An OpenStack Deployment Success Story

At a glance
The OpenStack movement has gained critical mass, and many Fortune 1000 companies are either deploying or considering deploying an OpenStack based solution. OpenStack allows an enterprise to treat its IT infrastructure as an on-demand, on-call service. The elasticity and agility to provision, develop, deploy, maintain, monitor and scale-out provides capital expenditure (CAPEX) and operational expenditure (OPEX) advantages. A network adapter is a critical component for OpenStack deployments and selecting the right one can deliver even better performance and reduced CAPEX and OPEX.

This paper provides an overview of a recent OpenStack deployment using Red Hat OpenStack Platform (OSPv6), Lenovo ThinkServer servers and storage, the Midokura software-defined networking (SDN) solution Midonet, and Emulex 10 and 40Gb Ethernet (GbE) Converged Network Adapters (CNAs).

The OpenStack PRISM
OpenStack software is like a prism – it translates a complex set of IT infrastructure into a simple and refined view with three primary components: compute, networking and storage. The abstracted simplification provides the user applications with more resiliency, high availability and scalability.

Figure 1. The OpenStack Prism.

Proof of concept (PoC) OpenStack deployment for healthcare enterprise
A leading $25B healthcare enterprise was looking for a private cloud solution to leverage the benefits of the cloud including rapid and flexible deployment, scalability and reduced total cost of ownership (TCO). An expert solutions team put together an OpenStack PoC concept and a successful deployment followed.

Figure 2 provides an overview of the OpenStack environment design. It comprises of 16 Lenovo ThinkServer RD550 compute nodes, 8 Lenovo ThinkServer RD650 storage nodes and 4 ThinkServer RD650 controller nodes. Emulex OcE14102 10GbE CNAs and Lenovo RackSwitch G8264 10GbE switches provide the 10GbE network connectivity.
An OpenStack Deployment Success Story

In Figure 2 above, the compute nodes labelled CPT1 to CPT16 are used to provision the cloud application workload, they have access to a central Ceph-based storage, comprised of nodes Ceph1 to Ceph8, sized at ~500TB storage.

Red Hat Enterprise OpenStack Platform 6 is an open source based platform based on the OpenStack Juno release and is used as the software framework to build the private cloud.

Midonet is an open source based software solution that provides tenant isolation and various networking functions in a multi-tenant deployment. Midonet has a Neutron (OpenStack networking component) plug-in and it allows tenants to create logical topologies via virtual routers, networks, NAT and load balancing. The distributed software installs Midonet agents on compute nodes and the controller gateway labelled GW1 to GW4 in Figure 2.

The Baseboard Management Controller (BMC) and advanced management functionality is provided by the Lenovo xCAT tool. The highly scalable and OpenStack network is a redundant network using a leaf-spine topology as shown in figure 3. More details of the setup, topology and configuration can be found in the Lenovo and Midokura PoC document.

Emulex OneConnect® OCe14102 is a 4th generation 2x10GbE port CNA. The OneConnect family of adapters has an install base numbered in the millions of ports worldwide and has a proven track record for network and storage hardware acceleration in bare metal and hypervisor virtualized environments.

The network adapter is an important element in any OpenStack deployment, as it provides the network plumbing to compute servers, switches and storage. OCe14102 provides not only the connectivity, but also the following benefits:

- **Converged Network**—Networking and storage traffic can run on the same 10GbE port, thereby reducing the port counts, power and cabling requirement.
- **Diskless boot using PXE**—PXE boot allows the server to boot over network using a remote boot image, an important feature also provisioned by OpenStack Ironic.
- **Higher VM density**—The CNA has a multi-processor architecture, and supports various types of hardware accelerations. Several protocols and tasks can be offloaded from the CPU to hardware, saving CPU cycles that in turn allow higher virtual machines (VMs) per server to be supported.
- **Virtualization**—The CNA provides multiple levels of virtualization support. In the simplest forms, Receive Side Scaling (RSS) allows pinning traffic flows to dedicated cores on a multi-core, multi-socket server. Single Root I/O Virtualization (SR-IOV) provides VMs direct access to hardware resources and saves the costly data copies in hypervisor. Network Virtualization using Generic Routing Encapsulation (NVGRE), Virtual Extensible Local Area Network (VXLAN) offload and Virtual Ethernet Bridging (VEB) provide network level virtualization.

* Lenovo and Midokura OpenStack PoC  https://lenovopress.com/redp5233
An OpenStack Deployment Success Story

- **Quality of Service** — The CNA provides sophisticated quality of service (QoS) knobs, in that the individual traffic flows are uniquely identified and dynamically controlled using fine grained rate limiting in hardware. The support for Priority Flow Control (PFC), Link Layer Discovery Protocol (LLDP) and Data Center Bridging Exchange (DCBX) allows for lossless and out of band QoS configuration.

- **Virtual Extensible Local Area Network (VXLAN)** — VXLAN is a network technology that extends the reach of Layer 2 traffic over an IP network by slapping a VXLAN header on top of a Layer 2 frame. VXLAN provides network isolation and scalability in a multi-tenant deployment.

As shown in figure 3, Midonet agents on the compute server and the controller gateway use VXLAN for traffic isolation in the OpenStack deployment. The VXLAN gateway terminates the VXLAN tunnel for external network access.

VXLAN adds an overhead of about 50 bytes on each packet, in both transmit and receive path, for the compute and gateway nodes. The 50 byte overhead and packet processing times on the node adds visible overhead for various OpenStack workloads. Emulex OCE14102 simplifies VXLAN deployment by offloading VXLAN from the hypervisor kernel to hardware adapter, a feature that is supported across the majority of operating systems (OS), including inbox support on RHEL 7.1 (inbox).

When VXLAN offload is enabled in the OCE14102, it can result in significant savings of server CPU resources compared to a standard adapter without VXLAN offload capability. This translates into more free CPU cycles, increased VM density and lower power consumption.

In a scale-out architecture, the savings scale-out proportionate to the number of nodes and adapters in the OpenStack environment. Figure 4 shows the benefit of VXLAN offload in comparison to VXLAN offload disabled on the same OCE14102 adapter, as well as showing a standard adapter that does not support VXLAN. The circle size indicates CPU utilization, therefore, the smaller the size, the lower the CPU utilization.

In summary, by deploying an OpenStack solution, this leading healthcare enterprise is benefiting from flexibility and cloud scalability as well as achieving CAPEX and OPEX savings. Put in Automotive Industry terms, selecting the right tires is important for a high performance car, and so is selecting the right network adapter for OpenStack deployments. Emulex OneConnect OCE14102 provides the flexibility and performance benefits for a diverse set of OpenStack workloads.

Contact the author Saurabh Sureka - saurabh.sureka@avagotech.com for OpenStack related queries or any feedback.