

QxStack NFV Infrastructure with Red Hat OpenStack Platform

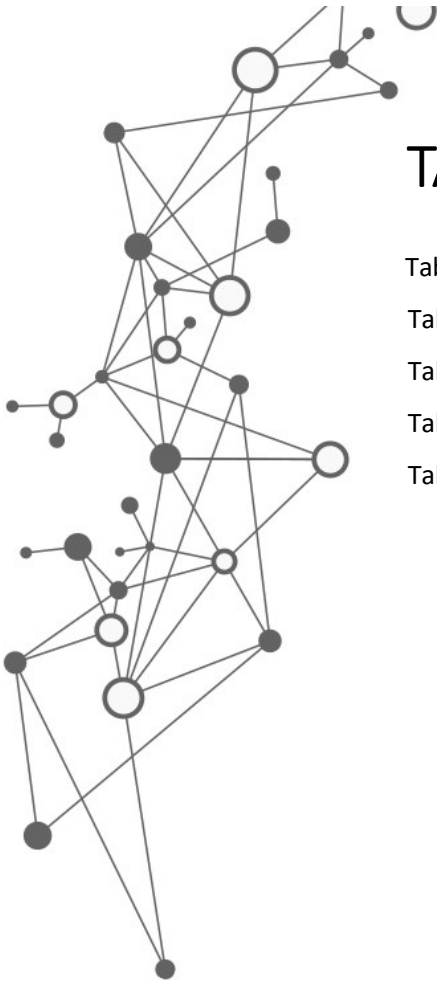
A Carrier-Grade Infrastructure Pre-Integrated and Validated
for Network Service Providers





CONTENT

CONTENT.....	1
TABLES	2
FIGURES.....	3
1. OVERVIEW.....	4
2. STANDARD ARCHITECTURE	5
2.1. HARDWARE.....	5
2.1.1. QUANTAGRID D52B-1U: DIRECTOR NODE	5
2.1.2. QUANTAGRID D52BQ-2U: CONTROLLER AND COMPUTE NODE.....	6
2.1.3. QUANTAGRID D52T-1ULH: STORAGE NODE	6
2.1.4. QUANTAMESH T1048-LY4R: MANAGEMENT SWITCH	7
2.1.5. QUANTAMESH BMS T4048-IX8D: TOP-OF-RACK (TOR) SWITCH.....	8
2.2. SOLUTION SOFTWARE	9
2.2.1. RED HAT OPENSTACK PLATFORM 13	9
2.2.2. NETWORK TOPOLOGY	9
2.2.3. EPA FEATURES	10
2.2.4. HIGH AVAILABILITY	13
3. SELECTED FEATURES	14
3.1. QUICK-ASSIST TECHNOLOGY	14
3.2. REAL-TIME KVM	14
3.3. QCT OPERATIONAL TOOLS	15
3.3.1. QXSTACK AUTO-DEPLOYMENT TOOL	15
3.3.2. TESTBOX	16
3.3.3. QCT SERVICE ASSURANCE FRAMEWORK.....	18
4. PERFORMANCE EVALUATION	21
4.1. TESTBOX PERFORMANCE RESULTS	21
4.2. ADDITIONAL ENHANCEMENT.....	23
5. CONCLUSION.....	26
LEGAL DISCLAIMER.....	26



TABLES

Table 1. Solution hardware - QuantaGrid D52B-1U	5
Table 2. Solution hardware - QuantaGrid D52BQ-2U	6
Table 3. Solution hardware - QuantaGrid D52T-1ULH	7
Table 4. Solution hardware - QuantaMesh T1048-LY4R	7
Table 5. Solution hardware - QuantaMesh T4048-IX8D.....	8



FIGURES

Figure 1. Example of Network Layout.	10
Figure 2. CPU Topology and Core Allocation.	11
Figure 3. NUMA-Balanced Hardware Design.	11
Figure 4. Comparison among Network Virtualization Stacks.	12
Figure 5. RT-KVM Architecture.	14
Figure 6. QxStack Auto-Deployment Tool – Surveyform.	15
Figure 7. QxStack Auto-Deployment Tool – Workflow.	16
Figure 8. TestBox Architecture.	17
Figure 9. QCT Service Assurance Framework.	19
Figure 10. Grafana Panels with Telemetry Metrics.	20
Figure 11. Functional Healthcheck Results.	21
Figure 12. East-West and North-South Testing Topologies.	21
Figure 13. Comparison between with/without EPA Features based on OVS-DPDK using Yardstick.	22
Figure 14. Comparison between with/without EPA Features based on OVS-DPDK using NFVBench.	22
Figure 15. Comparison between with/without EPA Features based on SR-IOV using Yardstick.	23
Figure 16. Comparison between with/without EPA Features based on SR-IOV using NFVBench.	23
Figure 17. QAT Validation Topology.	24
Figure 18. Comparison between with/without QAT on SR-IOV Compute Nodes.	24
Figure 19. Comparison of Interrupt Latency between with/without real-time KVM.	25
Figure 20. Comparison of Computing Jitter with/without real-time KVM.	25

1. Overview

The development of high-speed network has been changing the world through enhancing data transmission and information analysis. Since data is explosively generated, the need of new infrastructure and the establishment of 5G network are accelerated. Facing the inevitable digital transformation, Communication Service Providers (CSPs) are looking for emerging software-defined infrastructure to easily manage network resources and keep a dominant position in the telco market.

Network Function Virtualization (NFV) is recognized as a key, which decouples network service from proprietary hardware and runs workloads on common hardware and open source software. Featuring the modularized software-defined infrastructure, OpenStack is considered to be a matching foundation for NFV for its flexibility on cloud infrastructure and management. CSPs and telecommunications companies have been gradually adopting an agile and scalable OpenStack platform with compelling technical and business benefits.

Quanta Cloud Technology (QCT) are dedicated to delivering performance-optimized OpenStack solutions for different scenarios. The OpenStack-based NFV Infrastructure (NFVI) solution is provided to meet the carrier-grade requirements of datacenter and network innovations. Based on QCT's latest-released NFVI solution, this solution delivers updated system design with exceptional network performance to satisfy digital demands and accelerate cloud transformation.

In this reference architecture, QCT will address the recommended hardware configuration, solution software, and network planning of the QxStack NFV Infrastructure with Red Hat OpenStack Platform solution. Meanwhile, this paper also introduces the integration of network virtualization technologies, the resource allocation strategies, and performance evaluation with industry-leading Open Platform for NFV (OPNFV) test suite. Some add-on features including the deployment tool and service assurance framework will also be introduced for CSPs to efficiently utilize the designated NFVI platform.

2. Standard Architecture

This section will illustrate the recommended hardware and software options, and best practices as a standard architecture. To accelerate time to value for CSPs, a large number of advantages are introduced for planning a large-scale NFVI environment.


2.1. Hardware

NFVI is considered to be a turnkey solution to decouple network functions from proprietary appliances and bring a flexible, high-performance platform based on industry-standard servers. The QxStack NFV Infrastructure with Red Hat OpenStack Platform solution is built up with high expandable x86 hardware, supporting a fully virtualized infrastructure including servers, storage, and networking.

2.1.1. QuantaGrid D52B-1U: Director node

*QuantaGrid D52B-1U*¹ is an ultra-dense server with up to 5 PCIe expansion slots per chassis. It features flexible I/O options, including a variety of SAS Mezzanine and OCP NIC/ PHY Mezzanine cards in three different SKUs. This 1U server with Intel® Xeon® Processor Scalable Family is designed to enhance computing power, provide faster socket interconnect, and strengthen memory bandwidth.

Table 1. Solution Hardware - QuantaGrid D52B-1U.

		QuantaGrid D52B-1U Default Configuration
Processor	Intel® Xeon® Processor Scalable Family	
Storage	SFF tiered SKU: <ul style="list-style-type: none"> • 8x 2.5" hot-plug SATA/SAS drives • 4x 2.5" hot-plug NVMe/SATA/SAS drives 	
Memory	Total slots: 24 Capacity: Up to 3TB of memory for RDIMM/LRDIMM Memory type: 2666 MHz DDR4 RDIMM	
Network	LOM: Dedicated 1x GbE management port	
Expansion Slot	SFF tiered SKU, option 1: <ul style="list-style-type: none"> • 1x PCIe Gen3 x16 SAS mezzanine slot • 1x PCIe Gen3 x16 OCP 2.0 mezzanine slot or PHY card • 1x PCIe Gen3 x16 LP MD-2 • 1x PCIe Gen3 x8 LP MD-2 	


¹ QCT QuantaGrid D52B-1U:

<https://www.qct.io/product/index/Server/rackmount-server/1U-Rackmount-Server/QuantaGrid-D52B-1U>

2.1.2. QuantaGrid D52BQ-2U: controller and compute node

In the same product family with QuantaGrid D52B-1U, *QuantaGrid D52BQ-2U*² provides up to 10 PCIe expansion slots in a NUMA-balanced designed 2U chassis. It is a multiple configuration design, featuring flexible I/O options, a variety of SAS Mezzanine, and OCP NIC/ PHY Mezzanine options. This 2U server with Intel® Xeon® Processor Scalable Family can unleash computing power, accelerate socket interconnect, and strengthen memory bandwidth.

Table 2. Solution Hardware - QuantaGrid D52BQ-2U.

		QuantaGrid D52BQ-2U Default Configuration
Processor	Intel® Xeon® Processor Scalable Family	
Storage	SFF tiered SKU: <ul style="list-style-type: none"> • 8x 2.5" hot-plug SATA/SAS drives • 4x 2.5" hot-plug NVMe/SATA/SAS drives 	
Memory	Total slots: 24 Capacity: Up to 3TB of memory for RDIMM/LRDIMM Memory type: 2666 MHz DDR4 RDIMM	
Network	LOM: Dedicated 1x GbE management port	
Expansion Slot	SFF tiered SKU, option 1: <ul style="list-style-type: none"> • 1x PCIe Gen3 x16 OCP 2.0 mezzanine slot or PHY card • 2x PCIe Gen3 x 8 FHHL or 1x PCIe Gen3 x16 FHHL • 1x PCIe Gen3 x16 FHHL • 1x PCIe Gen3 x16 LP MD-2 (must be with SAS add-on card) 	

2.1.3. QuantaGrid D52T-1ULH: storage node

*QuantaGrid D52T-1ULH*³ features hybrid-tiered storage architecture in an ultra-dense 1U platform. It is tailored for hyperscale data centers and software-defined storage with up to 1TB memory capacity. Equipped with 12 hot-swappable 3.5" disk drives and 4 hot-swappable 2.5" NVMe SSDs, QuantaGrid D52T-1ULH is ideal for tier storage planning to accelerate IOPs and throughput without sacrificing large data storage capacity.


² QCT QuantaGrid D52B-2U:

<https://www.qct.io/product/index/Server/rackmount-server/2U-Rackmount-Server/QuantaGrid-D52BQ-2U#overview>

³ QCT QuantaGrid D52T-1ULH:

<https://www.qct.io/product/index/Storage/Storage-Server/1U-Storage-Server/QuantaGrid-D52T-1ULH#specifications>


Table 3. Solution Hardware - QuantaGrid D52T-1ULH.

		QuantaGrid D52T-1ULH Default Configuration
Processor	Intel® Xeon® Processor Scalable Family	
Storage	<ul style="list-style-type: none"> • 12x 3.5"/2.5" hot-plug 12Gb/s SAS or 6Gb/s SATA HDD/SSD • 4x 2.5" hot-plug 7mm 6Gb/s SATA SSD • 2x PCIe M.2 support with M.2 adapter for boot disk 	
Memory	Total slots: 12 Capacity: Up to 768GB RDIMM Memory type: 2666 MHz DDR4 RDIMM	
Expansion Slot	<ul style="list-style-type: none"> • 1x PCIe Gen3 x16 SAS mezz slot • 1x PCIe Gen3 x16 OCP mezz slot • 1x PCIe Gen3 x16 LP MD-2 	

2.1.4. QuantaMesh T1048-LY4R: management switch

QuantaMesh T1048-LY4R⁴ supports 48 10BASE-T/100BASE-TX/1000BASE-T and 4 1/10G SFP+ ports equipped with BMC in a compact rack unit size. T1048-LY4R plays a management role in a server rack with not only Ethernet switch functions but also IPMI 2.0 compliance. With Open Network Installation Environment (ONIE) pre-loaded, it can be used with any OS which supports the ONIE installer to achieve agile installation and fast response for today's IT demands.

Table 4. Solution Hardware - QuantaMesh T1048-LY4R.

		QuantaMesh T1048-LY4R Default Configuration
Physical ports	Port configuration: <ul style="list-style-type: none"> • 48x 10/100/1000BAE-T • 4x 1/10GbE SFP+ ports 	
Performance	Switching capacity: 176Gbps Maximum forwarding rate: 131Mpps MAC: 16K	
High Availability	Features: <ul style="list-style-type: none"> • Redundant power supply: 1+1 • Hot-swappable fan tray: N+2 • Multi-chassis Link Aggregation (MLAG) 	


⁴ QCT QuantaMesh T1048-LY4R:

<https://www.qct.io/product/index/Networking/Bare-Metal-Switch/Management-Switch/QuantaMesh-BMS-T1048-LY4R#specifications>

2.1.5. QuantaMesh BMS T4048-IX8D: Top-of-Rack (ToR) switch

*QuantaMesh T4048-IX8D*⁵ supports 48 SFP28 and 8 QSFP28 (10/25/40/50/100GbE speed) ports and is equipped with BMC in a compact rack unit size. By leveraging merchant silicon chip, T4048-IX8D is fully compliant with IEEE 802.3by and is a high-density Ethernet switch with advanced features such as smart table, dynamic load balancing, and VxLAN/RIOT support. With ONIE pre-loaded, T4048-IX8D is compatible with any network operating system which supports ONIE installer to achieve agile installation and fast response.

Table 5. Solution Hardware - QuantaMesh T4048-IX8D.

		QuantaMesh T4048-IX8D Default Configuration
Physical ports	Port configuration: <ul style="list-style-type: none"> • 48x SFP28 (10/25GbE) • 8x QSFP28 ports (10/40 or 25/50/100GbE) 	
Performance	Switching capacity: 4Tbps Maximum forwarding rate: 2Bpps MAC: Unified Forwarding Table to dynamically allocate the L2/L3 tables	
High Availability	Features: <ul style="list-style-type: none"> • Redundant power supply: 1+1 • Hot-swappable fan tray: N+2 • Multi-chassis Link Aggregation (MLAG) 	

⁵ QCT QuantaMesh T4048-IX8D:
<https://www.qct.io/product/index/Networking/Bare-Metal-Switch/Leaf-Switch/QuantaMesh-BMS-T4048-IX8D#specifications>

2.2. Solution Software

2.2.1. Red Hat OpenStack Platform 13

Red Hat is one of the leading provider of enterprise open source solutions. It has been in a leading position in the OpenStack and Linux communities. Red Hat OpenStack Platform delivers core Infrastructure-as-a-Service (IaaS) and provides a powerful foundation for next-generation datacenters. Based on the upstream Queens release, Red Hat OpenStack Platform 13 not only supports NFV deployments with the automated OVS-DPDK and SR-IOV configurations but also enables Real-time KVM to align with the needs of NFV workloads. Red Hat OpenStack Platform 13 fully supports containerized OpenStack service deployment to bring greater flexibility and agility into IT operations. Moreover, Red Hat OpenStack Platform 13 can be tuned to deliver a private or public cloud platform including:

- Web browser-based interface accessible to users and administrators.
- Authentication and authorization mechanisms.
- Integrated networking.
- Virtual machine provisioning engine and image storage.
- Fully distributed object storage.
- Persistent block-level storage.

With these advantages, QCT adopts Red Hat OpenStack Platform to be the foundation of Virtual Infrastructure Manager (VIM).

2.2.2. Network Topology

Red Hat OpenStack Platform uses Neutron network service to manage the software-based networks, static and floating IP addresses, and the DHCP service. A well-designed network topology is indispensable to ensure the efficiency, correctness, and availability of the communication among OpenStack services. In order to optimize the network performance and fulfill CSPs' strict requirements, the network topology of the QxStack NFV Infrastructure with Red Hat OpenStack Platform solution is designed according to the Red Hat OpenStack Platform 13 Network Functions Virtualization Planning and Configuration Guide⁶.

The Red Hat OpenStack Platform director provides multiple node types, including controller, SR-IOV compute, OVS-DPDK compute, and Red Hat Ceph Storage to build the overcloud environment as NFV infrastructure. Different services are assigned to each node for easy management and optimal system resource utilization. Likewise, network types are attached to the overcloud nodes based on the assigned services. The example of the logical network topology consists of the OpenStack standard networks and two data plane provider networks, as shown in Figure 1.

⁶ Network Functions Virtualization Planning Guide: https://access.redhat.com/documentation/en-us/red_hat_openstack_platform/13/html/network_functions_virtualization_planning_and_configuration_guide/

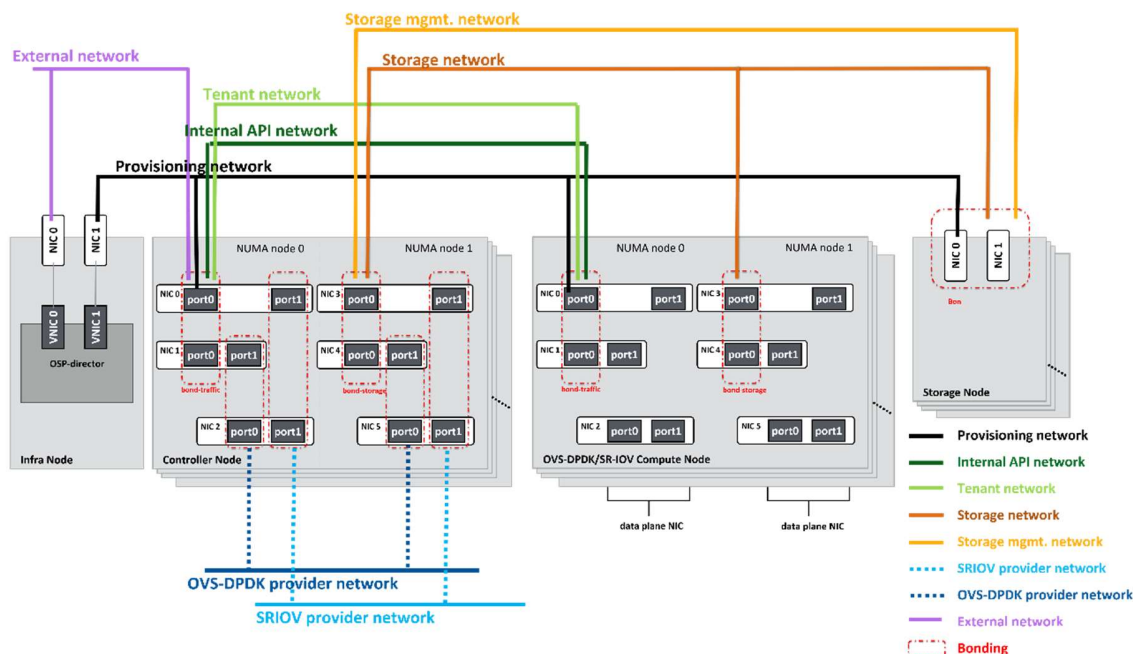


Figure 1. Example of Network Layout.

The standard network segments of Red Hat OpenStack Platform 13 environment shown in the upper half of the figure include provisioning network, internal API network, tenant network, storage network, storage management network, and external network. The director deploys and monitors the overcloud via the provisioning network. The OpenStack services in the overcloud communicate with each other with the Internal API network. The OpenStack Neutron uses the tenant network to isolate each tenant with their own network environment. The overcloud nodes access the storage nodes by the storage network and the overcloud controller nodes use the storage management network to monitor the status of storage cluster. The external network not only hosts the OpenStack Dashboard (Horizon) for graphical system management but also handles the public API for OpenStack services.

Unlike OpenStack standard networks, the two data plane provider networks, OVS-DPDK provider network and SR-IOV provider network, in the lower half of the figure only host the data plane traffics. With the data plane provider networks, the QxStack NFV Infrastructure with Red Hat OpenStack Platform solution is able to provide high-efficient packet processing to fulfill the requirements of network-sensitive workloads in NFV.

2.2.3. EPA Features

Resource Allocation Strategy

To optimize resource allocation, the QxStack NFV Infrastructure with Red Hat OpenStack Platform solution provides CPU pinning and features a NUMA-balanced design. By default, multiple virtual CPUs (vCPU) share a pool of physical CPUs (pCPUs) resource and allows over-commitment of CPU resource on a host in virtualized infrastructures. CPU pinning enables one-to-one mapping between vCPUs and pCPUs. It is beneficial for some latency-sensitive workloads to have dedicated CPU resources, which can prevent performance degradation from resource contention. As shown in Figure 2, the CPU pinning dedicates specific compute resources to specific VMs and increases cache efficiency.

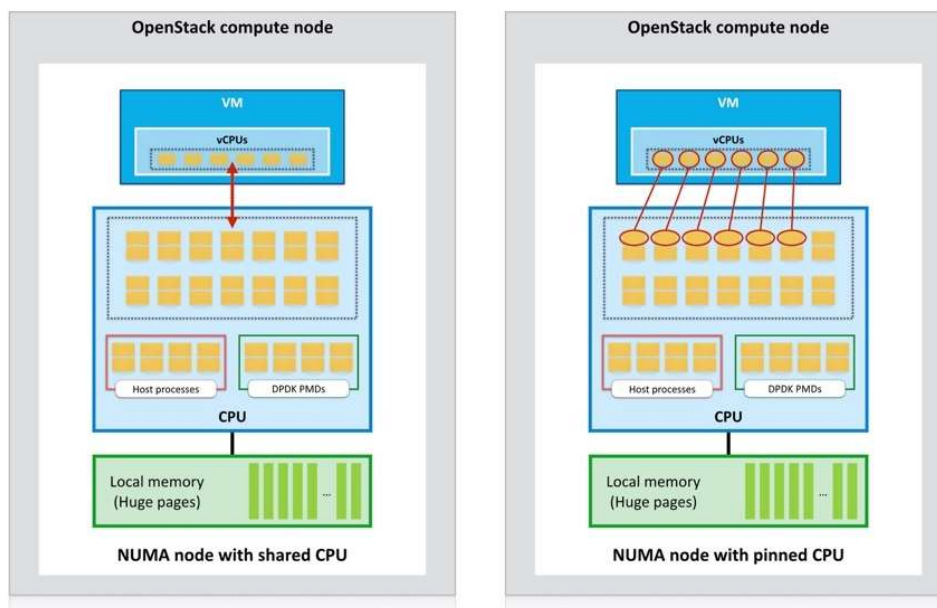


Figure 2. CPU Topology and Core Allocation.

Traditional uniform memory access (UMA) architecture models share memory resources evenly across all CPUs and sockets in a multiprocessor system. This often results in long memory access time, regardless of the location of the memory in relation to the CPU or socket. NUMA architecture models geographically-distributed system memory by considering its location in relation to each CPU, which speed the access to the memory closer to the CPU. Processes can then access local CPU memory—rather than another CPU’s local memory or shared memory—to improve computational performance. In Red Hat OpenStack Platform, OpenStack Compute (Nova) intelligently schedules and assigns memory resources when the instances are launched. Administrators can create instance configurations customized for specific performance levels to target specialized workloads. QCT proposed an architecture with a NUMA-balanced design which supports local memory access and distributes NICs across CPUs and sockets, as shown in Figure 3.

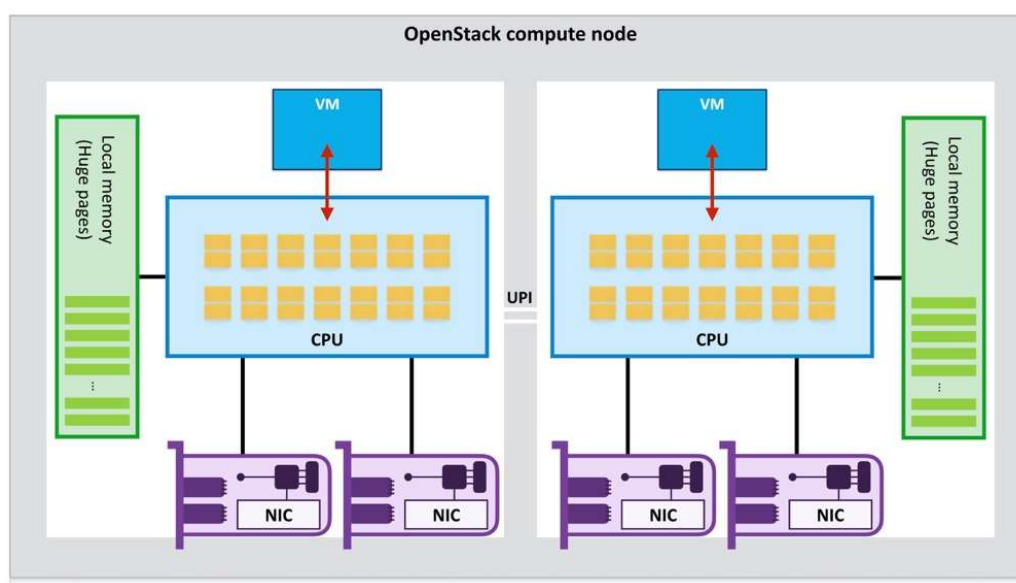


Figure 3. NUMA-Balanced Hardware Design.

Data Plane Acceleration

In generic network stack, the network traffic from instances flows through a complicated network virtualization stack, including TAP device, Linux bridge, veth pair, and Open vSwitch to access the Internet. To accelerate the packet processing, the QxStack NFV Infrastructure with Red Hat OpenStack Platform solution not only supports memory huge pages in the compute nodes but also provides both OVS-DPDK and SR-IOV data plane acceleration. The comparison of network virtualization stacks among generic network stack, OVS-DPDK network stack, and SR-IOV network stack is shown in Figure 4.

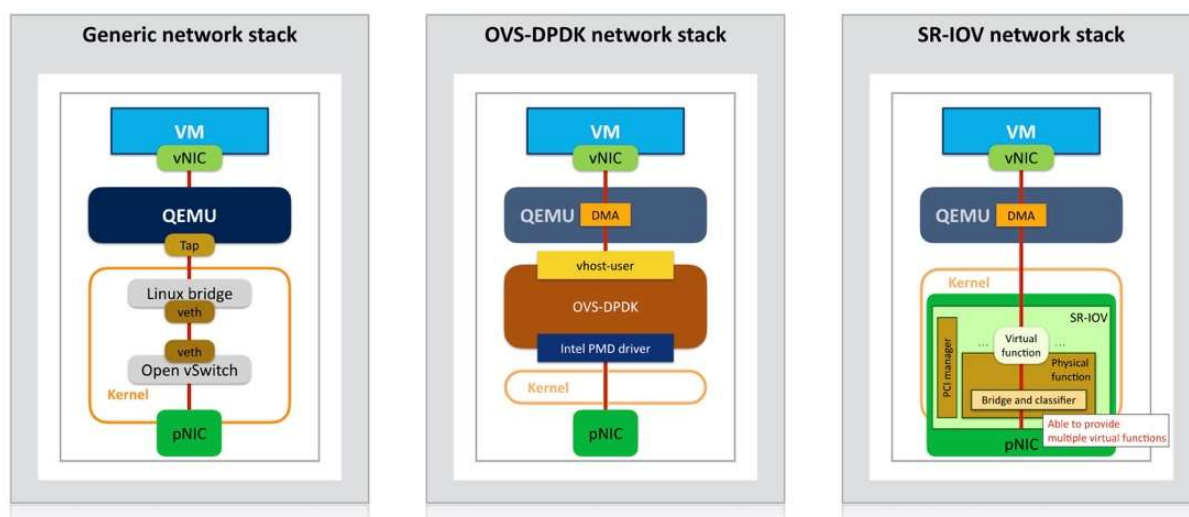


Figure 4. Comparison among Network Virtualization Stacks.

Intel Data Plane Development Kit (DPDK) consists of a set of data plane libraries and user-space network drivers for accelerating packet processing. It provides a programmable framework that enables applications to operate the NIC directly. Open vSwitch with DPDK (OVS-DPDK) replaces the standard Open vSwitch (OVS) kernel-space datapath with DPDK user-space datapath. It performs user-space packet processing on top of the dedicated CPU cores isolated from kernel processes. With OVS-DPDK adopted, the QxStack NFV Infrastructure with Red Hat OpenStack Platform provides enhanced network performance. In the meanwhile, it remains the transparency in overall architecture and exposes the features—including OpenFlow, Open vSwitch Database (OVSDb), and command lines—as the standard OVS.

Single Root I/O Virtualization (SR-IOV) is a hardware-based technology that improves both performance and scalability, and its specification is defined by Peripheral Component Interconnect Special Interest Group⁷ (PCI-SIG). An SR-IOV-enabled PCIe device can appear as multiple physical functions (PFs) and virtual functions (VFs). The PFs with full-featured PCIe functions is capable of configuring and managing the SR-IOV functionality by assigning VFs.

The VFs with simple PCIe functions only process I/O operations and appear as PCI devices. Both PFs and VFs are able to provide near-native performance and can be attached to a guest through hypervisor. SR-IOV provides an independent memory space and Direct Memory Access (DMA) streams for data transfer and management. With

⁷ PCI-SIG: <https://pcsig.com/>

such I/O virtualization technology, SR-IOV-enabled PCIe device is able to efficiently share a single device to multiple guests without hypervisor involvement.

2.2.4. High Availability

Based on Red Hat OpenStack Platform 13, QCT designed a hardware-software integrated solution with high availability. The QxStack NFV Infrastructure with Red Hat OpenStack Platform solution incorporates high availability architecture to avoid single point of failure in the infrastructure, and to provide stable and reliable services. The solution provides the network with high availability through cross-NIC bonding of the network interfaces and the adoption of Multi-Chassis Link Aggregation Group (MC-LAG) between Top-of-Rack (ToR) data switches. Besides, the solution also provides the high availability at Operating System (OS) level by creating Redundant Array of Independent Disks (RAID) on disks.

3. Selected Features

In addition to the standard features, the QxStack NFV Infrastructure with Red Hat OpenStack Platform solution provides several add-on functions to create more advantages. Quick-Assist Technology (QAT) and real-time KVM can be respectively adopted to provide better security and latency performance for NFV workloads. The QCT operational tools can be adopted to efficiently deploy and monitor the environment, and to reduce time and cost for CSPs.

3.1. Quick-Assist Technology

The security function is inevitable in the network architecture to avoid data leakage and data theft attack in cloud. The optimized and streamlined security functions need to be designed to the integration of hardware and software. Intel® QuickAssist Technology (QAT) is a core platform technology using a hardware-assisted security engine for implementing major security functions such as IPsec and SSL/TLS. This security engine works to reserve processor cycles for application processing which not only relieves CPU burden but also improves overall system performance. QAT is also a hardware accelerator which can accelerate and compress cryptographic workloads by offloading the data to the server. The architecture of the QxStack NFV Infrastructure with Red Hat OpenStack Platform solution leverages Intel QAT to enhance the security application performance.

3.2. Real-Time KVM

The latency can be normally reduced by the standard tuning of workloads; however, it still fails to completely achieve the real-time processing. Real-time KVM is a technology for virtual machine to achieve real-time performance, similar to the performance of real-time OS in physical server. By setting up the real-time KVM, the priority of the tasks in the virtual machine is visible to the host which can prevent the prioritized real-time task from interfering, as shown in Figure 5. The QxStack NFV Infrastructure with Red Hat OpenStack Platform solution can fully support real-time KVM. The compute nodes with data plane accelerations, either OVS-DPDK or SR-IOV, can integrate with the real-time OS to take the advantages of data plane acceleration and real-time technology, providing crucial-determined latency and lower jitter for real-time processing.

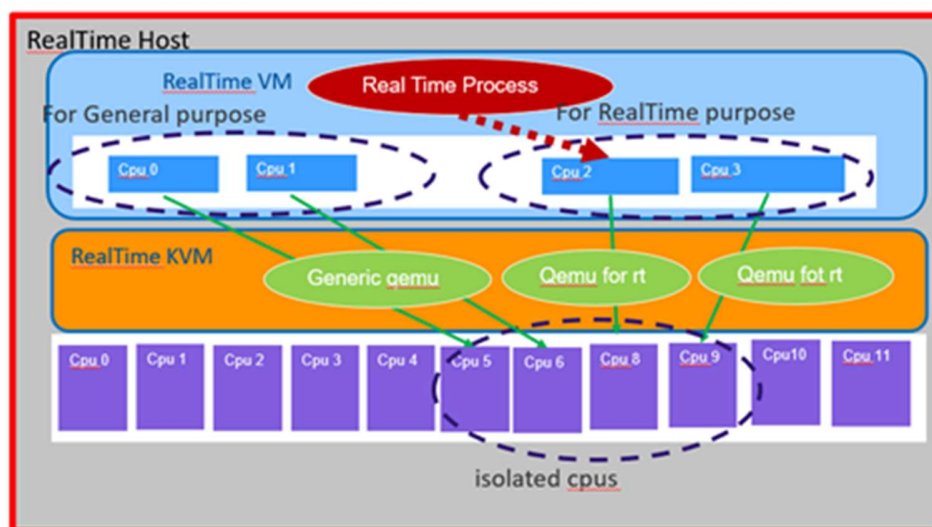


Figure 5. RT-KVM Architecture.

3.3. QCT Operational Tools

3.3.1. QxStack Auto-Deployment Tool

Deploying NFVI can be complicated and challenging. NFVI is composed by hundreds of services. Cloud managers need to configure each service for hardware specification, design network topology, and tune overall performance. Additionally, cloud managers also need to adjust BIOS settings on each server, configure switches on each port based on network topology, and collect hardware data for deploying NFVI. Manual intervention can possibly cause NFVI in an unstable status or even result in system malfunction. Understanding the complexity of the deployment process, QCT developed an automatic tool, QxStack Auto-Deployment Tool, to provide the following benefits for cloud managers.

Simplified the planning of NFV Infrastructure

To deploy NFVI, both hardware resource allocation and software settings need to be planned and further to be configured. However, these configurations are in different format and structure. Cloud managers need to be careful and experienced in handling these differentiations. To effectively configure the QxStack NFV Infrastructure with Red Hat OpenStack Platform, QxStack Auto-Deployment Tool can be adopted to plan the settings of NFVI through user-friendly interfaces, as shown in Figure 6. Cloud manager can follow the steps and complete the survey form to generate a configuration file. With this file, QxStack Auto-Deployment Tool will automatically configure all the hardware and software settings of the QxStack NFV Infrastructure with Red Hat OpenStack Platform solution.

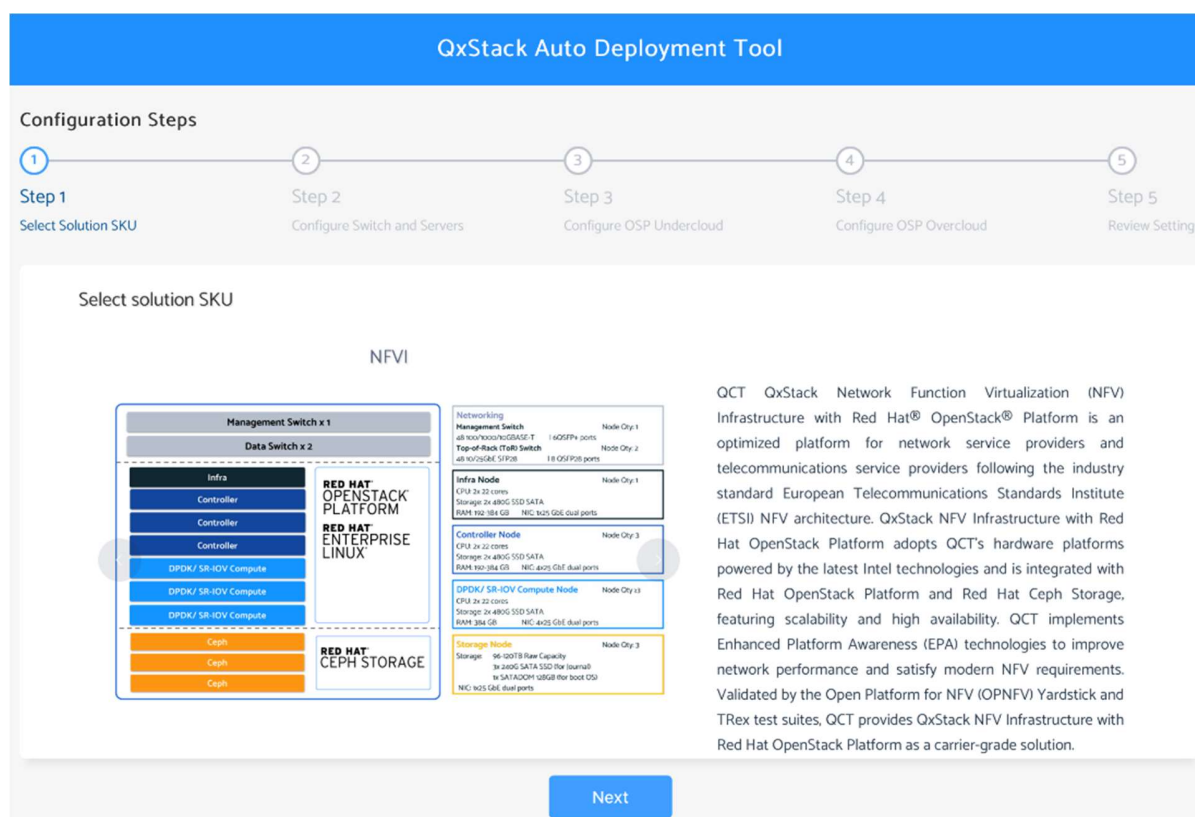


Figure 6. QxStack Auto-Deployment Tool – Surveyform.

Automated hardware provisioning

To deploy NFVI, provisioning hardware resources for OpenStack platform is a huge and heavy task. For example, to generate the server list to be imported to the OpenStack platform, cloud managers need to power on all the compute servers, query the NIC mac addresses, and update the server list one by one. After collecting the hardware resources, the operating system (OS) should be installed on all the compute servers. Generally, most deployment tools provide their features to provision OS on compute servers. However, network switch needs to be manually configured to ensure the network is connected; otherwise, the OS provisioning tasks will be failed.

QxStack Auto-Deployment Tool provides a series of features to automatically control and provision hardware resources. Cloud managers only need to complete the basic settings (e.g., the network IP range, external network VLAN id) and the QxStack Auto-Deployment Tool will automatically take the heavy tasks such as applying the settings to compute servers and network switches, configuring BIOS, and assigning IP for BMC to complete the deployment. Besides, QxStack Auto-Deployment Tool can be also adopted to simultaneously monitor the deployment progress and status from the web UI, as shown in Figure 7.

QxStack Auto-Deployment Tool can effectively provide a streamlined and simplified process including architecture planning, hardware provisioning, and solution deployment for the QxStack NFV Infrastructure with Red Hat OpenStack Platform solution.

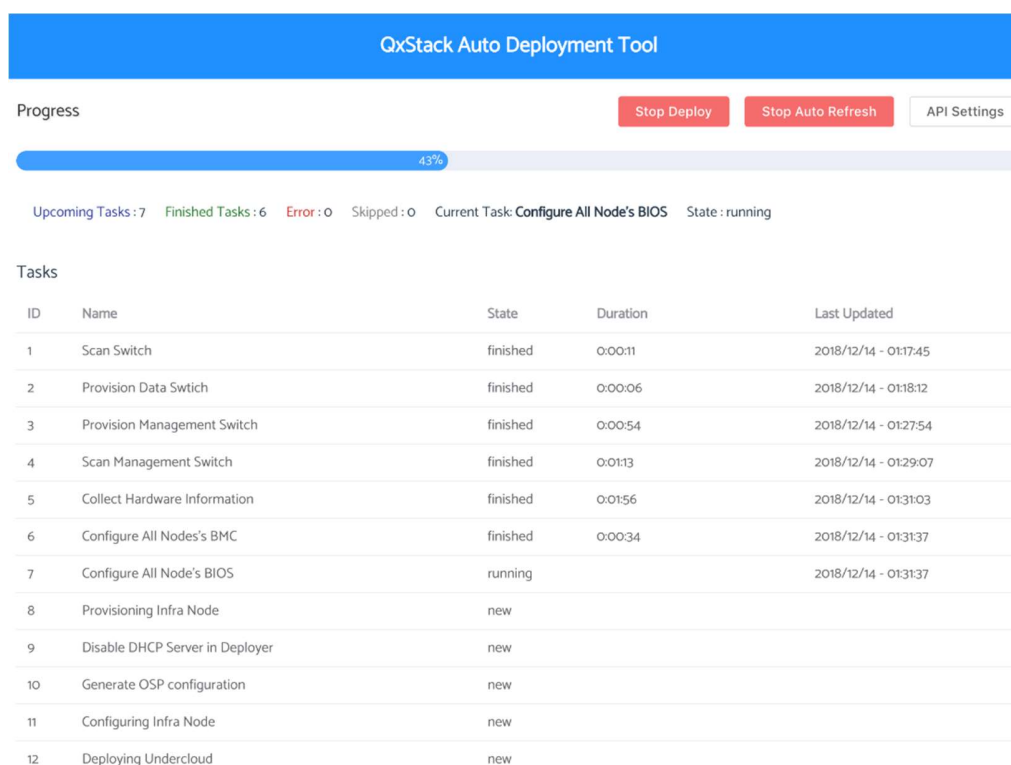


Figure 7. QxStack Auto-Deployment Tool – Workflow.

3.3.2. TestBox

In order to validate the QxStack NFV Infrastructure with Red Hat OpenStack Platform solution, QCT leverages the test ecosystem in OPNFV which is an open source community under the Linux Foundation. OPNFV aims at

By leveraging OPNFV test projects, TestBox is able to validate the two network performance scenarios, VM-to-VM test scenario and Phy-VM-Phy test scenario. In the VM-to-VM test scenario, TestBox leverages Yardstick to launch two VM instances on the OpenStack platform in which one instance acts as the traffic generator (e.g., Pktgen, Ping, iPerf3) and the other one is the target VM to measure the east-west traffics. In the Phy-VM-Phy test scenario, TestBox leverages NFVbench to launch only the target VM on the OpenStack platform and Trex is triggered as a traffic generator on a standalone server to measure the north-south traffics. To further simplify the complicated testing configurations, TestBox facilitates a single Test Description File (TDF) to validate the projects on the same NFVI solution. The TDF is a YAML file for TestBox to distinguish general configurations (e.g., openrc information, external network, availability zone) and customized settings (e.g., provider network, VLAN ID, NIC information) for each test project. According to TDF configurations, TestBox can simply set up the test environment, run the selected test suite, and collect the testing results.

The diagram illustrates the QxStack NFVI Solution architecture, organized into several layers:

- Test in a Box:** The top layer, containing a **Testbox** which is divided into **Setup**, **Test**, and **Result** phases. A **TDF** (Traffic Description File) is associated with the Setup phase.
- Test Suite:** The middle layer, which includes **Functest**, **Yardstick**, and **NFVbench**. **NFVbench** is connected to a **Traffic Gen** component.
- VM-to-VM Traffic:** A section showing **TrafficGen VM** and **Target VM** connected by a bidirectional arrow, representing **VM-to-VM Traffic**.
- Phy-VM-Phy Traffic:** A section showing a **Target VM** connected to a **Traffic Gen** component, representing **Phy-VM-Phy Traffic**.
- Enhanced Platform Awareness:** A layer below the traffic sections, containing **Red Hat® OpenStack® Platform** and **Red Hat® Ceph Storage**.
- Red Hat® Enterprise Linux®:** The operating system layer.
- QxStack NFVI Solution:** The overall solution, encompassing all the layers above.

Figure 8. TestBox Architecture.

Functest

Functest is one of the OPNFV test projects, targeting on functional tests of the NFVI solution. The comprehensive test covers healthcheck, smoke, feature, component, and VNF tiers. To validate basic functions, QCT leverages connection_check and api_check test cases in the healthcheck tier to develop functional tests in TestBox.

Yardstick

Yardstick is another OPNFV test project, targeting on implementing a variety of test cases aligned with the European Telecommunications Standards Institute (ETSI) TST 001 specification⁸ to characterize the underlying NFV infrastructure. Each test case is configured by a YAML file and deployed on the underlying OpenStack by the Heat template. Among comprehensive test cases in Yardstick, QCT selects the performance metrics related to NFVI network to measure the east-west traffic in VM-to-VM test scenario.

NFVbench

NFVbench is also an OPNFV test project, targeting to provide a measurement toolkit for the data plane benchmark of a fully OpenStack-based NFVI production. Both Phy-VM-Phy and Phy-VM-VM-Phy test scenarios are pre-built in the NFVbench. The test process is automatically launched and configured via standard OpenStack APIs. To verify the QxStack NFV Infrastructure with Red Hat OpenStack Platform solution, QCT selects the Phy-VM-Phy scenario with a L2 forwarder to measure the network performance of north-south traffic.

TestBox can be utilized to validate both OVS-DPDK and SR-IOV provider networks in either VM-to-VM or Phy-VM-Phy scenario. The performance results in terms of network latency and throughput are displayed on the Grafana dashboard.

3.3.3. QCT Service Assurance Framework

Based on Advanced Message Queuing Protocol (AMQP), QCT proposed a Service Assurance Framework (hereafter called Telemetry) which is a scalable telemetry architecture to deal with large amount of performance data (metrics) and implement service assurance for administrators to monitor resources. The metrics collected by collectd are sent to Qpid Dispatch Router (QDR) based on AMQP 1.0 and consumed by Smart Agent to be stored in Prometheus Time-Series Database (TSDB). Grafana dashboard subsequently queries the Prometheus TSDB to show the visualized data on web UI, as shown in Figure 9.

⁸ ETSI GS NFV-TST 001: https://www.etsi.org/deliver/etsi_gs/NFV-TST/001_099/001/01.01.01_60/gs_nfv-tst001v010101p.pdf

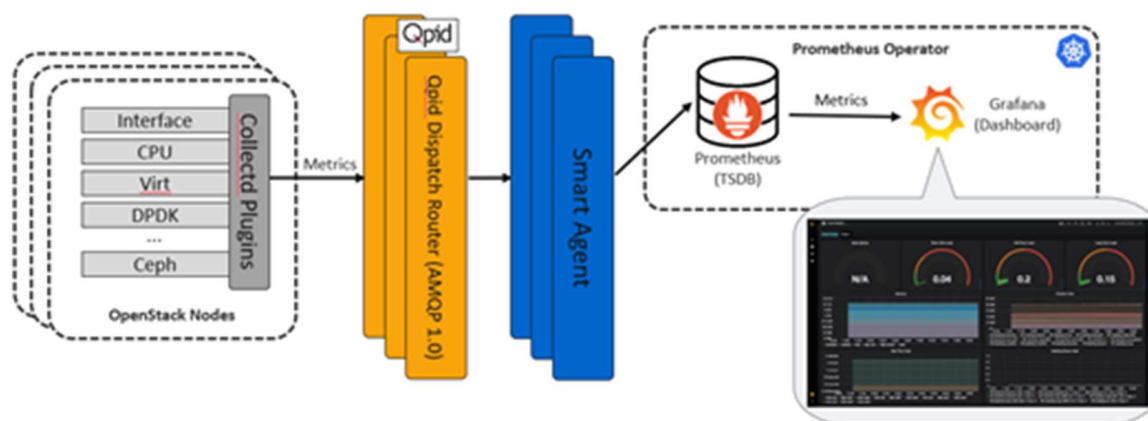


Figure 9. QCT Service Assurance Framework.

Collectd

Collectd periodically collects system information (metrics) from target nodes as a daemon. In QxStack NfV Infrastructure with Red Hat OpenStack Platform solution, collectd uses several plugins to monitor CPU, disk, memory, network resources in all nodes. Additionally, virt plugin and ceph plugins are also respectively applied to compute nodes and ceph nodes. Collectd sends these metrics as JSON-format data to QDR based on AMQP 1.0.

Qpid Dispatch Router (QDR)

QDR is a scalable message router based on AMQP 1.0 interconnected between AMQP endpoints such as clients, brokers, or other AMQP-enabled services. QDR provides the redundancy to ensures the reliability and scalability of route transmission.

Smart Agent

Smart Agent is a metrics consumer and HTTP server. It consumes metrics from QDR, converts those data into Prometheus TSDB formats, and then stores the data in cache. In the meantime, Smart Agent also provides a HTTP server for Prometheus Server to scrape the metrics.

Prometheus

Prometheus is a monitoring system which scrapes metrics via a pull model over HTTP and stores data in Time-Series Database (TSDB). To easily deploy and manage containerized Prometheus in kubernetes cluster, Prometheus operator here is applied to automatically create and destroy the Prometheus instance, generate configurations obtained from ServiceMonitor, and watch the real-time resource status related to Prometheus.

Grafana

Grafana is a dashboard for metrics visualization. It can query metrics from a variety of databases, including Prometheus TSDB. With Grafana, administrators can easily monitor system resources, find faults, and get alerts. Figure 10 shows the Grafana panels with QCT Telemetry metrics.



Figure 10. Grafana Panels with Telemetry Metrics.

The scalable AMQP Telemetry architecture can be adopted for NFVI administrators to watch the real-time infrastructure status, aware disruptions, and simultaneously handle these events for shortening service downtime.

4. Performance Evaluation

This section shows the detailed results of traffic performance evaluation based on different testing topologies with different tools. The operational tool, TestBox, is mainly used to verify the benefits of EPA features. The selected features, including QAT and RT-KVM are adopted to effectively enhance the performances.

4.1. TestBox Performance Results

Before validating the network performance, it is prerequisite to check the health of the QxStack NFV Infrastructure with Red Hat OpenStack solution. QCT leverages OPNFV Functest in TestBox to verify the two test cases, basic connectivity and API access to core OpenStack services (e.g., Keystone, Glance, Neutron, Nova) and the external network, as shown in Figure 11.

TEST CASE	PROJECT	TIER	DURATION	RESULT
connection_check	functest	healthcheck	00:08	PASS
api_check	functest	healthcheck	07:13	PASS

Figure 11. Functional Healthcheck Results.

The test topology of dataplane performance evaluation includes east-west (VM-to-VM) traffic and north-south (Phy-VM-Phy) traffic, as shown in Figure 12. The tests for east-west traffic evaluate the dataplane performance of underlying infrastructure while the tests for north-south traffic evaluate the network performance between the infrastructure and the outer network.

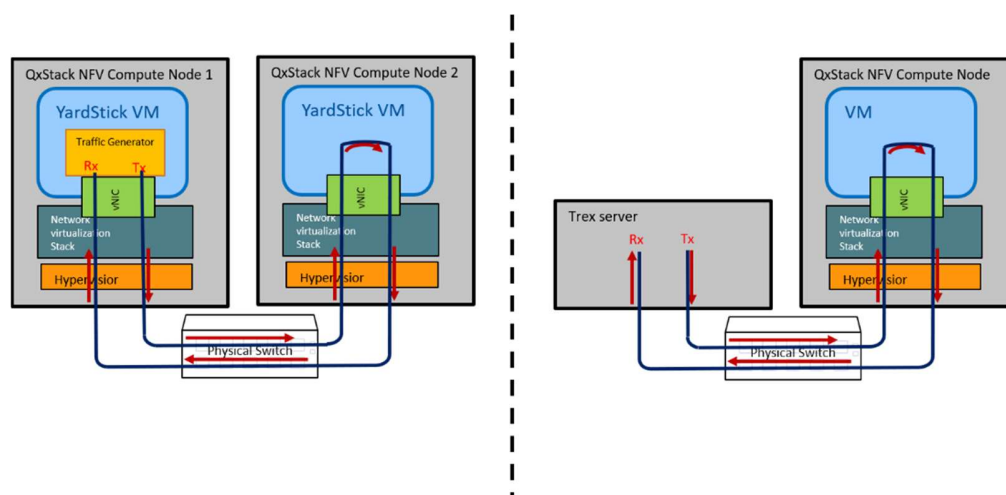


Figure 12. East-West and North-South Testing Topologies.

Network performance of OVS-DPDK with and without EPA features

Figure 13 shows the performance results of VM-to-VM tests based on OVS-DPDK network. In this case, TestBox leverages OPNFV Yardstick to evaluate the network performance in terms of throughput and latency between enabling and non-enabling EPA features. The “w/o EPA” scenario is tested under the topology without NUMA-balanced hardware design and the VM is launched with default CPU-shared policy. On the other hand, the “with

EPA" scenario is tested under the topology with NUMA-balanced hardware design and the VM is launched with CPU pinning policy.

Compared with the network performance without the EPA feature enabled, the throughput and latency performances with EPA enabled are generally improved. For example, the throughput with EPA enabled in 1024 packet size is improved up to 6.07% and the latency with EPA enabled in 1024 packet size is improved up to 22.67%.

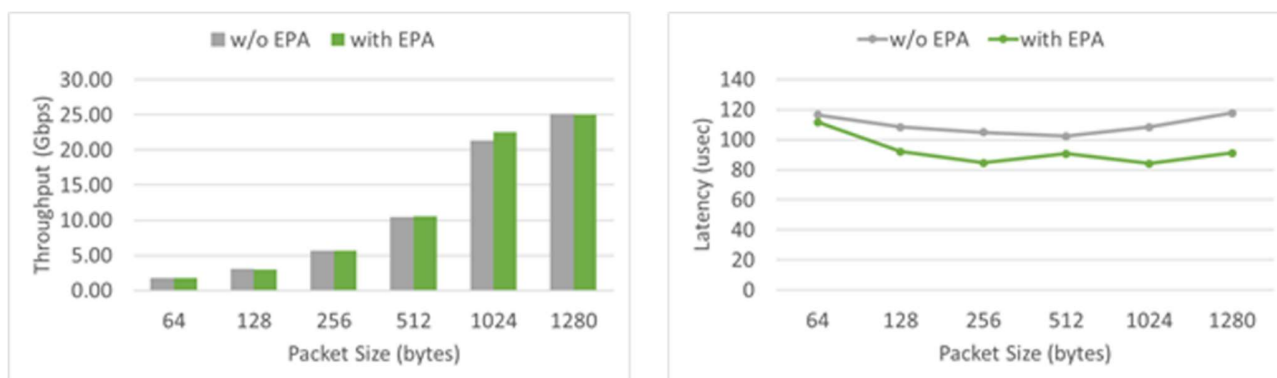


Figure 13. Comparison between with/without EPA Features based on OVS-DPDK using Yardstick.

Figure 14 shows the performance results of Phy-VM-Phy tests based on OVS-DPDK network. In this case, TestBox leverages OPNFV NFVbench to evaluate the network performance in terms of throughput and latency between enabling and non-enabling EPA features. The test results show that applying EPA features can generally improve the performance. For example, the throughput is improved up to 65.56% in 512 packet size and the latency is improved up to 38.73% in 512 packet size.

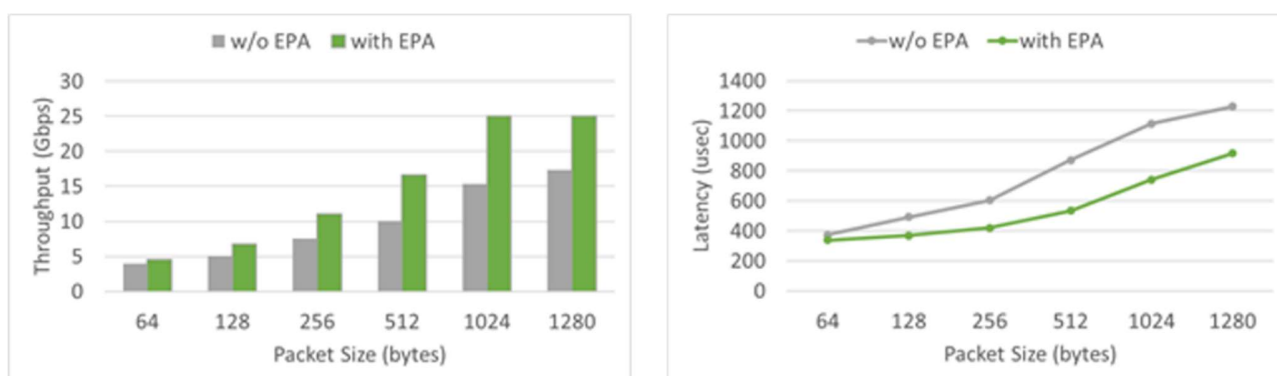


Figure 14. Comparison between with/without EPA Features based on OVS-DPDK using NFVBench.

Network performance of SR-IOV with and without EPA features

Figure 15 shows the performance results of VM-to-VM tests based on SR-IOV network. By adopting the inherent design of SR-IOV acceleration technique, a VM is resided in the CPU socket in which the VFs are assigned to the VM for the direct access to hardware devices. Hence, both the “w/o EPA” and the “with EPA” scenarios are tested under the topology with NUMA awareness. The former scenario launches VMs with default CPU-shared policy while the latter one launches VMs with CPU pinning policy. The throughput performance with EPA feature enabled is generally improved. For example, the throughput is improved by 2.77% in 64 packet size and the latency is improved by 28.5% in 1024 packet size when EPA features are enabled.

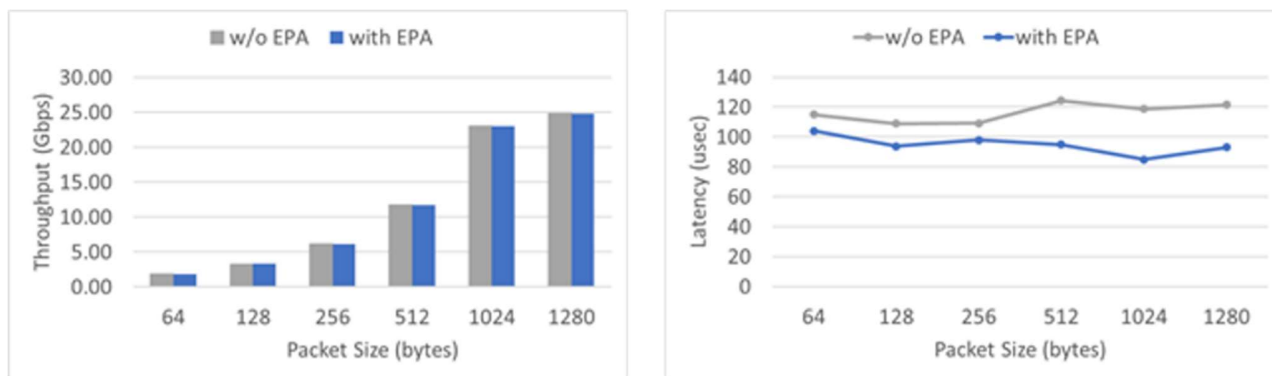


Figure 15. Comparison between with/without EPA Features based on SR-IOV using Yardstick.

Figure 16 shows the performance results of Phy-VM-Phy tests based on SR-IOV network. As aforementioned, both the “w/o EPA” and the “with EPA” scenarios over SR-IOV network are tested under the topology with NUMA-balanced hardware design. The results show that the tests over SR-IOV network is able to achieve near line-rate performance when the packet size is larger and equal to 256 bytes. The overall latency shows no significant difference between the “w/o EPA” and the “with EPA” scenarios.

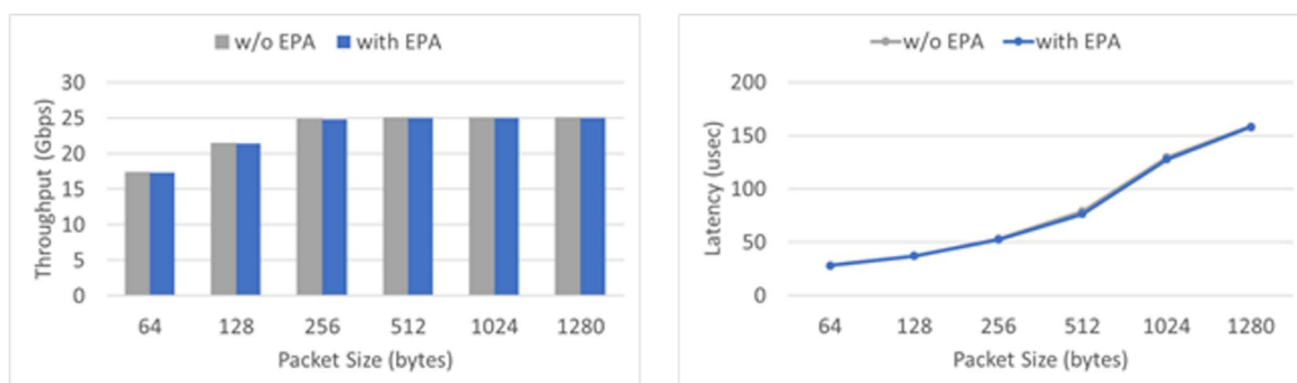


Figure 16. Comparison between with/without EPA Features based on SR-IOV using NFVBench.

4.2. Additional Enhancement

QAT

The QxStack NFV Infrastructure with Red Hat OpenStack Platform solution adopts SR-IOV technology to create multiple QAT VFs for the instances. The testing methodology of QAT performance adopts TRex as the packet generator and DPDK-IPSec application to verify the influence of QAT VFs, as shown in Figure 17.

DPDK-IPSec is an application developed based on DPDK. The huge pages must be enabled in the instances to execute this application. The DPDK-IPSec creates a tunnel between two instances for transmitting the encrypted packets. The first instance uses the DPDK-IPSec to encrypt and transmit the data via the tunnel. The second one also uses the DPDK-IPSec to receive and decode these packets.

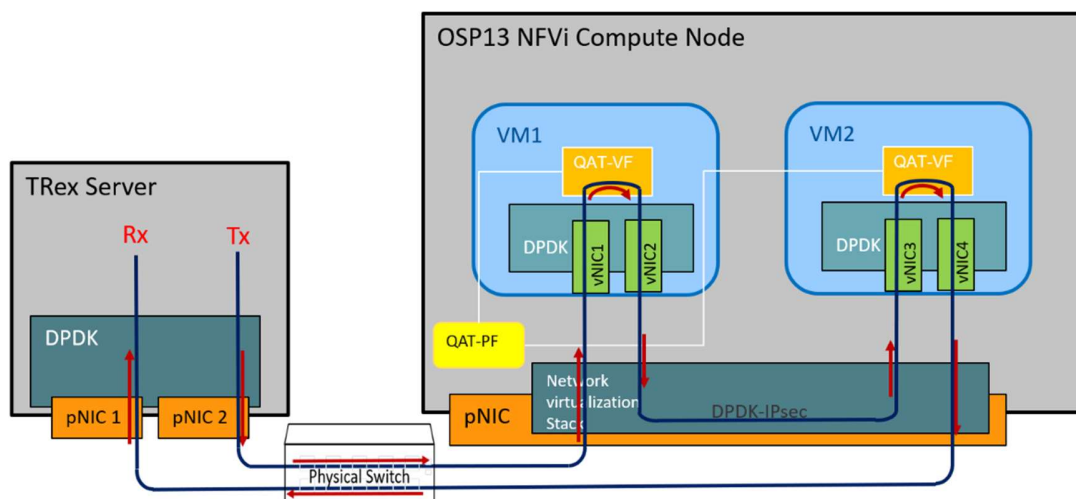


Figure 17. QAT Validation Topology.

Take SR-IOV compute nodes with EPA features enabled as an example, the performance of the application with QAT VFs is improved by 30% when the packet size is larger and equal to 512 bytes, as shown in Figure 18.

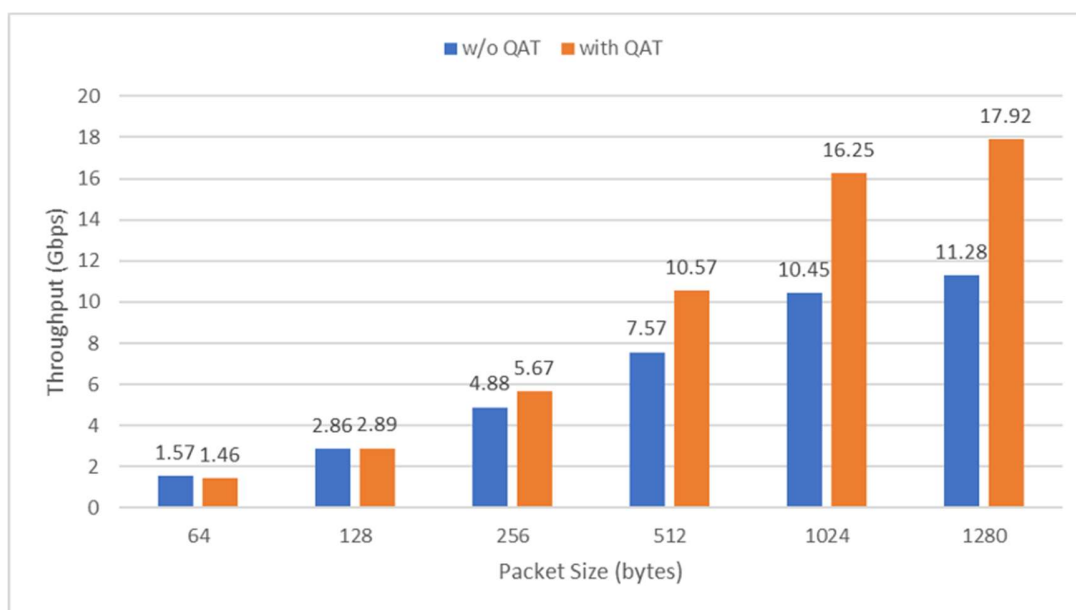


Figure 18. Comparison between with/without QAT on SR-IOV Compute Nodes.

Real-Time KVM

Cyclicttest is a tool used to measure the interrupt latency in the environment which requires the determined response time. By adopting Cyclicttest, the interrupt latency with real-time KVM is lower than that of with generic KVM; moreover, the interrupt latency with real-time KVM is reduced by 22% in average and 33% in maximum, as shown in Figure 19.

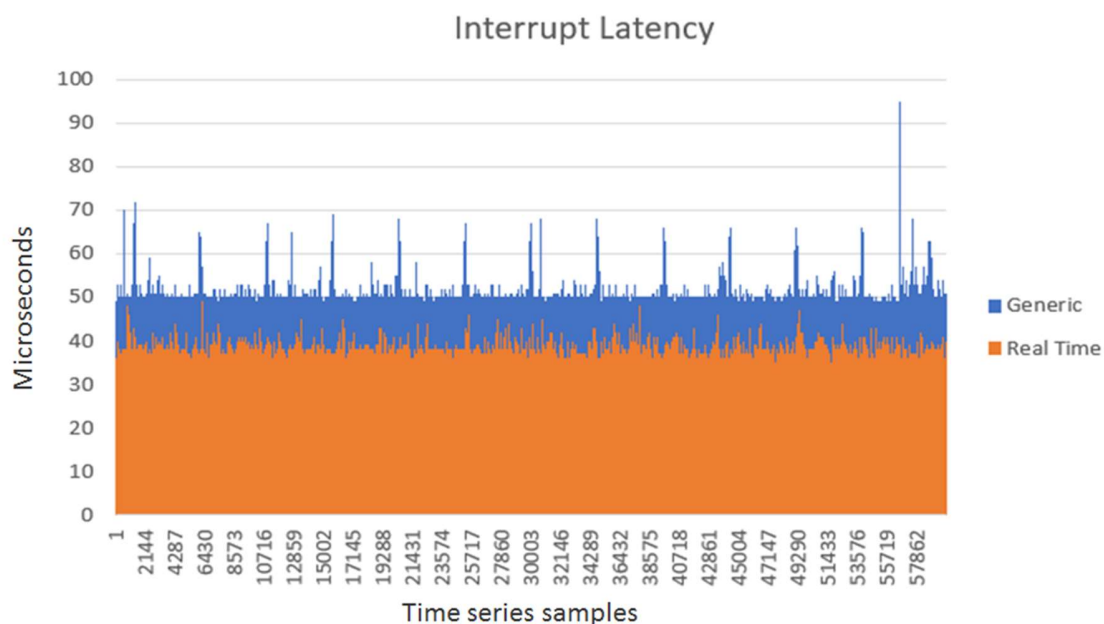


Figure 19. Comparison of Interrupt Latency between with/without real-time KVM.

Computing Jitter refers to the variance in processing time when underline hardware or OS interferences occur. Performance Measurement and Analysis (PMA) is a benchmark tool used to evaluate the computing jitter. By adopting PMA, the variance in computing jitter with real-time KVM is lower than that of with generic KVM; moreover, the computing jitter with real-time KVM is reduced by 15% in average and 75% in maximum, as show in Figure 20.

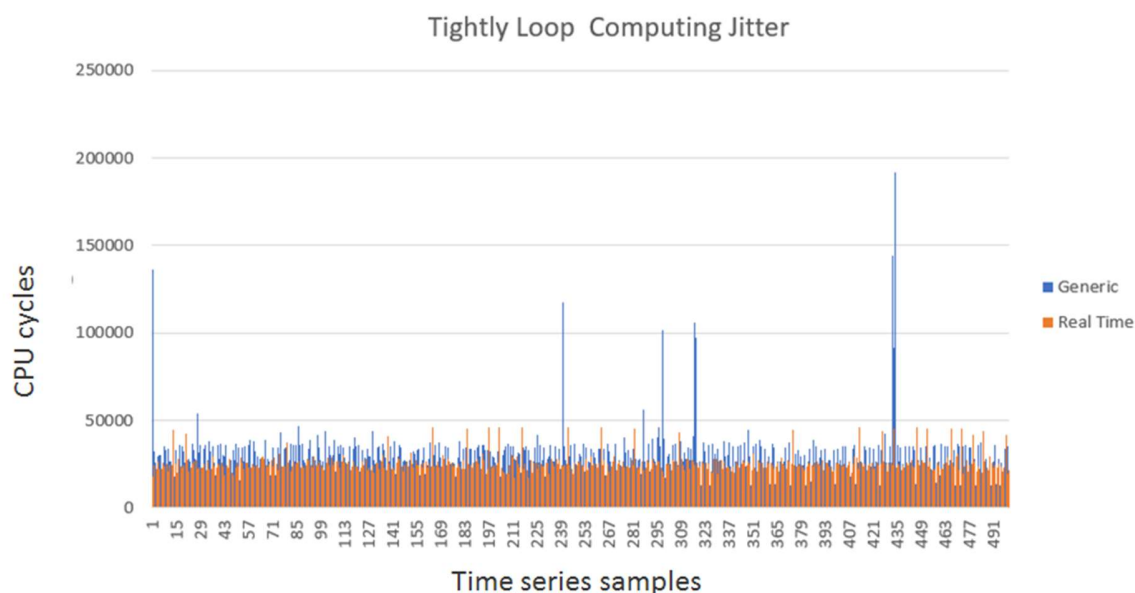


Figure 20. Comparison of Computing Jitter with/without real-time KVM.

5. Conclusion

To digitally transform the legacy network infrastructure is considered to be an immediate need for CSPs today. The QxStack NFV Infrastructure with Red Hat OpenStack Platform solution provides a flexible, reliable, high-performance NFV foundation to realize the cloud transformation.

The building block pre-integrated with QCT NUMA-balanced hardware and Red Hat open source software brings the benefits of agility and flexibility. To minimize the efforts and simplify the deployment process for CSPs, a well-designed QxStack Auto-Deployment Tool is equipped with NFVI foundations for business expansion.

This reference architecture recommends the network planning with EPA features enabled to ensure outstanding network performance and high availability. With the comprehensive validation based on industry-standard OPNFV test framework, CSPs could rely on the solution and build up an NFV infrastructure on the fly. To learn more about the QxStack NFV Infrastructure with Red Hat OpenStack Platform solution, please visit www.qct.io/q/NFVI.

The OpenStack® Word Mark and OpenStack Logo are either registered trademarks / service marks or trademarks / service marks of the OpenStack Foundation, in the United States and other countries and are used with the OpenStack Foundation's permission. We are not affiliated with, endorsed or sponsored by the OpenStack Foundation or the OpenStack community.

LEGAL DISCLAIMER

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH QUANTA CLOUD TECHNOLOGY (QCT) PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN QCT'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, QCT ASSUMES NO LIABILITY WHATSOEVER AND QCT DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF QCT PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

UNLESS OTHERWISE AGREED IN WRITING BY QCT, THE QCT PRODUCTS ARE NOT DESIGNED NOR INTENDED FOR ANY APPLICATION IN WHICH THE FAILURE OF THE QCT PRODUCT COULD CREATE A SITUATION WHERE PERSONAL INJURY OR DEATH MAY OCCUR.

Quanta Cloud Technology (QCT) may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined." QCT reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

All products, computer systems, dates, and figures specified are preliminary based on current expectations, and are subject to change without notice. Contact your local QCT sales office or your distributor to obtain the latest specifications and before placing your product order.



ABOUT QCT

QCT (Quanta Cloud Technology) is a global datacenter solution provider extending the power of hyperscale datacenter design in standard and open SKUs to all datacenter customers.

Product lines include servers, storage, network switches, integrated rack systems and cloud solutions, all delivering hyperscale efficiency, scalability, reliability, manageability, serviceability and optimized performance for each workload.

QCT offers a full spectrum of datacenter products and services from engineering, integration and optimization to global supply chain support, all under one roof.

The parent of QCT is Quanta Computer Inc., a Fortune Global 500 technology engineering and manufacturing company.

<http://www.QCT.io>



UNITED STATES

QCT LLC., Silicon Valley office
1010 Rincon Circle, San Jose, CA 95131
TOLL-FREE: 1-855-QCT-MUST
TEL: +1-510-270-6111
FAX: +1-510-270-6161
Support: +1-510-270-6216

QCT LLC., Seattle office
13810 SE Eastgate Way, Suite 190, Building 1,
Bellevue, WA 98005
TEL: +1-425-633-1620
FAX: +1-425-633-1621



CHINA

云达科技, 北京办公室 (Quanta Cloud Technology)
北京市朝阳区东大桥路 12 号润诚中心 2 号楼
TEL +86-10-5920-7600
FAX +86-10-5981-7958

云达科技, 杭州办公室 (Quanta Cloud Technology)
浙江省杭州市西湖区古墩路浙商财富中心 4 号楼 303 室
TEL +86-571-2819-8650



JAPAN

Quanta Cloud Technology Japan 株式会社
東京都港区芝大門 2-5-8 芝大門牧田ビル 3F, 105-0012
TEL +81-3-5777-0818
FAX +81-3-5777-0819



GERMANY

Quanta Cloud Technology Germany GmbH
Hamborner Str. 55, 40472 Düsseldorf
TEL +492405-4083-1



TAIWAN

雲達科技 (Quanta Cloud Technology)
桃園市龜山區文化二路 211 號 1 樓
1F, No. 211 Wenhua 2nd Rd., Guishan Dist., Taoyuan City 33377,
Taiwan

All specifications and figures are subject to change without prior notice. Actual products may look different from the photos.

QCT, the QCT logo, Rackgo, Quanta, and the Quanta logo are trademarks or registered trademarks of Quanta Computer Inc.

All trademarks and logos are the properties of their representative holders.

Copyright © 2017-2018 Quanta Computer Inc. All rights reserved.