LACP Configuration Guide

Using Intel® Ethernet 500 and 700 Series Network Adapters and Various Server Operating Systems

Technical Brief

Ethernet Networking Division (ND)

April 2018

Revision History

Revision	Date	Comments
1.0	April 27, 2018	Initial release (Intel public).

Contents

NOTE: This page intentionally left blank.

1.0   Introduction

Since its inception in 1973, Intel has been vital to the development of the Ethernet, and continues to be the Industry leader. For over 40 years, the Ethernet has been growing to accommodate increasing bandwidth needs for complex multimedia; streaming video, music, and voice data, for example. Beginning with 1 GbE, and expanding to 10 GbE, 25 GbE and 40 GbE, computing and storage resource needs continue to grow.

1.1   Intel® Ethernet Network Adapters

For the purposes of this technical brief, the following adapters were used:

- Intel® Ethernet 500 Series Network Adapters: X520, X540, and X550
- Intel® Ethernet 700 Series Network Adapters: X710, XXV710, and XL710

1.2   NIC Teaming with Link Aggregation Control Protocol

NIC Teaming is a crucial component of meeting today's "Always On" Data Center demands. Intel's Network I/O products support a variety of methods and standards that provide fault tolerance, load balancing, and bandwidth aggregation. This guide focuses on the configuration of NIC Teaming via the Link Aggregation Control Protocol (LACP) using Intel® Ethernet 500 and 700 Series Network Adapters.

LACP is a vendor-independent standard, which is defined in IEEE standard 802.1ax (previously 802.3ad). It uses active detection of LACP-enabled devices to form and maintain a single logical link comprised of 1 to 8 physical links. This logical link is known as a Link Aggregation Group (LAG).

The members of a LAG communicate among themselves using multicast packets called LACPDUs. This single logical link provides fault tolerance should a physical link fail, and also provides bandwidth aggregation. The LAG detects a failed physical link and moves conversations as necessary.

This guide explains how to configure Layer 2 LACP LAGs from servers running various operating systems (see Section 1.3) to various Cisco* switches.

Section 2.0 provides four sets of instructions for completing the LACP configuration: one for the switch configuration, and one for each of the classes of operating systems listed above. These configurations were tested using a representative sampling of the Intel® Ethernet 500 and 700 Series Network Adapters, as well as most of the speeds available on each of those adapters, as shown in Table 1 on page 18.

1.3   Server Software

The following server software was used in the testing:

- Microsoft* Windows Server 2012R2U1 and 2016.
- Red Hat Enterprise Linux* 7.4 (RHEL 7.4)
- VMware* vSphere 6.5 U1 (commonly known as VMware ESXi 6.5 U1)

1.4   Server Hardware Platform

The server hardware support for each NIC depends on the version of PCI Express. This ranges from PCI Express v2.0 5.0 GT/s x8 lanes to PCI Express v3.0 8.0 GT/s x 8 lanes. The server hardware used for testing these configurations was the Dell* R730. Refer to the Product Guide for Intel® Ethernet Network Adapters to select the right adapter for your server hardware.

1.5   Document Conventions

The following conventions are used in this document.

- The tilde '~' indicates an entry to be made on the command line.
- **Bold** text indicates any of the following:
  - Buttons to be clicked.
  - Categories to be selected.
  - Keys to be pressed on the keyboard.
  - Fields in a dialog box.
  - Command names.
- *Italic* text indicates any of the following:
  - Filenames, directory paths, or URLs.
  - Topics or values that have special significance.

2.0 Configuration Guides

2.1 Generic Outline of Configuration Steps

This section provides a high-level outline of the steps to configure LACP from a server to a switch.

On the switch(s):

1. Create a Layer 2 channel-group from a collection of physical ports

On the server:

1. Create a NIC Team from a collection of physical ports.

2. Configure the network settings on the resultant virtual device.

Assumption: These guidelines assume that there is a management network connection to the server.

2.2 Cisco Switch Configuration

The tested configuration steps are identical for all the Cisco switches listed in Table 1 on page 18.

Reference: The steps outlined here were derived from product documentation on *cisco.com*. For a complete list of limitations, options, and other details, refer to the product's documentation on the *cisco.com* website

2.2.1 LACP Feature Enablement

To enable the LACP feature on the Cisco switch, follows these steps:

1. Connect to the switch with **ssh** or **telnet** as appropriate.

2. Enable EXEC mode with the following command:

   ```
   ~ enable
   ```

3. Enter password when prompted.

4. The LACP feature must be globally enabled on the switch(s). To verify this, display the running configuration with the following command:

   ```
   ~ show run
   ```

5. Look for "feature lacp" in the output.

   If you do not see it, enable LACP with following commands:

   ```
   ~ config terminal
   ~ feature lacp
   ~ exit
   ```

2.2.2 Layer 2 Channel-Group Setup

To setup the Layer 2 channel-group on the Cisco switch, follows these steps:

1. Before you begin, make a note of the physical interfaces to be placed in the channel-group. To see a list of interfaces with a brief summary of their current configuration, display the running configuration with the following command:

   ```
   ~ show run
   ```

   Note: All of the interfaces to be added to the channel-group must be configured identically prior to adding them to the group.

2. Enter Configuration Mode with the following command:

   ```
   ~ config terminal
   ```

3. Select the first interface (e1/19) to add to the channel group.

   ```
   (config) ~ interface Ethernet1/19
   ```

   Note: In this example, channel-group #1 is created, and interfaces e1/19 and e1/20 are used.

4. Create the channel-group and add the selected interface with one command.

   For Windows and Linux:

   ```
   (config-if) ~ channel-group 1 mode active
   ```

   For VMware ESXi Host:

   ```
   (config-if) ~ channel-group 1 mode on
   ```

5. Repeat Step 3 and Step 4 to add the next interface (e1/20) to the channel-group.

6. Repeat Step 3 and Step 4 until you have added all interfaces intended for this channel-group.

7. Exit configuration mode with the following commands:

   ```
   (config-if) ~ exit
   (config) ~ exit
   ```

8. To verify the channel-port is created, display the running configuration with the following command:

   ```
   ~ show run
   ```

   The virtual interface that represents the channel-port should appear at the top of the interface list.

9. Display full details about the interface:

```
~ show interface port-channel1
```

```
port-channel1 is down (No operational members)
  Hardware: Port-Channel, address: 0000.0000.0000 (bia 0000.0000.0000)
  MTU 1500 bytes, BW 100000 Kbit, DLY 10 usec
  reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation ARPA
  Port mode is access
  auto-duplex, auto-speed
  Input flow-control is off, output flow-control is off
  Switchport monitor is off
  EtherType is 0x8100
  Members in this channel: Eth1/19, Eth1/20
  Last clearing of "show interface" counters never
  30 seconds input rate 0 bits/sec, 0 packets/sec
  30 seconds output rate 0 bits/sec, 0 packets/sec
  Load-Interval #2: 5 minute (300 seconds)
    input rate 0 bps, 0 pps; output rate 0 bps, 0 pps
  RX
    0 unicast packets  0 multicast packets  0 broadcast packets
    0 input packets  0 bytes
    0 jumbo packets  0 storm suppression packets
    0 runts  0 giants  0 CRC  0 no buffer
    0 input error  0 short frame  0 overrun  0 underrun  0 ignored
    0 watchdog  0 bad etype drop  0 bad proto drop  0 if down drop
    0 input with dribble  0 input discard
    0 Rx pause
  TX
    0 unicast packets  0 multicast packets  0 broadcast packets
    0 output packets  0 bytes
    0 jumbo packets
    0 output errors  0 collision  0 deferred  0 late collision
    0 lost carrier  0 no carrier  0 babble 0 output discard
    0 Tx pause
    0 interface resets
```

Note: The interface shows as down until the underlying physical interfaces are cabled to ports on the server that have been configured as a team set to use LACP.

2.3   Microsoft Windows NIC Teaming Configuration

The tested configuration steps are identical for both Windows Server 2012R2U1 and 2016. In both cases the *Server Manager* is used to complete the configuration.

Assumption: The following steps assume that a channel-group has already been configured on the switch(s) to which this NIC Team will be connected.

Note: All of the ports to be added to the NIC team must share the same configuration, speed, and duplex settings. Additionally, settings and capacities must match or be compatible with the interfaces of the switch to which this NIC Team will be connected.

Reference: The steps outlined here were derived from product documentation on *microsoft.com*. For a complete list of limitations, options, and other details, refer to the product's documentation on the *micrososft.com* website.

To configure NIC Teaming with Microsoft Windows, follow these steps:

1. Connect the NIC ports from the server to the ports on the switch that are members of the channel-port created in Section 2.2.2. The type of connectors and cables vary based on the environment.

2. Open the Server Manager.

3. Select **Local Server**.

4. Click the link to the right of **NIC Teaming**.
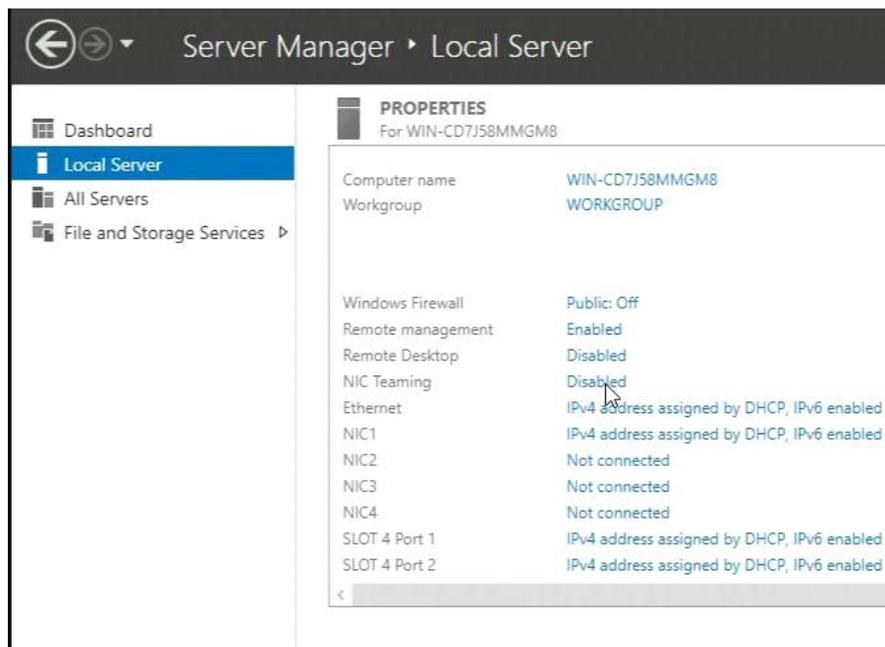
5. Under **ADAPTERS and INTERFACES**, control-click on the NIC ports you want to add to the team.
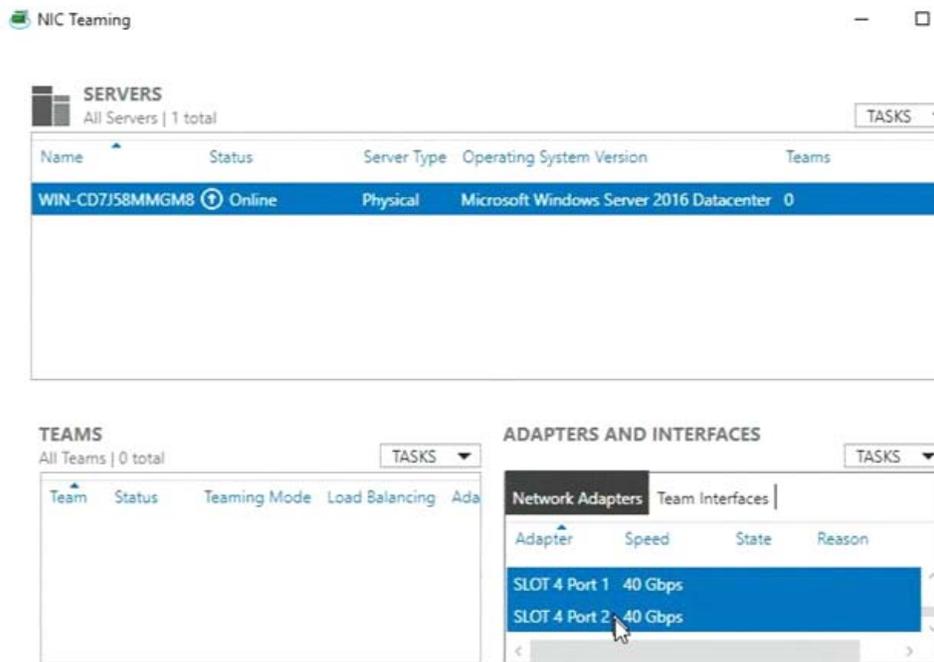
6. Right-click the selected ports and choose **Add to New Team** from the pop-up menu.
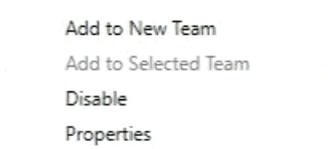
7. Expand **Additional properties** and set the **Teaming mode** to LACP.
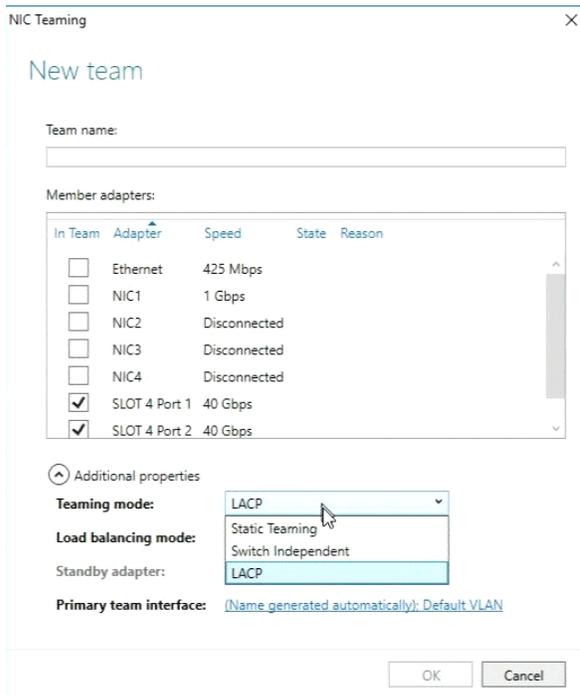
8. Enter a name for the team under **Team name** and click **OK**.

9. Close the dialog box.

10. Open the Windows Control Panel and click **View network status and task** under the **Network and Internet** section.

11. Click **Change adapter settings**.

12. The team that was just setup caused a virtual adapter to be created, named after the team name supplied in Step 8. Right-click that virtual adapter and choose **Properties**.

13. Configure the networking settings as appropriate for your network. For this example IPv6 was disabled and IPv4 was set to use DHCP for configuration.

2.4   RHEL 7.4 NIC Teaming Configuration

NIC Teaming in RHEL is achieved using the Bonding Kernel Module. The configuration is achieved by creating or editing network scripts. You can accomplish these edits using a GUI or command line text editor so long as you run them as **root** or via **sudo**.

Assumption: The following steps assume that a channel-group has already been configured on the switch to which this NIC Team will be connected.

Note: All of the ports to be added to the NIC team must share the same configuration, speed, and duplex settings. Additionally, settings and capacities must match or be compatible with the interfaces of the switch to which this NIC Team will be connected.

Reference: The steps outlined here were derived from product documentation on *redhat.com*. For a complete list of limitations, options, and other details, refer to the product's documentation on the *redhat.com* website.

2.4.1   Bonding Kernel Module

The *Bonding Kernel* module is the driver that provides NIC teaming with Linux. Detailed information about the module and its configuration can be found on the *kernel.org* website.

With RHEL 7.4, the bonding kernel module is loaded dynamically as a result of these configuration steps. Hence, there is no reason to manually load the module. This information is included because older configuration instructions require the user to manually ensure the bonding kernel module was loaded.

2.4.2   NIC Team Setup

To configure NIC Teaming with RHEL 7.4, follow these steps:

1. Connect the NIC ports from the server to the ports on the switch that are members of the channel-port created in Section 2.2.2. The type of connectors and cables vary based on the environment.

2. Navigate to the *network-scripts* directory.

   ```
   ~ cd /etc/sysconfig/network-scripts/
   ```

3. Start a new file in the text editor of your choice. The editor must have **root** privileges to write and edit the files in the *network-scripts* directory.

4. Create the virtual device for the NIC Team. The file name is *ifcfg-bond0*.

5. Enter the following text into the file:

   ```
   NAME=bond0
   DEVICE=bond0
   BONDING_MASTER=yes
   TYPE=bond
   IPADDR=192.168.100.100
   NETMASK=255.255.255.0
   GATEWAY=192.168.100.1
   BONDING_OPTS="miimon=100 mode=4 lacp_rate=1"
   #BOOTPROTO=none
   NM_CONTROLLED=no
   ```

   Note: This example demonstrates configuring a static address. You can also configure the virtual device to use DHCP. For DHCP, omit the `IPADDR`, `NETMASK`, and `GATEWAY` entries, then uncomment the `BOOTPROTO` option and set it to `DHCP`.

> **Note:** With a static IP Address, you must enter the commented `BOOTPROTO` option to work around a problem where NetworkManger interferes with the configuration.

6. Link the physical ports to the virtual device.

   For this example, the physical devices p4p1 and p4p2 are used. Open the files named *ifcfg-p4p1* and *ifcfg-p4p2*, then make the following edits:

   ```
   NAME="p4p1"
   DEVICE="p4p1"
   ONBOOT=yes
   UUID="ea032961-e2ea-4be1-a78f-d6685a66d917"
   BOOTPROTO=none
   TYPE=Ethernet
   MASTER=bond0
   SLAVE=yes
   NM CONTROLLED=no
   ```

   > **Note:** Your UUID will be different, or you may have a `MACADDR` or `HWADDR` in its place.

7. Repeat Step 6 for each physical port to be added to the team.

8. Enter the following command to restart the network services to activate the team:

   ```
   ~ systemctl restart network
   ```

9. Enter the following command to verify that the device is running correctly:

   ```
   ~ cat /proc/net/bonding/bond0
   ```

   Your results should similar to the following:

   ```
   Bonding Mode: IEEE 802.3ad Dynamic link aggregation
   Transmit Hash Policy: layer2 (0)
   MII Status: up
   MII Polling Interval (ms): 100
   Up Delay (ms): 0
   Down Delay (ms): 0

   802.3ad info
   LACP rate: fast
   Min links: 0
   Aggregator selection policy (ad_select): stable
   System priority: 65535
   System MAC address: 90:e2:ba:17:39:e8
   Active Aggregator Info:
            Aggregator ID: 1
            Number of ports: 2
            Actor Key: 13
            Partner Key: 32768
            Partner Mac Address: 00:0d:ec:f6:c4:bc

   Slave Interface: p4p1
   MII Status: up
   Speed: 10000 Mbps
   Duplex: full
   Link Failure Count: 2
   Permanent HW addr: 90:e2:ba:17:39:e8
   Slave queue ID: 0
   Aggregator ID: 1
   Actor Churn State: none
   Partner Churn State: none
   Actor Churned Count: 0
   Partner Churned Count: 0
   details actor lacp pdu:
   ```

```
              system priority: 65535
              system mac address: 90:e2:ba:17:39:e8
              port key: 1
              port priority: 25
              port number: 1
              port state: 63
details partner lacp pdu:
              system priority: 32768
              system mac address: 00:0d:ec:f6:c4:bc
              oper key: 32768
              port priority: 32768
              port number: 271
              port state: 61

Slave Interface: p4p2
MII Status: up
Speed: 10000 Mbps
Duplex: full
Link Failure Count:
Permanent HW addr: 90:e2:ba:17:39:e9
Slave queue ID: 0
Aggregator ID: 1
```

2.5 VMware ESXi 6.5 U1 NIC Teaming Configuration

Assumption: The following steps assume that a channel-group has already been configured on the switch to which this NIC Team will be connected.

Note: All of the ports to be added to the NIC team must share the same configuration, speed, and duplex settings. Additionally, settings and capacities must match or be compatible with the interfaces of the switch to which this NIC Team will be connected.

Reference: The steps outlined here were derived from product documentation on *vmware.com*. For a complete list of limitations, options, and other details, refer to the product's documentation on the *vmware.com* website

To configure NIC Teaming with VMware ESXi 6.5 U1, follow these steps:

1. Connect the NIC ports from the server to the ports on the switch that are members of the channel-port created in Section 2.2.2. The type of connectors and cables vary based on the environment.

2. On the host system, access the shell.

3. Add a new Standard vSwitch. The following example uses the name `lag1`:

   ```
   ~ esxcli network vswitch standard add -v lag1
   ```

4. Create a Port Group. The following example names the Port Group `PROD`:

   ```
   ~ esxcli network vswitch standard portgroup add -p PROD -v lag1
   ```

5. Add uplinks to the vSwitch. The following example uses the NICs named `vmnic4` and `vmnic5`.

   ```
   ~ esxcli network vswitch standard uplink add -u vmnic4 -v lag1
   ~ esxcli network vswitch standard uplink add -u vmnic5 -v lag1
   ```

6. Configure the Teaming Properties

   ```
   ~ esxcli network vswitch standard policy failover set -v lag1 -l iphash -a 'vmnic4,vmnic5'
   ```

Note: The load balancing should be set to **Route Based on IP Hash**, and both links should be active.

2.6   Intel® Ethernet 700 Series Network Adapters - Special Considerations

Each Intel® Ethernet 700 Series Network Adapter has a built-in hardware LLDP engine, which is enabled by default. The LLDP Engine is responsible for receiving and consuming LLDP frames, and also replies to the LLDP frames that it receives. The LLDP engine does not forward LLDP frames to the network stack of the Operating System.

LACP may not function correctly in certain environments that require LLDP frames containing LCAP information to be forwarded to the network stack. To avoid this situation, the user must disable the Intel® Ethernet 700 Series Network Adapter's hardware LLDP engine.

The following subsections provide instructions on how to disable the Intel® Ethernet 700 Series Network Adapter's hardware LLDP engine.

2.6.1   RHEL 7.4

To disable the Intel® Ethernet 700 Series Network Adapter's hardware LLDP engine in a RHEL 7.4 environment, execute the following command in Linux Shell:

```
ethtool –set-priv-flags <interface name> disable-fw-lldp on
```

2.6.2   VMware ESXi 6.5 U2

Disabling the Intel® Ethernet 700 Series Network Adapter's hardware LLDP engine for ESXi 6.5 U2 uses an i40en module parameter option. Execute the following command on VMware ESXi console to disable the LLDP engine:

```
esxcfg-module -s LLDP=0,0 i40en
```

The above command disables the LLDP engine on Port 0 and Port 1 of the Intel® Ethernet Converged Network Adapter X710-DA2. The ESXi host must be rebooted for the command to take effect.

3.0 Test Results

The following Intel® Ethernet Network Adapters were successfully tested with all of the configurations listed in Section 2.0.

Iperf3 was used to test the ability of the LAG to fail-over without dropping an active connection. Continuous pings were used to test how long it took to the recover the LAG if all of its physical members were disconnected and then reconnected. In all cases, the LAG maintained the **Iperf3** conversation on the first disconnect, and continuous pings showed that the LAG recovered in approximately 30 seconds when the links were re-established after a total disconnect of cables.

The testing of these configurations used a single switch for each channel-group. It is possible (and in some cases preferable) to span the channel-group across separate switches that are on the same network segment. Please consult your product manual for such configurations. VMware's ESXi host is the exception to this possibility, in that the channel-group members are required to reside on the same physical switch.

The switches used during testing where selected based on what was readily available in the test environment and would provide the speed needed for the current test case. As long as the target switch supports LACP, there is no reason why a switch other than what is listed should not work.

In all cases the drivers for the network adapters where updated to the latest version available on intel.com prior to testing.

Table 1 lists all of the hardware and software used in the testing.

Table 1. Test Matrix

Adapter	NVM Version	Cisco Switch	Speed	Operating System	NIC Driver Version
Intel® Ethernet Converged Network Adapter X520-SR2	N/A	Nexus 5010	10 GbE	RHEL 7.4	5.3.6
				Windows Server 2016	4.1.77.0
				Windows Server 2012 R2	4.1.77.0
				ESXi 6.5 U1	Ixgben_1.6.5-1OEM.600.0.0.2768847
Intel® Ethernet Converged Network Adapter X540-T2	N/A	Catalyst 4948	1 GbE	RHEL 7.4	5.3.6
				Windows Server 2016	4.1.76.0
				Windows Server 2012 R2	4.1.76.0
				ESXi 6.5 U1	Ixgben_1.6.5-1OEM.600.0.0.2768847
		Nexus 5596T	10 GbE	RHEL 7.4	5.3.6
				Windows Server 2016	4.1.76.0
				Windows Server 2012 R2	4.1.76.0
				ESXi 6.5 U1	Ixgben_1.6.5-1OEM.600.0.0.2768847
Intel® Ethernet Converged Network Adapter X550-T2	N/A	Catalyst 4948	1 GbE	RHEL 7.4	5.3.6
				Windows Server 2016	4.1.74.0
				Windows Server 2012 R2	4.1.74.0
				ESXi 6.5 U1	Ixgben_1.6.5-1OEM.600.0.0.2768847
		Nexus 5596T	10 GbE	RHEL 7.4	5.3.6
				Windows Server 2016	4.1.74.0
				Windows Server 2012 R2	4.1.74.0
				ESXi 6.5 U1	Ixgben_1.6.5-1OEM.600.0.0.2768847

Table 1. Test Matrix (Continued)

Adapter	NVM Version	Cisco Switch	Speed	Operating System	NIC Driver Version
Intel® Ethernet Converged Network Adapter X710-DA2	6.01	Nexus 5010	1 GbE and 10 GbE	RHEL 7.4	2.4.6
				Windows Server 2016	1.8.94.0
				Windows Server 2012 R2	1.8.94.0
				ESXi 6.5 U1	i40en-1.5.6-1OEM.650.0.0.4598673
Intel® Ethernet Converged Network Adapter X710-DA4	6.01	Catalyst 4948	1 GbE and 10 GbE	RHEL 7.4	2.4.6
				Windows Server 2016	1.8.94.0
				Windows Server 2012 R2	1.8.94.0
				ESXi 6.5 U1	i40en-1.5.6-1OEM.650.0.0.4598673
Intel® Ethernet Network Adapter XXV710-DA2	6.02	Nexus 5010	10 GbE	RHEL 7.4	2.4.6
				Windows Server 2016	1.8.94.0
				Windows Server 2012 R2	1.8.94.0
				ESXi 6.5 U1	i40en-1.5.6-1OEM.650.0.0.4598673
		Nexus 3232C	25 GbE	RHEL 7.4	2.4.6
				Windows Server 2016	1.8.94.0
				Windows Server 2012 R2	1.8.94.0
				ESXi 6.5 U1	i40en-1.5.6-1OEM.650.0.0.4598673
Intel® Ethernet Converged Network Adapter XL710-QDA2	6.01	Nexus 5672UP	40 GbE	RHEL 7.4	2.4.6
				Windows Server 2016	1.8.94.0
				Windows Server 2012 R2	1.8.94.0
				ESXi 6.5 U1	i40en-1.5.6-1OEM.650.0.0.4598673

Note: Make sure to use Native Mode API based ESX drivers for Intel® Ethernet Network Adapters. The latest drivers are available from the VMware Compatibility Site.

4.0   Summary

This document, along with the accompanying test results, demonstrate best-of-breed Intel® Ethernet 500 and 700 Series Network Adapters, providing (alongside its partners) a high-availability network I/O solution that modern IT environments demand.

4.1   Driver Download Locations

Operating System	Link
Windows	*https://downloadcenter.intel.com/product/46828/10-25-40-Gigabit-Ethernet-Adapters*
Linux	*https://downloadcenter.intel.com/product/46828/10-25-40-Gigabit-Ethernet-Adapters*
VMware ESXi 6.5 U1	*https://my.vmware.com/web/vmware/details?downloadGroup=DT-ESXI65-INTEL-I40EN-156&productId=614*
	https://my.vmware.com/web/vmware/details?downloadGroup=DT-ESX60-INTEL-NIC-IXGBEN-165&productId=491

4.2   Product Information

To see the full line of Intel® Ethernet Network Adapters for PCI Express, visit:

www.intel.com/ethernet

To speak to a customer service representative regarding Intel products, call 1-800-538-3373 (U.S. and Canada) or visit the following for the telephone number in your area:

www.intel.com/support

NOTE: **This page intentionally left blank.**

LEGAL

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.

This document (and any related software) is Intel copyrighted material, and your use is governed by the express license under which it is provided to you. Unless the license provides otherwise, you may not use, modify, copy, publish, distribute, disclose or transmit this document (and related materials) without Intel's prior written permission. This document (and related materials) is provided as is, with no express or implied warranties, other than those that are expressly stated in the license.

Intel disclaims all express and implied warranties, including without limitation, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, as well as any warranty arising from course of performance, course of dealing, or usage in trade.

This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest forecast, schedule, specifications and roadmaps.

The products and services described may contain defects or errors which may cause deviations from published specifications.

Copies of documents which have an order number and are referenced in this document may be obtained by calling 1-800-548-4725 or by visiting www.intel.com/design/literature.htm.

Intel and the Intel logo are trademarks of Intel Corporation in the U.S. and/or other countries.

* Other names and brands may be claimed as the property of others.

© 2018 Intel Corporation.