Intel® Open Network Platform Server (Release 1.4)

Release Notes

Document Revision 1.3
July 2015

Revision History

Revision	Date	Comments
1.3	July 19, 2015	Updated section 3.2, "Release 1.4 Limitations," to provide a workaround for a Keystone change since release of 1.4 of Intel® Open Network Platform Server
1.2	June 9, 2015	Document updated for release of 1.4 of Intel® Open Network Platform Server
1.1	February 19, 2015	Document updated for release of 1.3 of Intel® Open Network Platform Server
1.0	December 15, 2014	Initial release (Intel public)

Contents

Figures

Tables

1.0 Introduction

This document describes Release 1.4 of the Intel® Open Network Platform Server (Intel ONP Server). The Intel ONP Server is a platform-level reference architecture designed to make it easier for software and hardware vendors, and service providers to develop and deploy Software-Defined Networking (SDN) and Network Function Virtualization (NFV) solutions. The deployment of this reference architecture is done mainly using DevStack. DevStack does not make the deployment production ready, but it does afford the developer a good option to experiment with Intel's software and hardware stack.

Intel ONP Server defines the integration of hardware and software components, providing a framework to deliver the many benefits of Intel architecture to SDN and NFV. The underlying foundation of this reference architecture is comprised of IA-based Standard High Volume Servers (SHVS). The software stack is based on open source software originated by open standard communities like Open vSwitch, Data Plane Development Kit (DPDK), OpenDaylight, OpenStack, and KVM. Intel is working closely with these communities and contributing to the evolution and advancement of their open standard projects.

Intel ONP Server provides a SDN/NFV reference solution that defines three main network elements: networking nodes, OpenDaylight controller, and OpenStack manager. An SDN/NFV lab environment comprised of those network elements is the basis for ONP Server software integration and customer use case validation.

This document describes the main functionalities contributed by Intel to open source community projects that are now integrated into Intel ONP Server Release 1.4. It also details major solution limitations that remain.

Note: Intel has decided to increase investment in Open vSwitch community projects focusing on DPDK and advance hardware acceleration beginning with Release 1.4.

2.0 System Overview

2.1 Release 1.4 Highlights

Intel ONP Server, Release 1.4, introduces new capabilities and some software upgrades from the previous release:

- Integration with the latest Open Source software releases:
  - Upgrade to Openstack Kilo 2015.1 release with the following key feature enhancements:
    - Enhanced Platform Awareness (EPA) capabilities
    - Improved CPU pinning to virtual machines
    - I/O-based Non-Uniform Memory Architecture (NUMA) aware scheduling
  - Integration with Data Plane Development Kit release 1.8
  - Upgrade to a newer Open vSwitch version that supports VFIO enabling the use of standard kernel modules and improved performance
  - Upgrade to OpenDaylight Helium SR3 release
- Updated to support Real-Time Linux Kernel (v3.14.36-rt34) installation

2.2 Intel ONP Server Release Distribution

Intel ONP Server, Release 1.4, is delivered in the form of a Reference Architecture Guide. Access to the documentation is under: Intel ONP Servers Release 1.4 Software.

Intel® ONP Server Reference Architecture Guide — This document provides instructions on how to build Intel ONP Server software, set the functionality test environment, and perform the tests.

2.3 Reference Architecture Environment

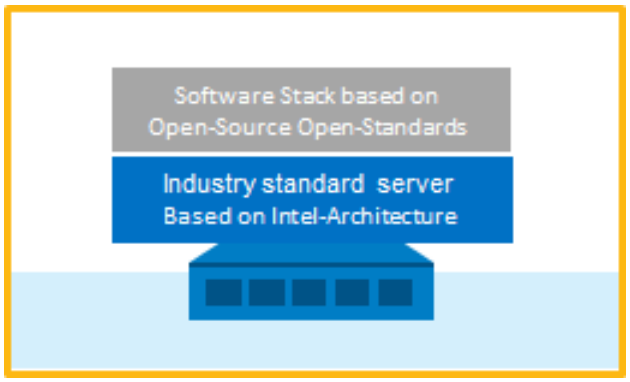

Figure 2–1. Intel ONP Server Node View

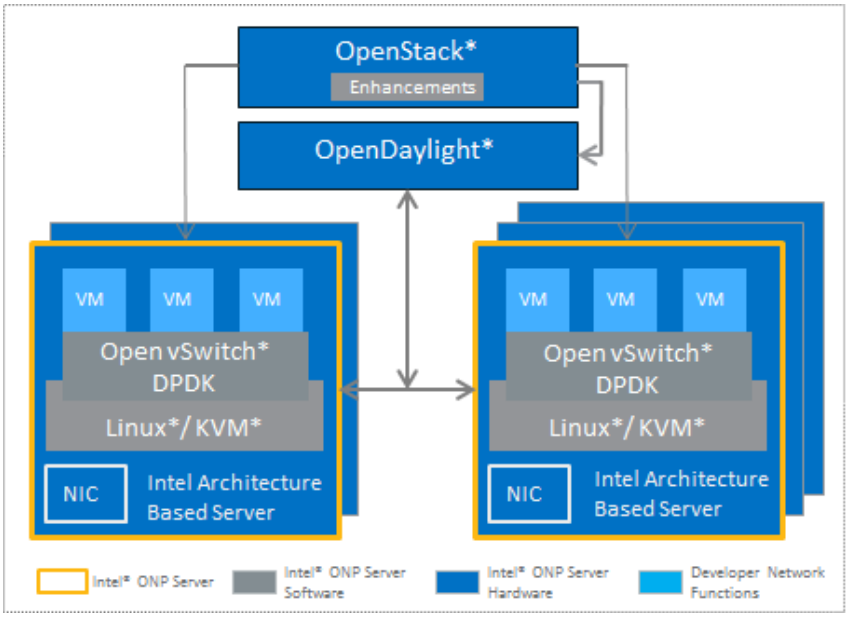

Figure 2–2. Intel ONP Server Test and Validation Environment

2.4 Network Elements

2.4.1 Compute Node Software

The following table shows the compute node software components that were used during the testing of VNF components —vIPS, vBNG, NUMA placement and SR-IOV pass-through for OpenStack.

Table 2–1. Compute Node Software

Software	Notes
Fedora 21 x86_64 (Server Version)	Underlying system-level OS, based on the 3.18.8-201.fc21. x86_64 kernel.
Linux* Real-Time Kernel	Real-Time Linux Kernel 3.14.36-rt34: Provides a baseline to give real-time aspects to the compute nodes.
DPDK 1.8.0	Software libraries are used to dramatically accelerate packet processing, increasing throughput and scalability.
OpenvSwitch with: – Kernel datapath: v2.3.1-git4750c96 – DPDK datapath: v2.3.90	Includes support for OVS with DPDK-netdev
Libvirt v1.2.9.2.fc21.x86_64	Toolkit and API are used by QEMU-KVM to manage virtual machines. OpenStack (Nova) also uses it to manage the compute resources of the host.
QEMU-KVM v2.1.3-3.fc21.x86_64	Open Source machine emulator and virtualizer: Includes KVM that is used to enable hardware accelerations in Intel® platforms

2.4.2 Controller Node Software

The following table shows the controller node software components that were used during the testing of VNF components — vIPS, vBNG and NUMA placement, and SR-IOV passthrough for OpenStack.

Table 2–2. Controller Node Software

Software	Notes
Fedora 21 x86_64 (Server Version)	Underlying system OS: Upgraded to the 3.18.8-201.fc21. x86_ 64 kernel
OpenStack* Kilo	OpenStack-related tools for building and managing clouds. Includes the DevStack shell script for automating development-environment builds.
OpenvSwitch v2.3.1-git4750c96	Includes support for OVS with DPDK-netdev
OpenDaylight Helium SR3	Use OpenDaylight as the OpenStack network management provider through the ML2 plug-in.

2.4.3 Reference Platform Hardware

Haswell-based Platforms

The ONP Server uses the latest and greatest Haswell-based platform. More hardware details are provided in Table 2–3.

Table 2–3. Hardware Ingredients (Code-named Wildcat Pass)

Item	Description	Notes
Platform	Intel® Server Board S2600WTT 1100 W power supply	Intel® Xeon® processor-based DP server (Formerly code-named Wildcat Pass) 120 GB SSD 2.5in SATA 6GB/s Intel Wolfsville SSDSC2BB120G4. Supports SR-IOV
Processors	Intel® Dual Xeon® Processor E5-2697 V3 2.6 GHz, 35 MB, 145 W, 14 cores	(Formerly code-named Haswell) 14 cores, 2.60 GHz, 145 W, 35 MB total cache per processor, 9.6 GT/s QPI, DDR4-1600/1866/2133, 28 hyper-threaded cores per CPU for 56 total cores
	Intel® Dual Xeon® Processor Series E5-2699 v3 2.3 GHz, 45 MB, 145 W, 18 cores	(Formerly code-named Haswell) 18 cores, 2.3 GHz, 145 W, 45 MB total cache per processor, 9.6 GT/s QPI, DDR4-1600/1866/2133, 36 hyper-threaded cores per CPU for 72 total cores
Memory	8 GB DDR4 RDIMM Crucial CT8G4RFS423	64 GB RAM (8 x 8 GB)
NICs (XL710)	Intel® Ethernet Controller XL710 4x 10GbE has been tested with Intel FTLX8571D3BCV-IT and Intel AFBR-703sDZ-IN2 850nm SFPs	(Code-named Fortville) NICs are on socket zero.
NICs (R520)	Intel® Ethernet Controller R520 2x 10GbE Firmware version 0x 18bf0001	(Code-named Niantic) NICs are on socket zero.
BIOS	SE5C610.86B.01.01.0008.031920151331 Release Date: 03/19/2015	Intel® Virtualization Technology for Direct I/O (Intel® VT-d) enabled only for SR-IOV PCI passthrough tests, hyper-threading enabled.
Quick Assist Technology	Intel® Communications Chipset 8950 (Coleto Creek)	Walnut Hill PCIe card 1x Coleto Creek; supports SR-IOV.

3.0 Functionality Highlights

3.1 New Functionalities Contributed by Intel to Open Source Projects

The following table summarizes new software features/bug fixes that Intel has contributed to open source projects.

Table 3–1. New Features/Bug Fixes Contributed to Open Source Projects

No.	Feature Name	Category	Ingredient	Description	IA Value	Commit ID if Upstreamed, Link in Posted as Patch
1.	CPU pinning	Feature/ bug fix	OpenStack	Support for pinning VMs to physical CPUs	Ability to pin workloads to physical CPU cores critical for DPDK-enabled applications	https://review.openstack.org/#/c/170190/ – libvirt: Add version check when pinning guest CPUs
2.	I/O (PCIe) based NUMA aware schedul-ing	Feature	OpenStack	Support affinitizing VM to the same NUMA nodes as the PCI device that is passed into the VM	Limit QPI traffic when using PCI passthrough, optimal performance when passing through NICS and QA devices	https://review.openstack.org/#/q/status:merged+project:openstack/nova+branch:master+topic:bp/input-output-based-numa-scheduling,n,z
3.	NUMA topology/ NUMA I/O schedul-ing bug fixes	Bug fix	OpenStack	Bug fixes for numa topology and io numa scheduling		https://bugs.launchpad.net/nova/+bug/1444021 – HostState.consume_from_instance fails when instance has numa topology https://bugs.launchpad.net/nova/+bug/1441169 – can't schedule vm with numa topology and pci device https://bugs.launchpad.net/nova/+bug/1445040 – InstancePCIRequests.obj_from_db fails to get requests from db https://bugs.launchpad.net/nova/+bug/1397381 – numa cell ids need to be normal-ized before creating xml

Table 3–1. (Cont'd)

No.	Feature Name	Category	Ingredient	Description	IA Value	Commit ID if Upstreamed, Link in Posted as Patch
4.	PCI pass-through bug fix	Bug fix	OpenStack	Bug fix for PCI passthrough		https://bugs.launchpad.net/nova/+bug/1435981 – Pci passthrough filter throws exception, if host has no pci devices. https://review.openstack.org/#/c/167817/
5.	Add support for VFIO	Feature	OpenStack/ DevStack plugin and vSwitch	VFIO was added in DPDK 1.7. This epic makes SRT use VFIO instead of UIO. VFIO provides a more secure user-space driver environment than UIO. Refer to https://www.kernel.org/doc/Documentation/vfio.txt	VFIO helps provide a more secure user-space driver environment than UIO. It allows the use of DPDK ports without having to insert a non-standard Linux kernel module.	https://review.openstack.org/#/c/171388/ – implements configurable dpdk phyical port drivers vswitch commit ID:491c2ea3230f53ecbe65a556a0a1cc68647d7b99
6.	ovs-dpdk deploy-ment via a Dev-Stack plugin	Feature	OpenStack/ DevStack plugin	ovs-dpdk is deployed using a devstack plugin. Users no longer need to patch DevStack; enabling the plugin is a configuration step.	N/A	https://review.openstack.org/#/c/155354/ – implements DevStack plugin support. https://review.openstack.org/168442 - Support third party (non-ovsdpdk) q_agents.

Table 3–1. (Cont'd)

No.	Feature Name	Category	Ingredient	Description	IA Value	Commit ID if Upstreamed, Link in Posted as Patch
7.	Various improvements to networking-ovs-dpdk	Feature/ bug fix	DevStack plugin	Various bug fixes and usability improvements to the DevStack plugin		https://review.openstack.org/175345 – Runs tests on neutron stable/kilo branch. https://review.openstack.org/173464 – fixes install of ovs when not rebuilding from source https://review.openstack.org/173367 – Add config option for disabling libvhost https://review.openstack.org/173385 – Cap the version of neutron in test requirements. https://review.openstack.org/171729 – Monkey patch Eventlet before starting the agent. https://review.openstack.org/171696 – Added check for lib/neutron -> lib/neutron-legacy https://review.openstack.org/171190 – Stop rebuilding of ovs/dpdk every time stack.sh is run https://review.openstack.org/169857 – Implements refactoring of codebase for consistant formating https://review.openstack.org/168431 -– hugepage mounts/kmodules discovery false positives issues https://review.openstack.org/166890 - Support launching VM's whitout root priviliges https://review.openstack.org/166871 – refactors ovs to free hugepages on failed start https://review.openstack.org/166862 – implements configable hugetlb mount point pagesize https://review.openstack.org/166844 – implements serial vswitch initalisation. https://review.openstack.org/175003 – Installation and usage documentation

Table 3–1. (Cont'd)

No.	Feature Name	Category	Ingredient	Description	IA Value	Commit ID if Upstreamed, Link in Posted as Patch
8.	40 GbE NIC sup port	Feature	DPDK	DPDK 1.8.0 added additional feature support for 40 GbE (4 x 10) (Fortville) NICs. It added enhanced Flow Director, Ethertype filtering, and RSS support.	Allows DPDK enabled applications to take advantage of Intel 40 GbE NICs.	Multiple commits under v1.8.0 tag. Major Commit: 8db9e2a1b2327ca0b9178e47f92e81b729cf0702
9.	DPDK 1.8 support	Feature	vSwitch	This patch ports the Open vSwitch datapath to use DPDK 1.8.0	Update Open vSwitch to use the latest version of DPDK. This allows users to take advantage of the latest DPDK features and performance improve-ments.	b8e57534ece5f620af7d7fa0278c8e9746dd719c
10.	Version update	SW Version	OpenDay-light	Use latest released version of OpenDaylight Helium – SR3. This is pulled in auto-matically with DevStack.	N/A	N/A

Table 3–1. (Cont'd)

No.	Feature Name	Category	Ingredient	Description	IA Value	Commit ID if Upstreamed, Link in Posted as Patch
11.	Java and Fedora versions	SW Version	OpenDay-light	The Java version required by OpenDaylight remains 1.7. The transition to using Fedora 21 requires specific installation steps to install Java 1.7 and change the default Java version from 1.8 (default in Fedora 21) to 1.7.	N/A	N/A

3.2 Release 1.4 Limitations

Table 3–2. Release 1.4 Limitations

No.	Open Software	Limitation
1.	OpenStack	After stack compilation is successful on controller and compute node, log in to OpenStack UI to spawn VM. Observe error: `Failed to connect to server (code 1006).` The workaround is to flush the iptables on the compute node with the command `iptables –F`.
2.	DPDK	The DPDK driver for XL710 will fail to load when the platform has more than 64 logical cores. The solution is to disable CPU cores through a BIOS change to make the total CPU cores 64 or less (e.g., for a 72-core system, disable 8 or more cores in the BIOS).
3.	ODL interoperability with OpenStack	For Fedora 21, the native Java version from the yum repository is Java v1.8. OpenDaylight Helium SR3, the current released version and used in this release, however, only supports Java v1.7. It is necessary to manually install Java 1.7 on Fedora 21 system and make it the default Java version. Refer to *Intel® Open Network Platform Server Reference Architecture (Release 1.4)*, section 6.4.2, for details of the installation procedure.
4.	ODL interoperability with OpenStack	On the ODL controller node, a bug exists in the */opt/stack/networking-odl/devstack/plugin.py* file that tries to install openjava-1.7 from the Fedora 21 repository, but it does not exist. Refer to *Intel® Open Network Platform Server Reference Architecture (Release 1.4)*, section 6.4.2, for details of the workaround.

5.	ONP integration script	It is a known issue that Intel ONP Server scripts have external dependencies (like Keystone) that change after a specific release of the ONP Server, and these changes might break the ONP Server installation. Intel is evaluating options to provide a more resilient and longer term solution than just documenting these issues and their workarounds as they occur.
		
		After ONP1.4 was released on June 9th, 2015, one of its dependencies (Keystone) was updated (to v1.5.2 on June 29th), which broke the OpenStack installation (i.e., the ***prepare_stack.sh*** script would fail on the controller node). For now, the following commands can be executed to work around this issue. A more resilient and longer term solution will be provided soon.
		
		Workaround Option 1. If you already executed ***prepare_stack.sh*** and experienced the failure, then execute the following commands:
		`cd /opt/stack/keystone`
		`git checkout stable/kilo`
		`cd /home/stack/devstack`
		`./unstack.sh`
		`./stack.sh`
		
		Workaround Option 2. Alternatively, if you have not already ran ***prepare_stack.sh***, then go to the folder */home/stack/* and edit the file *onps_commit_ids* as follows:
		1. Change the following line:
		*keystone git://git.openstack.org/openstack/keystone.git master **2015.1.0***
		To:
		*keystone git://git.openstack.org/openstack/keystone.git master **stable/kilo***
		
		2. Run `./prepare_stack.sh`

4.0 Installation Instructions

4.1 Node Software

This is a high level description of the installation process. Read the README file described below for more details:

Instructions:

1. Download the Intel ONP Server, Release 1.4 Scripts tarball from:

   https://01.org/packet-processing/intel®-onp-servers

   The bundle contains the following files:

Files	Description	Notes
README	A script to do the initial system configuration and pull the right updates from the Fedora repos	—
prepare_system.sh	A script to do the initial system configuration and pull the right updates from the Fedora repos	—
prepare_stack.sh	A wrapper to DevStack scripts	—
onps_config	The main configuration file with the necessary setup parameters	—
onps_config_stand_alone_compute	A version of onps_config which will set up a standalone compute node without any controllers involved	—
local.conf-controller	A DevStack template configuration file	Do not modify.
local.conf-compute	A DevStack template configuration file	Do not modify.
onps_commit_ids	Commit ids or tags for OpenStack components	Do not modify.

2. Plan ahead to decide what hosts you are going to use and what interfaces of your hosts will belong to management and/or the data plane network.
3. Install a fresh Fedora 21 installation on each host, and install the tarball in step 1 on each host.
4. Untar the tarball and confirm the above described files are present.
5. Manually edit the `onps_server` configuration file on each host based on the planned role of that host (controller/node), type of interfaces, type of vSwitch desired, etc.
6. Execute `prepare_system.sh` and reboot. The script will parse `onps_server` and prepare the system accordingly.
7. Execute `prepare_stack.sh`. The script will parse `onps_server` again and will kick off DevStack in the background.

After successfully executiing the above procedure in all the hosts, a fully deployed set of compute and controller nodes should be available with the version information detailed in section 2.4.

Note: This set of scripts and configuration files are for evaluation and facilitation only.

Legal Information

By using this document, in addition to any agreements you have with Intel, you accept the terms set forth below.

You may not use or facilitate the use of this document in connection with any infringement or other legal analysis concerning Intel products described herein. You agree to grant Intel a non-exclusive, royalty-free license to any patent claim thereafter drafted which includes subject matter disclosed herein.

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.

The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request. Contact your local Intel sales office or your distributor to obtain the latest specifications and before placing your product order.

Intel technologies may require enabled hardware, specific software, or services activation. Check with your system manufacturer or retailer. Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit http://www.intel.com/performance.

All products, computer systems, dates and figures specified are preliminary based on current expectations, and are subject to change without notice. Results have been estimated or simulated using internal Intel analysis or architecture simulation or modeling, and provided to you for informational purposes. Any differences in your system hardware, software or configuration may affect your actual performance.

No computer system can be absolutely secure. Intel does not assume any liability for lost or stolen data or systems or any damages resulting from such losses.

Intel does not control or audit third-party web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.

Intel Corporation may have patents or pending patent applications, trademarks, copyrights, or other intellectual property rights that relate to the presented subject matter. The furnishing of documents and other materials and information does not provide any license, express or implied, by estoppel or otherwise, to any such patents, trademarks, copyrights, or other intellectual property rights.