PHD House, 4th Floor, Ramakrishna Dalmia Wing
4/2, Siri Institutional Area, August Kranti Marg, New Delhi – 110016,
Tel# 9355144223 E-mail: dg@mait.com Website: http://www.mait.com

Ref.No.MAIT/PY/2291 August 09, 2023

Shri R R Mittar
Advisor & Head, TEC
Department of Telecommunications

Subject: MAIT Representation on MTCTE Essential Requirement (ER) for Hypervisor TEC30032201

Respected Sir,

Greetings from MAIT!

At the outset, MAIT would like to thank TEC for understanding industry's concern and supporting the industry in every possible aspect especially in terms of MTCTE regime. This is in further reference to TEC's Essential Requirement (TEC30032201) which states requirements for hypervisor products sold in India for Telecommunications purposes. MAIT would like to suggest some inputs for the current ER.

MAIT's interpretation of TEC30032201 ER:

TEC30032201 and Annexure-H2 are requiring that hypervisor products sold for telecommunications deployments to pass a selected list of OPNVF 2018.09 test specifications.

OPNFV Background:

The Open Platform for NFV is a Linux Foundation project that aims to deliver NFV Infrastructure (NFVI) and Virtualized Infrastructure Management (VIM) by integrating components from upstream projects such as SDN controllers, cloud/container orchestration software, virtualization software and dataplane acceleration technologies. The OPNFV project provides test harnesses and test suites to validate if a solution stack based on the combination of open-source projects for each OPNFV release is properly behaving and interoperating with each other.

OPNFV is based on OpenStack and related technologies. OpenStack is an open-source cloud computing platform consisting of interrelated components that control diverse, multi-vendor hardware pools of processing, storage, and networking resources throughout a data center. Many of the test cases in the OPNFV test specification are exercising and testing Open Stack APIs against OpenStack based management solutions.

Examples of Hypervisor:

- **POWER Hypervisor:**

  Enterprise Power systems ship with a hypervisor referred to as the Power Hypervisor. The Power Hypervisor is shipped as part of firmware and performs the function of hardware partitioning and virtualization. Its functionality is similar to other embedded hypervisors such as PR/SM.

The primary purpose of the Power Hypervisor is to manage the partitioning and virtualization of hardware resources in the server and enforce isolation between Logical Partitions (LPARs). Some examples of the functionality it provides for each type of resource would be:

- Compute: CPU state is virtualized and managed to ensure that each LPAR's CPU state is isolated and is not visible to any other LPAR on the server. The hypervisor also manages the enforcement of compute capacity allocation to LPARs.
- Memory: Memory address space is divided into chunks that are assignable to each LPAR. Using translation table mechanisms, the hypervisor ensures that each LPAR can only access the memory assigned to it or authorized to access.
- IO: The hypervisor manages the IO fabric down to the PCI endpoint level. PCI endpoints are then assigned to LPARs which are responsible for initializing and managing the PCI device.
- Accelerators: Some Power Processor generations contain accelerator units for functions such as compression or encryption. The hypervisor manages allocation of these shared resources to LPARs and ensures isolation of state between LPARs using the shared resource.

It should be noted that the Power Hypervisor does not own any IO devices that are assigned to LPARs and does not contain any device drivers for IO devices assigned to LPARs. For example, as with any modern server a variety of network adapters are supported by Power servers, however, the hypervisor itself does not contain any drivers for those devices that are capable of establishing a network connection and initiating network traffic. That functionality is part of the Operating Systems that run in the LPARs. Additionally, since the hypervisor does not own and manage any IO devices, the hypervisor does not own or manage the contents of any storage devices.

In addition to providing hardware isolation, the Power Hypervisor also provides advanced enterprise functionality such as Reliability, Availability, Serviceability (RAS) features. Some examples of the RAS features it provides would be managing IO fabric error recovery, controlling LEDs used for serviceability, detecting failed hardware and generating error logs calling for service, etc.

It is important to note that the Power Hypervisor does not contain any sort of a user interface environment that users can interact with. The control and management of it is performed by other products or components such as PowerVC, Hardware Management Console, Novalink, or service processor menus.

- **KVM Hypervisor:**

All Linux Kernels have the KVM hypervisor included. This Type-1 hypervisor (utilizing CPU technologies) is being used in almost every Cloud offering worldwide and as such is the most used hypervisor worldwide.

The primary purpose of KVM is to isolate virtual machines from the underlying hardware as well as running multiple different Operating Systems on the same hardware platform simultaneously.

KVM operates on the host as a driver for the Linux kernel, which runs on the host as a bare metal operating system. Some examples of the functionality provided by the host kernel would be:

- Compute: CPU state is virtualized and managed to ensure that each virtual CPU state is isolated and is not visible to any other processes on the hypervisor. The hypervisor also manages the enforcement of compute capacity allocation to VMs.
- The Linux memory management is used to assign memory to the process that is running the VM (usually QEMU). With that the same memory isolation applies as for any other process. In addition, AMD and Intel provide memory encryption technologies (SEV, SEV-ES, and in the future SEV-SNP and TDX) that can be used to encrypt the memory for the Virtual Machine.
- IO: The bare metal operating system manages the IO fabric down to the PCI endpoint level. PCI endpoints can either use drivers provided by the host, or they can be assigned to VMs which are responsible for initializing and managing the PCI device. In the latter case, the configuration performed by the virtual machine is mediated by another device driver, VFIO, which will emulate/virtualize as necessary the configuration steps; VFIO also ensures that DMA and interrupt injection are protected by an IOMMU.
- Privilege separation: KVM code running in supervisor mode only includes functionality for virtualization of CPUs and memory. The operation of virtual devices (such as virtio-blk) is managed by a userland process called QEMU that translates the virtual machine's requests to system calls on the host. This process is confined through Linux functionality such as cgroups, SELinux, namespaces and seccomp filters.

Conclusion:

1. <u>The testing scope of TEC ER TEC30032201 is significantly broader than the hypervisor's function and is not an adequate standard for Hypervisor. TEC shall instead release a different ER that is constrained to Hypervisor layer for public consultation before enforcement.</u>

   MAIT believes that intention of TEC ER TEC30032201 is to ensure that products purchased for telecommunications deployments using NFV type functionality in India are interoperable with OPNVF specifications based on Open Stack, however the hypervisor is not the software entity that implements the functionality that OPNVF is specifying. The hypervisor is just one part of an overall solution stack that is defined by OPNVF. Furthermore, the scope of the functionality the OPNFV test suites target is not functionality that is implemented by hypervisors or any single entity in a solution stack, but instead functionality that is implemented by cloud management solution stacks and the various pieces of software and networking devices that make up those solution stacks.

2. Hypervisor can perform its function only when installed in a server product. Server is currently excluded from MTCTE scope according to notification 6-8/2020/TC/TEC/75 dated 2022/03/22. Thus, Hypervisor shall not be regulated separately.

3. Many Hypervisors in the market are not designed based on Open Stack. Those non-Open Stack Hypervisors can only be installed on specific hardware products. Testing such Hypervisor also requires this specific hardware. TEC shall not force all Hypervisor to go with Open Stack.

Opinion by ER and Annexure No:

1.1.1 Image Services:

OPNFV Test Specification 8.1.7:
https://docs.opnfv.org/projects/dovetail/en/latest/testing/user/testspecification/tempest_image/index.html

This test suite is focused on testing the VM image management services. VM image management is not functionality that is provided by a hypervisor. It is functionality that is provided by a cloud orchestrator.

1.1.2 Interoperabiity:

OPNFV Test Specification 8.1.13:
https://docs.opnfv.org/projects/dovetail/en/latest/testing/user/testspecification/tempest_osinterop/index.html

This test suite is focused on testing OpenStack API availability to ensure interoperability between various components in a solution stack. This is not functionality that is provided by a hypervisor. It is functionality that is provided by a cloud orchestrator and the various OpenStack drivers in the cloud orchestrator solution.

1.1.3 Life Cycle Management:

OPNFV Test Specification 8.1.15:
https://docs.opnfv.org/projects/dovetail/en/latest/testing/user/testspecification/tempest_vm_lifecycle/index.html

This test suite is focused on VM management operations such as starting/stopping a VM, managing the storage assigned to it, migrate it between servers, etc. While many of these things are functions that a hypervisor is responsible for initiating, it is the role of the software managing the hypervisor, such as cloud orchestration software, to coordinate these sorts of activities. For example, this test suite includes things such network configuration of routers and subnets for VMs, which the hypervisor has no knowledge of, only the operating systems running in the VMs.

1.1.4 Network Security Group:

OPNFV Test Specification 8.1.5:
https://docs.opnfv.org/projects/dovetail/en/latest/testing/user/testspecification/tempest_compute/index.html

This test suite is focused on the network security related operations, such as creating security groups and the networking configuration associated with them, such as router, subnet, IPs, security keys, virtual ports, etc. The Power hypervisor does not have knowledge of anything beyond L2 networking.

1.1.5 Networking:

OPNFV Test Specification 8.1.8, 8.1.10. 8.1.11, 8.1.14, 8.1.16. 8.1.18 & 8.1.19:
https://docs.opnfv.org/projects/dovetail/en/latest/testing/user/testspecification/tempest_ipv6/index.html

https://docs.opnfv.org/projects/dovetail/en/latest/testing/user/testspecification/tempest_network_api/index.html

https://docs.opnfv.org/projects/dovetail/en/latest/testing/user/testspecification/tempest_network_scenario/index.html

https://docs.opnfv.org/projects/dovetail/en/latest/testing/user/testspecification/tempest_trunk_ports/index.html

https://docs.opnfv.org/projects/dovetail/en/latest/testing/user/testspecification/tempest_volume/index.html

https://docs.opnfv.org/projects/dovetail/en/latest/testing/user/testspecification/vping/index.html

https://docs.opnfv.org/projects/dovetail/en/latest/testing/user/testspecification/vpn/index.html

This suite of tests is focused on IP network configuration in the following areas:

8.1.8 – IPv6 Support

8.1.10 – IPv4 Support

8.1.11 – IPv4 Support

8.1.14 – Trunking Support

8.1.16 – SAN Volume Support

8.1.18 – Network connectivity between VMs

8.1.19 – VPN Support

This suite of tests is focused on testing various network and IP configuration functionality for the VMs and network infrastructure being managed by the cloud orchestration software. The Power Hypervisor is not involved in networking at the IP layer, SAN volume management, or VPN support. This set of tests appears to be intended to validate that an OpenStack implementation properly configures the cloud infrastructure providing the compute, network and storage and is outside the scope of responsibilities of a hypervisor.

1.1.6 Resiliency:

OPNFV Test Specification 8.1.1 & 8.1.4:
https://docs.opnfv.org/projects/dovetail/en/latest/testing/user/testspecification/highavailability/index.html
https://docs.opnfv.org/projects/dovetail/en/latest/testing/user/testspecification/stress/index.html

These test suites are focused on the ability of the solution to properly function in conditions where servers in the infrastructure being managed experience high load stress and or failures. This is not the functionality that a hypervisor provides. It is the functionality that cloud orchestration software provides.

1.1.7 Role Based Access Control:

OPNFV Test Specification 8.1.2, 8.1.3, 8.1.6:
https://docs.opnfv.org/projects/dovetail/en/latest/testing/user/testspecification/security_patrole/index.html
https://docs.opnfv.org/projects/dovetail/en/latest/testing/user/testspecification/security_patrole_vxlan_dependent/index.html

https://docs.opnfv.org/projects/dovetail/en/latest/testing/user/testspecification/tempest_identity_v3/index.html

These test suites are focused on user and group management for the purposes of managing a cloud infrastructure. This is not functionality that is provided by hypervisors such as the Power Hypervisors, where all user authorization management and tracking is handled by the management entity outside of the hypervisor.

1.1.8 Scheduling:

OPNFV Test Specification 8.1.9:
https://docs.opnfv.org/projects/dovetail/en/latest/testing/user/testspecification/tempest_multi_node_scheduling/index.html

This test suite is focused on the capabilities of a cloud orchestrator to manage the scheduling of VMs on different servers subject to various affinity, security group, and network availability polities. This is not the functionality that a hypervisor provides. It is the functionality that cloud orchestration software i.e. managing hypervisor provides.

We look forward to your positive consideration of our request.

Warm regards

Col Suhail Zaidi (Retd)
Director General

CC: Shri Anand Katoch, DDG (TC), Telecom Engineering Centre