JULY 11-13, 2017
BUDAPEST, HUNGARY

Xen Project

DEVELOPER AND DESIGN SUMMIT

Xen Project
DEVELOPER AND DESIGN SUMMIT

JULY 11-13, 2017
BUDAPEST, HUNGARY

How to abstract hardware acceleration device in cloud environment

Maciej Grochowski
Intel DCG Ireland

Xen Project
DEVELOPER AND DESIGN SUMMIT

Outline

- Introduction to Hardware Accelerators
- Intel® QuickAssist Technology (Intel® QAT) as example of Accelerator
- Accelerator capacity and QoS
- Virtualizing an Accelerator
- Challenges with Accelerators in the Cloud
- Porting Intel® QAT to Xen*
- Future challenges

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

Hardware Accelerator

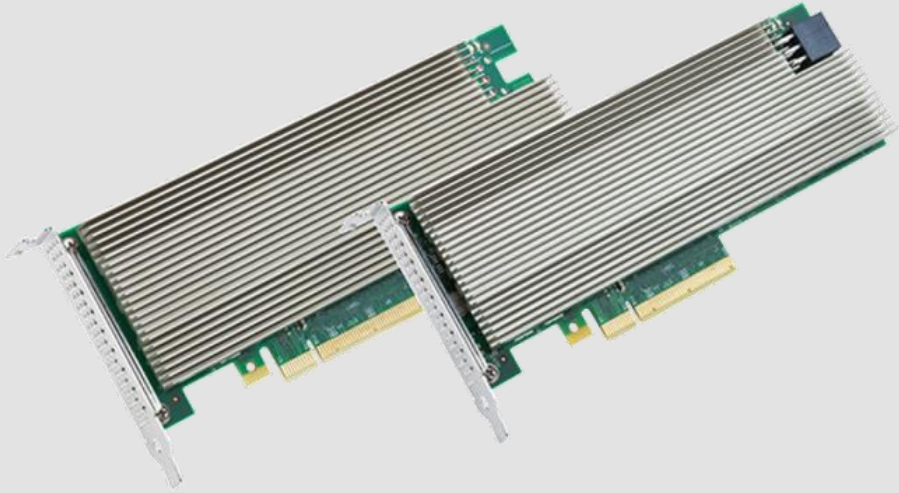
Focus is on fixed function accelerators exposed as PCI devices

Example is Intel® QAT, which offers acceleration for

- Cryptography (both symmetric and public key)
- Compression (lossless)

Comes in a few form factors

- PCI-Express* plugin card
- Integrated in platform (e.g. in chipset or SOC)



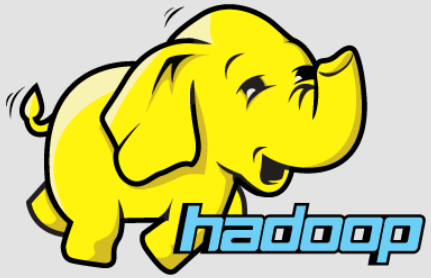
Other types of HW Accelerators (FPGA, ML Chips, vector processors) not considered here

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

Why Hardware Accelerators?

Acceleration is helpful when:

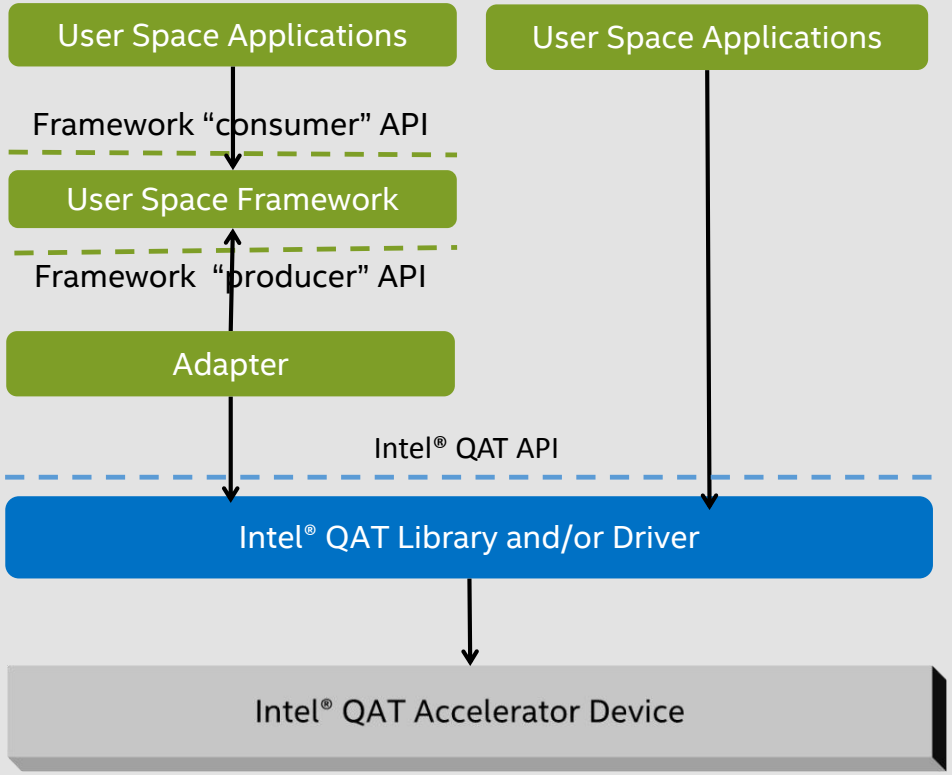
- Application is consuming multiple CPU cores performing compute-intensive workload, e.g. cryptography (TLS, IPsec), compression (Big Data, Storage)
- Need to scale vertically: in addition to CPU we can add acceleration resources to offload



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

Acceleration Device Software Stack

- Typical accelerators expose services to software via request/response rings or queues
- Multiple processes can access the device via different request rings
- Software abstracts these rings via an API
- This can also be plugged in to various cryptographic or compression frameworks



QoS Requirements

Resources to be shared include:

- Acceleration capacity
- Bandwidth to and from accelerator (PCIe upstream and downstream)

QoS Requirements

Each VM can be assigned:

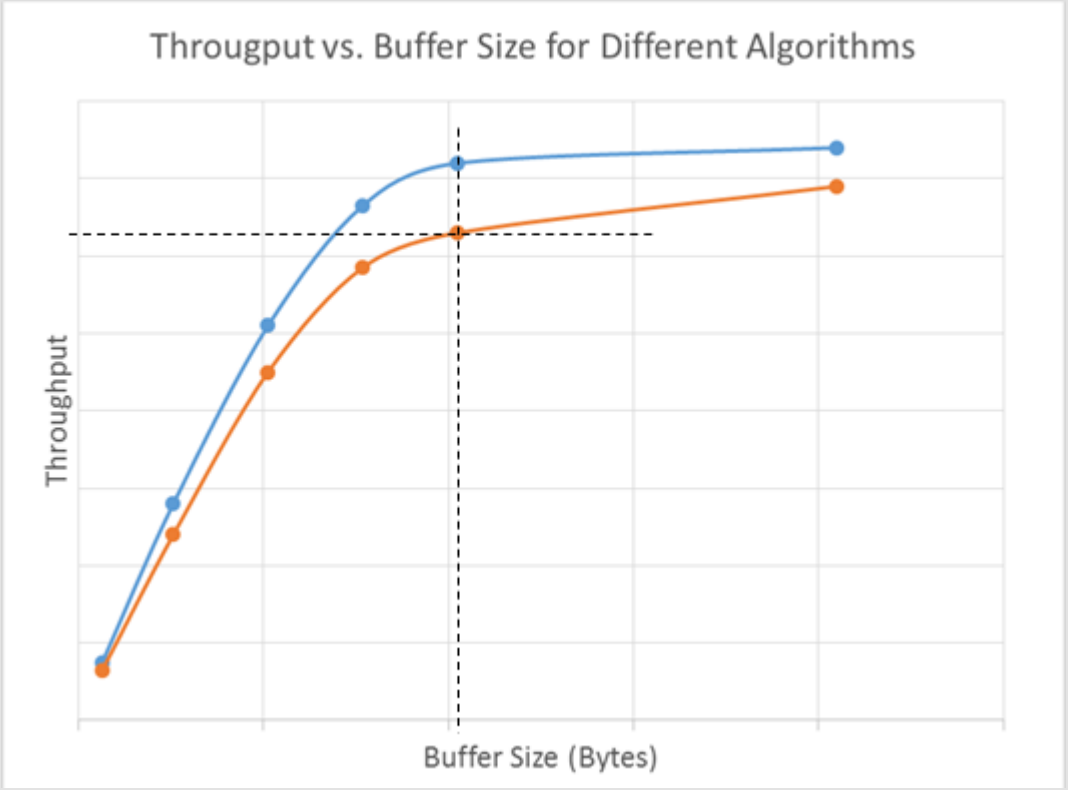
1. CIR (Committed Information Rate) (guarantee/minimum)
2. PIR (Peak Information Rate) (limit/maximum)

Typically:

- Cloud service users want to be guaranteed some minimum
- Cloud service providers wants to limit to some maximum

Defining Acceleration Capacity

- NICs typically measure QoS in terms of throughput/bandwidth
- For accelerators, this does not work well: throughput is highly dependent on algorithm, packet size, compression level, and other factors
- Instead we offer throughput for a specified **reference operation**
  - Specified algorithm, buffer size, compression ratio, etc.
  - Throughput for different algorithms, buffer sizes, etc. will be different

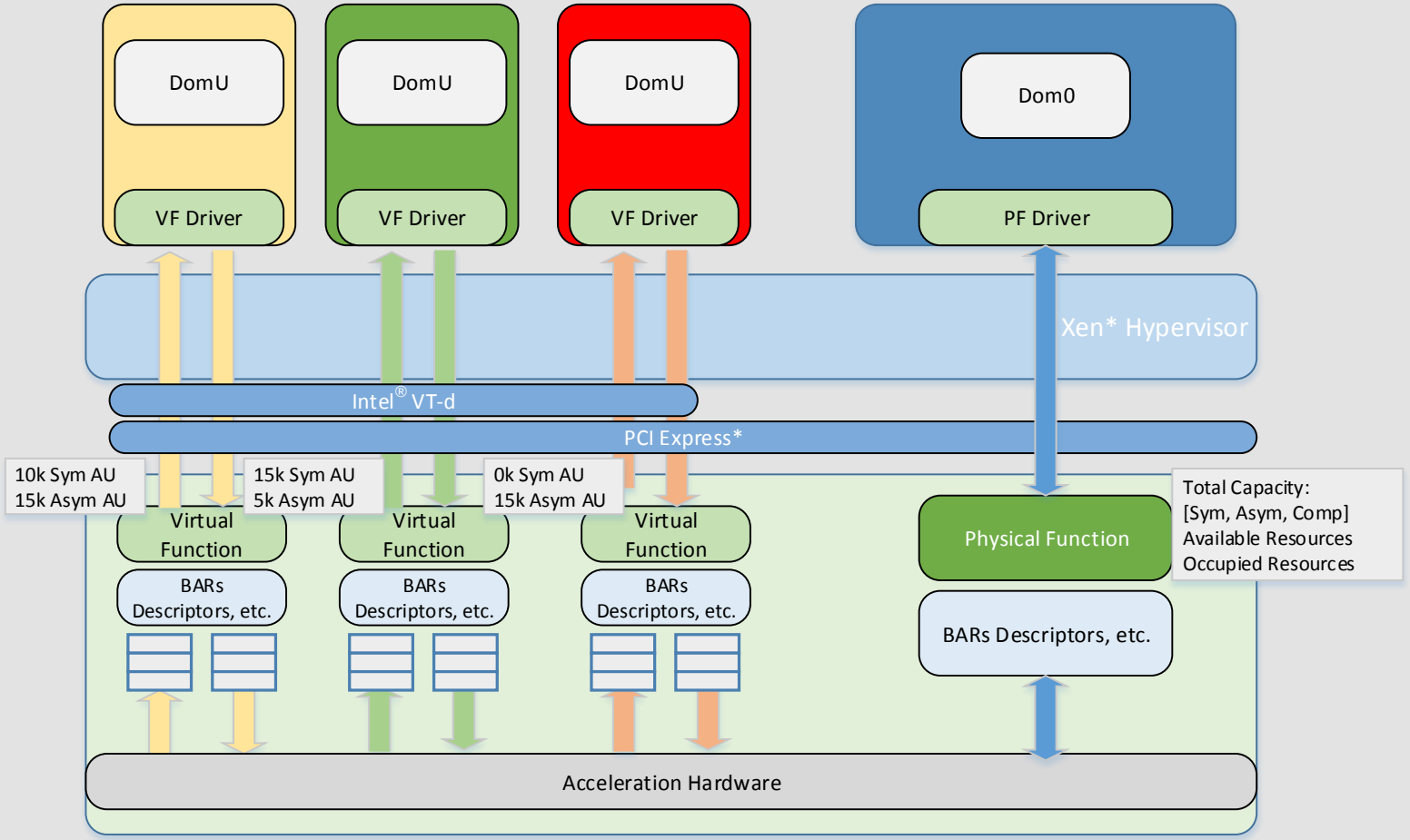


Acceleration Units

Throughput for specific reference operation we used to name Acceleration Unit

- Asymmetric crypto Acceleration unit:
  - 1ops of reference algorithm
  - Reference operation
    - Algorithm: RSA
    - Key Size: 2048 bit
    - Direction: Decrypt with CRT
- Symmetric crypto Acceleration unit:
  - 1Mbps of reference algorithm
  - Reference operation
    - Algorithm: AES-128-CBC + SHA256-HMAC
    - Buffer Size: 1024B
    - Direction: Encrypt and Generate MAC

Example: Acceleration Units Management



Virtualizing the Accelerator

	Hardware-Based (SR-IOV)	Software-Based (PV, emulation)
Performance	Optimal: guest can write directly to device	Adds overhead/offload cost (VM exits, adaptation, etc.)
Scalability	Limited by number of VFs/rings available in hardware	Unlimited
Portability	Requires device driver for specific device and vendor	Portable across devices and vendors

For NICs, cloud tends to prefer para-virtualization for portability
For an accelerator, performance tends to be the most important factor

Moving Accelerator Device to Cloud

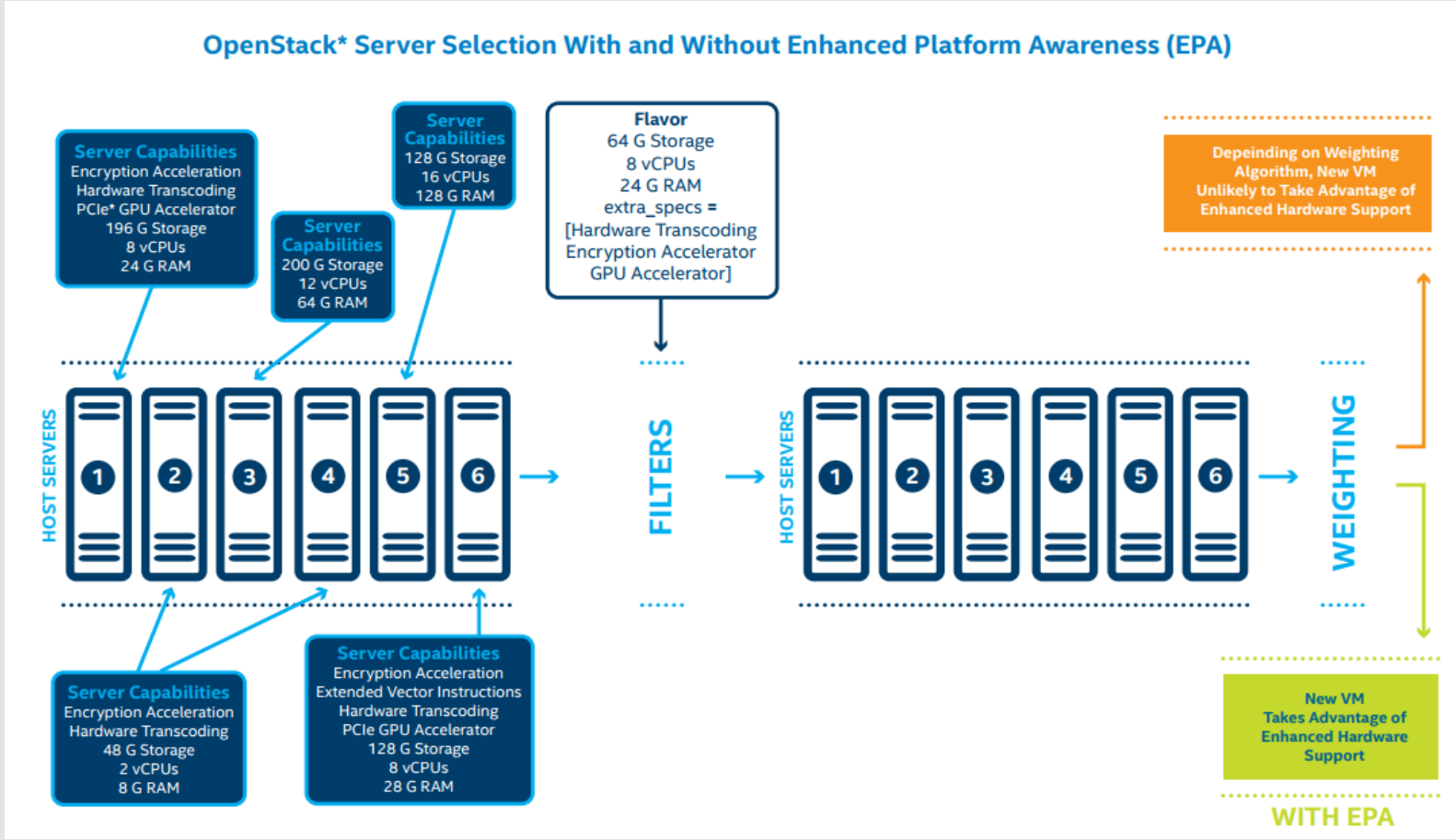
Cloud generally prefers generic “virtual device” (PV model)

For performance, trend is to provide dedicated hardware

- SSD drive vs. HDD
- NIC VFs (SR-IOV)
- Specific CPU model or capabilities

Intel is driving Enhanced Platform Awareness

- EPA makes capabilities visible
- Orchestrator or Virtual Infrastructure Manager can know capabilities during spawning VM



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

Porting Intel® QAT to Xen*

Virtualizing Acceleration Device

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

Moving Intel® QAT to Xen*

Interrupt Latency:

- Drivers which rely on mailboxes for communication between PF driver and VF driver incur higher latency
- Timeouts need to be tuned for Xen*

VF and PF drivers typically need to communicate

- Configuration information at startup and on certain events
- During normal operation to gather statistics

Future challenges

- PV driver for Accelerators
  - Abstract virtual device for crypto being defined, see virtio-crypto on qemu-devel mailing list
  - PV driver for SR-IOV to exchange statistics data
- Expose common information of capabilities in generic way
  - Define common way to manage SLA for acceleration devices

Summary

- Visible trend in the market of platform awareness
- Acceleration Devices can benefit from SR-IOV performance
- QoS need to consider throughput at a reference operation

Other Resources

Documentation and software:

https://01.org/intel-quickassist-technology

https://01.org/packet-processing/intel%C2%AE-quickassist-technology-drivers-and-patches

Intel® QAT chipset reference:

Intel.com/quickassist

http://ark.intel.com/products/80372/Intel-Communications-Chipset-8955

Openstack* Enhanced platform awareness:

https://01.org/sites/default/files/page/openstack-epa_wp_fin.pdf

Any other questions:

Maciej.Grochowski@intel.com

Legal Disclaimer

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. 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.

Intel may make changes to specifications, product descriptions, and plans at any time, without notice.

All dates provided are subject to change without notice.

Intel is a trademark of Intel Corporation in the U.S. and other countries.

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

Copyright © 2017, Intel Corporation. All rights are protected.

Q&A?