Open-E JovianDSS

Intel® E810-CQDA2 for OCP3.0 100GbE NIC Certification Report

Release date: 2021.12.29

Table of Contents

1. Introduction

The aim of this report is to present the methodology and results of the certification process conducted on the Intel E810-CQDA2 100 Gbps Network Interface Card when used with Open-E JovianDSS software. Functional testing was carried out for both the Single node and High Availability storage cluster configurations together with performance measurements to ensure full compatibility.

2. Device Under Test description

When performing the certification process, an Intel E810-CQDA2 for OCP3.0 Network Adapter was used. A detailed description of the adapter can be found in Table 1.

*Table 1. Intel E810-*CQDA2 *for OCP3.0 Network Adapter specifications.*

Product name	Intel Ethernet Network Adapter E810-CQDA2 for OCP 3.0
Data Rate Per Port	100/50/25/10GbE
Port Configuration	Dual
Interface	PCIe 4.0 (16 GT/s)
Intel® Data Direct I/O Technology	Yes
RDMA	Yes
Intelligent Offloads	Yes

3. Test environment description

Hardware specifications of the environments used during the certification process are included in Table 2. This applies to the Single node configuration as well as both the HA storage cluster nodes.

Table 2. Hardware specification of test environment.

System name	Intel® Server System M50CYP2UR208
Motherboard	Intel M50CYP2SBSTD
CPU	2x Intel Xeon Gold 6334 3.60GHz
RAM	6x 16 GB 3200 MHz DDR4
NIC	Intel E810-CQDA2 for OCP3.0
Storage devices	6x Intel DC P4510 2 TB (12x for Single node)
System	Open-E JovianDSS up29r1 Arch Linux-2021.11.01 (client side)

Performance measurements were done using fio for linux, v3.28.

4. Functional tests

To ensure the proper operation of the tested device when used with Open-E JovianDSS software, functional testing was done for both the Single node and HA storage cluster configurations. The performed tests, along with their results, are described in Tables 3 and 4 respectively.

Table 3. Single node functional tests.

Tested functionality	Result
Hardware detection and presentation on the UI	passed
Network configuration	passed
Link state reporting	passed
Bonding	passed
System stability under load over extended period of time	passed
Network statistics	passed
RDMA (RoCE/iWARP)	not supported[1]

Table 4. HA storage cluster functional tests.

Tested functionality	Result
Manual Failover	passed
Automatic Failover triggered after network failure	passed
Automatic Failover triggered after system shutdown	passed
Automatic Failover triggered after system reboot	passed
Automatic Failover triggered after system power-off	passed
Automatic Failover triggered after I/O failure	passed
System stability under load over extended period of time	passed
Cluster network configuration	passed

[1] RDMA support for Intel network adapters will be added in future releases.

5. Single node performance test

The following test was intended to ensure that the Intel E810 network interface card is fully utilized when used as the connection to the client machine.

5.1. Test description

Open-E JovianDSS was configured as a Single node using the storage parameters described in Table 5. The Intel E810 network adapter was used as a single port connection on both the storage server and client machine. Fio was run on the client side, as described in Table 6, for every test case shown in Table 7.

Table 5. Storage configuration for Single node test.

zpool data groups	12x mirrored Intel DC P4510 2 TB
volblocksize	4 kB
ZFS ARC size	200 GB
zvol sync	disabled
zvol compression	none
zvol provisioning	thick
connection	100 Gbps single port connection

Table 6. Fio parameters used for Single node test.

version	3.28
ioengine	libiscsi
run time	90 s
ramp time	10 s
test size	10 GB for every thread
direct	yes
threads count	1, 4, 8, 16
queue depth	1, 16, 64, 128

Table 7. Test profiles description for Single node performance test.

Test profiles	IO pattern	Read to write %	Block size
Sequential read	sequential	100/0	1 MB
Sequential write	sequential	0/100	1 MB

5.2. Performance results

Figures 1 and 2 show sequential read and write performance results. For better visualization, the charts' vertical axis were scaled to the theoretical maximum bandwidth available for the tested device.

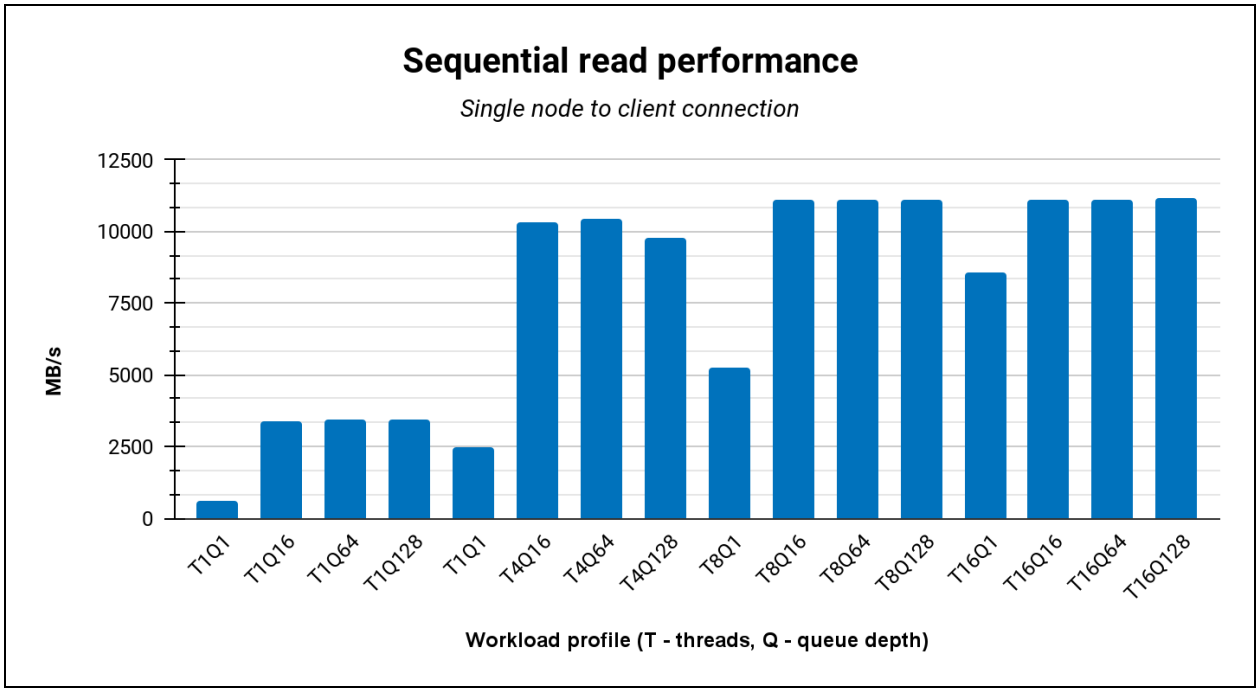

Fig. 1. Sequential read performance results of a Single node test.

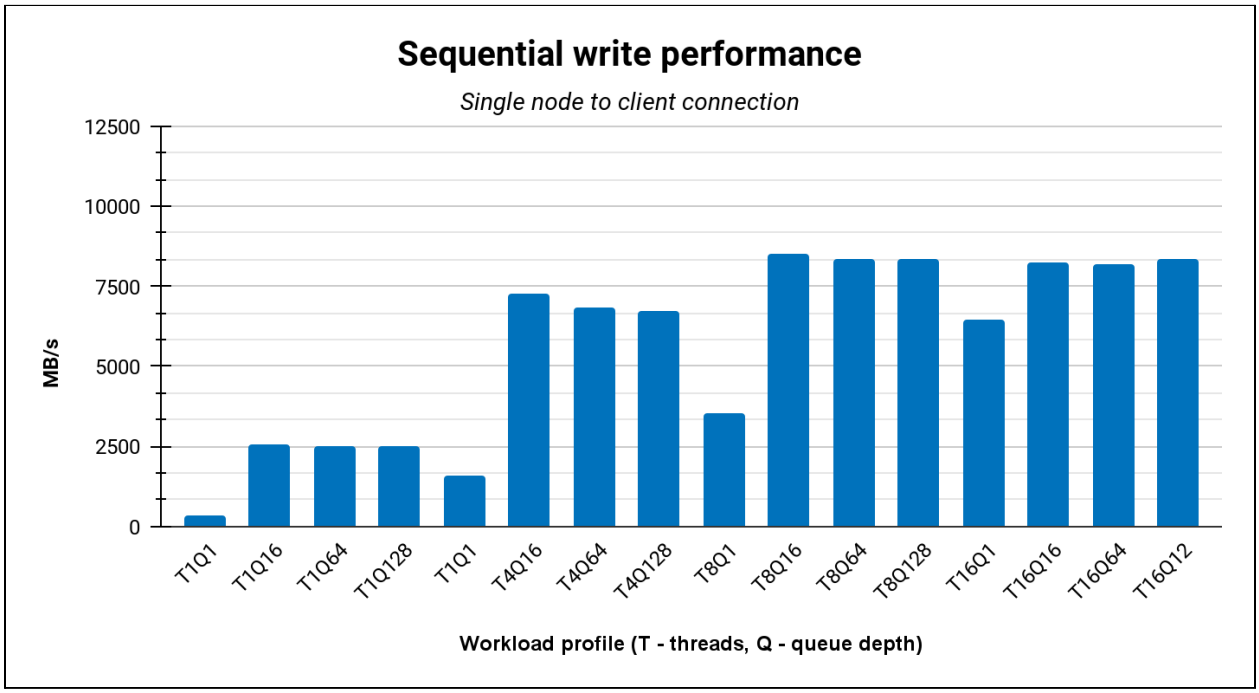

Fig. 2. Sequential write performance results of a Single node test.

5.3. Test conclusions

When running on a client machine, fio was able to generate a sequential read workload with a throughput of around 11 GB/s. Taking into account the overhead, this value is very close to the theoretical maximum data transfer rate of the tested device. As for the sequential write, a maximum throughput of only around 8 GB/s was achieved, due to the storage limitations.
The aforementioned results suggest that the tested network adapter is fully utilized by Open-E JovianDSS only when used to access storage from the client side.

6. HA cluster mirroring path performance test

The mirroring path is a crucial element of every High Availability non-shared storage cluster, as it is responsible for data redundancy, which allows synchronization between the paired cluster nodes' disks. In modern all-flash-based storage architectures, where high bandwidths are involved, it's important to take full advantage of the network adapter capabilities. To ensure that the declared NIC speed is achieved in Open-E JovianDSS mirroring path, the following test was conducted.

6.1. Test description

The Open-E JovianDSS non-shared storage cluster was configured using two nodes, described in Table 4. The tested Intel E810 network adapter was used for a mirroring path with a single port connection. Configured storage is shown in Table 8. Fio was run locally on one of the nodes with the parameters shown in Table 9. As a result of the storage being mirrored over the network, its overall throughput was bound to the performance of the tested network adapter. The applied test cases are described in Table 10.
Maximum ZFS ARC size was decreased to 10 GB so as to increase the number of IO requests being served by reading the storage disks synchronized via mirroring path and not the data cached in RAM, when read workload was applied. Otherwise almost all data would be read from RAM, leaving the mirroring path unutilized.

Table 8. Storage configuration for non-shared storage cluster test.

zpool data groups	6x mirrored Intel DC P4510 2 TB
volblocksize	4 kB
ZFS ARC size	10 GB
zvol sync	always
zvol compression	none
zvol provisioning	thick
mirroring path	100 Gbps single port connection

Table 9. Fio parameters used for non-shared storage cluster tests.

version	3.28
ioengine	libaio
run time	90 s
ramp time	10 s
test size	10 GB for every thread
direct	yes
threads count	1, 4, 8, 16
queue depth	1, 16, 64, 128

Table 10. Test profiles description for Cluster over Ethernet performance test.

Test profile	IO pattern	Read to write %	Block size
Sequential read	sequential	100/0	1 MB
Sequential write	sequential	0/100	1 MB

6.2. Performance results

Figures 3 and 4 show sequential read and write performance results. For better visualization, the charts' vertical axis were scaled to the theoretical maximum bandwidth available for the tested device.

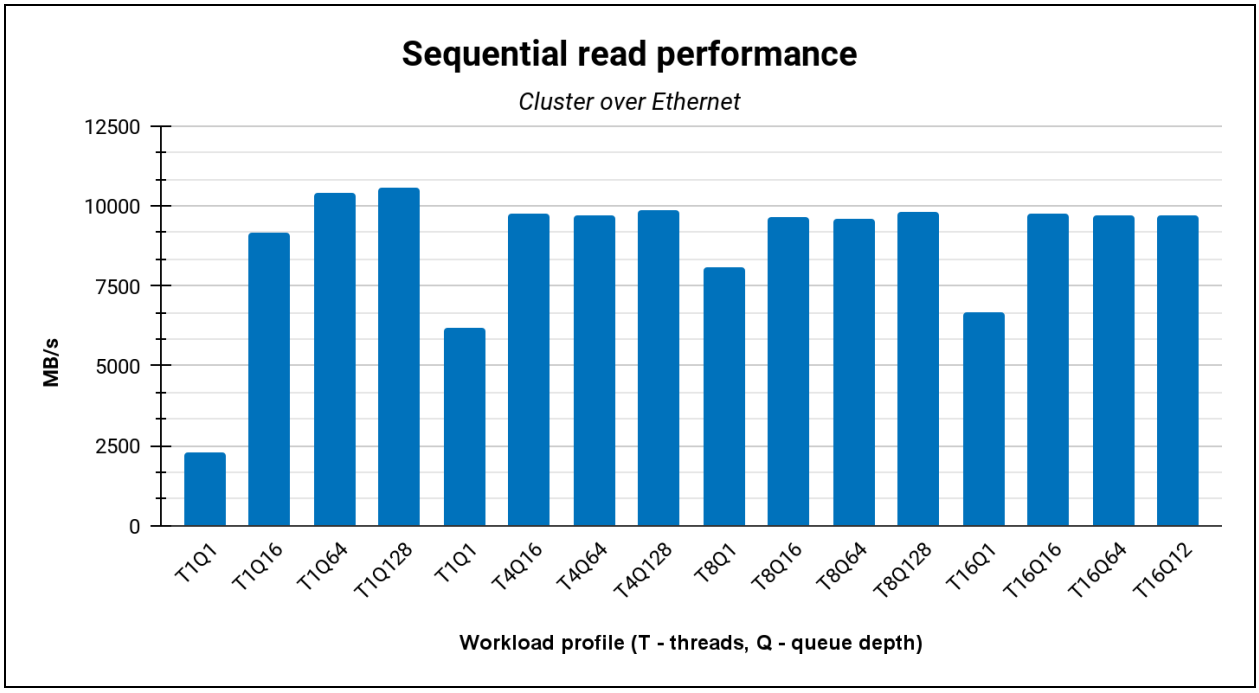

Fig. 3. Sequential read performance results of a Cluster over Ethernet test.

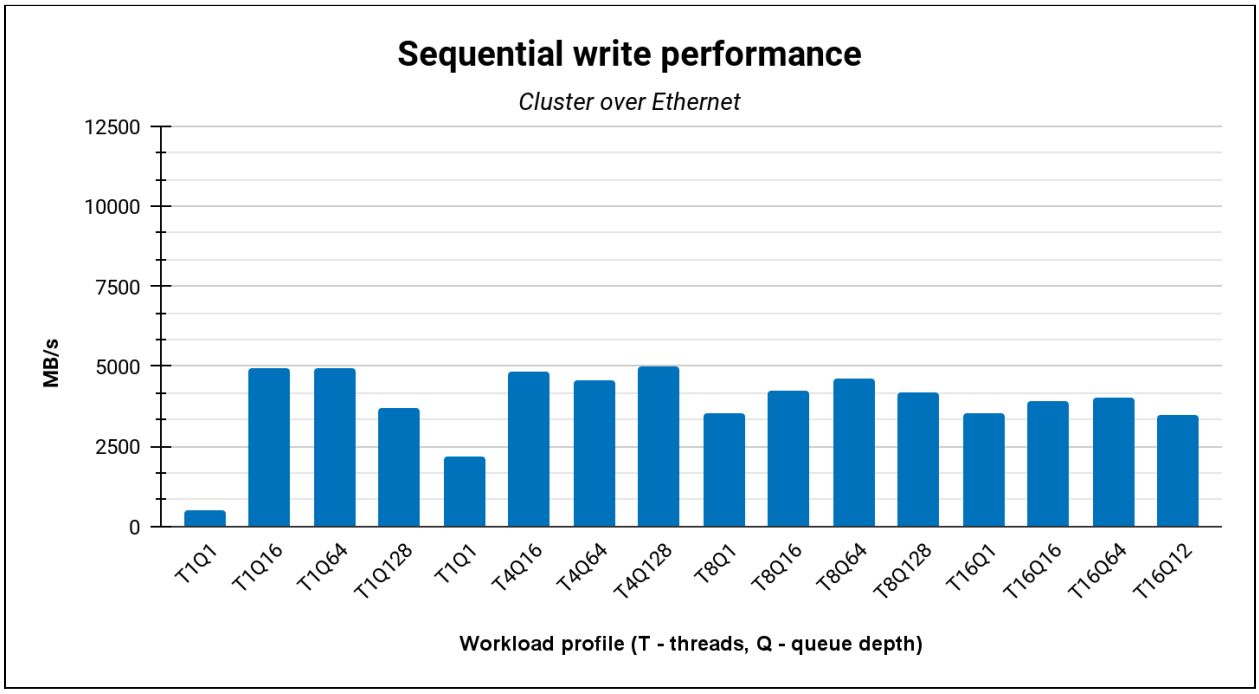

Fig. 4. Sequential write performance results of a Cluster over Ethernet test.

6.3. Test conclusions

The locally run fio was able to generate a sequential write workload with a throughput of around 5 GB/s, which was close to the limit of the underlying storage. When a sequential read was performed, a maximum of 10 GB/s was achieved. After accounting for the overhead, this is close to the maximum theoretical data transfer rate of the tested device. This confirms that the tested network adapter is fully utilized when used for mirroring path in an Open-E JovianDSS non-shared storage cluster.

7. Summary

The Intel E810-CQDA2 network adapter was comprehensively tested for full functional compatibility with Open-E JovianDSS. Performance characteristics were also tested in several use cases. Both Single node and HA cluster operations were taken into consideration. The tests were designed to find any abnormalities in the tested device. Given the results achieved in testing, the examined device can now safely be added to the Hardware Certification List and granted “Certified by Open-E” status.