Simulate Intel’s Next generation communication platform data plane solutions

November 2012
Executive Summary

Today’s telecom equipment depends on a complex mix of CPUs and network processors. These heterogeneous architectures are difficult to program and scale poorly. The cost of sustaining and improving such a solution and making sure multiple software programming models and tool chains all work flawlessly together for a long period throughout product life cycle is quite high. Intel next generation communication platform code named Crystal Forest combines essential computing ingredients into one platform. It is capable of handling multiple types of workloads such as control plane, application plane and data plane concurrently – all with high throughputs. Previously, customers had to come up with multiple platforms and designs in order to support different types of operations. With Crystal Forest and follow-on platforms, all the computing needs can be handled with a single platform with a simpler and more consistent system and software architecture. This translates to faster time to market and lower Total Cost of Ownership (TCO).

This white paper discusses how to run Intel® DPDK software unmodified in a Simics Crystal forest environment. It describes the steps needed to run the Intel DPDK software stack on virtual platforms and compares the time needed typically to get to the same starting point on a real hardware platform.

In addition to hardware features and improvements, an optimized software stack: Intel® Data Plane Development Kit (DPDK) has been developed for Crystal Forest based designs and follow-on product lines to further boost the performance of data plane applications on general purpose CPU. Intel® DPDK provides a building block for customers to develop data plane solutions on Intel® Architecture with substantial
performance benefit. Commercial solutions with Intel® DPDK built-in are also available to ensure maximum of flexibility and meet demands from potential diverse usage models.

Traditionally, software development and test work begins when hardware becomes available. Simics, a full system functional simulation technology, completely changes that paradigm. With Simics, developers get to work on virtual systems that reflect accurate hardware behaviors and can use the environments to eliminate defects early and help all phases of product life cycle. Intel and Wind River® collaborated and provided simulation support for Crystal Forest platform (Ref [1]). This provides customers with a powerful design weapon that can further reduce time to market and improve quality. This new virtual platform provides developers unprecedented access and is ideal for developing complicated communication solutions.

This white paper discusses how to run Intel® DPDK software unmodified in a Simics Crystal forest environment. It describes the steps needed to run the Intel DPDK software stack on virtual platforms and compares the time needed typically to get to the same starting point on a real hardware platform. The paper also touches on potential usage models and benefits to customers.

The Intel® Embedded Design Center provides qualified developers with web-based access to technical resources. Access Intel Confidential design materials, step-by step guidance, application reference solutions, training, Intel’s tool loaner program, and connect with an e-help desk and the embedded community. Design Fast. Design Smart. Get started today.

Simulate Intel’s Next generation communication platform data plane solution

Crystal Forest - next generation communication platform

As more cores, features, and I/O are integrated into the CPU, some of the tasks that traditionally require multiple platforms can now co-exist in the same CPU. This means reduced design cost, easier software configuration management, and faster time to market. Intel’s Crystal Forest platform (Ref [1]) is one of Intel's next generation communication platforms that is capable of consolidating application processing, control processing and data processing tasks on a single CPU architecture.

For customers, this means significant cost-saving. Now they can support multiple workloads all with a single architecture and a consistent programming model. More focus can be dedicated to their unique value-adds and usage scenario. Working with Intel® Architecture, the customers know that the features and architecture are consistent and backward compatible, and their solutions will be supported and sustained many years down the road.

An example of Crystal Forest platforms, the Crystal Forest ATCA board is shown in Figure 1. The Intel® Xeon® processor based CPU can scale from 1 core to 16 core based on performance requirements. It also comes with 10G Ethernet capabilities with advanced virtualization technology built in. The Crystal Forest platforms adopt the same cores that are used for server, desktop and laptops however enhance requirement so they meet unique demanding environment for embedded applications. Tools and software solutions can be leveraged for embedded system development. The Crystal Forest platform is the first new IA platform to announce Simics support after VirtuTech acquisition (Ref [1]).
Multiple chipset versions are available so customers can make the selection based on their unique needs on I/O, speed, and also security features. Most of the chipsets have integrated Intel® QuickAssist technology, which processes and accelerates specialized packet workloads - cryptography, compression and deep packet inspection included - on standard Intel platforms. Using this technology, secure Internet transactions can be accelerated up to 100Gbps on the platform to give service providers the ability to handle many more secure transactions and without the cost of specialized solutions. The network will also be able to evolve to provide "always-on" secure Internet connections, as opposed to the opt-in connections currently used on select applications or for financial transactions online.

**Intel® DPDK: jump-start data plane solution development**

A general purpose operating system such a Linux* is designed to support a vast set of applications. The decisions and choices often do not favor a particular workload. Intel® Data plane Development kit (DPDK) is a set of low level software libraries and algorithms that improve the performance and throughput of packets on Intel® architecture platforms. It offers a single software programming model that can be used for the current generation of IA CPUs and next generations as well. It provides a set of programming interfaces (APIs) that can be used from the Linux* user space, offering low-overhead alternatives. The library provides a set of building blocks that customers can use to create or enhance their data plane solutions. There are multiple ways to use Intel DPDK, for example, it can be used as a standalone
solution for integration with customers applications or as part of commercial data plane solutions from proved ecosystem partners.

This is specifically suited for high speed networking and data I/O intensive applications. On the Crystal Forest platform with Intel DPDK, it delivers 5 times more performance than previous generation, for example, it is capable of 160 million packets per second for L3 forwarding running on IA. The performance will continue to leap as IA tick-tocks its architecture and process technology. Customers will enjoy these performance and feature increases without losing any of the benefits in backward compatibility, consistency in instruction sets and tool chains and a strong ecosystem support.

Intel DPDK offers a set of data plane libraries such as buffer management, queue management, packet flow classification and poll mode drivers (Figure 2). It also provides an Environmental Abstraction Layer (EAL) as well as a Bare Metal Execution layer.

**Figure 2 Intel DPDK services**

![Diagram of Intel DPDK services]

Intel DPDK is a set of libraries that can be used to optimize or improve performance over traditional & general purpose Linux. For example, it provides a scheme to remove or reduce performance issues commonly associated with interrupt handler penalty, context switching, data copying and Linux scheduler. These areas may be acceptable for generate purpose transactions but can be an issue for data I/O intensive workload when we start to talk about 10G/40G workload.

Intel DPDK can be used as a tool kit and users can integrate desired libraries, functions, and techniques into their unique data plane solutions. Multiple
vendors now offer integrated DPDK services and commercial data plane software stacks. These further reduce time to market so customers can focus on their key value-add services.

**Simics* Overview**

Developers can accelerate software development, testing and integration by utilizing a simulation model of the Crystal Forest platform provided by Wind River Simics*. With Simics, users can model any Crystal Forest target configuration and then run unmodified target software on that model. Wind River Simics enables developers to do BIOS bring-up, operating-system optimization and application development more efficiently.

Wind River Simics* provides full system simulation solutions for Intel’s Crystal Forest reference platforms and customer-specific situations that may even contain hundreds of boards in large networked systems. With Simics, OEMs can run the exact target software as runs on the physical board from BIOS to application. Several examples of Simics Crystal Forest benefit are highlighted in Figure 3. As Simics is a full product life cycle technology, it has many other benefits.

**Figure 3 Simics* Crystal Forest Benefits**

<table>
<thead>
<tr>
<th>Run Unmodified Software</th>
<th>Simulate an entire rack or network</th>
<th>Analyze new technology on legacy system</th>
<th>Decouple 5G dev. from hardware availability</th>
<th>Build OEM-specific virtual platforms</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Run exact same software as runs on physical boards</td>
<td>- Easily scalable</td>
<td>- Prototype and study the impact</td>
<td>- Pre-silicon, post launch, sustaining</td>
<td>- Extend OEM-specific modules or features</td>
</tr>
</tbody>
</table>

Figure 4 shows a sample screen shot of a Crystal Forest networking demo. Using Simics* technology, users not only can create a network between devices, but also can use tools to instrument the network and inject errors for test purposes. This feature enables testing the network in simulation and studying the behavior of the network.
Simulation of Intel’s Next generation communication platform data plane solution

Figure 4 Simics Crystal Forest networking demo

Installation of Simics* Crystal Forest

Step 1: Apply for Simics Crystal Forest Evaluation

Wind River Simics*: Crystal Forest evaluation packages and license can be requested at http://windriver.com/products/simics/simics_intel_architecture/. Please follow “Request your Crystal Forest Evaluation” link on the webpage. Please provide the information as requested by the evaluation form.

Simics Intel platform models are available for the following host platforms:

- Linux32
- Linux64
- Windows32
- Windows64

It is advised to install DPDK demo setup on a host machine with at least 4(8 recommended) Gb of RAM and 500 MB disk space available.

Step 2: Installing Simics packages

Once your evaluation request is approved, you will be provided with Simics packages.
The Wind River Simics installation contains a User Manual providing detailed instructions on Simics usage and basic configuration. As a result of the installation and setup process workspace is created with a basic directory structure in it.

**Step 3: Download Intel® Data Plane Development Kit**


The downloaded Wind River Linux disk image with DPDK installed (WRL43.CRF-DPDK-1GB-INITRD-1.img) must be put to the workspace directory.

The workspace contains several guest machines setup scripts in the targets/CrystalForest directory.

**Step 4: Preconfigured setup scripts, you are ready for Simulating DPDK!**

The demo package contains several pre-configured scripts users can try out to jumpstart the process:

- crystal-forest-wrl-dpdk.simics – single machine setup script
- crystal-forest-wrl-dpdk-001-dual-boot.simics - booting dual board setup with networking enabled
- crystal-forest-wrl-dpdk-002-setup.simics - script executing packet generation setup
- crystal-forest-wrl-dpdk-003-run.simics - script actually executing DPDK packets routing

The machines used in this whitepaper were pre-set according to DPDK evaluation guidelines having 10GigE 82599 plugged in to the North cluster PCI Express ports.

**Networking with 10GigE 82599**

Basic communication function can be simulated by opening crystal-forest-wrl-dpdk-001-dual-boot.simics machine script form Simics GUI (File->New Session From Script).

Once this script is selected simulated machines will be loaded and ready for execution.
This configuration contains two 4 core x 2 threads Crystal Forest boards with two Ethernet ports connected through Ethernet switches. Once simulation starts two boards will start booting. It may take some time to expand the compressed file system used in the demo.

As soon as boot process is done the full system state is saved to the Simics checkpoint directory for prompt returning to this state later on.

Simics provides flexible capabilities to connect the real network from the host machines to the guest and vice versa. A detailed network configuration guide is available within Simics framework.

To provide network infrastructure to the simulated environment, Simics contains an abstraction called “service node”. The service node is an internal virtual server which provides basic network capabilities such as DHCP, FTP server, NAT, etc... Demo configuration contains two service nodes (eth_servicenode_0, eth_servicenode_1), one for each subnet. We are going to demonstrate two use cases – bringing up DHCP server in the virtual network to enable guest machines communications and providing NAT service for guest machines connecting Internet.

Enabling the DHCP server at the service node:

1. Open Simics command line window(Menu->Tools->Command Line Window)
2. type eth_servicenode_0.dhcp-add-pool pool-size = 10 ip = "10.10.0.2"
3. Restart the network on the both guest machines: /etc/init.d/network restart

4. The hosts should get IP addresses 10.10.0.2 and 10.10.0.1 respectively

5. Internal network is up and running

Figure 6 Simics Crystal Forest network connection enabled

Connecting to the internet from a simulated virtual network:

1. Open Simics command line window (Menu->Tools->Command Line Window)

2. Type connect-real-network 10.10.0.2 ethernet-link = eth_switch_0

3. Now we are ready to browse the internet
Running unmodified Intel® DPDK on Simics Crystal Forest

The Intel® Data Plane Development Kit evaluation image configures two guest machines. One of them is running the DPDK binary while the other runs the PKTGEN binary. The machine state saved in the previous section is used. Setup of both systems is done according "GettingStartedGuide_WR Linux_DI PKD_CRF_Romley.pdf" covering real board setup and software installation.

In Simics Crystal Forest necessary console input is scripted using commands of Simics internal interpreter, so no additional manual typing is necessary. Two machine scripts are included in the Crystal Forest package to pre-configure and run the simulation:

1. Run crystal-forest-wrl-dpdk-002-setup.simics machine script from the Simics GUI (File->New Session From Script). This script will run the necessary binaries and configure packets forwarding on the simulated machine "crystal". The full system state will be saved right after setup is done.

2. Run crystal-forest-wrl-dpdk-003-run.simics machine script from the Simics GUI (File->New Session From Script). This script will load a previously saved state and start packet generation with the specified parameters. In this whitepaper 64 bytes packets are used for data exchange between machines.

3. In order to visualize PKTGEN statistics type “screen on” command at the serial console of simulated machine “light”.

Figure 7 Simics Crystal Forest connected to the Internet
Debug examples

The Wind River Simics framework provides network debugging capabilities similar to actual hardware: capturing network traffic, running gdb, etc.

In addition to this set several unique debugging options exist at the virtual platform simulator.

- Synchronous machines stop. Multi machine configuration debugging is always a tricky task because machines act independently and their sequence of events is not synchronized. Once an event appears on one of the machines it is impossible to stop and analyze the state of the rest system. That is not the case with the Simics environment. In the Simics framework all the
machines are stopped at once, by either setting a breakpoint or pressing the “pause” button in the Simics Control window.

- Full system state non-intrusive inspection. The Simics framework allows users to inspect all registers of all devices available in the simulated platform (Menu->Debug->Device Registers), CPU registers of all cores for each simulated machine (Menu->Debug->CPU Registers), and data stored in the system DRAM memory (Menu -> Debug -> Memory Contents).

A complete list of Wind River Simics features boosting software debugging is included in the User Guides available with the Simics base package.

Q & As

How to get access to Simics Crystal Forest?

The easiest way for enterprise users to access Simics Crystal Forest is to contact Wind River Simics (Ref [5]). Follow the “Request Evaluation for Crystal Forest” link and steps provided by the website.

Where to access Intel DPDK?

At the time this white paper is published, the standard way to get access to DPDK source code is to contact your Intel sales representatives. This website contains some reference documents for DPDK:

Wind River Linux and Intel® Data Plane Development Kit evaluation are available for download from Intel: http://downloadcenter.intel.com/Detail_Desc.aspx?agr=Y&DwnIdID=20305

Where to get commercial Solutions for DPDK

At the time this white paper is published, the list of companies that provide commercial support for integrated DPDK include: Wind River*, Tieto*, 6WIND* (Ref [10]). The list is expected to grow as more and more customers find usage models to take advantage of data plane solutions.

Conclusions

In this white paper, we announced that a full data plane software stack (Intel® DPDK) can run in the Simics IA environment. This is an excellent example of the capability of Simics. Using simulation, users can jump start their learning process for a software stack such as DPDK and then start focus on new software and features that are specific to their application and usage models.
Intel and Wind River have come up with the full system simulation technology that supports next generation communication platform code named Crystal Forest. With the diversity of the platform, it is quite possible that users will use the real platforms as well as the simulation technology to their advantage and speed up the time to market and provide solutions to end users faster. If you are considering or using Crystal Forest platforms, or future generation platforms, it is time to utilize this new weapon in Simics to help your product development.

References

1. Intel’s Next generation Communication Platform Code Named Crystal Forest,
   http://newsroom.intel.com/community/intel_newsroom/blog/2012/02/14/intels-next-generation-communications-platform-key-to-accelerated-network-services

2. White paper: Crystal Forest platform,

3. Press Briefing: Crystal Forest platform,

4. Packet processing on IA,

5. Wind River* Simics* for Intel Architecture with Eval link for Crystal Forest,
   http://www.windriver.com/products/simics/simics_intel_architecture/

6. Wind River* Simics* for Intel Architecture Overview,

7. White Paper: Improving the Embedded IA design process with Simics*,
   http://download.intel.com/design/intarch/papers/326341.pdf?iid=5715

8. Intel’s Simics* Website,

9. Jakob Engblom, Crystal Forest on Simics,
   http://blogs.windriver.com/engblom/2012/03/crystal-forest-on-simics.html

10. Software tool support for DPDK,

**Authors**

**Tian Tian** is a Senior Platform Architecture Engineer for the Communications and Storage Group at Intel Corporation. Tian is involved with next generation of IA communication product lines and leads Simics GTM effort.

**Alexander Belousov** is an Engineering Manager for Software Service Group, Simics Technology Center at Intel Corporation. Alex leads a team focusing on simulation solution for IA CPUs.
Simulate Intel's Next generation communication platform data plane solution

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.

A "Mission Critical Application" is any application in which failure of the Intel Product could result, directly or indirectly, in personal injury or death. SHOULD YOU PURCHASE OR USE INTEL’S PRODUCTS FOR ANY SUCH MISSION CRITICAL APPLICATION, YOU SHALL INDEMNIFY AND HOLD INTEL AND ITS SUBSIDIARIES, SUBCONTRACTORS AND AFFILIATES, AND THE DIRECTORS, OFFICERS, AND EMPLOYEES OF EACH, HARMLESS AGAINST ALL CLAIMS COSTS, DAMAGES, AND EXPENSES AND REASONABLE ATTORNEYS' FEES ARISING OUT OF, DIRECTLY OR INDIRECTLY, ANY CLAIM OF PRODUCT LIABILITY, PERSONAL INJURY, OR DEATH ARISING IN ANY WAY OUT OF SUCH MISSION CRITICAL APPLICATION, WHETHER OR NOT INTEL OR ITS SUBCONTRACTOR WAS NEGLIGENCE IN THE DESIGN, MANUFACTURE, OR WARNING OF THE INTEL PRODUCT OR ANY OF ITS PARTS.

Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined". Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

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.

Copies of documents which have an order number and are referenced in this document, or other Intel literature, may be obtained by calling 1-800-548-4725, or go to:
http://www.intel.com/design/literature.htm%20

Intel, and the Intel logo 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 © 2012 Intel Corporation. All rights reserved.