Intel Verified Reference Configuration for Network Function Virtualization Infrastructure (NFVI) v4 and Secure Access Service Edge on Red Hat* OpenShift Container Platform (RHOCP*)

Intel Accelerated Solution

Authors
Jonathan Tsai, Timothy Miskell, Ai Bee Lim

Key Contributors
David Lu, Georgii Tkachuk

Revision 1.0

March 2023

Document Number: 777240

Contents

Figures

Tables

Revision History

Document Number	Revision Number	Description	Revision Date
TBD	1.0	Initial release	March 2023

§

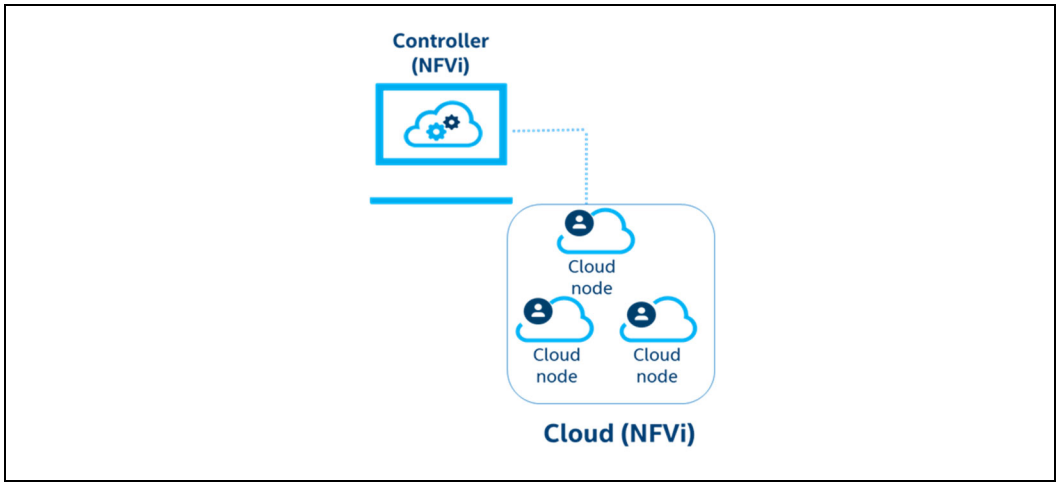
1 Introduction

Intel Accelerated Solutions are a family of workload-optimized, infrastructure solutions, based on the Intel® Xeon® Scalable processor family targeting complex workloads of today. This document describes a reference implementation for the 4th Gen Intel® Xeon® Scalable processor family.

When network operators, service providers, cloud service providers, or enterprise infrastructure companies choose an Intel Verified Reference Configuration (VRC) for Network Function Virtualization Infrastructure (NFVI) deployment based upon an Intel® Xeon® Processor Scalable Family-based Intel VRC, they should be able to deploy various network function virtualized applications more securely and efficiently than ever before. End users spend less time, effort, and expense evaluating hardware and software options. Intel VRCs help end users simplify design choices by bundling hardware and software pieces together while making the high performance more predictable.

Intel VRCs for NFVI v4 are based on a multi-node architecture that consists of a controller, a cloud node and storage at a minimum. Thus, it is expected that the Intel VRC for NFVI provides all required resources to implement a software-defined infrastructure that resides within each cloud server instance and is controlled by the hypervisor. The Controller Node is intended to be used for control, signaling and management implementing the Virtual/Container Network Function (VNF) Management (VNFM) and Virtualization Infrastructure Management (VIM). Thus it may not require additional local storage and hardware acceleration.

Figure 1. Intel® Verified Reference Configurations for NFVI Environment



All Intel VRCs feature a workload-optimized stack tuned to take full advantage of an Intel® Architecture (IA) foundation. To be certified as an Intel VRC, Original Equipment Manufacturer (OEM) systems must meet a performance threshold that represents a premium customer experience.

There are two configurations for Intel VRCs for NFVI reference designs for the Cloud Node:

- Intel VRCs for NFVI Plus configuration for the Cloud Node is defined with at least a 32-core 4th Generation Intel® Xeon® Scalable processor and high-performance network, with storage and integrated platform acceleration products from Intel® for maximum virtual machine density.
- Intel VRCs for NFVI Base configuration for the Cloud Node is defined with a 24-core or higher 4th Generation Intel® Xeon® Scalable processor and network, with storage and add-in platform acceleration products from Intel® targeting for optimized value and performance-based solutions.

There is also a configuration for Intel VRCs for NFVI reference designs for the Controller Node:

Intel VRCs for NFVI configuration for the Controller Node for NFVI is defined with at least a 24-core or higher 4th Generation Intel® Xeon® Scalable processor and network adapters to be able to communicate with all the cloud nodes. Application storage and add-in platform acceleration products may not be required for bare bone controller functionality to manage and communicate with Cloud nodes.

Bill of Materials (BOM) requirement details for the configurations are provided later section of this document.

Intel VRCs for NFVI are defined in collaboration with Communication Service Provider and ecosystem partners to demonstrate the value of an I/O Balanced Architecture for Network Function Virtualization. The solution leverages the hardened hardware, firmware, and software to allow customers to integrate on top of this known good foundation.

With the steady rise of cloud-based network traffic, edge computing has become increasingly more important to process this larger amount of data efficiently. By processing data at the network edge, latency for applications and services is reduced while potentially heavy traffic loads at the data center level, are alleviated by the distribution of traffic among the edge branches. However, this expanded network radius also corresponds to a wider attack range for a potential security threat. With these challenges in mind, Secure Access Service Edge (SASE) is intended to deliver a comprehensive solution for providing strong computational capability at the edge, securing the larger service area, and easily deploying new branches of the network.

Intel VRC for NFVI with SASE provides numerous benefits to ensure end users have excellent performance for their network applications. Some of the key benefits of the Intel VRC based on the 4th Generation Intel® Xeon® Scalable Processor Family processor include:

- High core counts and per-core performance
- Compact, power-efficient system-on-chip (Soc) platform
- Streamlined path to cloud-native operations
- Accelerated AI inference
- Accelerated encryption and compression
- Platform-level security enhancements

Figure 2. SASE Use Cases

1.1 Terminology

Table 1. Terminology

Term	Description
AIC	Add-In Card
API	Application Program interface
AGF	Access Gateway Function
BIOS	Basic Input/Output System
BOM	Bill of Materials
BtG	Boot Guard Technology
CUPS	Control Plane and User Plane Separation
DC	Data Center
DIMM	Dual Inline Memory Module
DPDK	Data Plane Development Kit
DRAM	Dynamic Random Access Memory
DUT	Device Under Test
ECDH	Elliptic Curve Diffie-Hellman protocol
ECDSA	Elliptic Curve Digital Signature Algorithm
GbE	Gigabit Ethernet
HQoS	Hierarchical Quality of Service
Intel® QAT	Intel® QuickAssist Technology
Intel® TXT	Intel® Trusted Execution Technology
Intel® UPI	Intel® Ultra Path Interconnect
Intel® VT	Intel® Virtualization Technology
Intel VRC	Intel Verified Reference Configuration
MLC	Memory Latency Checker
NAT	Network Address Translation
NFVI	Network Function Virtualization Infrastructure

Term	Description
NIC	Network Interface Controller
NUMA	Non-Uniform Memory Access
NVMe*	Non-Volatile Memory Express*
OAM	Operation, Administration and Management
OCP	Open Compute Project
OEM	Original Equipment Manufacturer
PCIe*	Peripheral Component Interconnect express*
PHY	Physical Layer
PXE	Pre-boot Execution Environment
QinQ	A standard that allows multiple VLAN headers in an Ethernet frame
RAS	Reliability, Availability, and Serviceability
SASE	Secure Access Service Edge
SR-IOV	Single Root Input/Output Virtualization
SSD	Solid State Drive
TCO	Total Cost of Ownership
TPM	Trusted Platform Module
Intel® TXT	Intel® Trusted Execution Technology
TPM	Trusted Platform Module
Intel® UPI	Intel® Ultra Path Interconnect
VIM	Virtualization Infrastructure Management
VMX	Virtual Machine Extension
VNFM	Virtual Network Function Management
VRC	Verified Reference Configuration
Intel® VT	Intel® Virtualization Technology

1.2 Reference Documents and Resources

Table 2. Reference Documents and Resources

Document	Document Number/Location
Intel® QuickAssist Technology Software for Linux* - Getting Started Guide – HW Version 2.0	https://developer.intel.com/quickassist
Intel® Network Platform Xeon-SP (NPX-SP) VRC for NFVI High Level Design Specification	736017
Red Hat's Certified Guest Operating System policy	https://access.redhat.com/articles/973163

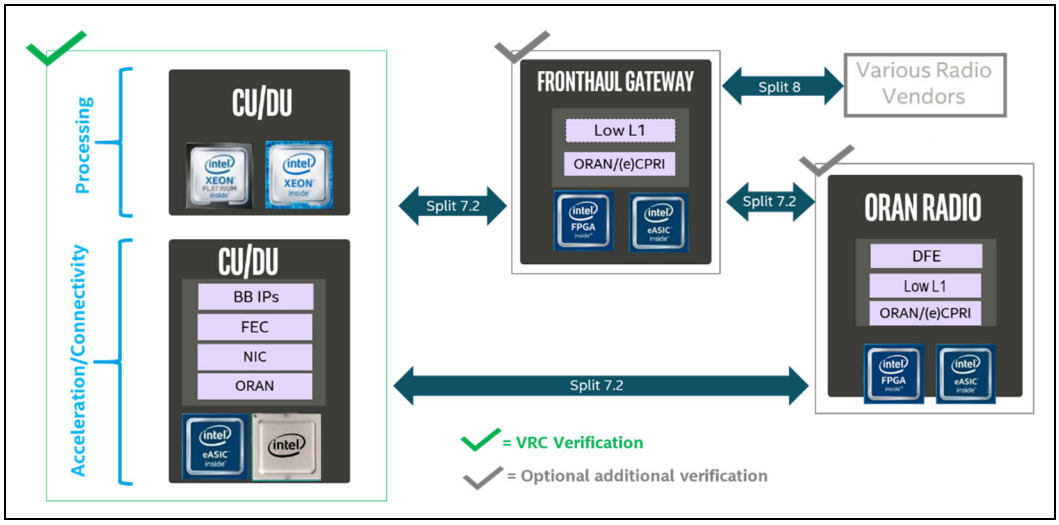
Document	Document Number/Location
Red Hat* OpenShift Container Platform version 4.12	https://docs.openshift.com/container-platform/4.12/welcome/index.html
Intel® Ethernet Connection E810 Application Device Queues (ADQ) Configuration Guide	609008
Intel® Deep Learning Boost - Boost Network Security AI Inference Performance in Google Cloud Platform (GCP) Technology Guide	https://networkbuilders.intel.com/solutionslibrary/intel-deep-learning-boost-boost-network-security-ai-inference-performance-in-google-cloud-platform-gcp-technology-guide
BIOS Settings for Intel® Wireline, Cable, Wireless and Converged Access Platform	https://cdrdv2.intel.com/v1/dl/getContent/747130

§

2 Solution Components

Intel VRCs consists of select hardware, various Intel® Xeon® processor technologies along with optimized software and firmware configurations. It consists of the components below.

Figure 2. Solution Overview



2.1 Intel® Xeon® Processor Scalable Performance Family

Intel® Xeon® Scalable processors are designed to accelerate performance across the fastest-growing workloads. These processors have the most built-in accelerators of any CPU on the market to help maximize performance efficiency for emerging workloads, especially those powered by AI.

In addition to delivering outstanding general-purpose performance, Intel® Xeon® drives efficiency with built-in accelerators. Data center operators can leverage built-in AI, telemetry, and power management tools for intelligent control electricity usage.

Intel's innovative workload accelerators enable end users to do more with less reducing TCO by delivering performance, power, resource, and cost efficiency as well as providing advanced security technologies.

The 4th Gen Intel® Xeon® Scalable Processors (formerly codenamed Sapphire Rapids) are the latest processors for Datacenter workloads that offer:

- **Enhanced Per Core Performance** with up to 60 cores in a standard socket
- **Enhanced Memory Performance** with support for up to 4800MT/s DIMMs (2 DPC)
- Increased Memory Capacity with up to 8 channels
- **Breakthrough System Memory & Storage** with Intel® Optane™ persistent memory 200 series

- **Built-in AI Acceleration** with enhanced performance of Intel® Deep Learning Boost
- **Faster UPI** with 3 Intel® Ultra Path Interconnect (Intel® UPI) at 11.2 GT/s
- **More, Faster I/O** with PCI Express 4 and up to 64 lanes (per socket) at 16 GT/s
- **New Hardware-Enhanced Security** delivering security technologies leadership with Intel® SGX, Intel® TME, Intel® PFR etc.
- **Enhanced Intel® Speed Select Technology (Intel® SST)** with three capabilities supported on the majority of Gold CPUs

2.2 Intel® C741 Chipset (codenamed: Emmitsburg) Platform Controller Hub

The Emmitsburg PCH provides extensive I/O support. The functions and capabilities are as follows:

- ACPI Power Management Logic Support, Revision 4.0a
- PCI Express Base Specification, Revision 4.0
- Integrated Serial ATA host controller supports data transfer rates of up to 6 b/s on all ports.
- xHCI USB controller with 10 USB 3.2 Gen 1 and 14 USB2 ports
- Serial Peripheral Interface
- Enhanced Serial Peripheral Interface
- Flexible I/O—Allows some high speed I/O signals to be configured as PCIe root ports, SATA, or USB 3.2 Gen 1
- General Purpose Input Output (GPIO)
- Interrupt controller, and timer functions
- System Management Bus Specification, Version 2.0
- Integrated Clock Controller/Real Time Clock Controller
- Intel® High-Definition Audio (Intel® HD Audio)
- Integrated 10/100/1000 Mbps Ethernet MAC
- Supports Intel® Rapid Storage Technology enterprise (Intel® RSTe)
- Supports Intel® Active Management Technology and Intel® Server Platform Services (Intel® SPS)
- Supports Intel® Virtualization Technology (Intel® VT) for Directed I/O (Intel® VT-d)
- Supports Intel® Trusted Execution Technology (Intel® TXT)
- JTAG Boundary Scan support
- Intel® Trace Hub (Intel® TH) for debug
- ADR Support

2.3 Intel® Ethernet 800 Series

Intel® Ethernet 800 Series offers:

- **Higher Bandwidth** as Intel's first NIC with PCIe* 4.0 and 50Gb PAM4 SerDes
- **Improved Application Efficiency** with Application Device Queues (ADQ), Dynamic Device Personalization (DDP)
- **Versatility** with Flexible speeds: 2x100/50/25/10GbE, 4x25/10GbE, or 8x10GbE
- **RDMA support** for both iWARP and RoCEv2 providing a choice in hyper-converged networks

2.3.1 Intel® Network Adapter with Data Plane Development Kit (DPDK)

Intel® Network Products offer continuously innovative solutions for high throughput and performance for networking infrastructure. The Intel® Network Adapter with Data Plane Development Kit (DPDK) provides highly optimized Network Virtualization and fast data path packet processing. DPDK offers many use cases that are hardened on this Intel® Accelerated Solution for NFVI Forwarding Platform.

2.3.2 Intel® Ethernet 800 Series Dynamic Device Personalization (DDP)

Dynamic Device Personalization (DDP) usage to reconfigure network controllers for different network functions on-demand, without the need for migrating all VMs from the server, avoids unnecessary loss of compute for VMs during server cold restart. It also improves packet processing performance for applications/VMs by adding the capability to process new protocols in the network controller at run-time.

This kind of on-demand reconfiguration is offered in the Intel® Ethernet 800 Series NICs.

Dynamic Device Personalization describes the capability of the Intel® Ethernet 800 Series devices to load an additional firmware profile on top of the device's default firmware image, enabling parsing and classification of additional specified packet types that can be distributed to specific queues on the NIC's host interface using standard filters. Software applies these custom profiles in a non-permanent, transaction-like mode so that the original network controller's configuration is restored after NIC reset or by rolling back profile changes by software. Using APIs provided by drivers, personality profiles can be applied by the DPDK. Support for kernel drivers and integration with higher level management/orchestration tools is in progress.

DDP can be used to optimize packet processing performance for different network functions, native or running in virtual environment. By applying a DDP profile to the network controller, the following use cases could be addressed.

A general purpose, OS-default DDP package is automatically installed with all supported Intel® Ethernet Controller 800 Series drivers on Microsoft* Windows*, ESX*, FreeBSD*, and Linux* operating systems. Additional DDP packages are available to address needs for specific market segments. For example, a telecommunications (Comms) DDP package is available to support certain market-specific protocols in addition to the protocols in the OS-default package.

- The OS-default DDP package supports the following:
  - MAC, EtherType, VLAN
  - IPv4, IPv6, TCP, ARP, UDP
  - SCTP, ICMP, ICMPv6, CTRL
  - LLDP, VXLAN-GPE, VXLAN (non-GPE), Geneve, GRE, NVGRE, RoCEv2
  - MPLS (up to 5 consecutive MPLS labels in the outermost Layer 2 header group)
- In addition to the previous list, the Comms DDP package also supports the following protocols:
  - GTP
  - PPPOE
  - L2TPv3
  - IPSec
  - PFCP

2.4 Intel® Xeon® Scalable Platform Technologies

2.4.1 Intel® Hyper-Threading Technology (Intel® HT Technology)

Enables multiple threads to run on each core, which ensures that systems use processor resources more efficiently. Intel® HT Technology also increases processor throughput, improving overall performance on threaded software.

2.4.2 Intel® Turbo Boost Technology

Accelerates processor and graphics performance for peak loads, automatically allowing processor cores to run faster than the rated operating frequency if they’re operating below power, current, and temperature specification limits.

2.4.3 Intel® Speed Select Technology (Intel® SST)

Intel® Speed Select Technology is a collection of features that improve performance and optimize TCO by providing more control over CPU performance. With Intel® SST, one server can do more.

- **Intel® SST- BF**: Trade off base frequency between higher and lower priority cores
- **Intel® SST- TF**: Capability to configure turbo frequencies beyond traditional max frequency limit
- **Intel® SST-CP**: Capability to allocate Surplus frequency to highest priority cores
- **Intel® SST-PP**: Capability to configure the CPU to run at 3 distinct operating points or profiles

2.4.4 Intel® Virtualization Technology (Intel® VT)

Intel® Virtualization Technology (Intel® VT) provides hardware abstraction to allow multiple workloads to co-exist and share common resources while maintaining full isolation.

The Intel® VT portfolio includes the capability in the CPU to implement Virtual Machine Extension (VMX) instructions. These allow all software in the Virtual Machine to run natively without performance impact from operations such as instruction translation and page table swaps with memory virtualization support.

Intel® VT also includes input/output (I/O) virtualization that supports offloading packet processing to network adapters, directly assigning virtual functions to virtual machines with Single Root I/O Virtualization (SR-IOV) and Intel® Data Direct I/O Technology Enhancement (Intel® DDIO), providing native network performance in the VM.

To obtain all the benefits of Intel® VT, Intel VRCs for Network Function Virtualization Infrastructure (NFVI) are required to have all virtualization technology features enabled.

2.4.5 Intel® Resource Director Technology (Intel® RDT)

Intel® Resource Director Technology (Intel® RDT) are collection of advanced technologies providing control over shared platform resources

Cache Monitoring Tech (CMT): Per-thread L3 Occupancy Monitoring

Cache Allocation Tech (CAT): Per-thread L3 Occupancy Control, New Code/Data Prioritization (CDP) extension

Memory BW Monitoring (MBM): Per-thread Memory Bandwidth Monitoring

Memory Bandwidth Allocation (MBA): Per-core Bandwidth Control – more precisely control "noisy neighbors"

2.4.6 Intel® On Demand

Intel® On Demand (formerly SDSI) is a new capability offering on demand one-time activation of incremental features during the Intel® Xeon CPU Life Cycle. Once enabled through a backend infrastructure, a hardened SW architecture for license delivery, with automated billing. License provisioning on the platform via out of band and in band communication for the one-time activation of features. CPU Features include: SGX, IAA, QAT, DLB, DSA.

2.4.7 Intel® Accelerator Interfacing Architecture (AiA)

Accelerator Interfacing Architecture (AiA) includes work submission, new instructions that provides low latency user-space work dispatch and synchronization between software and accelerators. AiA improves overall performance of the system for use in conjunction with the integrated accelerators in CPU, as well as discrete accelerators like NIC, GPU, FPGA-instantiated accelerators.

2.4.8 Intel® Deep Learning Boost

Intel® Deep Learning Boost includes extensive hardware inside CPU to provide set of instructions to accelerate Artificial Intelligence (AI) or Deep Learning Inference and training workload. Instructions include intel Deep Learning Boost i.e. VNNI/INT8, BF16 and with 4th Gen Intel® Xeon Scalable Processor the new Intel® AMX/TMUL instructions with INT8 and BF16 support.

2.4.9 Intel® Dynamic Load Balancer (DLB)

Intel® Dynamic Load Balancer previously known as Hardware Queue Manager improves the system performance related to handling network data on multi-core Intel® Xeon Scalable processors:

1. Distributed Processing - Intel® DLB enables the efficient distribution of network processing across multiple CPU cores/threads.
2. Dynamic Load Balancing - Intel® DLB dynamically distributes network data across multiple CPU cores for processing as the system load varies.
3. Dynamic Network Processing Reordering- Intel® DLB restores the order of networking data packets processed simultaneously on CPU cores.

2.4.10 Next-Gen Intel® QuickAssist Technology (Intel® QAT)

This product contains Intel® QuickAssist Technology which is integrated into the Processor with the following functions:

- Cryptographic Functions
  - Cipher Operations
  - NULL, AES
  - Snow3G UEA2
  - ZUC/128-EEA3(64B/1024B)
  - SM4
  - AES-GCM: Single Pass
  -
  - CHACHA20-POLYHash Operation
  - NULL, SHA-1
  - SHA-224/256
  - SHA-384/512
  - AES-(X)CBC-MAC
  - Galois Hash 64/128
  - UIA2 (Snow3G)
  - SHA3-224(64B/1024B)
  - SHA3-256(64B/1024B)
  - SHA3-384(64B/1024B)
  - SHA3-384(64B/1024B)
  - SHA3-512(64B/1024B)
  - ZUC/EIA3(64B/1024B)
  - SM3(64B/1024B)
- Authentication Operation
  - SHA1, SHA-256, SHA-512, SHA-224, SHA-384, SHA3-256, SHA3-512,
  - SHA3-224, SHA3-384, SM3, All HMAC variations
  - AES-CBC-MAC, AES-XCBC-MAC, GHASH64 (GMAC), GHASH128 (GMAC)
- Wireless Authentication Operation
  - AES-CBC-MAC, AES F9, SNOW3G UIA2, ZUC (128-EIA3)
- Cipher-Hash Combined Operation
- Key Derivation Operation
- Wireless Cryptography
  - AES, SM4, SNOW 3G*(UEA2), ZUC(128-EEA3)
  - Public Key Functions

- RSA Operation
- Diffie-Helman Operation
- Digital Signature Standard Operation
- Key Derivation Operation
- Elliptic Curve Cryptography: ECDSA* and ECDH*
  - Compression/Decompression Functions
- DEFLATE, LZ4s, LZ4

2.4.11 Intel® Data Streaming Accelerator (DSA)

4th Gen Intel® Xeon® Scalable Processors incorporate Intel® Data Streaming Accelerator Technology version 1.0 through a data accelerator for improving the performance of storage, networking, and various other I/O applications. The DMA engine is optimized for moving data between memory. The Intel® Data Streaming Accelerator replaces Intel® QuickData Technology used in previous Server processor generations. For more details, refer to Intel® Data Streaming Accelerator Specification, https://software.intel.com/content/www/us/en/develop/articles/intel-data-streamingaccelerator-architecture-specification.html

2.4.12 Intel® In-Memory Analytics Accelerator (Intel® IAA)

Intel® IAA **increases query throughput** and **decreases memory footprint** via:

1. Deeper compression than software-only techniques
2. More effective bandwidth, as deeply compressed data consumes less bandwidth
3. Core offload, as IAA performs computationally demanding scan and filter operations in place of cores.

2.4.13 Intel® Volume Management Device (VMD) & Intel® Virtual RAID on CPU (VROC)

Intel® VMD and Intel® VROC together provide RAS features for NVMe storage with maximum system up-time, ease of maintenance, cost-effective RAID solution and better NVMe performance in virtualization applications.

2.4.14 Intel® Seamless Firmware Update Technology

Intel® Seamless Firmware Update Technology is a cutting-edge solution from Intel that updates microcode with no perceived degradation to the services running on the platform. This technology is able to aggregate Intel firmware updates during run time without interrupting system operation. Also. it maintains server uptime while updating and activating firmware components.

2.4.15 Hardware Enhanced Security Features

4th Gen Intel® Xeon® Scalable Processors offer new hardware enhanced security features:

- **Intel® Software Guard Extensions (Intel® SGX) with additional integrity feature:** Provides fine grain data protection via application isolation in memory. Integrity provides greater resistance against physical attacks
- **Intel® Platform Firmware Resilience (Intel® PFR):** Verification of platform firmware images now extended to peripherals
- **Software Hardening Execution Controls**: Intel Virtualization Technology-Redirect Protection (formerly HLAT) - Enhanced page-table protections for OS kernel. Control flow Enforcement Tech (CET) - Return Oriented Programming (ROP), Jump Oriented Programming (JOP), Call Oriented Programming (COP) attack prevention with Shadow Stack and ENDBRANCH.
- Intel® Total Memory Encryption (**Intel® TME-MK**): Hardware-enabled memory encryption designed for multi-tenant server platforms. Also supports full memory encryption via single key

2.4.16 Trusted Platform Module (TPM)

TPM protects the system start-up process by ensuring it is tamper-free before releasing system control to the operating system. TPM 1.2 also provides secured storage for sensitive data, such as security keys and passwords, and performs encryption and hash functions.

Intel® Trusted Execution Technology (Intel® TXT) utilizes this technology.

2.4.17 Intel® Trusted Execution Technology (Intel® TXT)

Intel® TXT provides the foundation for highly scalable platform security in physical and virtual infrastructures. It helps harden servers at the hardware level against threats of hypervisor, BIOS, or other firmware attacks, malicious rootkit installations, and other types of attacks or misconfiguration to firmware and operating systems.

2.4.18 Intel® Hyper-Threading Technology (Intel® HT Technology)

Intel® HT Technology enables multiple threads to run on each core, which ensures that systems use processor resources more efficiently. Intel® HT Technology also increases processor throughput, improving overall performance on threaded software.

2.4.19 Intel® Boot Guard (Security)

Hardware-based boot integrity protection prevents unauthorized software and malware takeover of boot blocks critical to a system's function, thus providing added level of platform security based on hardware.

3 Design Compliance Requirements

This chapter focuses on the design requirements for Intel VRC for NFVI.

3.1 Intel VRC Hardware Requirements

The checklists in this chapter provide guidance for assessing the conformance to the Intel® Verified Reference Configuration for NFVI hardware platform requirement for the Cloud Node Base Configuration, Cloud Node Plus Configuration, Controller Node. For the platform to conform to the desired Intel® Verified Reference Configuration for NFVI, all requirements in the checklist must be met.

Table 3. Intel VRC for NFVI Plus Configuration- the Cloud Node HW Configuration

Ingredient	Requirement	Required/ Recommended	Quantity
Processor	Intel® Xeon® Gold 6438N Processor at 1.8GHz, 32C/64T, 205W or higher number SKU	Required	2
Memory	Option 1: DRAM only configuration: 512 GB (16x 32 GB DDR5, 4800 MHz)	Required	16
	Option 2: DRAM only configuration: 512 GB (32x 16 GB DDR5, 4800 MHz)		32
Network [1]	Intel® Ethernet Network Adapter E810-2CQDA2	Required	4 (2 per NUMA node)
Storage (Boot Drive)	480 GB or equivalent boot drive	Required	1
Storage (Capacity)	3.84 TB or equivalent drive (recommended NUMA aligned)	Recommended	4 (2 per NUMA node)
LAN on Motherboard (LOM)	10 Gbps or 25 Gbps port for Pre-boot Execution Environment (PXE) and Operation, Administration and Management (OAM)	Required	4 (2 per NUMA node)
	1/10 Gbps port for Management NIC	Required	1

Table 4. Intel VRC for NFVI Base Configuration – the Cloud Node HW Configuration

Ingredient	Requirement	Required/ Recommended	Quantity
Processor	Intel® Xeon® Gold 5418N processor at 1.8 GHz, 24C/48T, 165W or higher number SKU	Required	2
Memory[1]	DRAM only configuration: 256 GB (16x 16 GB DDR5, 4800 MHz)	Required	16

Ingredient	Requirement	Required/ Recommended	Quantity
Network [2]	Option 1 – Intel® Ethernet Network Adapter E810-CQDA2	Required	4 (2 per NUMA node)
	Option 2 – Intel® Ethernet Network Adapter E810-2CQDA2		2 (1 per NUMA node)
Storage (Boot Drive)	480 GB or equivalent boot drive	Required	2
Storage (Capacity)	3.84 TB or equivalent drive (recommended NUMA aligned)	Recommended	2 (1 per NUMA node)
LAN on Motherboard (LOM)	10 Gbps or 25 Gbps port for PXE/OAM	Required	2
	1/10 Gbps port for Management NIC	Required	1

Table 5. Intel VRC for NFVI – Controller Node HW Configuration

Ingredient	Requirement	Required/ Recommended	Quantity per Node
Processor[1]	Intel® Xeon® Gold 5418N processor at 1.8 GHz, 24C/48T, 165W (SST-PP config 2) or higher number SKU	Required	2
Memory[2]	DRAM only configuration: 256 GB (16x 16 GB DDR5, 4800 MHz)	Required	16
Network [3]	Option 1 – Intel® Ethernet Network Adapter E810-CQDA2	Required	2
	Option 2 – Intel® Ethernet Network Adapter E810-CQDA2 with Ethernet Port Configuration Tool (EPCT) to break down interface to 4x25G on single port or 8x10G mode (4x10 on single port)		
	Option 3 – Intel® Ethernet Network Adapter E810-XXVDA2		
Storage (Boot Drive)	480 GB or equivalent boot drive	Required	2
Storage (Capacity)	3.84 TB or equivalent drive (recommended NUMA aligned)	Required	2 (1 per NUMA node)
LAN on Motherboard (LOM)	10 Gbps or 25Gbps port for PXE/OAM	Recommended	2
	1/10 Gbps port for Management NIC	Required	1

3.2 Intel VRC Software Requirements

The table below is a guide for assessing the conformance to Intel Accelerated Solution for NFVI or SASE Intel Accelerated Solution for NFVI or SASE software requirements.

For the platform to conform the desired Intel Accelerated Solution for NFVI or SASE, all requirements listed in the checklist below must be satisfied.

Table 5. Intel Accelerated Solution for NFVI or SASE – SW Configuration

Ingredient	SW Version Details
DPDK	21.11.2
Intel® QAT	QAT 20.L.1.0.10-00005
Intel® Ethernet Network Adapter E810-CQDA2	CVL 4.10 0x800151a9 ice 4.18.0-372.40.1.el8_6.x86_64 iavf 4.18.0-372.40.1.el8_6.x86_64
Intel® Ethernet Network Adapter E810-2CQDA2	CVL 4.10 0x800151a9 ice 4.18.0-372.40.1.el8_6.x86_64 iavf 4.18.0-372.40.1.el8_6.x86_64
Red Hat	OpenShift version 4.12.1
Node Feature Discovery Operator	4.10.6
SR-IOV Network Operator	4.12.0-202303231115
Intel Device Plugins Operator	0.24.1
Kernel	4.18.0-372.40.1.el8_6.x86_64

Note:

1. Intel® recommends checking your system's exposure to the "Spectre" and "Meltdown" exploits.
2. Intel® QAT Software Package available from https://developer.intel.com/quickassist
3. The Intel® QAT driver is required to achieve the performance KPI's needed for certification. In the event of an issue, Red Hat* may request the user reproduce the issue with the Intel® QAT In-Tree driver to rule out interaction with the Intel® QAT driver from https://developer.intel.com/quickassist. Refer to https://access.redhat.com/articles/1067 which explains Red Hat* Support policy for Out Of Tree (OOT) drivers.
4. The customer should be aware of Red Hat's Certified Guest Operating System policy here (tier 3 vs. tier 1): https://access.redhat.com/articles/973163
5. Intel® Ethernet Adapter E810-CQDA2 Non-Volatile Memory (NVM) Update Utility for Intel® Ethernet Network Adapter 810 Series can be found at the following link: https://www.intel.com/content/www/us/en/download/19624/non-volatile-memory-nvm-update-utility-for-intel-ethernet-network-adapter-e810-series.html?wapkw=nvm%20update

3.3 BIOS Settings

To meet the performance requirements for an Intel VRC for NFVI platform solution, Intel® recommends using the BIOS settings for enabling processor p-state and c-state with Intel® Turbo Boost Technology ("turbo mode") enabled. Hyper threading is

recommended to provide higher thread density. Intel® also recommends using the BIOS settings for on demand Performance with power consideration.

Refer to the document *BIOS Settings for Intel® Wireline, Cable, Wireless and Converged Access Platform (#747130)* ***chapter 3*** for information on the BIOS settings.

Note: BIOS settings differ from vendor to vendor. Please contact your Intel Representative if you do not see the exact setting in your BIOS.

3.4 Platform Technology Requirements

This section lists the requirements for Intel's advanced platform technologies.

NFVI requires Intel® VT and Intel® Scalable I/O Virtualization (Intel® Scalable IOV) to be enabled to reap the benefits of hardware virtualization. Either Intel® Boot Guard or Intel® Trusted Execution Technology establishes the firmware verification, allowing for platform static root of trust.

Table 4. Platform Technology Requirements

Platform Technologies		Enable/Disable	Required/Recommended
Intel® VT	Intel® CPU VMX Support	Enable	Required
	Intel® I/O Virtualization	Enable	Required
Intel® Boot Guard	Intel® Boot Guard	Enable	Required
Intel® TXT	Intel® Trusted Execution Technology	Enable	Recommended

3.5 Platform Security

Intel Accelerated Solutions for vRAN must enable Intel® Boot Guard Technology to verify that the platform firmware is suitable during the boot phase.

In addition to protecting against the known attacks, all Intel Accelerated Solutions recommend installing the Trusted Platform Module (TPM). The TPM enables administrators to secure platforms for a trusted (measured) boot with known trustworthy (measured) firmware and OS. This allows local and remote verification by third parties to advertise known safe conditions for these platforms through implementation of Intel® Trusted Execution Technology (Intel® TXT).

3.6 Side Channel Mitigation

Reference intimation protection is verified with Spectre and Meltdown exposure using the latest Spectre and Meltdown Mitigation Detection Tool, which confirms the effectiveness of firmware and operating system updates against known attacks.

§

4 Platform Tuning for Worker Node

4.1 Boot Parameter Setup

For the workload testing, it is first necessary to setup the worker node with appropriate boot parameters as well as 1GB hugepages. Create the following file called "99-worker-custom.bu":

```
variant: openshift
version: 4.12.0
metadata:
  name: 99-worker-custom
  labels:
    machineconfiguration.openshift.io/role: worker
openshift:
  kernel_arguments:
    - iommu=pt
    - intel_iommu=on
    - hugepagesz=1G
    - default_hugepagesz=1G
    - hugepages=150
    - vfio-pci.ids=8086:4943
```

Run the following commands:

```
butane 99-worker-custom.bu -o ./99-worker-custom.yaml
oc create -f 99-worker-custom.yaml
```

The worker node should automatically reboot to apply the changes.

4.2 Building QAT Driver and Using QAT Device Plugin

Follow the instructions provided below to build the QAT2.0 OOT Driver:

1. Download the QAT2.0 driver (QAT20.L.1.0.10-00005.tar.gz) from https://www.intel.com/content/www/us/en/developer/topic-technology/open/quick-assist-technology/overview.html by accessing the Linux* Hardware v2.0 driver page.
2. Download the following RPM packages from the Red Hat Package Browser (https://access.redhat.com/downloads/content/package-browser).

Note: You need to have a Red Hat account to access this page

boost-1.66.0-10.el8.x86_64.rpm
boost-atomic-1.66.0-10.el8.x86_64.rpm
boost-chrono-1.66.0-10.el8.x86_64.rpm
boost-container-1.66.0-10.el8.x86_64.rpm
boost-context-1.66.0-10.el8.x86_64.rpm
boost-coroutine-1.66.0-10.el8.x86_64.rpm
boost-date-time-1.66.0-10.el8.x86_64.rpm
boost-devel-1.66.0-10.el8.x86_64.rpm
boost-fiber-1.66.0-10.el8.x86_64.rpm
boost-filesystem-1.66.0-10.el8.x86_64.rpm
boost-graph-1.66.0-10.el8.x86_64.rpm
boost-iostreams-1.66.0-10.el8.x86_64.rpm

boost-locale-1.66.0-10.el8.x86_64.rpm
boost-log-1.66.0-10.el8.x86_64.rpm
boost-math-1.66.0-10.el8.x86_64.rpm
boost-program-options-1.66.0-10.el8.x86_64.rpm
boost-random-1.66.0-10.el8.x86_64.rpm
boost-regex-1.66.0-10.el8.x86_64.rpm
boost-serialization-1.66.0-10.el8.x86_64.rpm
boost-signals-1.66.0-10.el8.x86_64.rpm
boost-stacktrace-1.66.0-10.el8.x86_64.rpm
boost-system-1.66.0-10.el8.x86_64.rpm
boost-test-1.66.0-10.el8.x86_64.rpm
boost-thread-1.66.0-10.el8.x86_64.rpm
boost-timer-1.66.0-10.el8.x86_64.rpm
boost-type_erasure-1.66.0-10.el8.x86_64.rpm
boost-wave-1.66.0-10.el8.x86_64.rpm
libicu-60.3-2.el8_1.x86_64.rpm
libicu-devel-60.3-2.el8_1.x86_64.rpm
libnl3-cli-3.5.0-1.el8.x86_64.rpm
libnl3-devel-3.5.0-1.el8.x86_64.rpm
libquadmath-8.5.0-16.el8_7.x86_64.rpm
libquadmath-devel-8.5.0-16.el8_7.x86_64.rpm
yasm-1.3.0-7.el8.x86_64.rpm

3. Create a tar archive of all the rpm files called "rpm_packages.tar.gz":

```
tar -cf rpm_packages.tar.gz *.rpm
```

4. Create the following script, naming the file "make_qat_driver.sh":

```
    #!/bin/bash

    yum update -y --allowerasing
    yum install -y gcc gcc-c++ systemd-devel pciutils kmod
    cd /home
    tar -xvf rpm_packages.tar.gz

    rpm -ivh yasm-1.3.0-7.el8.x86_64.rpm libicu-60.3-2.el8_1.x86_64.rpm
libquadmath-8.5.0-16.el8_7.x86_64.rpm
    rpm -ivh libicu-devel-60.3-2.el8_1.x86_64.rpm libquadmath-devel-8.5.0-
16.el8_7.x86_64.rpm
    rpm -ivh boost-atomic-1.66.0-10.el8.x86_64.rpm boost-container-1.66.0-
10.el8.x86_64.rpm boost-context-1.66.0-10.el8.x86_64.rpm boost-date-time-
1.66.0-10.el8.x86_64.rpm boost-iostreams-1.66.0-10.el8.x86_64.rpm
    rpm -ivh boost-math-1.66.0-10.el8.x86_64.rpm boost-program-options-1.66.0-
10.el8.x86_64.rpm boost-regex-1.66.0-10.el8.x86_64.rpm boost-serialization-
1.66.0-10.el8.x86_64.rpm boost-signals-1.66.0-10.el8.x86_64.rpm boost-
stacktrace-1.66.0-10.el8.x86_64.rpm
    rpm -ivh boost-system-1.66.0-10.el8.x86_64.rpm boost-thread-1.66.0-
10.el8.x86_64.rpm
    rpm -ivh boost-chrono-1.66.0-10.el8.x86_64.rpm
    rpm -ivh boost-coroutine-1.66.0-10.el8.x86_64.rpm boost-filesystem-1.66.0-
10.el8.x86_64.rpm boost-graph-1.66.0-10.el8.x86_64.rpm boost-locale-1.66.0-
10.el8.x86_64.rpm boost-log-1.66.0-10.el8.x86_64.rpm boost-random-1.66.0-
10.el8.x86_64.rpm boost-timer-1.66.0-10.el8.x86_64.rpm
    rpm -ivh boost-type_erasure-1.66.0-10.el8.x86_64.rpm boost-wave-1.66.0-
10.el8.x86_64.rpm boost-test-1.66.0-10.el8.x86_64.rpm
    rpm -ivh boost-fiber-1.66.0-10.el8.x86_64.rpm
    rpm -ivh boost-1.66.0-10.el8.x86_64.rpm
    rpm -ivh boost-devel-1.66.0-10.el8.x86_64.rpm
    rpm -ivh libnl3-cli-3.5.0-1.el8.x86_64.rpm
    rpm -ivh libnl3-devel-3.5.0-1.el8.x86_64.rpm

    # Make QAT2.0
    mkdir /root/QAT
    mv QAT20* /root/QAT
    cd /root/QAT
    tar -xf *
```

```
./configure --enable-icp-sriov=host
make
make install
```

5. Using your Red Hat login information, run the following command:
```
podman login registry.redhat.io
```

6. Create the following file called "qat_driver_spec.yaml":
```
apiVersion: v1
kind: Pod
metadata:
  labels:
    run: qat-driver0
  name: qat-driver0
spec:
  containers:
  - args:
    - /usr/lib/systemd/systemd
    image: registry.redhat.io/openshift4/driver-toolkit-rhel8:v4.12.0-
202301171655.p0.ge31abf2.assembly.stream
    name: qat-driver
    stdin: true
    securityContext:
      privileged: true
  nodeName: worker-0
  restartPolicy: Always
```
   a. Replace "worker-0" with the name of your worker node being used

7. Create and run the following script, naming the file "build_qat.sh":
```
#!/bin/bash

for i in 1 2
do
    oc delete pod qat-driver
    oc create -f ./qat_driver_spec.yaml
   sleep 90
   oc cp ./QAT20.L.1.0.10-00005.tar.gz qat-driver:/home
   oc cp ./rpm_packages.tar.gz qat-driver:/home
   oc cp ./make_qat_driver.sh qat-driver:/home
   oc exec -it qat-driver -- /home/make_qat_driver.sh
done
```

8. Run the following commands:
```
oc cp qat-driver:/root/QAT/build/qat_4xxx.bin ./qat_4xxx.bin
oc cp qat-driver:/root/QAT/build/qat_4xxx_mmp.bin ./qat_4xxx_mmp.bin
```

9. Follow the steps in the "Loading custom firmware blobs in the machine config manifest" section at https://docs.openshift.com/container-platform/4.12/post_installation_configuration/machine-configuration-tasks.html#rhcos-load-firmware-blobs_post-install-machine-configuration-tasks to configure the "qat_4xxx.bin" and "qat_4xxx_mmp.bin" files to be placed in the "/var/lib/firmware" directory
   a. Use the following for the file "98-worker-firmware-blob.bu":
```
variant: openshift
version: 4.12.0
metadata:
  labels:
    machineconfiguration.openshift.io/role: worker
  name: 98-master-firmware-blob
storage:
  files:
  - path: /var/lib/firmware/qat_c4xxx.bin
    contents:
      local: qat_4xxx.bin
    mode: 0644
```

```
      - path: /var/lib/firmware/qat_c4xxx_mmp.bin
        contents:
          local: qat_4xxx_mmp.bin
        mode: 0644
    openshift:
      kernel_arguments:
        - 'firmware_class.path=/var/lib/firmware'
```

b. To install butane, follow the steps at this link: https://docs.openshift.com/container-platform/4.12/installing/install_config/installing-customizing.html

c. The system should automatically reboot after this step

10. Every time the system is restarted, the "build_qat.sh" script must be run.

    To deploy the QAT Device Plugin in order to utilize QAT resources in Pods, follow the steps here: https://github.com/mregmi/intel-device-plugins-for-kubernetes/tree/qat_enable/cmd/operator/ocp_quickstart_guide.

5 Performance Verification

This chapter aims to verify the performance metrics for the reference intimation for NFVI to ensure that there is no anomaly seen. Refer to information in this section to ensure that the performance baseline for the platform is as expected.

The Plus solution was tested on March 29, 2023, with the following hardware and software configurations:

- 2 NUMA nodes
- 2x Intel® Xeon® Gold 6438N processors
- Total Memory: 512 GB, 16 slots/32 GB/4800 MT/s DDR5 RDIMM
- Hyperthreading: Enable
- Turbo: Enable
- C-State: Enable
- Storage: 1.8T INTEL SSDSC2KG01
- Network devices: 4x Single port Intel® Ethernet Network Adapter E810-CQDA2
- Network speed: 50 GbE
- BIOS: EGSDCRB1.SYS.9409.P09.2212301305
- Microcode: 0xab000190
- OS/Software: Red Hat CoreOS 412.86.202301191053-0 / Red Hat Enterprise Linux* 8.6 (kernel 4.18.0-372.40.1.el8_6.x86_64)

5.1 Memory Latency Checker (MLC)

The first application is the Memory Latency Checker which can be downloaded from https://www.intel.com/content/www/us/en/developer/articles/tool/intelr-memory-latency-checker.html

Download the latest version and execute this application, unzip the tarball package and go into Linux* folder and execute `./mlc or ./mlc_avx512`.

Table 5-1. Memory Latency Checker

Key Performance Metric	Local Socket
Idle Latency	100 ns
Memory Bandwidths between nodes within the system (using read-only traffic type) MB/s	260000

Table 5-6. Peak Injection Memory Bandwidth (1 MB/sec) Using All Threads

Peak Injection Memory Bandwidth (1 MB/sec) using all threads	Plus
All Reads	507000
3:1 Reads-Writes	469000
2:1 Reads-Writes	464000
1:1 Reads-Writes	447000
STREAM-Triad	436000
Loaded Latencies using Read-only traffic type with Delay=0 (ns)	200
L2-L2 HIT latency (ns)	62
L2-L2 HITM latency (ns)	63

Note: If the latency performance and memory bandwidth performance is outside the range, please verify the validity of the Platform components, BIOS settings, kernel power performance profile used and other software components.

5.2 NGINX*

Intel® QAT hardware acceleration helps to offload public key exchange for SSL layer 7 application. Intel VRC for NFVI – Plus platform must be able to demonstrate the minimum 7700 Connection Per Second (CPS) with the full software stack of NGINX* application with Intel® QAT HW (2 devices).

Intel VRC for NFVI - Plus Platform without Intel® QAT must demonstrate a minimum of 400 Connections Per Second (CPS) with the full software stack of NGINX* application.

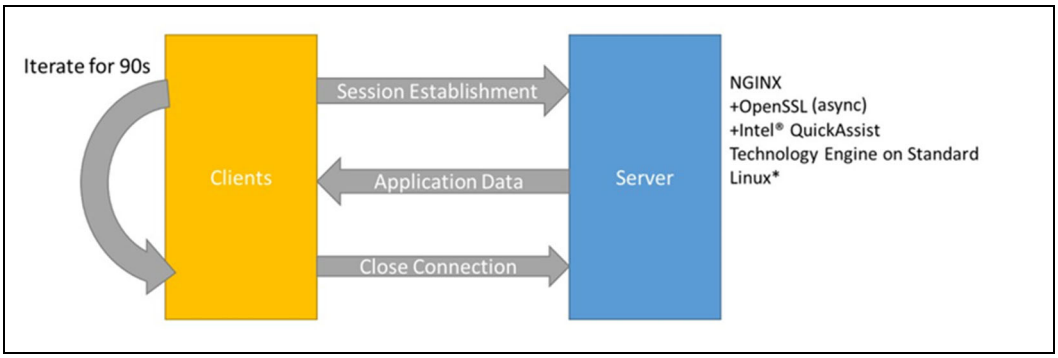
Table 7. NGINX* Workload Configuration

Ingredient	Software Version Details
Async mode nginx	v0.4.7
OpenSSL	3.0.2
QAT Engine	v0.6.19
IPP Crypto	Ippcp_2021.7
IPSec MB	1.3

Table 8. NGINX* Performance Requirements

Configuration[1]	Intel® QAT HW (2 devices)	Intel® QAT HW (1 device)	Intel® QAT-SW	Core Software
Intel® Verified Reference Configuration for NFVI - Plus platform	7700 CPS[2,3]	5400 CPS	1000 CPS	400 CPS
	[1] Uses ECDHE-X448-RSA4K TLS 1.3 handshakes [2] Intel® QAT HW (2 devices) is showing up to 16.8x performance over Core Software. [3] Intel® QAT HW (2 devices) is showing up to 7.2x performance over QAT-SW.			

Figure 3. Test Methodology for SSL with NGINX*



Note: Test Methodology and Procedures can be found in Appendix F.

The test methodology implements the following to measure the maximum CPS that the system can sustain:

- NGINX* Server Stack Utilizing Intel® QAT Engine and OpenSSL* 3.0.2+
- OpenSSL* scripts running in a container simulate client traffic sending 0-byte requests to the server (HTTPS workloads).
- The test measures the number of connection requests per second that the server can sustain.

5.2.1 NGINX Test Methodology

Follow these steps in order to run the NGINX workload testing:

1. Make sure that the QAT driver has been setup as described in Section 5.2 The QAT driver container created there will be used to run the NGINX application.
2. Create the following file called "install_nginx.sh":

```
#!/bin/bash

export ASYNC_NGINX_VERSION="v0.4.7"

yum update -y
```

```
yum install -y libudev-devel make gcc gcc-c++ pkg-config openssl-devel zlib-
devel wget git autoconf cmake libtool
yum install -y nasm # This is the only one that doesn't work

cd /home
git clone --depth 1 -b $ASYNC_NGINX_VERSION
https://github.com/intel/asynch_mode_nginx.git

cd /home/asynch_mode_nginx
./configure --prefix=/var/www --conf-path=/usr/share/nginx/conf/nginx.conf --
sbin-path=/usr/bin/nginx --pid-path=/run/nginx.pid --lock-
path=/run/lock/nginx.lock --modules-path=/usr/lib64/nginx --without-
http_rewrite_module --with-http_ssl_module --add-dynamic-
module=modules/nginx_qat_module/ --with-cc-opt="-DNGX_SECURE_MEM -I/include -
Wno-error=deprecated-declarations" --with-ld-opt="-L/src"
make && make install
```

3. Create the following file called "qat_hw_setup.sh":

```
#!/bin/bash

cd /home
git clone --depth 1 https://github.com/intel/QAT_Engine
yum install -y autoconf automake libtool openssl-devel

cd /home/QAT_Engine
./autogen.sh
./configure --enable-qat_hw --with-qat_hw_dir=/root/QAT
make && make install
```

4. The QAT configuration files need to be modified to use with the NGINX workload. Create the following file "4xxx_dev0.conf":

```
[GENERAL]
ServicesEnabled = asym
ConfigVersion = 2
FirmwareAuthEnabled = 1
CyNumConcurrentSymRequests = 512
CyNumConcurrentAsymRequests = 64
statsGeneral = 1
statsDc = 1
statsDh = 1
statsDrbg = 1
statsDsa = 1
statsEcc = 1
statsKeyGen = 1
statsLn = 1
statsPrime = 1
statsRsa = 1
statsSym = 1
StorageEnabled = 0
PkeServiceDisabled = 0
AutoResetOnError = 0
PmIdleInterruptDelay = 0
PmIdleSupport = 1
KptEnabled = 1
KptMaxSWKPerFn = 1
KptMaxSWKPerPASID = 1
KptMaxSWKLifetime = 31536000
KptSWKShared = 1
[KERNEL]
NumberCyInstances = 0
```

```
NumberDcInstances = 0
[SHIM]
NumberCyInstances = 1
NumberDcInstances = 0
NumProcesses = 1
LimitDevAccess = 1
Cy0Name = "UserCY0"
Cy0IsPolled = 1
Cy0CoreAffinity = 2
```

5. Create copies of the file:

```
cp 4xxx_dev0.conf 4xxx_dev1.conf ; cp 4xxx_dev0.conf 4xxx_dev2.conf ;
cp 4xxx_dev0.conf 4xxx_dev3.conf
```

   a. For "4xxx_dev1.conf", change the value of "Cy0CoreAffinity" to 3
   b. For "4xxx_dev2.conf", change the value of "Cy0CoreAffinity" to 4
   c. For "4xxx_dev3.conf", change the value of "Cy0CoreAffinity" to 5

6. Create the following file called "4xxxvf_dev0.conf":

```
[GENERAL]
ServicesEnabled = asym
ConfigVersion = 2
FirmwareAuthEnabled = 1
CyNumConcurrentSymRequests = 512
CyNumConcurrentAsymRequests = 64
statsGeneral = 1
statsDc = 1
statsDh = 1
statsDrbg = 1
statsDsa = 1
statsEcc = 1
statsKeyGen = 1
statsLn = 1
statsPrime = 1
statsRsa = 1
statsSym = 1
StorageEnabled = 0
PkeServiceDisabled = 0
AutoResetOnError = 0
KptEnabled = 1
KptMaxSWKPerFn = 1
KptMaxSWKPerPASID = 1
KptMaxSWKLifetime = 31536000
KptSWKShared = 1
[KERNEL]
NumberCyInstances = 0
NumberDcInstances = 0
[SHIM]
NumberCyInstances = 1
NumberDcInstances = 0
NumProcesses = 1
LimitDevAccess = 1
Cy0Name = "UserCY0"
Cy0IsPolled = 1
Cy0CoreAffinity = 2
```

7. Run the following commands:

```
for i in {1..63}
do
    cp 4xxxvf_dev0.conf 4xxxvf_dev${i}.conf
done
tar -cf mod_qat_conf.tar.gz 4xxx*conf
```

8. Remove the "vfio_pci" module of your worker node using "ssh core@<name_of_worker> sudo modprobe -r vfio_pci"
9. Run the following commands to setup the QAT changes:

```
oc cp mod_qat_conf.tar.gz qat-driver:/etc
oc exec qat-driver -- tar -C /etc -xf /etc/mod_qat_conf.tar.gz
oc exec qat-driver -- service qat_service stop
oc exec qat-driver -- service qat_service start
```

10. Run the following commands to setup NGINX and QAT Engine with QAT HW:

```
oc cp install_nginx.sh qat-driver:/home
oc cp qat_hw_setup.sh qat-driver:/home
oc exec qat-driver -- /home/install_nginx.sh
oc exec qat-driver -- /home/qat_hw_setup.sh
```

11. Use the following command to create the certificate and key to use for NGINX:

```
openssl req -x509 -newkey rsa:4096 -keyout server.key -out server.crt -days
365 -nodes
```

12. Create the following file called "nginx_test_qatengine.conf":

```
user root;
load_module /usr/lib64/nginx/ngx_ssl_engine_qat_module.so;
worker_processes 32;
#ssl_engine qat;
worker_rlimit_nofile 30000;

events
{
        use epoll;
        worker_connections 200000;
        #multi_accept on;
}
ssl_engine {
        use_engine qatengine;
        default_algorithms RSA,EC,DH,PKEY_CRYPTO;
        qat engine {
            qat_offload_mode async;
            qat_notify_mode poll;
            #qat_poll_mode heuristic;
            #qat_shutting_down_release on;
        }
}
http
{
          ssl_buffer_size 64k;
          keepalive_timeout  100;
          include       /usr/share/nginx/conf/mime.types;
          default_type  application/octet-stream;
          sendfile on;

        server
        {
                listen       4400 ssl reuseport backlog=200000;
                server_name  localhost;
                access_log off;
```

```
                sendfile on;
                #ssl                  on;
                ssl_asynch           on;
                ssl_certificate      /home/server.crt;
                ssl_certificate_key  /home/server.key;

                ssl_session_timeout  5m;
                ssl_ecdh_curve X448;

                ssl_protocols  TLSv1.3;
                ssl_ciphers ALL;

                ssl_prefer_server_ciphers   on;

                location /
                {
                    root   html;
                    index  index.html index.htm;
            }
        }
}
```

13. Create the following file called "nginx_test_software.conf":

```
user root;
worker_processes 32;
#ssl_engine qat;
worker_rlimit_nofile 30000;
#pid  /usr/local/nginx/logs/nginx.pid;

events
{
        use epoll;
        worker_connections 200000;
        #multi_accept on;
}
http
{
          ssl_buffer_size 64k;
          keepalive_timeout  100;
          include       /usr/share/nginx/conf/mime.types;
          default_type  application/octet-stream;
          sendfile on;

        server
        {
                listen       4400 ssl reuseport backlog=200000;
                server_name  localhost;
                access_log off;
                sendfile on;
```

```
                #ssl                    on;
                #ssl_asynch             on;
                ssl_certificate         /home/server.crt;
                ssl_certificate_key     /home/server.key;

                ssl_session_timeout     5m;
                ssl_ecdh_curve X448;

                ssl_protocols   TLSv1.3;
                ssl_ciphers ALL;

                ssl_prefer_server_ciphers    on;

                location /
                {
                    root    html;
                    index   index.html index.htm;
                }
        }
}
```

14. Create the following file called “openssl_clients.sh”:

```
#!/bin/bash
#
# Copyright 2017 Intel Corporation
#
# Permission is hereby granted, free of charge, to any person obtaining a copy
# of this software and associated documentation files (the "Software"),
# to deal in the Software without restriction, including without limitation
# the rights to use, copy, modify, merge, publish, distribute, sublicense,
# and/or sell copies of the Software, and to permit persons to whom
# the Software is furnished to do so, subject to the following conditions:
#
# The above copyright notice and this permission notice shall be included in
all
# copies or substantial portions of the Software.
#
# THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR
# IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY,
# FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT.
# IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM,
# DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT,
# TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH
# THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.
#
# SPDX-License-Identifier: MIT
#

#######################################
```

```
############# USER INPUT #############
######################################
ip_address=localhost
_time=90
clients=950
portbase=4400
cipher=TLS_AES_128_GCM_SHA256;
OPENSSL_DIR=/usr

#########################################################################
###LOGIC TO FIND HYPERTHREAD CORE CORRESPONDING TO PHYSICAL CORE ####
#########################################################################


# Cores to use for client traffic generation.  When
#   running the CPS tests on one system, use every core that is
#   not being used for nginx threads.
#

_cores_used=""

total_cores=$(lscpu | grep -E '^Thread|^Core|^Socket|^CPU\(' | sed -n '1p' |
awk '{print $2}')
thread_per_core=$( lscpu | grep -E '^Thread|^Core|^Socket|^CPU\(' | sed -n
'2p' |awk '{print $4}' )
cores_per_socket=$(  lscpu | grep -E '^Thread|^Core|^Socket|^CPU\(' | sed -n
'3p' |awk '{print $4}' )
socket_count=$( lscpu | grep -E '^Thread|^Core|^Socket|^CPU\(' | sed -n '4p'
|awk '{print $2}' )
total_physical_cores=$(( $cores_per_socket * $socket_count ))

function htForPhy (){

phy_core=$1

if [ "${phy_core}" -eq "${total_physical_cores}" ]; then
 return $total_cores

else
for ((i=$(($total_physical_cores +1)) ; i<=$total_cores; i++));
do
 if [ "$(($i%$total_physical_cores))" -eq "$phy_core" ];then
        return $i
 fi
done

fi

}
```

```
function findCoresUsed () {

core_begin0=$(( $cores_per_socket - 1 ))
core_end0=$(( $cores_per_socket  - 1 ))
core_begin1=$(( $cores_per_socket  + 1 ))
core_end1=$(( $total_cores - 1 ))
_cores_used_nginx=”-c 1-6”

# Check if plus configuration
if [ $cores_per_socket  -ge 14 ] ; then
        core_begin0=$(( $cores_per_socket - 2 ))
     _cores_used_nginx="-c 1-11"

fi

_cores_used="-c $core_begin0-$core_end0,$core_begin1-$core_end1"
}

findCoresUsed

# Running the NGINX
taskset $_cores_used_nginx nginx -c /home/nginx_test_qatengine.conf


######################################
############# USER INPUT #############
######################################

#Check for OpenSSL Directory
if [ ! -d $OPENSSL_DIR ];
then
    printf "\n$OPENSSL_DIR does not exist.\n\n"
    printf "Please modify the OPENSSL_DIR variable in the User Input
section!\n\n"
    exit 0
fi

helpAndError () {
    printf "\nThis script is to run the CPS testing HTTPS.\n"
    printf "\nTo use this script: ./connection_test_update1.sh \n"
    printf "\nTo do a dry-run, use the emulation flag:\n"
    printf "./connection_test_update1.sh --emulation\n\n"
    exit 0
}
emulation=0

#Check for h flag or no command line args
if [[ $1 == *"h"* ]]; then
    helpAndError
    exit 0
```

```
fi

#Check for emulation flag
if [[ $@ == **emulation** ]]
then
    emulation=1
fi

#The total commandline will be cmd1 + "192.168.1.1:4400" + cmd2
cmd1="$OPENSSL_DIR/bin/openssl s_time -connect"
cmd2="-new -ciphersuites $cipher -time $_time"

#Print out variables to check
printf "\n Location of OpenSSL:           $OPENSSL_DIR\n"
printf " IP Addresses:                   $ip_address\n"
printf " Time:                           $_time\n"
printf " Clients:                        $clients\n"
printf " Port Base:                      $portbase\n"
printf " Cipher:                         $cipher\n"
printf " Cores Used:                     $_cores_used\n"


#Remove previous .test files
rm -rf ./.test_*

#Get starttime
starttime=$(date +%s)

#Kick off the tests after checking for emulation
if [[ $emulation -eq 1 ]]
then
    for (( i = 0; i < ${clients}; i++ )); do
        printf "$cmd1 $ip_address:$(($portbase)) $cmd2 >
.test_$(($portbase))_$i &\n"
    done
    exit 0
else
    for (( i = 0; i < ${clients}; i++ )); do
        taskset $_cores_used $cmd1 $ip_address:$(($portbase)) $cmd2 >
.test_$(($portbase))_$i &
    done
fi

waitstarttime=$(date +%s)
# wait until all processes complete
while [ $(ps -ef | grep "openssl s_time" | wc -l) != 1 ];
do
    sleep 1
done
```

```
total=$(cat ./.test_$(($portbase))* | awk '(/^[0-9]* connections in [0-9]*
real/){ total += $1/$4 } END {print total}')
echo $total >> .test_sum
sumTotal=$(cat .test_sum | awk '{total += $1 } END { print total }')
printf "Connections per second:      $sumTotal CPS\n"
printf "Finished in %d seconds (%d seconds waiting for procs to start)\n"
$(($(date +%s) - $starttime)) $(($waitstarttime - $starttime))
rm -rf ./.test_*
```

15. Create the following file called "openssl_pod.yaml":

```
apiVersion: v1
kind: Pod
metadata:
  labels:
    run: openssl-pod
  name: openssl-pod
spec:
  containers:
  - args:
    - /bin/bash
    image: ubuntu:22.04
    name: openssl-pod
    stdin: true
  nodeName: worker-0
  restartPolicy: Always
```

    a. Replace "worker-0" with the name of the worker node you are using

16. Create the following file called "run_nginx_test.sh":

```
#!/bin/bash

read -p $'Please enter number of cores to test with\n' num_cores

# Modify worker processes for NGINX
oc cp nginx_test_qatengine.conf qat-driver:/home
oc cp nginx_test_software.conf qat-driver:/home
oc cp server.crt qat-driver:/home
oc cp server.key qat-driver:/home
#oc cp driver_nginx_setup.sh qat-driver:/home
#oc exec qat-driver -- /home/driver_nginx_setup.sh
oc cp openssl_clients.sh openssl-pod:/home

nginx_ip=$(oc exec qat-driver -- ip a | grep -A 3 eth | grep "inet " | awk
'{print $2;}' | awk -F"/" '{print $1;}')
oc exec openssl-pod -- sed -i "s/localhost/${nginx_ip}/"
/home/openssl_clients.sh

oc exec openssl-pod -- sed -i "s/5-31,69-95/$(( ${num_cores} + 2 ))-
31,$(( ${num_cores} + 66 ))-95/" /home/openssl_clients.sh

echo "Running the core SW test"

oc exec qat-driver -- taskset -c "1-${num_cores},65-$(( ${num_cores} + 64 ))"
nginx -c /home/nginx_test_software.conf

oc exec openssl-pod -- /home/nginx_test_sw.sh > coresw_results.log &

sleep 150
```

```
oc exec qat-driver -- pkill nginx

sleep 10

echo "Running the QAT test"

oc exec qat-driver -- taskset -c "1-${num_cores},65-$(( ${num_cores} + 64 ))"
nginx -c /home/nginx_test_qatengine.conf

oc exec openssl-pod -- /home/nginx_test_qengine.sh > qat_results.log &

sleep 150

oc exec qat-driver -- pkill nginx
```

17. Run the file "run_nginx_test.sh" to perform the testing and generate the results for the out-of-box SW and QAT HW (2 devices) configurations.
18. To generate the results for the QAT HW (1 device) configuration, change the number of worker processes in the "nginx_test_qatengine.conf" from 32 to 16 and rerun the "run_nginx_test.sh" file.
    a. Rename the "coresw_results.log" and "qat_results.log" files from the previous step first to prevent them from being overwritten
    b. The "qat_results.log" file in this run will correspond to the results of this configuration
    c. After this test is complete, revert the number of worker processes in the "nginx_test_qatengine.conf" back to 32
19. Create the following file called "qat_sw_setup.sh":

```
#!/bin/bash

export QAT_ENGINE_VERSION="v0.6.18"

cd /home
git clone --depth 1 -b $QAT_ENGINE_VERSION https://github.com/intel/QAT_Engine

git clone --depth 1 https://github.com/intel/ipp-crypto
git clone --depth 1 https://github.com/intel/intel-ipsec-mb

# Install QAT Engine

yum install -y autoconf automake libtool openssl-devel
rpm -ivh nasm-2.15.03-3.el8.x86_64.rpm

cd /home/ipp-crypto/sources/ippcp/crypto_mb
cmake . -B"../build" -DOPENSSL_INCLUDE_DIR=/usr/include/openssl -
DOPENSSL_LIBRARIES=/usr/lib64 -DOPENSSL_ROOT_DIR=/usr/bin/openssl
cd ../build
make crypto_mb && make install

cd /home/intel-ipsec-mb
make && make install LIB_INSTALL_DIR=/usr/lib64

cd /home/QAT_Engine
./autogen.sh
./configure --disable-qat_hw --enable-qat_sw
make && make install
```

20. Download "nasm-2.15.03-3.el8.x86_64.rpm" from the Red Hat Package Browser (https://access.redhat.com/downloads/content/package-browser). Note that you will need to have a Red Hat account to access this page
21. For setting up the QAT SW configuration, follow these commands:

```
oc cp qat_sw_setup.sh qat-driver:/home
oc cp nasm-2.15.03-3.el8.x86_64.rpm qat-driver:/home
oc exec -it qat-driver -- /bin/bash
cd /home/QAT_Engine
make uninstall ; make clean
cd /home
rm -rf QAT_Engine
/home/qat_sw_setup.sh
exit
```

22. Run the "run_nginx_test.sh" file to generate the QAT SW results in the "qat_results.log" file.
    a. Rename the "qat_results.log" file from step 18 first to prevent it from being overwritten

5.3 QATzip

QATzip is a user space library which builds on top of the Intel® QuickAssist Technology user space library, to provide extended accelerated compression and decompression services by offloading the actual compression and decompression request(s). QATzip produces data using the standard gzip* format (RFC1952) with extended headers or lz4 blocks with lz4 frame format. The data can be decompressed with a compliant gzip* or lz4 implementation. QATzip is designed to take full advantage of the performance provided by Intel® QuickAssist Technology.

In order to compare the value proposition for QATzip, the compression throughput, decompression throughput, and compression ratio can be gathered to review the value proposition of QAT offload in this use case.

5.4 VPP IPSec

Vector Packet Processor (VPP) Internet Protocol Security (IPSec) is generally used for firewall or VPN applications and provides secure remote access to onsite servers. For a given platform, the VPP IPSec workload can demonstrate the effectiveness of its crypto processing capabilities.

For the Intel VRC for NFVI Plus Platform, ensure that the results of the system follow the expected results as shown in Table 10, to baseline the performance of the platform.

Table 9. VPP IPSec Workload Configuration

Ingredient	Software Version Details
VPP	23.02
DPDK	21.11.2
T-Rex	v3.00
IPSec MB	1.3

Table 10. Plus Platform VPP IPSec Performance Requirements

Packet Size (bytes)	Optimized SW[1] Throughput (Gbps)	Out-of-box Software[2] Throughput (Gbps)
64	19.00	0.40
128	28.98	0.50
256	44.47	1.00
512	62.48	2.50
1024	91.54	5.00
1280	99.42	5.99
1450	99.96	6.99

Note: [1] Uses IPSecMB crypto handler

[2] Uses openssl crypto handler

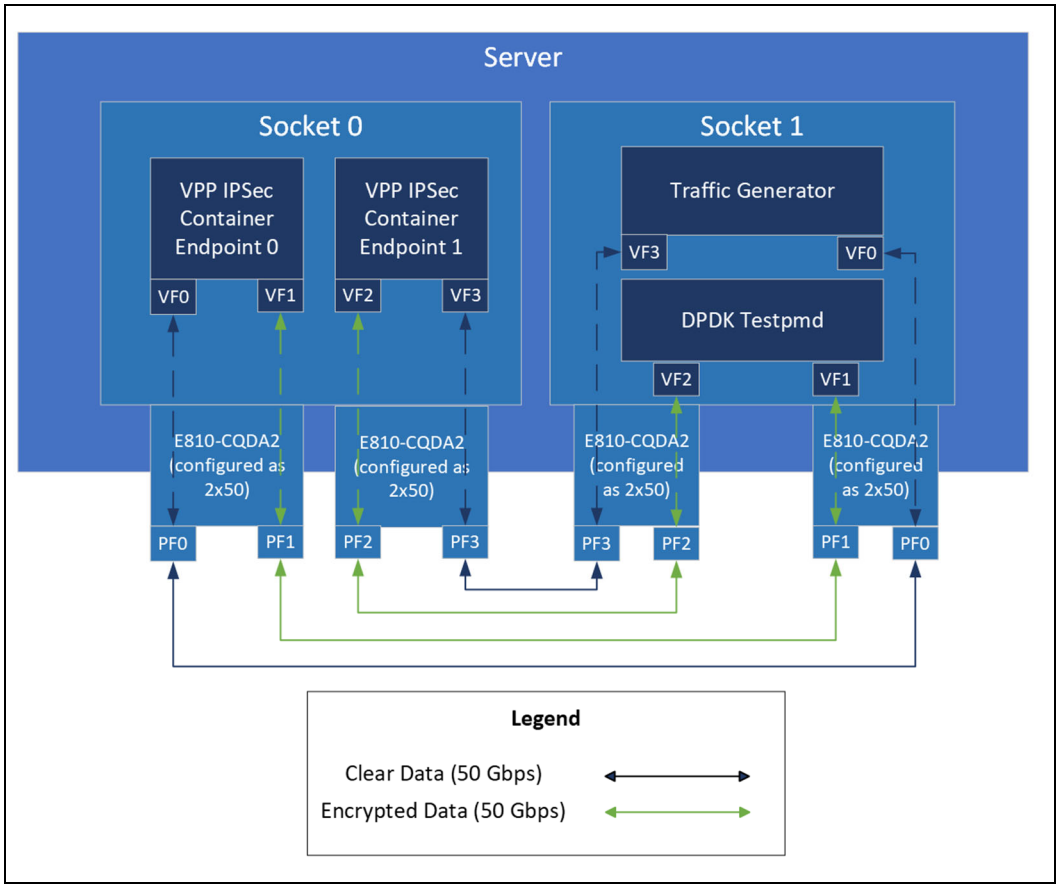
5.4.1 VPP IPSec Test Methodology

5.4.1.1 Setup

In order to perform the VPP IPSec benchmark testing, the system under test must be setup as shown in Figure 4.

Each socket of the system must contain four 50Gbps NIC ports to be connected to each other. Socket 0 will host the VPP IPSec container endpoints while socket 1 will host the traffic generator and also use the DPDK testpmd application as a simple switch. The generated traffic will flow bidirectionally throughout the setup. On socket 1, the two ports receiving the clear data should be used for the traffic generator, while the two ports receiving the encrypted data should run the DPDK testpmd application to forward the traffic to the other port.

Figure 4. Test Setup for VPP IPSec



To run the VPP IPSec testing, the first step is to configure the NIC devices to have a total of four 50Gbps ports available for each socket of the worker node server. Download the Ethernet Port Configuration Tool (EPCT) from this link: https://www.intel.com/content/www/us/en/download/19437/ethernet-port-configuration-tool-linux.html?wapkw=ethernet%20port%20configuration%20tool. Follow the EPCT instructions to configure the NIC devices.

Following this, it is necessary to utilize the SR-IOV Network Operator to create VFs of the NIC ports to be used in the VPP IPSec endpoints. From the Red Hat OpenShift web console of the cluster, navigate to the OperatorHub and search for the SR-IOV Network Operator. Install the operator, then navigate to the "Installed Operators" tab.

Click on the "SR-IOV Network Operator" and create an instance of the Sriov Network Node Policy. Under "nicSelector," select "rootDevices" and provide the PCI address of the first NIC port that will be used. Change the "resourceName" to the name of the network device followed by "_vf", and change "numVfs" to 1. Press "Create" and repeat this step for the other 7 NIC ports that will be used. Once completed, the worker node will automatically restart to apply the settings.

5.4.1.2 Testing Procedure

Follow these steps in order to run the VPP IPSec testing workload:

1. Create the following file called "setup_vpp.sh":

```
#!/bin/bash

read -p $'Before proceeding, please make sure you have modified the ipsec.cli
files in the ep0 and ep1 directories to use the appropriate MAC address in your
setup (lines 1033 and 1035).\nAdditionally, please modify the vpp_ep0.yaml and
vpp_ep1.yaml files to use the appropriate network devices VFs for your setup.\nIt
is also necessary to run the commands "sudo ip link set <nic_device> vf 0 spoof
off" and "sudo ip link set <nic_device> vf 0 trust on" for each NIC VF that will
be used.\nPress Enter to continue.\n'

oc delete pod vpp-ep0
oc delete pod vpp-ep1
oc delete pod dpdk-pod
oc delete pod trex-pod

for i in 0 1
do
    oc create -f vpp_ep${i}.yaml
    sleep 120
    oc cp get-vpp.sh vpp-ep${i}:/home
    oc cp install_vpp.sh vpp-ep${i}:/home
    oc cp ep${i} vpp-ep${i}:/home
    oc exec vpp-ep${i} -- /home/install_vpp.sh
done

./setup_dpdk_pod.sh
./setup_trex_pod.sh
```

2. Create the following file called "install_vpp.sh":

```
#!/bin/bash

apt-get update
apt-get install -y --no-install-recommends apt-transport-https ca-certificates
curl gnupg iproute2 iputils-ping kmod pciutils
rm -rf /var/lib/apt/lists/*

mkdir /vpp
mv /home/get-vpp.sh /vpp
cd /vpp
set -eux; ./get-vpp.sh
apt-get update
apt-get install -y -V ./*.deb
dpkg-query -f '${Version}\n' -W vpp > /vpp/version
rm -rf vom*.deb vpp-dbg*.deb
rm -rf /var/lib/apt/lists/*

mkdir -p /var/log/vpp
```

```
cd /home
```

3. Create the following file called “get_vpp.sh”:

```
    #!/bin/bash

    [ -z "$REPO URL" ] &&
REPO_URL="https://packagecloud.io/install/repositories/fdio/${REPO:=release}"

    # the code below comes from FDio's CSIT project.
    function get_vpp () {
    # Get and/or install Ubuntu VPP artifacts from packagecloud.io.
    #
    # Variables read:
    # - REPO_URL - FD.io Packagecloud repository.
    # - VPP_VERSION - VPP version.
    # - INSTALL - If install packages or download only. Default: download

    ls "*.deb" 2>/dev/null && { die "remove existing *.deb files"; }

    set -exuo pipefail
    trap '' PIPE

    curl -sS "${REPO_URL}"/script.deb.sh | bash || {
    die "Packagecloud FD.io repo fetch failed."
    }

    # If version is set we will add suffix.
    artifacts=()
    both_quotes='"'"'"
    match="[^${both quotes}]*"
    qmatch="[${both_quotes}]\?"
    sed_command="s#.*apt_source_path=${qmatch}\(${match}\)${qmatch}#\1#p"
    apt_fdio_repo_file=$(curl -s "${REPO_URL}"/script.deb.sh | \
    sed -n ${sed_command}) || {
    die "Local fdio repo file path fetch failed."
    }

    if [ ! -f ${apt_fdio_repo_file} ]; then
    die "${apt_fdio_repo_file} not found, \
    repository installation was not successful."
    fi

    packages=$(apt-cache -o Dir::Etc::SourceList=${apt_fdio_repo_file} \
    -o Dir::Etc::SourceParts=${apt_fdio_repo_file} dumpavail \
    | grep Package: | cut -d " " -f 2) || {
    die "Retrieval of available VPP packages failed."
    }
    if [ -z "${VPP_VERSION-}" ]; then
    allVersions=$(apt-cache -o Dir::Etc::SourceList=${apt_fdio_repo_file} \
    -o Dir::Etc::SourceParts=${apt_fdio_repo_file} \
```

```
    show vpp | grep Version: | cut -d " " -f 2) || {
    die "Retrieval of available VPP versions failed."
    }
    if [ "${REPO}" != "master" ]; then
    nonRcVersions=$(echo "$allVersions" | grep -v "\-rc[0-9]") || true
    [ -n "${nonRcVersions}" ] && allVersions=$nonRcVersions
    fi
    VPP_VERSION=$(echo "$allVersions" | head -n1) || true
    fi

    set +x
    echo "Finding packages with version: ${VPP_VERSION-}"
    for package in ${packages}; do
    # Filter packages with given version
    pkg_info=$(apt-cache show -- ${package}) || {
    die "apt-cache show on ${package} failed."
    }
    ver=$(echo ${pkg_info} | grep -o "Version: ${VPP_VERSION-}[^ ]*" | head -1) ||
true
    if [ -n "${ver-}" ]; then
    if [ "${package}" == "vom" ]; then
    echo " x '${package}' skipped"
    else
    echo "+++'${package}' found"
    ver=$(echo "$ver" | cut -d " " -f 2)
    artifacts+=(${package[@]/%/=${ver-}})
    fi
    else
    echo " - '${package}'"
    fi
    done
    set -x

    if [ "${INSTALL:-false}" = true ]; then
    apt-get -y install "${artifacts[@]}" || {
    die "Install VPP artifacts failed."
    }
    else
    apt-get -y download "${artifacts[@]}" || {
    die "Download VPP artifacts failed."
    }
    fi
    }

    function die () {
    # Print the message to standard error end exit with error code specified
    # by the second argument.
    #
    # Hardcoded values:
    # - The default error message.
```

```
    # Arguments:
    # - ${1} - The whole error message, be sure to quote. Optional
    # - ${2} - the code to exit with, default: 1.

    set -x
    set +eu
    echo "${1:-Unspecified run-time error occurred!}"
    exit "${2:-1}"
    }

    get_vpp
```

4. Create the following file called “vpp_ep0.yaml”:

```
apiVersion: v1
kind: Pod
metadata:
  labels:
    run: vpp-ep0
  name: vpp-ep0
spec:
  containers:
  - args:
    - /bin/bash
    image: ubuntu
    name: vpp-ep0
    stdin: true
    securityContext:
      runAsUser: 0
      capabilities:
        add: [“IPC LOCK”,”SYS RESOURCE”,”NET RAW”]
    volumeMounts:
    - mountPath: /dev/hugepages
      name: hugepage
    resources:
      limits:
        memory: "50Gi"
        hugepages-1Gi: "50Gi"
        openshift.io/<sock0_port_0>: 1
        openshift.io/<sock0_port_1>: 1
      requests:
        memory: "50Gi"
        hugepages-1Gi: "50Gi"
        openshift.io/<sock0_port_0>: 1
        openshift.io/<sock0_port_1>: 1
  volumes:
  - name: hugepage
    emptyDir:
      medium: HugePages
  nodeName: worker-0
  restartPolicy: Always
```

a. Replace "<sock0_port_0>" and "<sock0_port_1>" with the names of the corresponding Sriov Network Node policies created in the section above for PF0 and PF1 of socket 0. Replace "worker-0" with the name of your worker node being used.

5. Create the following file called "vpp_ep1.yaml":

```
apiVersion: v1
kind: Pod
metadata:
  labels:
    run: vpp-ep1
  name: vpp-ep1
spec:
  containers:
  - args:
    - /bin/bash
    image: ubuntu
    name: vpp-ep1
    stdin: true
    securityContext:
      runAsUser: 0
      capabilities:
        add: ["IPC_LOCK","SYS_RESOURCE","NET_RAW"]
    volumeMounts:
    - mountPath: /dev/hugepages
      name: hugepage
    resources:
      limits:
        memory: "50Gi"
        hugepages-1Gi: "50Gi"
        openshift.io/<sock0_port_2>_vf: 1
        openshift.io/<sock0_port_3>_vf: 1
      requests:
        memory: "50Gi"
        hugepages-1Gi: "50Gi"
        openshift.io/<sock0 port 2> vf: 1
        openshift.io/<sock0_port_3>_vf: 1
  volumes:
  - name: hugepage
    emptyDir:
      medium: HugePages
  nodeName: worker-0
  restartPolicy: Always
```

a. Replace "<sock0_port_2>" and "<sock0_port_3>" with the names of the corresponding Sriov Network Node policies created in the section above for PF2 and PF3 of socket 0. Replace "worker-0" with the name of your worker node being used.

6. Create the following file called "dpdk_pod.yaml":

```
apiVersion: v1
kind: Pod
```

```
metadata:
  labels:
    run: dpdk-pod
  name: dpdk-pod
spec:
  containers:
  - args:
    - /bin/bash
    image: ubuntu
    name: dpdk-pod
    stdin: true
    securityContext:
      runAsUser: 0
      capabilities:
        add: [“IPC_LOCK”,”SYS_RESOURCE”,”NET_RAW”]
    volumeMounts:
    - mountPath: /dev/hugepages
      name: hugepage
    resources:
      limits:
        memory: “30Gi”
        hugepages-1Gi: “10Gi”
        openshift.io/<sock1_port_1>_vf: 1
        openshift.io/<sock1_port_2>_vf: 1
      requests:
        memory: “30Gi”
        hugepages-1Gi: “10Gi”
        openshift.io/<sock1_port_1>_vf: 1
        openshift.io/<sock1_port_2>_vf: 1
  volumes:
  - name: hugepage
    emptyDir:
      medium: HugePages
  nodeName: worker-0
  restartPolicy: Always
```

a. Replace “<sock1_port_1>” and “<sock1_port_2>” with the names of the corresponding Sriov Network Node policies created in the section above for PF1 and PF2 of socket 1. Replace “worker-0” with the name of your worker node being used.

7. Create the following file called “setup_dpdk_pod.sh”:

```
#!/bin/bash

oc create -f dpdk_pod.yaml
wget http://fast.dpdk.org/rel/dpdk-21.11.2.tar.xz
sleep 120
oc cp dpdk-21.11.2.tar.xz dpdk-pod:/home
oc cp run_testpmd.sh dpdk-pod:/home
oc cp install_dpdk.sh dpdk-pod:/home
oc exec dpdk-pod -- /home/install_dpdk.sh
```

8. Create the following file called “install_dpdk.sh”:

```
#!/bin/bash

apt update -y
apt install -y kmod pciutils iproute2 meson python3-pyelftools libnuma-dev

cd /home
tar -xf dpdk-21.11.2.tar.xz
cd dpdk-stable-21.11.2
meson setup -Dexamples=all build
cd build
ninja
ninja install
ldconfig
```

9. Create the following file called “run_testpmd.sh”:

```
#!/bin/bash

/home/dpdk-stable-21.11.2/build/app/dpdk-testpmd -l 42-47 -n 4 --socket-mem=0,4096
-a <sock1_port_1> -a <sock1_port_2> --main-lcore=42 --in-memory
```

   a. Replace “<sock1_port_1>” and “<sock1_port_2>” with the PCI addresses (ex: 81:09.0) of the Ethernet Adaptive Virtual Functions created for the NIC ports corresponding to PF1 and PF2 of socket 1.

10. Create the following file called “trex_pod.yaml”:

```
apiVersion: v1
kind: Pod
metadata:
  labels:
    run: trex-pod
  name: trex-pod
spec:
  containers:
  - args:
    - /bin/bash
    image: python
    name: trex-pod
    stdin: true
    securityContext:
      runAsUser: 0
      capabilities:
        add: [“IPC_LOCK”,”SYS_RESOURCE”,”NET_RAW”]
    volumeMounts:
    - mountPath: /dev/hugepages
      name: hugepage
    resources:
      limits:
        memory: “10Gi”
        hugepages-1Gi: “10Gi”
```

```
        openshift.io/<sock1_port_0>: 1
        openshift.io/<sock1_port_3>: 1
      requests:
        memory: "10Gi"
        hugepages-1Gi: "10Gi"
        openshift.io/<sock1_port_0>: 1
        openshift.io/<sock1_port_3>: 1
  volumes:
  - name: hugepage
    emptyDir:
      medium: HugePages
  nodeName: worker-0
  restartPolicy: Always
```

a. Replace "<sock1_port_0>" and "<sock1_port_3>" with the names of the corresponding Sriov Network Node policies created in the section above for PF0 and PF3 of socket 1. Replace "worker-0" with the name of your worker node being used.

11. Create the following file called "setup_trex_pod.sh":

```
#!/bin/bash

oc create -f trex_pod.yaml
wget --no-check-certificate https://trex-tgn.cisco.com/trex/release/v3.00.tar.gz
sleep 30
oc cp v3.00.tar.gz trex-pod:/home
oc cp vpp_packets_p0.py trex-pod:/home
oc cp vpp_packets_p1.py trex-pod:/home
oc cp install_trex.sh trex-pod:/home
oc cp trex_cfg.yaml trex-pod:/etc

oc run toolkit-pod --image=registry.redhat.io/openshift4/driver-toolkit-
rhel8:v4.12.0-202301171655.p0.ge31abf2.assembly.stream -- sleep infinity
sleep 30
oc cp toolkit-pod:/lib/modules ./modules
oc cp toolkit-pod:/usr/src/kernels ./kernels
oc cp modules trex-pod:/lib/
oc cp kernels trex-pod:/usr/src/
oc delete pod toolkit-pod

oc exec trex-pod -- /home/install_trex.sh
```

a. Note that this step assumes that the kernel version being used is 4.18.0-372.40.1.el8_6.x86_64. If a different version is being used, the image version being used in line 12 needs to be changed so that the kernel source files in the image are the same as the kernel version being used. The different versions of this image can be found at: https://catalog.redhat.com/software/containers/openshift4/driver-toolkit-rhel8/604009d6122bd89307e00865?container-tabs=gti

12. Create the following file called "trex_cfg.yaml":

```
- version       : 2
  interfaces    : ["<sock1_port_0>","<sock1_port_3>"]
```

```
  port_limit    : 2

  enable_zmq_pub : true
  zmq_pub_port : 4500
  zmq_rpc_port : 4501
  limit_memory: 8192
  rx_desc : 1024
  tx_desc : 1024
  c : 7
  port_bandwidth_gb : 50
  services_core : 32
  port_info:
    - dst_mac: "<sock0_port0_vf_mac>" # Socket 0 Port 0 MAC
      src_mac : "<sock1_port0_vf_mac>" # Socket 1 Port 0 MAC
    - dst_mac: "<sock0_port3_vf_mac>" # Socket 0 Port 3 MAC
      src_mac : "<sock1_port3_vf_mac>" # Socket 1 Port 3 MAC

  platform :
    master_thread_id : 33
    latency_thread_id : 34
    dual_if :
      - socket : 1
        threads : [35,36,37,38,39,40,41]
```

a. Replace <sock1_port_0> and <sock1_port_3> with the corresponding PCI address of the VFs for PF0 and PF3 of socket 1. Replace <sock0_port0_vf_mac>, <sock1_port0_vf_mac>, <sock0_port3_vf_mac>, and <sock1_port3_vf_mac> with the corresponding MAC address of the VF for each port.

13. Create the following file called "vpp_packets_p0.py":

```
from trex_stl_lib.api import *
import argparse

class STLS1(object):

    def __init__ (self):
        self.mode = 0
        self.fsize = 64; # the size of the packet
        self.tunnels = 128; # number of VPP IPSec tunnels being used

    def create_pkt_base (self,addr):
        pkt_dst = "10.64.0." + str(addr)
        t=[

Ether(dst="<sock0_port0_vf_mac>")/IP(src="192.168.105.3",dst=pkt_dst)/UDP(dport=10
24,sport=49000),

Ether()/Dot1Q(vlan=12)/IP(src="16.0.0.1",dst="48.0.0.1")/UDP(dport=12,sport=1025),

Ether()/Dot1Q(vlan=12)/Dot1Q(vlan=12)/IP(src="16.0.0.1",dst="48.0.0.1")/UDP(dport=
12,sport=1025),

Ether()/Dot1Q(vlan=12)/IP(src="16.0.0.1",dst="48.0.0.1")/TCP(dport=12,sport=1025),
```

```
                Ether()/Dot1Q(vlan=12)/Ipv6(src="::5")/TCP(dport=12,sport=1025),
                Ether()/IP()/UDP()/Ipv6(src="::5")/TCP(dport=12,sport=1025)
                ];
        return t[self.mode]

    def create_stream (self):

        # Create base packet and pad it to size
        size = self.fsize - 4; # HW will add 4 bytes ethernet FCS
        profile = []
        for i in range(1,self.tunnels + 1):
            base_pkt =  self.create_pkt_base(i)
            pad = max(0, size - len(base_pkt)) * 'x'
            pkt = STLPktBuilder(pkt = base_pkt/pad,
                vm = [])
            profile.append(STLStream(packet = pkt, mode = STLTXCont()))
        return STLProfile(profile).get_streams()

    def get_streams (self, tunables, **kwargs):
        parser = argparse.ArgumentParser(description='Argparser for
{}'.format(os.path.basename(__file__)),

formatter_class=argparse.ArgumentDefaultsHelpFormatter)

        args = parser.parse_args(tunables)
        # create 1 stream
        return self.create_stream()


# dynamic load - used for trex console or simulator
def register():
    return STLS1()
```

a. Replace <sock0_port0_vf_mac> with the corresponding MAC address of the VF for port 0 in socket 0

14. Create the following file called "vpp_packets_p1.py":

```
from trex_stl_lib.api import *
import argparse

class STLS1(object):

    def __init__ (self):
        self.mode = 0
        self.fsize = 64; # the size of the packet
        self.tunnels = 128; # number of VPP IPSec tunnels being used

    def create_pkt_base (self,addr):
        pkt_dst = "20.64.0." + str(addr)
        t=[
```

```
Ether(dst="<sock0_port3_vf_mac>")/IP(src="192.168.115.3",dst=pkt_dst)/UDP(dport=10
24,sport=49000),

Ether()/Dot1Q(vlan=12)/IP(src="16.0.0.1",dst="48.0.0.1")/UDP(dport=12,sport=1025),

Ether()/Dot1Q(vlan=12)/Dot1Q(vlan=12)/IP(src="16.0.0.1",dst="48.0.0.1")/UDP(dport=
12,sport=1025),

Ether()/Dot1Q(vlan=12)/IP(src="16.0.0.1",dst="48.0.0.1")/TCP(dport=12,sport=1025),
                Ether()/Dot1Q(vlan=12)/Ipv6(src="::5")/TCP(dport=12,sport=1025),
                Ether()/IP()/UDP()/Ipv6(src="::5")/TCP(dport=12,sport=1025)
                ];
        return t[self.mode]

    def create_stream (self):

        # Create base packet and pad it to size
        size = self.fsize - 4; # HW will add 4 bytes ethernet FCS
        profile = []
        for i in range(1,self.tunnels + 1):
            base_pkt =  self.create_pkt_base(i)
            pad = max(0, size - len(base_pkt)) * 'x'
            pkt = STLPktBuilder(pkt = base_pkt/pad,
                vm = [])
            profile.append(STLStream(packet = pkt, mode = STLTXCont()))
        return STLProfile(profile).get_streams()

    def get_streams (self, tunables, **kwargs):
        parser = argparse.ArgumentParser(description='Argparser for
{}'.format(os.path.basename(__file__)),

formatter_class=argparse.ArgumentDefaultsHelpFormatter)

        args = parser.parse_args(tunables)
        # create 1 stream
        return self.create_stream()


# dynamic load - used for trex console or simulator
def register():
    return STLS1()
```

a. Replace <sock0_port3_vf_mac> with the corresponding MAC address of the VF for port 3 in socket 0

15. Create the following file called "install_trex.sh":

```
#!/bin/bash

apt update -y ; apt install -y vim kmod pciutils file tmux iproute2 meson python3-
pyelftools

cd /home
```

```
tar -xf v3.00.tar.gz

mv vpp_packets_p0.py vpp_packets_p1.py v3.00/stl

cd /usr/lib/x86_64-linux-gnu/
ln -s -f libc.a liblibc.a

sed -i 's/collections.Hashable/collections.abc.Hashable/'
/home/v3.00/external_libs/pyyaml-3.11/python3/yaml/constructor.py
```

16. Create the directories ep0 and ep1:

```
mkdir ep0
mkdir ep1
```

17. Inside directory ep0, create the following file called "start_vpp_native.sh"

```
#!/bin/bash
vpp -c /home/ep0/startup_native.conf
exit 0
```

18. In the directory ep0, create the following file called "start_vpp_qat_sw.sh":

```
#!/usr/bin/env bash
vpp -c /home/ep0/startup_qat_sw.conf
exit 0
```

19. Inside directory ep0, create the following file called "startup_native.conf":

```
unix {
   nodaemon
   log /var/log/vpp/vpp.log
   full-coredump
   cli-listen /run/vpp/cli.sock
   gid vpp
   interactive
   exec /home/ep0/ipsec.cli
}

logging {
  size 4096
  default-log-level debug
  default-syslog-log-level debug
}

api-trace {
  ## This stanza controls binary API tracing. Unless there is a very strong
reason,
  ## please leave this feature enabled.
  On
  ## Additional parameters:
  ##
  ## To set the number of binary API trace records in the circular buffer,
configure nitems
```

```
  ##
  ## nitems <nnn>
  ##
  ## To save the api message table decode tables, configure a filename. Results in
/tmp/<filename>
  ## Very handy for understanding api message changes between versions,
identifying missing
  ## plugins, and so forth.
  ##
  ## save-api-table <filename>
}

api-segment {
  gid vpp
}

socksvr {
  default
}

cpu {
        ## In the VPP there is one main thread and optionally the user can create
worker(s)
        ## The main thread and worker thread(s) can be pinned to CPU core(s)
manually or automatically

        ## Manual pinning of thread(s) to CPU core(s)

        ## Set logical CPU core where main thread runs, if main core is not set
        ## VPP will use core 1 if available
        main-core 0

        ## Set logical CPU core(s) where worker threads are running
        corelist-workers 1-8

        ## Automatic pinning of thread(s) to CPU core(s)

        ## Sets number of CPU core(s) to be skipped (1 … N-1)
        ## Skipped CPU core(s) are not used for pinning main thread and working
thread(s).
        ## The main thread is automatically pinned to the first available CPU core
and worker(s)
        ## are pinned to next free CPU core(s) after core assigned to main thread
        # skip-cores 4

        ## Specify a number of workers to be created
        ## Workers are pinned to N consecutive CPU cores while skipping "skip-
cores" CPU core(s)
        ## and main thread's CPU core
        # workers 4

        ## Set scheduling policy and priority of main and worker threads
```

```
       ## Scheduling policy options are: other (SCHED_OTHER), batch (SCHED_BATCH)
       ## idle (SCHED_IDLE), fifo (SCHED_FIFO), rr (SCHED_RR)
       scheduler-policy fifo

       ## Scheduling priority is used only for "real-time policies (fifo and rr),
       ## and has to be in the range of priorities supported for a particular
policy
       # scheduler-priority 50
}

buffers {
       ## Increase number of buffers allocated, needed only in scenarios with
       ## large number of interfaces and worker threads. Value is per numa node.
       ## Default is 16384 (8192 if running unpriviledged)
       # buffers-per-numa 128000

       ## Size of buffer data area
       ## Default is 2048
       default data-size 2048

       ## Size of the memory pages allocated for buffer data
       ## Default will try 'default-hugepage' then 'default'
       ## you can also pass a size in K/M/G e.g. '8M'
       # page-size default-hugepage
}

dpdk {
       ## Change default settings for all interfaces
       dev default {
          ## Number of receive queues, enables RSS
          ## Default is 1
          num-rx-queues 4

          ## Number of transmit queues, Default is equal
          ## to number of worker threads or 1 if no workers treads
          num-tx-queues 4

          ## Number of descriptors in transmit and receive rings
          ## increasing or reducing number can impact performance
          ## Default is 1024 for both rx and tx
          num-rx-desc 4096
          num-tx-desc 4096

          ## VLAN strip offload mode for interface
          ## Default is off
          #vlan-strip-offload off

          ## TCP Segment Offload
          ## Default is off
```

```
    ## To enable TSO`, 'enable-tcp-udp-checksum' must be set
    tso off

    ## Devargs
    ## device specific init args
    ## Default is NULL
    # devargs safe-mode-support=1,pipeline-mode-support=1

    ## rss-queues
    ## set valid rss steering queues
    rss-queues 0-4
}

## Whitelist specific interface by specifying PCI address and in
## addition specify custom parameters for this interface
dev 0000:3d:01.0 {
     workers 1-4
     name eth1
     num-rx-queues 4
}
dev 0000:3d:09.0 {
     workers 5-8
     name eth2
     num-rx-queues 4
}

# QAT VFs
#dev 0000:81:00.1
#dev 0000:81:00.2
#dev 0000:81:00.3

## Blacklist specific device type by specifying PCI vendor:device
## Whitelist entries take precedence
# blacklist 8086:10fb

## Disable multi-segment buffers, improves performance but
## disables Jumbo MTU support
no-multi-seg
## Change hugepages allocation per-socket, needed only if there is need for
## larger number of mbufs. Default is 256M on each detected CPU socket
socket-mem 8192
## Disables UDP / TCP TX checksum offload. Typically needed for use
## faster vector PMDs (together with no-multi-seg)
no-tx-checksum-offload
## Enable UDP / TCP TX checksum offload
## This is the reversed option of 'no-tx-checksum-offload'
# enable-tcp-udp-checksum
## Change UIO driver used by VPP, Options are: igb_uio, vfio-pci,
## uio_pci_generic or auto (default)
uio-driver vfio-pci
```

```
       log-level debug
}

memory {
## Set the main heap size, default is 1G
        main-heap-size 1G

        ## Set the main heap page size. Default page size is OS default page
        ## which is in most cases 4K. if different page size is specified VPP
        ## will try to allocate main heap by using specified page size.
        ## special keyword 'default-hugepage' will use system default hugepage
        ## size
        main-heap-page-size 1G
}

## node variant defaults
node {
  ## specify the preferred default variant
  #       default { variant avx512 }

  ## specify the preferred variant, for a given node
  #        ip4-rewrite { variant avx2 }
}


plugins {
        ## Adjusting the plugin path depending on where the VPP plugins are
        #       path /ws/vpp/build-root/install-vpp-native/vpp/lib/vpp_plugins
        path /usr/lib/x86_64-linux-gnu/vpp_plugins

        ## Disable all plugins by default and then selectively enable specific
plugins
        # plugin default { disable }
        plugin dpdk_plugin.so { enable }
        # plugin acl_plugin.so { enable }

        ## For QAT: Comment out all plugins below
        ## For IPSecMB: Uncomment plugin crypto_native_plugin.so
        plugin crypto_ipsecmb_plugin.so { disable }
        # plugin crypto_native_plugin.so { disable }
        # plugin crypto_sw_scheduler_plugin.so { disable }
}

## Statistics Segment
statseg {
    # socket-name <filename>, name of the stats segment socket
    #     defaults to /run/vpp/stats.sock
    # size <nnn>[KMG], size of the stats segment, defaults to 32mb
    size 32M
    # per-node-counters on | off, defaults to none
```

```
    per-node-counters off
    # update-interval <f64-seconds>, sets the segment scrape / update interval
}
```

20. In directory ep0, create the following file called "startup_qat_sw.conf":

```
   nodaemon
   log /var/log/vpp/vpp.log
   full-coredump
   cli-listen /run/vpp/cli.sock
   gid vpp
   interactive
   exec /home/ep0/ipsec.cli
}

logging {
  size 4096
  default-log-level debug
  default-syslog-log-level debug
}

api-trace {
  ## This stanza controls binary API tracing. Unless there is a very strong
reason,
  ## please leave this feature enabled.
  On
  ## Additional parameters:
  ##
  ## To set the number of binary API trace records in the circular buffer,
configure nitems
  ##
  ## nitems <nnn>
  ##
  ## To save the api message table decode tables, configure a filename.
Results in /tmp/<filename>
  ## Very handy for understanding api message changes between versions,
identifying missing
  ## plugins, and so forth.
  ##
  ## save-api-table <filename>
}

api-segment {
  gid vpp
}

socksvr {
  default
}

cpu {
```

```
        ## In the VPP there is one main thread and optionally the user can
create worker(s)
        ## The main thread and worker thread(s) can be pinned to CPU core(s)
manually or automatically

        ## Manual pinning of thread(s) to CPU core(s)

        ## Set logical CPU core where main thread runs, if main core is not set
        ## VPP will use core 1 if available
        main-core 0

        ## Set logical CPU core(s) where worker threads are running
        corelist-workers 1-8

        ## Automatic pinning of thread(s) to CPU core(s)

        ## Sets number of CPU core(s) to be skipped (1 … N-1)
        ## Skipped CPU core(s) are not used for pinning main thread and working
thread(s).
        ## The main thread is automatically pinned to the first available CPU
core and worker(s)
        ## are pinned to next free CPU core(s) after core assigned to main
thread
        # skip-cores 4

        ## Specify a number of workers to be created
        ## Workers are pinned to N consecutive CPU cores while skippi"g "skip-
co"es" CPU core(s)
        ## and main thr'ad's CPU core
        # workers 4

        ## Set scheduling policy and priority of main and worker threads

        ## Scheduling policy options are: other (SCHED_OTHER), batch
(SCHED_BATCH)
        ## idle (SCHED_IDLE), fifo (SCHED_FIFO), rr (SCHED_RR)
        scheduler-policy fifo

        ## Scheduling priority is used only f"r "real-time policies (fifo and
rr),
        ## and has to be in the range of priorities supported for a particular
policy
        # scheduler-priority 50
}

buffers {
        ## Increase number of buffers allocated, needed only in scenarios with
        ## large number of interfaces and worker threads. Value is per numa
node.
        ## Default is 16384 (8192 if running unpriviledged)
        # buffers-per-numa 128000
```

```
        ## Size of buffer data area
        ## Default is 2048
        default data-size 2048

        ## Size of the memory pages allocated for buffer data
        ## Default will t‘y 'default-hugep’ge' th‘n 'defa’lt'
        ## you can also pass a size in K/M/G e.‘. ’8M'
        # page-size default-hugepage
}

dpdk {
        ## Change default settings for all interfaces
        dev default {
           ## Number of receive queues, enables RSS
           ## Default is 1
           num-rx-queues 4

           ## Number of transmit queues, Default is equal
           ## to number of worker threads or 1 if no workers treads
           num-tx-queues 4

           ## Number of descriptors in transmit and receive rings
           ## increasing or reducing number can impact performance
           ## Default is 1024 for both rx and tx
           num-rx-desc 4096
           num-tx-desc 4096

           ## VLAN strip offload mode for interface
           ## Default is off
           #vlan-strip-offload off

           ## TCP Segment Offload
           ## Default is off
           ## To enable TS‘, 'enable-tcp-udp-check’um' must be set
           tso off

           ## Devargs
           ## device specific init args
           ## Default is NULL
           # devargs safe-mode-support=1,pipeline-mode-support=1

           ## rss-queues
           ## set valid rss steering queues
           rss-queues 0-4
        }

        ## Whitelist specific interface by specifying PCI address and in
        ## addition specify custom parameters for this interface
        dev 0000:3d:01.0 {
             workers 1-4
```

```
            name eth1
            num-rx-queues 4
      }
      dev 0000:3d:09.0 {
            workers 5-8
            name eth2
            num-rx-queues 4
      }

      # QAT VFs
      #dev 0000:81:00.1
      #dev 0000:81:00.2
      #dev 0000:81:00.3

      ## Blacklist specific device type by specifying PCI vendor:device
      ## Whitelist entries take precedence
      # blacklist 8086:10fb

      ## Disable multi-segment buffers, improves performance but
      ## disables Jumbo MTU support
      no-multi-seg
      ## Change hugepages allocation per-socket, needed only if there is need
for
      ## larger number of mbufs. Default is 256M on each detected CPU socket
      socket-mem 8192
      ## Disables UDP / TCP TX checksum offload. Typically needed for use
      ## faster vector PMDs (together with no-multi-seg)
      no-tx-checksum-offload
      ## Enable UDP / TCP TX checksum offload
      ## This is the reversed option 'f 'no-tx-checksum-offl'ad'
      # enable-tcp-udp-checksum
      ## Change UIO driver used by VPP, Options are: igb_uio, vfio-pci,
      ## uio_pci_generic or auto (default)
      uio-driver vfio-pci
      log-level debug
}

memory {
## Set the main heap size, default is 1G
       main-heap-size 1G

       ## Set the main heap page size. Default page size is OS default page
       ## which is in most cases 4K. if different page size is specified VPP
       ## will try to allocate main heap by using specified page size.
       ## special keywo'd 'default-hugep'ge' will use system default hugepage
       ## size
       main-heap-page-size 1G
}

## node variant defaults
```

```
node {
  ## specify the preferred default variant
  #       default { variant avx512 }

  ## specify the preferred variant, for a given node
  #        ip4-rewrite { variant avx2 }
}


plugins {
        ## Adjusting the plugin path depending on where the VPP plugins are
        #       path /ws/vpp/build-root/install-vpp-native/vpp/lib/vpp_plugins
        path /usr/lib/x86_64-linux-gnu/vpp_plugins

        ## Disable all plugins by default and then selectively enable specific
plugins
        # plugin default { disable }
        plugin dpdk_plugin.so { enable }
        # plugin acl_plugin.so { enable }

        ## For QAT: Comment out all plugins below
        ## For IPSecMB: Uncomment plugin crypto_native_plugin.so
        # plugin crypto_ipsecmb_plugin.so { disable }
         plugin crypto_native_plugin.so { disable }
        # plugin crypto_sw_scheduler_plugin.so { disable }
}

## Statistics Segment
statseg {
    # socket-name <filename>, name of the stats segment socket
    #     defaults to /run/vpp/stats.sock
    # size <nnn>[KMG], size of the stats segment, defaults to 32mb
    size 32M
    # per-node-counters on | off, defaults to none
    per-node-counters off
    # update-interval <f64-seconds>, sets the segment scrape / update interval
}
```

21. Inside the directory ep0, create the following file called “create_ipsec_cli_ep0.py”:

```
#!/usr/bin/env python3
import sys

def main( argv ):

    crypto_handler       = argv[1]
    crypto_alg           = argv[2]
    integ_alg            = argv[3]
    num_ipsec_tunnels    = argv[4]
```

```
if not crypto_handler in ["ipsecmb", "native"]:
    print( "Unknown crypto handler: %s" % (crypto_handler) )
    return 1

if not crypto_alg in ["aes-cbc-128", "aes-gcm-128"]:
    print( "Unknown crypto algorithm: %s" % (crypto_alg) )
    return 1

if not integ_alg in ["none", "sha1-96"]:
    print( "Unknown integrity algorithm: %s" % (integ_alg) )
    return 1

try:
    num_ipsec_tunnels = int( argv[4] )
except:
    print( "Invalid number of IPSec tunnels: %s" % (num_ipsec_tunnels) )
    return 1

if num_ipsec_tunnels <= 0:
    print( "Invalid number of IPSec tunnels: %d" % (num_ipsec_tunnels) )
    return 1

if num_ipsec_tunnels >= 256:
    print( "Invalid number of IPSec tunnels: %d" % (num_ipsec_tunnels) )
    return 1

output_f_name="ipsec.cli"
num_interfaces       = 2
mtu                  = 1518
intf_ip_addrs        = ["192.168.105.2", "172.16.10.2"]
neigh_ip_addrs       = ["192.168.105.3", "172.16.10.3"]
neigh_mac_addrs      = ["<sock1_port0_vf_mac>", "<sock1_port1_vf_mac>"]
neigh_subnets        = ["20.64.0.0/10", "10.192.0.0/10"]
crypto_key           = "4339314b55523947594d6d3547666b45"
integ_key            = "4339314b55523947594d6d3547666b45"
start_sa             = 20
start_spi            = 1000
tunnel_src_prefix    = "10.128.0."
tunnel_dst_prefix    = "10.192.0."
ext_ip_prefix        = "10.64.0."
reverse_spi          = False

lines = []

# Set MTU on interfaces
for i in range( 1, num_interfaces + 1 ):
    lines.append( "set interface mtu %d eth%d\n" % (mtu, i) )
lines.append( "\n" )

# Set IP addresses on interfaces and enable promiscuous mode
```

```
    for i in range( 1, num_interfaces + 1 ):
        next_intf_ip_addr = intf_ip_addrs[i-1]
        lines.append( "set interface ip address eth%d %s/24\n" % (i, next_intf_i
p_addr) )
        lines.append( "set interface promiscuous on eth%d\n" % (i) )
    lines.append( "\n" )

    # Create IPSec Tunnels]
    next_sa = start_sa
    next_spi = start_spi
    for i in range( 1, num_ipsec_tunnels + 1 ):
        next_ip_tunnel_name = "ipip%d" % ( i - 1 )
        lines.append( "create ipip tunnel src %s%d dst %s%d\n" % (tunnel_src_pre
fix, i, tunnel_dst_prefix, i) )
        lines.append( "set interface state ipip%d up\n" % (i - 1) )
        if reverse_spi == True:
            lines.append( "ipsec sa add %d spi %d crypto-alg %s crypto-key %s in
teg-alg %s" % (next_sa, next_spi + 1, crypto_alg, crypto_key, integ_alg) )
        else:
            lines.append( "ipsec sa add %d spi %d crypto-alg %s crypto-key %s in
teg-alg %s" % (next_sa, next_spi, crypto_alg, crypto_key, integ_alg) )
        if integ_alg != "none":
            lines[-1] += " integ-key %s" % ( integ_key )
        lines[-1] += "\n"
        next_sa += 1
        next_spi += 1
        if reverse_spi == True:
            lines.append( "ipsec sa add %d spi %d crypto-alg %s crypto-key %s in
teg-alg %s" % (next_sa, next_spi - 1, crypto_alg, crypto_key, integ_alg) )
        else:
            lines.append( "ipsec sa add %d spi %d crypto-alg %s crypto-key %s in
teg-alg %s" % (next_sa, next_spi, crypto_alg, crypto_key, integ_alg) )
        if integ_alg != "none":
            lines[-1] += " integ-key %s" % ( integ_key )
        lines[-1] += "\n"
        lines.append( "ipsec tunnel protect %s sa-in %d sa-out %d\n" % (next_ip_
tunnel_name, next_sa-1, next_sa) )
        lines.append( "set interface ip address %s %s%d/32\n" % (next_ip_tunnel_
name, tunnel_src_prefix, i) )
        lines.append( "ip route add %s%d/32 via %s\n" % (ext_ip_prefix, i, next_
ip_tunnel_name) )
        lines.append( "\n" )
        next_sa += 1
        next_spi += 1

    # Set ARP table entries and IP route entries
    for i in range( 1, num_interfaces + 1 ):
        next_neigh_ip_addr = neigh_ip_addrs[i-1]
        next_neigh_mac_addr = neigh_mac_addrs[i-1]
        next_neigh_subnet = neigh_subnets[i-1]
        lines.append( "set ip neighbor eth%d %s %s\n" % (i, next_neigh_ip_addr,
next_neigh_mac_addr) )
```

```
        lines.append( "ip route add %s via %s eth%d\n" % (next_neigh_subnet, nex
t_neigh_ip_addr, i) )
    lines.append( "\n" )

    # Set interface state up
    for i in range( 1, num_interfaces + 1 ):
        lines.append( "set interface state eth%d up\n" % (i) )
    lines.append( "\n" )

    lines.append( "set crypto handler all openssl\n")
    lines.append( "set crypto handler all %s\n" % (crypto_handler) )

    with open( output_f_name, 'w' ) as output_file:
        output_file.writelines( lines )

if __name__ == "__main__":
sys.exit( main(sys.argv) )
```

a. Replace <sock1_port0_vf_mac> and <sock1_port1_vf_mac> with the MAC addresses of the VFs of the corresponding ports.

22. In the directory ep0, run the following command to generate the "ipsec.cli" file to use for VPP:

```
python3 create_ipsec_cli_ep0.py ipsecmb aes-cbc-128 sha1-96 128
```

23. In the directory ep1, create the following file called "start_vpp_native.sh":

```
#!/usr/bin/env bash
vpp -c /home/ep1/startup_native.conf
exit 0
```

24. In the directory ep1, create the following file called "start_vpp_qat_sw.sh":

```
#!/usr/bin/env bash
vpp -c /home/ep1/startup_qat_sw.conf
exit 0
```

25. In the directory ep1, create the following file called "startup_native.conf":

```
unix {
   nodaemon
   log /var/log/vpp/vpp.log
   full-coredump
   cli-listen /run/vpp/cli.sock
   gid vpp
   interactive
   exec /home/ep1/ipsec.cli
}

logging {
  size 4096
  default-log-level debug
  default-syslog-log-level debug
}
```

```
api-trace {
  ## This stanza controls binary API tracing. Unless there is a very strong
reason,
  ## please leave this feature enabled.
  on
  ## Additional parameters:
  ##
  ## To set the number of binary API trace records in the circular buffer,
configure nitems
  ##
  ## nitems <nnn>
  ##
  ## To save the api message table decode tables, configure a filename. Results in
/tmp/<filename>
  ## Very handy for understanding api message changes between versions,
identifying missing
  ## plugins, and so forth.
  ##
  ## save-api-table <filename>
}

api-segment {
  gid vpp
}

socksvr {
  default
}

cpu {
       ## In the VPP there is one main thread and optionally the user can create
worker(s)
       ## The main thread and worker thread(s) can be pinned to CPU core(s)
manually or automatically

       ## Manual pinning of thread(s) to CPU core(s)

       ## Set logical CPU core where main thread runs, if main core is not set
       ## VPP will use core 1 if available
       main-core 64

       ## Set logical CPU core(s) where worker threads are running
       corelist-workers 65-72

       ## Automatic pinning of thread(s) to CPU core(s)

       ## Sets number of CPU core(s) to be skipped (1 ... N-1)
       ## Skipped CPU core(s) are not used for pinning main thread and working
thread(s).
       ## The main thread is automatically pinned to the first available CPU core
and worker(s)
```

```
       ## are pinned to next free CPU core(s) after core assigned to main thread
       # skip-cores 4

       ## Specify a number of workers to be created
       ## Workers are pinned to N consecutive CPU cores while skipping "skip-
cores" CPU core(s)
       ## and main thread's CPU core
       # workers 4

       ## Set scheduling policy and priority of main and worker threads

       ## Scheduling policy options are: other (SCHED_OTHER), batch (SCHED_BATCH)
       ## idle (SCHED_IDLE), fifo (SCHED_FIFO), rr (SCHED_RR)
       scheduler-policy fifo

       ## Scheduling priority is used only for "real-time policies (fifo and rr),
       ## and has to be in the range of priorities supported for a particular
policy
       # scheduler-priority 50
}

buffers {
       ## Increase number of buffers allocated, needed only in scenarios with
       ## large number of interfaces and worker threads. Value is per numa node.
       ## Default is 16384 (8192 if running unpriviledged)
       # buffers-per-numa 128000

       ## Size of buffer data area
       ## Default is 2048
       default data-size 2048

       ## Size of the memory pages allocated for buffer data
       ## Default will try 'default-hugepage' then 'default'
       ## you can also pass a size in K/M/G e.g. '8M'
       # page-size default-hugepage
}

dpdk {
       ## Change default settings for all interfaces
       dev default {
          ## Number of receive queues, enables RSS
          ## Default is 1
          num-rx-queues 4

          ## Number of transmit queues, Default is equal
          ## to number of worker threads or 1 if no workers treads
          num-tx-queues 4

          ## Number of descriptors in transmit and receive rings
          ## increasing or reducing number can impact performance
```

```
        ## Default is 1024 for both rx and tx
        num-rx-desc 4096
        num-tx-desc 4096

        ## VLAN strip offload mode for interface
        ## Default is off
        #vlan-strip-offload off

        ## TCP Segment Offload
        ## Default is off
        ## To enable TSO, 'enable-tcp-udp-checksum' must be set
        tso off

        ## Devargs
        ## device specific init args
        ## Default is NULL
        # devargs safe-mode-support=1,pipeline-mode-support=1

        ## rss-queues
        ## set valid rss steering queues
        rss-queues 0-4
    }

    ## Whitelist specific interface by specifying PCI address and in
    ## addition specify custom parameters for this interface
    dev 0000:3d:11.0 {
         workers 65-68
         name eth1
         num-rx-queues 4
    }
    dev 0000:3d:19.0 {
         workers 69-72
         name eth2
         num-rx-queues 4
    }

    # QAT VFs
    #dev 0000:81:00.4
    #dev 0000:81:00.5
    #dev 0000:81:00.6

    ## Blacklist specific device type by specifying PCI vendor:device
    ## Whitelist entries take precedence
    # blacklist 8086:10fb

    ## Disable multi-segment buffers, improves performance but
    ## disables Jumbo MTU support
    no-multi-seg
    ## Change hugepages allocation per-socket, needed only if there is need for
    ## larger number of mbufs. Default is 256M on each detected CPU socket
```

```
       socket-mem 8192
       ## Disables UDP / TCP TX checksum offload. Typically needed for use
       ## faster vector PMDs (together with no-multi-seg)
       no-tx-checksum-offload
       ## Enable UDP / TCP TX checksum offload
       ## This is the reversed option of 'no-tx-checksum-offload'
       # enable-tcp-udp-checksum
       ## Change UIO driver used by VPP, Options are: igb_uio, vfio-pci,
       ## uio_pci_generic or auto (default)
       uio-driver vfio-pci
       log-level debug
}

memory {
## Set the main heap size, default is 1G
        main-heap-size 1G

        ## Set the main heap page size. Default page size is OS default page
        ## which is in most cases 4K. if different page size is specified VPP
        ## will try to allocate main heap by using specified page size.
        ## special keyword 'default-hugepage' will use system default hugepage
        ## size
        main-heap-page-size 1G
}

## node variant defaults
node {
  ## specify the preferred default variant
  #       default { variant avx512 }

  ## specify the preferred variant, for a given node
  #        ip4-rewrite { variant avx2 }
}


plugins {
        ## Adjusting the plugin path depending on where the VPP plugins are
        #       path /ws/vpp/build-root/install-vpp-native/vpp/lib/vpp_plugins
        path /usr/lib/x86_64-linux-gnu/vpp_plugins

        ## Disable all plugins by default and then selectively enable specific
plugins
        # plugin default { disable }
        plugin dpdk_plugin.so { enable }
        # plugin acl_plugin.so { enable }

        ## For QAT: Comment out all plugins below
        ## For IPSecMB: Uncomment plugin crypto_native_plugin.so
        plugin crypto_ipsecmb_plugin.so { disable }
        # plugin crypto_native_plugin.so { disable }
```

```
        # plugin crypto_sw_scheduler_plugin.so { disable }
}

## Statistics Segment
statseg {
    # socket-name <filename>, name of the stats segment socket
    #     defaults to /run/vpp/stats.sock
    # size <nnn>[KMG], size of the stats segment, defaults to 32mb
    size 32M
    # per-node-counters on | off, defaults to none
    per-node-counters off
    # update-interval <f64-seconds>, sets the segment scrape / update interval
}
```

26. In the directory ep1, create the following file called "startup_qat_sw.conf":

```
unix {
   nodaemon
   log /var/log/vpp/vpp.log
   full-coredump
   cli-listen /run/vpp/cli.sock
   gid vpp
   interactive
   exec /home/ep1/ipsec.cli
}

logging {
  size 4096
  default-log-level debug
  default-syslog-log-level debug
}

api-trace {
  ## This stanza controls binary API tracing. Unless there is a very strong
reason,
  ## please leave this feature enabled.
  on
  ## Additional parameters:
  ##
  ## To set the number of binary API trace records in the circular buffer,
configure nitems
  ##
  ## nitems <nnn>
  ##
  ## To save the api message table decode tables, configure a filename. Results in
/tmp/<filename>
  ## Very handy for understanding api message changes between versions,
identifying missing
  ## plugins, and so forth.
  ##
  ## save-api-table <filename>
}
```

```
api-segment {
  gid vpp
}

socksvr {
  default
}

cpu {
       ## In the VPP there is one main thread and optionally the user can create
worker(s)
       ## The main thread and worker thread(s) can be pinned to CPU core(s)
manually or automatically

       ## Manual pinning of thread(s) to CPU core(s)

       ## Set logical CPU core where main thread runs, if main core is not set
       ## VPP will use core 1 if available
       main-core 64

       ## Set logical CPU core(s) where worker threads are running
       corelist-workers 65-72

       ## Automatic pinning of thread(s) to CPU core(s)

       ## Sets number of CPU core(s) to be skipped (1 ... N-1)
       ## Skipped CPU core(s) are not used for pinning main thread and working
thread(s).
       ## The main thread is automatically pinned to the first available CPU core
and worker(s)
       ## are pinned to next free CPU core(s) after core assigned to main thread
       # skip-cores 4

       ## Specify a number of workers to be created
       ## Workers are pinned to N consecutive CPU cores while skipping "skip-
cores" CPU core(s)
       ## and main thread's CPU core
       # workers 4

       ## Set scheduling policy and priority of main and worker threads

       ## Scheduling policy options are: other (SCHED_OTHER), batch (SCHED_BATCH)
       ## idle (SCHED_IDLE), fifo (SCHED_FIFO), rr (SCHED_RR)
       scheduler-policy fifo

       ## Scheduling priority is used only for "real-time policies (fifo and rr),
       ## and has to be in the range of priorities supported for a particular
policy
       # scheduler-priority 50
}
```

```
buffers {
       ## Increase number of buffers allocated, needed only in scenarios with
       ## large number of interfaces and worker threads. Value is per numa node.
       ## Default is 16384 (8192 if running unpriviledged)
       # buffers-per-numa 128000

       ## Size of buffer data area
       ## Default is 2048
       default data-size 2048

       ## Size of the memory pages allocated for buffer data
       ## Default will try 'default-hugepage' then 'default'
       ## you can also pass a size in K/M/G e.g. '8M'
       # page-size default-hugepage
}

dpdk {
       ## Change default settings for all interfaces
       dev default {
          ## Number of receive queues, enables RSS
          ## Default is 1
          num-rx-queues 4

          ## Number of transmit queues, Default is equal
          ## to number of worker threads or 1 if no workers treads
          num-tx-queues 4

          ## Number of descriptors in transmit and receive rings
          ## increasing or reducing number can impact performance
          ## Default is 1024 for both rx and tx
          num-rx-desc 4096
          num-tx-desc 4096

          ## VLAN strip offload mode for interface
          ## Default is off
          #vlan-strip-offload off

          ## TCP Segment Offload
          ## Default is off
          ## To enable TSO, 'enable-tcp-udp-checksum' must be set
          tso off

          ## Devargs
          ## device specific init args
          ## Default is NULL
          # devargs safe-mode-support=1,pipeline-mode-support=1

          ## rss-queues
          ## set valid rss steering queues
```

```
          rss-queues 0-4
       }

       ## Whitelist specific interface by specifying PCI address and in
       ## addition specify custom parameters for this interface
       dev 0000:3d:11.0 {
            workers 65-68
            name eth1
            num-rx-queues 4
       }
       dev 0000:3d:19.0 {
            workers 69-72
            name eth2
            num-rx-queues 4
       }

       # QAT VFs
       #dev 0000:81:00.4
       #dev 0000:81:00.5
       #dev 0000:81:00.6

       ## Blacklist specific device type by specifying PCI vendor:device
       ## Whitelist entries take precedence
       # blacklist 8086:10fb

       ## Disable multi-segment buffers, improves performance but
       ## disables Jumbo MTU support
       no-multi-seg
       ## Change hugepages allocation per-socket, needed only if there is need for
       ## larger number of mbufs. Default is 256M on each detected CPU socket
       socket-mem 8192
       ## Disables UDP / TCP TX checksum offload. Typically needed for use
       ## faster vector PMDs (together with no-multi-seg)
       no-tx-checksum-offload
       ## Enable UDP / TCP TX checksum offload
       ## This is the reversed option of 'no-tx-checksum-offload'
       # enable-tcp-udp-checksum
       ## Change UIO driver used by VPP, Options are: igb_uio, vfio-pci,
       ## uio_pci_generic or auto (default)
       uio-driver vfio-pci
       log-level debug
}

memory {
## Set the main heap size, default is 1G
        main-heap-size 1G

        ## Set the main heap page size. Default page size is OS default page
        ## which is in most cases 4K. if different page size is specified VPP
        ## will try to allocate main heap by using specified page size.
```

```
        ## special keyword 'default-hugepage' will use system default hugepage
        ## size
        main-heap-page-size 1G
}

## node variant defaults
node {
  ## specify the preferred default variant
  #       default { variant avx512 }

  ## specify the preferred variant, for a given node
  #        ip4-rewrite { variant avx2 }
}


plugins {
        ## Adjusting the plugin path depending on where the VPP plugins are
        #       path /ws/vpp/build-root/install-vpp-native/vpp/lib/vpp_plugins
        path /usr/lib/x86_64-linux-gnu/vpp_plugins

        ## Disable all plugins by default and then selectively enable specific
plugins
        # plugin default { disable }
        plugin dpdk_plugin.so { enable }
        # plugin acl_plugin.so { enable }

        ## For QAT: Comment out all plugins below
        ## For IPSecMB: Uncomment plugin crypto_native_plugin.so
        # plugin crypto_ipsecmb_plugin.so { disable }
         plugin crypto_native_plugin.so { disable }
        # plugin crypto_sw_scheduler_plugin.so { disable }
}

## Statistics Segment
statseg {
    # socket-name <filename>, name of the stats segment socket
    #     defaults to /run/vpp/stats.sock
    # size <nnn>[KMG], size of the stats segment, defaults to 32mb
    size 32M
    # per-node-counters on | off, defaults to none
    per-node-counters off
    # update-interval <f64-seconds>, sets the segment scrape / update interval
}
```

27. In the directory ep1, create the following file called “create_ipsec_cli_ep1.py”:

```
#!/usr/bin/env python3
import sys

def main( argv ):
```

```
crypto_handler       = argv[1]
crypto_alg           = argv[2]
integ_alg            = argv[3]
num_ipsec_tunnels    = argv[4]

if not crypto_handler in ["ipsecmb", "native"]:
    print( "Unknown crypto handler: %s" % (crypto_handler) )
    return 1

if not crypto_alg in ["aes-cbc-128", "aes-gcm-128"]:
    print( "Unknown crypto algorithm: %s" % (crypto_alg) )
    return 1

if not integ_alg in ["none", "sha1-96"]:
    print( "Unknown integrity algorithm: %s" % (integ_alg) )
    return 1

try:
    num_ipsec_tunnels = int( argv[4] )
except:
    print( "Invalid number of IPSec tunnels: %s" % (num_ipsec_tunnels) )
    return 1

if num_ipsec_tunnels <= 0:
    print( "Invalid number of IPSec tunnels: %d" % (num_ipsec_tunnels) )
    return 1

if num_ipsec_tunnels >= 256:
    print( "Invalid number of IPSec tunnels: %d" % (num_ipsec_tunnels) )
    return 1

output_f_name="ipsec.cli"
num_interfaces       = 2
mtu                  = 1518
intf_ip_addrs        = ["172.16.10.3", "192.168.115.2"]
neigh_ip_addrs       = ["172.16.10.2", "192.168.115.3"]
neigh_mac_addrs      = ["<sock1_port2_vf_mac>", "<sock1_port3_vf_mac>"]
neigh_subnets        = ["10.128.0.0/10", "10.64.0.0/10"]
crypto_key           = "4339314b55523947594d6d3547666b45"
integ_key            = "4339314b55523947594d6d3547666b45"
start_sa             = 20
start_spi            = 1000
tunnel_src_prefix    = "10.192.0."
tunnel_dst_prefix    = "10.128.0."
ext_ip_prefix        = "20.64.0."
reverse_spi          = True

lines = []

# Set MTU on interfaces
```

```
    for i in range( 1, num_interfaces + 1 ):
        lines.append( "set interface mtu %d eth%d\n" % (mtu, i) )
    lines.append( "\n" )

    # Set IP addresses on interfaces and enable promiscuous mode
    for i in range( 1, num_interfaces + 1 ):
        next_intf_ip_addr = intf_ip_addrs[i-1]
        lines.append( "set interface ip address eth%d %s/24\n" % (i,
next_intf_ip_addr) )
        lines.append( "set interface promiscuous on eth%d\n" % (i) )
    lines.append( "\n" )

    # Create IPSec Tunnels]
    next_sa = start_sa
    next_spi = start_spi
    for i in range( 1, num_ipsec_tunnels + 1 ):
        next_ip_tunnel_name = "ipip%d" % ( i - 1 )
        lines.append( "create ipip tunnel src %s%d dst %s%d\n" %
(tunnel_src_prefix, i, tunnel_dst_prefix, i) )
        lines.append( "set interface state ipip%d up\n" % (i - 1) )
        if reverse_spi == True:
            lines.append( "ipsec sa add %d spi %d crypto-alg %s crypto-key %s
integ-alg %s" % (next_sa, next_spi + 1, crypto_alg, crypto_key, integ_alg) )
        else:
            lines.append( "ipsec sa add %d spi %d crypto-alg %s crypto-key %s
integ-alg %s" % (next_sa, next_spi, crypto_alg, crypto_key, integ_alg) )
        if integ_alg != "none":
            lines[-1] += " integ-key %s" % ( integ_key )
        lines[-1] += "\n"
        next_sa += 1
        next_spi += 1
        if reverse_spi == True:
            lines.append( "ipsec sa add %d spi %d crypto-alg %s crypto-key %s
integ-alg %s" % (next_sa, next_spi - 1, crypto_alg, crypto_key, integ_alg) )
        else:
            lines.append( "ipsec sa add %d spi %d crypto-alg %s crypto-key %s
integ-alg %s" % (next_sa, next_spi, crypto_alg, crypto_key, integ_alg) )
        if integ_alg != "none":
            lines[-1] += " integ-key %s" % ( integ_key )
        lines[-1] += "\n"
        lines.append( "ipsec tunnel protect %s sa-in %d sa-out %d\n" %
(next_ip_tunnel_name, next_sa-1, next_sa) )
        lines.append( "set interface ip address %s %s%d/32\n" %
(next_ip_tunnel_name, tunnel_src_prefix, i) )
        lines.append( "ip route add %s%d/32 via %s\n" % (ext_ip_prefix, i,
next_ip_tunnel_name) )
        lines.append( "\n" )
        next_sa += 1
        next_spi += 1

    # Set ARP table entries and IP route entries
    for i in range( 1, num_interfaces + 1 ):
```

```
        next_neigh_ip_addr = neigh_ip_addrs[i-1]
        next_neigh_mac_addr = neigh_mac_addrs[i-1]
        next_neigh_subnet = neigh_subnets[i-1]
        lines.append( "set ip neighbor eth%d %s %s\n" % (i, next_neigh_ip_addr,
next_neigh_mac_addr) )
        lines.append( "ip route add %s via %s eth%d\n" % (next_neigh_subnet,
next_neigh_ip_addr, i) )
    lines.append( "\n" )

    # Set interface state up
    for i in range( 1, num_interfaces + 1 ):
        lines.append( "set interface state eth%d up\n" % (i) )
    lines.append( "\n" )

    lines.append( "set crypto handler all openssl\n")
    lines.append( "set crypto handler all %s\n" % (crypto_handler) )

    with open( output_f_name, 'w' ) as output_file:
        output_file.writelines( lines )

if __name__ == "__main__":
    sys.exit( main(sys.argv) )
```

a. Replace <sock1_port2_vf_mac> and <sock1_port3_vf_mac> with the MAC addresses of the VFs of the corresponding NIC ports.

28. In the directory ep1, run the following command to generate the "ipsec.cli" file to use for VPP:

    ```
    python3 create_ipsec_cli_ep1.py ipsecmb aes-cbc-128 sha1-96 128
    ```

29. Run the "setup_vpp.sh" script.

30. In one console window, run the following commands to start the VPP endpoints:

    ```
    oc exec vpp-ep0 -- /home/ep0/start_vpp_native.sh &
    oc exec vpp-ep1 -- /home/ep1/start_vpp_native.sh &
    ```

31. In a separate console window, run the following commands to start the DPDK testpmd application:

    ```
    oc exec -it dpdk-pod /bin/bash
    ./run_testpmd.sh
    ```

32. In a separate console window, run the following commands to start the T-Rex traffic generator:

    ```
    oc exec -it trex-pod /bin/bash
    cd /home/v3.00
    tmux new-session
    ./t-rex-64 -i
    <detach from the tmux-session>
    ./trex-console
    ```

33. In the T-Rex console interface, use the following commands:

```
tui
start -f stl/vpp_packets_p0.py -p 0 -m 5%
start -f stl/vpp_packets_p1.py -p 1 -m 5%
```

34. Repeat step 31 by using "stop" and changing the line rate parameter (-m 5%) to find the highest rate with zero packet loss at steady state. Observe the throughput results and record the total Tx pps amount.
35. To stop the traffic, use the command "stop" and exit from the T-Rex console interface with the "quit" command twice.
36. Repeat steps 33-35 after modifying the "stl/vpp_packets_p0.py" and "stl_vpp_packets_p1.py" files to use different packet sizes according to the baseline results provided in Section 6.4.
37. Use the following commands to stop the VPP endpoints:

```
oc exec vpp-ep0 -- pkill vpp
oc exec vpp-ep1 -- pkill vpp
```

38. Repeat the testing from steps 33-36 after using the following commands to start the second VPP configuration:

```
oc exec vpp-ep0 -- /home/ep0/start_vpp_qat_sw.sh &
oc exec vpp-ep1 -- /home/ep1/start_vpp_qat_sw.sh &
```

5.5 Security AI

AI inference is used in network/security to help prevent advanced cyber-attacks. In order to improve the latency associated with this application, the Intel® Xeon® Scalable Processor contains technologies to accelerate AI inference such as AVX-512 and Vector Neural Network Instructions. The Security AI workload utilizes the TensorFlow deep-learning framework, Intel® oneAPI Deep Neural Network Library (oneDNN), and Intel® Neural Compressor to improve the performance of the AI inference model.

The starting model for the security AI is an open-source deep-learning model called Malconv (https://github.com/elastic/ember/tree/master/malconv) which is given as a pre-trained Keras H5 format file. This model is used to detect malware within files. The performance of the model can be improved by converting it to an FP32 frozen graph model, and further still by enabling Intel® oneDNN. Finally, by using the Intel® Neural Compressor for post-training quantization, the model can become a frozen INT8 model.

For a full explanation of the security AI workload and test procedure, please refer to the following document "Intel® Deep Learning Boost – Boost Network Security AI Inference Performance in Google Cloud Platform (GCP) Technology Guide." Ensure that the results of the tests follow the expected results as shown in the following tables to baseline the performance of the platform.

Table 11. Security AI Workload Configuration

Table 12. Plus Platform Security AI Performance Requirements

Model Type	Inference Time (ms)
Keras h5	86.9
FP32 Frozen without oneDNN	65.0

FP32 Frozen with oneDNN	51.3
INT8 Frozen with oneDNN	25.6

5.5.1 Security AI Test Methodology

Follow the instructions provided below to run the Security AI inference models:

1. Create a requirements.txt file containing the required package dependencies:

```
TensorFlow
TensorFlow-estimator
keras
neural-compressor
progress
numpy
```

2. Create a script called h5_to_savedmodel.py to first convert the Keras h5 format to the SavedModel format:

```
import tensorflow as tf
model = tf.keras.models.load_model("malconv.h5")
model.save("malconv_saved_model")
```

3. Create a script called saved_model_to_frozen.py to convert the SavedModel to the FP32 frozen graph model:

```
import sys
import tensorflow as tf
from tensorflow.python.saved_model import signature_constants
from tensorflow.python.training import saver
from tensorflow.python.framework import convert_to_constants
from tensorflow.core.protobuf import config_pb2
from tensorflow.python.grappler import tf_optimizer
from tensorflow.core.protobuf import meta_graph_pb2
from tensorflow.python.platform import gfile
from tensorflow.python.eager import context
assert context.executing_eagerly()
if len(sys.argv) != 3:
    print('Usage:')
    print(f'\tpython3 {sys.argv[0]} model_path output_pbfile')
    sys.exit(1)

model = tf.keras.models.load_model(sys.argv[1])
model.summary()

func = model.signatures[signature_constants.DEFAULT_SERVING_SIGNATURE_DEF_KEY]
frozen_func = convert_to_constants.convert_variables_to_constants_v2(func)

grappler_meta_graph_def = saver.export_meta_graph(
    graph_def=frozen_func.graph.as_graph_def(), graph=frozen_func.graph)

fetch_collection = meta_graph_pb2.CollectionDef()
for array in frozen_func.inputs + frozen_func.outputs:
    fetch_collection.node_list.value.append(array.name)
grappler_meta_graph_def.collection_def["train_op"].CopyFrom(fetch_collection)

grappler_session_config = config_pb2.ConfigProto()
rewrite_options = grappler_session_config.graph_options.rewrite_options
rewrite_options.min_graph_nodes = -1
opt = tf_optimizer.OptimizeGraph(grappler_session_config,
grappler_meta_graph_def, graph_id=b"tf_graph")

f = gfile.GFile(sys.argv[2], 'wb')
f.write(opt.SerializeToString())
```

4. Create the following YAML config file called malconv.yaml:

```
model:
  name: malconv
  framework: tensorflow
  inputs: input_1
  outputs: Identity

tuning:
  accuracy_criterion:
    relative:  0.01                                    # optional. default value
is relative, other value is absolute. this example allows relative accuracy
loss: 1%.
  exit_policy:
    timeout: 0                                         # optional. tuning
timeout (seconds). default value is 0 which means early stop. combine with
max_trials field to decide when to exit.
    max trials: 1
  random_seed: 9527                                    # optional. random seed
for deterministic tuning.
```

5. Create the following script called quantize.py to quantize the FP32 frozen model with Intel® Neural Compressor:

```
import os
import argparse
import numpy as np
from neural_compressor.experimental import Quantization, common

def parse_args():
    parser = argparse.ArgumentParser()
    parser.add_argument(
        '-m', '--input_model', type=str, dest='input_model', help='frozen fp32
model', required=True)
    parser.add_argument(
        '-c', '--input_config', type=str, dest='input_config', help='yaml
config file', required=True)
    parser.add_argument(
        '-i', '--input_path', type=str, dest='input_path', help='input
dataset', required=True)
    parser.add_argument(
        '-o', '--output_file', type=str, dest='output_file', help='output
file', required=True)
    args = parser.parse_args()
    return args

def load_dataset(input_path):
    result = []
    mal_path = os.path.join(input_path, 'MALICIOUS')
    if os.path.exists(mal_path):
        mal files = [(1, os.path.join(mal path, fp)) for fp in
os.listdir(mal_path)]
        result.extend(mal_files)
    clean_path = os.path.join(input_path, 'KNOWN')
    if os.path.exists(clean_path):
        clean_files = [(0, os.path.join(clean_path, fp)) for fp in
os.listdir(clean_path)]
        result.extend(clean_files)

    return result
def read_file(filepath: str, expect_size: int):
    if filepath[-4:] == '.npy':
        data = np.load(filepath, allow_pickle=True)
    else:
        data = np.fromfile(filepath, np.ubyte)

    if data.size < expect_size:
```

```
        data = np.pad(data, (0, expect_size - data.size), 'constant',
constant_values=(0, 0))
    else:
        data = data[:expect_size]

    return np.array([data])

class Dataset:
    def __init__(self, input_path):
        self.batch_size = 32
        self.dataset = load_dataset(input_path)

    def __iter__(self):
        for label, filepath in self.dataset:
            data = read_file(filepath, expect_size=1048576)
            yield data, label

    def __len__(self):
        return len(self.dataset)

if __name__ == '__main__':
    os.environ['TF_ENABLE_ONEDNN_OPTS'] = '1'
    args = parse_args()
    quantizer = Quantization(args.input_config)
    quantizer.model = common.Model(args.input_model)
    quantizer.calib_dataloader = Dataset(args.input_path)
    quantizer().save(args.output_file)
```

6. Use your Docker login information to run the following command:
   `podman login docker.io`

7. Create the file "analyze_scores.py":

```
"""
analyze_scores.py
*** Code has been significantly cropped for public demo release ***
@author: Brody Kutt (bkutt@paloaltonetworks.com)
"""

import os
import csv
import argparse
import numpy as np
from collections import defaultdict
from sklearn.metrics import f1_score, roc_curve, roc_auc_score,
confusion_matrix


def float2string(inp):
    return ('%.15f' % inp).rstrip('0').rstrip('.')


def format_predict_data(fields, prediction):
    result = defaultdict(list)

    for row in prediction:
        for field, value in zip(fields, row):
            result[field].append(value)

    result['Score'] = [float(x) for x in result['Score']]
    result['Predict'] = [int(x) for x in result['Predict']]
    result['Actual'] = [int(x) for x in result['Actual']]

    return result


def read_predict_file(path: str) -> dict:
```

```
    with open(path, 'r') as csv_file:
        csv_reader = csv.reader(csv_file)
        _ = next(csv_reader)
        fields = next(csv_reader)
        return format_predict_data(fields, csv_reader)


def apply_threshold(scores: list, threshold: float) -> list:
    return [int(score >= threshold) for score in scores]


def recall_specificity_at_thresh(y_scores, y_test, threshold,
adjusted_ben=None, adjusted_mal=None):
    """
    Return the highest recall possible when the specificity is set to 100%
    i.e. there are no FPs. If 100% specificity isn't possible even with a
    maximum threshold, -1 will be returned.
        Parameters:
        y_scores: array-like, shape (n_samples), ensemble scores
        y_test: array-like, shape (n_samples), test labels
        threshold: float, the decision threshold
        adjusted_ben: int, adjusted total number of benign files (when
                           measuring adjusted performance)
        adjusted_ben: int, adjusted total number of malicious files (when
                           measuring adjusted performance)
        Returns:
        recall, specificity: float, float; the recall and specificity
    """
    predict_discrete = apply_threshold(y_scores, threshold)
    cm = confusion_matrix(y_test, predict_discrete, labels=[0, 1])

    if adjusted_ben:
        assert adjusted_ben >= (cm[0][0] + cm[0][1])
        cm[0][0] += (adjusted_ben - (cm[0][0] + cm[0][1]))  # Count them as
TNs

    if adjusted_mal:
        assert adjusted_mal >= (cm[1][0] + cm[1][1])
        cm[1][0] += (adjusted_mal - (cm[1][0] + cm[1][1]))  # Count them as
FNs

    tn, fp, fn, tp = cm.ravel()

    recall = (tp * 100.0) / float(fn + tp) if float(fn + tp) != 0 else 100.0
    specificity = (tn * 100.0) / float(fp + tn) if float(fp + tn) != 0 else
100.0

    return recall, specificity


def parse_args():
    parser = argparse.ArgumentParser(
        description='Do analysis on computed malicious class scores.')
    parser.add_argument(
        '--pred_fps',
        nargs='+',
        help=('All filepaths leading to prediction files you wish compare. '
              'Separate each with a space.'))
    parser.add_argument(
        '--labels',
        help=('Labels for prediction files you wish to see in the plot '
              'legend. Supply one for each prediction file. Separate each '
              'with a comma.'))
    parser.add_argument(
        '--cust_threshs',
        default='',
```

```
        required=False,
        metavar='thresh1,thresh2,...',
        help=('All custom threhsolds you would like to test. Separate each '
              'with a comma.'))
    parser.add_argument(
        '--ref_fprs',
        default='',
        required=False,
        metavar='fpr1,fpr2,...',
        help=('All FPRs which you want to discover corresponding recall. '
              'Separate each with a comma.'))
    return parser.parse_args()


class Analyzer:
    def __init__(self, **kwargs):
        self.roc_data = []
        self.labels = kwargs['labels']
        self.ref_fprs = kwargs['ref_fprs']
        self.custom_thresholds = kwargs['custom_thresholds'] if
'custom_thresholds' in kwargs else None

        if 'pred_fps' in kwargs:
            self.all_data = self._read_predict_files(kwargs['pred_fps'])
        else:
            self.all_data = kwargs['all_data']

    def run(self):
        self._print_header()
        self._compute_custom_threshold_stats()
        self._compute_roc_curves()
        self._compute_tprs()
        self._print_tail()

    @staticmethod
    def _print_header():
        print('-' * 80)

    @staticmethod
    def _print_tail():
        print('\nExiting...')
        print('-' * 80)

    def _read_predict_files(self, pred_fps):
        print('Reading in predictions files...')
        result = []
        for label, filepath in zip(self.labels, pred_fps):
            result.append(read_predict_file(filepath))
            print(f'\tRead in {label} predictions: {filepath}!')
        return result

    def _compute_custom_threshold_stats(self) -> None:
        if not self.custom_thresholds:
            return

        print('\nComputing custom threshold stats...')
        for label, data in enumerate(self.labels, self.all_data):
            print(f'\n--> Using predictions with label \'{label}\':')
            for threshold in self.custom_thresholds:
                print(f'----> Stats using custom threshold
{float2string(threshold)}...')
                r, s = recall_specificity_at_thresh(data['Score'],
data['Actual'], threshold)
                predict = apply_threshold(data['Score'], thresh)
                f1 = f1_score(data['Actual'], predict)
```

```
                print(f'------> Recall: {r:%.6f}, Specificity: {s:%.6f}, F1:
{f1:%.6f}')

    def _compute_roc_curves(self):
        print('\nComputing ROC curves...')
        for label, data in zip(self.labels, self.all_data):
            print(f'\n--> Using predictions with label \'{label}\':')
            fpr, tpr, thresholds = roc_curve(data['Actual'], data['Score'])
            self.roc_data.append((fpr, tpr, thresholds))
            new_data = data['Score']
            new_data = np.array(new_data, dtype=np.float16)
            auc = roc_auc_score(data['Actual'], new_data)
            print(f'----> ROC AUC: {float2string(auc)}')

    def _compute_tprs(self):
        if not self.ref_fprs:
            return

        print('\nComputing TPRs at reference FPRs...')
        for label, data, (fpr, tpr, thresholds) in zip(self.labels,
self.all_data, self.roc_data):
            print(f'\n--> Using predictions with label \'{label}\':')
            for ref_fpr in self.ref_fprs:
                print(f'----> Stats using reference FPR <=
{float2string(ref_fpr)}...')
                idx = np.sum(fpr <= ref_fpr) - 1
                print(f'------> Threshold: {thresholds[idx]:.10f}, FPR:
{fpr[idx]:.10f}, TPR {tpr[idx]:.10f}')


def analyze_scores(**kwargs):
    if 'pred_fps' in kwargs or 'all_data' in kwargs:
        Analyzer(**kwargs).run()
    else:
        raise Exception('No prediction file or data found!')


def main():
    args = parse_args()

    for fp in args.pred_fps:
        assert (os.path.isfile(fp))

    analyze_scores(
        pred_fps=args.pred_fps,
        labels=args.labels.strip().split(','),
        ref_fprs=[float(i) for i in args.ref_fprs.strip().split(',') if i !=
''],
        custom_thresholds=[float(i) for i in
args.cust_threshs.strip().split(',') if i != ''],
    )


if __name__ == '__main__':
    # from pudb import set_trace
    # set_trace()
    main()
```

8. Create the file "ai_run.sh":

```
#!/bin/bash

cd /home

set -e

num_loop=$(seq 3)
```

```
dataset=/home/datasets

export TF_CPP_MIN_LOG_LEVEL=3

function random_sleep {
    echo ""
    SEC=$[($RANDOM % 5) + 10 ]
    echo "[$(date +"%Y-%m-%d %T")] Sleep $SEC seconds..."
    eval "sleep $SEC"
}

function exec_cmd {
    echo "[$(date +"%Y-%m-%d %T")] Running CMD: $*"
    eval $*
    STATUS=$?
}

declare -a tf_models=("malconv_int8.pb" "malconv_fp32.pb" "malconv.h5")

for model in "${tf_models[@]}"
do
    random_sleep
    for core in {0..31}
    do
        exec_cmd "TF_ENABLE_ONEDNN_OPTS=1 numactl --physcpubind=$core python
test.py -i $dataset -m $model > logs/${model}_core${core}.log" &
    done

    # wait for all the commands to finish
    while [ $(ps -ef | grep "python test.py" | wc -l) != 1 ];
    do
        sleep 1
    done
done

random_sleep
for core in {0..31}
do
        exec_cmd "TF_ENABLE_ONEDNN_OPTS=0 numactl --physcpubind=$core python
test.py -i $dataset -m malconv_fp32.pb >
logs/malconv_fp32.pb_nodnn_core${core}.log" &
done
# wait for all the commands to finish
while [ $(ps -ef | grep "python test.py" | wc -l) != 1 ];
do
    sleep 1
done

cd logs
cat *h5*.log | grep "average inference time" | awk -v N=$4 '{print $4}' >>
h5_results.log
cat *fp32*nodnn*.log | grep "average inference time" | awk -v N=$4 '{print
$4}' >> nodnn_fp32_results.log
cat malconv_fp32.pb_core* | grep "average inference time" | awk -v N=$4
'{print $4}' >> fp32_results.log
cat *int8*.log | grep "average inference time" | awk -v N=$4 '{print $4}' >>
int8_results.log
rm *core*.log
```

a. In the lines "for core in {0..31}", change the 31 to the number of cores per socket in your system minus 1

9. Create the file "test.py":

```
import os
import time
import pickle
import socket
```

```
import argparse
import numpy as np
from progress.bar import Bar
from analyze_scores import analyze_scores


def load_dataset(input_path):
    result = []
    mal_path = os.path.join(input_path, 'MALICIOUS')
    if os.path.exists(mal_path):
        mal_files = [(1, os.path.join(mal_path, fp)) for fp in
os.listdir(mal_path)]
        result.extend(mal_files)

    clean_path = os.path.join(input_path, 'KNOWN')
    if os.path.exists(clean_path):
        clean_files = [(0, os.path.join(clean_path, fp)) for fp in
os.listdir(clean_path)]
        result.extend(clean_files)

    return result


def read_file(filepath, expect_size=1048576):
    if filepath[-4:] == '.npy':
        data = np.load(filepath, allow_pickle=True)
    else:
        data = np.fromfile(filepath, np.ubyte)
    if data.size < expect_size:
        data = np.pad(data, (0, expect_size - data.size), 'constant',
constant_values=(0, 0))
    else:
        data = data[:expect_size]

    return np.array([data])


class H5Model:
    def __init__(self, h5_path):
        import tensorflow as tf
        self.model = tf.keras.models.load_model(h5_path)

    def predict(self, input_data):
        start = time.time()
        result = self.model.predict(input_data)
        finish = time.time()
        return result[0], 1000 * (finish - start)


class SavedModel:
    def __init__(self, save_model_dir):
        import tensorflow as tf
        self.session = tf.compat.v1.Session()
        meta_graph_def = tf.compat.v1.saved_model.loader.load(self.session,
['serve'], save_model_dir)
        signature = meta_graph_def.signature_def
        serving_default = signature['serving_default']
        self.X = serving_default.inputs['input_1'].name
        self.y = serving_default.outputs['dense_2'].name

    def predict(self, input_data):
        start = time.time()
        result = self.session.run(self.y, feed_dict={self.X: input_data})
        finish = time.time()
        return result[0][0], 1000 * (finish - start)
```

```
class ONNXModel:
    def __init__(self, onnx_file, num_cores=-1):
        import onnxruntime as ort
        self.onnx_file = onnx_file
        if num_cores > 0:
            sess_options = ort.SessionOptions()
            sess_options.intra_op_num_threads = num_cores
            sess_options.execution_mode = ort.ExecutionMode.ORT_SEQUENTIAL
            sess_options.graph_optimization_level =
ort.GraphOptimizationLevel.ORT_ENABLE_ALL
            self.session = ort.InferenceSession(self.onnx_file,
sess_options=sess_options, providers=['CPUExecutionProvider'])
        else:
            print('ONNXModel: not specify session options')
            self.session = ort.InferenceSession(self.onnx_file,
providers=['CPUExecutionProvider'])
        self.input_name = self.session.get_inputs()[0].name
        self.output_name = self.session.get_outputs()[0].name

    def predict(self, input_data):
        float32_data = input_data.astype(np.float32)
        start = time.time()
        result = self.session.run([self.output_name], {self.input_name:
float32_data})
        finish = time.time()
        return result[0][0][0], 1000 * (finish - start)


class FrozenModel:
    def __init__(self, pb_filepath, intra_cores=-1, inter_cores=-1,
config=None):
        import tensorflow as tf
        graph = tf.Graph()
        if intra_cores > 0 and inter_cores > 0:
            print(f'FrozenModel: intra_op_parallelism_threads={intra_cores},
inter_op_parallelism_threads={inter_cores}')
            config =
tf.compat.v1.ConfigProto(intra_op_parallelism_threads=intra_cores,
inter_op_parallelism_threads=inter_cores)
        else:
            print('FrozenModel: not specify config options')
        with graph.as_default():
            graph_def = tf.compat.v1.GraphDef()
            with open(pb_filepath, "rb") as f:
                self.model_path = pb_filepath
                graph_def.ParseFromString(f.read())
                _ = tf.import_graph_def(graph_def, name='')
                self.session = tf.compat.v1.Session(config=config,
graph=graph)
                self.input_t1 = graph.get_tensor_by_name("input_1:0")
                self.output_data = graph.get_tensor_by_name("Identity:0")

    def predict(self, input_data):
        start = time.time()
        result = self.session.run(self.output_data, feed_dict={self.input_t1:
input_data})
        finish = time.time()
        return result[0][0], 1000 * (finish - start)


class TFLiteModel:
    def __init__(self, tflite_file):
        import tensorflow as tf
        self.interpreter = tf.lite.Interpreter(tflite_file)
        self.interpreter.allocate_tensors()
```

```
    def predict(self, input_data):
        start = time.time()
        input_data = input_data.astype(np.float32)
        input_details = self.interpreter.get_input_details()
        output_details = self.interpreter.get_output_details()
        self.interpreter.set_tensor(input_details[0]['index'], input_data)
        self.interpreter.invoke()
        result = self.interpreter.get_tensor(output_details[0]['index'])
        finish = time.time()
        return result[0][0], 1000 * (finish - start)


class TestOnDataset:
    def __init__(self, model, input_path):
        self.model = model
        self.threshold = 0.99
        self.avg_infer_time = None
        self.standard_deviation = None
        self.input_path = input_path

        if os.path.isdir(input_path):
            self.all_files = load_dataset(input_path)
            self.total = len(self.all_files)
        else:
            with open(input_path, 'rb') as file:
                self.data_list = pickle.load(file)
            self.total = len(self.data_list)

    def run(self):
        def create_generator():
            if os.path.isdir(self.input_path):
                for label, filepath in self.all_files:
                    yield label, read_file(filepath)
            else:
                for label, int8_data in self.data_list:
                    yield label, int8_data

        generator = create_generator()
        input_tensor = next(generator)[1]
        for _ in range(30):
            self.model.predict(input_tensor)

        all_infer_time = []
        files, scores, pred, all_y = [], [], [], []
        bar = Bar('Progress... ', max=self.total)
        generator = create_generator()
        for label, int8_data in generator:
            score, infer_time = self.model.predict(int8_data)
            all_infer_time.append(infer_time)
            files.append('-')
            scores.append(score)
            pred.append(int(score >= self.threshold))
            all_y.append(label)
            bar.next()
        bar.finish()

        self.avg_infer_time = np.mean(all_infer_time)
        self.standard_deviation = np.std(all_infer_time)

        print(f'average inference time: {self.avg_infer_time} ms')
        print(f'standard deviation: {self.standard_deviation} ms')
        print(f'filecount: {self.total}')

        analyze_scores(
```

```
            all_data=[{'Filename': files, 'Score': scores, 'Predict': pred,
'Actual': all_y}],
            labels=[socket.gethostname()],
            ref_fprs=[0.05, 0.01],
        )

    def dump_dataset(self, dump_path):
        if dump_path is None:
            return

        data_list = []
        for label, filepath in self.all_files:
            int8_data = read_file(filepath)
            data_list.append((label, int8_data))

        with open(dump_path, 'wb') as file:
            pickle.dump(data_list, file)


def load_model(model_path, args):
    if model_path[-2:] == 'h5':
        return H5Model(model_path)
    if model_path[-4:] == 'onnx':
        return ONNXModel(model_path, args.num_cores)
    if model_path[-6:] == 'tflite':
        return TFLiteModel(model_path)
    if os.path.isdir(model_path):
        return SavedModel(model_path)
    return FrozenModel(model_path, args.intra_cores, args.inter_cores)


def main():
    parser = argparse.ArgumentParser()
    parser.add_argument(
        '-m', '--model_path', type=str, dest='model_path', help='model path',
required=True)
    parser.add_argument(
        '-i', '--input_path', type=str, dest='input_path', help='input dataset
path', required=True)
    parser.add_argument(
        '-c', '--num_cores', type=int, dest='num_cores', help='number of cores
for ONNX runtime', default=-1)
    parser.add_argument(
        '-a', '--intra_cores', type=int, dest='intra_cores',
help='intra_op_parallelism_threads', default=-1)
    parser.add_argument(
        '-b', '--inter_cores', type=int, dest='inter_cores',
help='inter_op_parallelism_threads', default=-1)
    parser.add_argument(
        '-d', '--dump_path', type=str, dest='dump_path', help='save dataset
folder to pickle file', default=None)

    args = parser.parse_args()
    model = load_model(args.model_path, args)
    test = TestOnDataset(model, args.input_path)
    test.run()
    test.dump_dataset(args.dump_path)


if __name__ == '__main__':
    main()
```

10. Create the file “run_nfvi_ai.sh”:

```
#!/bin/bash

oc delete pod nfvi-ai
```

```
oc create -f nfvi_ai_pod.yaml
sleep 20

oc cp requirements.txt nfvi-ai:/home
oc cp h5_to_savedmodel.py nfvi-ai:/home
oc cp saved_model_to_frozen.py nfvi-ai:/home
oc cp quantize.py nfvi-ai:/home
oc cp malconv.yaml nfvi-ai:/home
oc cp datasets/ nfvi-ai:/home
oc cp setup_ai.sh nfvi-ai:/home
oc cp test.py nfvi-ai:/home
oc cp analyze_scores.py nfvi-ai:/home
oc cp ai_run.sh nfvi-ai:/home

oc exec nfvi-ai -- /home/setup_ai.sh
oc exec nfvi-ai -- /home/ai_run.sh

oc cp nfvi-ai:/home/logs logs/
echo "Security AI Mean Inference Times" > nfvi_ai.log
awk '{ sum += $1 } END { print "Keras h5:",sum/NR,"ms" }' logs/h5_results.log
>> nfvi_ai.log
awk '{ sum += $1 } END { print "FP32 without oneDNN:",sum/NR,"ms" }'
logs/nodnn_fp32_results.log >> nfvi_ai.log
awk '{ sum += $1 } END { print "FP32 with oneDNN:",sum/NR,"ms" }'
logs/fp32_results.log >> nfvi_ai.log
awk '{ sum += $1 } END { print "INT8 with oneDNN:",sum/NR,"ms" }'
logs/int8_results.log >> nfvi_ai.log
```

11. Create the file "nfvi_ai_pod.yaml":

```
apiVersion: v1
kind: Pod
metadata:
  labels:
    run: nfvi-ai
  name: nfvi-ai
spec:
  containers:
  - args:
    - /bin/bash
    image: python
    name: nfvi-ai
    stdin: true
  nodeName: worker-0
  restartPolicy: Always
```

    a. Replace "worker-0" with the name of the worker node you are using

12. Create the file "setup_ai.sh":

```
#!/bin/bash

apt update ; apt install -y libgl1 numactl
cd /home
mkdir logs
wget https://raw.githubusercontent.com/elastic/ember/master/malconv/malconv.h5
pip install --upgrade pip
pip install -r requirements.txt
pip install neural-compressor
pip install numpy
python h5_to_savedmodel.py
python saved_model_to_frozen.py malconv_saved_model/ malconv_fp32.pb
TF_ENABLE_ONEDNN_OPTS=1 python quantize.py -m malconv_fp32.pb -c malconv.yaml
-i ./datasets -o malconv_int8.pb
```

13. You will need to provide your own testing dataset to use. Create the following directories:

```
mkdir -p datasets/KNOWN
mkdir datasets/MALICIOUS
```

14. Place the benign files into the “datasets/KNOWN” directory, and place the malicious files in the “datasets/MALICIOUS” directory
15. Run the “run_nfvi_ai.sh” script. The generated “nfvi_ai.log” file will contain the mean inference time results of the four models tested.

§

6 Summary

The Intel NFVI v4 and SASE Verified Reference Configuration, is an Intel Accelerated Solution defined on 4th Gen Intel® Xeon® Scalable processors. This solution, combined with architectural improvements, feature enhancements, integrated Accelerators with high memory and IO bandwidth, provides a significant performance and scalability advantage in today's NFVI environments. These processors are optimized for network, cloud native, wireline, and wireless core-intensive workloads, and are especially suited for NFVI and SASE workload coupled using Intel® Ethernet E810-Network Controllers and Data Plane Development Kit.