SAMSUNG

Future of Data center and containerization with CXL

10. 29. 2020 | Memory Division / Product Planning Team / DC Platform Group

About us

Chon Yong Lee

- Product Planning Engineer at Samsung Electronics Memory Business
- Building memory solution software planning task
- Previously worked as a product manager at Cloud Labs, SK Telecom
- Architect of 5G Digital Twin DataCenter/Network Solution
- Session/Booth at Strata NY 2019, MWC 2016, 2018

- #Memory #CXL #SoftwarePlanning #DataCenter #5G #Network #SSD #Architecture #Cloud #DataVisualization #SDN #AR

Seok Jae Han

- Product Manager in the field from memory Solution development to strategy establishment
- Leading development of UFS Card, Samsung's 5G advanced removable storage
- Demonstrated 4K VR contents in HMD with UFS Card at CES 2017 & MWC 2018
- 5G Industrial specialist TF in SAMSUNG 2019

- #memorydevice #SolutionEngineer #VR #AR #5G #UFS #Datacenter

Agenda

1. Data Center/Cloud Trend

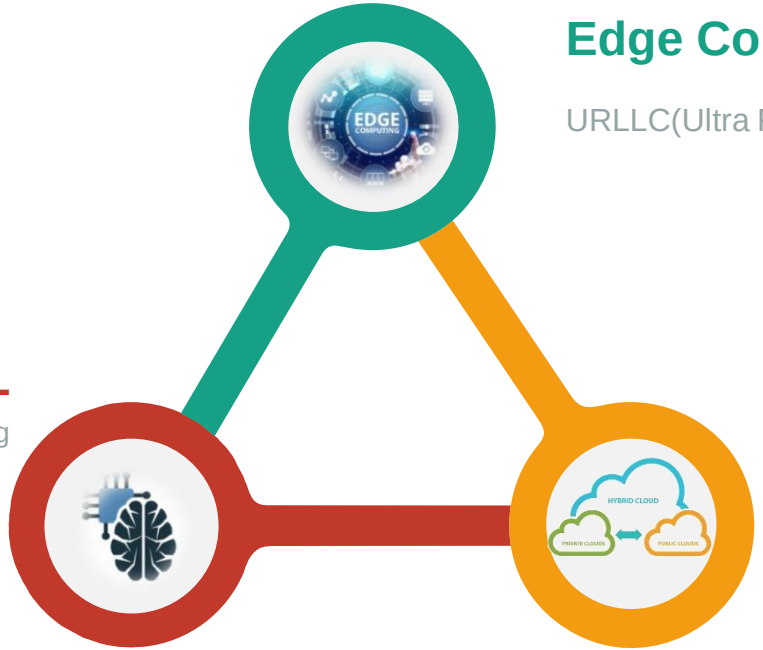
2. CXL (Compute eXpress Link) Overview

3. CXL Type3 Cloud Use Case

Data Center/Cloud Trend

Data Center Service Trend : Cloud & AI/ML & Edge

AI/ML, Public cloud and Edge computing



Edge Computing (5G)

URLLC(Ultra Reliable, Low Latency)

AI / ML

Heterogeneous Computing

Hybrid/Multi Cloud

Containerization, expand service area

Data Center Architecture Trend

1. Heterogeneous Computing

1) Special purpose Accelerator works with general purpose CPU in DC, HPC, AI, Image etc

2) CXL (Compute eXpress Link) : Intel lead open source base CPU ↔ Device I/F

- □ Link / Transaction Layer Modification in PCIe → supporting “cache coherency” & “low latency”
- □ Currently there are no Interconnection for Coherence between CPU and Accelerator

Data traffic explosion
Data Processing
Application enlarge
(Data : Value creation)

→

Various Accelerator
(Media, Image, language processing, AI...)
Heterogeneous Computing

→

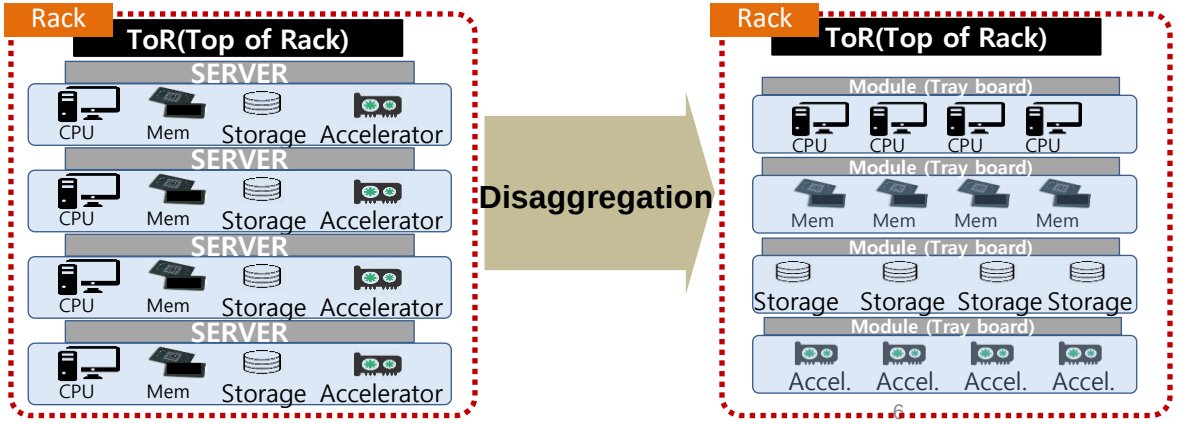
I/O Traffic
(CPU↔Accelerator)
increase on PCIe protocol

→

New Protocol requirement
CPU ↔ Device

(Gen-Z, OpenCAPI, CCIX, CXL)

2. The appearance of “Resource-centric” DC : Disaggregated DC with Memory centric computing



[3 Technology for Disaggregation]

- Device-to-device Interconnect Tech. in H/W
- Operation System for Shared memory pool
- New Application technology utilizing memory-centric computing

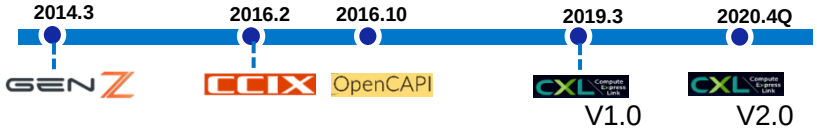
SAMSUNG

CXL Overview

Intel CXL(Compute eXpress Link)

1. **Intel formed the CXL Consortium in March 2019 and released CXL specification 1.0**
   1) Optimizes cache coherency and low latency through modification of PCIe link and transaction layers
   2) Supports cache coherent schemes for heterogeneous computing and resource disaggregation
2. **High prospects for standardization due to broad industry support**
   1) More than 130 participating companies, collaborating across 6 consortium work groups

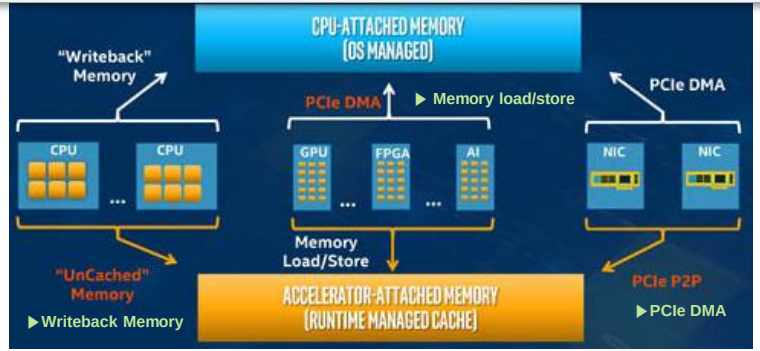
Interconnect Protocol History



구분	PHY	Topology	Coherency
Gen-Z	Ethernet	P2P and Switched	P22 Coherency
CCIX	PCIe Gen4	P2P and Switched	Full cache coherency (Processors↔accelerator)
OpenCAPI	Bluelink	P2P	Coherence access to memory
CXL	PCIe Gen5	P2P (**)	Full cache coherency (Processors↔accelerator)

CXL design goals

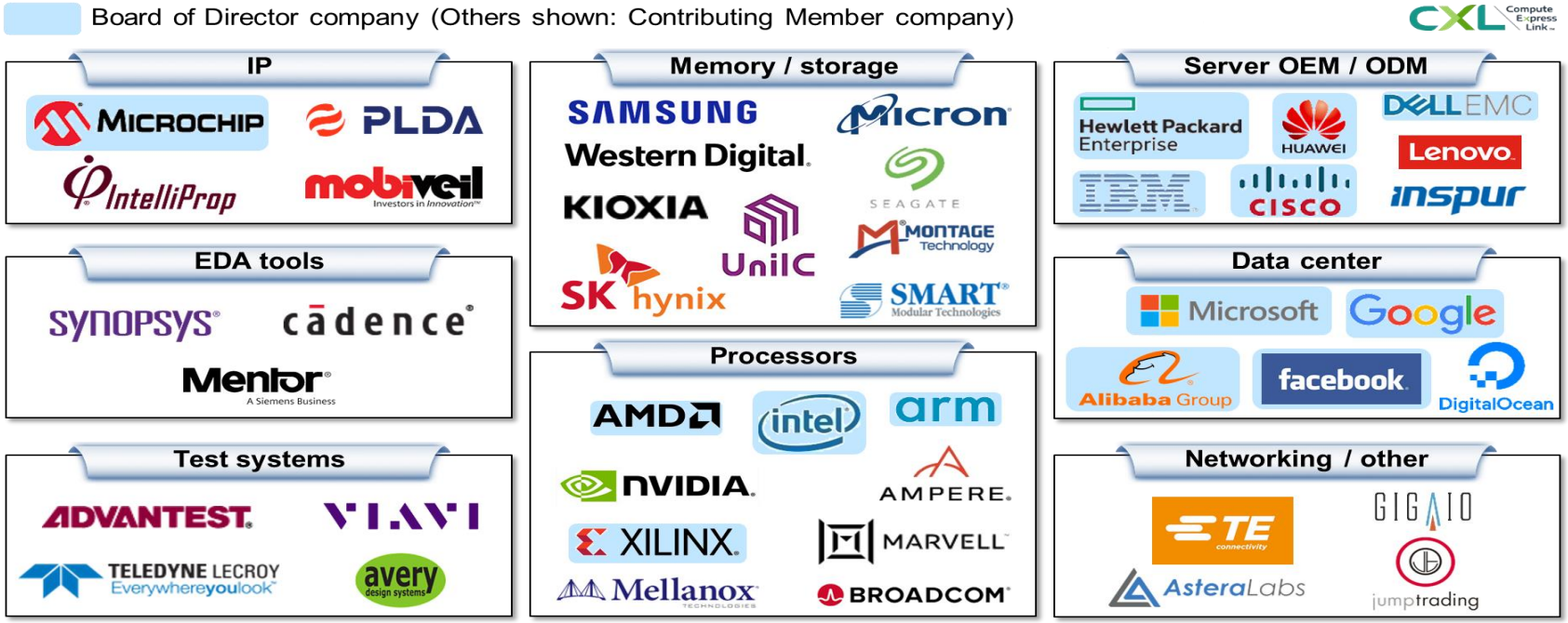
- Provide efficient access mechanism for heterogeneous computing
- Enable effective resource sharing as data movement expands
- Reduce latency for disaggregated memory/resources



SAMSUNG

CXL member

1. **Over 140 companies representing expertise in the semiconductor industry are participating in CXL**
   1) 14 Board of Director companies, over 130 Contributing Member companies

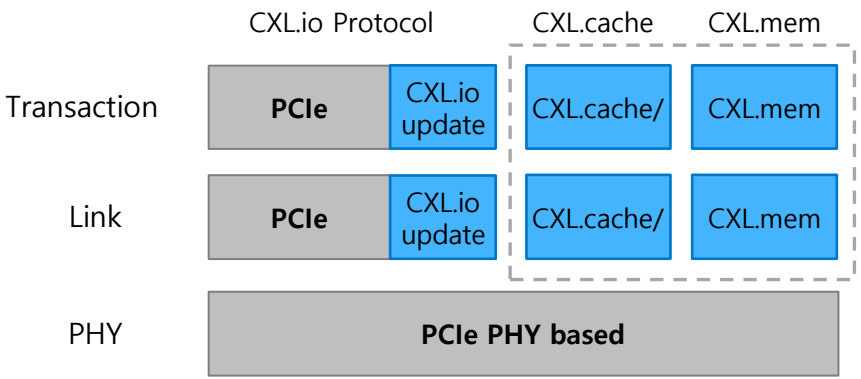


CXL Overview 1) Protocol

1. **CXL : Based on PCIe Physical layer, Add CXL "Memory Access" and "Cache Coherence"**
   1) CXL.io : discovery, configuration, register access, interrupts, and other PCIe functionalities
   2) CXL.cache: device access to processor memory
   3) CXL.memory: processor access to device attached memory

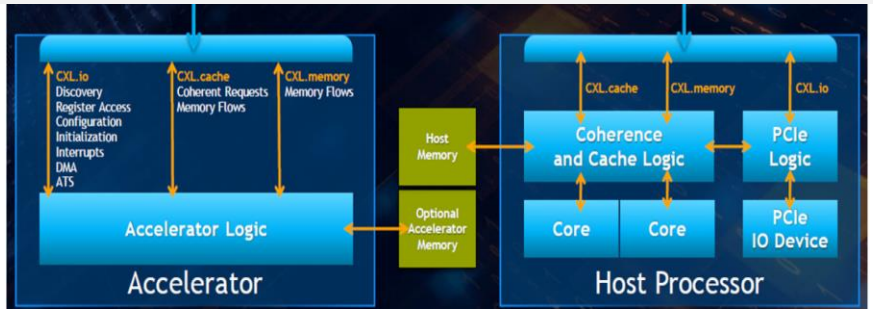
CXL Protocol Frame

- **CXL is based on PCIe Physical layer**
- **CXL.cache/CXL.mem → latency optimization**



CXL Protocol

1) Effective access mechanism of "shared memory pool" of Heterogeneous computing
2) Extended data flow and effective resource sharing between Accelerator and CXL devices
3) Reduced access latency of distributed memory



SAMSUNG

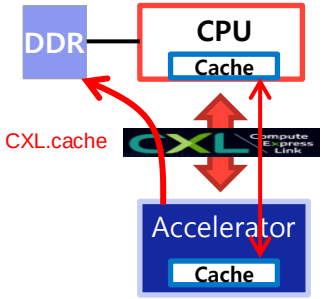
CXL Overview 2) Device type

1. **Classified into 3 types of CXL depending on whether Accelerator(CXL.cache) and Memory controller(CXL.mem) possessed**

Type 1

Caching / Accelerater

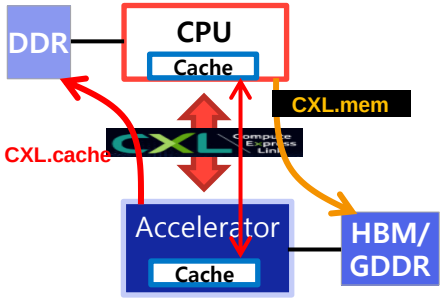
- Usage : NIC atomics
- Protocols : **CXL.io**, **CXL.cache**



Type 2

Accelerators with Memory

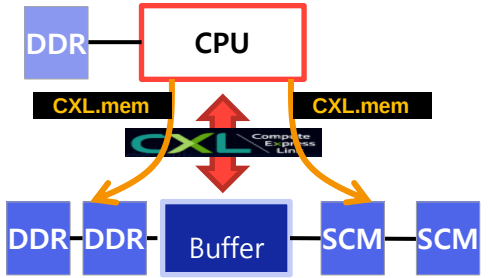
- Usage : GPU, Dense Computation
- Protocols : **CXL.io**, **CXL.cache**, **CXL.mem**



Type 3

Memory Expander

- Usage : Memory BW / Capacity Expansion, 2LM
- Protocols : **CXL.io**, **CXL.mem**



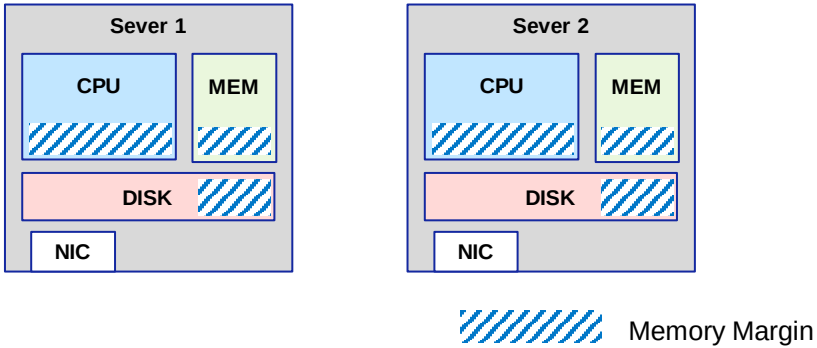
SAMSUNG

CXL Type3 Cloud Use Case

CXL use case 1-1) Memory margin sharing

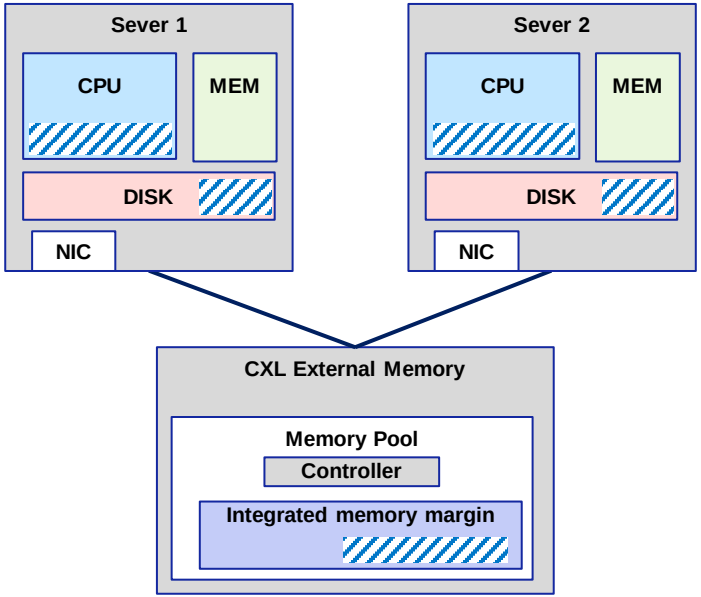
Sharing memory margin using CXL memory pool

As-is



- Integrate memory margin in CXL memory pool
- Use internal memory 100%
- CXL overhead is similar to UPI
- Physical range limits

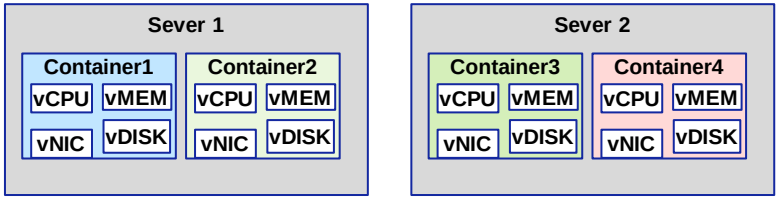
To-be



CXL use case 1-2) Increase memory utilization

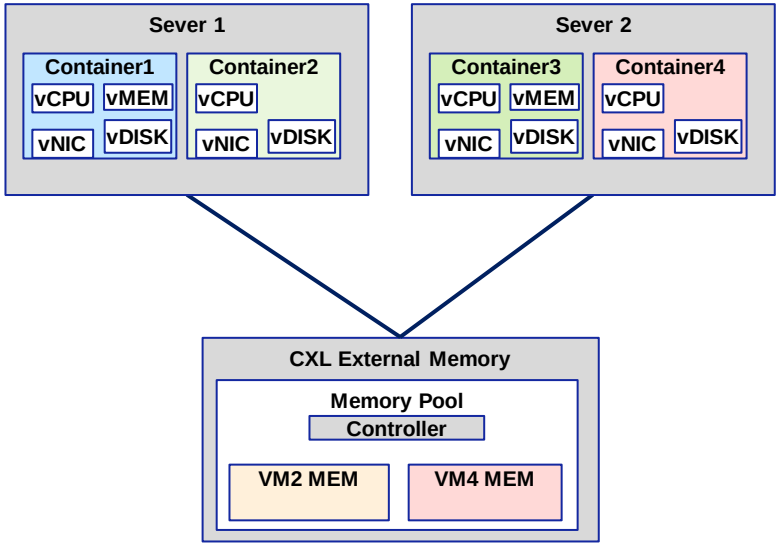
Container/VM uses external memory pool as its main memory

As-is



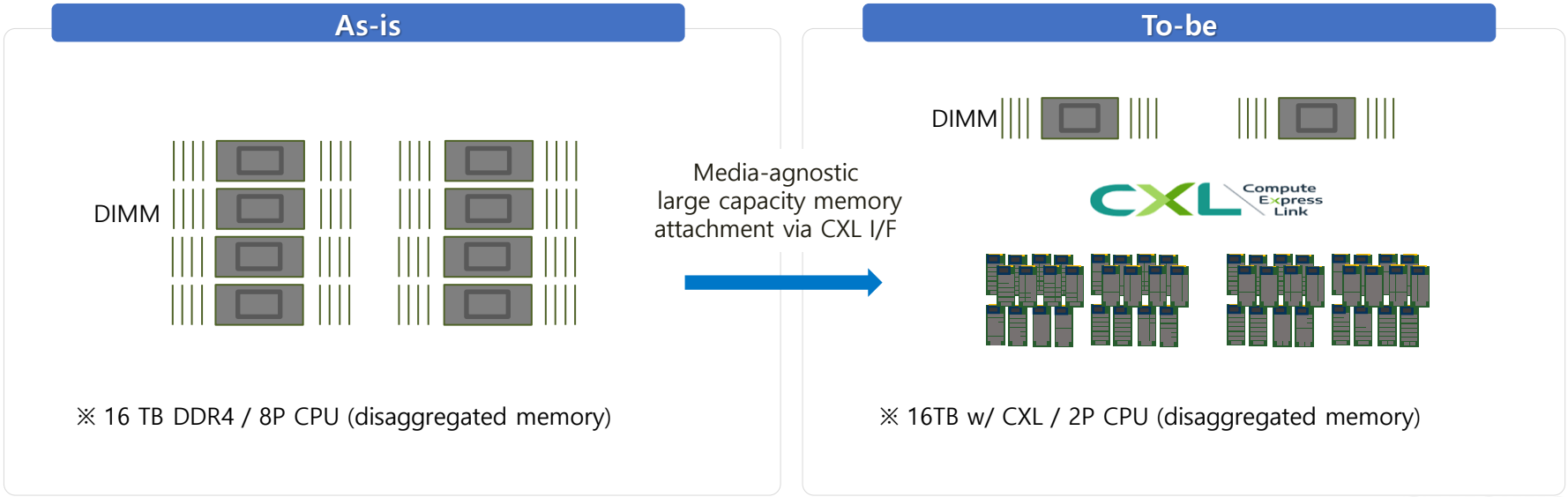
- Increase memory usage flexibility
- Memory centric containerization
- HA, Clustering, distributed processing

To-be



CXL use case 2) internal DRAM Memory pool for IMDB

1. **Memory expand using CXL, Server memory scale-up**
   1) Memory Pool / Disaggregation
   2) Add memory without CPU limitation
2. **Large/shared memory pool cloud change DB, Container and K8S services**

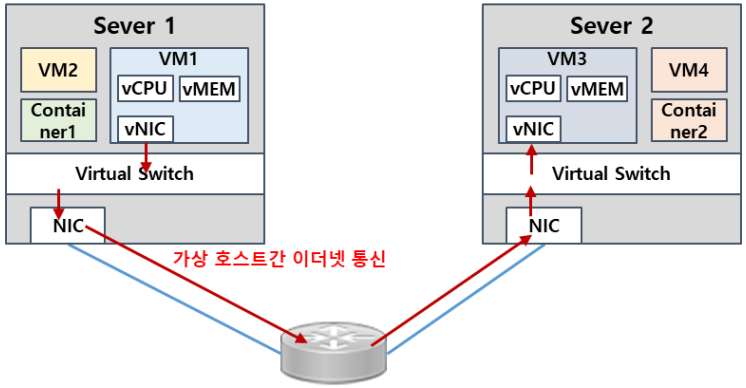


CXL use case 3) Network acceleration

CXL Memory pool 공유하는 Host간 데이터 교환하여 가상 호스트간 통신 가속

As-is

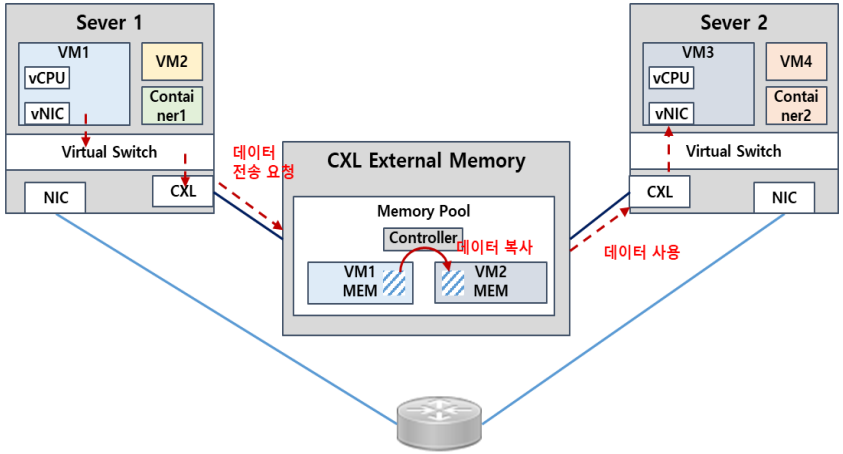
가상 호스트간 데이터 통신 방법



- MSA (Micro Serviced Architecture) 등으로 데이터 센터 내 East-West Traffic 증가
- 가상 스위치, Host NW 가속을 위해 Smart NIC, OVS-DPDK 등 다양한 기술 사용 중

To-be

CXL External Memory 활용 가상 호스트간 데이터 전달 – 데이터 복사



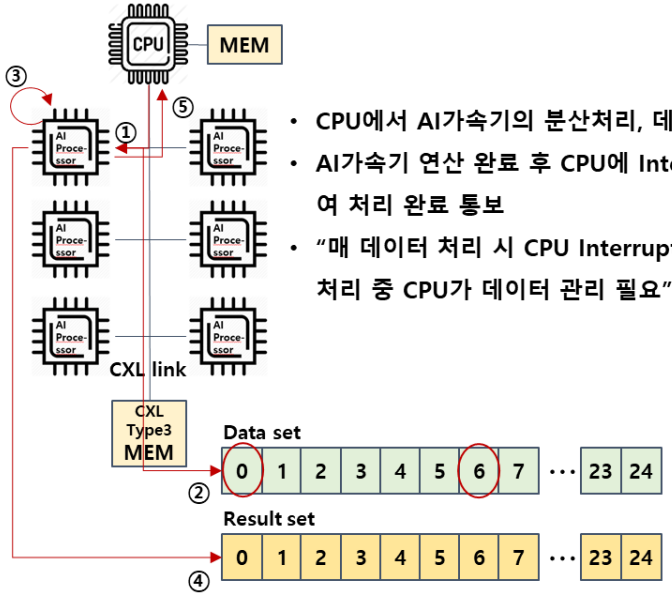
- 가상 스위치 S/W를 변경하여 CXL Memory Pool 內 데이터 복사로 데이터 통신 가능
- NW Stack, Security 등 이슈 있으나 구현 시 CXL 메모리 공유 그룹 내 데이터 전달 성능 대폭 향상

SAMSUNG

CXL use case 4) Heterogeneous Pub-sub Computing

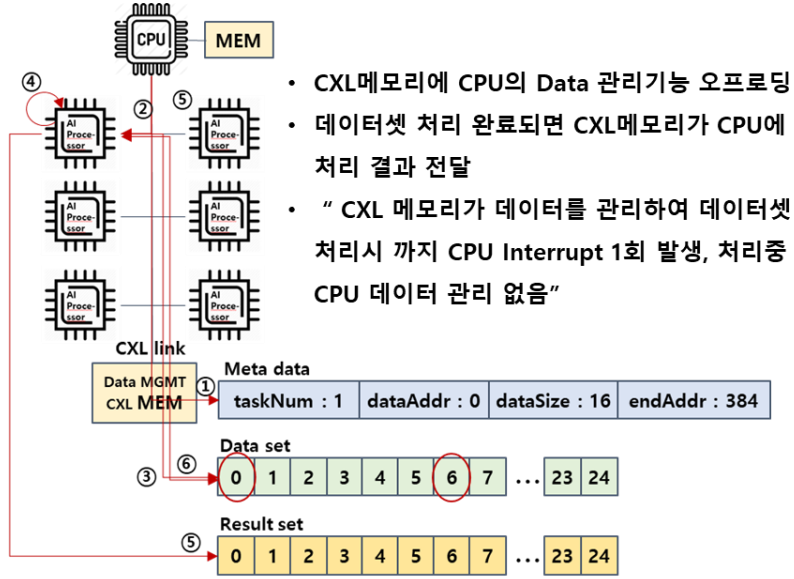
CPU와 AI 가속기간 CXL 메모리를 Message Queue 처럼 활용, Pub-sub 구조로 데이터 교환

As-is



- CPU에서 AI가속기의 분산처리, 데이터 관리
- AI가속기 연산 완료 후 CPU에 Interrupt 하여 처리 완료 통보
- "매 데이터 처리 시 CPU Interrupt 발생, 처리 중 CPU가 데이터 관리 필요"

To-be



- CXL메모리에 CPU의 Data 관리기능 오프로딩
- 데이터셋 처리 완료되면 CXL메모리가 CPU에 처리 결과 전달
- " CXL 메모리가 데이터를 관리하여 데이터셋 처리시 까지 CPU Interrupt 1회 발생, 처리중 CPU 데이터 관리 없음"

- 기존 방식은 매 데이터 프레임마다 CPU Interrupt, 데이터 관리 필요하나, 본 발명은 데이터 셋 처리 완료시 1회만 발생
- 다음 Stage 연산 및 다음 데이터 셋은 여전히 CPU가 관리

Next Step

- Heterogeneous Computing과 CXL Interface 환경에서 Container, Kubernetes의 변화?
- Container와 K8S를 돕기 위해서 메모리(DRAM/Storage) 솔루션은 어떻게 발전해야 할까요?
- S/W와 H/W의 co-dev, co-ops을 통한 차세대 데이터센터 제품 기획
- 오픈소스 커뮤니티와 같이 이야기 하고 싶습니다.
- 함께할 동료를 찾고 있습니다.

Q&A

chonyong.lee@samsung.com
sj0704.han@samsung.com

END