DIMM-Link: Enabling Efficient Inter-DIMM Communication for Near-Memory Processing

Zhe Zhou[*1,2,3], Cong Li[*1,3], Fan Yang[4], Guangyu Sun[†1,3]

[1]School of Integrated Circuits, [2]School of Computer Science, Peking University

[3] Beijing Advanced Innovation Center for Integrated Circuits

[4] School of Computer Science, Nankai University

{zhou.zhe, leesou, gsun}@pku.edu.cn, yangf@nbjl.nankai.edu.cn

HPCA23 Best paper

What is DIMM?
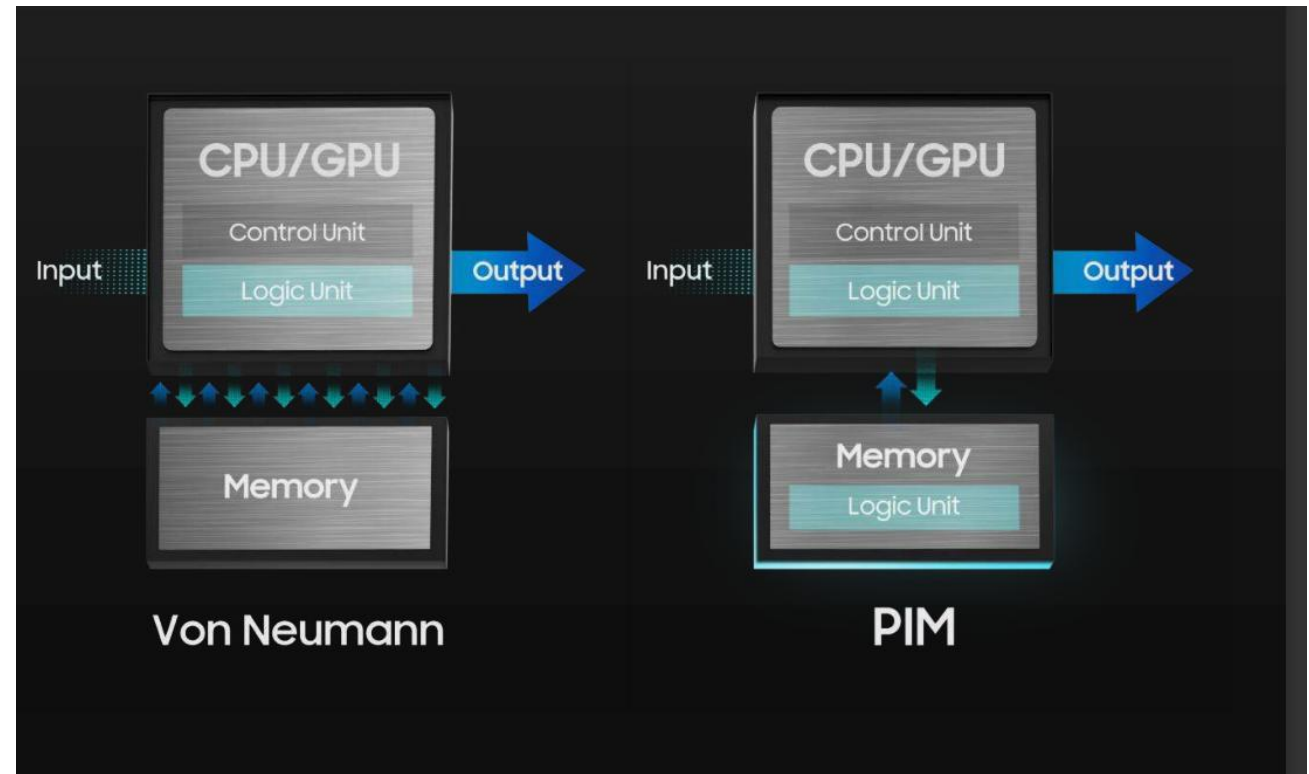- Dual-Inline Memory Module

What is DIMM-NMP?
- DIMM-based near-memory processing architectures
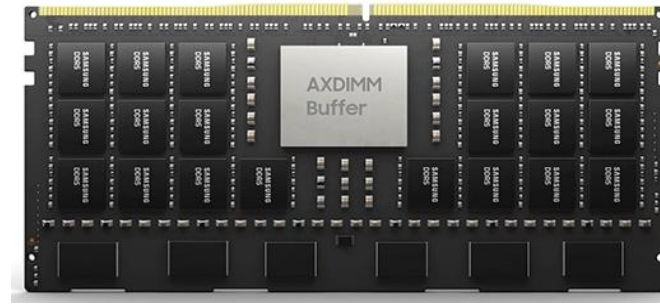
What is PIM?
- Processing in memory

Why DIMM-NMP(PIM)?
- Address the huge performance gap between CPU and main memory
- Offload memory-intensive operations
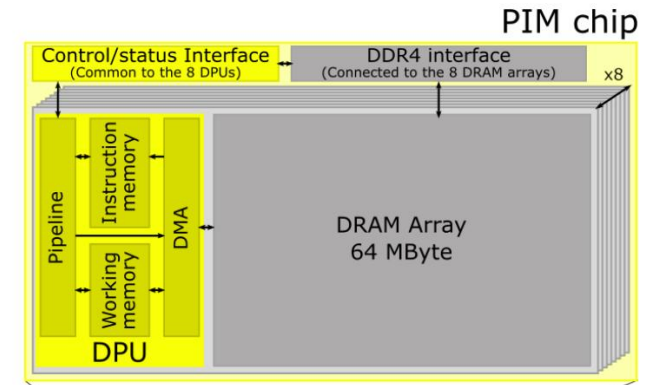- Mitigate energy-consuming off-chip/package data movement
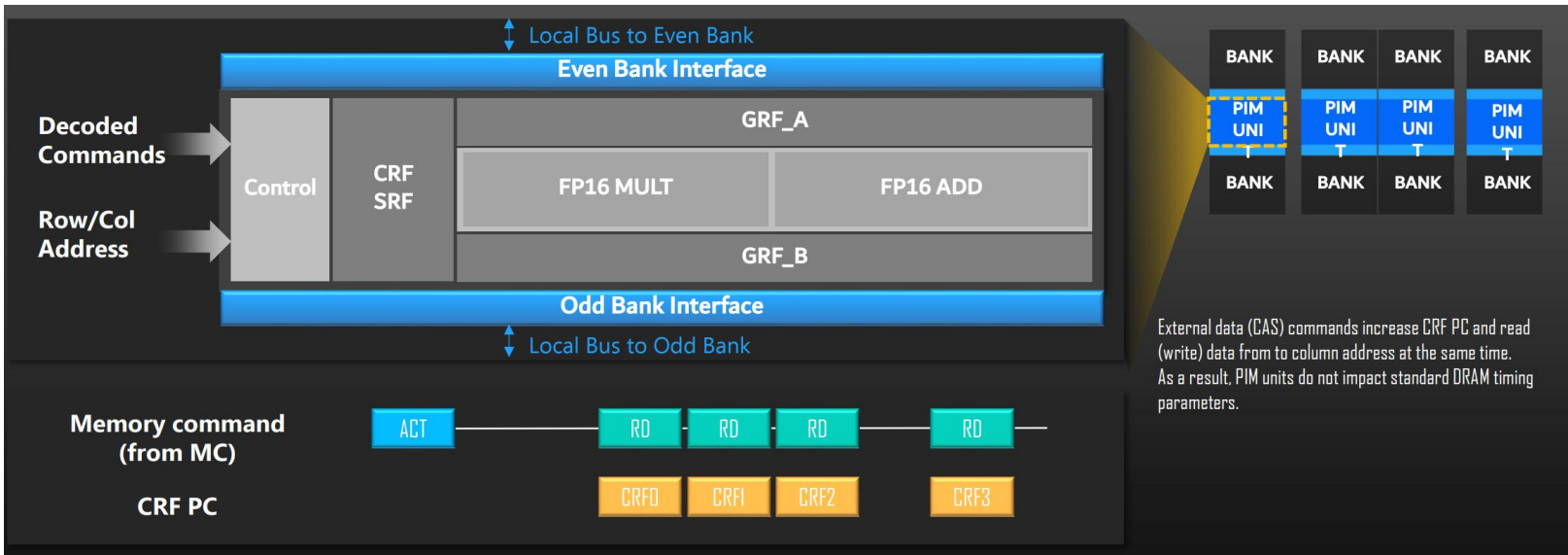
NMP(PIM) in industry

- Alibaba 3D HB-PNM[1]
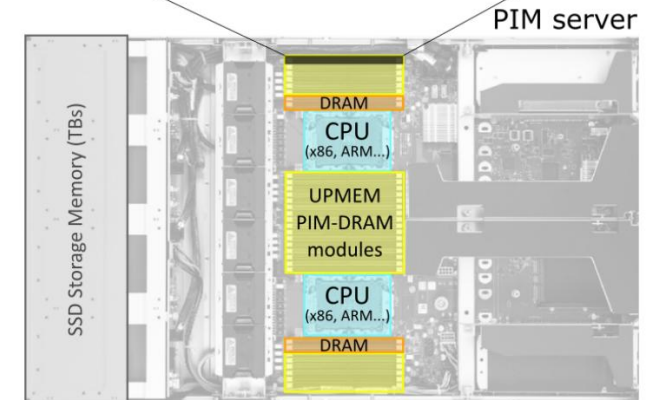- UPMEM PIM-DRAM[2]
- Samsung HBM2-PIM[3] and AxDIMM[4]

Samsung AxDIMM

PIM chip

Samsung HBM2-PIM

PIM module

PIM server

UPMEM PIM-DRAM

1. 3D Logic-to-DRAM Hybrid Bonding with Process-Near-Memory Engine for Recommendation System (ISSCC)
2. The true Processing In Memory accelerator (HOT CHIPS 31)
3. Aquabolt-XL: Samsung HBM2-PIM with in-memory processing for ML accelerators and beyond(HOT CHIPS 33)
4. Near-Memory Processing in Action: Accelerating Personalized Recommendation With AxDIMM (MICRO,Facebook)

Motivation

DIMM NMP
- cannot directly support inter-DIMM communication (IDC)
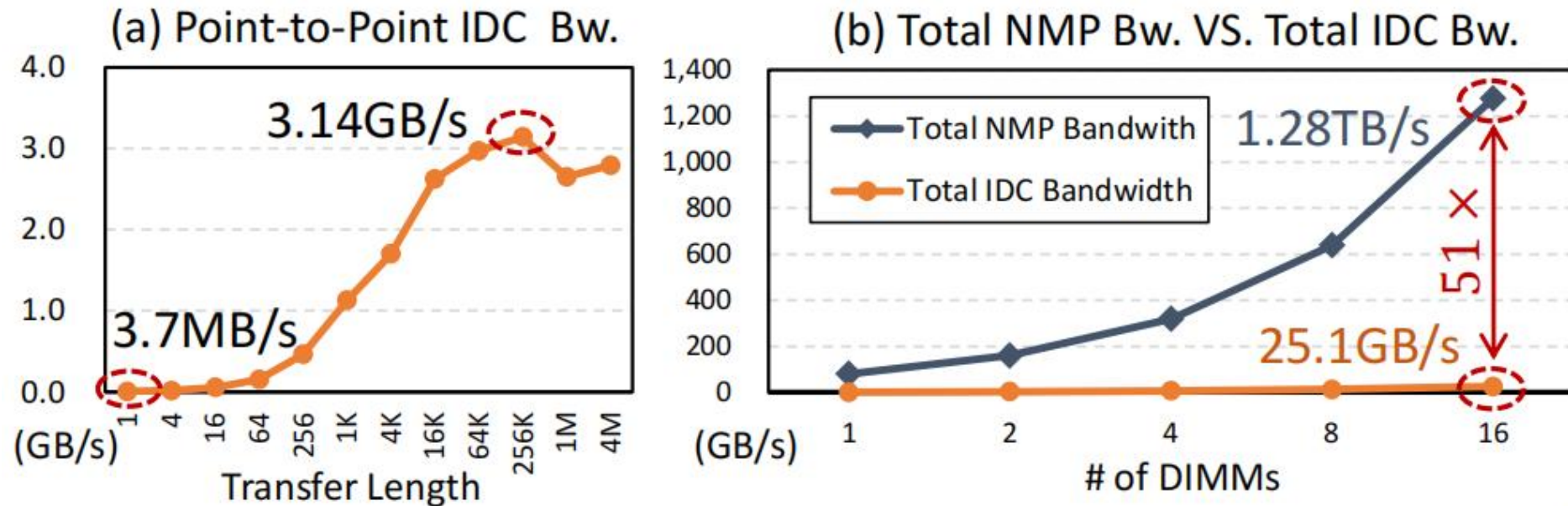- counts on the host CPU to forward inter-DIMM transactions

Fig. 1. IDC Performance Exploration on the Real DIMM-NMP Platform[1].

Motivation

TABLE I
COMPARISONS OF INTER-DIMM COMMUNICATION METHODS.

IDC Methods	CPU-Forwarding [3], [32]	Intra-Channel Broadcast [76]	Dedicated Bus [11]	DIMM-Link
Illustration				
Hardware Modification	DIMM Modules	Host CPU, DIMM Modules	DIMM Modules	DIMM Modules
Supported IDC Modes	Point-to-Point	Broadcast	Point-to-Point	Point-to-Point & Broadcast
Maximum Bandwidth	$\#Channel \times \beta/2$	$\#DIMM \times \beta$	β	$\#Link \times \beta$
Target NMP Apps	IDC-infrequent Applications	Sparse Tensor Algebra	Computational Genomics	Generic Applications

- CPU-Forwarding:

Expensive; DIMMs compete for the bandwidth; The periodical CPU polling occupies resources

- Intra-Channel Broadcast:

Limited DIMMs per channel; Many broadcast-unfriendly applications; Customized broadcast commands

- Dedicated Bus:

Unscalable IDC bandwidth; The timing and signal integrity issues
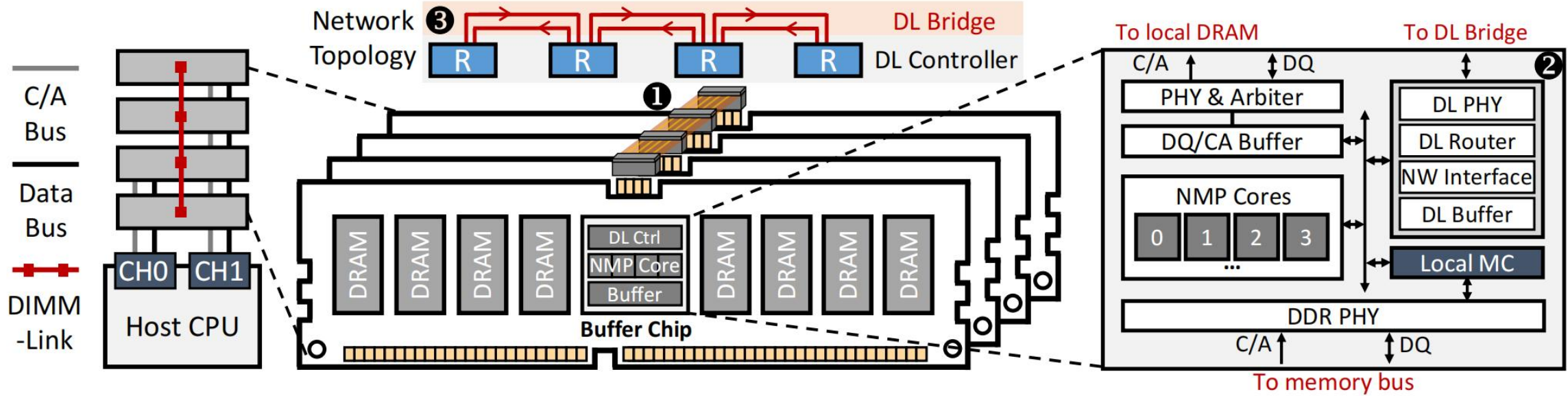
DIMM-LINK

Fig. 2. DIMM-Link Architecture.

DL Bridge(SerDes) and DL Controllers

DL-Bridge and the routers in the connected DIMMs form a network that allows concurrent packet-based data transmission
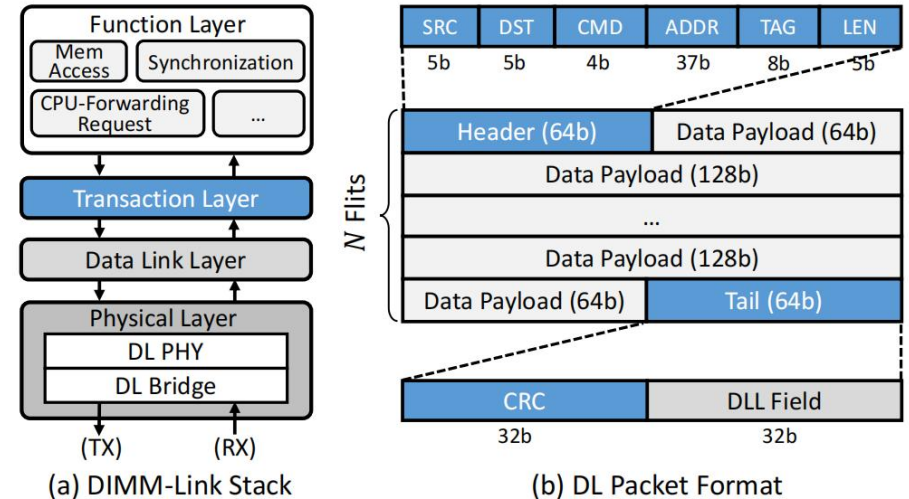
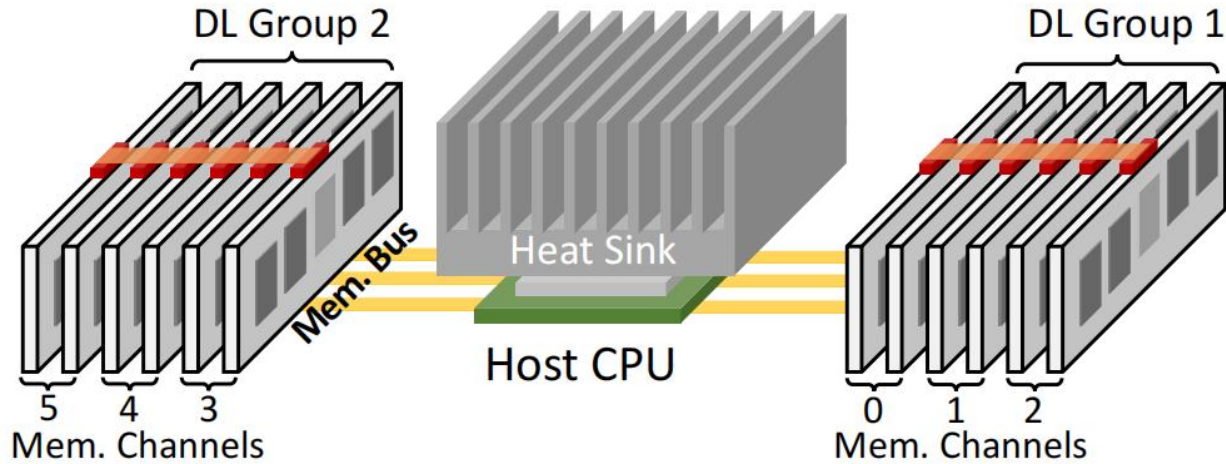## DIMM-LINK Protocol

Fig. 3. DIMM-Link Protocol.
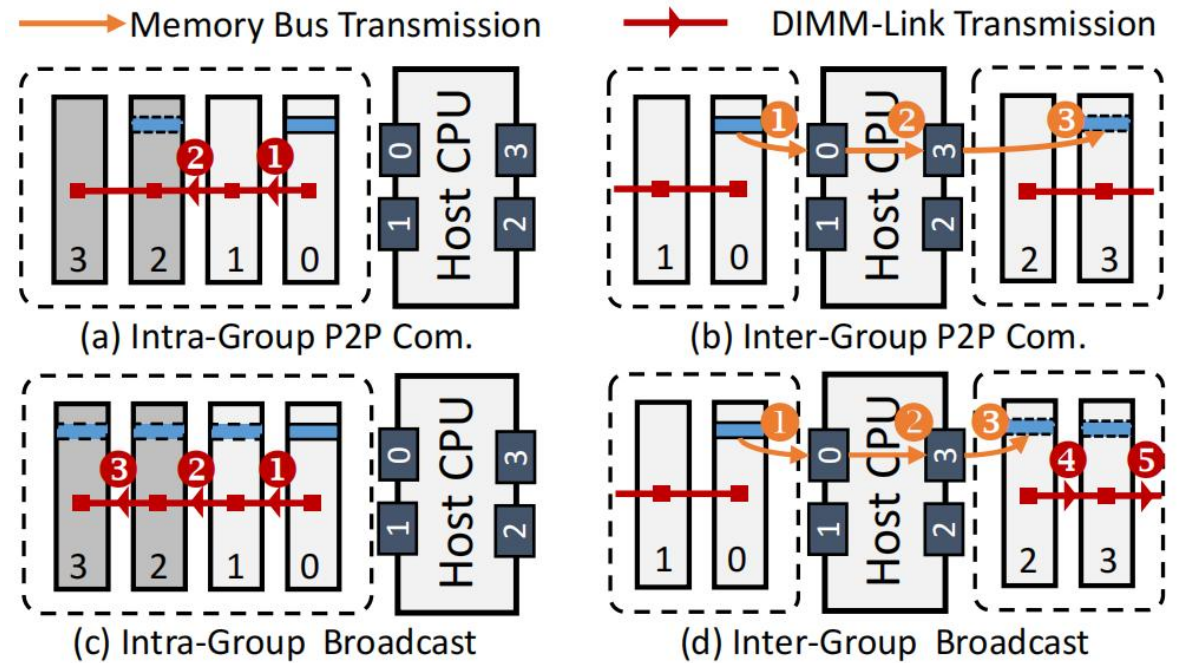
DIMM-LINK

Fig. 4. An Illustration of DIMM-Link Groups.

DL Group 2 · Heat Sink · Host CPU · DL Group 1

5 4 3 Mem. Channels · 0 1 2 Mem. Channels

Memory Bus Transmission DIMM-Link Transmission

(a) Intra-Group P2P Com. (b) Inter-Group P2P Com.
(c) Intra-Group Broadcast (d) Inter-Group Broadcast

Host CPU

Fig. 5. Four Inter-DIMM Communication Patterns.

Why we need DIMM-Link Group?

• the DIMMs on different sides are not directly connected

Hybrid Routing

TABLE II
SERDES TECHNIQUES COMPARISON

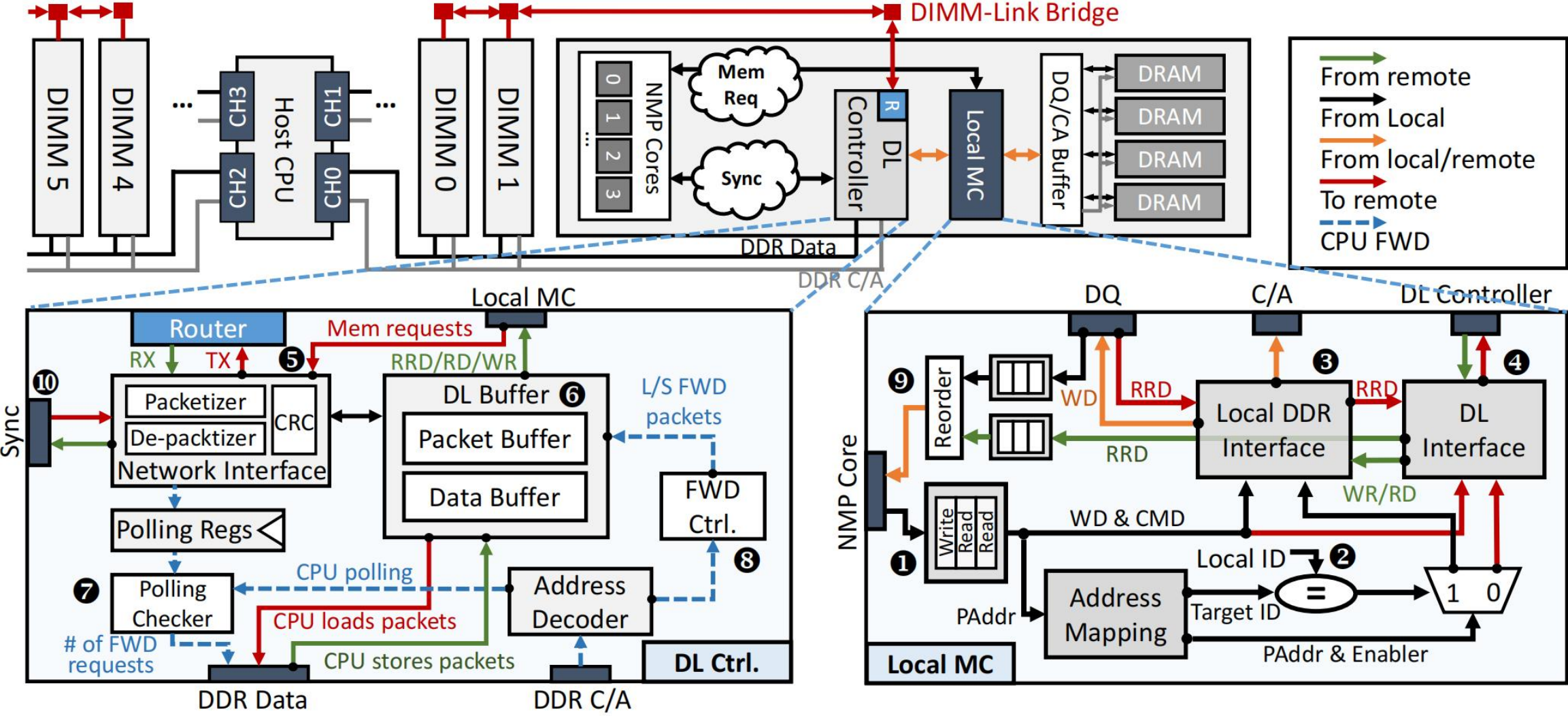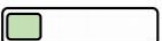| Reference	[10]	[25]	[69] (GRS)
Media	SMA Cable	Ribbon Cable	PCB
Singal Rate	6Gb/s/pin	16Gb/s/pin	25Gb/s/pin
Reach	953mm	500mm	80mm
Energy Eff. (pJ/b)	0.58	2.58	1.17

DIMM-LINK

Fig. 6. Detailed Architecture Design of DIMM-Link.

OPTIMIZATIONS

TABLE III
COMPARISONS OF POLLING MECHANISMS

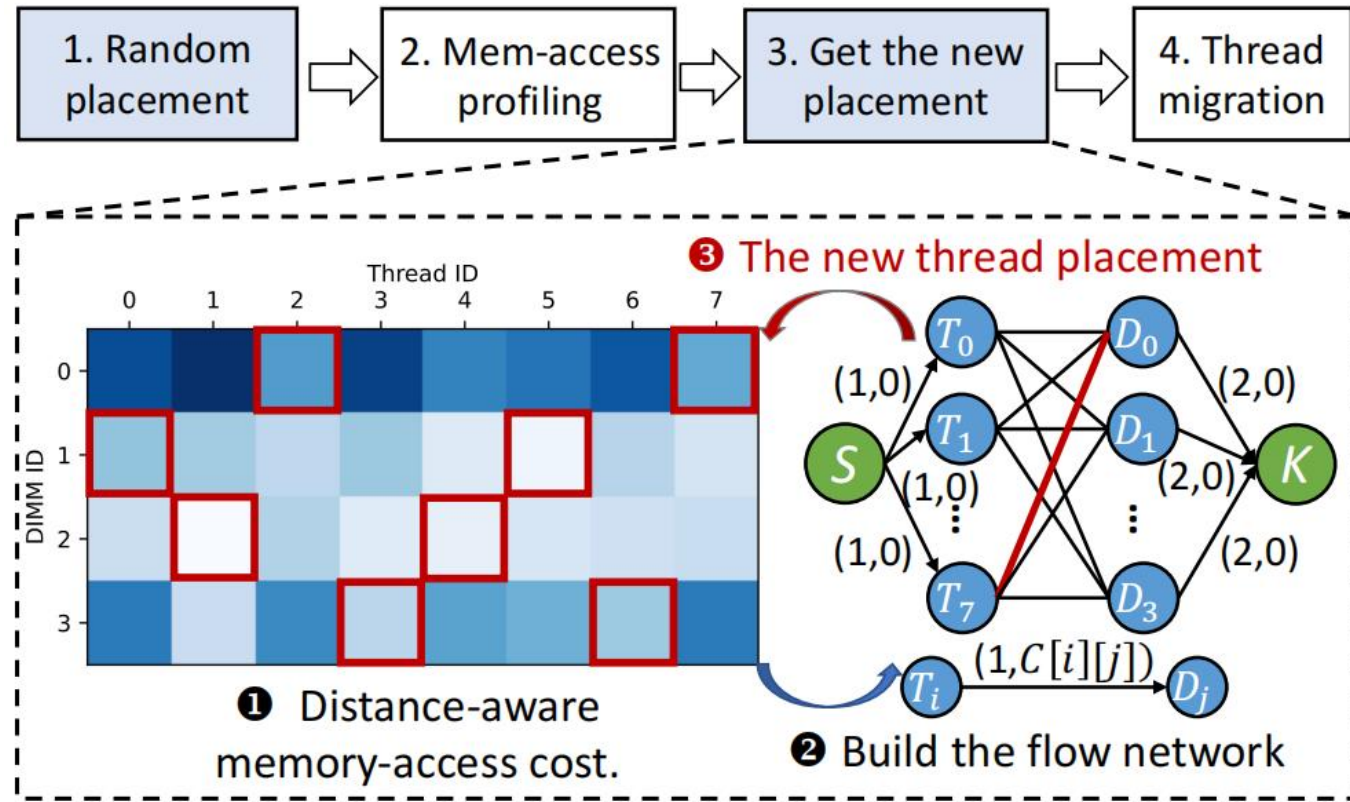| Methods	On-demand Polling?	CPU-Polling Range	Overhead	Latency
Baseline polling	✗	All DIMMs		
Baseline polling + Interrupt	✓	DIMMs in all interrupting channels		
Polling Proxy	✗	One DIMM per group		
Polling Proxy + Interrupt	✓	One DIMM per interrupting group		

Fig. 8. Distance-Aware Thread Placement.

minimum-cost maximum-flow

- The Polling Proxy Mechanism

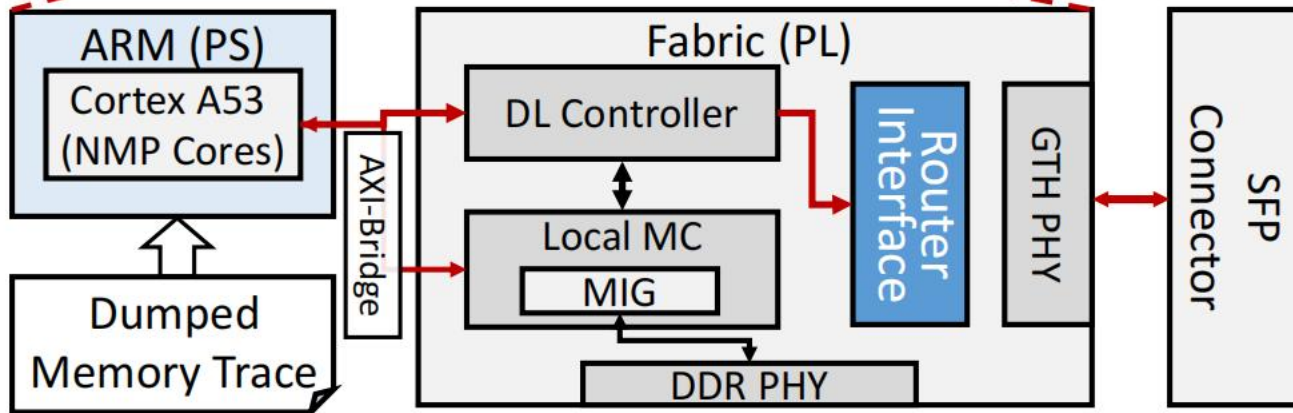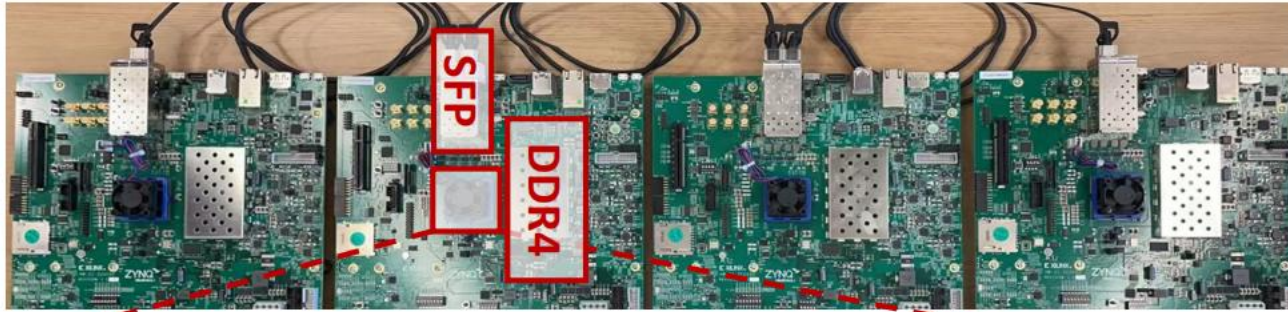- Distance-Aware Task Mapping

EVALUATION

Fig. 9. The FPGA Prototype.

Baselines	Description
MCN	CPU forwarding
AIM	dedicated memory bus
ABC-DIMM	broadcasts data within each memory channel

TABLE IV
BENCHMARKING APPLICATIONS

Name	Abbr.	Name	Abbr.
Breadth-First Search	BFS	Needleman-Wunsch	NW
Hotspot	HS	PageRank	PR
K-Means	KM	Single Source Shortest Path	SSSP

- FPGA-based prototype to validate its function

- Evaluate DIMM-Link's performance against three IDC baselines through simulation
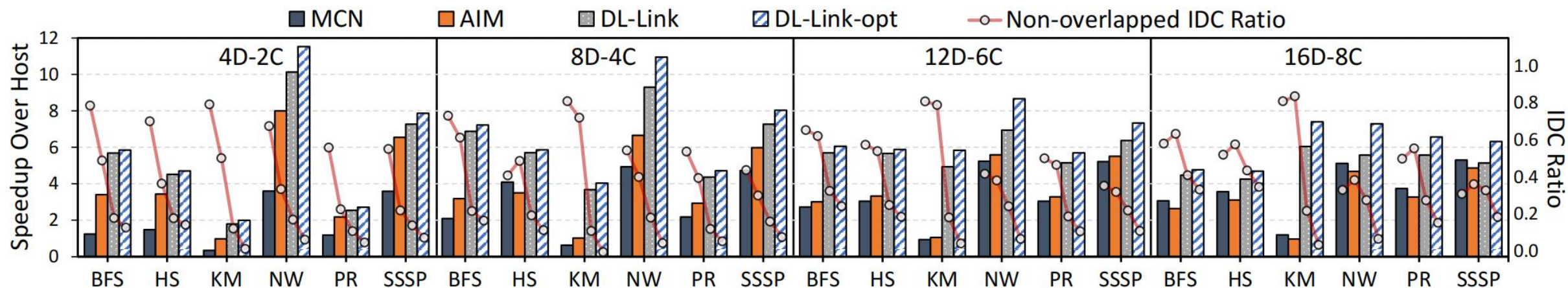
EVALUATION

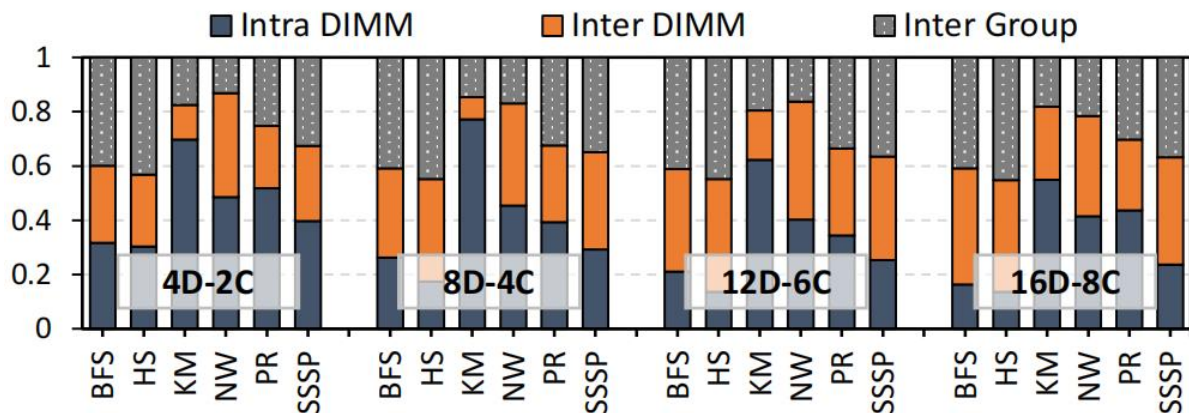Fig. 10. Performance and Non-overlapped IDC Ratio Comparisons.

Fig. 11. Data Transfer Breakdown of DL-Link-opt.

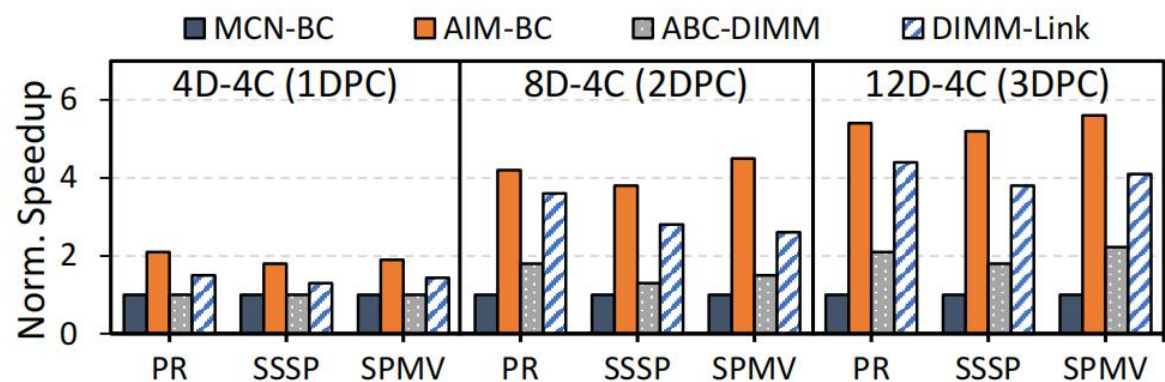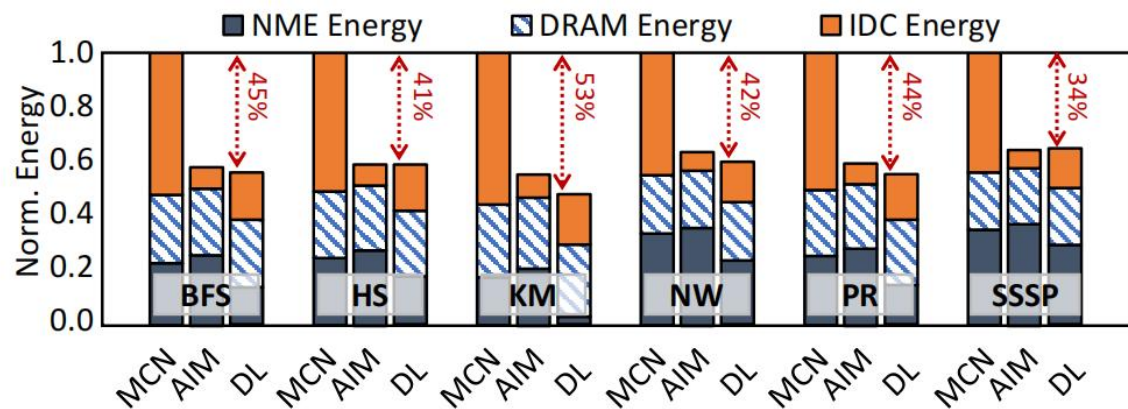Fig. 12. Broadcast Performance Comparison.

EVALUATION

Fig. 13. Energy Consumption Breakdown.
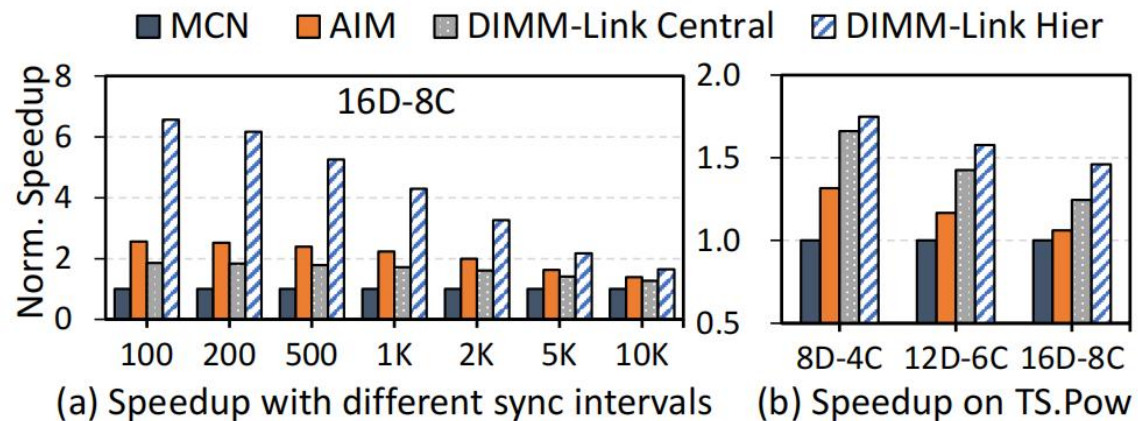

(a) End-to-End Performance   (b) Bus Occupation

Fig. 15. Performance with Different Polling Methods.

(a) Speedup with different sync intervals   (b) Speedup on TS.Pow
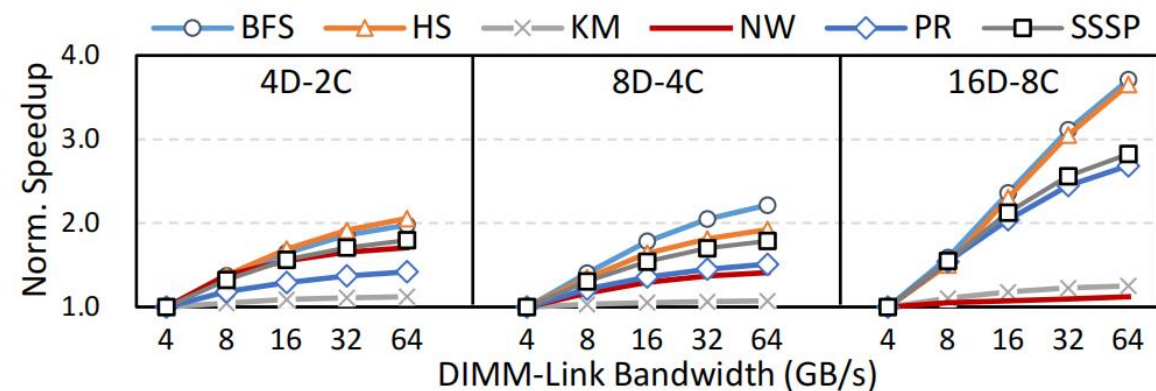
Fig. 14. Synchronization Performance.

Fig. 16. DIMM-Link Bandwidth Exploration

THANKS & QA