BKK19 TR-07
Enabling Seamless Acceleration with CCIX Technology
A Software Perspective

Thanu Rangarajan, Arm Limited
Kangkang Shen, Huawei Technologies
Hardware view and Use-cases

- Processor
  - Cache
  - Memory
- Accelerator
  - Cache
  - Memory
- Processor
  - Cache
  - Expansion Memory
  - Memory
A Simple CCIX System

- A simple CCIX system and its key components
A Simple CCIX System

- How the CCIX components map to a regular CPU-based world
Example System Topologies
CCIX as a NUMA System

- Best viewed as a NUMA system with initiators (RAs and processors) and targets (host memory and expansion memory)
NUMA...with heterogeneous Memory

- Viewed as a NUMA system with initiators (RAs and processors) and targets (host memory and expansion memory)
  - Different latency and bandwidth characteristics for different pairings of initiators and targets
UEFI/ACPI Description

• Acceleration Functions appear to software as heterogeneous processors
  o Best defined as ACPI *Generic Initiators*
    ■ Generic Initiators are non-CPU initiators, introduced in ACPI 6.3
    ■ Described in ACPI DSDT

• Expansion memory included as part of regular system memory
  o However, *preferred* for accelerator use
  o Kernel allocations and general-purpose applications steer away from expansion memory
  o Also called *specific-purpose memory*
  o Described in ACPI NUMA tables (SRAT, SLIT, HMAT)
CCIX over PCIe Transport

Protocol Layer

Transport Layer

CCIX Protocol

PCIe

AFC

Mem

RA

HA

CCIX

PCIe

AFC

Mem

RA

HA

CCIX

PCIe

AFC

Mem

RA

HA

CCIX

PCIe

AFC

Mem

RA

HA

CCIX

PCIe

AFC

Mem

RA

HA

CCIX

PCIe

AFC

Mem

RA

HA

CCIX

PCIe

AFC

Mem

RA

HA

CCIX

PCIe

AFC

Mem

RA

HA

CCIX

PCIe
CCIX Configuration, Capability and Status Registers

- Component Configuration and Status Register structures

1. Port structure
2. RA structure
3. HA structure
4. SA structure
5. AF Property Table
6. Memory Pool Descriptors
7. Link structure
8. Device-wide (global) structures
CCIX Configuration Space

- CCIX registers are located in PCIe configuration space:
  - Mapped into PCIe extended capability space as a DVSEC structure:
    - CCIX Transport Layer DVSEC for Transport Layer specific configuration and status registers
    - CCIX Protocol Layer DVSEC for Protocol Layer specific configuration and status registers
  - Includes all component structures
The CCIX Protocol Layer DVSEC contains all the component structures required to program the CCIX device, and query for device status.

Within the DVSEC region, component structures are logically arranged as:

- Capabilities and Status structures
- Control structures
Component Structure Layout

• Function 0 on the primary port holds the common (device-wide) component structures
• Non-0 functions can hold other, agent-specific component structures
• This form of partitioning enables a consistent discovery and configuration model for all CCIX devices
Overview of Software Layers

○ CCIX initialization based on self-description (PCIe-based)
  ▪ Description allows discovery of capabilities and topology
  ▪ Used to control the hardware configuration and mapping to system

○ CCIX configuration is performed in firmware (UEFI)

○ OS manages
  ▪ Placement of processes and data (NUMA / topology awareness)
  ▪ Data flows (Accelerator chaining)
  ▪ Minimal involvement with applications (eSVM)
  ▪ Power Management and RAS
Generic OS Architecture

- User-space
  - App
    - AF
  - App
    - AF

- Kernel space
  - OS - kernel
    - CCIX subsystem
    - PCIe subsystem
    - Expansion Memory
Boot Firmware Overview

- **Discovery and Configuration of CCIX Subsystem**
  - Discovery of CCIX devices
  - Discovery of CCIX device topology
  - Discovery and configuration of CCIX resources
    - Acceleration Functions
    - Caching Agents, Homing Agents, Memory Controllers
    - Expansion Memory

- **Advertisement of CCIX Resources**
  - Expansion Memory incorporated in the System Memory Map
  - NUMA Properties described (e.g. ACPI SRAT/HMAT)
  - Acceleration Functions are self-describing because they are PCIe functions
CCIX Boot Firmware Overview

1. Host Discovery
2. PCIe Enumeration
   - CCIX Device Enumeration (Discovery)
3. CCIX Device Enumeration (Discovery)
4. Component Discovery
5. Memory Discovery
6. Topology Discovery
7. Topology Configuration
8. Host-specific CCIX Configuration
9. Description of G-SAM to OS

Warm reset
Progress Thus Far...

• **UEFI and ACPI Enablement**
  ○ UEFI 2.8 and ACPI 6.3 specifications released with CCIX-specific updates

• **CCIX Ecosystem Software Project**
  ○ Ongoing joint-development of CCIX reference software stack – Huawei, Linaro and Arm
  ○ Open-source Software Development Model formalized, similar to PCIe

• **CCIX Specifications**
  ○ CCIX Software Developer’s Guide
    ■ Planned release Q2, ’19
  ○ CCIX Base Specification
    ■ Available to CCIX Consortium Members since Jan 2019
Thank You!