*Where IT perceptions are reality*

# Test Report

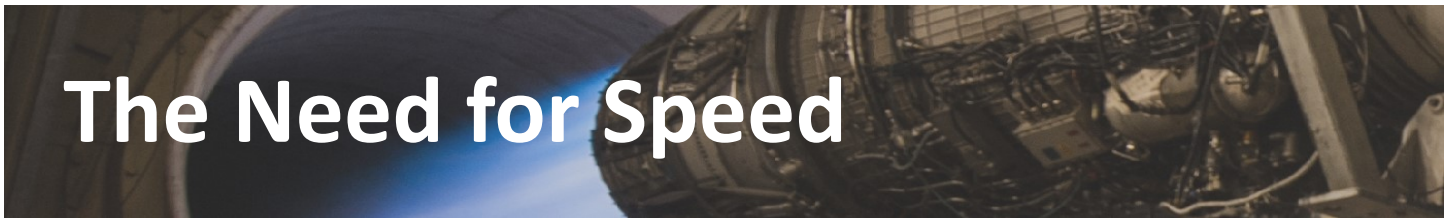
# OCe14000 Performance

## Featuring

Emulex OCe14102 Network Adapters

Emulex XE100 Offload Engine

Document # TEST2014001 v9, October 2014



# The Need for Speed

## Different Applications Need Different Types of Performance

A typical data center is filled with the servers hosting  dozens, hundreds, even thousands of applications. I/O to database driven business servers is most likely read-intensive, while I/O to a VoIP application is very write-intensive. Servers in clusters hosting Big Data Analytics need low-latency connections, while video servers need massive bandwidth. Server architects designing an efficient data center need network adapters capable of performing well in every dimension. Next-gen network adapters need to support higher bandwidth, lower latency, higher read throughput and higher write throughput than the last generation of adapters.

## A New Requirement: Accelerating the Performance of Overlay Networks

The explosive growth of company's network size is driving a requirement to create logical networks that are scalable and able to connect machines (nodes) across different networks.  For the data packets traveling across this vast series of networks, they need to find their way to the destination node without getting mixed or lost with the rest of the packets flowing through the network. This is done through Virtual Local Area Network (VLAN) Tagging.  VLAN Tagging is a method that identifies the packets travelling through networks and ensures delivery to the intended node.

Per the IEEE 802.1Q specification, the maximum limit for a Virtual Area Network (VLAN) is 4,094 tags and what happens when a network has reached that limit?  If a network administrator has a requirement to create logical networks that are scalable and links his virtual machines across different networks, then he needs additional tools.

Enter network overlay protocols Virtual eXtensible Local Area Networking (VXLAN) supported by VMware, and Network Virtualization using Generic Routing Protocol (NVGRE) supported by Microsoft. VXLAN land NVGRE allow the network administrator to create a Layer 2 network on top of a Layer 3 network using encapsulation, increase the number of available private networks to up to the theoretical maximum of 16 million.

## Test Objectives

The goal of the testing covered in this report is to validate the performance and efficiency advantages of next-gen Emulex OCe14102 10GbE adapters over last-gen Emulex OCe11102 10GbE adapters, in a variety of environments and application workloads.

To achieve this goal, three types of performance tests were run:

- VXLAN
- NIC packet rate
- iSCSI IOPs via HW iSCSI offload

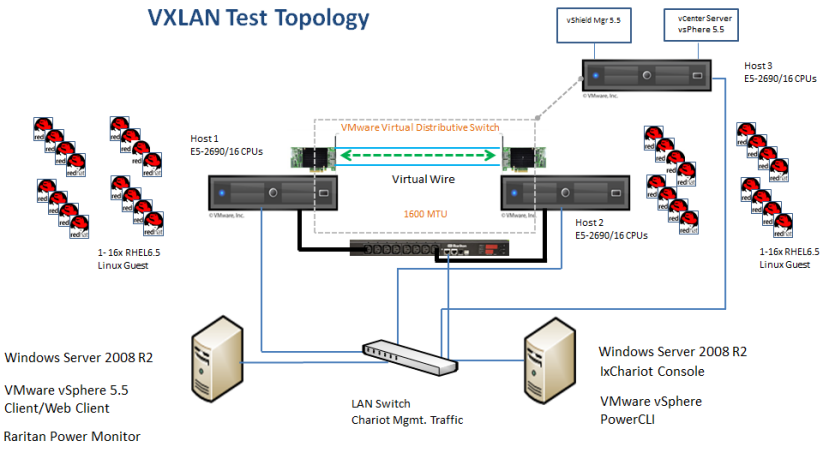
**VXLAN** | **A network virtualization technology that uses a VLAN-like encapsulation technique to encapsulate MAC-based OSI layer 2 Ethernet frames within layer 3 UDP packets.**

# VXLAN Offload Test Set-up

## Virtualized Server Environment

A VXLAN test topology was built within a Windows Server 2008 R2 and VMware vSphere 5.5 environment. Using industry standard plans and tools, the testing collected and documented data measurements including throughput, CPU utilization, CPU effectiveness and server power efficiency. CPU Test workloads included 1, 4, 8 and 16 thin-provisioned RHEL6.5 guest virtual machines configured on two ESXi hosts connected back to back. The RHEL6.5 virtual machines were configured with 2 vCPUs, 4 GB virtual memory, provisioned as servers only (no graphics), with limited services running inside each VM in order to achieve maximum performance. Ixia iXChariot benchmarking software was integrated with the VMware PowerShell CLI platform to start the IXIA network traffic generator, and to collect aggregated bandwidth and CPU utilization performance data from all ESX hosts UUT. Both hosts were connected to a Raritan PDU 120V power line to capture power usage during the tests.

## VXLAN Test Topology




The diagram below includes two ESXi5.5 hosts/ UUTs each with 16 E5-2690 CPUs 2.90 GHz, and 64 GBs memory. Hyper-threading was enabled and the RHEL 6.5 guest OS was configured with a recent version of VMware tools for Linux (version 9.4.5-1618308). VMware vShield Manager 5.5 and vCenter Server 5.5 appliances were configured to a third ESXi5.5 host to run all VMware management operations to a dedicated machine.

## Products Tested

1. Emulex OCe11102 adapters were tested using 4.6.100 VMware driver and 4.6.142 firmware

2. Emulex OCe14102 adapters were tested using 10.0.783 VMware driver and 10.0.803.19 firmware

## Test Methodology

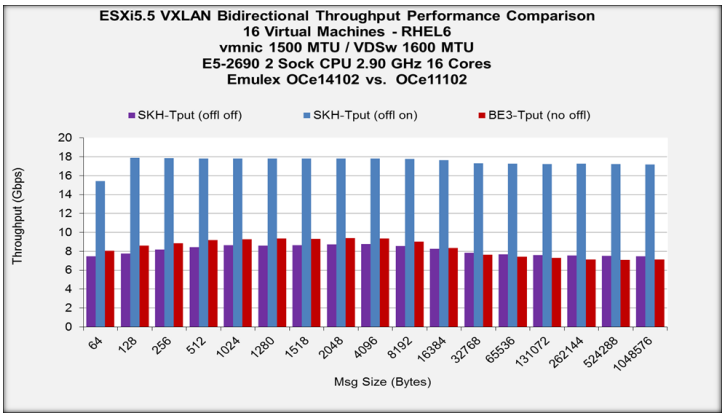
Both ESX hosts were connected to a Raritan 120V PDU, and power usage monitored from a different Windows machine. The tests were started from a Windows 2008 R2 server with VMware vSphereCLI scripts that used IxChariot bidirectional tests running 8 TCP threads per VM. After about 1 minute into the test, the vSphereCLI scripts collected 4 samples of PCPU utilization data from each host, for each message size tested, and saved the results to a csv file. The IxChariot throughput results were also collected and saved on a Windows 2008 R2 machine. Total test run time was approximately two minutes for each message size. The data collected included transmit and receive CPU utilization, plus throughput and power usage data for both incoming/outgoing VXLAN packets.

**NVGRE** | **A network virtualization technology that uses Generic Routing Encapsulation (GRE) to tunnel layer 2 packets over layer 3 networks.**

# VXLAN Offload Test Results

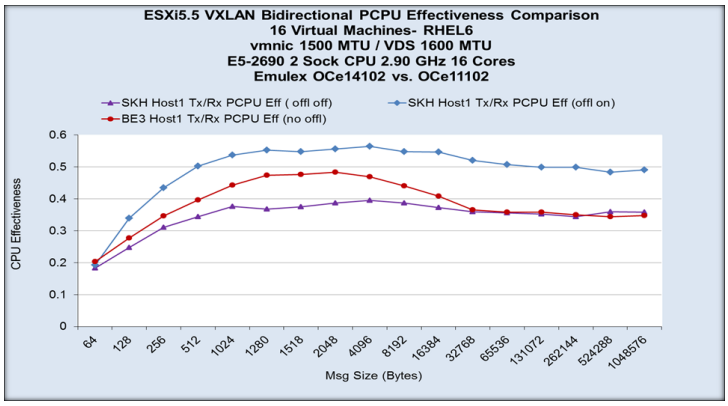
## Bidirectional Throughput (TPUT) Test Results

The chart below demonstrates a large test workload of 16 VMs configured on each ESX host includes TPUT results for the OCe14102 adapter with the VXLAN offload enabled/disabled and OCe11102 native VXLAN. The OCe14102 with VXLAN offload-enabled resulted in improved performance across a range of message sizes. For example, throughput for 4096 byte messages sent through the OCe14102 with offload enabled is 129% higher than with offload disabled. Similarly, throughput for messages sent through a OCe14102 was 143% better than the throughput with a OCe11102 .



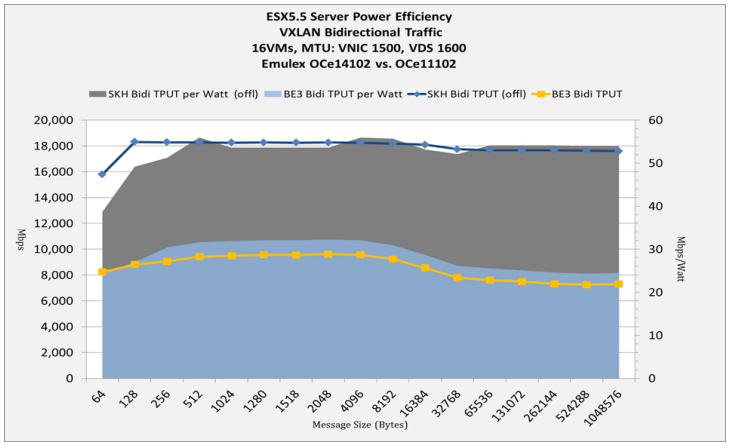

## CPU Effectiveness Test Results

CPU effectiveness tends to improve with larger packets but for a typical 4096 byte message size, the OCe14102 shows a 46% increase in CPU effectiveness with offload enabled. For larger message sizes, such as 32k or higher, throughput is over 42% better than the OCe11102 running the same workload.




## Server Power Efficiency Test Results

The server power efficiency of the OCe14102 with offload-enabled is as much as 83% better than the OCe11102 with small 512 byte message sizes. The server power efficiency of the OCe14102 is as much as 122% better with larger message sizes of 32k and higher, for the same VM workload.




**122%** | **Throughput for 4096 byte messages sent through a OCe14102 with offload enabled was 122% better than the throughput with offload disabled.**

# Small Packet Test Set-Up

## Small Packet Test Set-Up

The lab environment for the RFC 2544 tests consists of the Ixia IxAutomate Benchmarking test suite with pre-defined, automated tests for RFC 2544 Benchmarking, as well as a Ixia XM2 portable chassis with 10Gbps Universal Multi-Service modules connected back-to-back with a server running RHEL 6.4, with IP forwarding enabled. The server also was configured with Emulex OCe14102 and OC11102 adapters with a NIC personality. This is one of the recommended set up configurations for the RFC 2544.

## Products Tested

**1.** OCe14102: Firmware: 10.2.348.6, Driver: 10.2.348.0

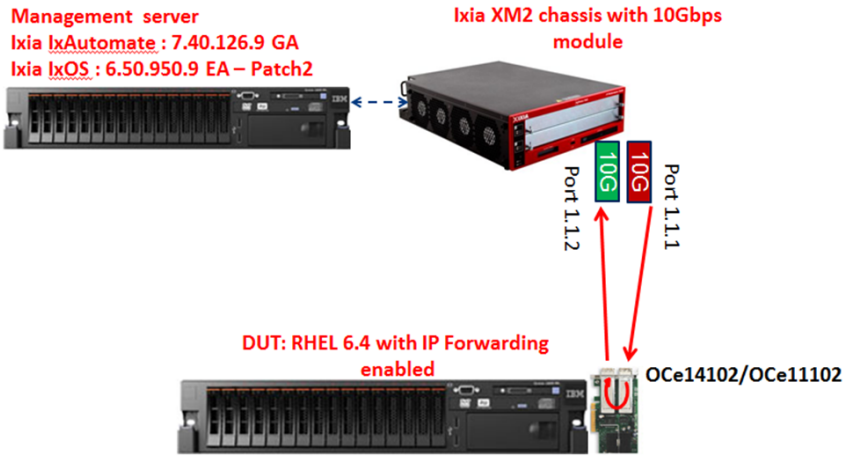
**2.** OCe11102: Firmware: 4.6.142.10, Driver: 4.6.142.5

## BIOS Performance Tuning:

Below are the changes made to the BIOS :

1. Maximum performance was selected for HP Power Profile.
2. Static High Performance mode was selected for the Power Regulator.
3. Optimized for Performance was selected for DIMM Voltage Preference.
4. Maximum Performance was selected for Memory Power Savings Mode.
5. Hyper-threading was disabled.
6. C-states was disabled.

**Small Packet Test Topology**




## OS Performance tuning:

A performance tuning script provided by Emulex was used to set up the Tx,Rx queue lengths and various other parameters.

## RFC 2544 Test Settings:

A traffic map was set up for automatic and unidirectional traffic. For Learning Frames, Frequency was set to 'Once per Test', Wait Time before Transmit was set to '1000 ms', and Number of Frames was set to '10'.

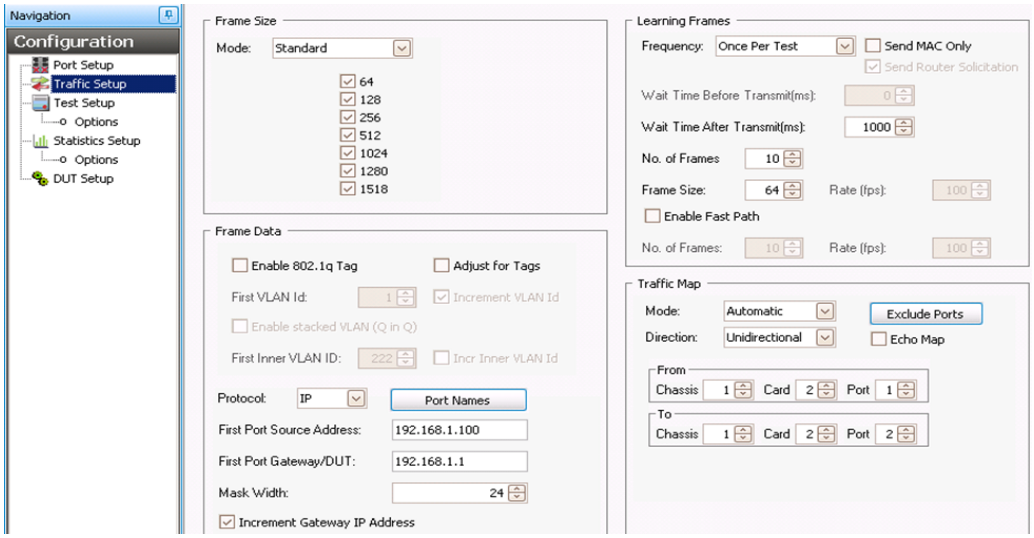
**CPU Effectiveness**

**CPU effectiveness is over 46% higher with offload enabled compared to offload enabled when running the same workload.**

# Small Packet Test Set-Up (cont.)

## Setup Configuration

Iteration was set to Linear with Loss Tolerance of '0%', Resolution was 1% and Initial Rate was '100%'. For Test parameters, No. of trials was set to '1', Addresses per port was set to '32' and Traffic Profile was set to 'Constant Loading'. The test duration was set to 30 seconds

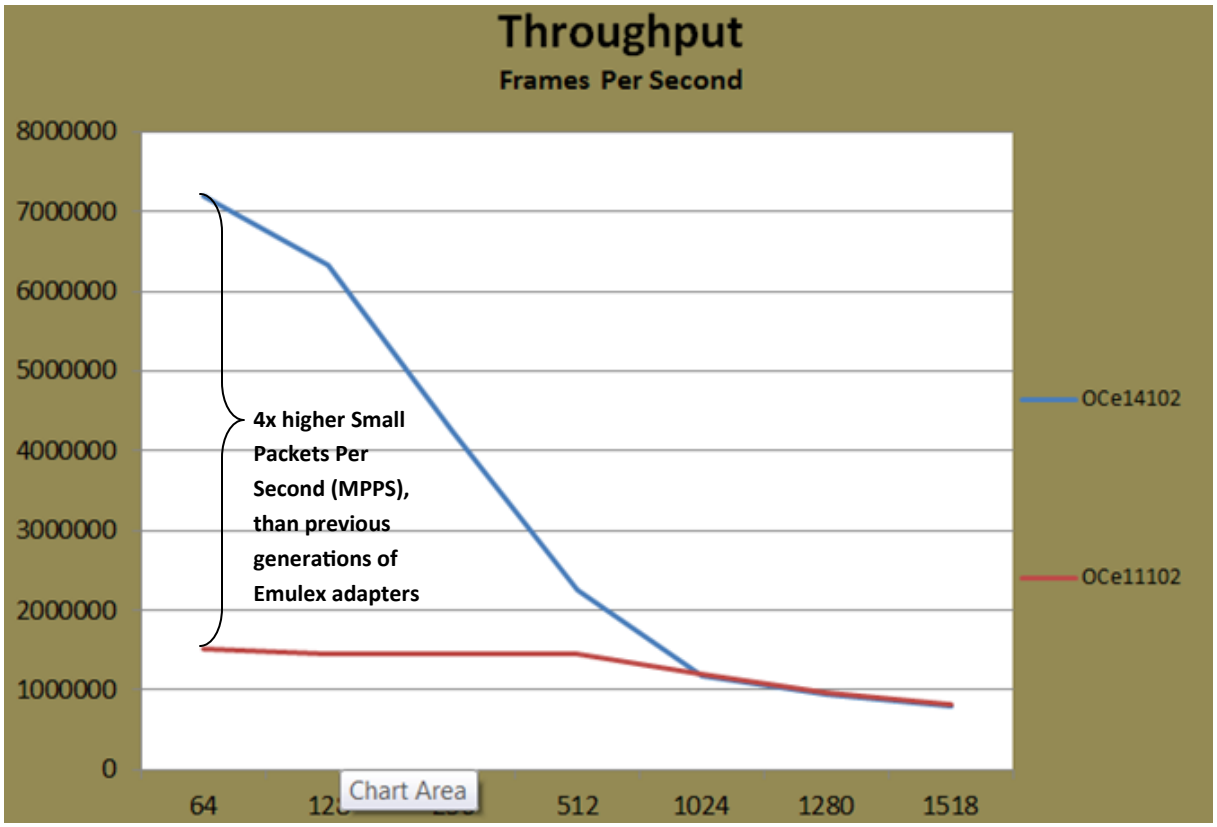

**Power Efficiency**

**The server power efficiency of the OCe14102 with offload enabled is as much as 122% better than the OCe11102 with no offload capability.**

# Small Packet Test Results

## Throughput in Packets Per Second

The small packet test was run for an MTU of 1500 for the Emulex OCe14102 and OCe11102 adapters. The results show that the packet per second (pps) performance for OCe14102 was 4 times better than OCe11102.




**PPS**

**The packet per second (pps) performance for OCe14102 was 4 times better than OCe11102.**

# iSCSI Block Storage Test Set-Up

## IOPS Performance and CPU Efficiency

Block storage testing was performed to demonstrate the IOPS performance and CPU utilization of the Emulex OCe14102 iSCSI CNA compared to the Emulex OCe11102 iSCSI CNA.

## Windows Server Environment

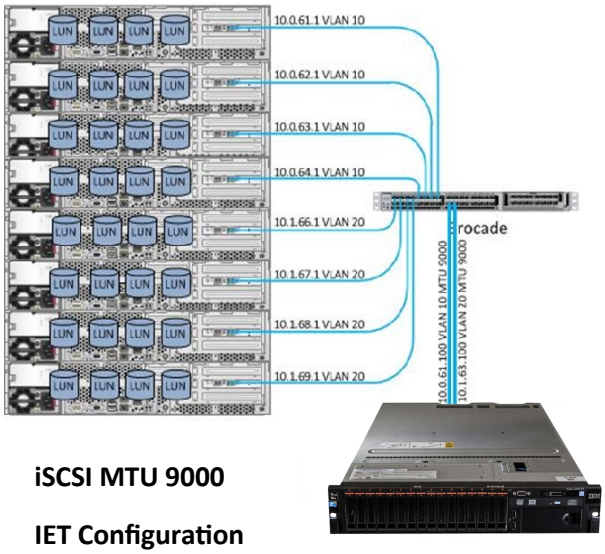
The IOPS test bed consisted of one dual processor x86 server running Windows 2012 R2 Enterprise Server. The server had one Emulex OCe14102 iSCSI dual-port adapter and one OCe11102 iSCSI adapter installed. A Brocade 8000 24 port 10Gb Ethernet switch was used to attach eight IET Targets as shown below in the architecture layout. The IET Targets had two 10GbE dual port Ethernet adapters running Red Hat Linux 6.2. Each target has 4 LUNs deployed at a 16MB size. For performance and bandwidth testing we used 4 LUNs per port and used 8 ports from each IET target.

## Workload

Multiple workers were set-up to compare the performance of a OCe14102 iSCSI CNA and OCe11102 iSCSI CNA. The IOmeter workers drove 4 target LUNs per target port with 32 outstanding I/Os. Sequential write I/Os were used to compare the throughput, IOPS and CPU utilization for different block sizes.

The same test configuration was used for the Emulex OCe11102 iSCSI adapter, but with 8 outstanding I/Os which resulted in better overall performance.

## Block Storage Test Topology




**iSCSI MTU 9000**

**IET Configuration**

**iSCSI Host**

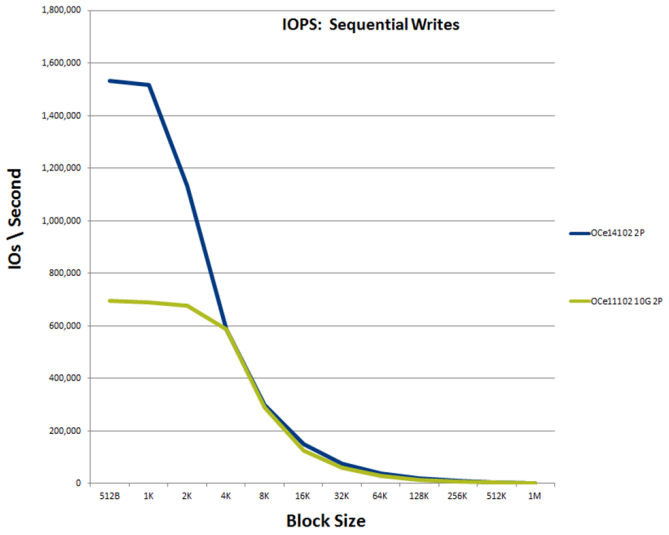
## Power Efficiency

**Server power efficiency measures the throughput achieved by the adapter in Mbps, per watt of server power consumption.**

## Sequential Write IOPS

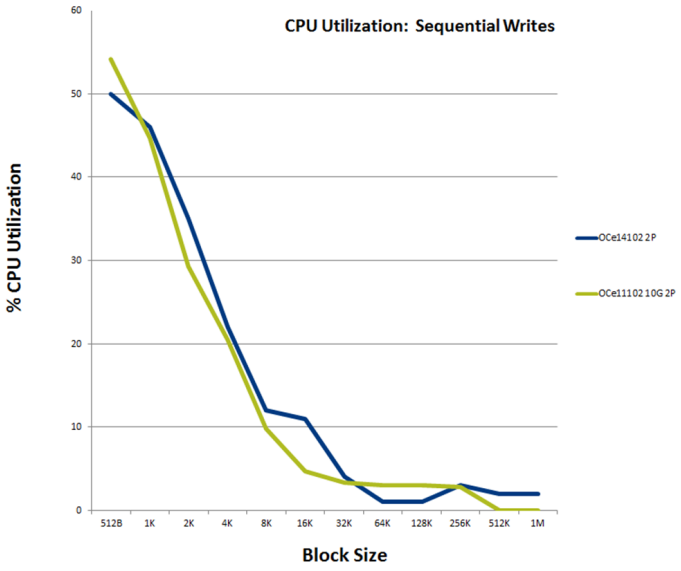
The following is the comparison of Emulex OCe14102 iSCSI to OCE11102 iSCSI for sequential writes I/Os using 2 ports.

The chart below shows in iSCSI testing using 512B and 1K block sizes, the Emulex OCe14102 achieved 1.5 Million IOPS, which is 2x better than the performance of the Emulex OCe11102.




## CPU Utilization

The chart below shows the CPU utilization of the OCe14102 was lower than OCe11102, even while the OCe14102 is showing 2x IOPS performance. Using 8K block sizes typically found in workloads such as MS Exchange, the OCe14102 utilizes 25% less CPU.



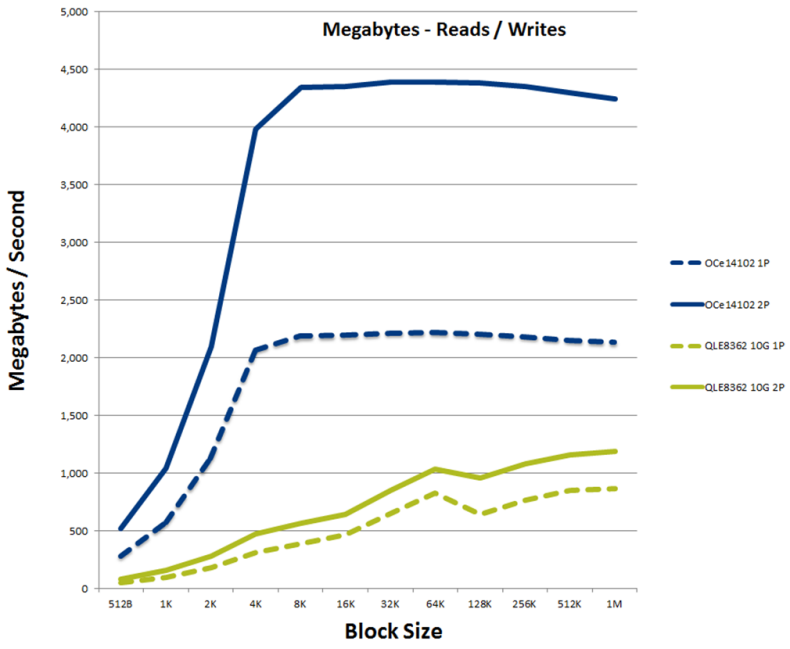

**Write Performance**

**The OCe14102 achieved 1.5 Million IOPS, which is 2x better than the performance of the Emulex OCe11102.**

# Comparing Emulex & QLogic Performance

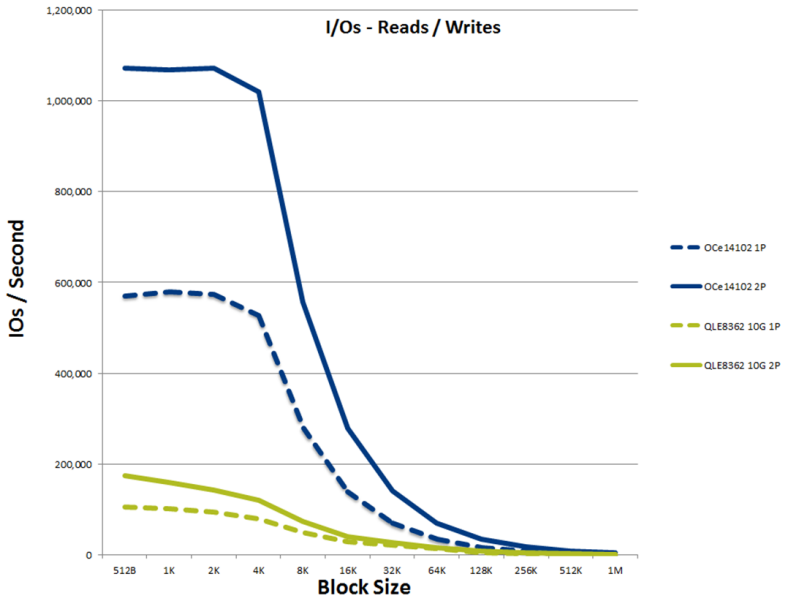
## Emulex OCe14102 vs QLogic QLE8362 Bandwidth (MB/s) with Windows 2012 R2

The following chart compares the performance in MB/s of the Emulex OCe14102 iSCSI and QLogic QLE8362 iSCSI for mixed reads and writes. The chart shows that in iSCSI testing with single and dual ports, the Emulex OCe14102 achieves 2 to 3x higher throughput at 4k and 8K block sizes.




## Emulex OCe14102 vs QLogic QLE8362 Transaction Performance (IOPS) with Windows

The following chart compares the performance in IOPS of the Emulex OCe14102 iSCSI and QLogic QLE8362 iSCSI for mixed reads and writes. The chart shows that once again the Emulex OCe14102 achieves 2 to 3x higher throughput at 4k and 8K block sizes.



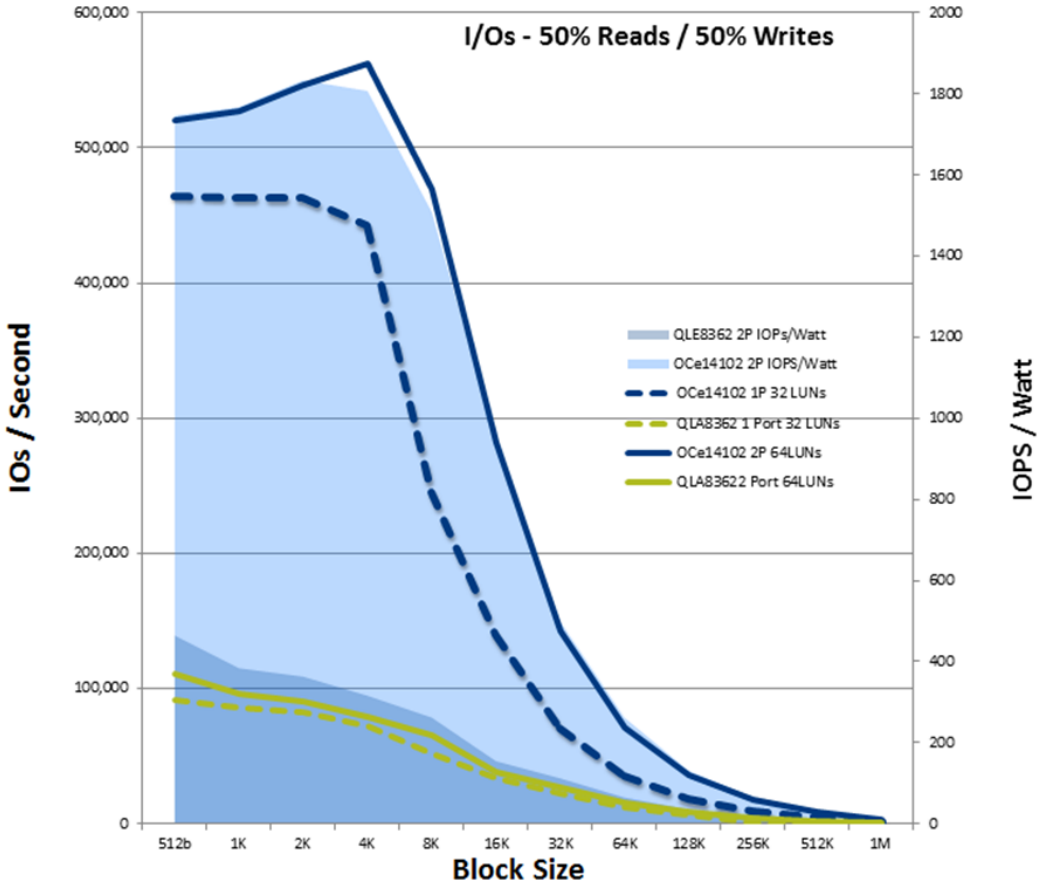

## Emulex Advantage

**In testing measuring bandwidth (MB/s) and transaction (IOPs) performance, the Emulex OCe14102 achieves 2 to 3x higher throughput at 4k and 8K block sizes.**

# Comparing Emulex & QLogic Power Efficiency

## Emulex OCe14102 vs QLogic QLE8362 Power Efficiency with Linux

The following chart compares the power efficiency (IOPS/Watt) of the Emulex OCe14102 iSCSI and QLogic QLE8362 for mixed iSCSI reads and writes. The chart shows that in testing with single and dual ports, the Emulex OCe14102 delivers up to 4x better performance & superior performance/watt.



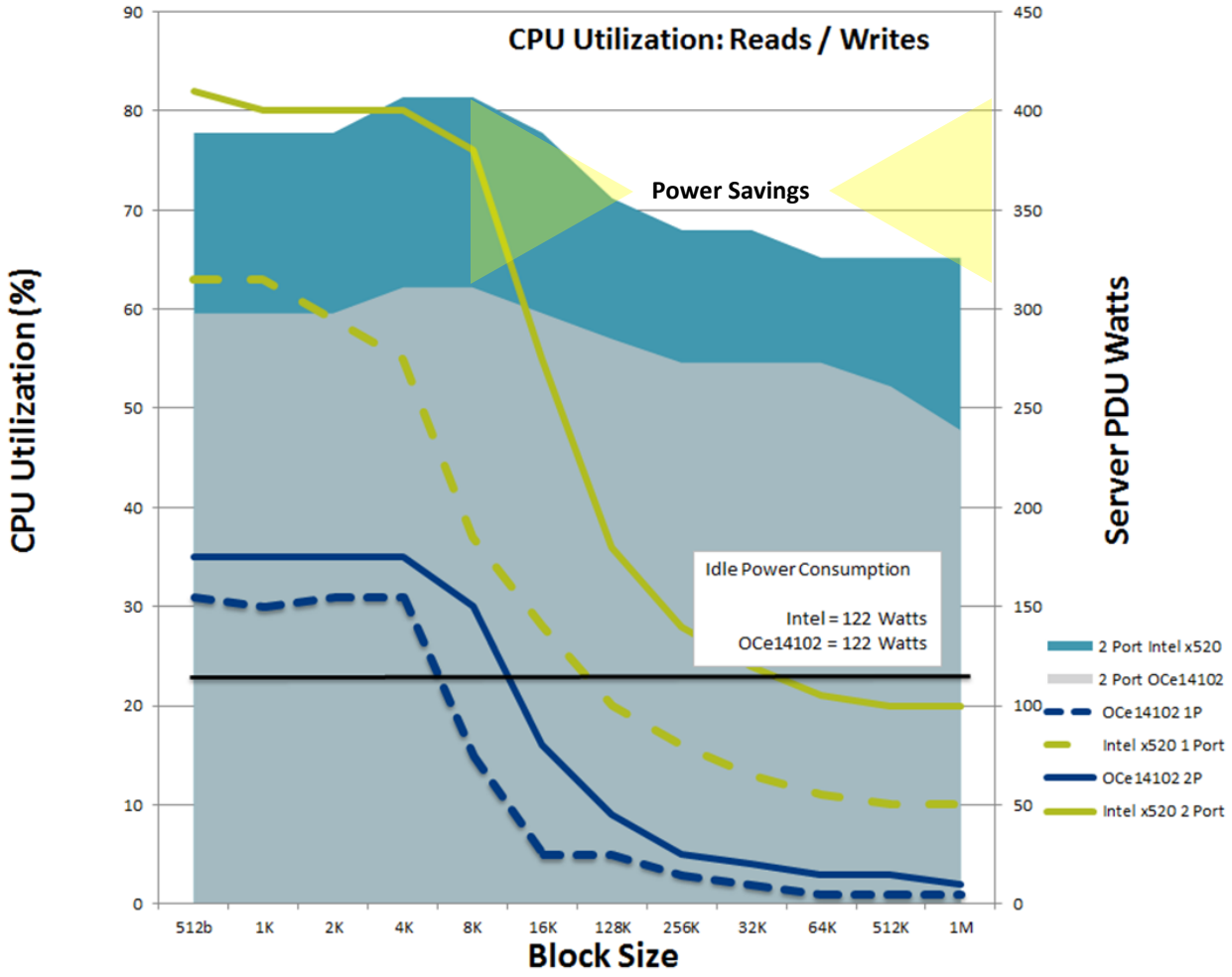

**Emulex Advantage**

**In testing measuring bandwidth (MB/s) and transaction (IOPs) performance, the Emulex OCe14102 delivers up to 4x better performance & superior performance/watt.**

# Comparing Emulex & Intel Power Savings

## Emulex OCe14102 vs Intel x520 CPU Utilization with Linux

The following chart compares the CPU utilization (%) and power consumption (Watts) of the Emulex OCe14102 and Intel x520 for mixed reads and writes. The chart demonstrates the Emulex OCe14102, with iSCSI offload to lower CPU usage, results in server power savings ranging from 50 Watts to almost 100 watts vs Intel, while still delivering full throughput.




**Emulex Advantage**

**The server power savings due to lower CPU usage ranges from 50 Watts to almost 100 watts vs Intel, while still delivering full throughput.**

# Summary

## Summary

The Emulex OCe14102 adapter performed dramatically better in the VXLAN, packet rate, and iSCSI SAN testing.

Key to overlay networking protocol performance improvements are hardware offloads for Virtual eXtensible Local Area Networking (VXLAN) supported by VMware and Network Virtualization using Generic Routing Protocol (NVGRE) supported by Microsoft. For both these protocols to deliver full value, CPU overhead and networking performance degradation are key issues that have to be addressed in a network adapter. The Emulex OCe14000 family of adapters utilizes their next generation XE100 controller with an RDMA architecture for superior packet throughput.

Higher performance for VXLAN and NVGRE overlay networking protocols is accomplished by offloading the overlay network processing to the XE100 controller ASIC, without breaking existing stateless TCP offloads that contribute to CPU efficient I/O processing, and have existed in adapters for over 10 years.

## The Bottom Line

The Emulex OCe14000 adapters with XE100 series controllers is a key component of a new high performance network, *and* can increase I/O to servers in existing networks. Along the way, the OCe14000 adapters delivers multiple system-level benefits including:

1. Up to 122% Increased I/O throughput with offload enabled through more efficient utilization of server CPU and 143% higher than with OCe11102 running the same workload
2. Increase in CPU effectiveness, (measured in Gbps / % of CPU utilized) by up to 46% with offload enabled and up to 42% higher than with OCe11102
3. Increase in host power efficiency, (measured in Mbps/watt) by 122% with offload enabled compared to OCE11102
4. Up to a 4x increase in small packet processing
5. Up to 2x improvement in IOP performance
6. Server power savings ranging from 50 Watts to almost 100 watts vs Intel, while still delivering full throughput.

## About the Author

Joe Kimpler is a senior analyst responsible for IT Brand Pulse Labs. Joe's team manages the delivery of technical services including hands-on testing, product reviews, total cost of ownership studies and product launch collateral. He has over 30 years of experience in information technology and has held senior engineering and marketing positions at Fujitsu, Rockwell Semiconductors, Quantum and QLogic. Joe holds an engineering degree from the University of Illinois and a MBA in marketing.