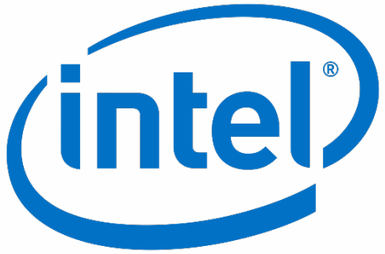
presented by



Compute Express Link (CXL) Update

UEFI 2020 Virtual Plugfest

May 19, 2020

Presented by Mahesh Natu (Intel) and Thanu Rangarajan (Arm)

Ack: Samer El-Haj-Mahmoud (Arm), Mike Rothman (Intel)

Meet the Presenters

Thanu Rangarajan
Principal Engineer
Member Company: Arm

Mahesh Natu
Principal Engineer, Data Platforms Group
Member Company: Intel

Agenda

- Introduction to CXL
- CXL Roadmap
- UEFI and ACPI Changes
- Summary and call to Action

Introduction to CXL

- Open industry standard for high bandwidth, low-latency coherent interconnect
- Connectivity between processor and accelerators/memory devices
- Addresses high-performance computational workloads across AI, ML, HPC, and Comms segments
  - Heterogeneous processing
  - Memory device connectivity
- Dynamic multiplexing of 3 protocols over PCI Express® (PCIe®) 5.0 Physical Layer
  - CXL.io – I/O semantics, similar to PCIe technology (mandatory)
    - CXL devices appear in PCIe config space, with additional register capabilities
  - CXL.cache – Caching Semantics (optional)
  - CXL.memory – Memory semantics (optional)

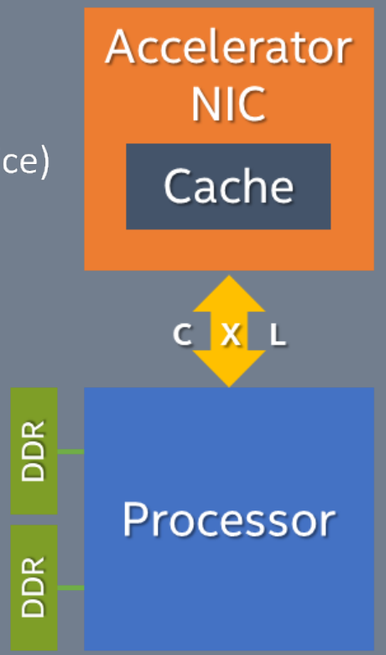
Representative CXL Usages

Caching Devices / Accelerators

Usages:
- PGAS NIC
- NIC atomics

Protocols:
- CXL.io
- CXL.cache

(Type 1 Device)

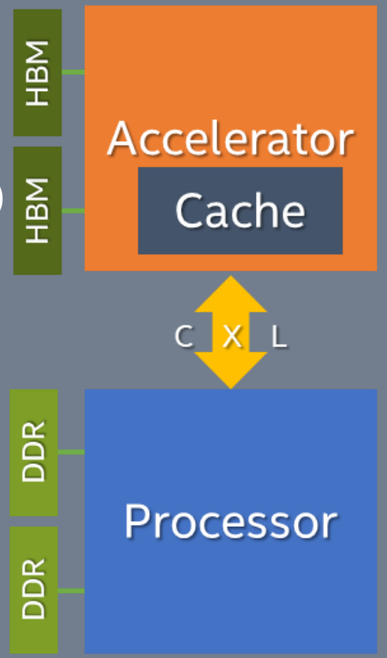


Accelerators with Memory

Usages:
- GPU
- FPGA
- Dense Computation

Protocols:
- CXL.io
- CXL.cache
- CXL.memory

(Type 2 Device)

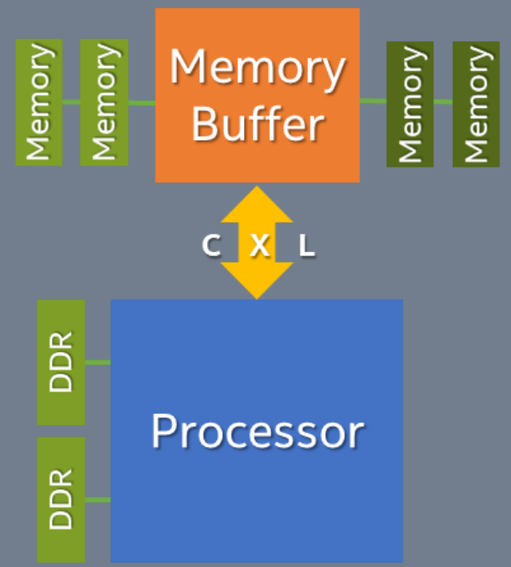


Memory Buffers

Usages:
- Memory BW expansion
- Memory capacity expansion
- 2LM

Protocols:
- CXL.io
- CXL.mem

(Type 3 Device)



CXL Roadmap

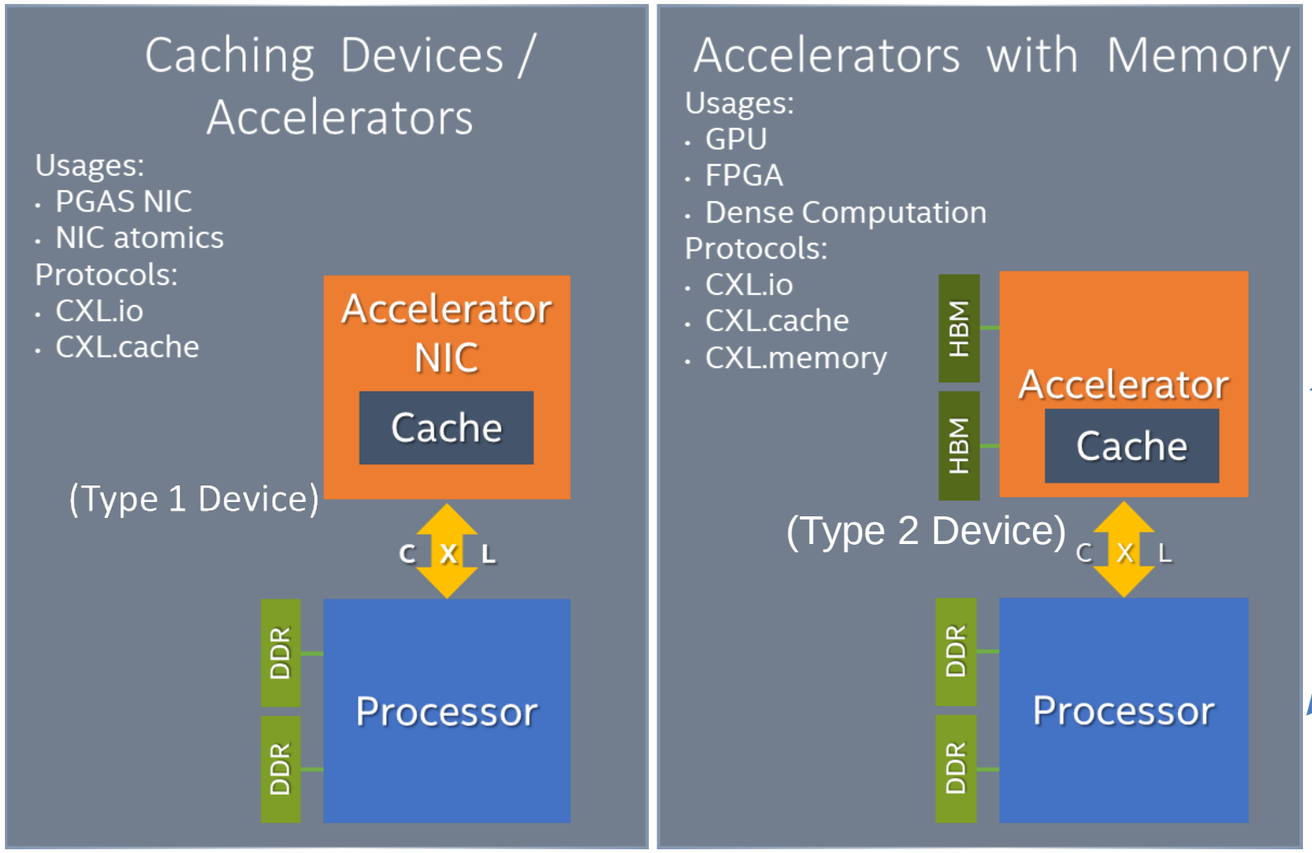
- CXL 1.1 specification available now
- CXL consortium is actively working on CXL 2.0
- CXL consortium has grown to 100+ members. If your company is not a member, consider joining.
- If your company is a member, consider joining various workgroups and contribute to future generation of CXL.
- https://computeexpresslink.org/

UEFI and ACPI Implications

Heterogeneous Computing Awareness

- Rethinking NUMA – hetero-memory and hetero-processors:
  - Generic Initiators
  - HMAT beyond heterogeneous memory
  - Coherent memory device characteristics and HMAT
- Redefining memory characteristics based on usage
  - Specific-purpose Memory (SPM)

ACPI Generic Initiator

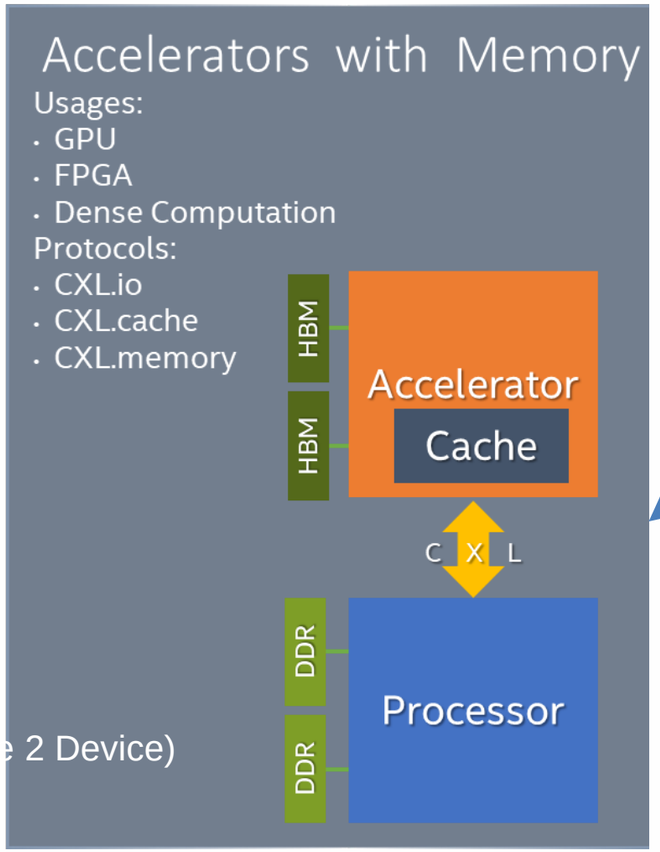


Initiators in NUMA Domains

- Non-processor initiators are classified as Generic Initiators
- Generic Initiators introduced in ACPI 6.3

UEFI Specific-Purpose Memory

- Is just regular EFI Conventional memory that has regular system memory behavior (i.e. WB, coherent, OS-managed)
- *Preferentially* used for acceleration or device-specific purposes
- Marked with the EFI_MEMORY_SP memory attribute in UEFI
- This memory attribute introduced in UEFI 2.8



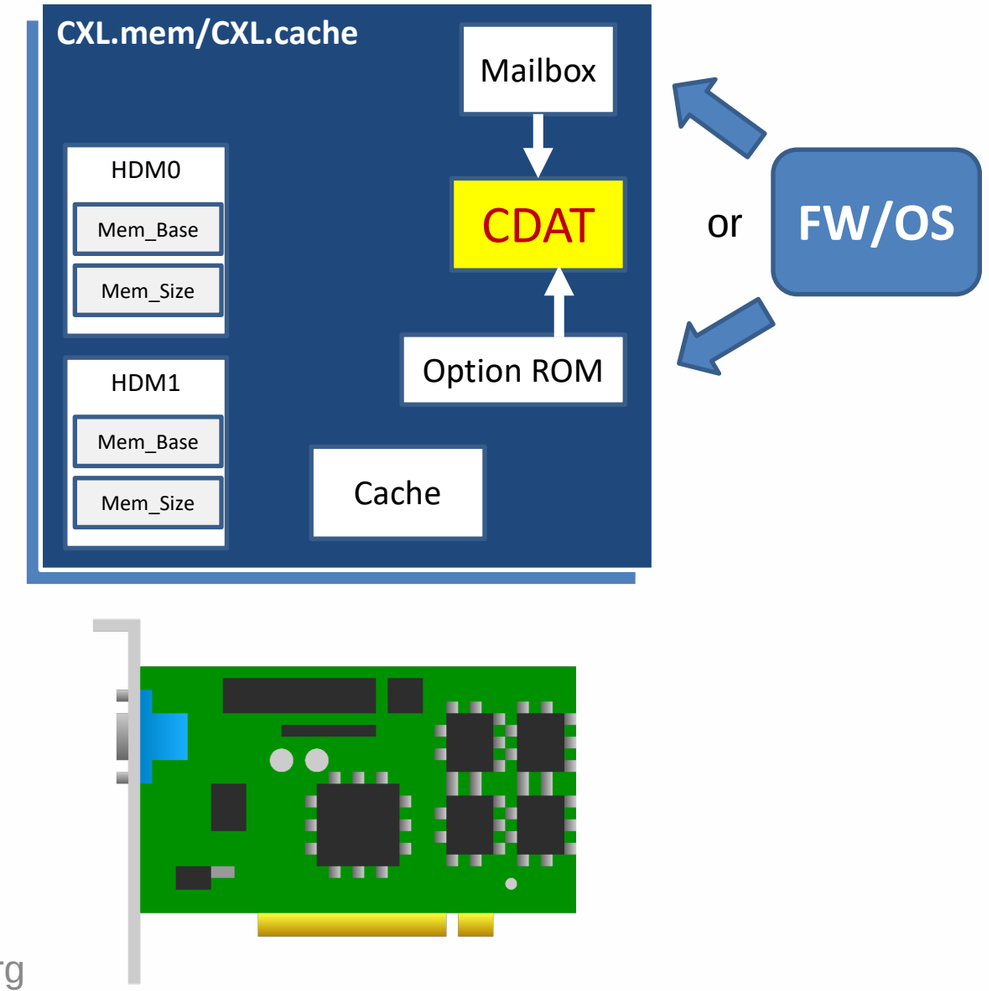
Accelerator or device-attached Memory

Coherent Device Memory and HMAT

- **CDAT**
  - Coherent Device Attributes Table
    - Coherent Accelerators
    - Accelerator-attached coherent memory
    - Coherent switch
  - Provides NUMA characteristics of the device:
    - Internal NUMA domains
    - Bandwidth
    - Latency
    - Presence of Generic Initiators
    - Memory Usage Recommendations (SPM)
    - Presence of Memory-side Caches on coherent devices
- CDAT provides NUMA data to assist creation of the HMAT table

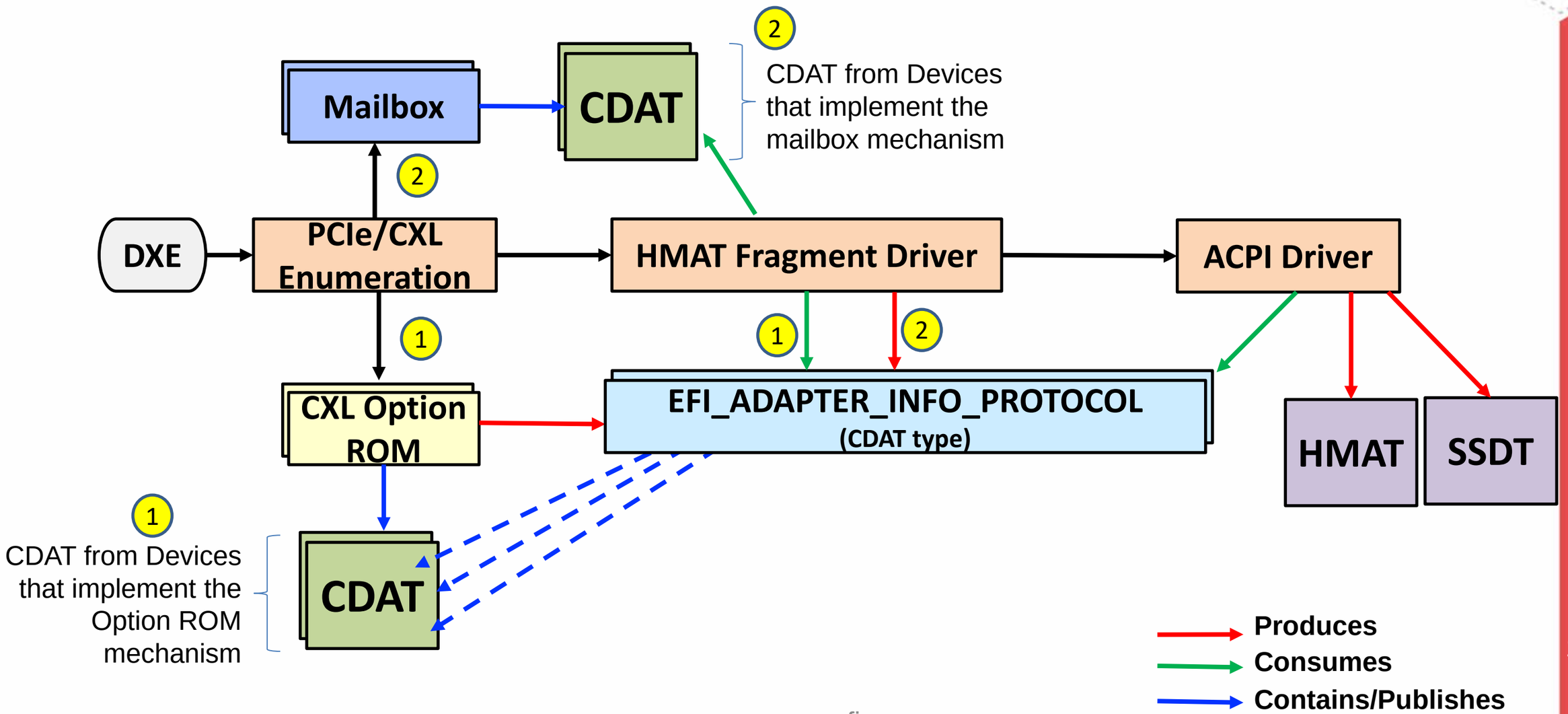
CDAT Discovery Mechanisms

- Option ROM (UEFI image)
  - Launched during CXL enumeration
- PCIe® Mailbox
  - Firmware/OS reads directly from device
- Device vendors can choose either option



CDAT Discovery and HMAT Creation

UEFI Flow (Proposed)



CXL Discovery

- CXL Host Bridges and associated registers can be discovered via CXL Early Discovery Table, prior to parsing of ACPI namespace.
  - CEDT has one entry for each CXL Host Bridge
  - CEDT format defined in CXL specification
- Next level of discovery is based on ACPI Namespace
  - CXL Host Bridge Hardware ID=“ACPI0016”
  - Compatibility ID of PCIe Host Bridge to enable enumeration by non-CXL enabled OSs
  - CXL _OSC method ensures OS and Firmware stay in sync, defined in CXL specification

Summary and Call to Action

- CXL may very well change how we compute
- UEFI and ACPI enablement for CXL is a work-in-progress
- Good progress has been registered in general on UEFI and ACPI enablement for heterogeneous computing systems
- We encourage everyone to participate in this industry effort
  - Please consider joining CXL
  - Please consider contributing to UEFI/ACPI definitions to support CXL and heterogeneous computing

Questions?

Thanks for attending the UEFI 2020 Virtual Plugfest

For more information on UEFI Forum and UEFI Specifications, visit http://www.uefi.org

presented by