Live Migration of vGPU

Aug 2016

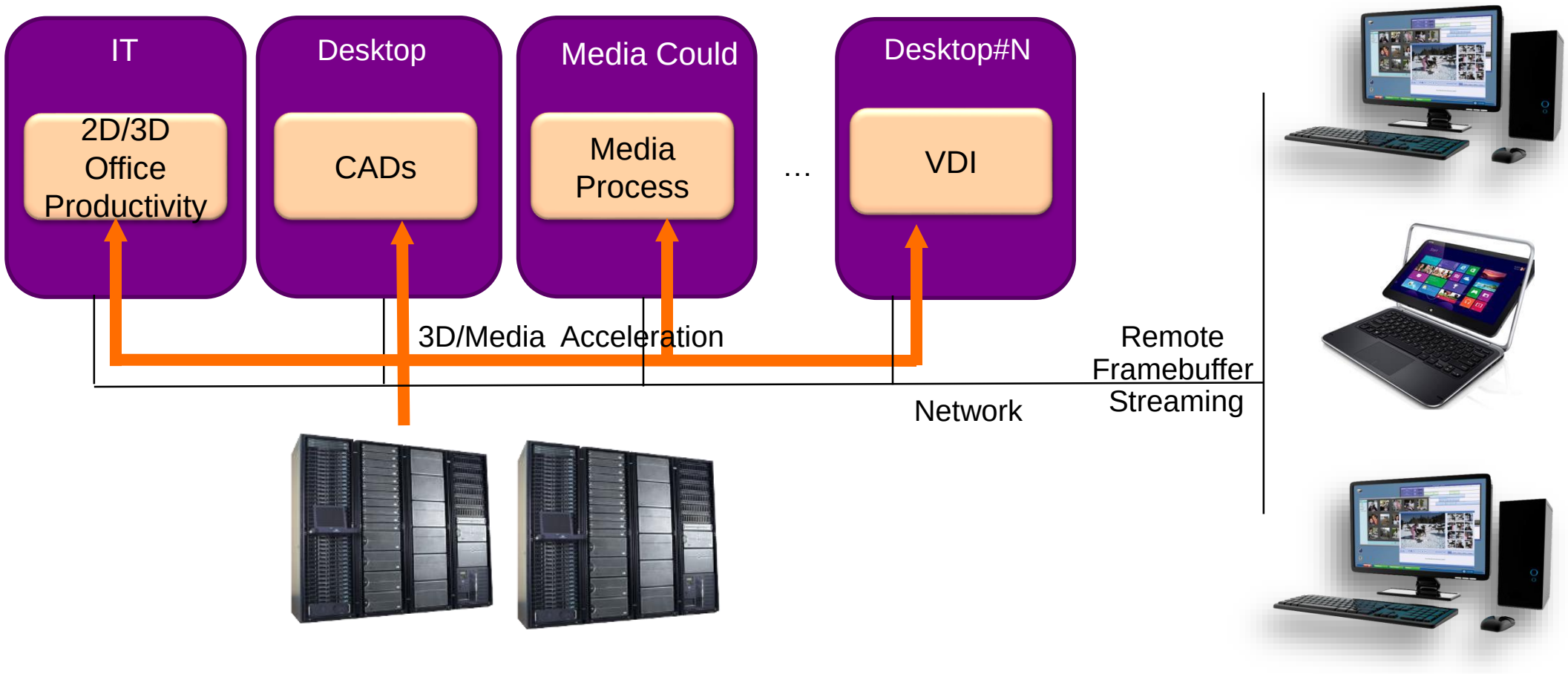
Xiao Zheng xiao.zheng@Intel.com

Kevin Tian kevin.tian@Intel.com

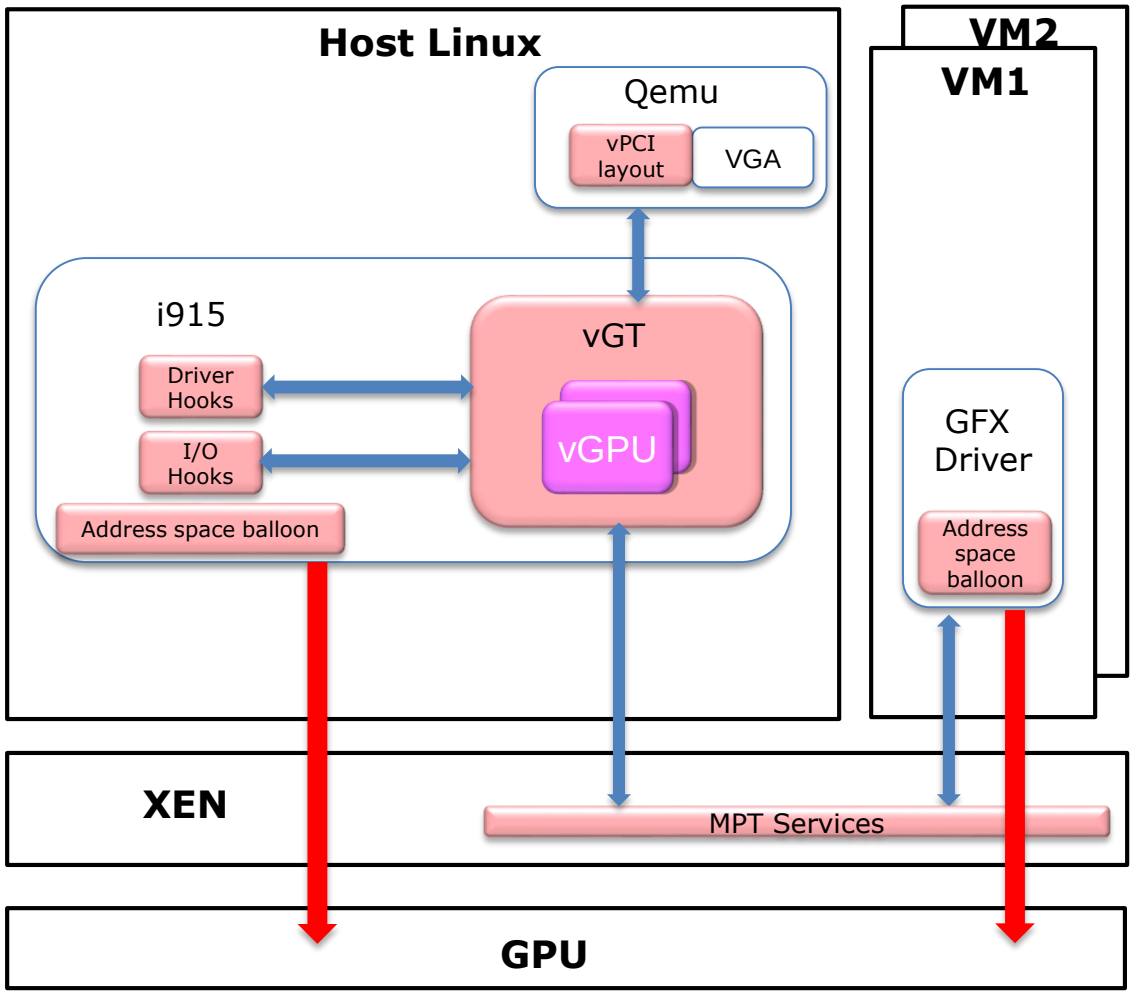
Agenda

- GPU Virtualization and vGPU Live Migration
- vGPU Resources
- Design and Solution
- Current Status
- Summary

GPU Virtualization Usage Cases



XENGT Architecture – Mediated Pass-through

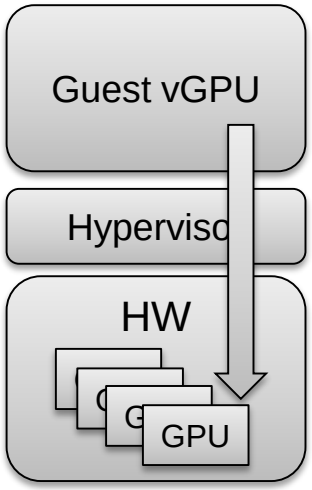


- pass-through for performance critical resource
- Trap and emulate for privileged resource
- Time-shared among VMs

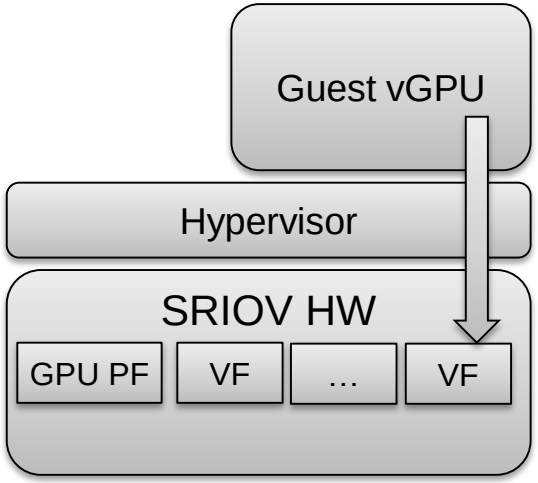
vGPU Live Migration

Live Migration: Load balance, Maintenance, Fault recovery
Unfortunately most of vGPU solutions do not support migration except GVT-g

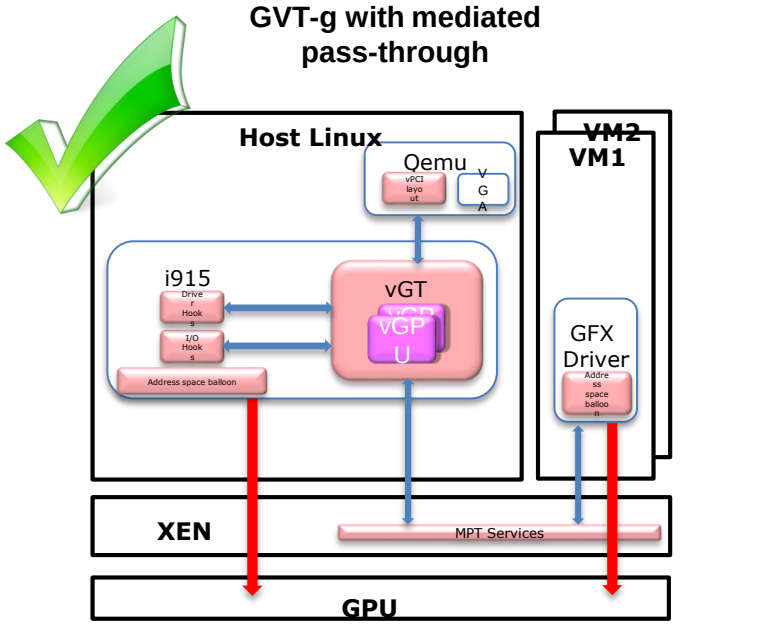
Typically a GPU pass-through solution



Typically a GPU SRIOV solution



GVT-g with mediated pass-through

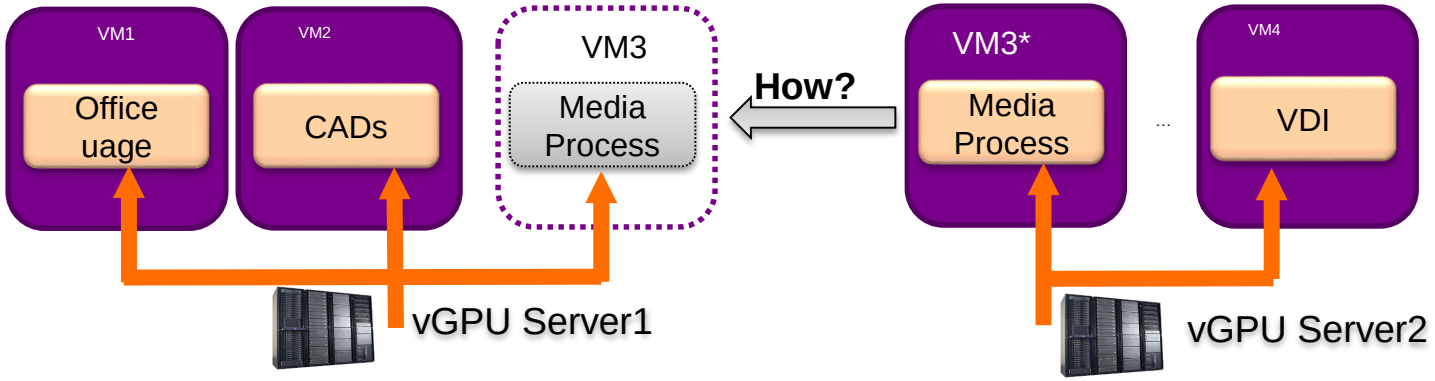


GVT-g architecture (Mediation) make it possible for seamless live migration

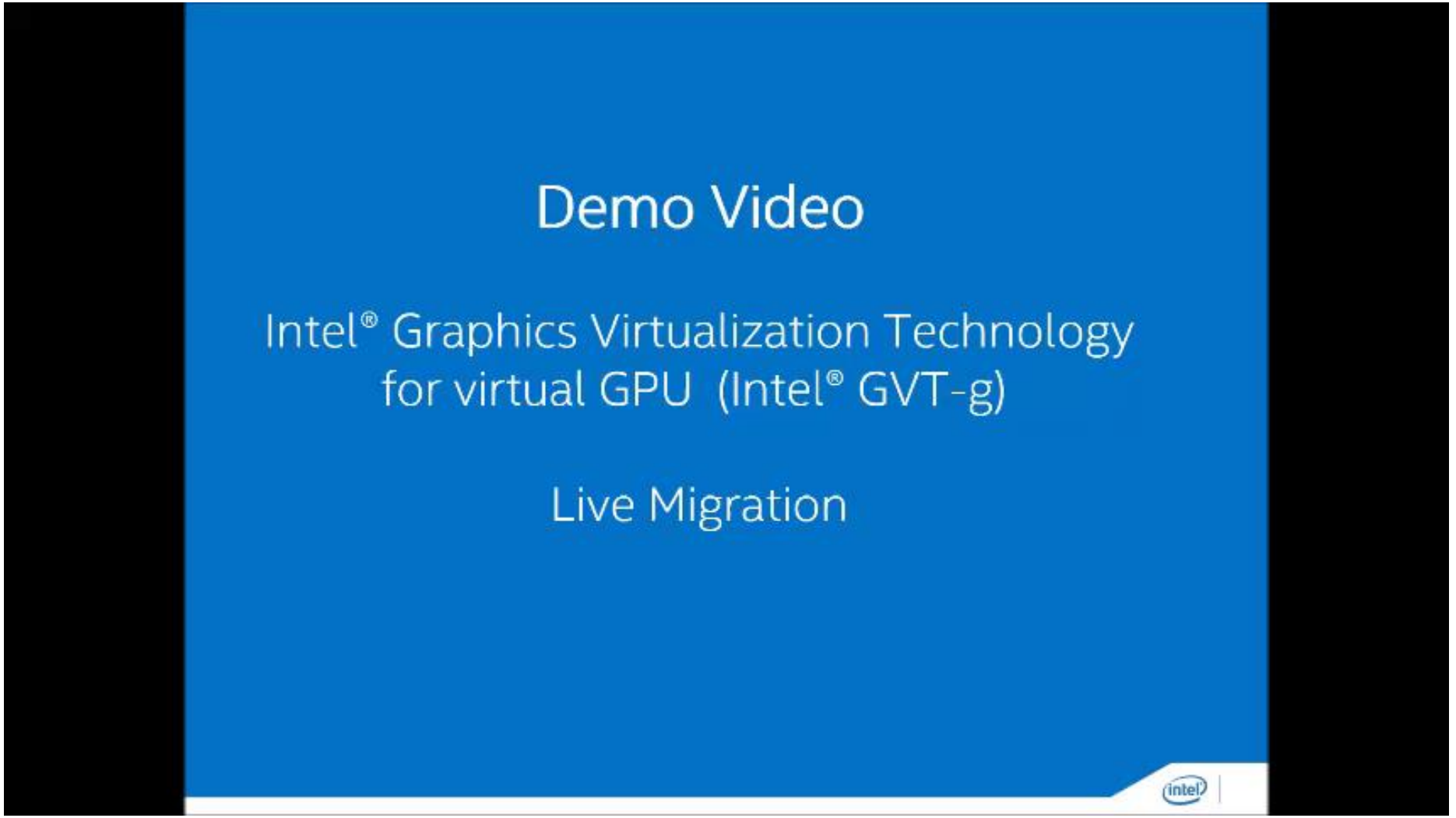
Live Migration of vGPU in GVT-g

Highlight feature:

- GVT-g is Open Source project, upstream ongoing
- vGPU Live Migration follows existing hypervisor migration flow
- 3D/2D/Media graphics workload seamless migrated between Servers or Local machine
- Support Linux/Windows Guest
- Live Migration Service downtime latency < 0.3 sec (Guest RAM 2GB, assigned 512MB vGPU memory, 10Gpbs adapter)



Demo: vGPU Live Migration with 3D workload



https://www.youtube.com/watch?v=y2SkU5JODIY

Agenda

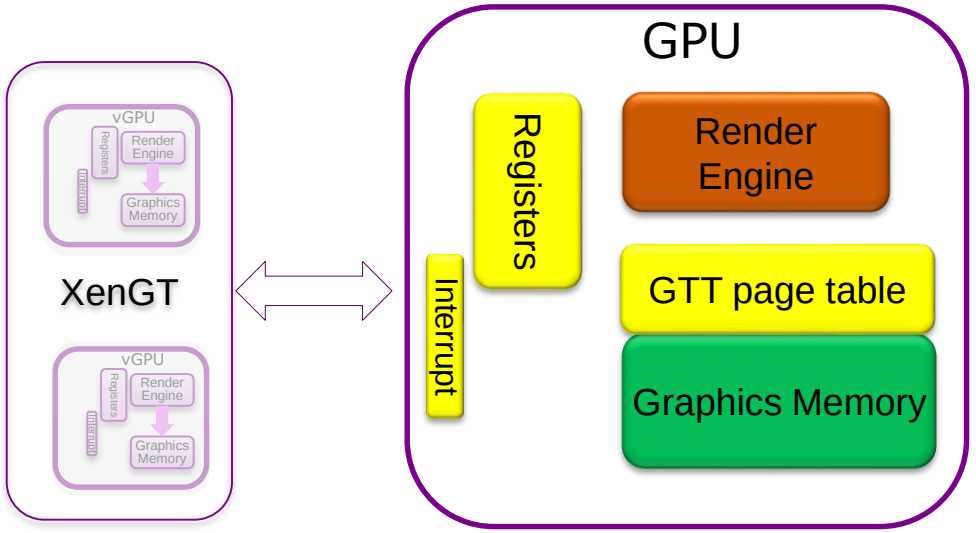
- GPU Virtualization and vGPU Live Migration
- vGPU Resources
- Design and Solution
- Current Status
- Summary

Inside of vGPU instance

pass-through for performance critical resource

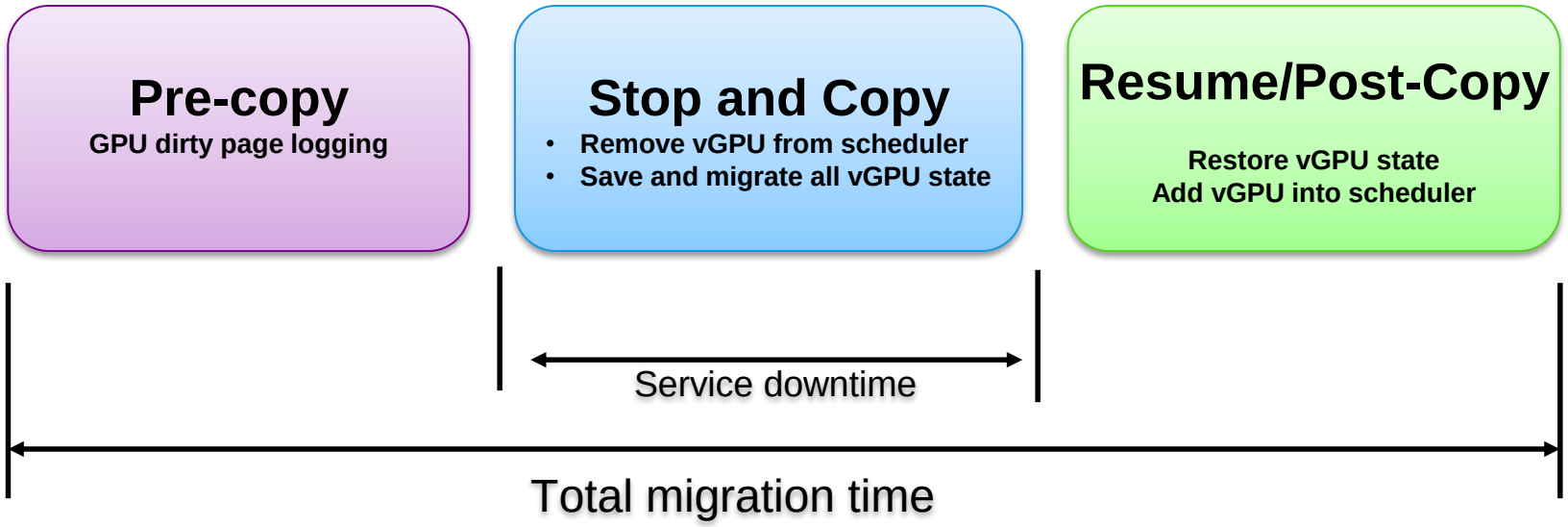
Trap and emulate for privileged resource

Time-shared among VMs

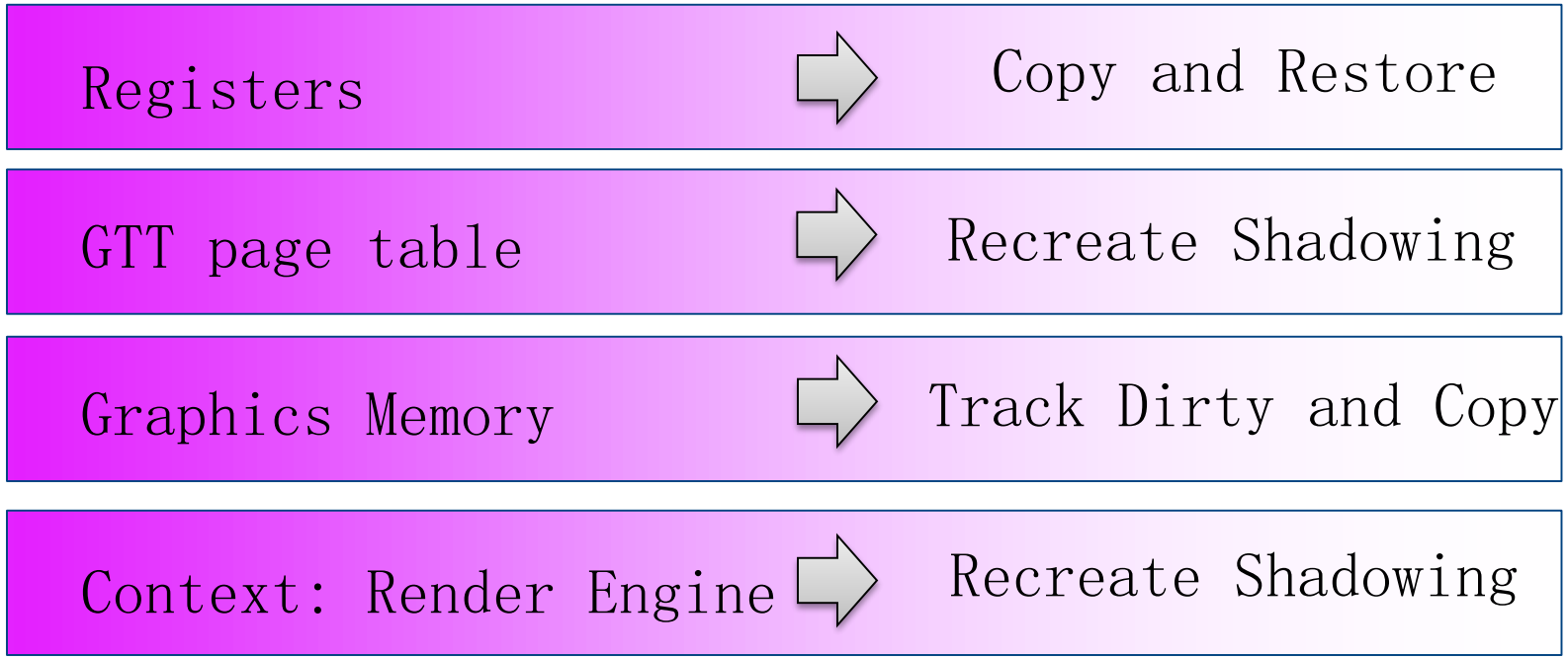


Challenge of Migrating vGPU Instance

- When and how to migrate Graphics Memory
- When and how to migrate Guest Graphics Page Table
- When and how to migrate Render Engine State



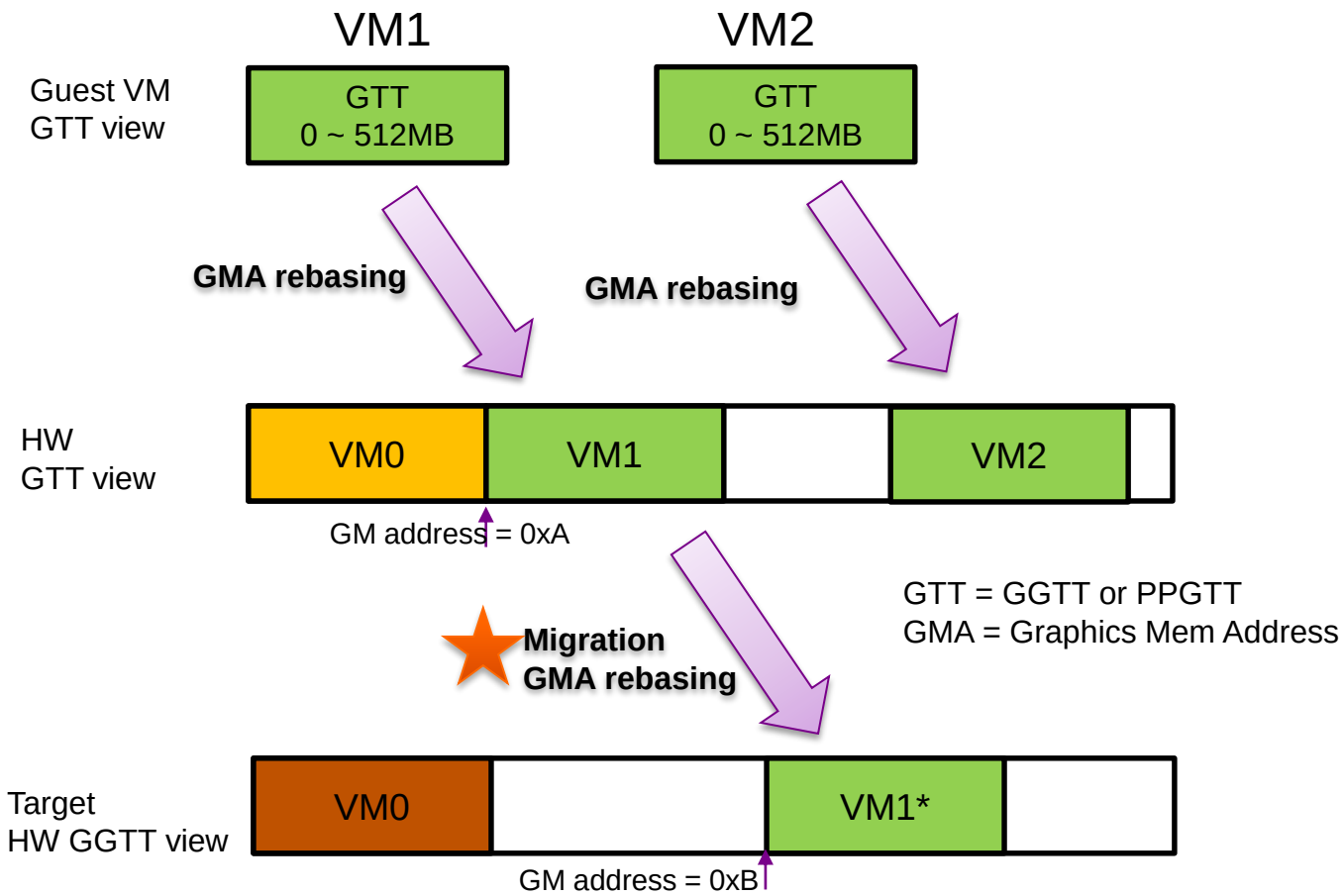
Migration Policies for Different vGPU Resources



Agenda

- GPU Virtualization and vGPU Live Migration
- vGPU Resources
- Design and Solution
- Current Status
- Summary

Guest GTT Page Table Migration



- Both GGTT and PPGTT are shadowed for Guest
- GGTT required rebasing due to GGTT partition among VMs
- Migration process actually:
  - A. Copy entire Guest GTT page table
  - B. Re-create the shadow page table for Guest on Target side
  - C. Rebasing GGTT for GPU commands

Graphics Memory Address rebasing:

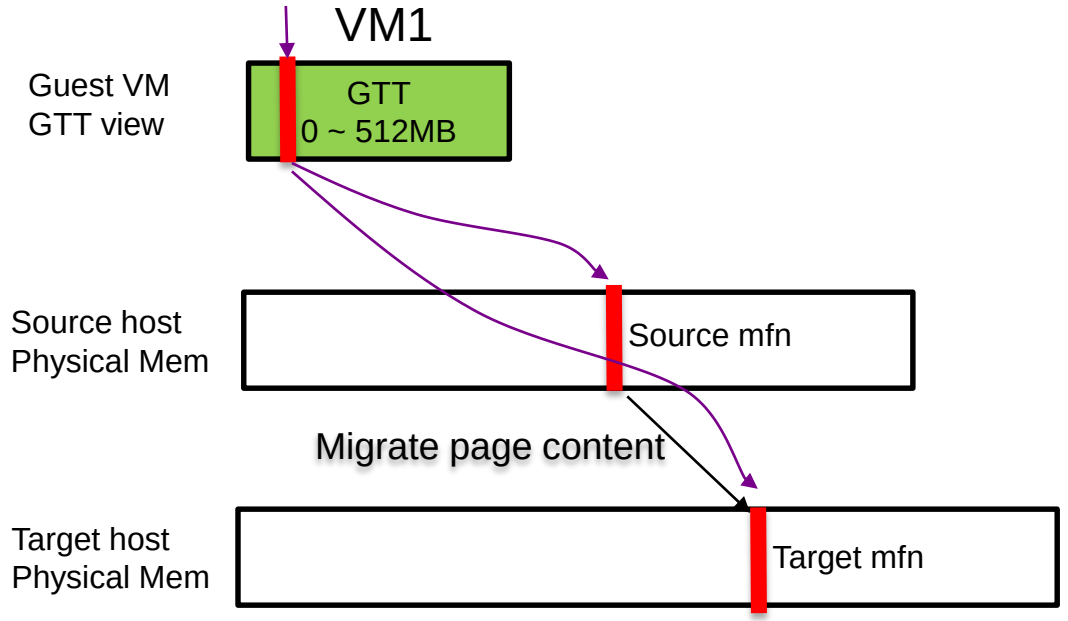
All vGPU cmds from Guest need to be rebased on new address in GVT-g before send to real GPU HW

Guest Graphics Memory Migration

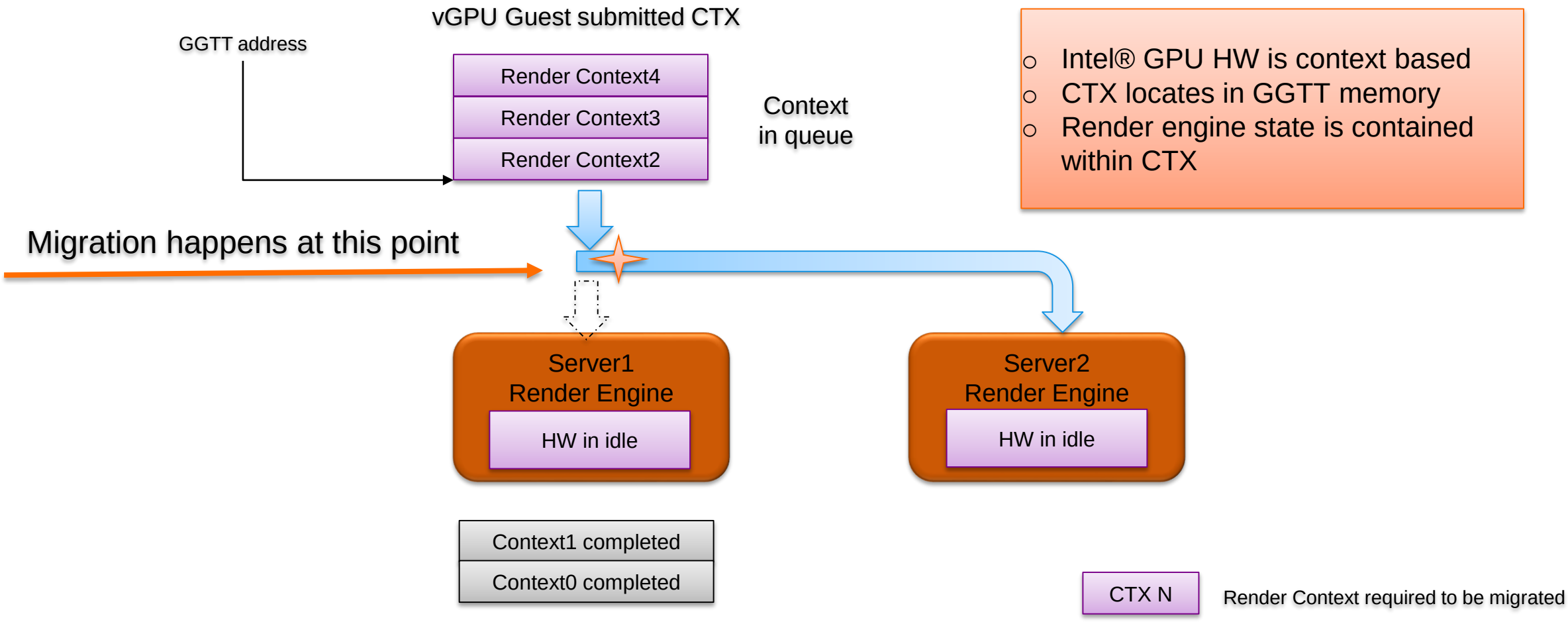
- **Pre-copy:** Logging dirty graphics memory pages
- **Stop-and-Copy:** Migrate contents to target
- **Resume/Post-copy:** Recreate GTT page table based on target mfn

Problem:
Intel® GPU page table entities has no Dirty or Accessed flags to track dirty pages

Solution:
Copy all used graphics memory to target.



Render Engine State Migration



- Intel® GPU HW is context based
- CTX locates in GGTT memory
- Render engine state is contained within CTX

Agenda

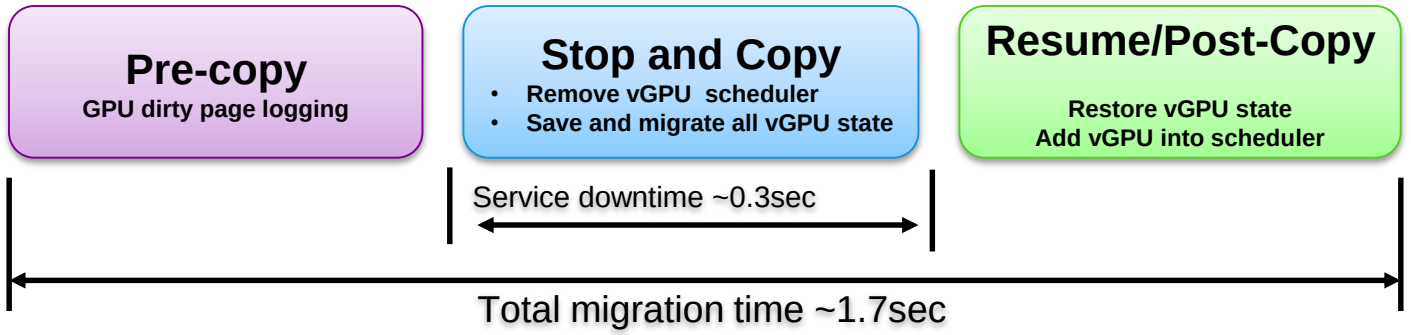
- GPU Virtualization and vGPU Live Migration
- vGPU Resources
- Design and Solution
- Current Status
- Summary

Current Status

- Experimental support both KVMGT and XENGT
- Platforms: Intel® 5th /6th Generation Intel® Core™ Processors
- Benchmarks covered:
  Windows guest: Heaven, 3Dmark06, Trophic, Media encoding/decoding, Linux guest: lightsmark, 2D
- Quality: 12hours overnight testing, migrating every 30sec
- Timing: (Guest RAM 2GB including 512MB Graphics memory, 10Gbps adapter)

  Service downtime ~0.3sec

  Total migration time: ~1.7sec



Summary

- Need 3D/2D/Media workload in virtualization?

  GVT-g is the choice

- Need GPU virtualization with migration support?

  GVT-g is the choice ☺

Resource Links

- Project webpage and release: https://01.org/igvt-g
- Project public papers and document: https://01.org/group/2230/documentation-list
- Intel® IDF: GVT-g in Media Cloud: https://01.org/sites/default/files/documentation/sz15_sfts002_100_engf.pdf
- XenGT introduction in summit in 2015: http://events.linuxfoundation.org/sites/events/files/slides/XenGT-Xen%20Summit-REWRITE%203RD%20v4.pdf
- XenGT introduction in summit in 2014: http://events.linuxfoundation.org/sites/events/files/slides/XenGT-LinuxCollaborationSummit-final_1.pdf

Notices and Disclaimers

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. INTEL PRODUCTS ARE NOT INTENDED FOR USE IN MEDICAL, LIFE SAVING, OR LIFE SUSTAINING APPLICATIONS.

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

All products, dates, and figures specified are preliminary based on current expectations, and are subject to change without notice.

Intel, processors, chipsets, and desktop boards 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.

No computer system can provide absolute security under all conditions. Intel® Trusted Execution Technology (Intel® TXT) requires a computer with Intel® Virtualization Technology, an Intel TXT-enabled processor, chipset, BIOS, Authenticated Code Modules and an Intel TXT-compatible measured launched environment (MLE). Intel TXT also requires the system to contain a TPM v1.s. For more information, visit http://www.intel.com/technology/security

Intel, Intel logo, Xeon, and Xeon Inside 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.