Design of PRAM-based Persistent NVDIMM Controllers to Prepare the Data Age

Myoungsoo Jung

Computer Architecture and MEmory systems Lab.

School of Integrated Technology

Yonsei University

Flash Memory Summit 2018
Santa Clara, CA

Data Explosion

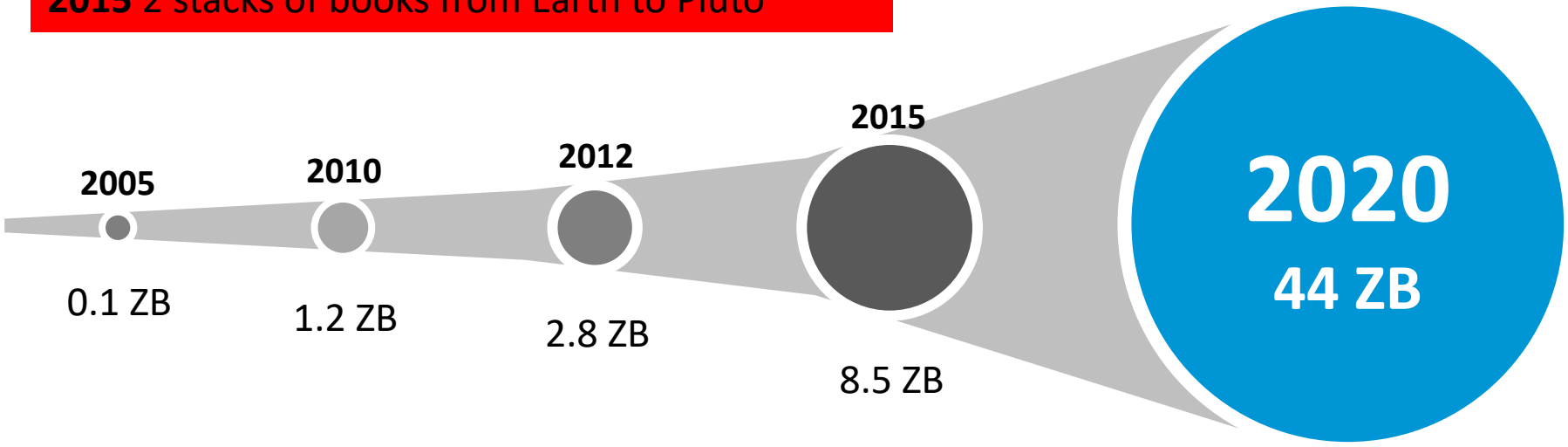
1900 London library (500,000 books)

1992 Internet explosion

2002 Less than 0.1% is stored on paper

2005 12 stacks of books from Earth to Sun

2015 2 stacks of books from Earth to Pluto



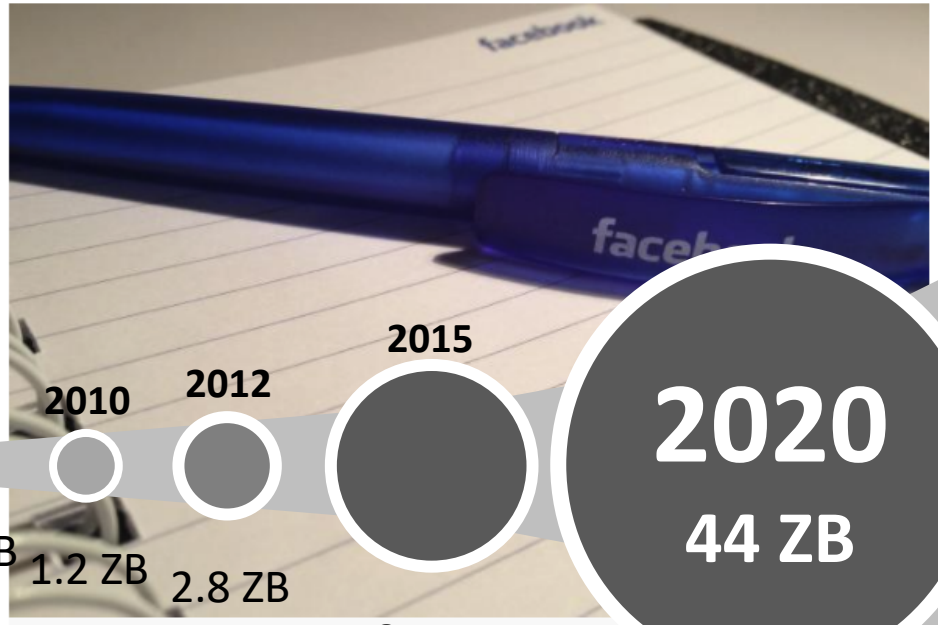
(1) IDC "Worldwide Internet of Things (IoT) 2013-2020 forecast" October 2013. (2) IDC "The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things" April 2014 (3) Global Smart Meter Forecasts, 2012-2020. Smart Grid Insights (Zypryme), November 2013 (4) http://en.wikipedia.org

Data Explosion

TECHNOLOGY

Facebook Privacy: Social Network Buys Data From Third-Party Brokers To Fill In User Profiles

BY **AJ DELLINGER** ON 12/28/16 AT 4:57 PM



Facebook pen and notebook Photo by [illegible]/Flickr

It comes as no surprise to any Facebook user that the social network gathers a considerable amount of information based on their actions and interests. But according to a report from ProPublica, the world's largest social network knows far more about its users than just what they do online.

Data

becomes the new currency

2020
44 ZB

2025
163 ZB

(IDC-2017)

Data Explosion

Flash Memory Summit

CAMELab.org
Computer Architecture and Memory Systems Laboratory

Storage capacity shipped **across** **dia types**
(HDD, Flash, SSD, Tape, Optic DRAM)

2005 0.1 ZB

2010 1.2 ZB

2012 2.8 ZB

2015 8.5 ZB

2020 44 ZB

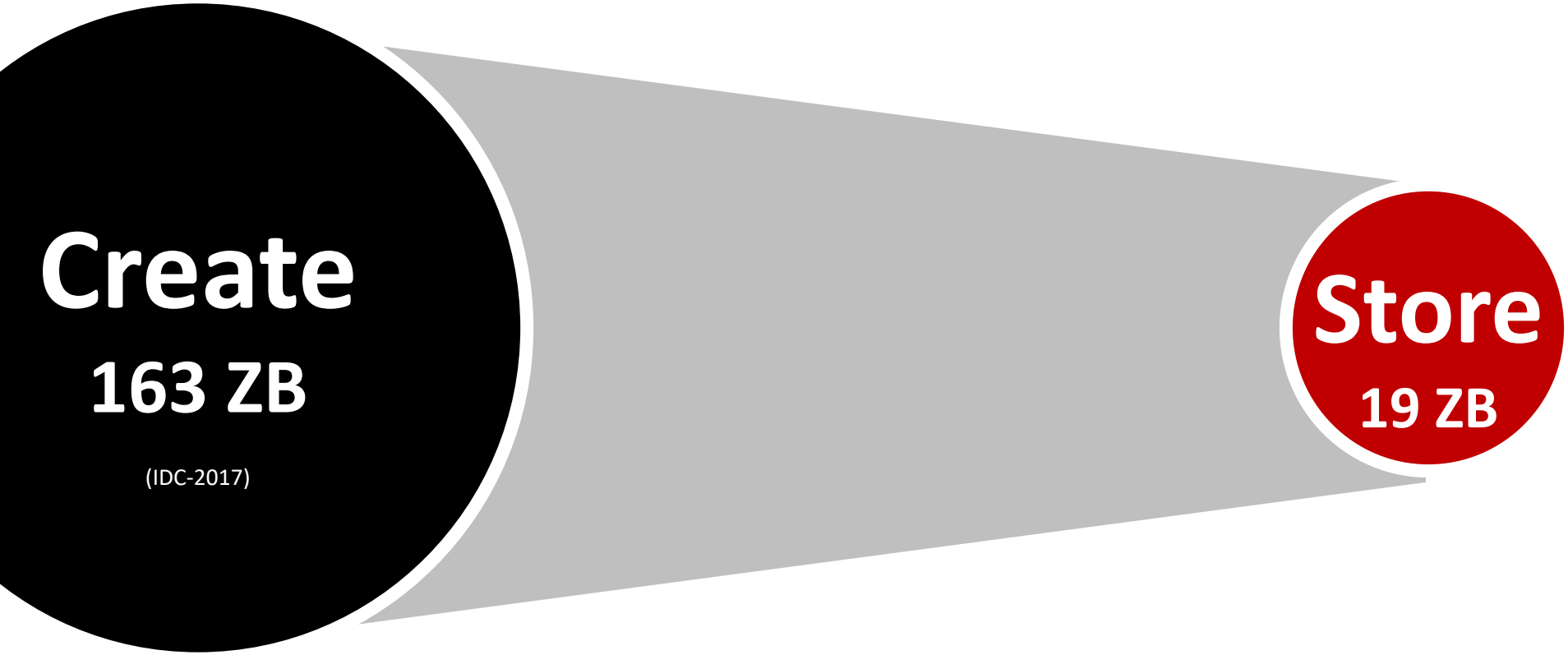
2025 163 ZB

(IDC-2017)

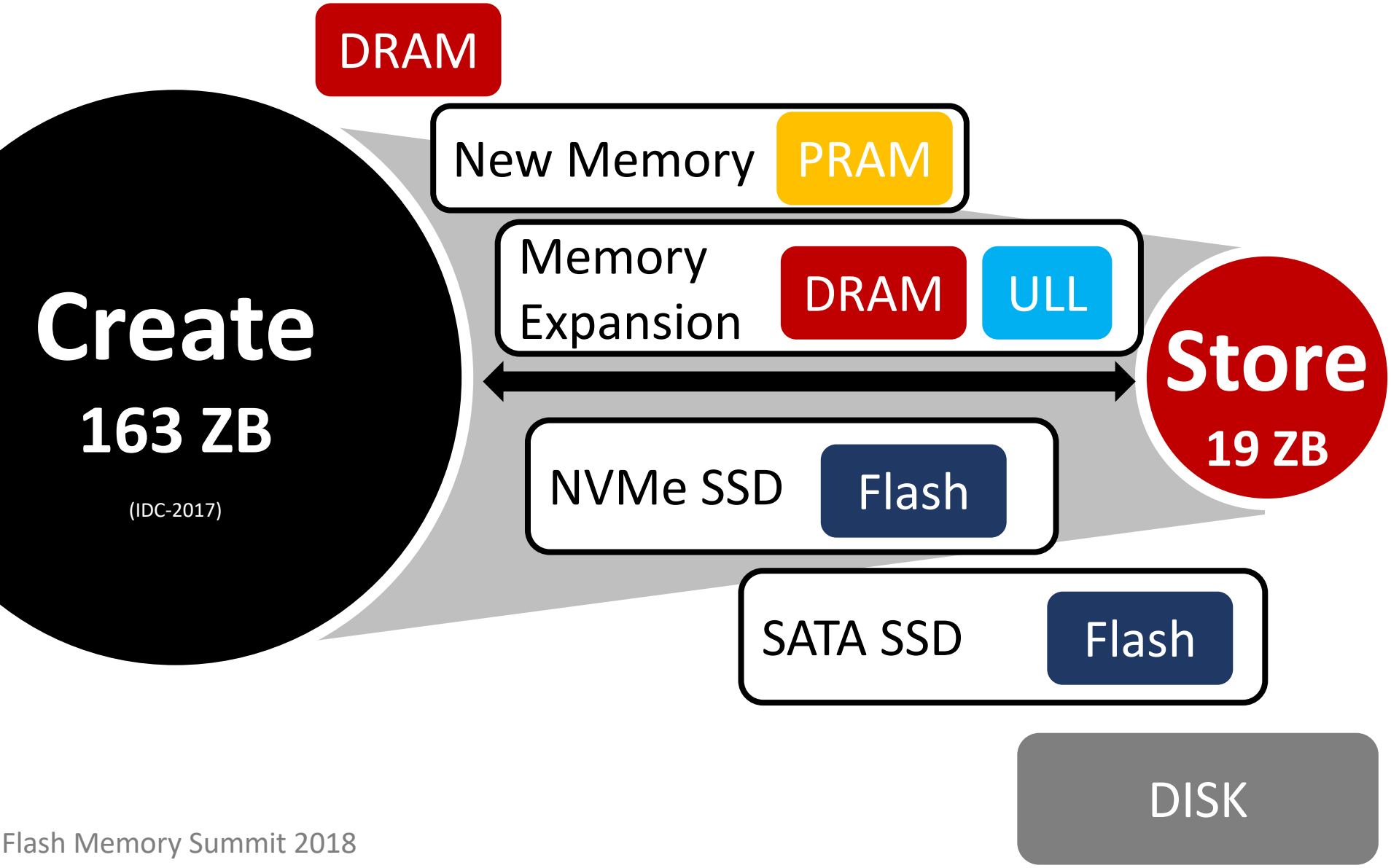
2025 19 ZB

GAP

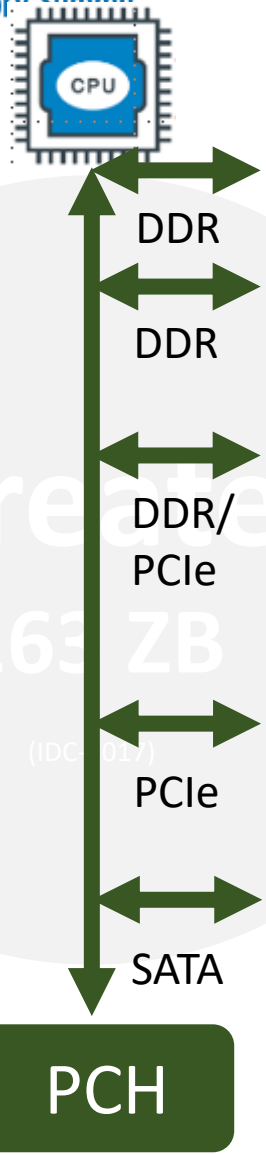
What Can System/Architecture Help ?



What Can System/Architecture Help?



What Can System/Architecture Help?



New Memory: PRAM

Memory Expansion: DRAM, ULL

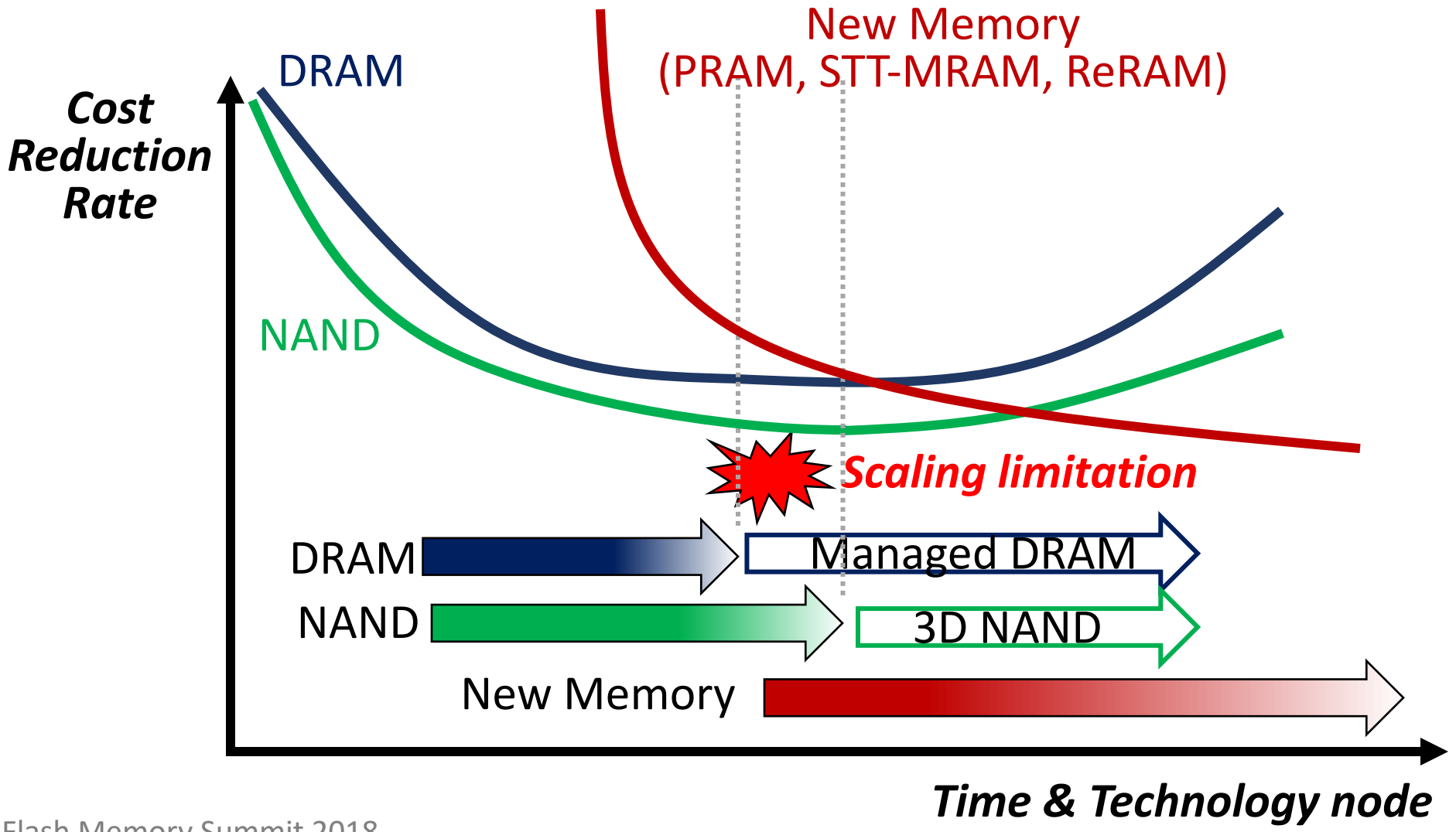
NVMe SSD: Flash

SATA SSD: Flash

- Make media more Reliable (ECC/LDPC)
- Addressing penalty
- Incarnating heterogeneous memory
- Hiding complexity
- Aggregating storage media
- Higher capacity
- Better performance
- Interface/Protocol

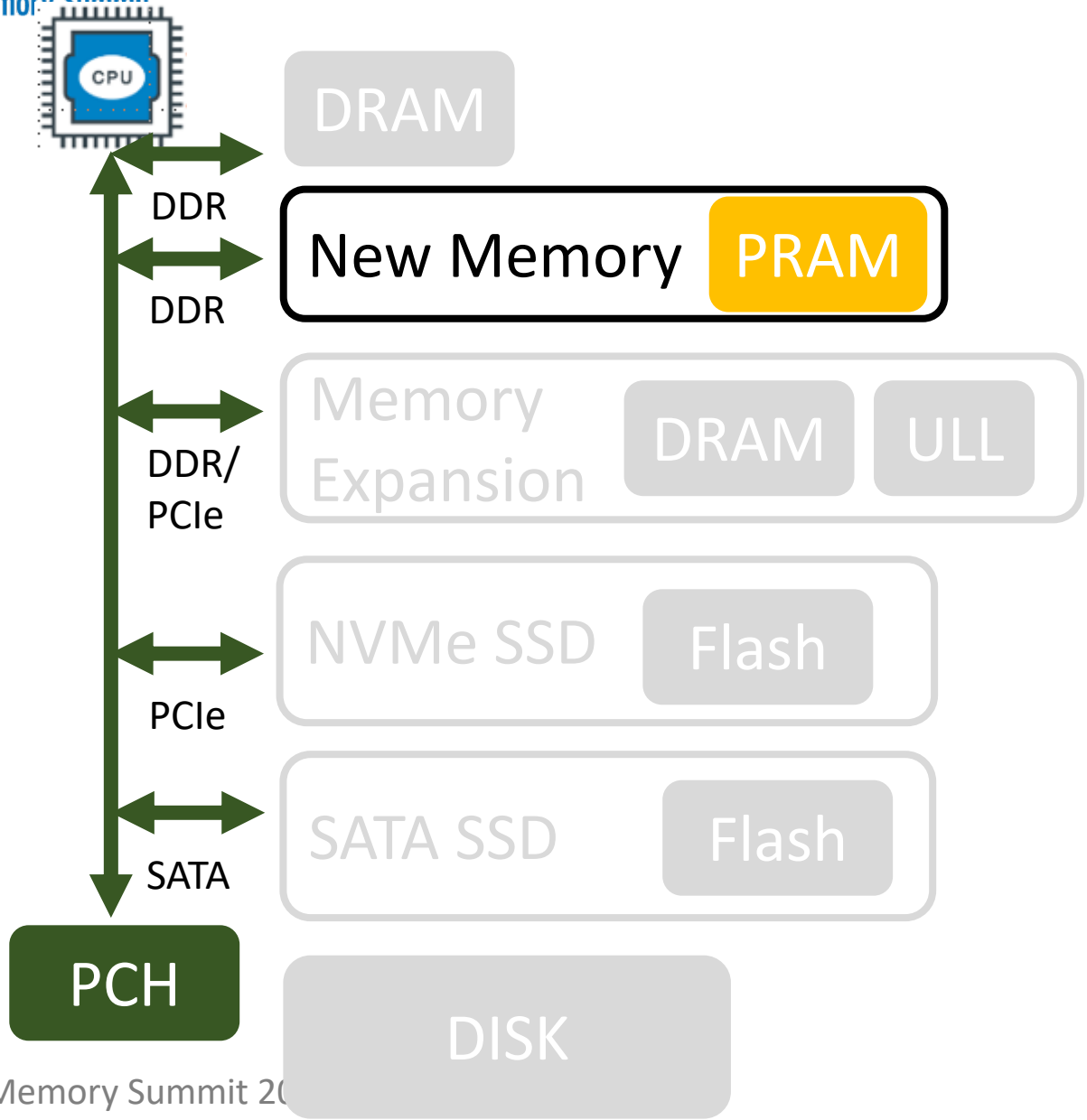
DISK

Cost scaling

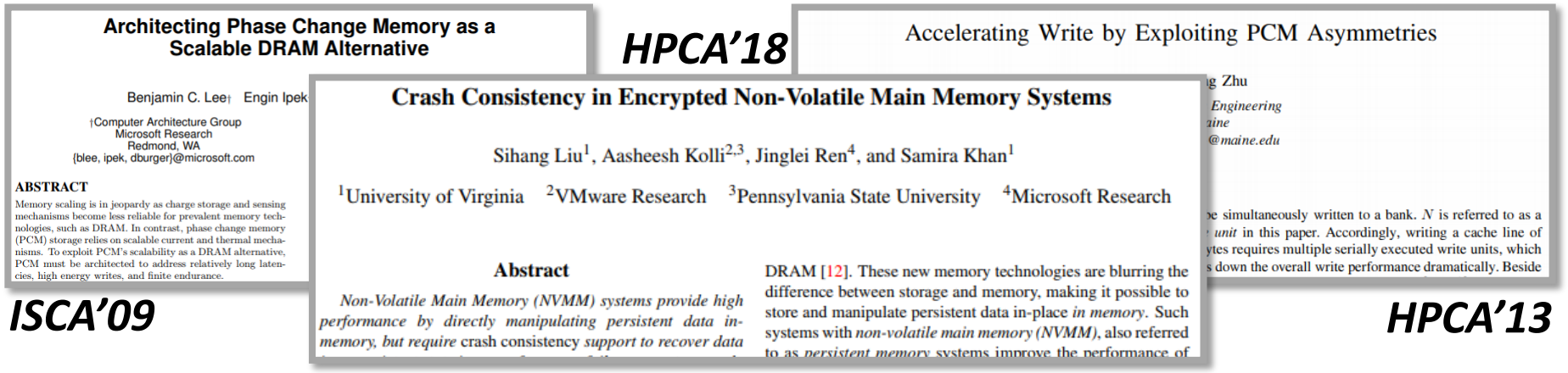


Source: SK Hynix

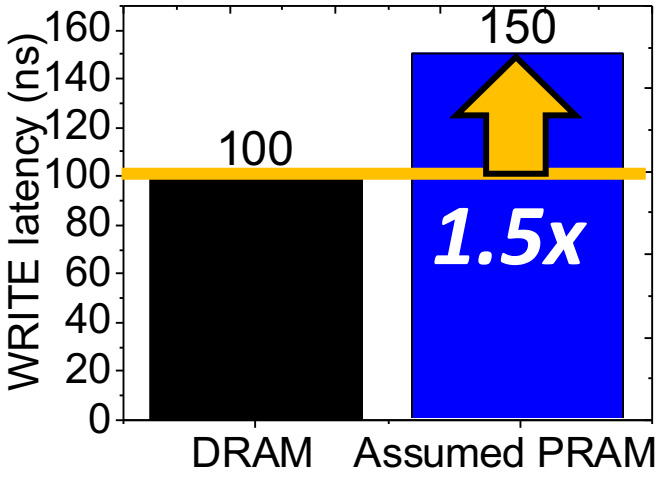
Memory Replacement



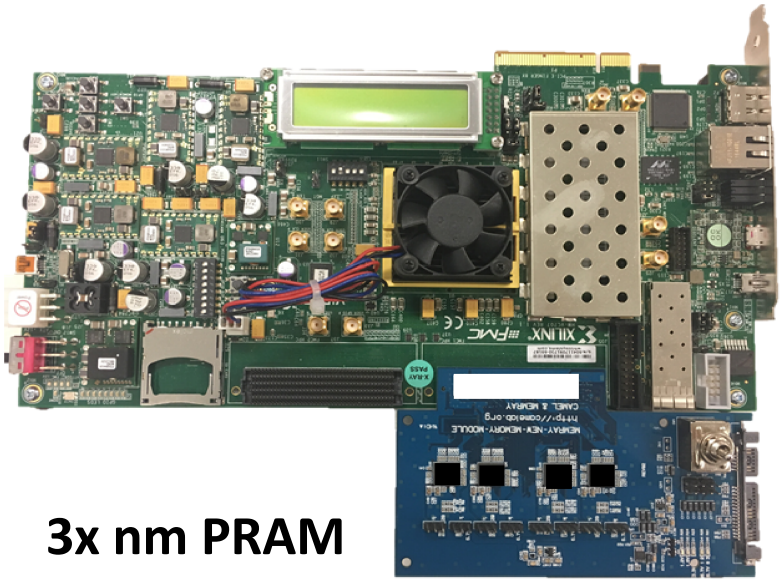
General Assumptions of PRAM Latency



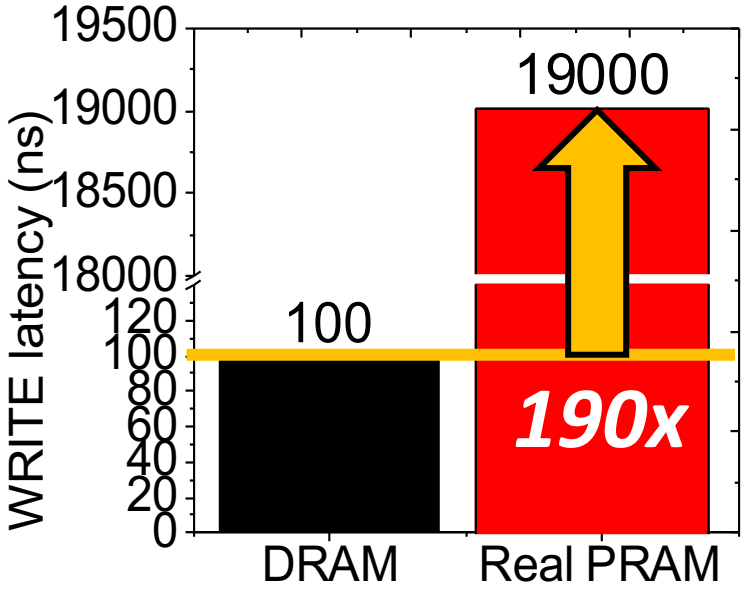
Many previous works assume PRAM's write latency as **similar to or slightly worse** than DRAM (**1.5x**)



PRAM Latency Measurement



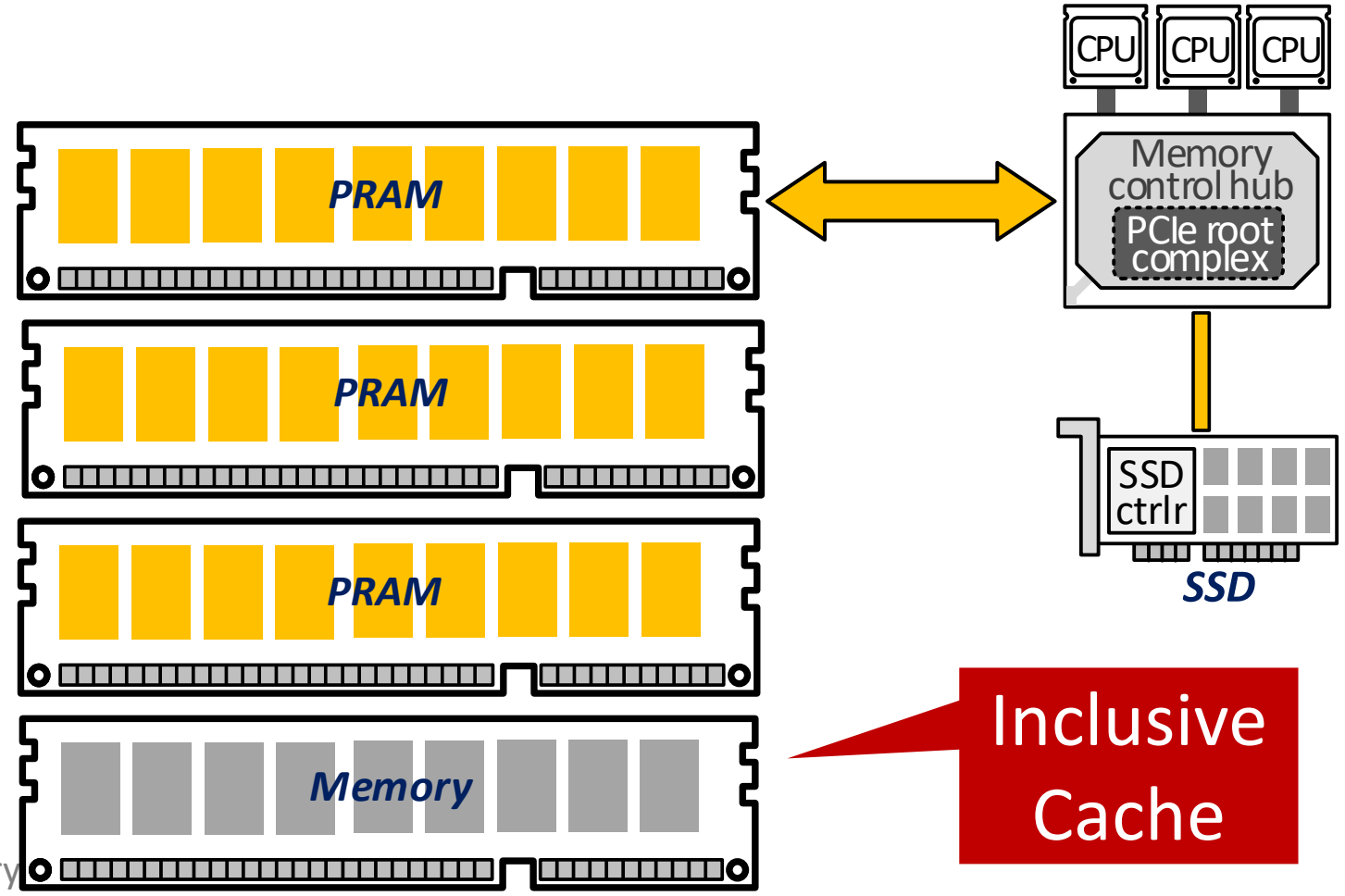
3x nm PRAM



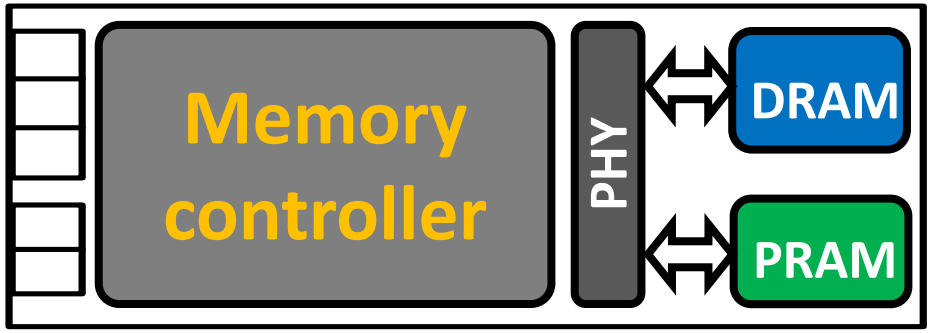
Our performance measurement on **real** 3x nm PRAM exhibits **expensive** write latency than DRAM (**190x**)

New Memory Placement

- PRAM offers promising read performance, but terrible write latency, compared to DRAM



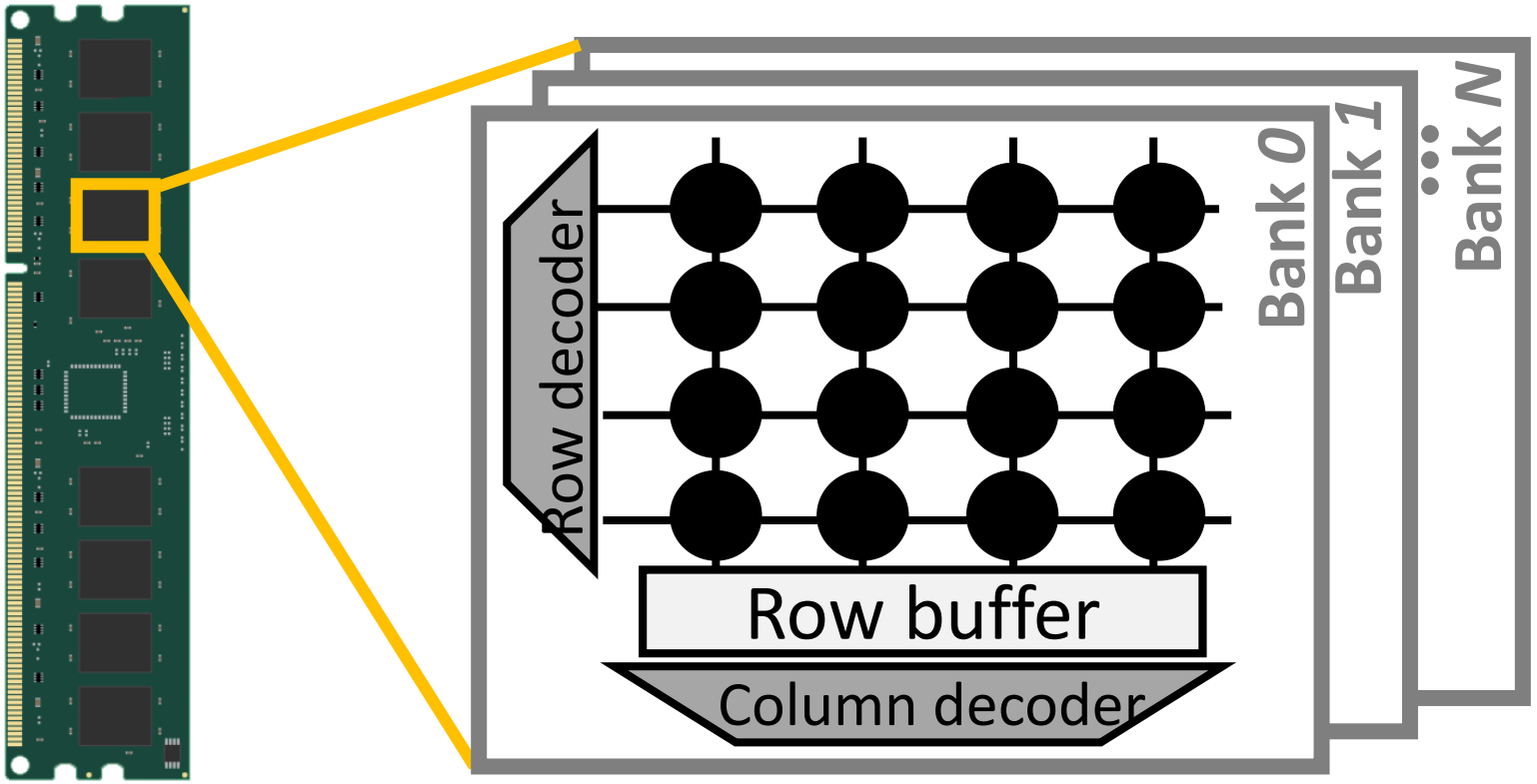
What Are The Considerations To Design Hybrid Memory (DRAM+PRAM) Controller?



To get insights of controller design, let's understand the details of DRAM and PRAM

DRAM's Multi-bank Architecture

- **Multiple banks** to serve multiple memory requests in parallel
- **Single row buffer** within a bank



Does PRAM have the Same Internal Architecture with DRAM?

Challenge1: PRAM's write latency is long

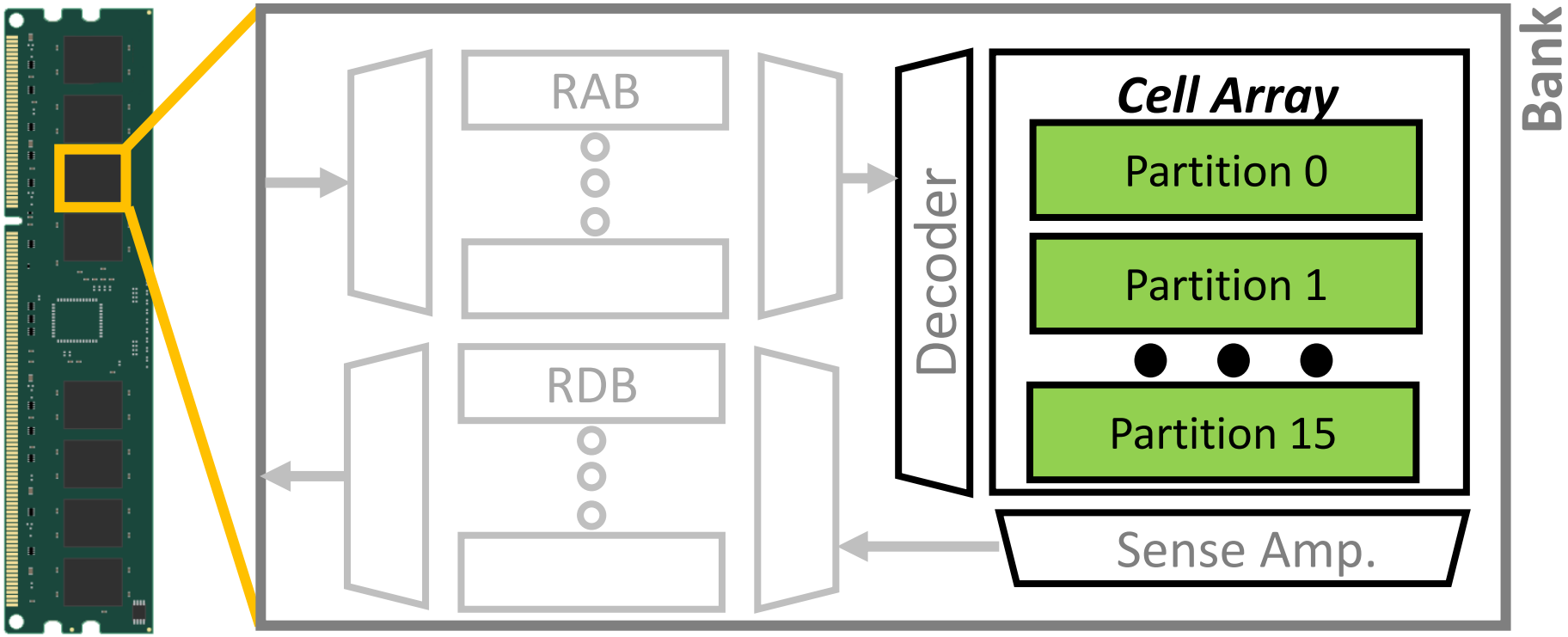
➔PRAM employs multiple row buffers

Challenge2: PRAM's asymmetric latency incurs lots of **bank conflicts**

➔PRAM uses multi-partition architecture

PRAM's Multi-partition Architecture

- **Multiple partitions within the bank** for partition-level parallelism
- **Multiple row buffers** to mitigate long write latency



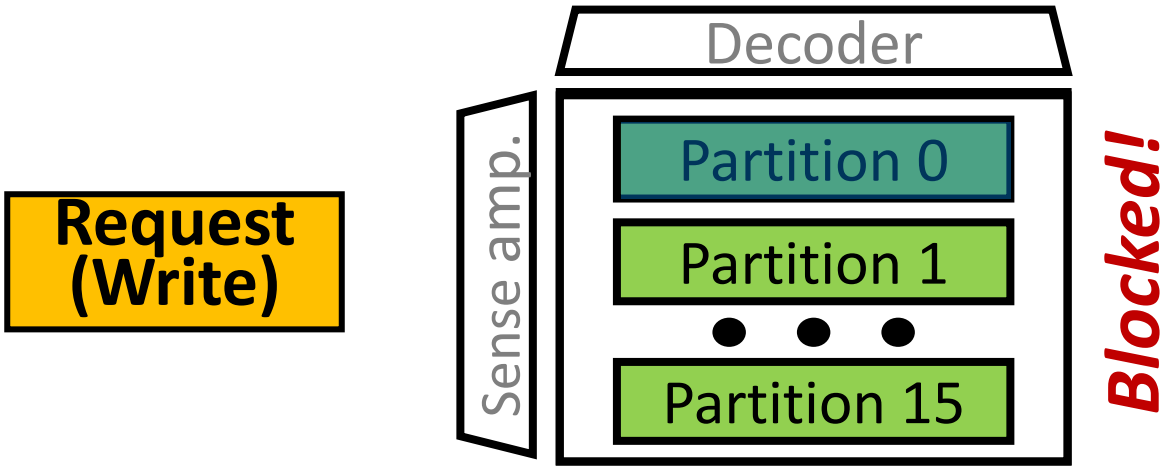
Can A Conventional DRAM Controller Be Aware of Multi-Partition? (Inside of A Bank)

(Revisited) Conventional DRAM scheduler just utilizes **bank-level parallelism**. Cannot see inside of bank!

∴ **Partition-level parallelism** should be supported

How Memory Requests Can Be Scheduled By Exploiting Multi-Partition Architecture?

Limitation of PRAM: In PRAM design, WRITE request blocks whole PRAM bank.

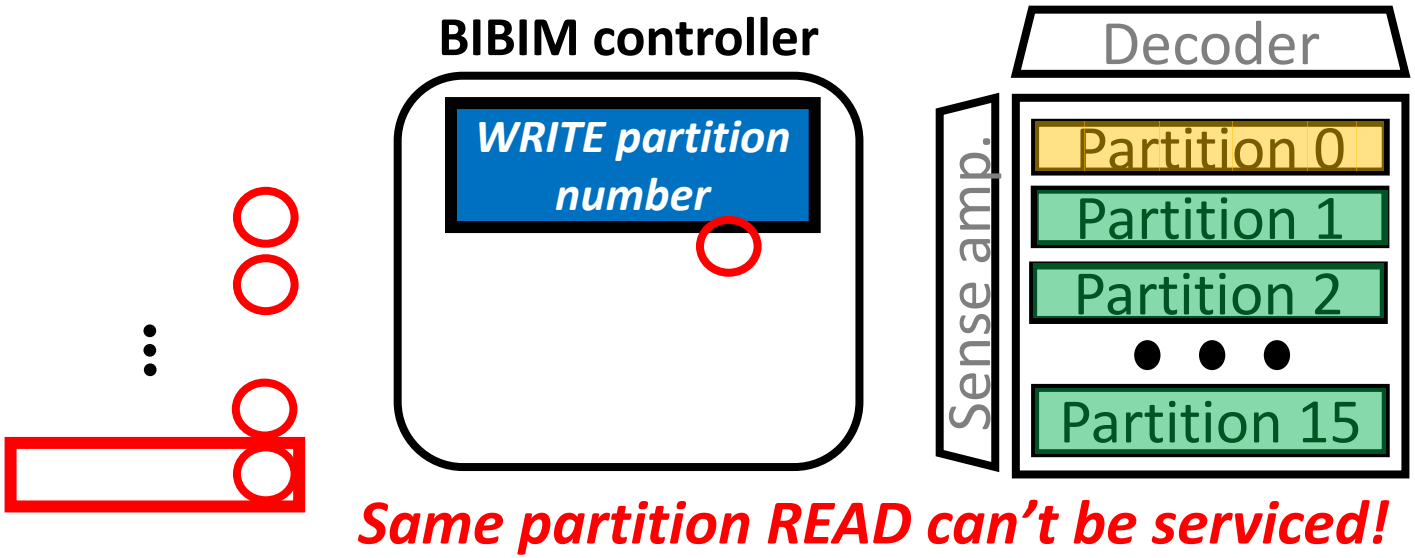


How Memory Requests Can Be Scheduled By Exploiting Multi-Partition Architecture?

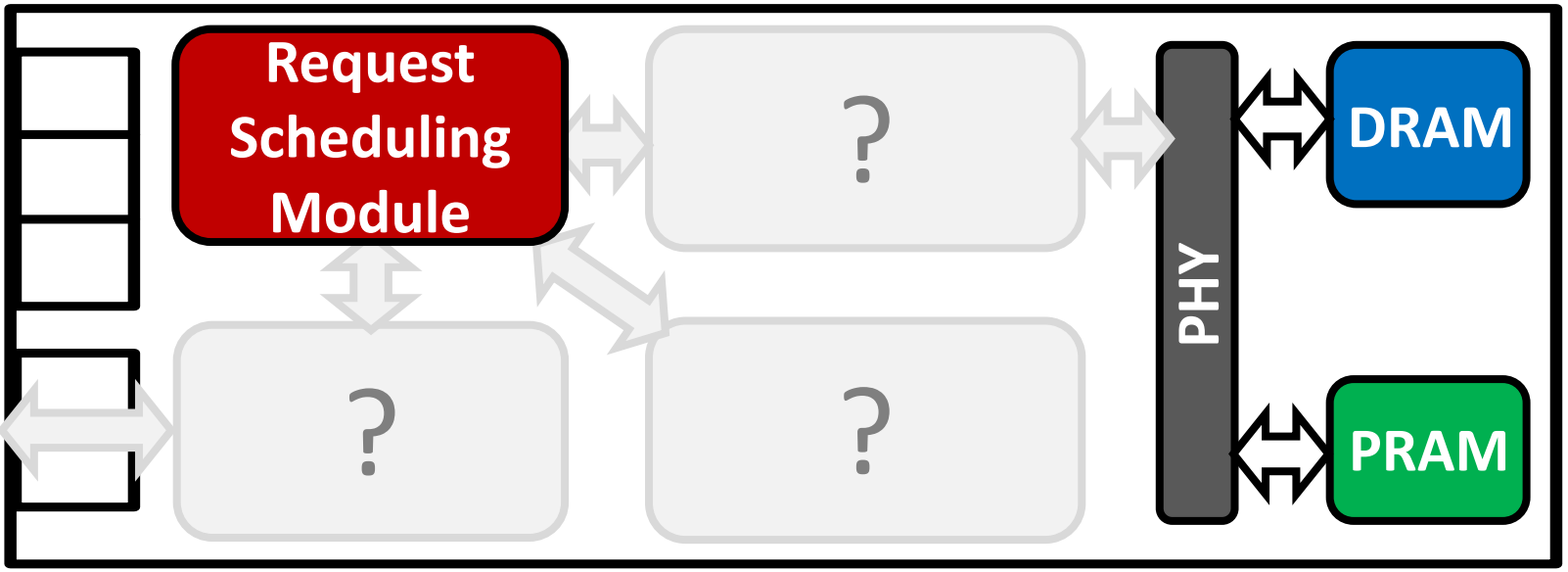
Key insight: Although WRITE cannot be serviced, READs can be serviced if partition number is different

Non-Blocking Read Service (NBRS)

Solution: Add a ***register*** to store 'partition number of WRITE' and compare it with partition number of incoming READ request



Design1: Scheduling Support Module for PRAM-aware New Scheduling Scheme

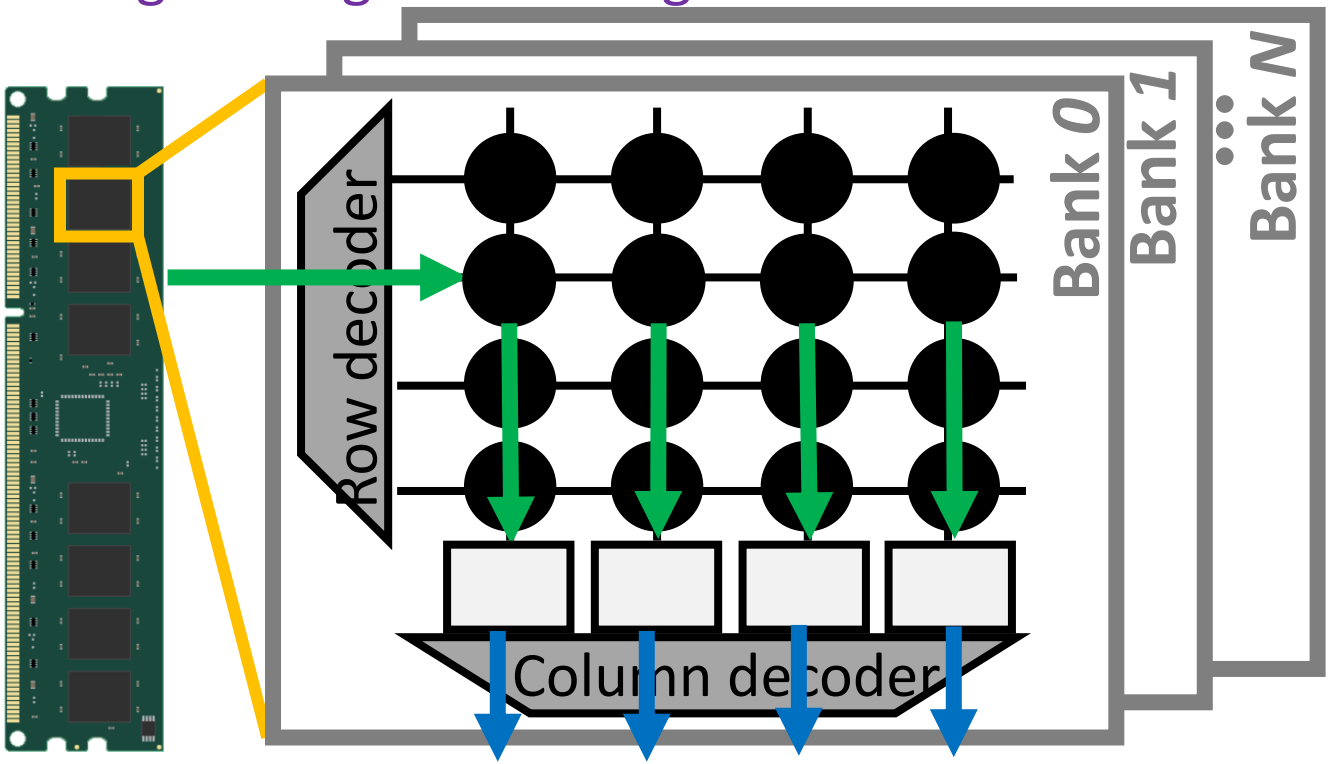


Now Requests Are Scheduled. Then, How Then Can It Serve to Hybrid Memory?

Firstly, as is generally known, LPDDR is JEDEC standard low-power memory interface (used for DRAM)

DRAM's Timing (LPDDR by JEDEC)

1) Activation: activate target row & write that data to row buffer
2) Read/Write: accessing row buffer with column address
3) Precharge: charge half-voltage of bit-line



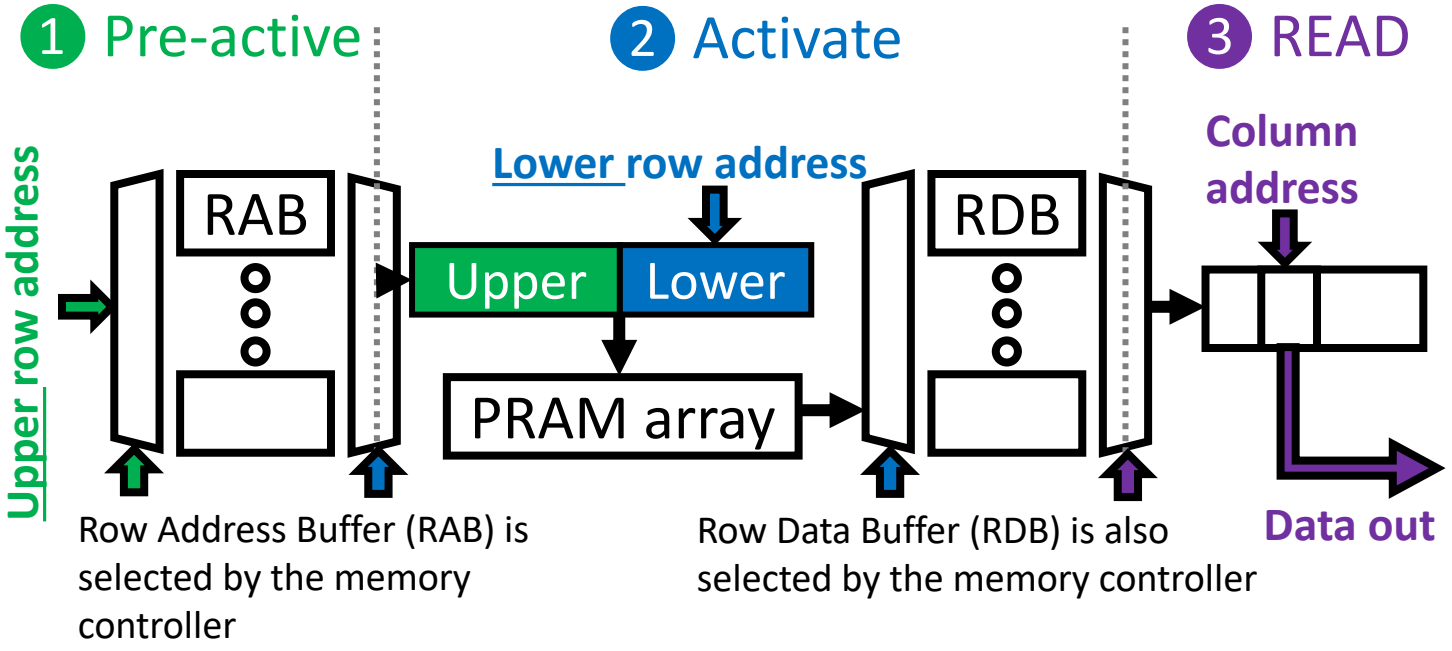
Does PRAM have the Same Memory Interface (LPDDR2) with DRAM?

(Revisited) PRAM has a different architecture with DRAM such as **Multiple row buffers** and **More larger capacity**

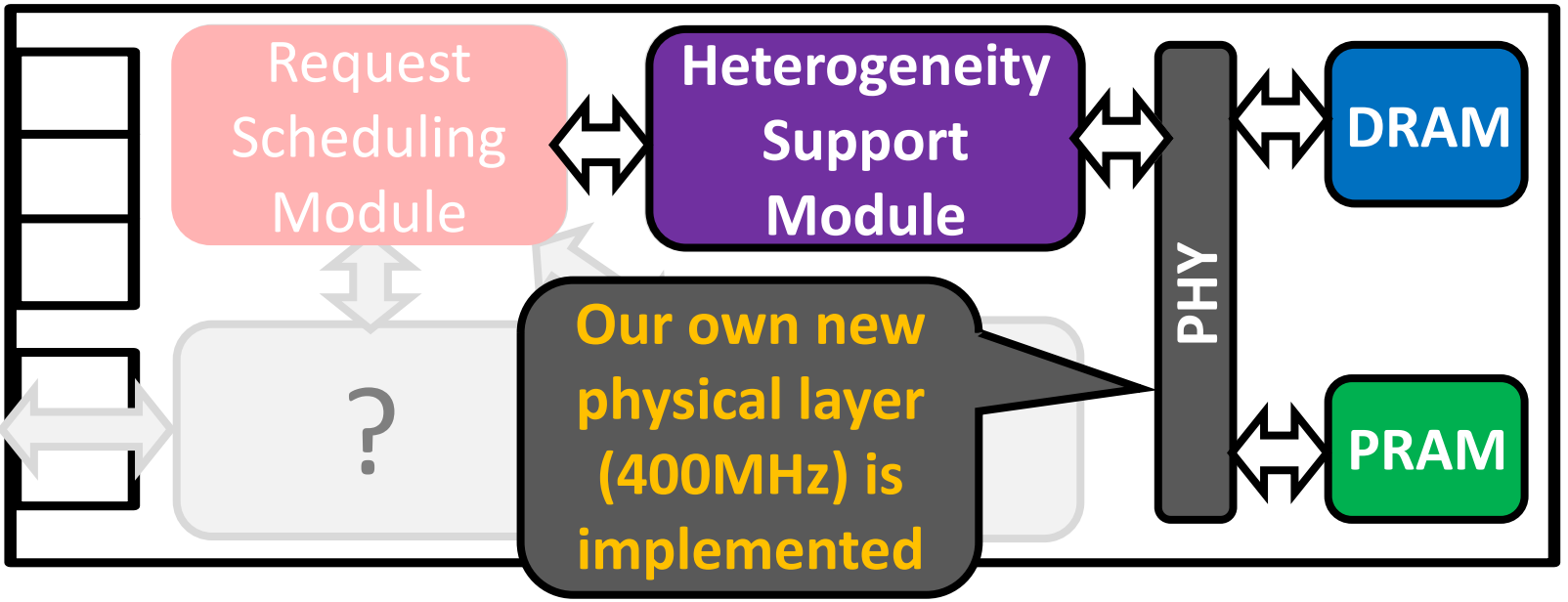
∴ Different interface is required

PRAM's Timing

- PRAM requires different timing model from DRAM
  - NVM memory space is much larger than a DRAM
- 3-Phase addressing (LPDDR-NVM by JEDEC)



Design2: Heterogeneity Support Module for both LPDDR & LPDDR-NVM



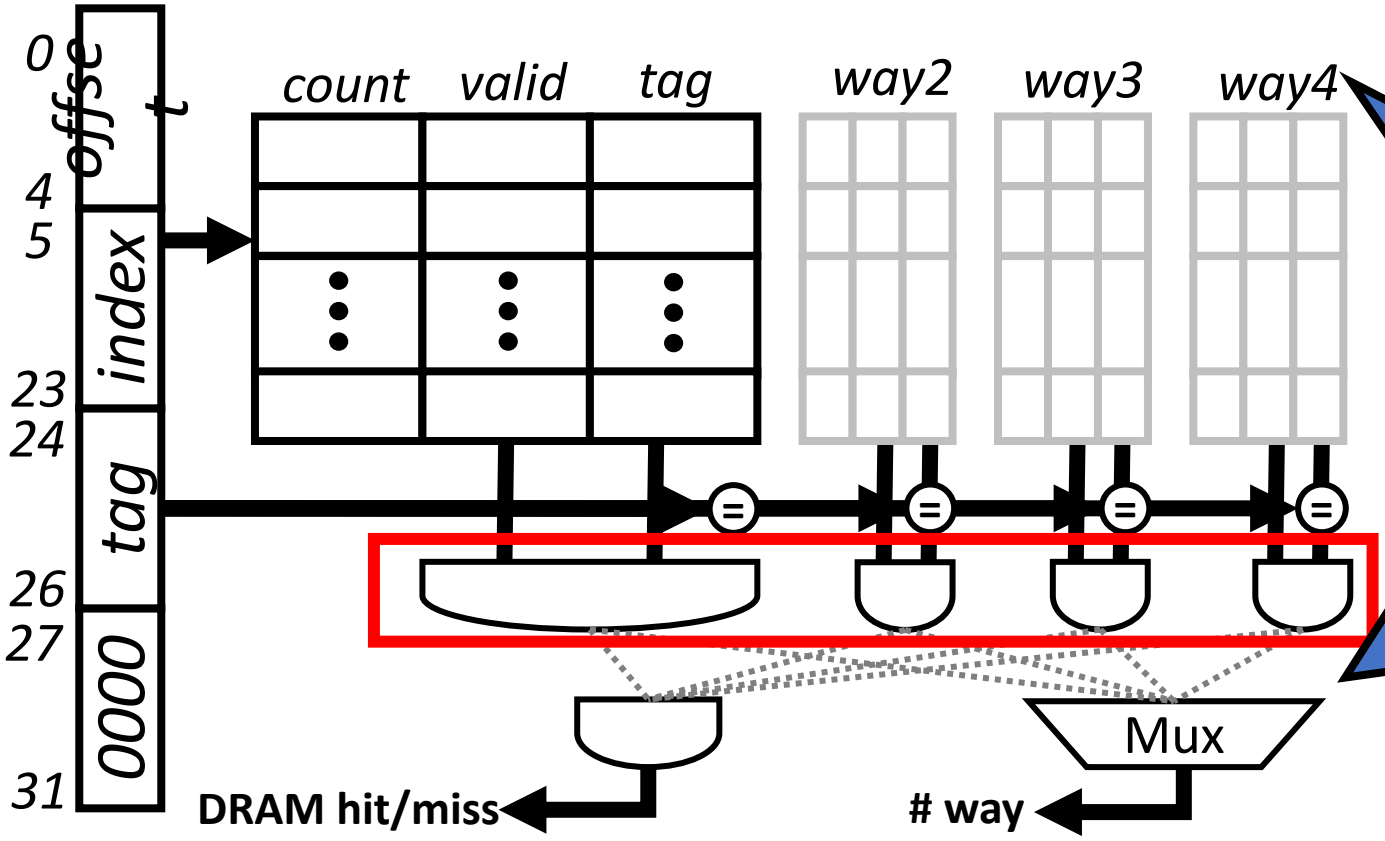
Don’t Forget DRAM is For Cache.
Then, How Caching Can Be Supported?

Solution: Keep which data exist in DRAM (caching info) in **lookup table**.

Moreover, like conventional cache, controller should have algorithms such as DRAM dataline update, eviction, and find empty dataline.

Hardware Support for Lookup Table

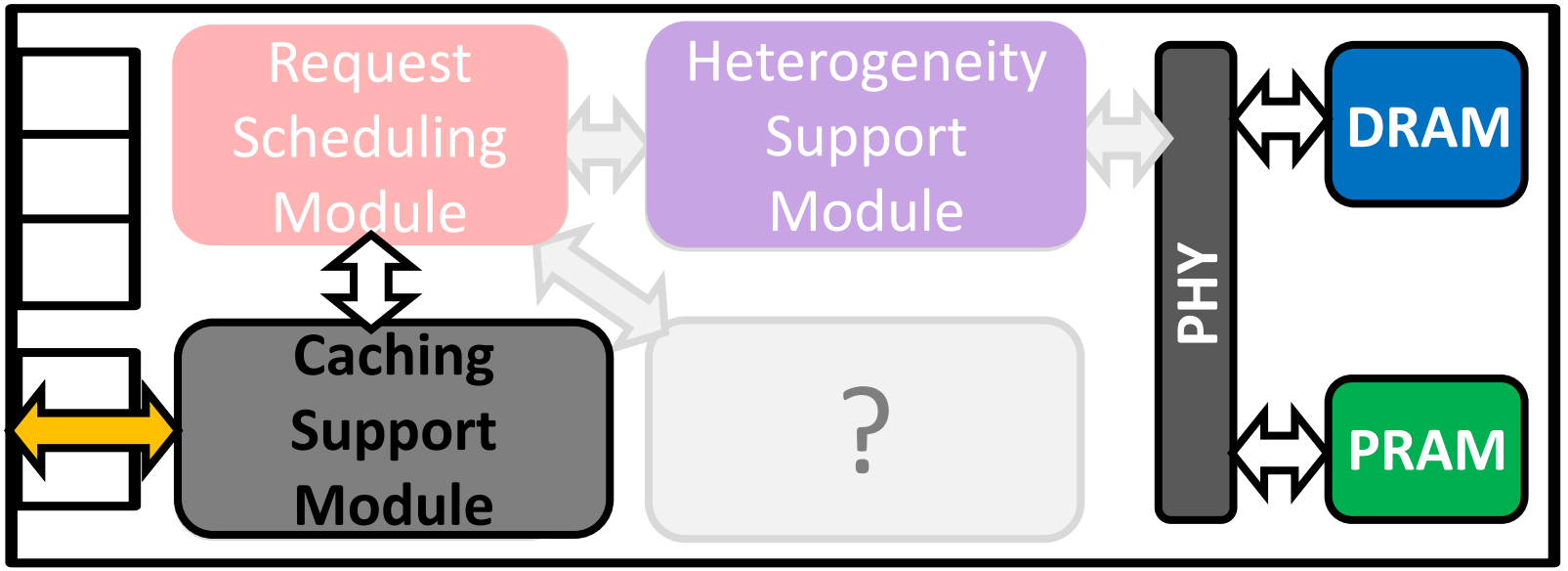
Lookup table do not include data value, includes address information



Only 512KB BRAM is used for lookup table

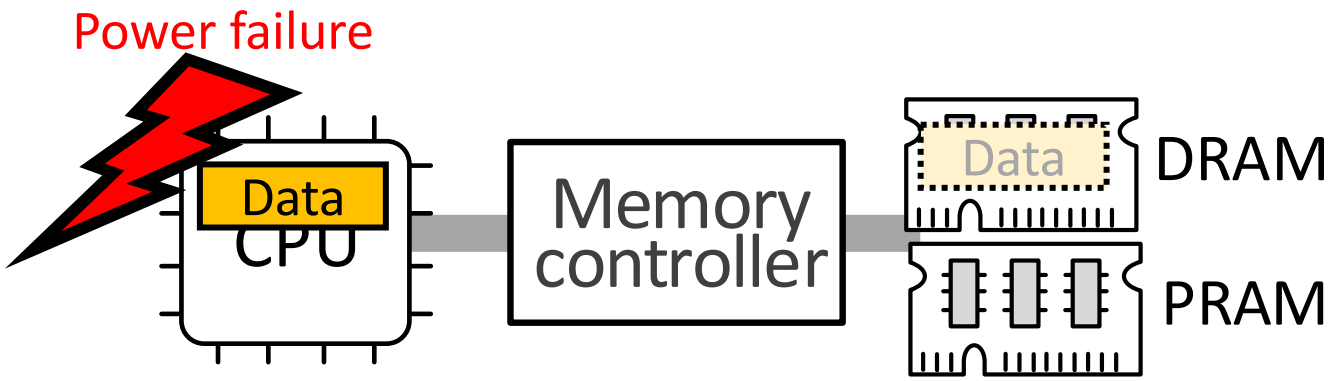
With multiple comparators, 4 ways can be parallelized

Design3: Caching Support Module for Use DRAM As Inclusive Cache of PRAM



BTW, How Non-Volatility of PRAM Can Be Maintained Although DRAM Is Integrated? (Hybrid)

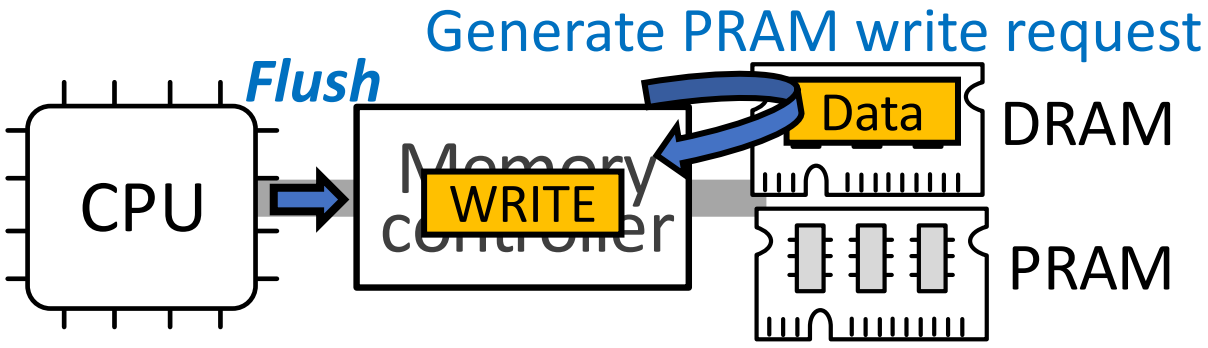
Challenge of hybrid memory: Data in DRAM will disappear when there is a power failure



FLUSH Operation

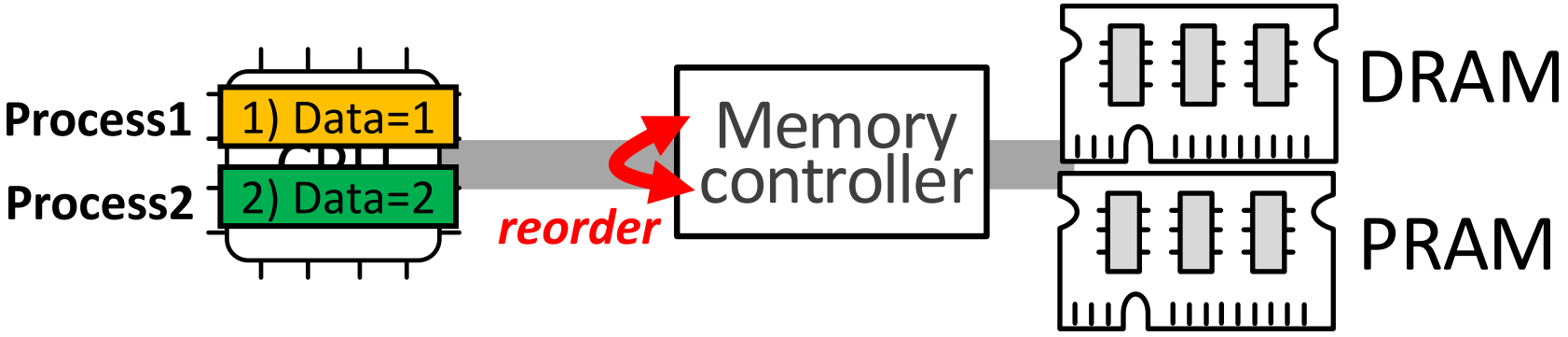
Solution: Provide ***Flush*** operation which moves DRAM data to PRAM. Memory controller generates 'PRAM write' request corresponding to the target DRAM row.

NOTE) 'PRAM write' will be stored in command queue which exists in memory controller. And DRAM dataline is invalidated.



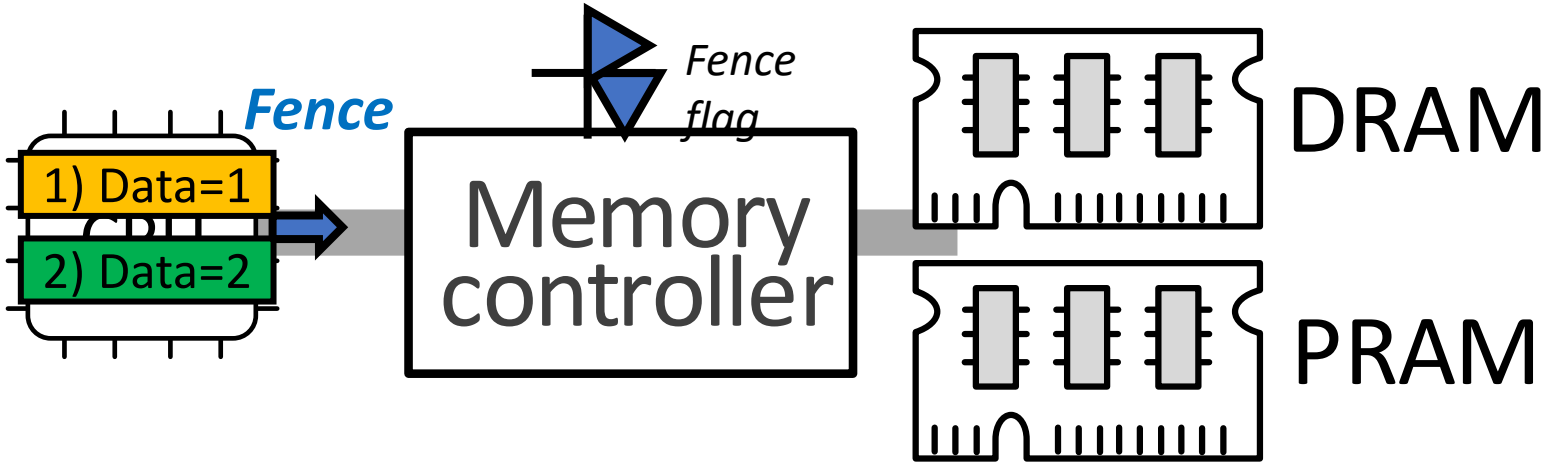
Okay, Data Delivery Is Guaranteed. Is It Good Enough?

Challenge of flush: User believes data has the latest value. But, the memory controller can reorder the order of memory request

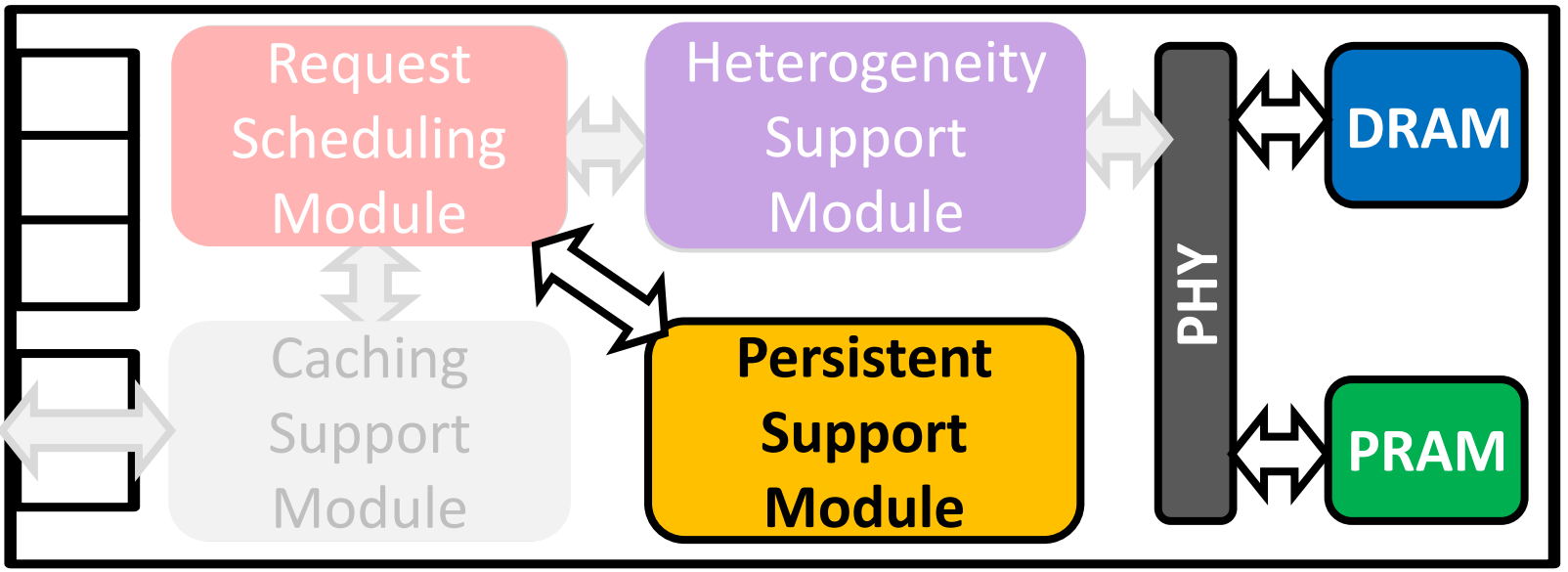


FENCE Operation

Solution: Provide ***Fence*** operation to enforce data delivery order of memory requests. The memory controller can simply add '*fence flag*' to check fenced or not



Design4: Persistent Support Module to Guarantee Data Delivery & Delivery Order

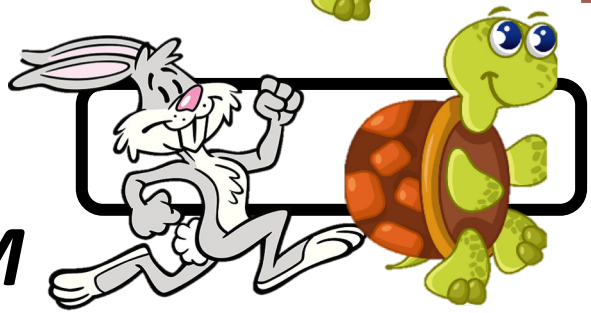


Demo (Track and Field)

DRAM

PRAM

P-NV DIMM

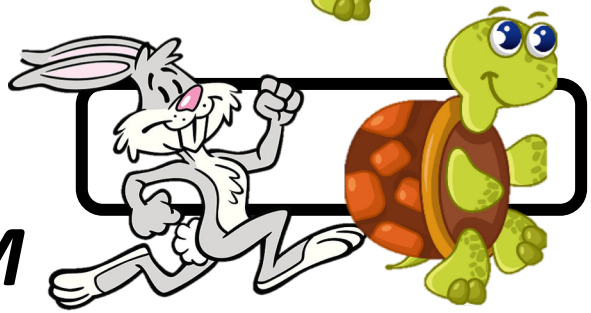


Performance comparison (memory access with synthetic, read-write inter-mixed trace)

Demo – Slow Version

DRAM

PRAM



P-NV DIMM

Demo – Normal Version

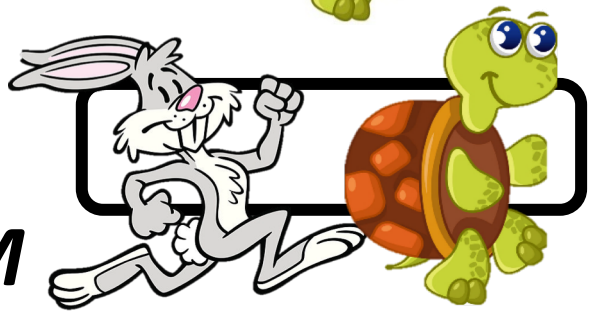
Flash Memory Summit

CAMELab.org
Computer Architecture and Memory Systems Laboratory

DRAM

PRAM

P-NV DIMM



0.105 sec

1.27 sec

0.468 sec

Further Enhancement

Flash Memory Summit

CAMELab.org
Computer Architecture and Memory Systems Laboratory

L2 cache
On-chip MC1
On-chip MC2

New Memory Controller

FPGA CTL (CH1)

FPGA CTL (CH0)

- ***Hybrid Cache Logic***
- ***Customized PRAM Phy***

New Memory Device

PRAM PHY

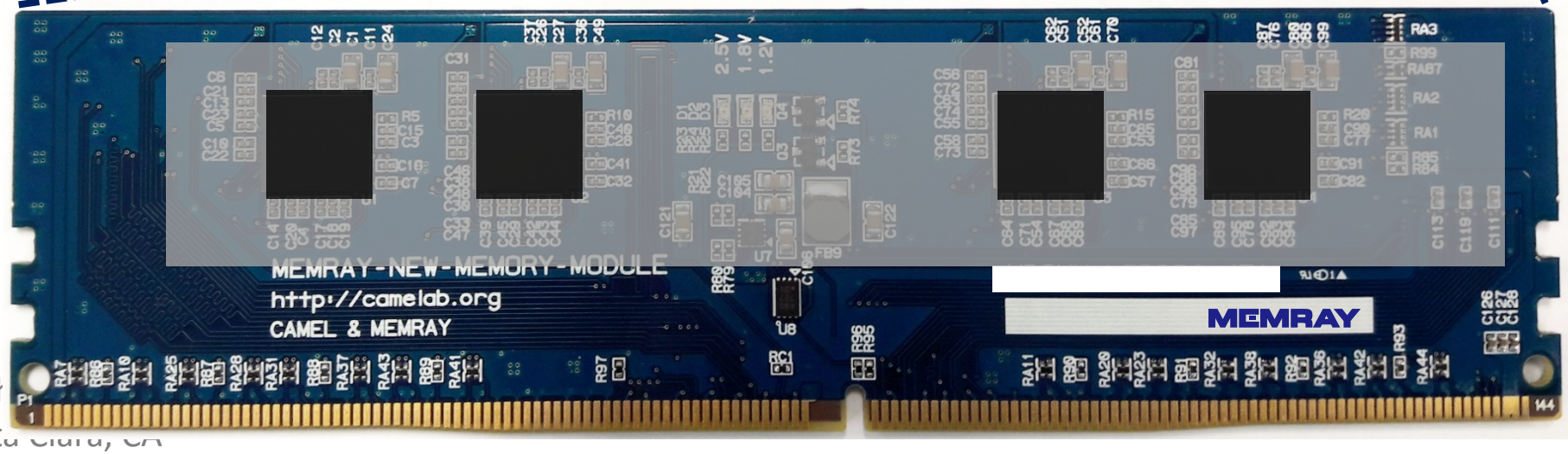
PRAM chip

PRAM chip

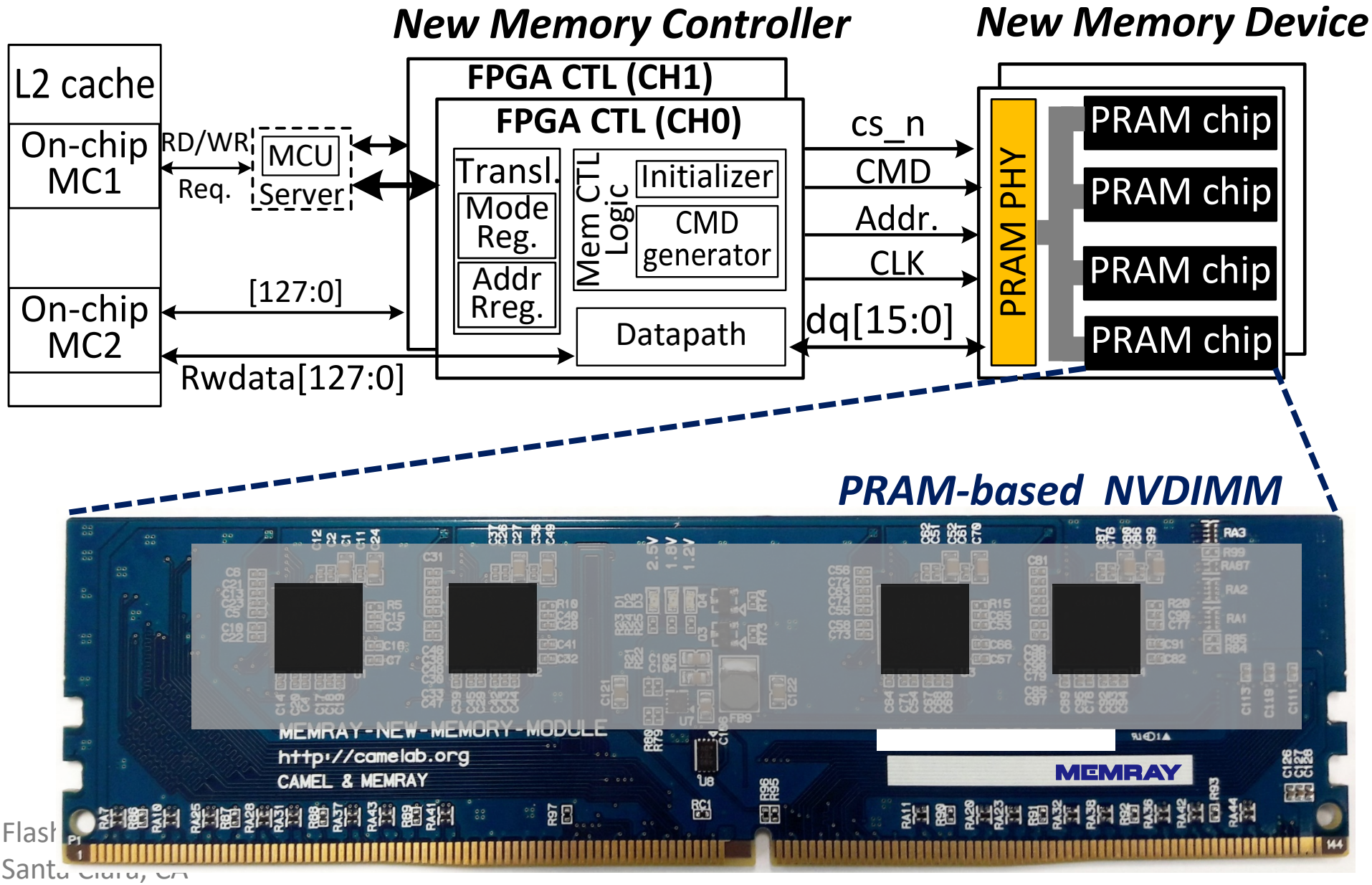
PRAM chip

PRAM chip

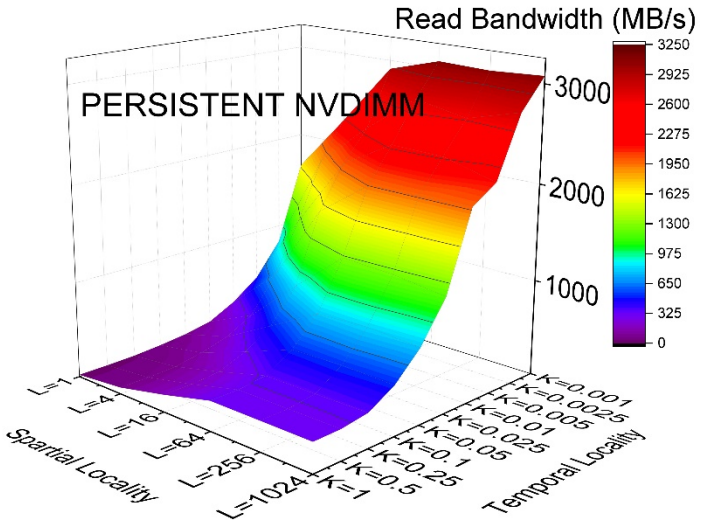
PRAM-based NVDIMM

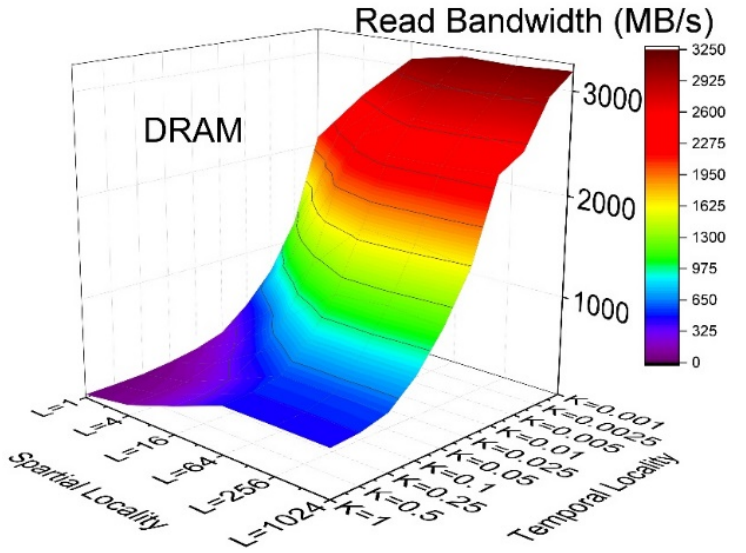


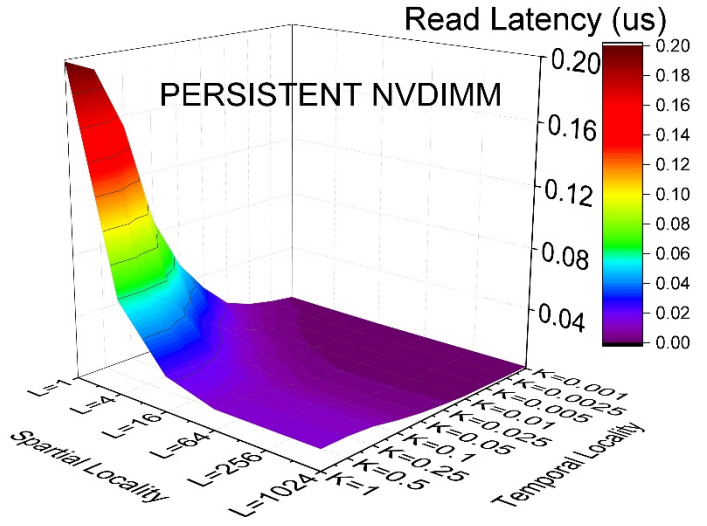
Further Enhancement

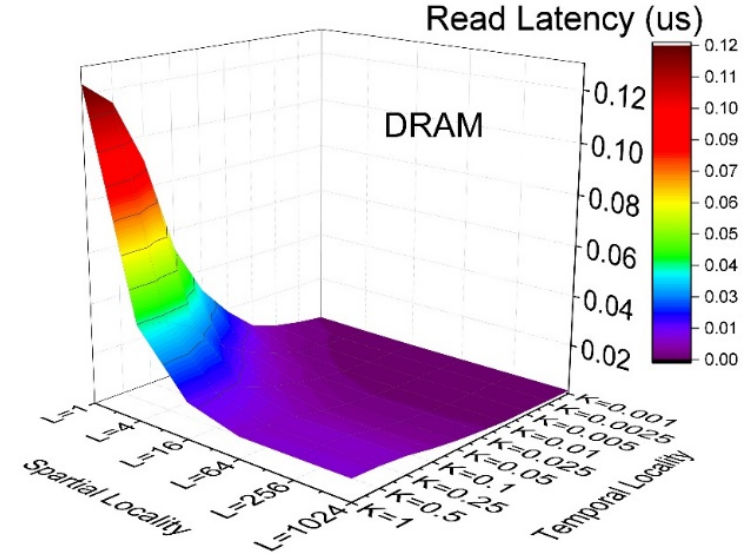


Prototype Results (APEX MAP for HPC)

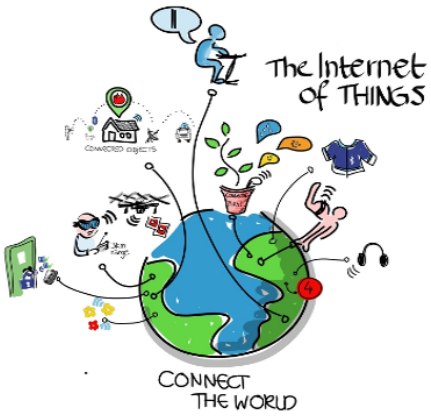


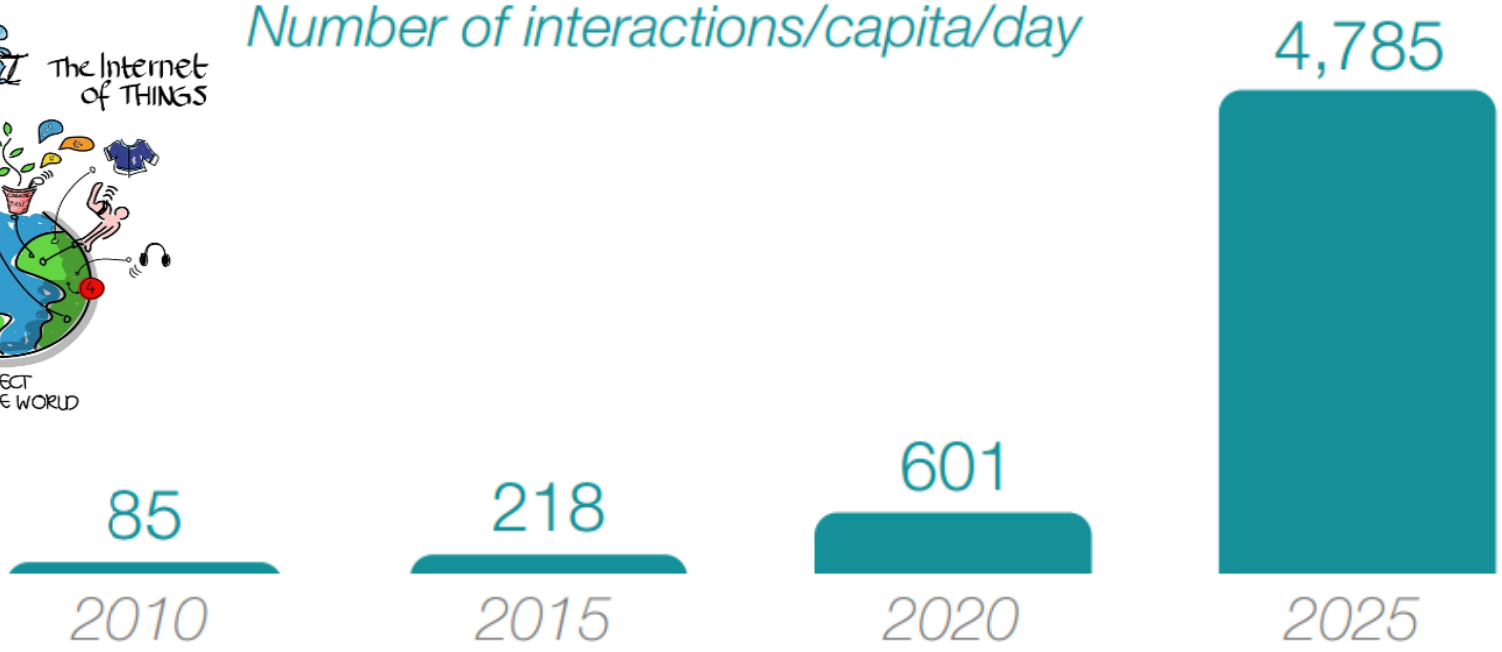


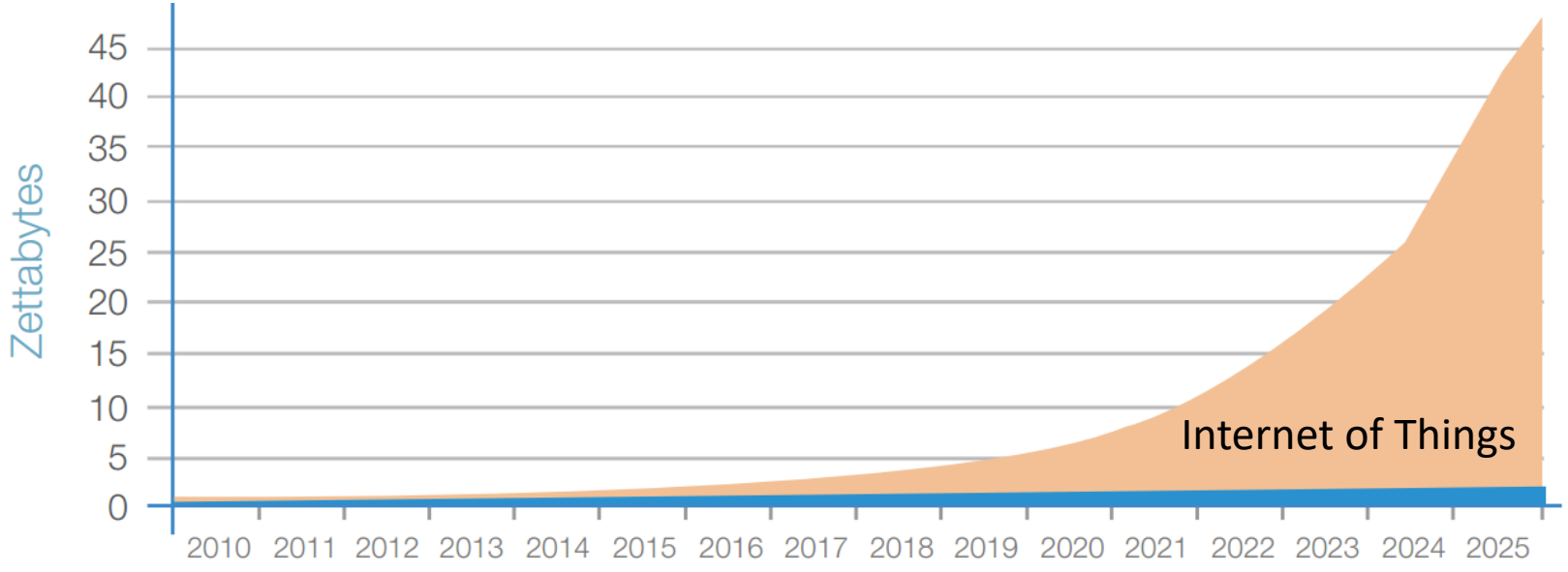




Real-Time Data



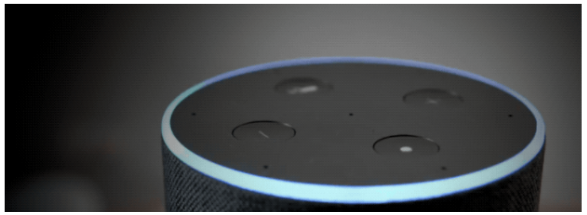




Amazon won't say if it hands your Echo data to the government

The retail, cloud, and device giant stands as the least transparent of transparent tech companies.

By Zack Whittaker for Zero Day | January 16, 2018 -- 21:36 GMT (05:36 GMT-08:00) | Topic: Cybersecurity in an IoT and Mobile World



Who owns connected car data?

28 Sep 2015
Matthew DeBord

Get ready for the biggest change to driving in the history of the automobile. Transportation and information are converging, at a rate much faster than anyone imagined possible even just a few years ago.

Every major automaker is grappling with the shift. General Motors has made a big bet on high-speed wireless connectivity throughout its vehicle fleet. Luxury carmakers such as BMW and Audi are rapidly enhancing the ability of the their cars to be as digitally enabled as smartphones. And Google and Apple are aggressively experimenting with both software and hardware, through Android Auto, self-driving cars, and Apple Car Play.

Here's just one example of how serious this business is for the auto industry. At the 2015 New York Auto Show, Ford CEO Mark Fields held a talk with a small

Amazon won't say if it hands your Echo data to the government

Amazon has a transpar

Three years ago, the re
reveal how many subpo
for customer data in a h
regularly published its g

Are Car Companies Going To Profit From Your Driving Data?

so-called advanced mobility community, which collectively has been anticipating

decades and you'll see a world in which cars collected data discretely, in

lace where data was freely flowing
huge shift from the days when cars
were in fact a rolling way to escape
tury: printed newspapers, daily mail

Who owns connected car data?

ears, as cars have become a crucial

Cars have gotten smart in a hurry. This has raised two major issues, one positive, one maybe not so positive.

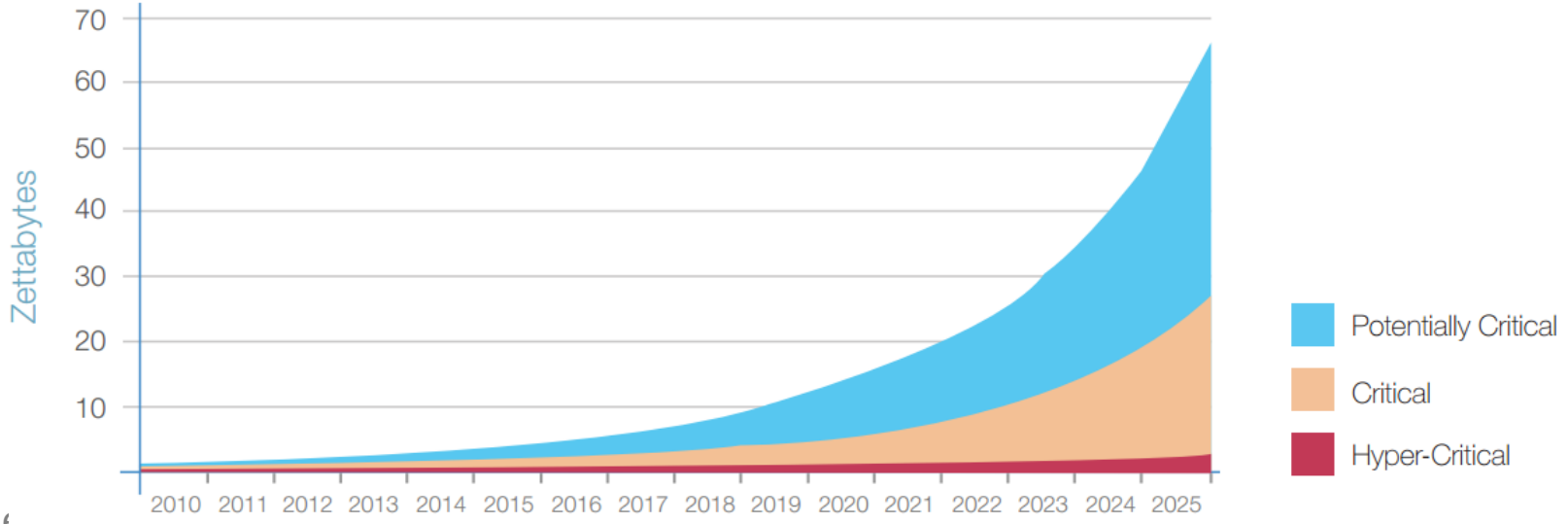
company's flagship Echo, an "always listening" speaker, collects vast amounts of customer data that's openly up for grabs by the government.

But Amazon's bi-annual transparency figures don't want you to know that.

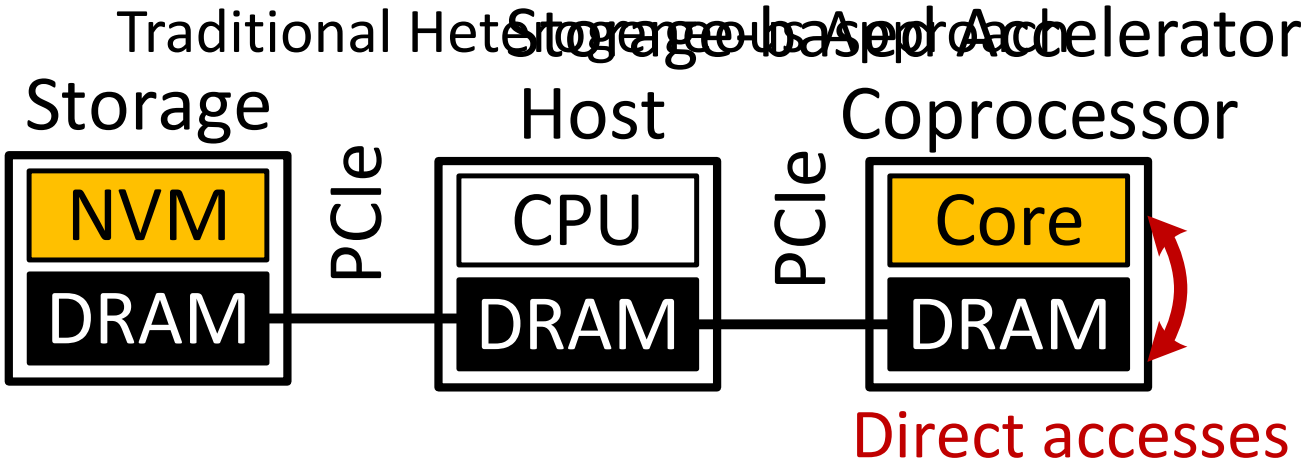
In fact, Amazon has been downright deceptive in how it presents the data

Endpoint

- Data should be tagged and classified by criticality
  - Security is matter – the private data should be managed by local or endpoint part
- Can we make each device as a standalone accelerator?
  - IoT leverages very power-limited devices
  - Even OS or file system can be a burden for IoT!



Storage-based Accelerator!

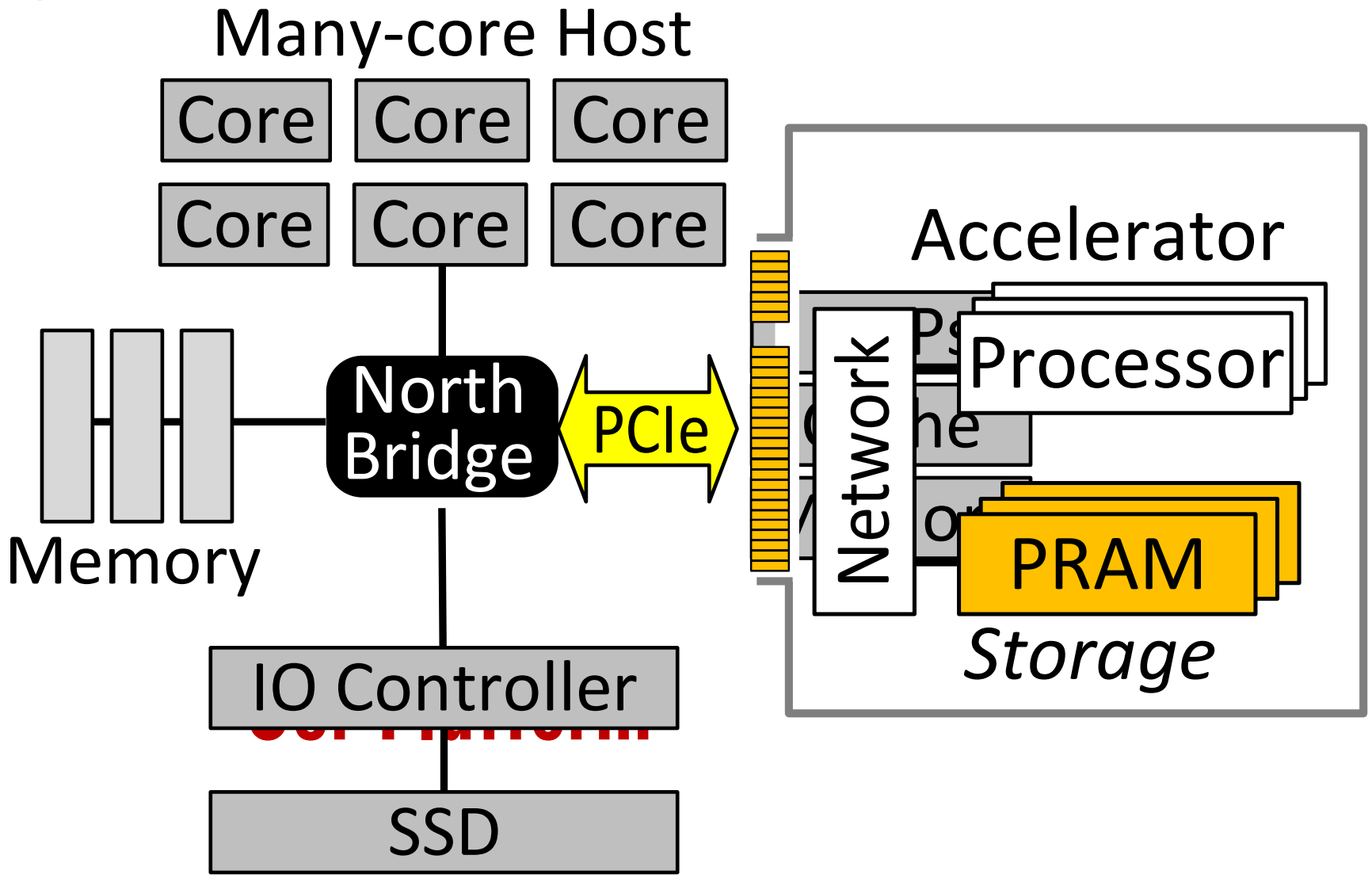


Integrate NVM chips into embedded coprocessor.

All cores are directly connected to storage.

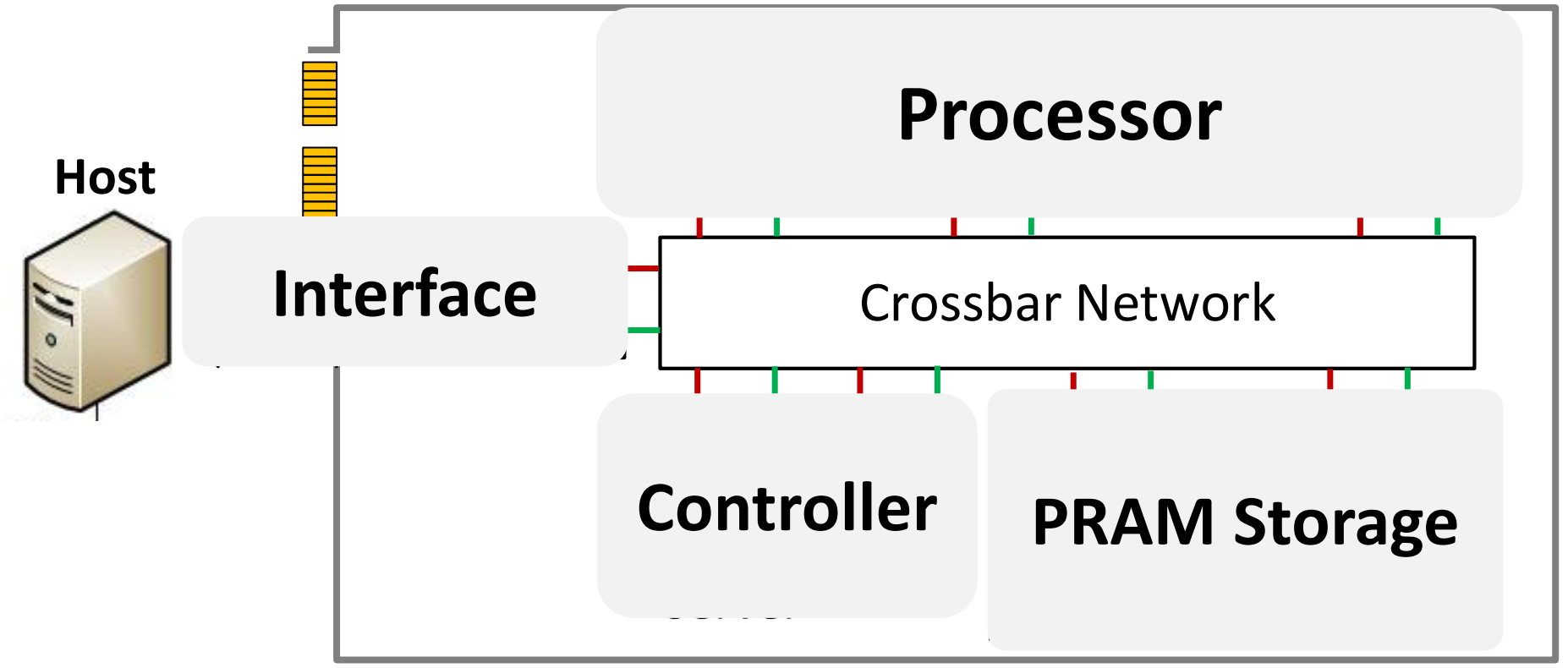
Process data within an accelerator.

Overall Architecture



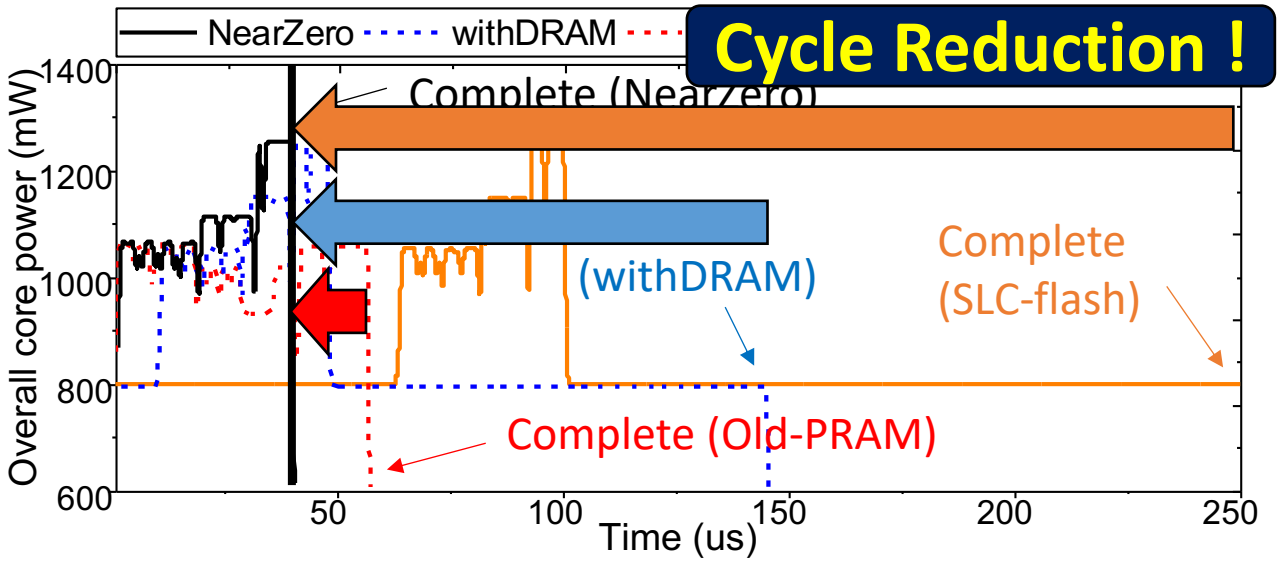
Heterogeneous Platform

NearZero architecture

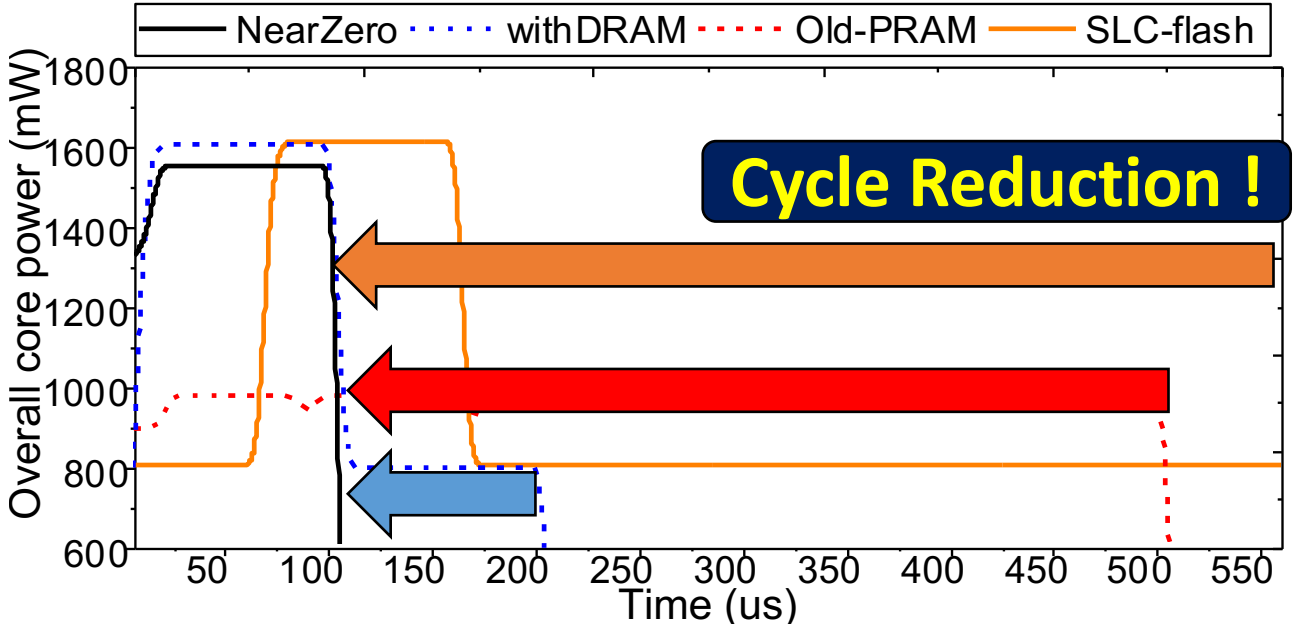


NearZero (Results)

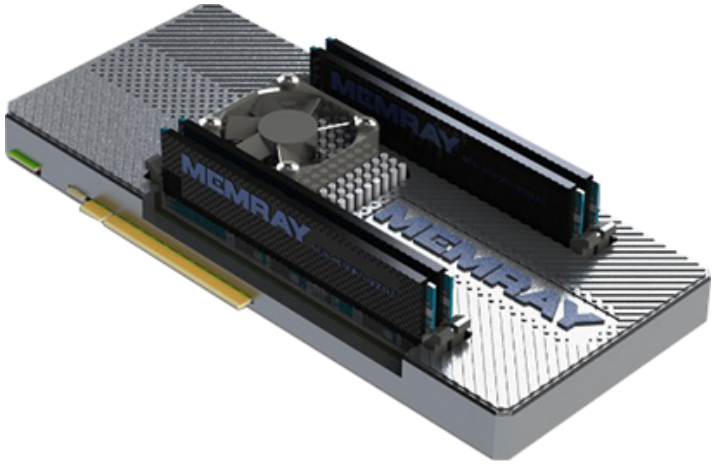
Read-intensive (processing power/energy)



Write-intensive (processing power/energy)



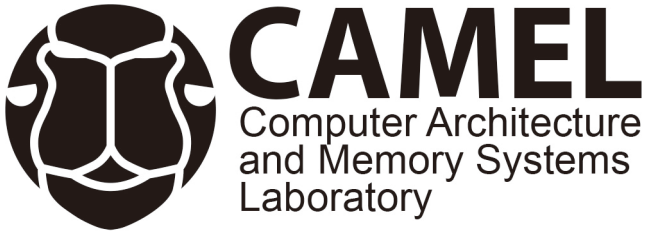
One More Direction w/ PRAM



- Full FPGA Automation for SCM-based Storage
- Performance
  Latency is sustainable for all random and sequential access patterns

Type	NVMe SSD (ASIC)	New Memory SSD (ASIC)	Our Prototype (FPGA)
Bandwidth (read/write)	1.1~2.3GB/sec	1.5~3.5GB/sec	2.5 ~ 5.2 GB/sec
Latency	15 ~ 150 us	8 us ~ 100 us	11~13 us

QnA


CAMEL
Computer Architecture
and Memory Systems
Laboratory