH CLI, use the following command to configure the Engine ID:

      ```
      snmp-server engine 80:00:01:37:03:0c:29:ef:d7:db:21
      ```

5. **Defect ID**: SONIC-67258

   Component: QoS-ACL

   Customer Probability: High

   Customer Symptom: Egress QoS policy-map application will be inactive. Egress ACL based QoS action like Policer, PCP remarking, DSCP remarking etc will not work at egress on TD4 devices.

   Customer Condition: QoS policy-map applied at the egress.

   Customer Workaround: None

6. **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

7. **Defect ID**: SONIC-73670

   Component: EVPN L2 VxLAN

   Customer Probability: Low

   Customer Symptom: Sometimes on performing warm-reboot data traffic loss of few packets can occur during warm-boot. The packet loss will stop after warm-boot completes.

   Customer Condition: Warm-boot performed on a device that has IP routes and active data traffic flows using these routes for forwarding

   Customer Workaround: There is no work around, however the packet loss may be seen only sometimes and will only for few packets

8. **Defect ID**: SONIC-73585

   Component: L3 - data plane

   Customer Probability: Medium

   Customer Symptom: After configuring a new switch-resource route-scale routes profile or switch-resource route-scale hosts profile, if config save followed by config reload is done instead of cold reboot, the GNMI get 'state' container fields will not reflect the right hardware profile configured on the device.

   Customer Condition: Customer should configure a new route-scale profile different from existing one on the device. Save the config and do a config reload.

   Customer Workaround: Klish cli 'show switch-resource route-scale' will give details about the hardware profile configured.

9. **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

10. **Defect ID**: SONIC-70088

    Component: L3 protocols - OSPF

    Customer Probability: Medium

    Customer Symptom: OSPF Virtual Link configured through REST or GNMI interface would not come up

    Customer Condition: Configuring OSPF Virtual Links through REST or GNMI interface.

    Customer Workaround: KLISH CLI interface for configuring OSPF Virtual Links.

11. **Defect ID**: SONIC-67390

    Component: SAI

    Customer Probability: Medium

    Customer Symptom: Following ERR message would be observed while collecting tech-support on TD4 based devices

    ```
    ERR syncd#syncd: :- processGetStatsEvent: Failed to get stats
    ```

    Customer Condition: Issue would be observed while collecting tech-support on TD4 based devices

    Customer Workaround: None - No functional impact

12. **Defect ID**: SONIC-66597

    Component: IP Forwarding

    Customer Probability: Low

    Customer Symptom: When Multi-hop BFD sessions over ECMP paths are configured, if more than one ECMP link goes down within a second, BFD sessions can flap.

    Customer Condition: Multi-hop BFD sessions over ECMP paths that have more than two links should be configured with default BFD timers (3 × 300 msec)

    - More than two ECMP links for BFD session nexthop should go down within a second

    Customer Workaround: Configure BFD receive or transmit intervals as 500ms and BFD detect multiplier as 5

13. **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

14. **Defect ID**: SONIC-61250

    Component: COPP

    Customer Probability: High

    Customer Symptom: On TD4, CPU queue pkt stats are not accurate for sFlow, MTU, NAT, and default traffic (Q0) queues.

    Customer Condition: When configure sFlow or MTU trap and configure associated policer in the associated trap group.

    When change associated trap group.

    When change trap group policer ID.

    When modify trap group policer rate.

    Customer Workaround: Use the KNET driver debug rx-queue stats for egress based traps.

15. **Defect ID**: SONIC-72065

    Component: EVPN L2 VxLAN

    Customer Probability: High

    Customer Symptom:

    – On TD4 based platforms, When a VXLAN frame carrying a VLAN Tag gets forwarded from internal to external tunnel, the VLAN Tag PCP gets overwritten
    – On TD3 based platforms the VLAN Tag is removed when traffic is forwarded from internal to external tunnel.

    Customer Condition: The Issue is seen in the Border Leaf when Traffic is L2 forwarded between internal and external VXLAN tunnels.

    Customer Workaround: None

16. **Defect ID**: SONIC-67389

    Component: L3 protocols - OSPF

    Customer Probability: Medium

    Customer Condition: OSPF configured with area-id related commands (which have area id as input), and all the related area commands are deleted.

    Customer Symptom: show running-config still shows area-id related configuration. No functional impact.

    Customer Workaround: Manually delete area-id related configurations under router OSPF once all the related area commands are deleted.

17. **Defect ID**: SONIC-71883

    Component: System

    Customer Probability: High

    Customer Condition: Copy/paste operation of a large set of commands using serial console session on old revision of DellEMC Z9664 platforms (rev X01).

    Customer Symptom: Random character drops during bulk command copy/paste resulting in command execution failures.

    Customer Workaround: Use SSH or use newer revision HW (X02 or newer).

18. **Defect ID**: SONIC-71339

    Component: System

    Customer Probability: High

    Customer Condition: When port speed is changed between 25G and 10G using port-group command for the ports connected with 5m DAC cable. The issue is only applicable to 25GBASE-CR-DAC-5.0M media when connected to DellEMC5248 or DellEMC5296.

    Customer Symptom: Link comes up with delay when the speed of the port is toggled from 25G to 10G and back to 25G.

    Customer Workaround: None.

Revision History

SONiC-RN411; July 24, 2023

Release Notes for Enterprise SONiC Distribution by Broadcom, Version 4.1.1.

BROADCOM®