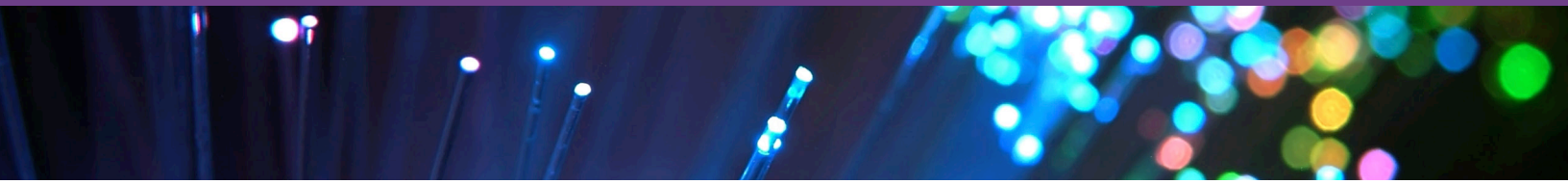


Achieving a High-Performance Virtual Network Infrastructure with PLUMgrid IO Visor™ & Mellanox ConnectX®-3 Pro

Whitepaper



What's wrong with today's clouds?

Compute and storage virtualization has enabled elastic creation and migration of applications within and across data centers. Unfortunately, on-demand network infrastructure provisioning in multi-tenant cloud environments is still very rudimentary due to the complex nature of Physical Network Infrastructure (PNI) and limited network functionality in hypervisors. Even after this complexity is somehow mastered by the cloud operations and IT teams, even simple reconfigurations (let alone complex upgrades) of the network remain largely error-prone.

The same networking problems percolate upstream into the tenants' virtual networks, where deployment, configuration and management of anything beyond a simple Virtual Private Connect bridge is still a significant challenge. These constraints result in wastage of thousands of dollars and months of effort for cloud consolidation or upgrade projects. In other words, the network has become the fundamental bottleneck in cloud deployments and operations.

PLUMgrid OpenStack Networking Suite 3.0: A one-stop shop for solutions to Operator and Tenant networking problems of your cloud

PLUMgrid OpenStack Networking Suite 3.0 [Figure 1] solves both operator and tenant networking problems. For the operator problems, all the core elements that currently exist in the physical infrastructure (IPMI, boot, CMS, etc.) are transitioned into the VNI, thus making it straightforward to deploy, manage and extend the underlying cloud infrastructure.

For tenant problems, PLUMgrid OpenStack Networking Suite 3.0 (Figure 1) enables users of a cloud to create and deploy *Virtual Domains* in seconds without any change to the physical network fabric. These Virtual Domains abstract away the underlying physical network and provide private switching, routing, policies, load balancing, NAT and DHCP services to tenants while ensuring total tenant-isolation and traffic security. Each tenant has full administrative control over the networks contained within its Virtual Domains. This is achieved through the PLUMgrid OpenStack Networking Suite 3.0 overlay network that leverages the VXLAN encapsulation protocol to decouple network services from the physical, device-based network.

Overlay networks standards, such as VXLAN, make the network as agile and dynamic as other parts of the cloud infrastructure. This agility in turn enables dynamic network segment provisioning for cloud workloads which induces dramatic increases in cloud resource utilization.

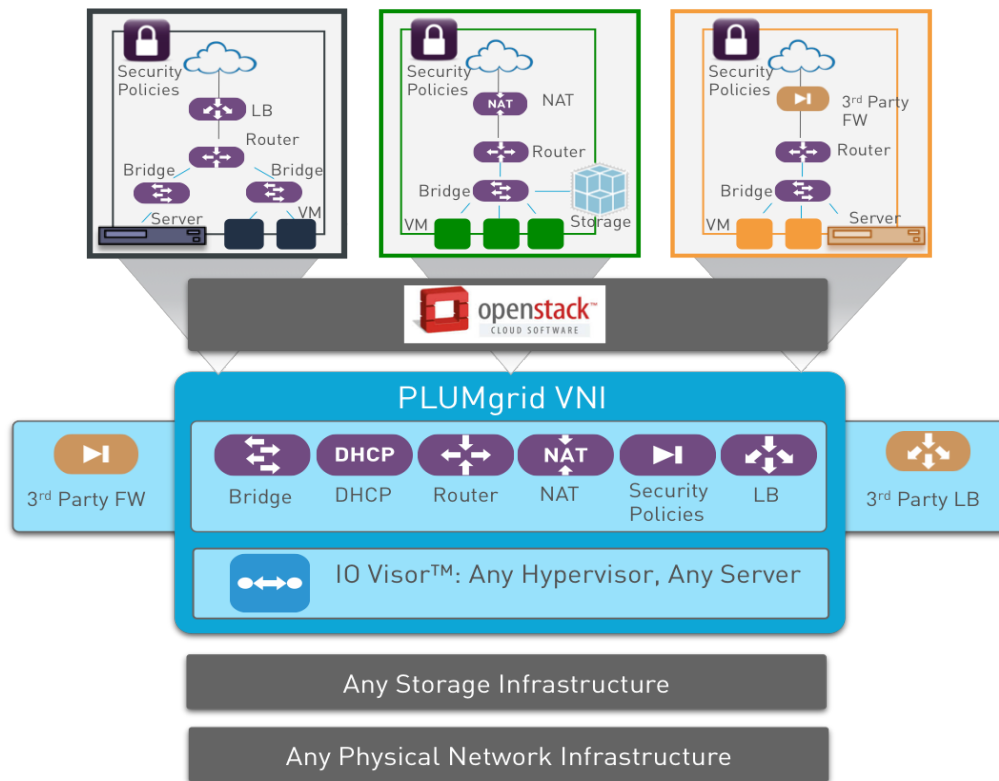


Figure 1: The PLUMgrid OpenStack Networking Suite 3.0

Mellanox ConnectX-3 Pro: Resolving VNI Performance Bottlenecks

While overlay-based VNI offers significant benefits in terms of flexibility, scalability and security for cloud infrastructure, the additional VXLAN processing remains a concern for high performance environments.

VXLAN uses an additional encapsulation layer, and therefore traditional NIC offloads cannot be utilized, resulting in high levels of CPU consumption. In addition to consuming expensive CPU processing resources, network throughput is also significantly reduced and a software-only solution simply does not scale with an ever-increasing number of workloads per server.

In order for VXLAN to be of real value, its inherent CPU overhead and network performance degradation must be eliminated. This can be achieved by offloading the overlay network processing to hardware embedded within the network controllers.

Mellanox's ConnectX-3 Pro is the only NIC that currently handles the traditional offloads despite the extra encapsulation layer. Furthermore, ConnectX-3 Pro enables RSS to steer and distribute traffic based on the inner packet, and not , like traditional NICs, only the outer packet. The primary benefit of this is that it allows multiple cores to handle traffic, which then enables the interconnect to run at the intended line-rate bandwidth, not at reduced bandwidth as occurs when VXLAN is used without RSS.

ConnectX-3 Pro addresses the unintended performance degradation seen when using VXLAN in both CPU usage and throughput. By using ConnectX-3 Pro, the scalability and security benefits of VXLAN can be utilized without any decrease in performance and without incurring an increase in CPU overhead.

Benchmarking results of PLUMgrid OpenStack Networking Suite 3.0 with Mellanox ConnectX-3 Pro

We benchmarked PLUMgrid OpenStack Networking Suite 3.0 with VXLAN encapsulation offloads on Mellanox ConnectX-3 Pro. Performance was measured for a single Virtual Domain with a Bridge topology configured on an IO Visor (running as a kernel module), with Mellanox ConnectX-3 Pro providing the underlying tunnel encapsulation/decapsulation functions.

All test results were collected with the following hardware and software configurations:

- Server: Dell PowerEdge R620.
- CPU: Dual Intel Xeon E5-2670 v2 @ 2.5GHz (10 cores with HT).
- RAM: 128GB RDIMM.
- NIC: Mellanox ConnectX-3 Pro.
- Hypervisor: KVM Ubuntu 12.04 with kernel version 3.14-rc8 (mlx4 VXLAN offload support).
- VNI: PLUMgrid OpenStack Networking Suite version 3.2 Build 8.
- Virtual Domain Configuration: Single Virtual Bridge.
- Traffic: `super_netperf [super_netperf 4 -H $IP -l 30 -- -m $SZ]`.
- VM pair: Two VMs on the same virtual bridge but on different hypervisors/servers.

Figure 2 shows the throughput comparison with a fixed packet size (9000 bytes) with and without VXLAN offloading. The improvement in performance with VXLAN offloads is significant (generally an order of magnitude) and the performance increases almost linearly until reaching the close-to-optimal 36 Gbps mark (with eight VM pairs generating traffic at maximum rates).

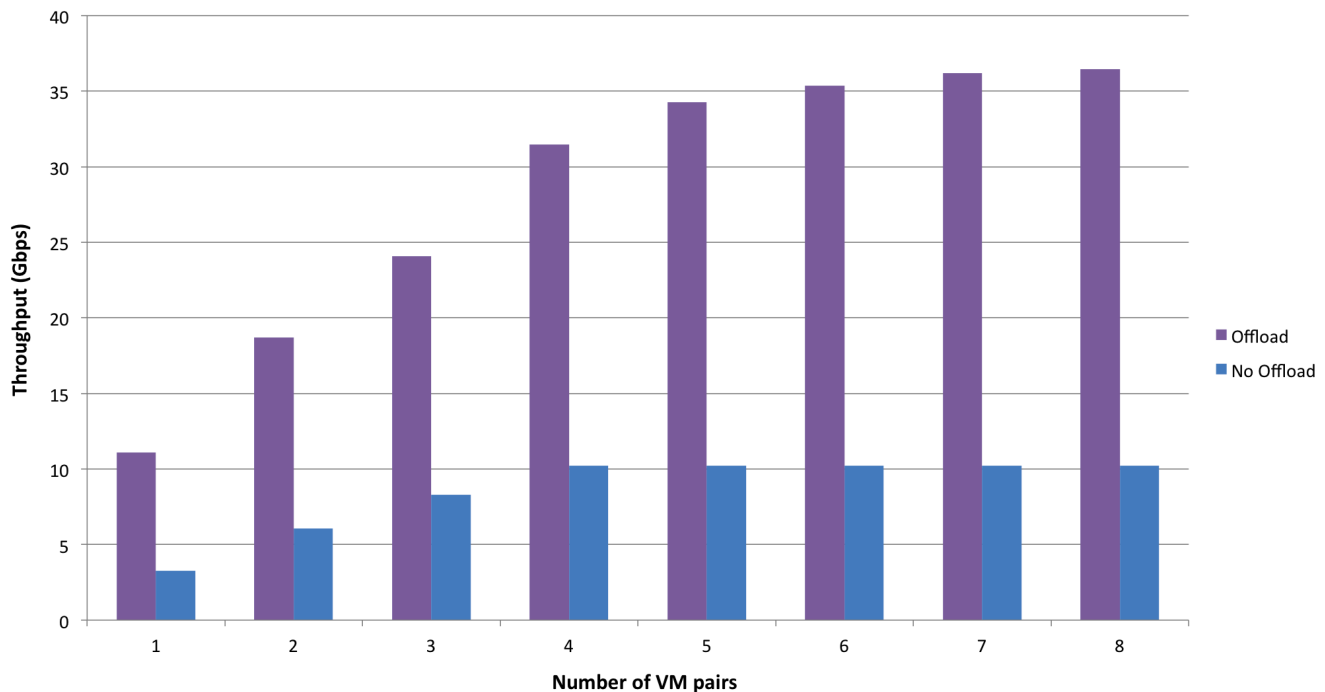


Figure 2: Throughput of PLUMgrid OpenStack Networking Suite 3.0 with and without Mellanox ConnectX-3 Pro VXLAN offloading

Figure 3 shows the PLUMgrid OpenStack Networking Suite 3.0 throughput with ConnectX-3 Pro's VXLAN off-loading under different packet sizes. It can be observed that for packets with sizes greater than 512 bytes, the traffic is processed at close to line rate. For very small packets (< 128 bytes), the throughput is lower but still reaches 5-10 Gbps for 4 or more VM pairs.

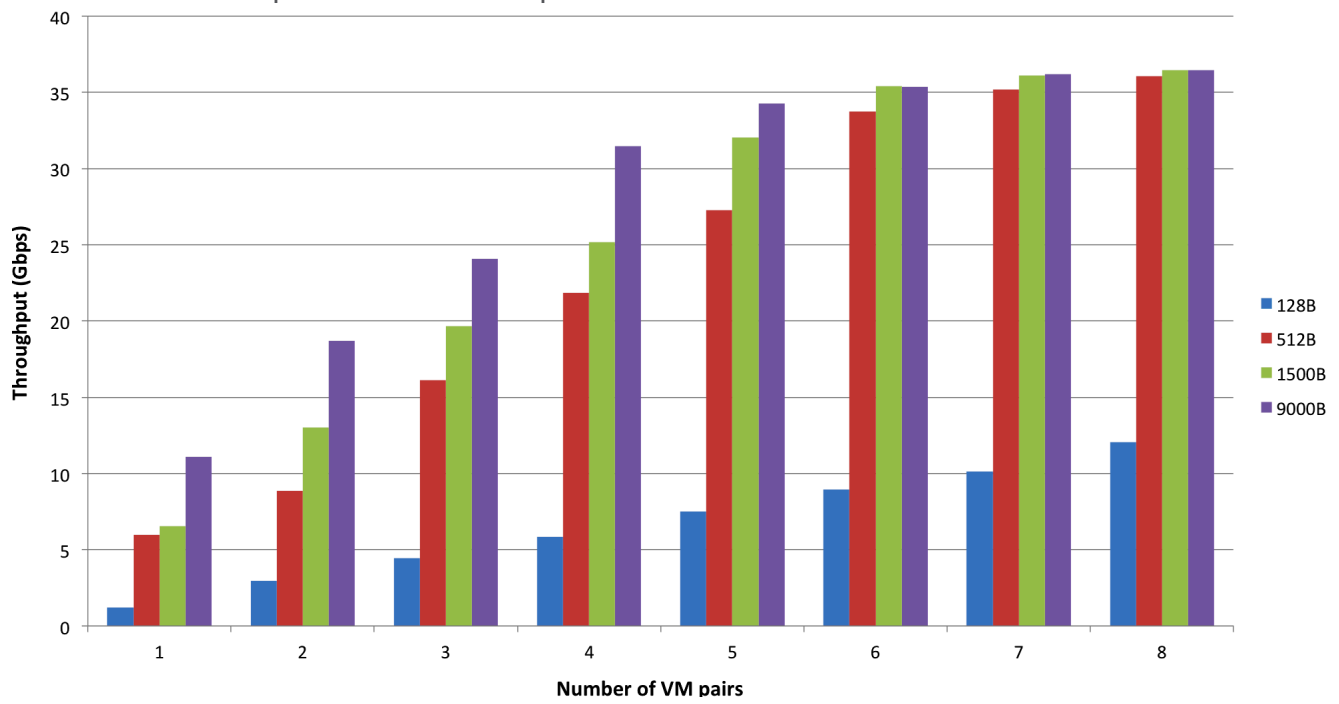


Figure 3: Throughput under varying packet sizes of PLUMgrid OpenStack Networking Suite 3.0 with Mellanox ConnectX-3 Pro VXLAN

Figure 4 shows the CPU utilization in comparison to the obtained throughput with a fixed message size (9000B) and VXLAN offload. It can be easily observed how the CPU idle percentage remains nearly constant on both TX and RX ends, while the throughput grows to 36 Gbps.

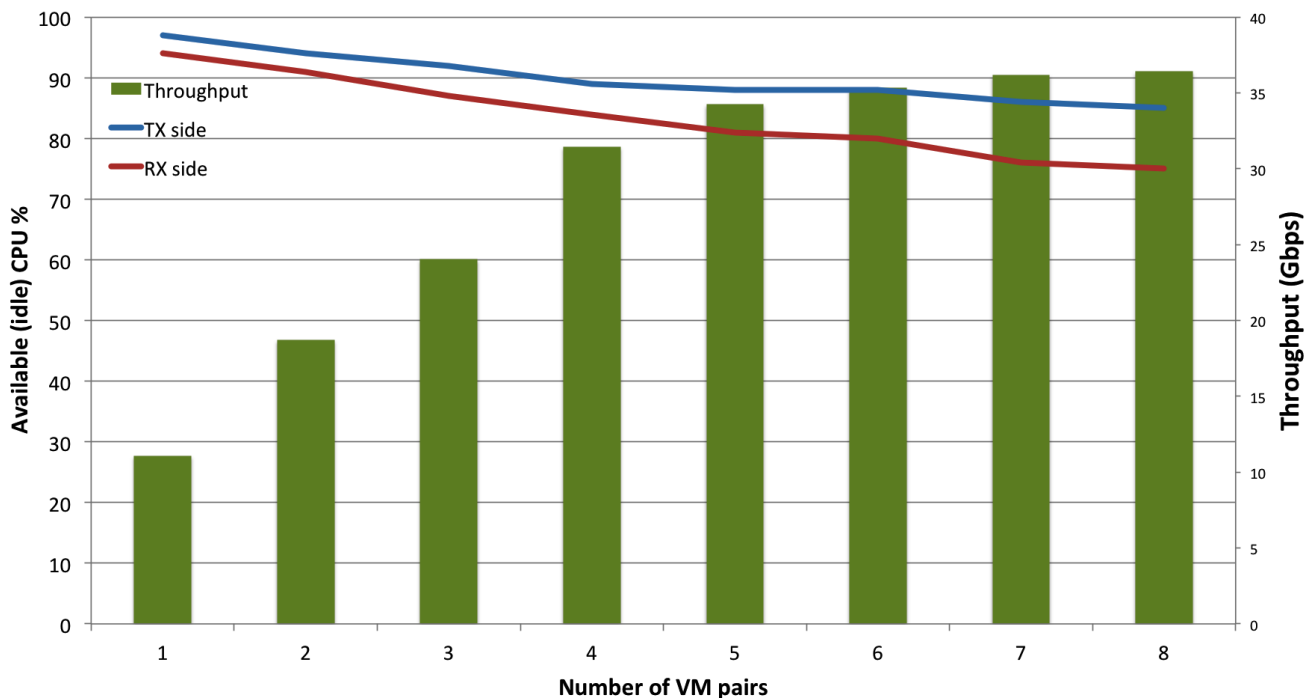


Figure 4: Available (Idle) CPU for PLUMgrid OpenStack Networking Suite 3.0 with Mellanox ConnectX-3 Pro VXLAN

Conclusion

PLUMgrid OpenStack Networking Suite 3.0 with Mellanox ConnectX-3 Pro provides a solution which allows you to enjoy the automation and flexibility of a VNI solution at wirespeed traffic processing rates. The performance benchmarking results of this whitepaper substantiate that the right VNI and NIC architectures can ensure that networking performance remains uncompromised as the number of tenants and workloads increase in your cloud.