



Intel® Cluster Ready 1.2
Intel® Server Board S5520UR
Platform* HPC 3.0.1
Red Hat* Enterprise Linux* 6.1
Configuration C1 (Westmere)

Version 1.0
10/24/2011



Legal Notices

The information contained in this document is provided for informational purposes only and represents the current view of Intel Corporation ("Intel") and its contributors ("Contributors") on, as of the date of publication. Intel and the Contributors make no commitment to update the information contained in this document, and Intel reserves the right to make changes at any time, without notice.

THIS DOCUMENT IS PROVIDED "AS IS" WITH NO WARRANTIES WHATSOEVER, INCLUDING ANY WARRANTY OF MERCHANTABILITY, NONINFRINGEMENT, FITNESS FOR ANY PARTICULAR PURPOSE, OR ANY WARRANTY OTHERWISE ARISING OUT OF ANY PROPOSAL, SPECIFICATION OR SAMPLE.

Intel disclaims all liability, including liability for infringement of any proprietary rights, relating to use of information in this specification. No license, express or implied, by estoppels or otherwise, to any intellectual property rights is granted herein.

Except that a license is hereby granted to copy and reproduce this Document for internal use only.

This document is provided under the terms of the Intel® Cluster Ready Program Agreement between Intel and your company.

This document is subject to change, as described in the Intel® Cluster Ready Program Agreement.

Intel, the Intel logo, and Intel Xeon are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

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

Copyright © 2006, 2007, 2008, 2009, 2010, 2011. Intel Corporation. All rights reserved.

Table of Contents

1. Hardware configuration.....	4
Required Hardware ingredients	4
Optional Hardware ingredients.....	4
BIOS Settings	4
Remote console configuration / KVMIP	5
Cluster configuration	5
2. Getting started.....	6
Introduction.....	6
Required Software Ingredients.....	6
Collateral Documentation	6
3. Red Hat* installation.....	8
Preparation.....	8
Install Red Hat* Enterprise Linux on the Installer Node	8
Environment information	8
Network configuration.....	8
Set disk partitioning schema and software set	9
4. Platform* HPC installation	11
Preparation.....	11
Prerequisites	11
Install Platform* HPC	11
5. Installer Node Customization	13
Starting Customization	13
Enabling Intel® Cluster Ready Software	13
Run the ICR configuration script.....	13
6. Build compute nodes	14
7. Verify the Cluster	15
Configure Intel® Cluster Checker tool automatically with icr_configure script	15

1. Hardware configuration

Required Hardware ingredients

Quantity	Item	Manufacturer	Model
32	Intel® server board Urbanna	Intel®	S5520UR
	Intel® server chassis	Intel®	SR1600URSASBPP
	Intel® HDD backplane	Intel®	ASR1600PASBP
	2 Intel® Xeon® Processors - Westmere.	Intel®	X5660 @ 2.80GHz Stepping: B0
	6x 2GB DDR3 PC3-10600	Micron*	MT18JSF25672PDZ-1G4D1
	500Gb SATA Hard Disk Drive 3 Gbs	Seagate*	Barracuda* ST3500320NS
	DVD/CDRW – Slimline SR1550/SR1560	Intel®	AXXDVDCDR
1	A low latency gigabit Ethernet switch.	Hewlett-Packard*	ProCurve* J4904A
	KVM over IP Solution	Avocent*	DSR8035

Optional Hardware ingredients

It is not required but advisable to install a low latency fabric for compute nodes interconnect.

Quantity	Item	Manufacturer	Model
32	InfiniBand* cabling / QSFP connector	Mellanox*	MCC4Q30C-002 Copper Cable 4XQSFP 30AWG 2m
	ConnectX IB - Dual-Port InfiniBand* Adapter Card QDR	Mellanox*	MHQH29-XTC Hw Revision: a0 Fw Version: 2.9.1000
1	MTS3600-BNC 36-port 20 and 40Gb/s InfiniBand* Switch System	Mellanox*	MTS3600Q-1BNC

BIOS Settings

- 1) The required firmware (S5500.86B.01.00.0059.082320111421) components for the S5520UR server board are:

- BIOS version 59
- BMC version 57
- ME version SPS_01.01.02.007.0_Production (1.12)
- FRU/SDR version 26

To verify the version installed on each server, enter the BIOS setup screen by pressing <F2> when prompted during BIOS POST, then navigate to System Management->System Information tab, and find the version information on that screen.

Please refer to the BIOS installation instructions on how to update the BIOS.

- 2) BIOS configuration requires some settings deviated from the default server configuration.

Execute the following steps to set the proper configuration:

- Enter the BIOS setup screen by pressing <F2> during BIOS POST.

- Load default BIOS settings by pressing <F9>.
- Disable graphical splash screen by setting Main -> Quiet Boot to Disabled.
- Disable processor throttling by unsetting Advanced->Processor Configuration->Enhanced Intel® Speedstep® Technology.
- Disable Intel® Hyper-Threading Technology by unsetting Advanced->Processor Configuration->Intel® Hyper-Threading Technology.
- Disable Intel® Virtualization support by unsetting Advanced-> Processor -> Configuration -> Intel® Virtualization Technology.
- Select Server Management -> Console Redirection. Select Console Redirection and choose "Serial Port A". Default settings are enough.
- Installer node:
 - a. Set the boot order to SATA3:TSSTcorp CDDVDW, #0500 ID01 LUN ATA, IBA GE Slot0100 (Intel® Boot Agent)
- Compute nodes:
 - a. Set the boot order to IBA GE Slot0100, #0500 ID01 LUN ATA drive.
- Press <F10> to save and reboot.

These BIOS settings must be set identically on all nodes of the cluster.

Remote console configuration / KVMIP

- 3) Using a KVMIP solution is one method for supplying remote console access. Connect and configure the KVMIP solution such that there is remote console access to each node in the cluster. Note: this Reference Design requires an Avocent KVM over IP solution. However, any other KVMIP or serial-over-LAN solutions will satisfy this requirement.

Cluster configuration

- 4) There must be a separate Installer Node plus the compute nodes. All the switching equipment must interconnect all nodes.
- 5) Ethernet Port 1 (eth0) on all nodes must be connected to the private network. This is the messaging network for Ethernet as well as the management and storage network. No systems outside the cluster should be connected to this network.
- 6) Ethernet Port 2 (eth1) on the Installer Node should be connected to the public network. This is the entry point to the cluster. Port 2 should remain disconnected on all compute nodes.
- 7) If an optional InfiniBand* fabric is installed in the cluster as a second messaging network, each node must connect port 1 (ib0) of its HCA to the InfiniBand* switch.
- 8) Serial port on the compute nodes should be connected to a terminal server if the terminal server is used for remote console logging on compute nodes.

2. Getting started

Introduction

This cluster recipe provides the instructions for installing and configuring Intel® servers with Platform* HPC Enterprise Edition to be compliant with the Intel® Cluster Ready architecture specification.

Required Software Ingredients

Distributed By	Description	Contact Information
Intel® Corporation	Reference Design Package The following files are included: * The Intel Cluster Checker files (config, head and node list) * The Reference Design Scripts * The Intel Cluster Ready Certificate * The Intel Cluster Ready Reference Implementation Release Notes	http://www.intel.com/go/cluster Intel_Cluster_Ready_Reference_Design-S5520UR-ICR1.2-HPC3.0.1-RH6.1-C1-v1.0.zip
Platform Computing*	Intel® Cluster Checker 1.8 Program registration is needed.	http://my.platform.com kit-intel-cluster-checker-1.8-1.x86_64.iso md5: 3cc90a7698518235b7c0d1e67ceb5c67
Platform Computing*	Platform HPC Enterprise Edition 3.0.1 Included kit list in the ISO image: * base * intel-vendor * nagios * os-ofed * pcm * platform-hpc-high-performance-computing-tools * platform-hpc-web-portal * platform-isf-ac * platform-lsf * platform-lsf-gpu * platform-mpi Site registration is needed.	http://my.platform.com hpc-3.0.1-intel-6135.rhel.iso md5: 7f326049371a5733aee3945c449d2035
Red Hat*	Red Hat* Enterprise Linux 6.1	https://www.redhat.com/apps/download/RHEL6.1-20110510.1-Server-x86_64-DVD1.iso SHA256: b10e41b8476e0ba4b9b5902591e4e294fe3e9e1a
Platform Computing*	Intel® Cluster Runtimes 3.3 Program registration is needed.	http://my.platform.com kit-intel-runtime-3.3-1.x86_64.iso md5: 58d1465c3a08d2cdd1b99aff07211a52

Collateral Documentation

Red Hat Operating System Website: <http://www.redhat.com>

Platform Provisioning system Website: <http://www.platform.com/>

Intel Server Boards: <http://www.intel.com/products/server/motherboard>
Intel Cluster Ready Program: <http://www.intel.com/go/cluster>

3. Red Hat* installation

Preparation

Obtain the Red Hat* Enterprise Linux version provided in the BOM as a physical DVD media. Platform* HPC and all Platform* kits will be requested as .iso images. Please contact the vendor in order to get the required media.

Prepare and be ready to provide the following data before continuing the deployment below - full domain name, fixed Head IP/netmask, private IP/netmask, gateway, DNS, and NTP addresses.

Install Red Hat* Enterprise Linux on the Installer Node

Insert the Red Hat* Enterprise Linux installation disk into the DVD drive of the Installer Node. Boot the system and configure BIOS settings if needed.

When the Red Hat* Enterprise Linux splash screen appears, select "Install or upgrade an existing system" and press <enter>. If nothing is entered within a timeout period the system will proceed with the installation process.

Environment information

9) Press <Skip> to skip the media test and start the installation.

10) On the Red Hat* Enterprise Linux - Installer window

On the "Welcome" screen, just press <Next>

On the "Language selection" screen, select "English" and press <Next>

On the "Keyboard selection" screen, select "U.S. English" and press <Next>

On the "Storage device selection" screen, select "Basic Storage Device" and press <Next>

Network configuration

11) Configure the Hostname. Enter a fully qualified public domain.

12) Press the "Configure Network" button at the bottom left corner.

13) Select "Wired" tab, select the "System eth1" and press <Edit>.

- Select the "Connect automatically" checkbox at upper left corner.
- Select "IPv4 Settings" tab, and change the "Method" to manual.
- Press "Add" button and enter the IP, netmask and gateway addresses.

Consider that the eth1 is the interface which interconnects the cluster with network outside the cluster, usually called "public network".

Gateway: add the desired gateway information.

- Enter a DNS server address.

If you have more than one address, enters each one separated by commas. The following parameters should be set according to the local Network configuration.

Primary DNS: A valid Primary DNS entry is required. Errors may occur if this value is not properly configured.

Secondary DNS: Secondary DNS entry is optional based on the network configuration.

- Finally press <Apply>.

14) Select "Wired" tab, select the "System eth0" and press <Edit>.

- Select the "Connect automatically" checkbox at upper left corner.
- Select "IPv4 Settings" tab, and change the "Method" to manual.
- Press "Add" button and enter the IP and netmask addresses.

Consider that the Eth0 is the interface for the private network. Compute nodes are provisioned from the Installer Node over the private network. For example, Eth0 could be configured with a private IP of 172.20.10.1, a netmask of 255.255.255.0.

- Finally press <Apply>.

15) To continue with the installation press <Close> and then press <Next>

16) Configure Time Zone and press <Next>

17) Set root password

Enter the root password that will be shared with every node within the cluster. If a window appears complaining the strength of the password press <Use Anyway> to continue, or press <Cancel> and set a stronger password.

- After configuring the root password, press <Next>

Set disk partitioning schema and software set

18) Select <Use All Space> and mark the "Review and modify partitioning layout" checkbox. Then press <Next> to continue.

19) At the "Partitioning schema" window, edit the existent partition schema according to next information:

- 51200MB ext4 partition for 'lv_home' logical volume mounted at '/home'.
- 102400MB ext4 partition for 'lv_root' logical volume mounted at '/'.
- 2000MB swap partition for 'lv_swap' logical volume.
- <remaining space> ext4 partition for 'lv_depot' logical volume mounted at '/depot'.

20) After finish creating/editing partition schema, press <Next> to continue.

- At the 'Warning' window press <Write changes to disk> to confirm the changes.

21) In the "Bootloader installation" screen, just press <Next> due the default configuration it's enough.

22) After the disk formatting you can choose a set of software to install. Select 'Basic Server' from the list and the 'Customize now' radio button at the bottom right corner to meet the required packages. Press <Next>.

23) In the customization screen select "Base System" on the right column and select "Base" on the left column. Press <Optional packages> and scroll down until "yum-plugin-

tmprepo” package. Select the checkbox, press <Close> and then <Next> to continue with the installation.

24) Press <Next> to continue and start the Installation process.

25) Press <Reboot> to restart the Installer Node once the installation completes.

4. Platform* HPC installation

Preparation

Place the Platform* HPC, the Intel® Cluster Checker kit and Intel® Cluster Runtimes kit ISOs in the /root folder of the Installer node.

The required ISOs are:

hpc-3.0.1-intel-6135.rhel.iso
kit-intel-cluster-checker-1.8-1.x86_64.iso
kit-intel-runtime-3.3-1.x86_64.iso

Note: It is strongly suggested to verify the files integrity with the 'md5sum' command before proceeding. The checksum of each file is available at the BOM of this document.

26) Once the Installer Node booted, log into the system as 'root'.

Note: In some cases, it may be possible that SSH questions are prompted for enabling password-less logging: just press <enter> to answer the questions.

Prerequisites

Before starting with the Platform* HPC installation it is required to accomplish some prerequisites.

27) Configure SELinux to be disabled after system reboot by running the following command.

```
# sed -i.orig -e 's/\(^SELINUX=\)enforcing$/\ldisabled/'  
/etc/selinux/config
```

28) Add MD5 authentication method by executing the following command.

```
# authconfig --enablemd5 --update  
# passwd
```

29) Finally reboot the cluster to let the configurations be applied.

```
# reboot
```

Install Platform* HPC

30) Mount the Platform* HPC ISO image into a folder

```
mkdir /mnt/platform_hpc  
mount -o loop /root/hpc-3.0.1-intel-6135.rhel.iso /mnt/platform_hpc
```

31) Execute pcm-installer

```
/mnt/platform_hpc/pcm-installer
```

32) Press 'enter' twice to configure "eth0" as the provisioning network and "eth1" as the public network.

33) Type down the Cluster Domain and then press 'enter'. This is the internal cluster domain and will be appended to cluster compute node names.

34) Press 'enter' to avoid configuring HPC HA.

35) The following Kits will be added automatically by the PCM* installer.

- base
- intel-vendor
- nagios
- os-ofed
- pcm
- platform-hpc-high-performance-computing-tools
- platform-hpc-web-portal
- platform-isf-ac
- platform-lsf
- platform-lsf-gpu
- platform-mpi

After all kits were added, the HPC installer will request the media to install the Operating System from. Insert the Red Hat Enterprise Linux* DVD and press 'enter'.

36) Press '1' to select "DVD Drive" and then press 'enter' to start adding the Operating System kit.

Note: this process may take several minutes to finish.

37) Press '3' to select "Add extra kits(s)", and then press 'enter'. Press '2' to select "ISO image or mount point", then type `"/root/kit-intel-cluster-checker-1.8-1.x86_64.iso"` and press 'enter'.

38) Repeat the previous step to add: `"kit-intel-runtime-3.3-1.x86_64.iso"`.

39) Press '4' and then 'enter' to continue installation.

Note: This step will take around 20 minutes to complete.

40) Specify the full path to the Platform HPC license file if available or press 'Enter' to skip this step and register the license at a later time.

Note: if the license file is not provided the entire set of the product features won't be available.

41) Reboot the system one last time in order to load the new kernel modules installed by the intel-vendor and os-ofed kit.

```
reboot
```

42) At the "Text Mode Setup Utility" screen, just press "Quit".

5. Installer Node Customization

Starting Customization

43) Once the system boots again, login back as 'root' again.

Make sure that the following files specified at Software components are located at /root directory. If these files are not available the ICR configuration script will raise an error.

- Intel® Cluster Checker License (check the Intel® Cluster Ready website)
- Intel® Cluster Ready Reference Design Package (zip file)

Enabling Intel® Cluster Ready Software

44) Extract the recipe package.

```
unzip Intel_Cluster_Ready_Reference_Design_S5520UR-ICR1.2-HPC3.0.1-  
RH6.1-C1-v1.0.zip
```

45) Install the Intel® Cluster Ready Reference Design Package:

```
cd S5520UR-ICR1.2-HPC3.0.1-RH6.1-C1-v1.0  
source install.sh
```

Run the ICR configuration script

The configuration script will perform all necessary steps to build an Intel Cluster Ready compliant cluster. All steps are executed automatically and will be logged in a log file at the /var/log/icr folder. The name of the log file follows the following naming convention: " icr- <YYYYMMDD>.<hhmmss>.log".

46) Execute the Intel® Cluster Ready configuration script:

```
source /opt/intel/icr/icrvars.sh  
icr_configure [--no-infiniband=True]
```

Once the configuration script completes the configuration, it will display "Configuration Succeeded".

Note: If the cluster has no Infiniband* fabric available, please instruct the configuration script to do not use such configuration by adding the --no-infiniband=True switch.

6. Build compute nodes

Please refer to the Platform* HPC documentation for detailed instructions on how to run addhost tool. Please refer to the board documentation on how to boot into PXE mode.

47) From a root prompt on the Installer Node run the command

```
kusu-addhost -i eth0 -n compute-rhel-6.1-x86_64 -t rack --rack=<rack name> --size=1 -b
```

48) Power on each compute node, ensuring the default BIOS boot is configured to boot over the network. To avoid overloading the Installer Node it is advisable to wait 30 seconds between each node power on.

Watch the Installer Node console for indication that the compute node installation has begun (an installing banner will appear).

Note:

An errant network device (e.g. local IP switch broadcasting a DHCP request) can be detected and registered as a compute node during this installation phase. This must be corrected to release the associated compute node name for a real compute node registration. If some node was powered off or rebooted a complete reboot of all nodes in the cluster should be done after the installation finishes. Not rebooting the cluster can lead to Intel® Cluster Checker tests to fail. The following steps will release the registered compute node:

```
ctrl-c
kusu-addhost -e '[compute-node-name]'
kusu-addhost -i eth0 -n compute-rhel-6.1-x86_64 -t rack --rack=<rack name> --size=1 -b
```

49) After all the nodes have been discovered, press “ctrl-c” to update the Provisioning System Database

```
ctrl-c
```

Note:

It is very important that addhost is closed through “ctrl-c” before the compute nodes boots for the first time. This step is important as the first boot configuration relies on information from the database that is updated only when addhost is closed. Failure to achieve this requirement will lead to failures detected by the “hostname” test module of Intel® Cluster Checker and can be solved through the reinstallation of the failing compute node.

50) Check that the compute nodes were installed with the command below. A list of nodes will be displayed. If all nodes were installed successfully, an “Installed” label must appear next to the node name. Wait until all the nodes are installed.

```
kusu-nghosts -l
```

51) Check that the compute nodes have booted successfully. Use the “pdsh” command to verify that all the nodes are available. Please wait until all the nodes are available.

```
pdsh -a hostname
```

52) Verify that all compute nodes have the hostname correctly configured. Execute the following command and verify that each node has a name assigned to the HOSTNAME parameter:

```
pdsh -a 'cat /etc/sysconfig/network|grep HOSTNAME'
```

For example, the output must look like:

```
compute000-eth0: HOSTNAME=compute000.pcm
compute001-eth0: HOSTNAME=compute001.pcm
compute002-eth0: HOSTNAME=compute002.pcm
```

If there is a node with the HOSTNAME empty, reinstall it using:

```
kusu-boothost -r -m '[compute-node-name]'
```

7. Verify the Cluster

For detailed instructions on configuring and running Intel® Cluster Checker Tool refer to its user documentation.

Configure Intel® Cluster Checker tool automatically with icr_configure script

The following steps will configure Intel® Cluster Checker, run it in certification mode and execute the sales report.

- 1) Execute the ICRCLCK section script in order to configure the necessary files to run Intel® Cluster Checker.

```
icr_configure --section=ICRCLCK [--no-infiniband=True] [--head-as-compute=True]
```

Note 1: If the cluster has no Infiniband* fabric available, please instruct the configuration script to do not use such configuration by adding the --no-infiniband=True switch.

Note 2: To configure the head node also as compute node please instruct the configuration script to use such configuration by adding the --head-as-compute=True switch.

- 2) Log in as the icr user

```
su - icr
```

- 3) Run Intel® Cluster Checker to verify the cluster configuration

```
cluster-check --certification
```

Note that after a successful execution of Intel® Cluster Checker the following message should be displayed:

Check has **Succeeded**.

- If Intel® Cluster Checker tests failed, then the following output will be printed on the screen:

Check has **FAILED**.

The user will need to review the logs located at /var/log/intel/clck to identify the issue. The log file name has the following form: "config-certification-YYYYMMDD.hhmmss.<out|xml>". If Intel® Cluster Checker is executed more than once, then the latest log should be analyzed.

For more information on Intel® Cluster Checker, please review Intel® Cluster Checker's documentation available through man pages:

```
man cluster-check
```

4) Afterwards, run Intel® Cluster Checker to generate the sales report file.

```
cluster-check --sales-report /var/log/intel/clck/config-  
certification-YYYYMMDD.hhmmss.xml
```

Note that after a successful execution of Intel® Cluster Checker the following message should be displayed:

Created file: ./summary-XXXXXX.csv