LMB: Augmenting PCIe Devices with CXL-Linked Memory Buffer

Presenter: Tao Lu (DapuStor)

Agenda

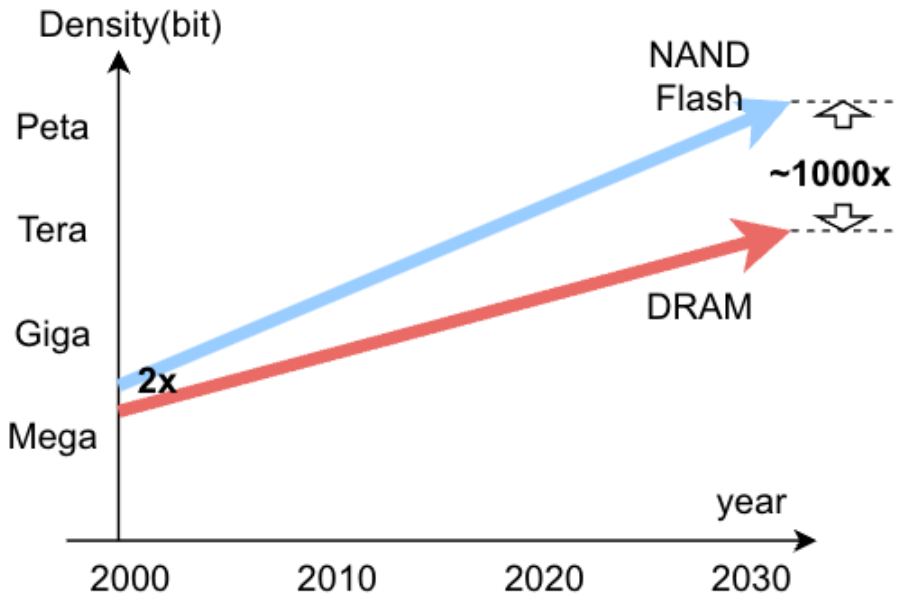
- Memory Shortage of PCIe Devices is the Question
- CXL Memory Expander is the Opportunity (if not the Answer)
- LMB Architecture
- Case Study
  - LMB for Ultra-capacity SSD 4KB Page Indexing
  - Discussion: LMB for KV-SSD, MS-SSD, and GPU
- Summary

Memory Shortage of SSD

SSD Constraint

On-board DRAM Shortage

- **Large Page (e.g. 16KB)**
- **LeaFTL (ASPLOS'23)**
- **LearnedFTL (HPCA'24)**



The trend of flash memory capacity (per SSD) and DRAM capacity (per DIMM)

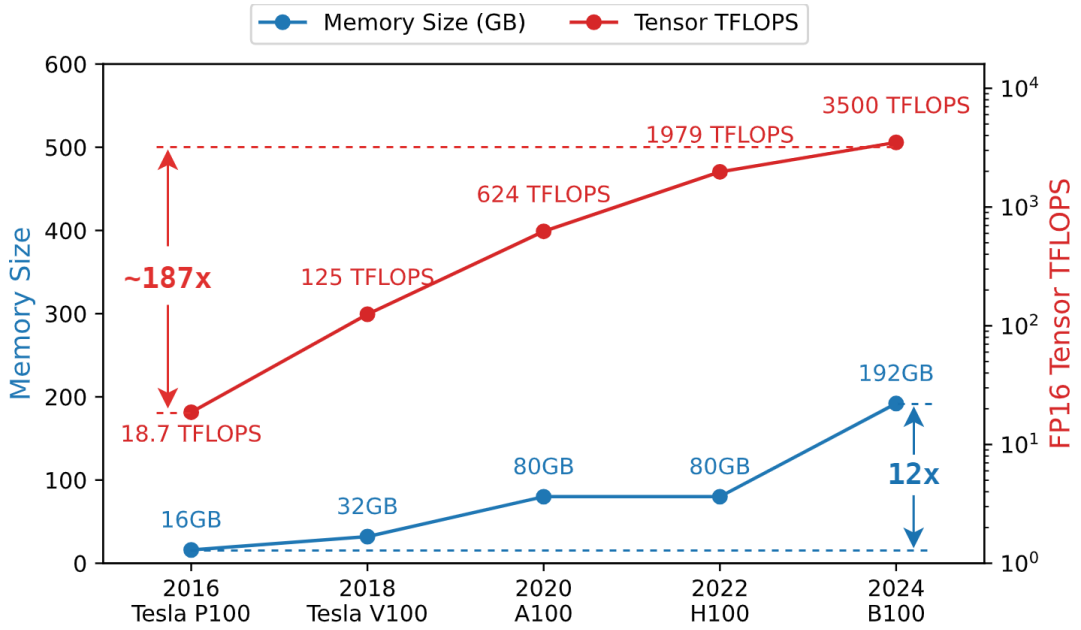
DapuStor

Memory Shortage of GPU

GPU Constraint

Limited HBM for AI models

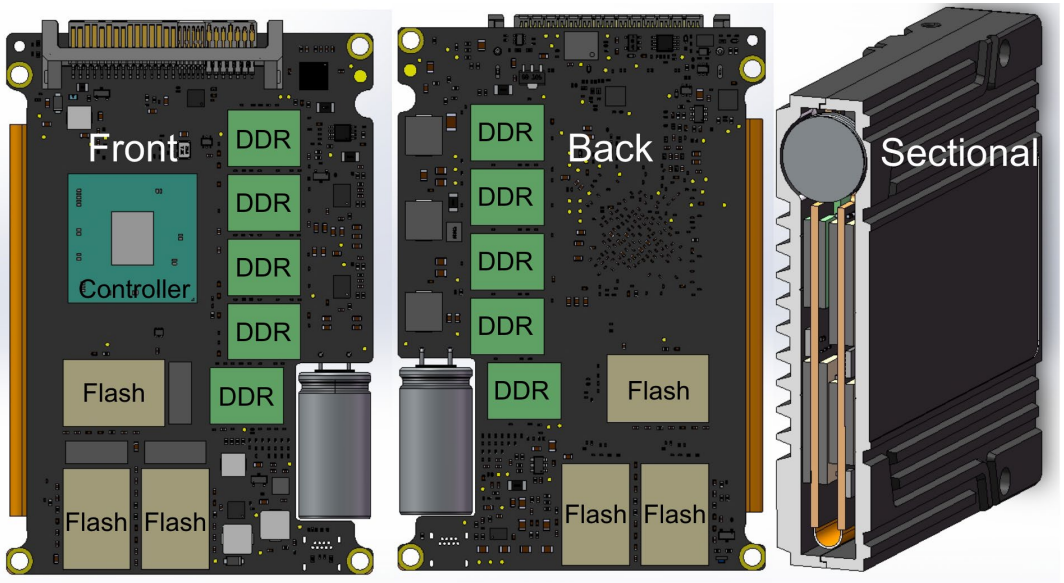
- **BaM (ASPLOS'23)**
- **G10 (MICRO'23)**
- **GMT (ASPLOS'24)**



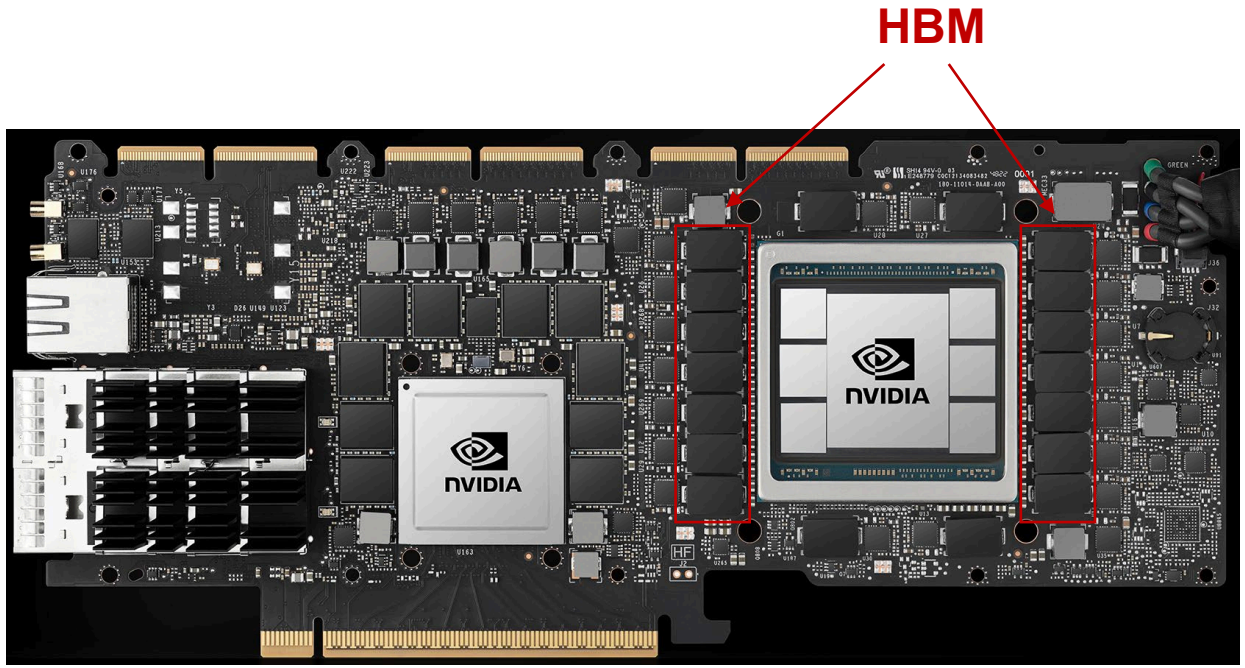
The growth trend of GPU computing power and HBM memory capacity

DapuStor

Root Cause of Memory Shortage Issue



SSD



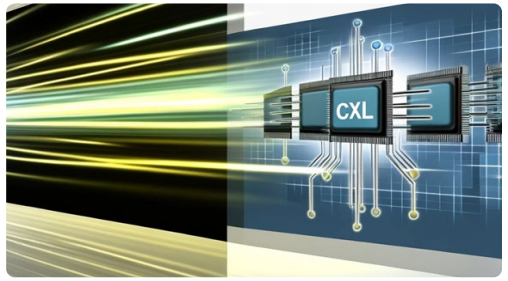
GPU

Root Cause: There is no room for more memory modules inside PCIe Devices

Source: https://www.nvidia.com/en-in/data-center/products/converged-accelerator/?ref=blog.paperspace.com

DapuStor

CXL Memory Expander is the Opportunity



Low Latency

CXL technology enables fast data transfer with **minimal latency**

Scalability

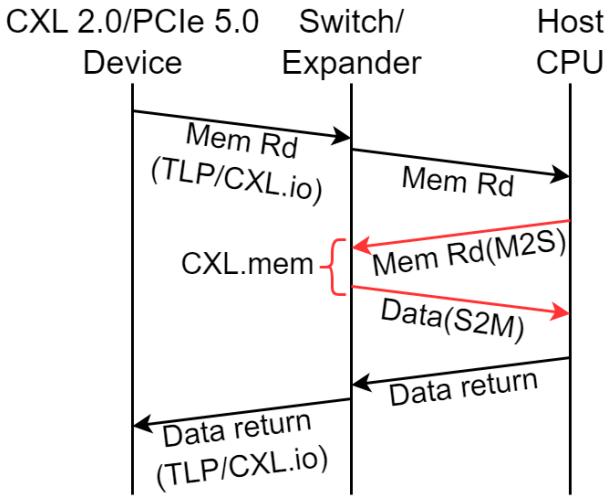
The flexible, **expandable memory architecture** supports growing demands for large systems.

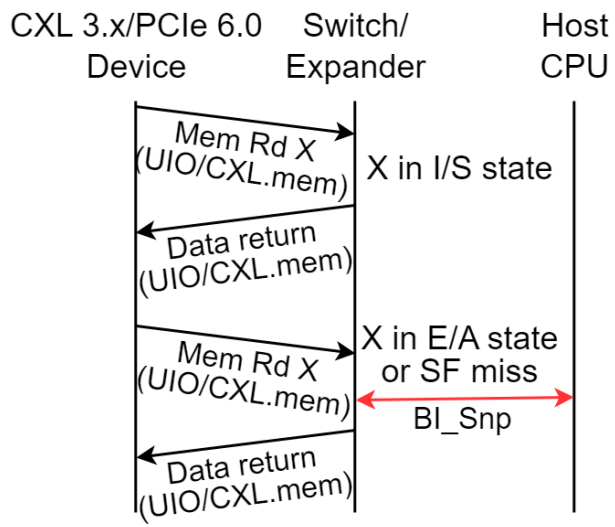
Memory Pooling

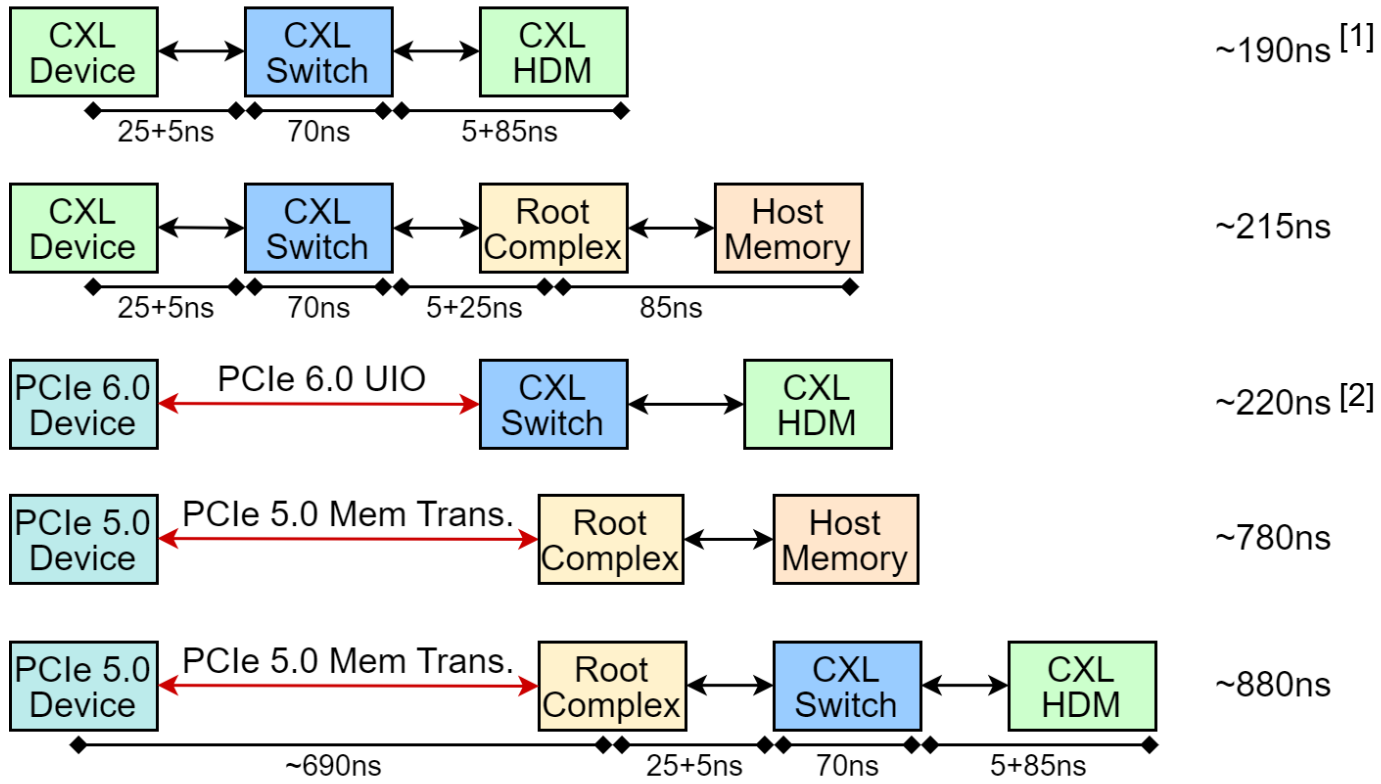
CXL allows for efficient **sharing of memory resources** between various devices

DapuStor

CXL Memory Expander is the Opportunity







[1]Huaicheng Li, Daniel S Berger, et al. 2023.Pond: CXL-based memory pooling systems for cloud platforms. ASPLOS, Volume 2. 574–587.
[2]Debendra Das Sharma. 2022. Compute Express Link®: An open industrystandard interconnect enabling heterogeneous data-centric computing. HOTI. IEEE, 5–12.

DapuStor

Linked Memory Buffers (LMB) Solution

CXL Memory Expander

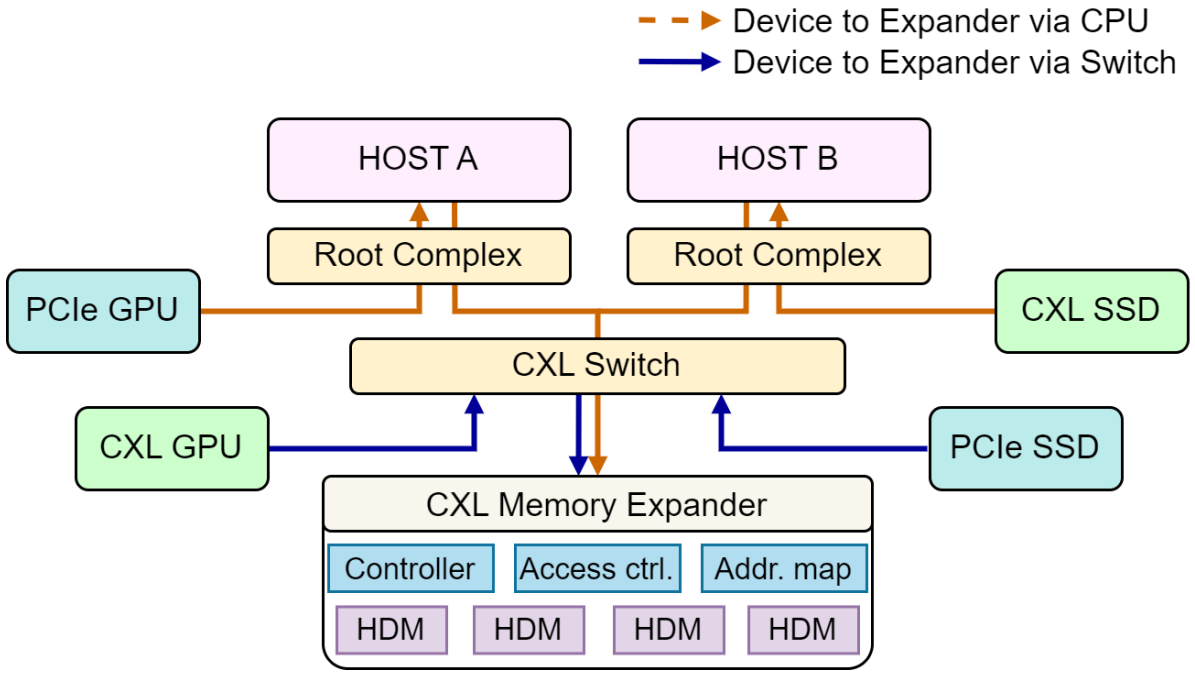
Physical source of additional DRAM

Unified Memory Interface

Seamless access to expanded memory

Dynamic Allocation

Efficient management of memory resources

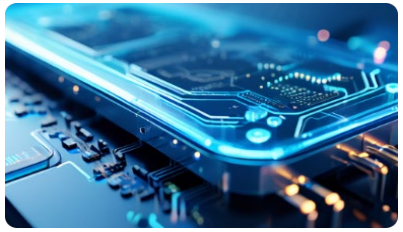


DapuStor

LMB Application Scenarios

Large Indexes

LMB improves indexing capabilities with low latency



Near-Data Processing

LMB enhances the efficiency of DPUs

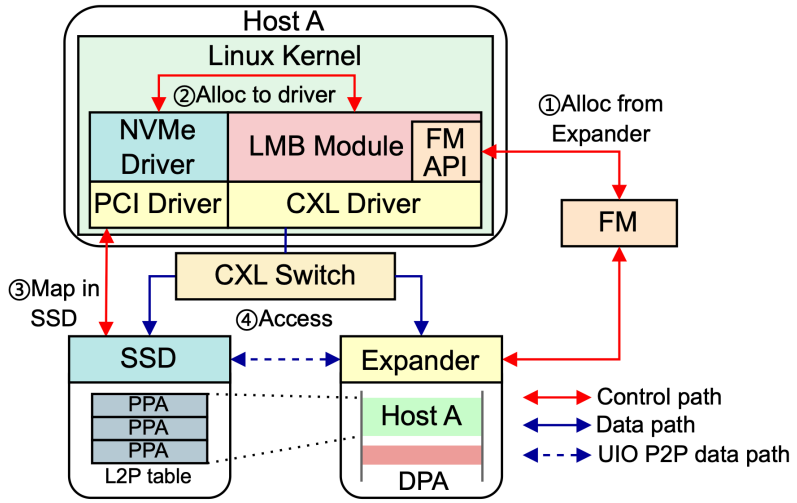
AI/ML

LMB enables large AI models to be deployed efficiently

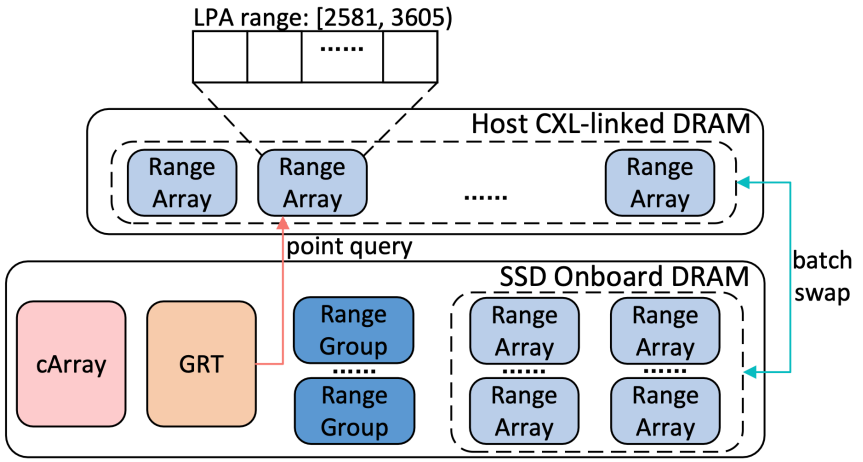
DapuStor

Case Study: 4KB L2P Page Mapping for SSDs

LMB for SSD Page Index



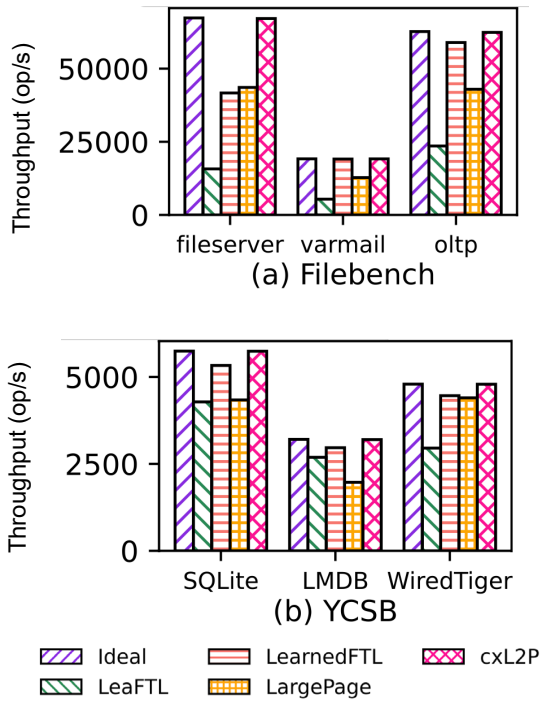
L2P Index data structures



DapuStor

Case Study: L2P (Evaluation)

Evaluation with FEMU emulation

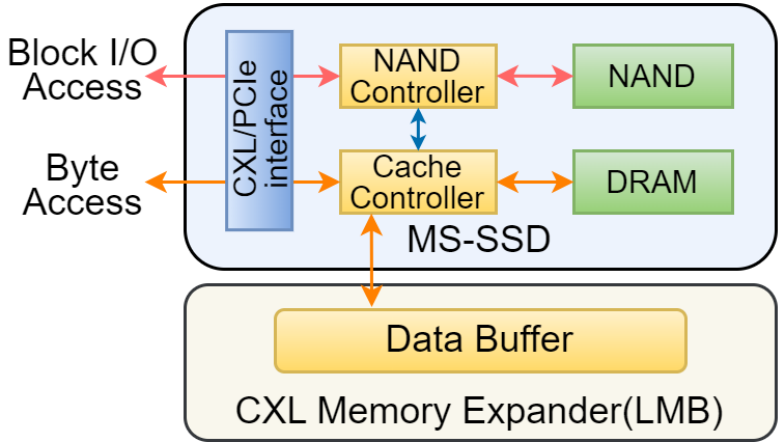


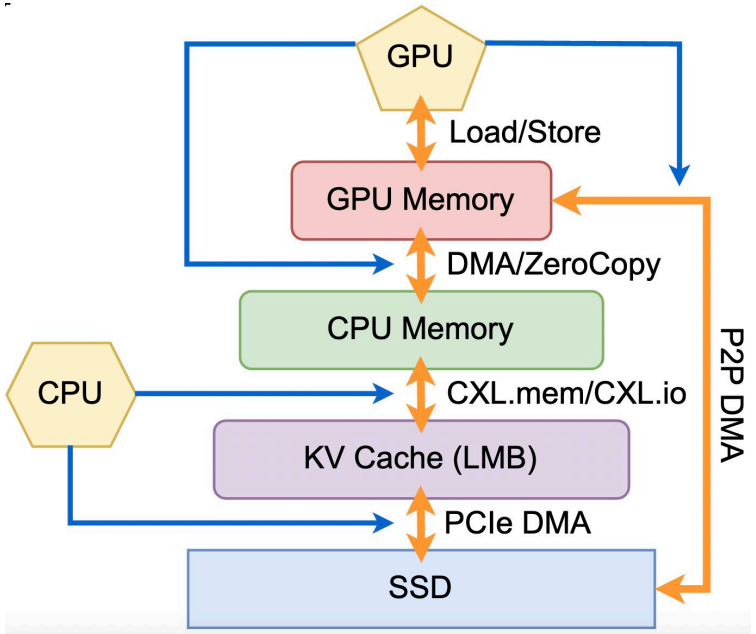
- Ultra-capacity SSD L2P solutions compromise performance: Large page, learned index
- LMB can provide enough DRAM for ultra-capacity SSD to implement 4KB fine-grained page mapping
- LMB-backed cxL2P enables SSDs to achieve the **ideal** performance

DapuStor

Discussion

- LMB for KV-SSD Index
- LMB for MS-SSD Data Buffer
- LMB for GPU Memory Tiering





DapuStor

Summary

01

Scalable Solution

LMB addresses onboard memory shortage issues of PCIe devices

02

Performance Boost

LMB Enhances SSD and GPU capabilities

03

Future Impact

LMB can widely enhance data centers and AI systems

DapuStor

Visit DapuStor at Booth#911

DapuStor