



OpenFabrics Alliance

Interoperability Working Group (OFA-IWG)

January 2010 Logo Validation Event Report

UNH-IOL – 121 Technology Drive, Suite 2 – Durham, NH 03824 – +1-603-862-0090
 OFILG – ofalab@iol.unh.edu

Chris Calandro
 Chelsio Communications
 370 San Aleso Avenue #100
 Sunnyvale, CA 94085

April 28, 2010
 Report Revision 1.1
 OFED Version 1.5

Enclosed are the results from OFA interoperability testing performed on the following devices under test (DUTs):

Chelsio Communications S310E-CXA Adapter

The test suite used for the tests described in this report is *OFA-IWG Interoperability Test Plan Release 1.30* (September 8, 2009 DRAFT). It is available at the OFA website :

iol.unh.edu/ofatestplan

Log files are available here:

[ftp.iol.unh.edu/pub/ofa](ftp://ftp.iol.unh.edu/pub/ofa)

The OFILG gratefully acknowledges the **iWARP Consortium** for developing and contributing the testing software used for test 11.5 iWARP Connectivity.

Results Overview - Mandatory Tests

Test Procedure	IWG Test Status	Result/Notes
11.5 iWARP Connectivity	Mandatory	PASS with Comments
12.12 Open MPI - homogeneous	Mandatory	Conditional PASS
12.13 MVAPICH2 - homogeneous	Mandatory	PASS

Testing Completed 2010-Apr-28

Jeff Laird
laird@iol.unh.edu

Review Completed 2010-Apr-28
 Mikkell Hagen

mhagen@iol.unh.edu

Results Overview - All Tests

Note: **Failing a beta test does not necessarily indicate a failure of the device under test.** There are several reasons that a device might fail a beta test that are beyond the control of the device manufacturer. Such reasons include unresolved problems with the OFED software, with other vendors' devices, or with the test plan procedures.

Test Procedure	IWG Test Status	Result/Notes
11.1 Ethernet Link Initialization	Beta	Warning
11.2 Ethernet Fabric Initialization	Beta	Not Available
11.3 Ethernet Fabric Reconvergence	Beta	Not Available
11.4 Ethernet Fabric Failover	Beta	Not Available
11.5 iWARP Connectivity	Mandatory	PASS with Comments
12.1 TI iSER	Beta	Not Available
12.2 TI NFS over RDMA	Beta	Not Tested
12.3 TI RDS	Beta	FAIL
12.4 TI SDP	N/A	Not Applicable
12.5 TI uDAPL	Beta	FAIL
12.6-7 TI Basic RDMA Interop	Beta	Not Tested
12.8-9 TI RDMA Operations	Beta	Not Tested
12.10 TI MPI - HP	Beta	Not Tested
12.11 TI MPI - Intel	Beta	Not Tested
12.12 TI MPI - Open MPI - homogeneous	Mandatory	Conditional PASS
12.12 TI MPI - Open MPI - heterogeneous	Beta	FAIL
12.13 TI MPI - MVAPICH2 - homogeneous	Mandatory	PASS
12.13 TI MPI - MVAPICH2 - heterogeneous	Beta	FAIL

For specific details regarding issues please see the corresponding test result.

Digital Signature Information

This document was created using an Adobe digital signature. A digital signature helps to ensure the authenticity of the document, but only in this digital format. For information on how to verify this document's integrity proceed to the following site:

http://www.iol.unh.edu/certifyDoc/certificates_and_fingerprints.php

If the document status still indicates "Validity of author NOT confirmed", then please contact the UNH-IOL to confirm the document's authenticity. To further validate the certificate integrity, Adobe 6.0 should report the following fingerprint information:

MD5 Fingerprint: EEE1 7A82 7806 EB21 AF94 F189 E4BE 361B

SHA-1 Fingerprint: ECFB 7FAF AB4A 0832 2408 E965 9F5C E3F2 D784 AAAB

Report Revision History

1.0 Initial release

1.1 Update for "Conditional PASS" in test 12.12.

Table 1: Result Key

The following table contains possible results and their meanings:

Result	Description
PASS	The Device Under Test (DUT) was observed to exhibit conformant behavior.
PASS with Comments	The DUT was observed to exhibit conformant behavior however an additional explanation of the situation is included, such as due to time limitations only a portion of the testing was performed.
Conditional PASS	The DUT was observed to exhibit conformant behavior under particular conditions.
FAIL	The DUT was observed to exhibit non-conformant behavior.
Warning	The DUT was observed to exhibit behavior that is not recommended.
Informative	Results are for informative purposes only and are not judged on a pass or fail basis.
Refer to Comments	From the observations, a valid pass or fail could not be determined. An additional explanation of the situation is included.
Not Applicable	The DUT does not support the technology required to perform this test.
Not Available	Due to testing station limitations or time limitations, the tests could not be performed.
Borderline	The observed values of the specified parameters are valid at one extreme and invalid at the other.
Not Tested	Not tested due to the time constraints of the test period.

Table 2: DUT Information

DUT #1 Details			
Manufacturer	Chelsio Communications	Firmware Version	T 7.4.0 TP 1.1.0
Model	S310E-CXA	Hardware Version	T3C
Driver	cxgb3 version 1.1.2-ofed	Located in Hosts	Anthe (8 cores), Fenrir (8 cores), Calypso (4 cores), Hyperion (4 cores), Mimas (4 cores)
Additional Comments/Notes			

Test Setup:



Device Test Summary Results:

Test Number and Name	Parts	Summary	Result
11.1 Ethernet Link Initialization		error message	Warning
Details			
rping successfully transmits data across the RDMA channel. However, an error message appears. Executing rping from Chelsio as client to Intel as server displays "cq completion failed status 5" on the Intel/server side. Executing rping from Intel as client to Chelsio as server displays "cq completion failed status 5" on the Intel/client side.			

Test Number and Name	Parts	Summary	Result
11.2 Ethernet Fabric Initialization		No switches to test	Not Available
Details			

Test Number and Name	Parts	Summary	Result
11.3 Ethernet Fabric Reconvergence		No switches to test	Not Available
Details			

Test Number and Name	Parts	Summary	Result
11.4 Ethernet Fabric Failover		No switches to test	Not Available
Details			

Test Number and Name	Parts	Summary	Result
11.5 iWARP Connectivity	Groups 1-7	unsupported features	PASS with Comments
Details			
Comments from the separate iWARP Connectivity report pertaining to several subtests within the iWARP Connectivity test:			
1.The OFED software stack does not support SendINV and SendSEINV.			
2.CX-4 and 10G optics do not support speed negotiation.			
3.The OFED software stack does not support the ability to configure Markers or CRC at runtime.			



Test Number and Name	Parts	Summary	Result
12.1 TI iSER		No iSER target	Not Available
Details			
There were no iSER targets in the test cluster.			

Test Number and Name	Parts	Summary	Result
12.2 TI NFS over RDMA		Insufficient time	Not Tested
Details			

Test Number and Name	Parts	Summary	Result
12.3 TI RDS	rds-ping	no output	FAIL
	rds-stress	no output	FAIL
Details			
<p>For this test the Chelsio parameter file /sys/module/iw_cxgb3/parameters/peer2peer contained "1" and the Intel parameter file /sys/module/iw_nes/parameters/send_first contained "0" on all hosts.</p> <p>rds-ping: Homogeneous passes. Heterogeneous produces no output but ends after the appropriate amount of time (3 seconds for a ping count of 3).</p> <p>rds-stress : Homogeneous passes. Heterogeneous behaves as if passing, but output values are all zero.</p>			

Test Number and Name	Parts	Summary	Result
12.4 TI SDP		no SDP support	Not Applicable
Details			
Current licensing restrictions prevent SDP support in iWARP RNIC solutions.			

Test Number and Name	Parts	Summary	Result
12.5 TI uDAPL		Hangs	FAIL
Details			
<p>For this test the Chelsio parameter file /sys/module/iw_cxgb3/parameters/peer2peer contained "1" and the Intel parameter file /sys/module/iw_nes/parameters/send_first contained "0" on all hosts.</p> <p>Homogeneous passes. Heterogeneous Chelsio (as client) to Intel (as server) hangs on test 3.2. Heterogeneous Intel (as client) to Chelsio (as server) hangs on test 1.1.</p>			

Test Number and Name	Parts	Summary	Result
12.6-7 TI Basic RDMA Interop	bw, lat		Not Tested
Details			

Test Number and Name	Parts	Summary	Result
12.8-9 TI RDMA Operations			Not Tested
Details			

Test Number and Name	Parts	Summary	Result
12.10 TI MPI - HP		Insufficient time	Not Tested
Details			

Test Number and Name	Parts	Summary	Result
12.11 TI MPI - Intel		Insufficient time	Not Tested
Details			

Test Number and Name	Parts	Summary	Result
12.12 TI MPI - Open MPI - homogeneous		Requires OFED patch	Conditional PASS
Details			
<p>MPI tests use all five hosts/RNICs and all 28 cores on those five hosts.</p> <p>During the test event:</p> <p>When executed normally this test bogged to a standstill during the Gather subtest. Adding the parameter "-mca mpi_leave_pinned 0" delayed the bogging until the end of the Bcast subtest. Reducing the core count to 27, 26, and 25 did not solve the problem, but reducing to 24 did. With 24 cores (one less than the maximum on four different hosts) and the "pinned" parameter set as above the test passed in 47 minutes.</p> <p>Slowdowns are sudden. For example, in the Alltoallv subtest with 16 processes, between the 4 kiB transmission and the 8 kiB transmission the times jump from 280 microseconds to 220,000 microseconds.</p> <p>After the test event:</p> <p>A firmware update, from the test-event version "T 7.4.0 TP 1.1.0" as recorded above on page 4 to "T 7.10.0 TP 1.1.0" corrects the performance problem mentioned in the previous paragraph. This change is for performance only and so is not required as a condition of passing.</p> <p>The problem described before that—bogging to a standstill with 25 or more cores—is corrected by an OFED patch. The need for this corrective action in the OFED software is the reason for the "Conditional PASS" test result. The OFED patch with installation instructions is available at ftp://iol.unh.edu/pub/ofa/Chelsio_patch/.</p> <p>Testing with OFED 1.5.1 succeeded without the patch.</p>			

Test Number and Name	Parts	Summary	Result
12.12 TI MPI - Open MPI - heterogeneous		Hangs at start.	FAIL
Details			
The test hangs with no output.			

Test Number and Name	Parts	Summary	Result
12.13 TI MPI - MVAPICH2 - homogeneous			PASS
Details			

Test Number and Name	Parts	Summary	Result
12.13 TI MPI - MVAPICH2 - heterogeneous		Unable to execute.	FAIL
Details			
<p>Homogeneous Intel requires both of the following command-line parameters for the IMB-MPI1 test to start.</p> <pre>MV2_VBUF_TOTAL_SIZE=9216 MV2_DEFAULT_MAX_CQ_SIZE=32766</pre> <p>Homogeneous Chelsio will run the IMB-MPI1 test with no parameters and with the former (VBUF) parameter, but setting the latter (CQ_SIZE) parameter causes Chelsio to not start. (This is the case not only with value 32,766 ($2^{15} - 2$) but also with values 16,382 ($2^{14} - 2$), 65,534 ($2^{16} - 2$), 131,070 ($2^{17} - 2$), and 262,142 ($2^{18} - 2$)).</p> <p>Therefore, for heterogeneous execution there is a question of whether or not to use the command-line parameters. As expected, all four possible combinations fail. When the CQ_SIZE parameter is present the failure is on the Chelsio host, and when it is not present the failure is on the Intel host.</p>			