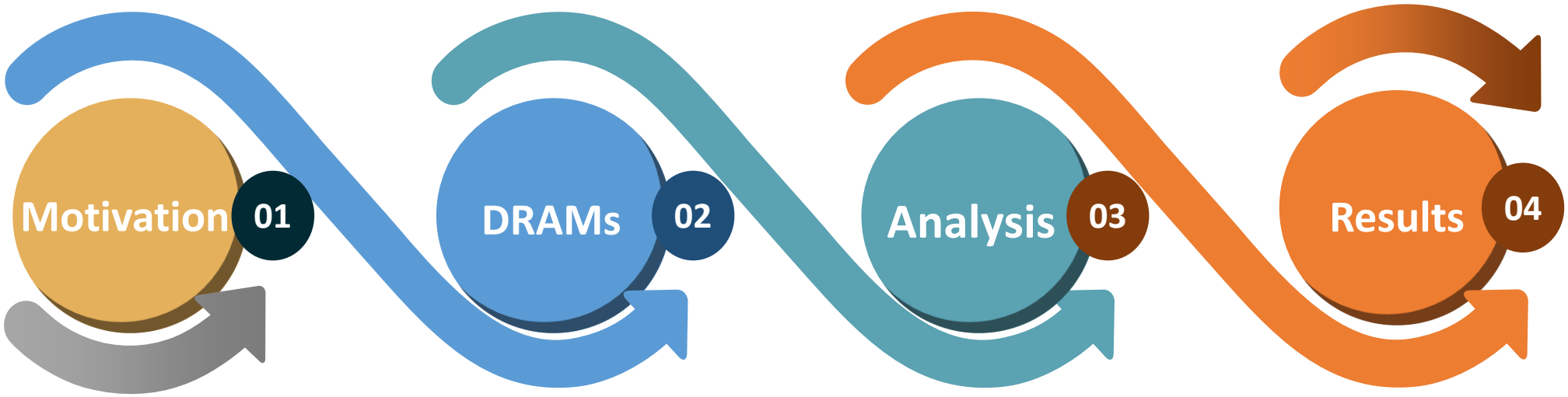
Managing DRAM Interference in Mixed Criticality Embedded Systems

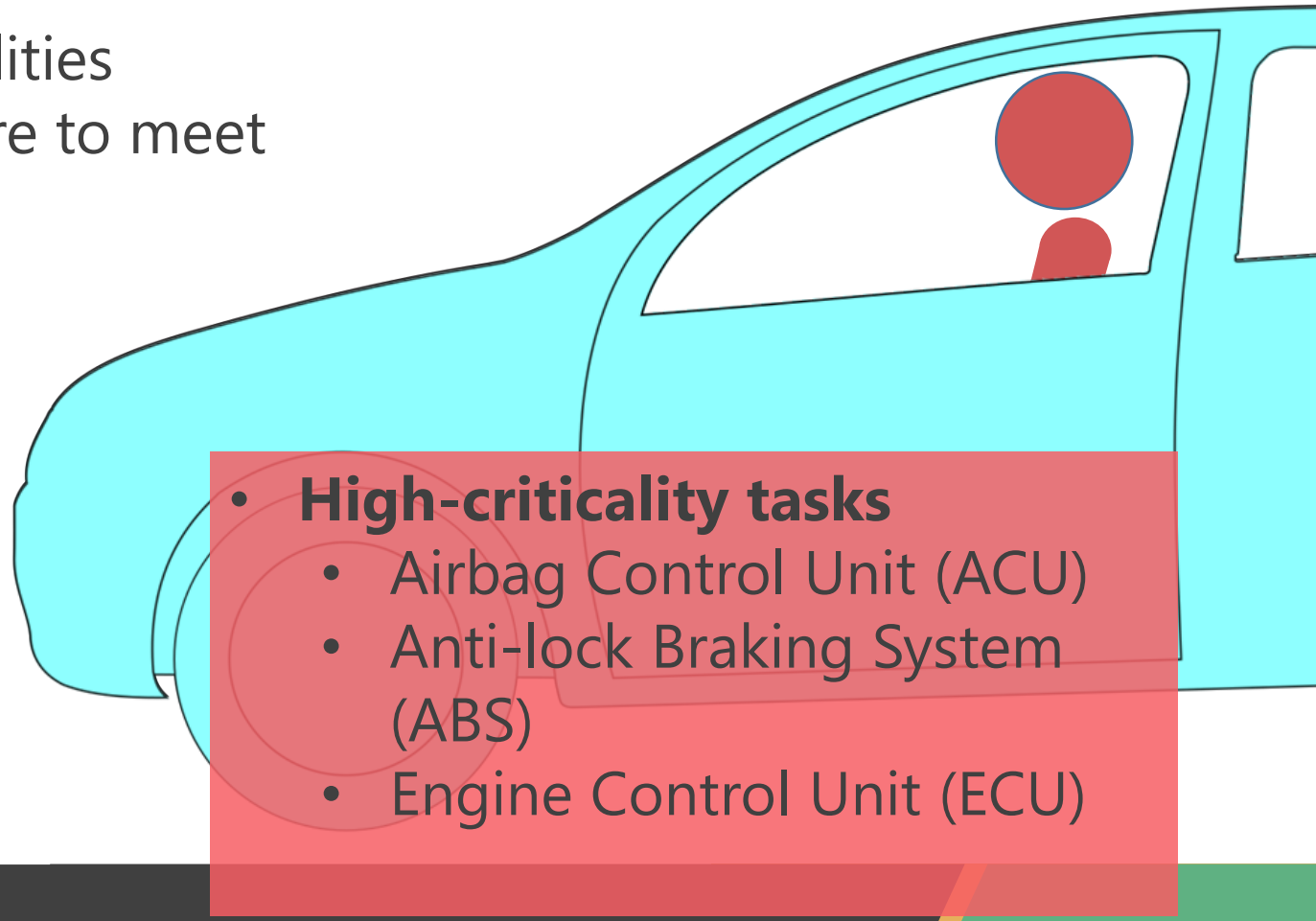
Mohamed Hassan

McMaster University



Outline

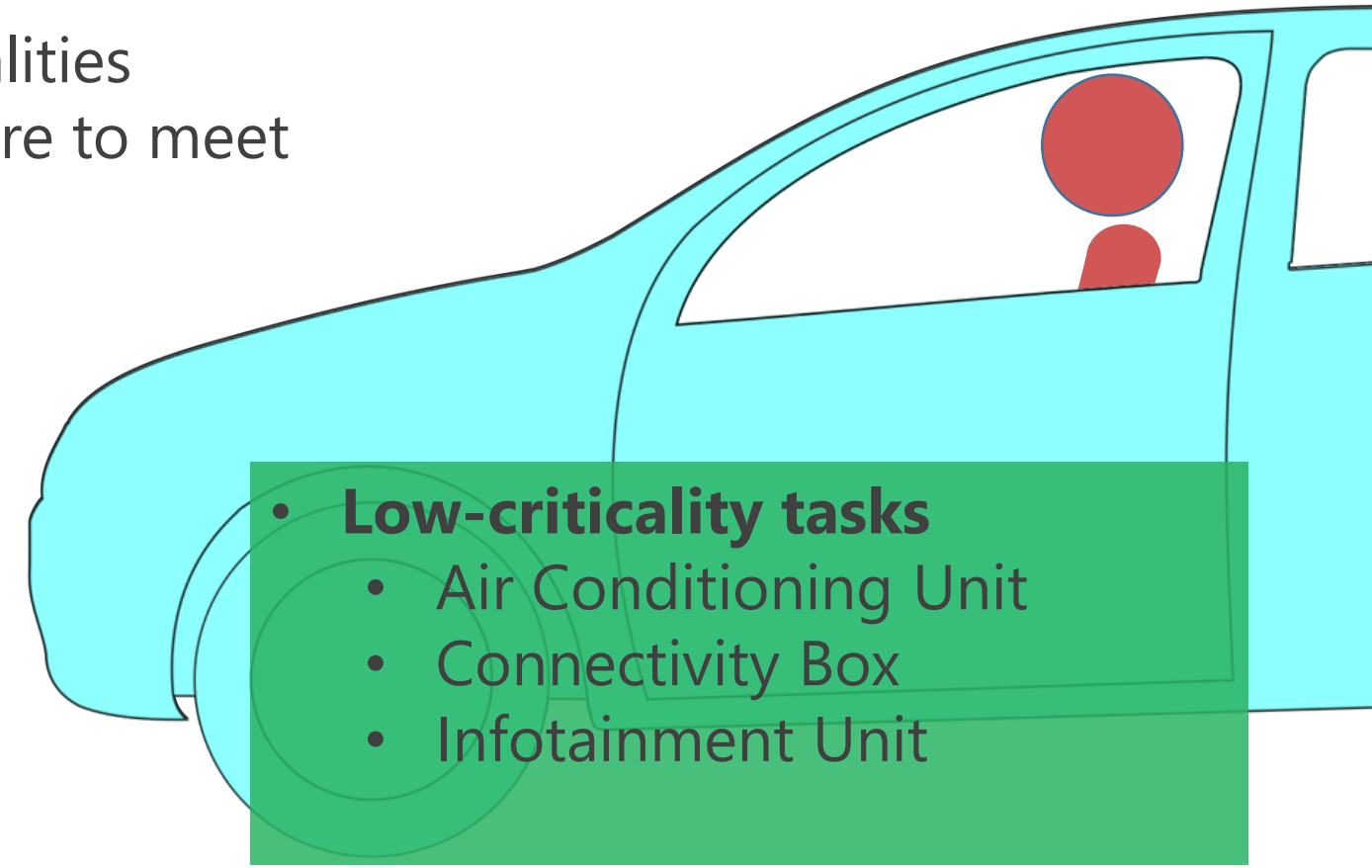
- **Emerging Systems No longer solely hosting isolated safety-critical tasks**
  - Execute tasks with different criticalities
  - Criticality α consequences of failure to meet requirements



- **High-criticality tasks**
  - Airbag Control Unit (ACU)
  - Anti-lock Braking System (ABS)
  - Engine Control Unit (ECU)

Mixed Criticality Systems

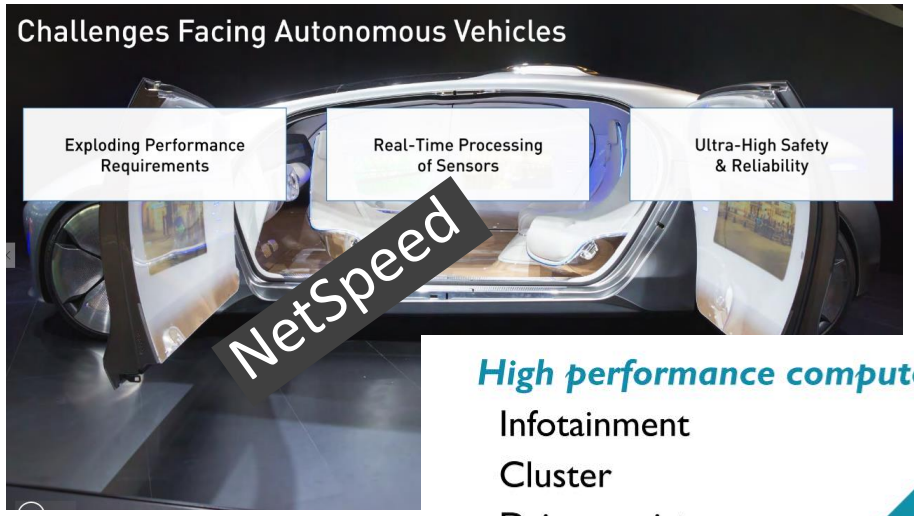
- **Emerging Systems No longer solely hosting isolated safety-critical tasks**
  - Execute tasks with different criticalities
  - Criticality α consequences of failure to meet requirements



- **Low-criticality tasks**
  - Air Conditioning Unit
  - Connectivity Box
  - Infotainment Unit

Mixed Criticality Systems

MOTIVATION



High performance compute

Infotainment
Cluster
Driver assist
Vehicle interface
User experience

Compute, Control, Sense

Real-time control

Safe
Secure
Responsive
Reliable
Fast boot

ARM

Cost Quality Ecosystem Temperature

ARM

Mixed Criticality Systems

MOTIVATION

Complementary SoC processor requirements

High performance compute

- Infotainment
- Cluster
- Driver assist
- Vehicle interface
- User experience

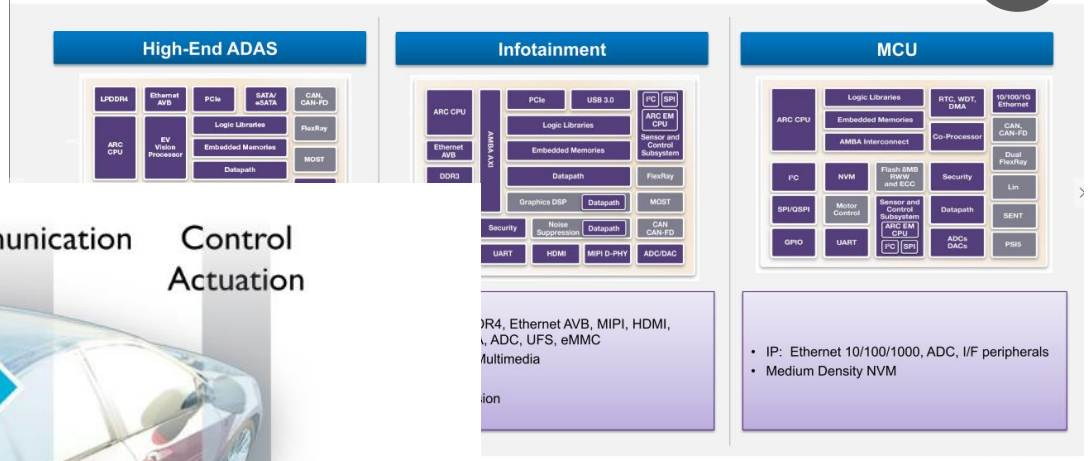
Compute, Control, Sense

Real-time control

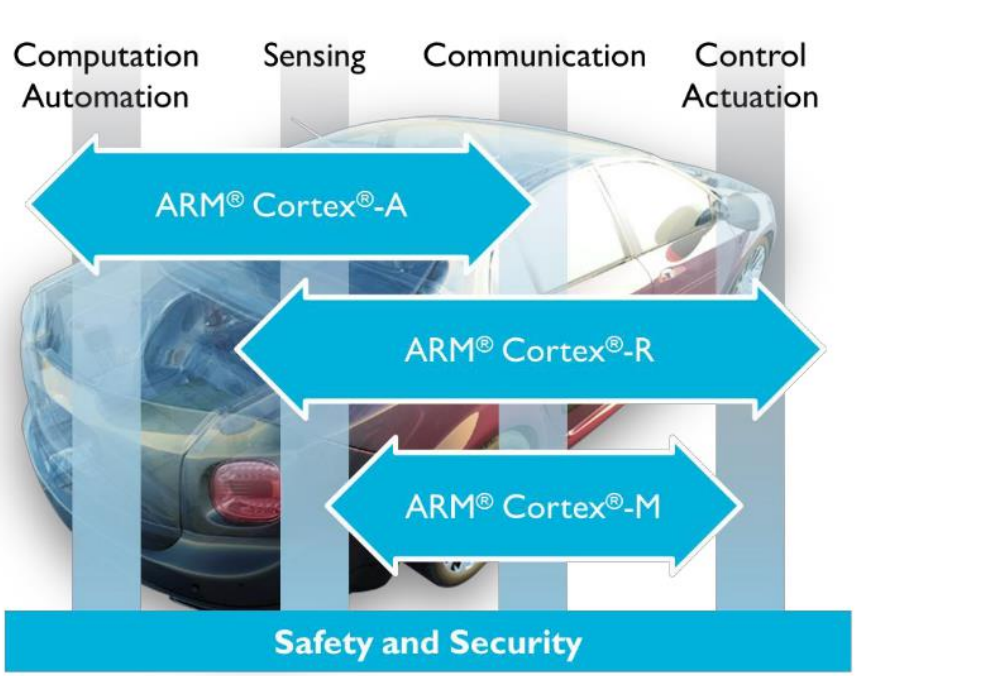
Cost **Quality** **Ecosystem**

Automotive Applications Require Different SoC Architectures

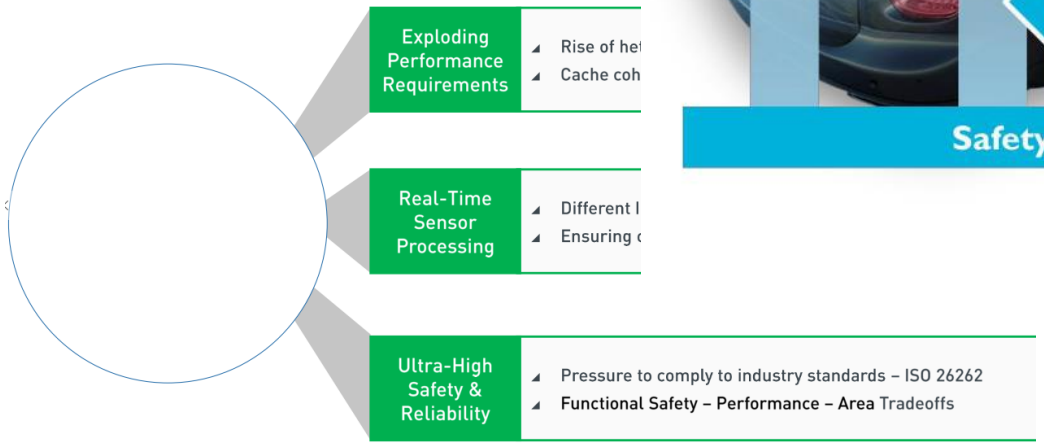
5



SYNOPSYS



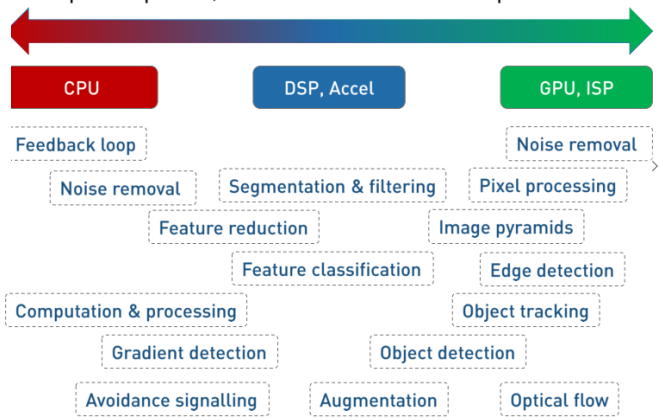
Translating System-Level Requirements



omputing

- Smaller amounts of data
- Highly structured data
- Complex computation/item

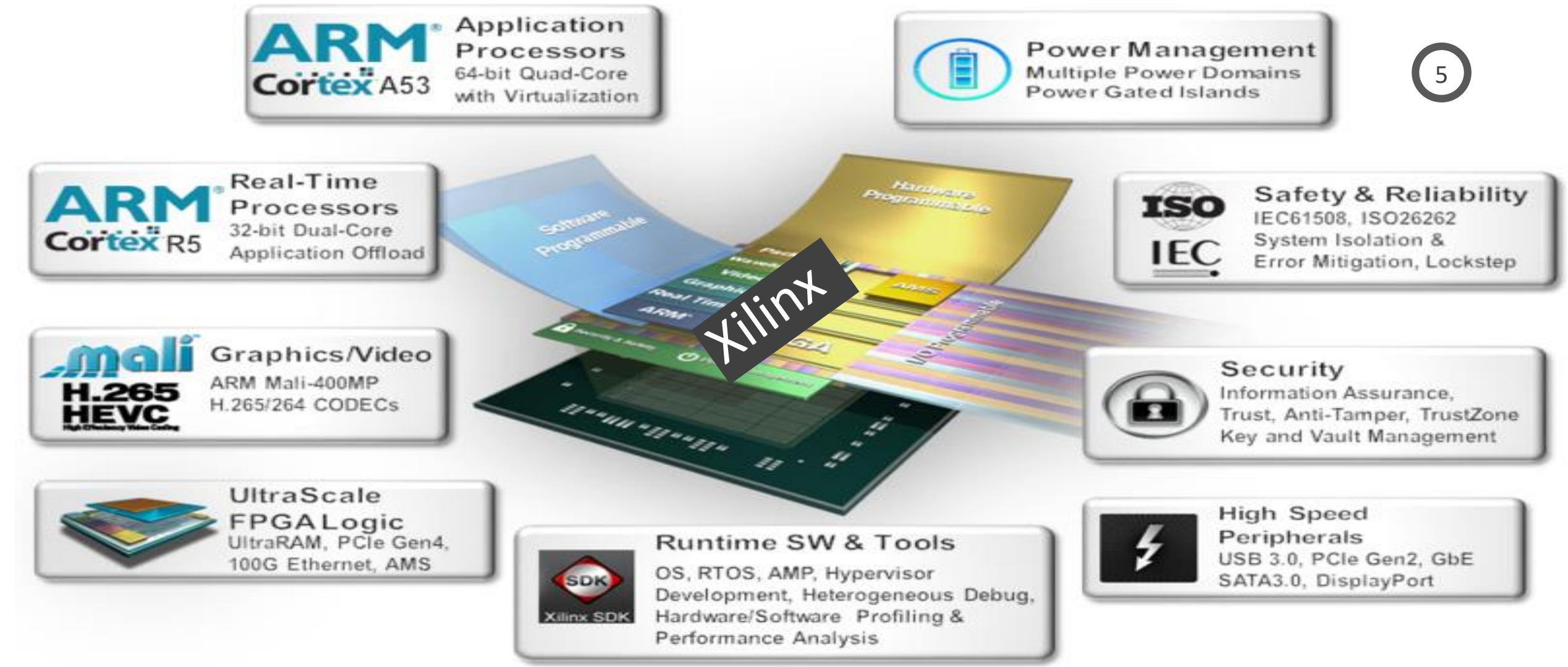
- Lots of data
- Simple computation/item
- Massive parallelism



ARM

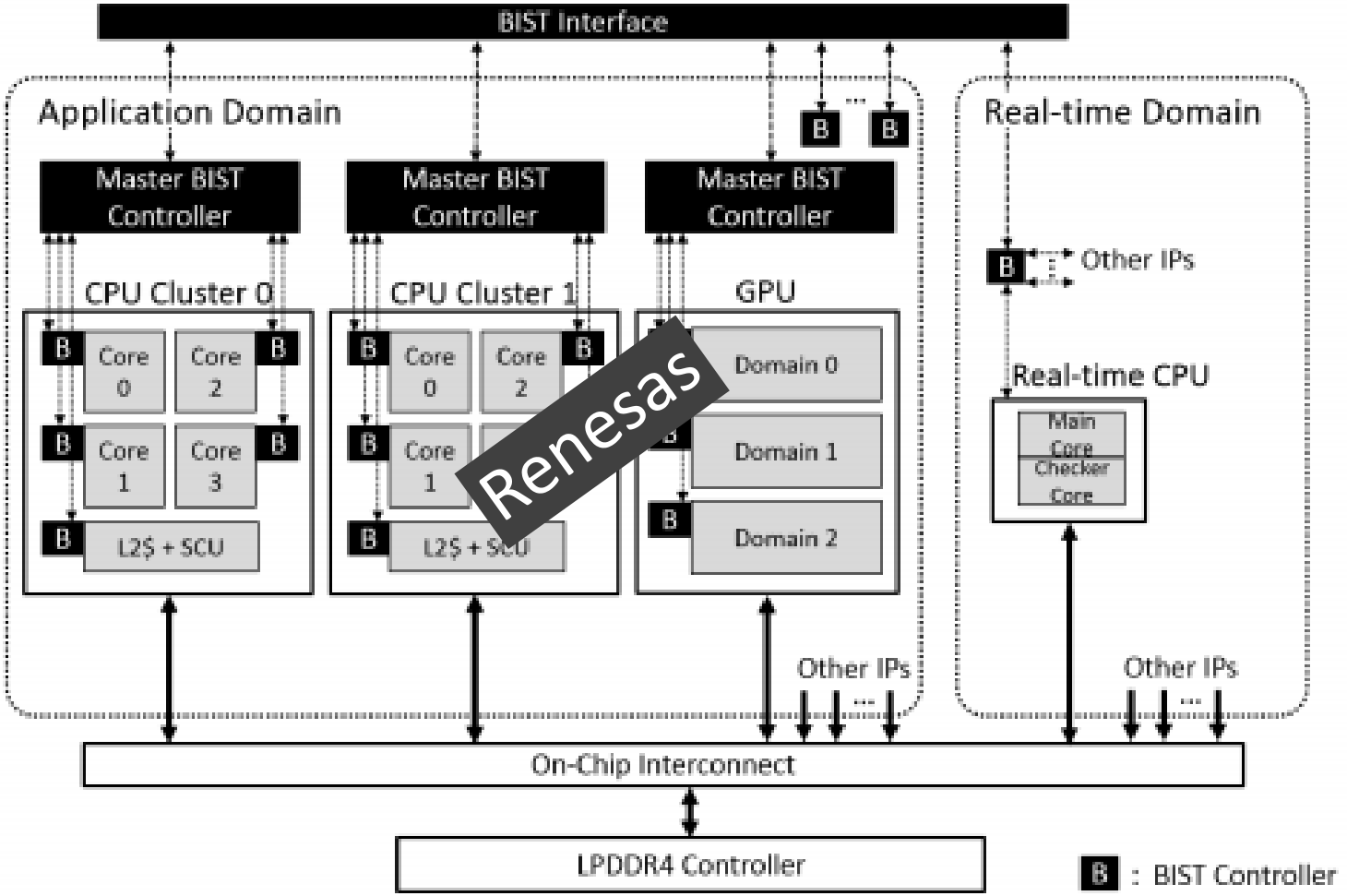
Pattern Recognition	Feature reduction
	Feature classification
	Augmentation
Feedback and Action	Computation & processing
	Feedback loop
	Avoidance signalling

Source: Extreme Tech, Google, ARM



Heterogenous MPSoCs with Real-time Processors

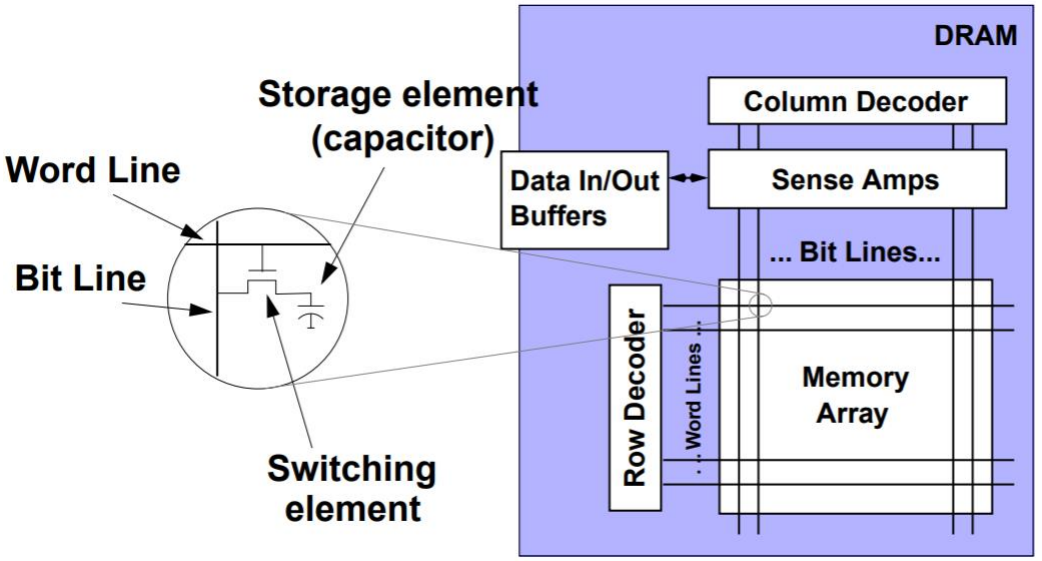
MOTIVATION



Heterogenous MPSoCs with Real-time Processors

MOTIVATION

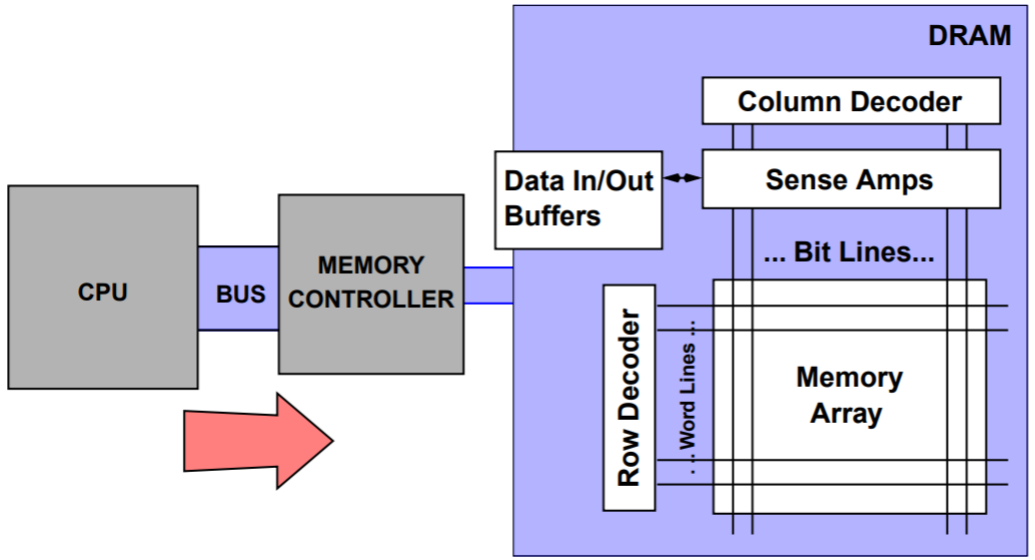
- DRAM Consists of multiple banks



Background

DRAM

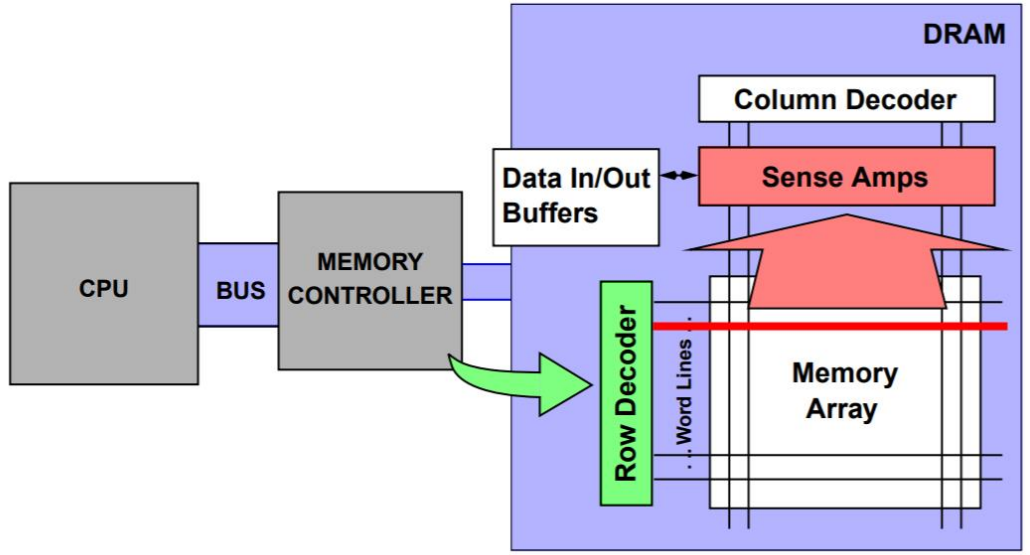
- DRAM Consists of multiple banks
- The memory controller (MC) manages accesses to DRAM



Background

DRAM

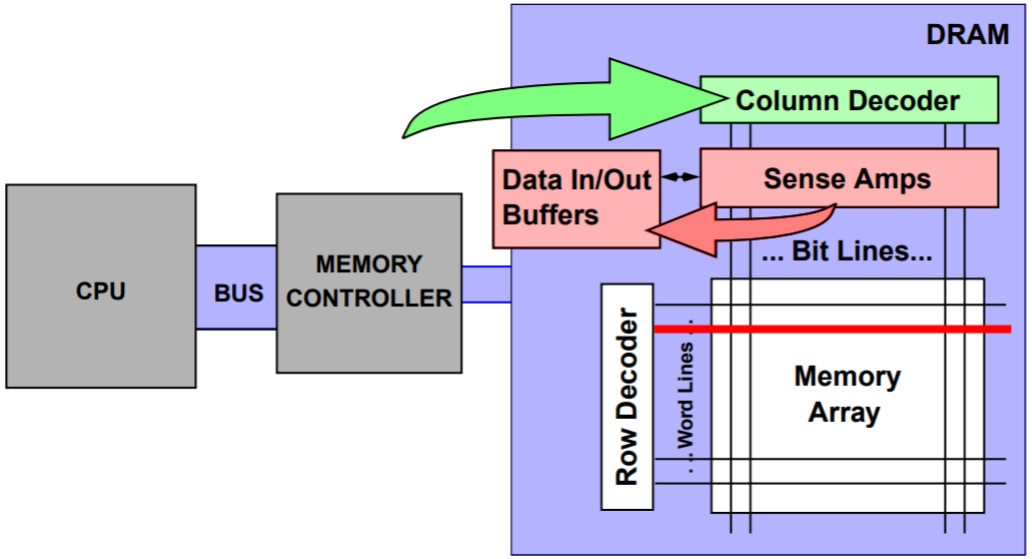
- DRAM Consists of multiple banks
- The memory controller (MC) manages accesses to DRAM
- A request in general consists of:
  - ACTIVATE command:
    - Bring data row from cells into sense amplifiers



Background

DRAM

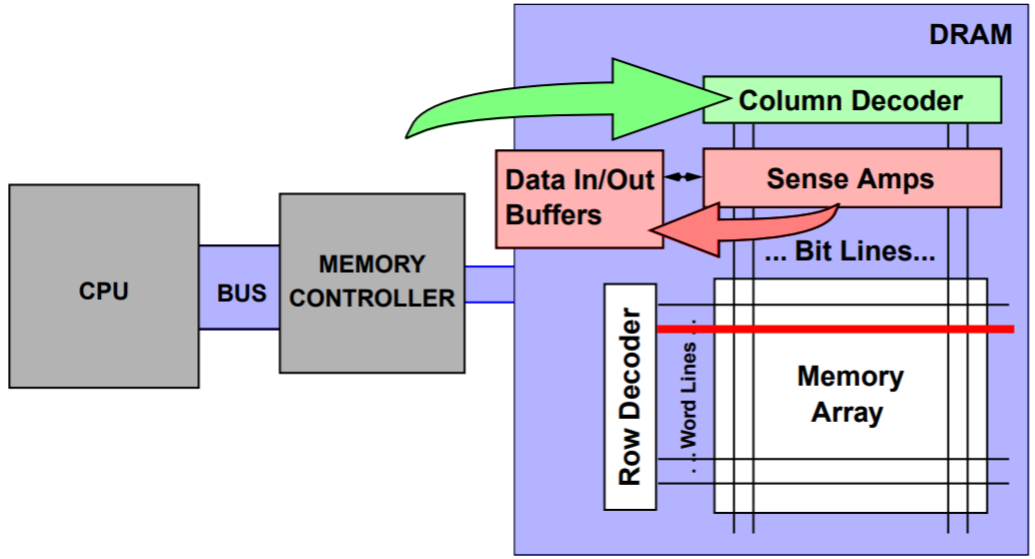
- DRAM Consists of multiple banks
- The memory controller (MC) manages accesses to DRAM
- A request in general consists of:
  - ACTIVATE command:
    - Bring data row from cells into sense amplifiers
  - RD/WR commands:
    - To read/write from specific columns in the sense amplifiers



Background

DRAM

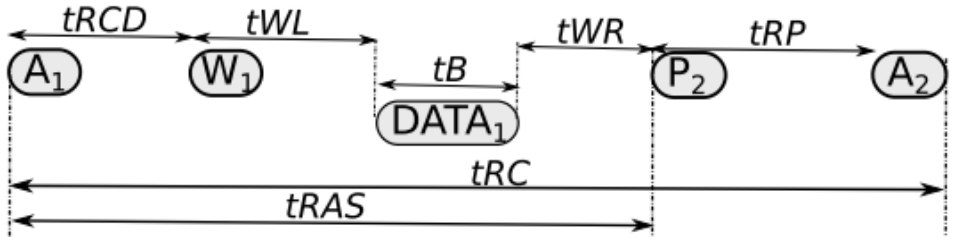
- DRAM Consists of multiple banks
- The memory controller (MC) manages accesses to DRAM
- A request in general consists of:
  - ACTIVATE command:
    - Bring data row from cells into sense amplifiers
  - RD/WR commands:
    - To read/write from specific columns in the sense amplifiers
  - PRECHARGE command:
    - to write back a previous row in the sense amplifiers before bringing the new one



Background

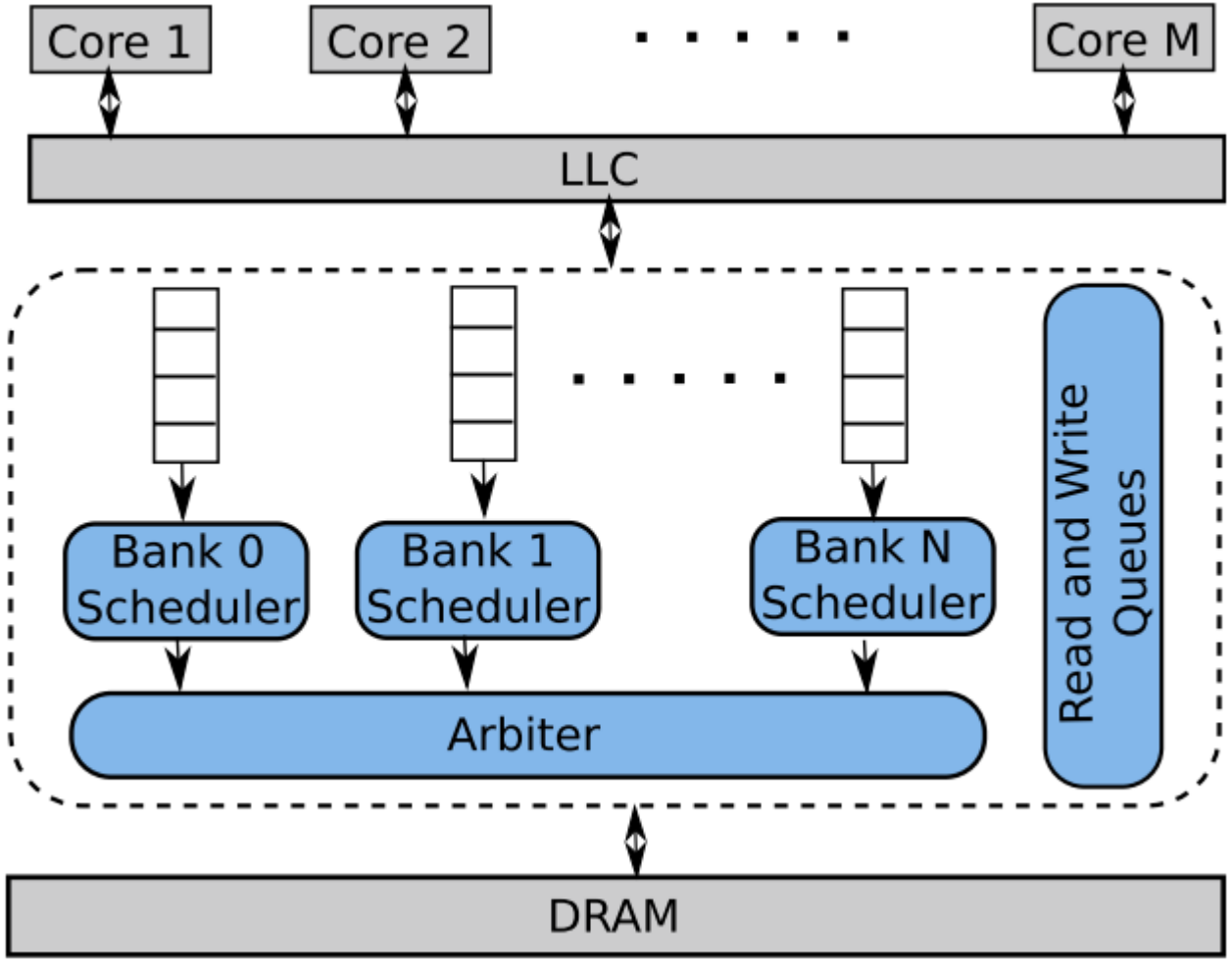
DRAM

- DRAM Consists of multiple banks
- The memory controller (MC) manages accesses to DRAM
- A request in general consists of:
  - ACTIVATE command:
    - Bring data row from cells into sense amplifiers
  - RD/WR commands:
    - To read/write from specific columns in the sense amplifiers
  - PRECHARGE command:
    - to write back a previous row in the sense amplifiers before bringing the new one
- All commands have associated timing constraints that have to be satisfied by the controller



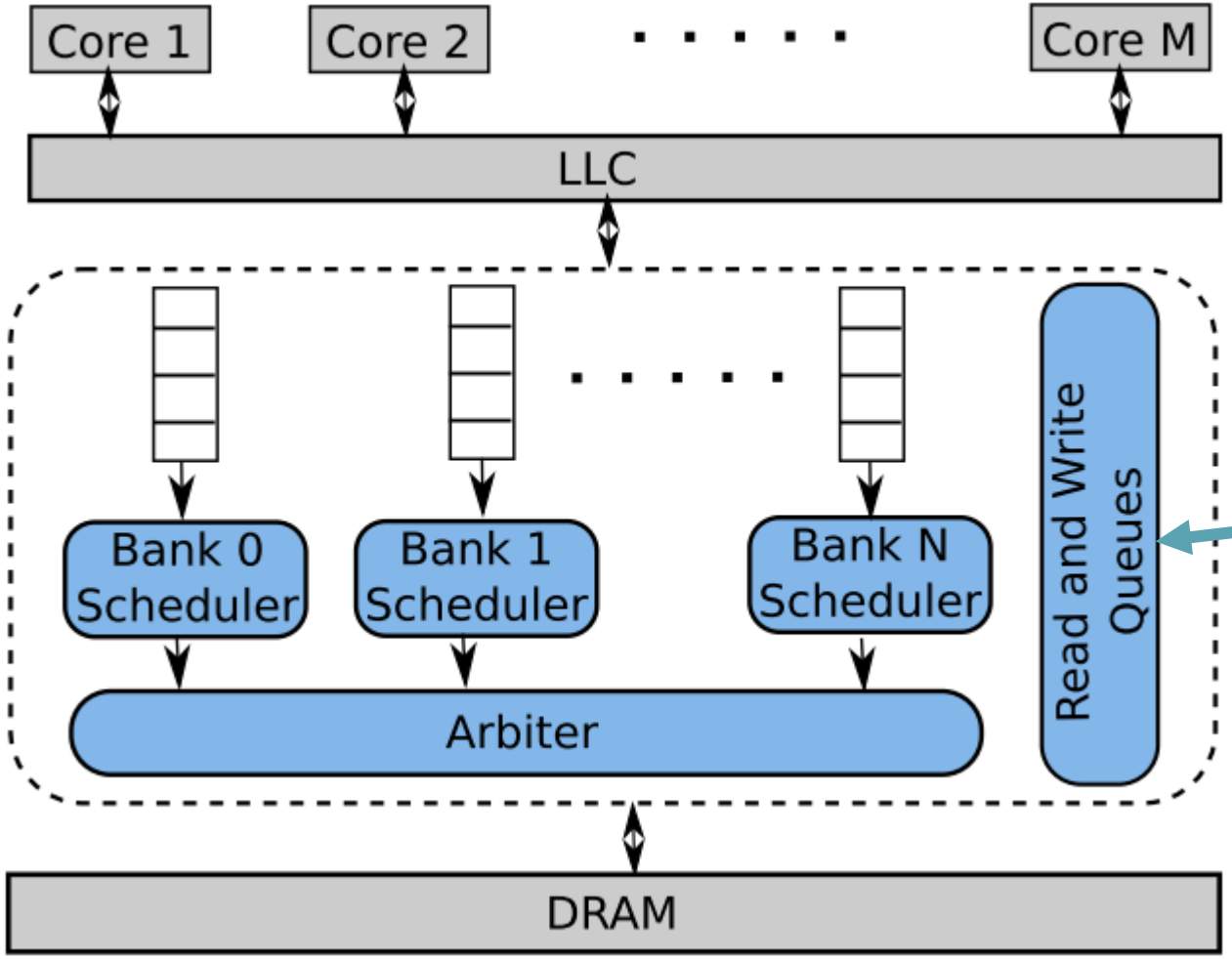
Background

DRAM



Background: Memory Controller

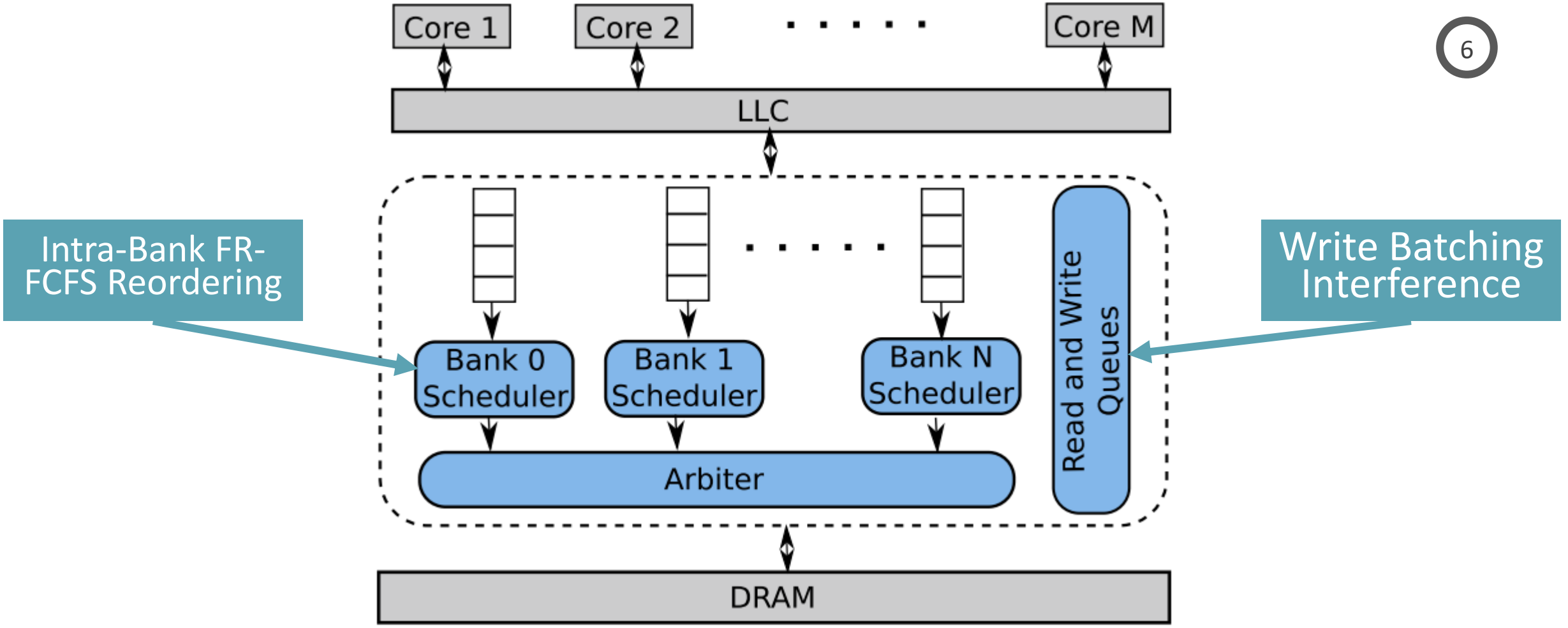
DRAM



Write Batching Interference

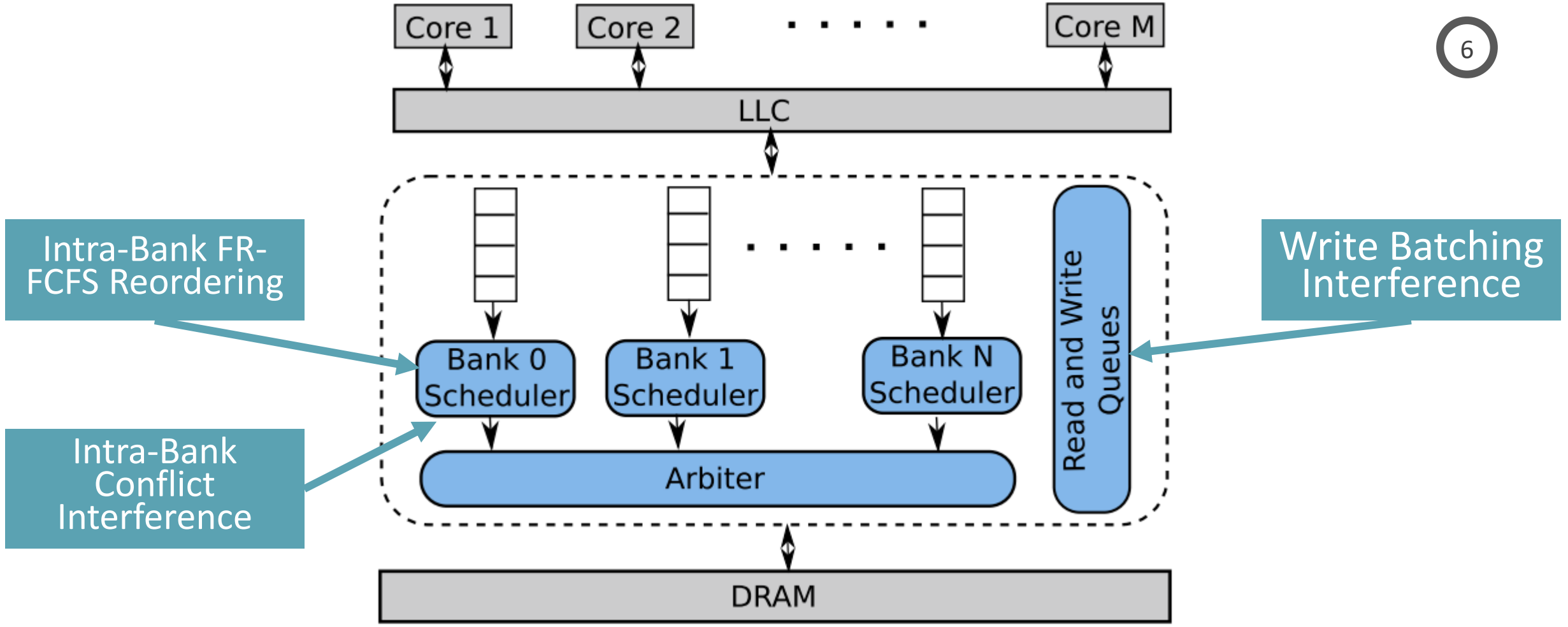
DRAM Interference

DRAM



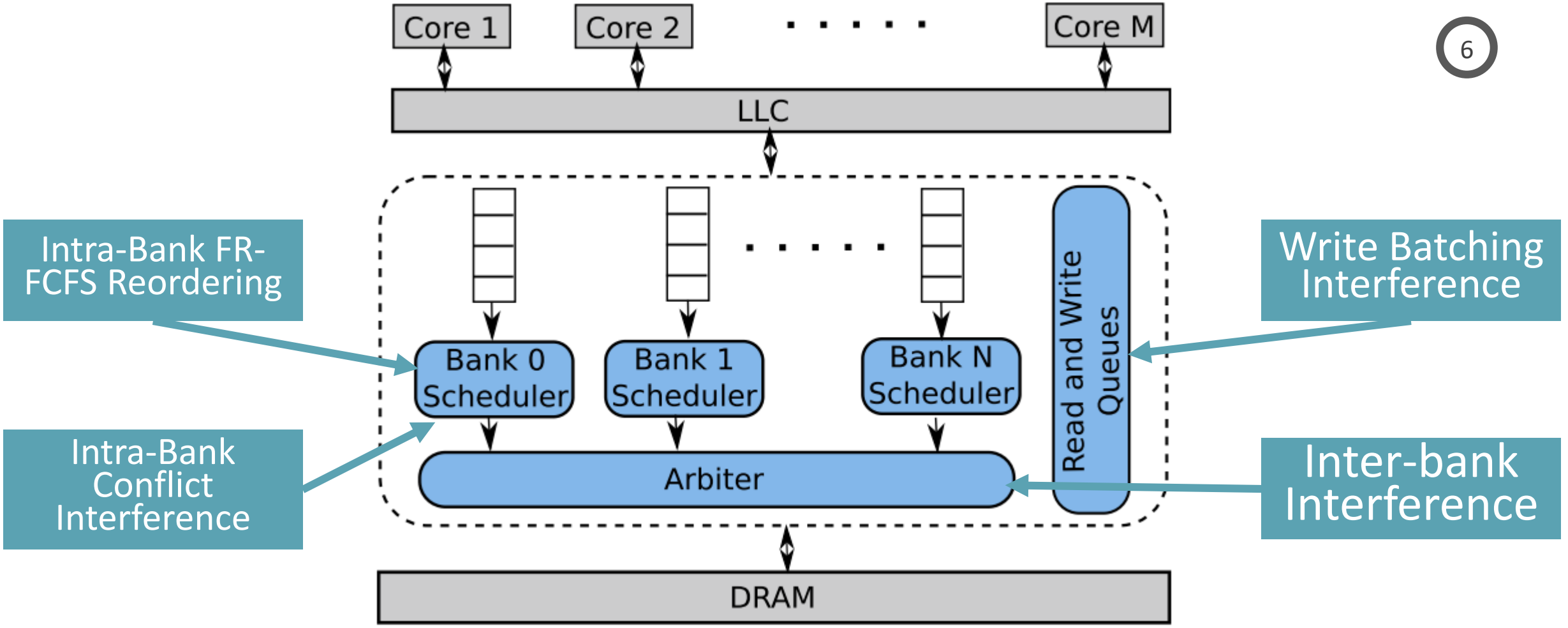
DRAM Interference

DRAM



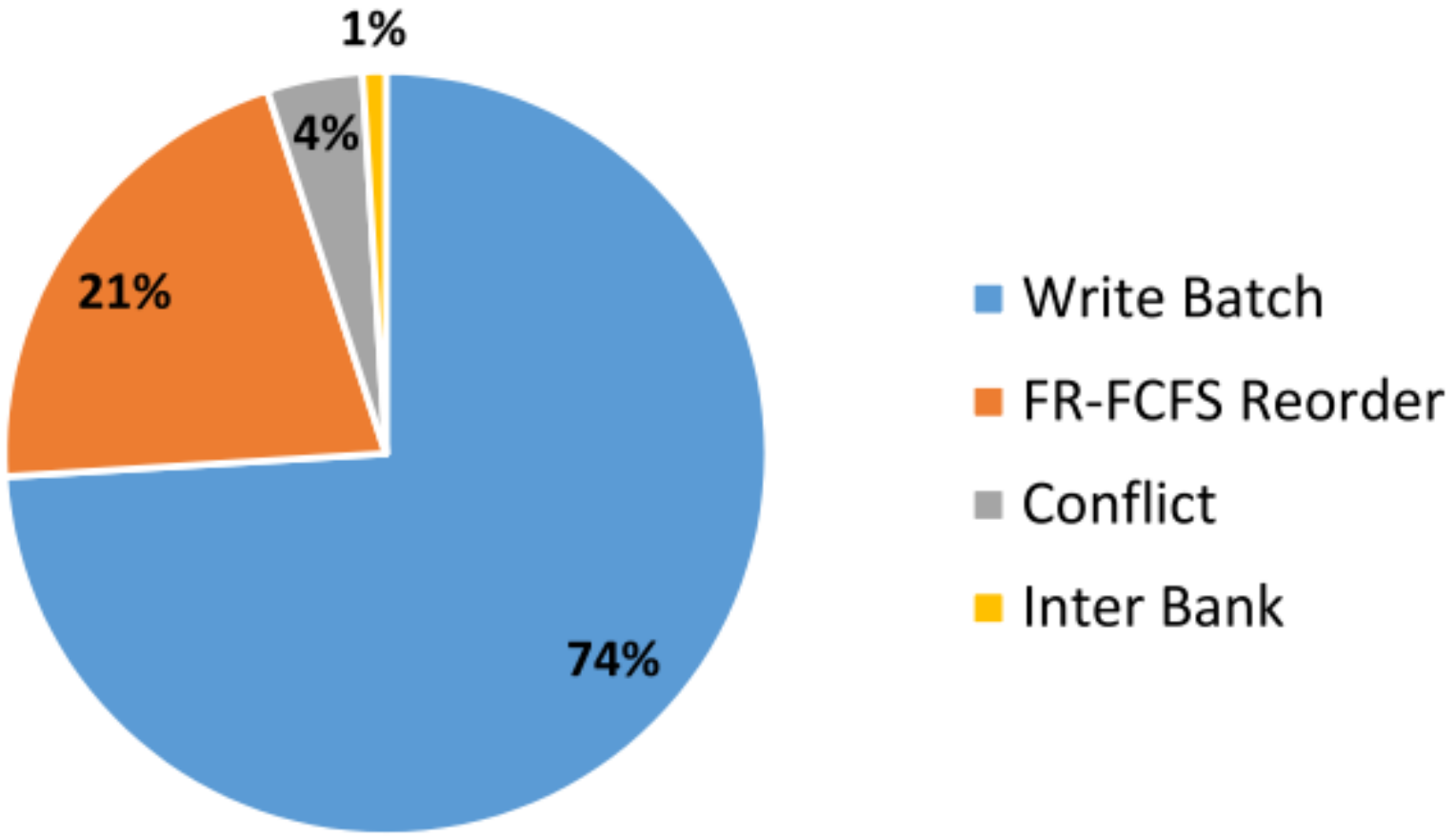
DRAM Interference

DRAM



DRAM Interference

DRAM



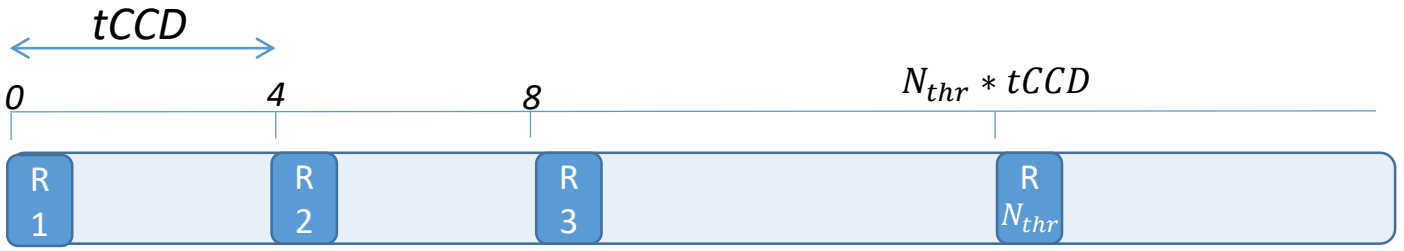
DRAM Delay Components

This paper:

- We study the effect of the largest two components in a mixed criticality system
- We derive both WCD bounds for critical cores and theoretical BW for non-critical cores

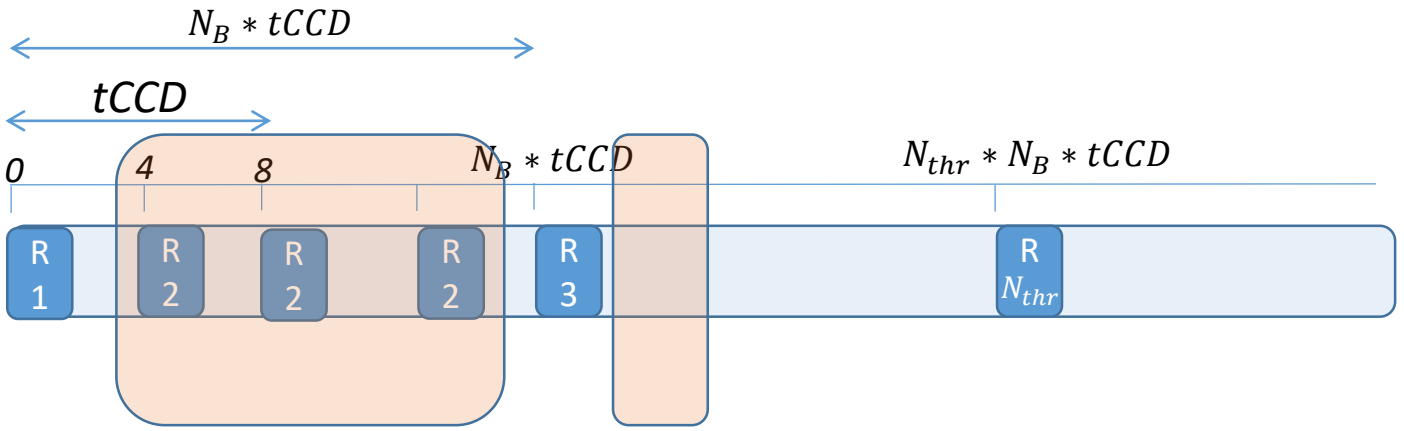
DRAM Delay Components

DRAM



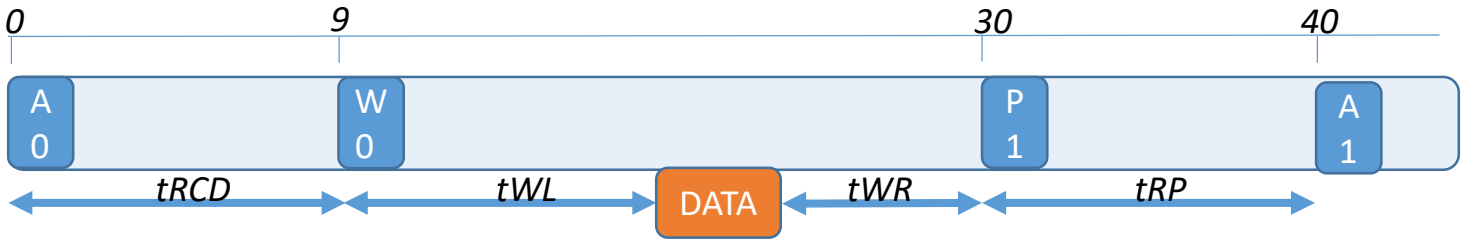
FR-FCFS WCD

DRAM



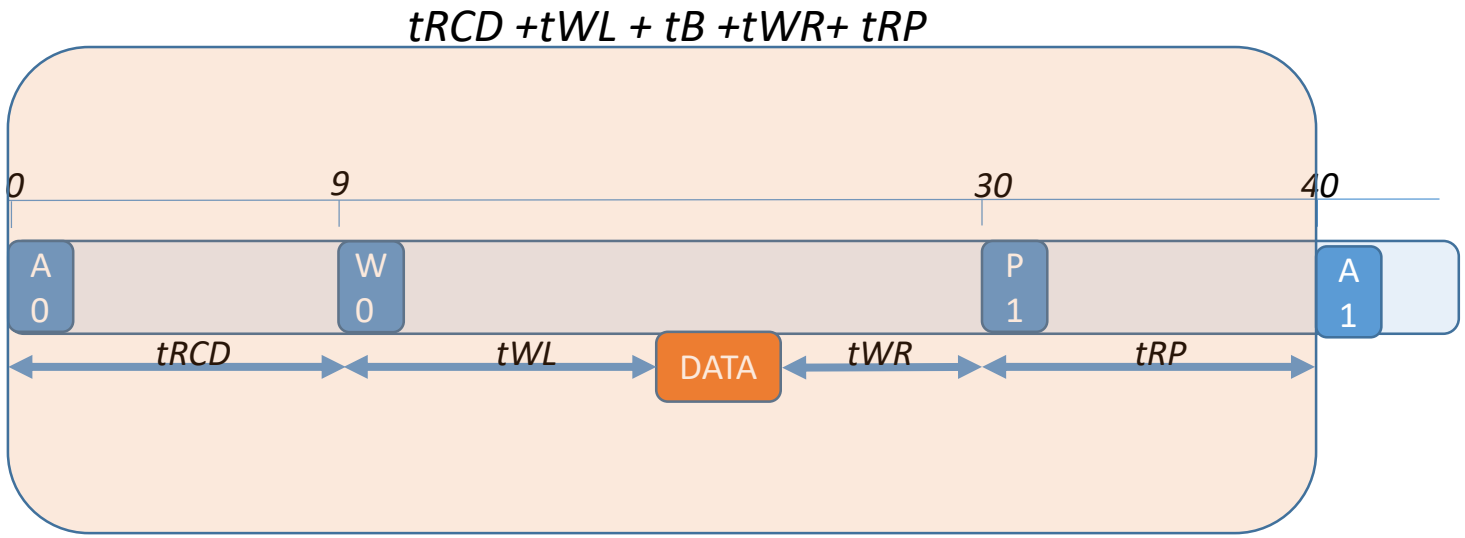
FR-FCFS WCD

DRAM



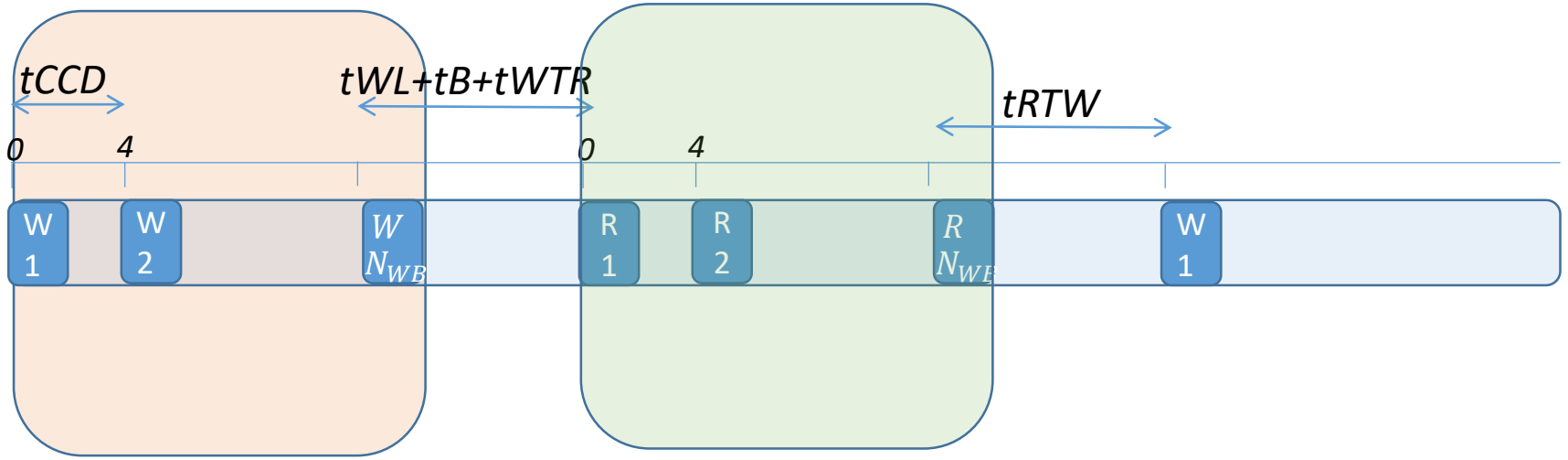
Write Batching WCD

DRAM



Write Batching WCD

DRAM



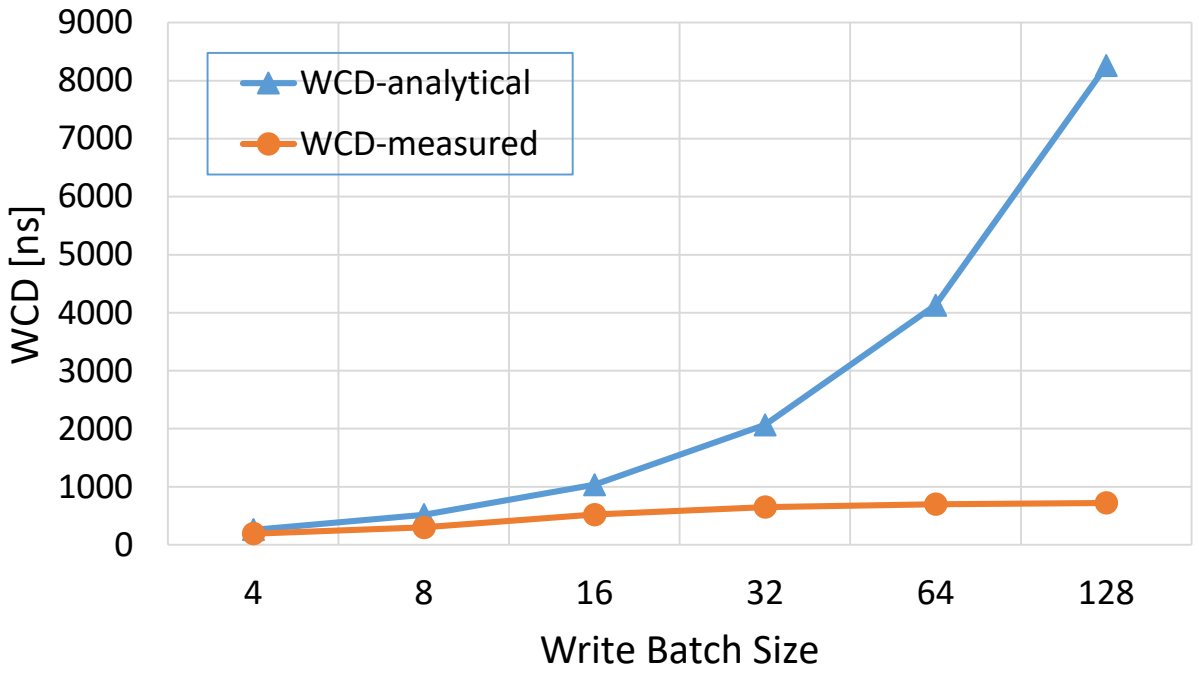
Background

DRAM

proc	2 Processors 1 OOO and 1 in-order pipeline a private 16KB L1 a shared 1MB L2 cache
DRAM	DDR3-1600
Benchmarks	Latency and BW benchmarks

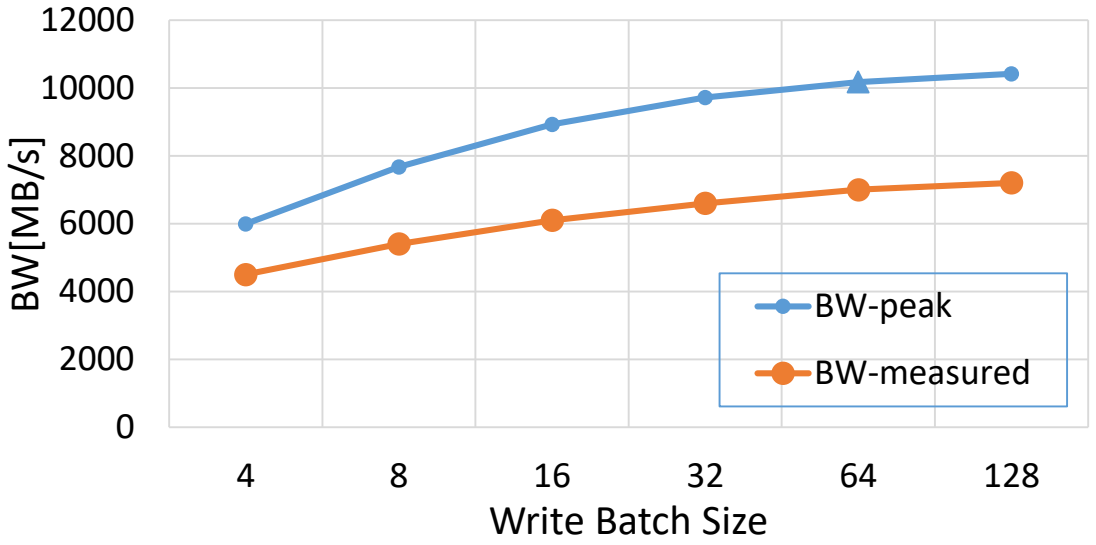
Evaluation Setup

RESULTS



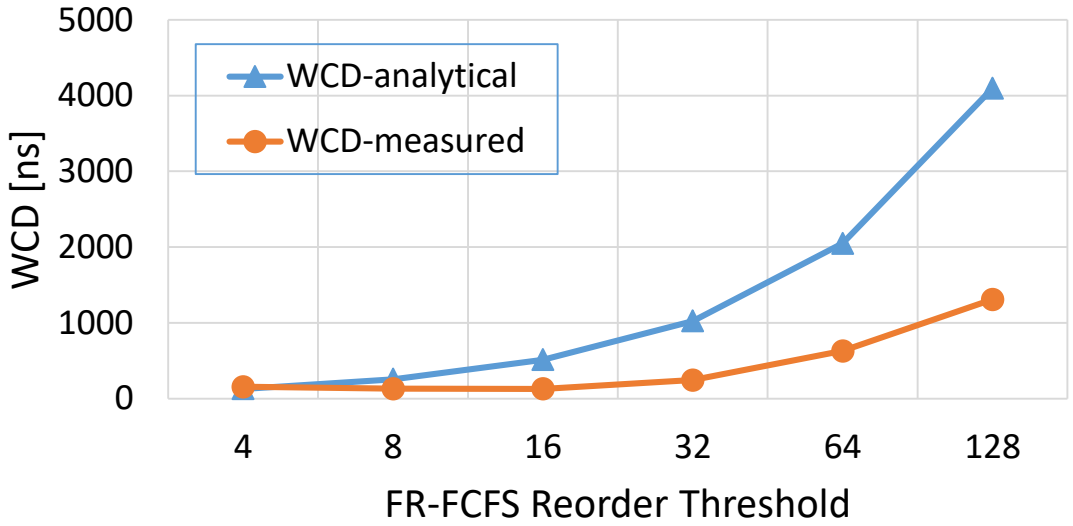
Write Batching WCD

RESULTS



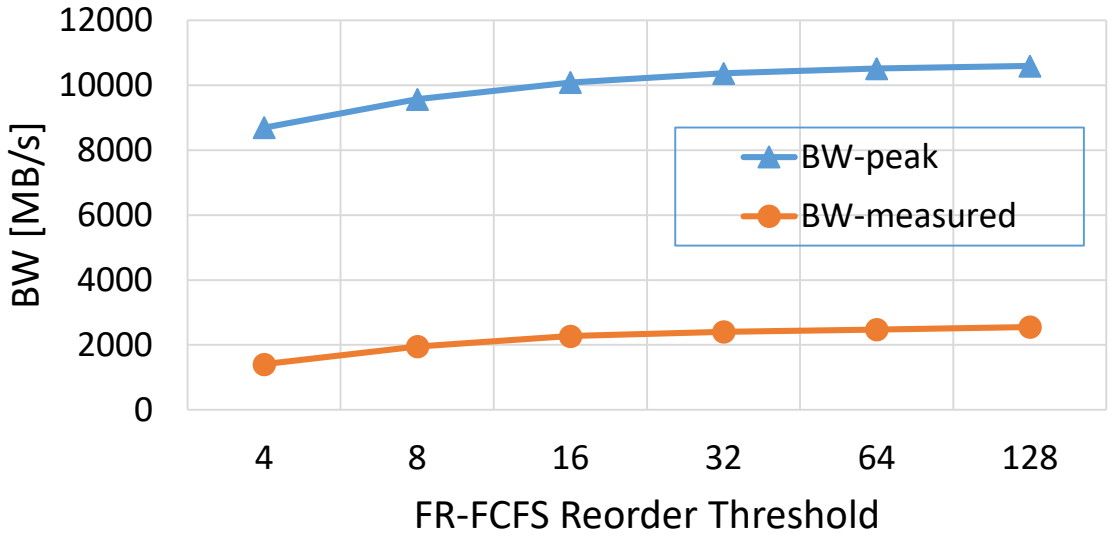
Write Batching BW

RESULTS



FR-FCFS WCD

RESULTS

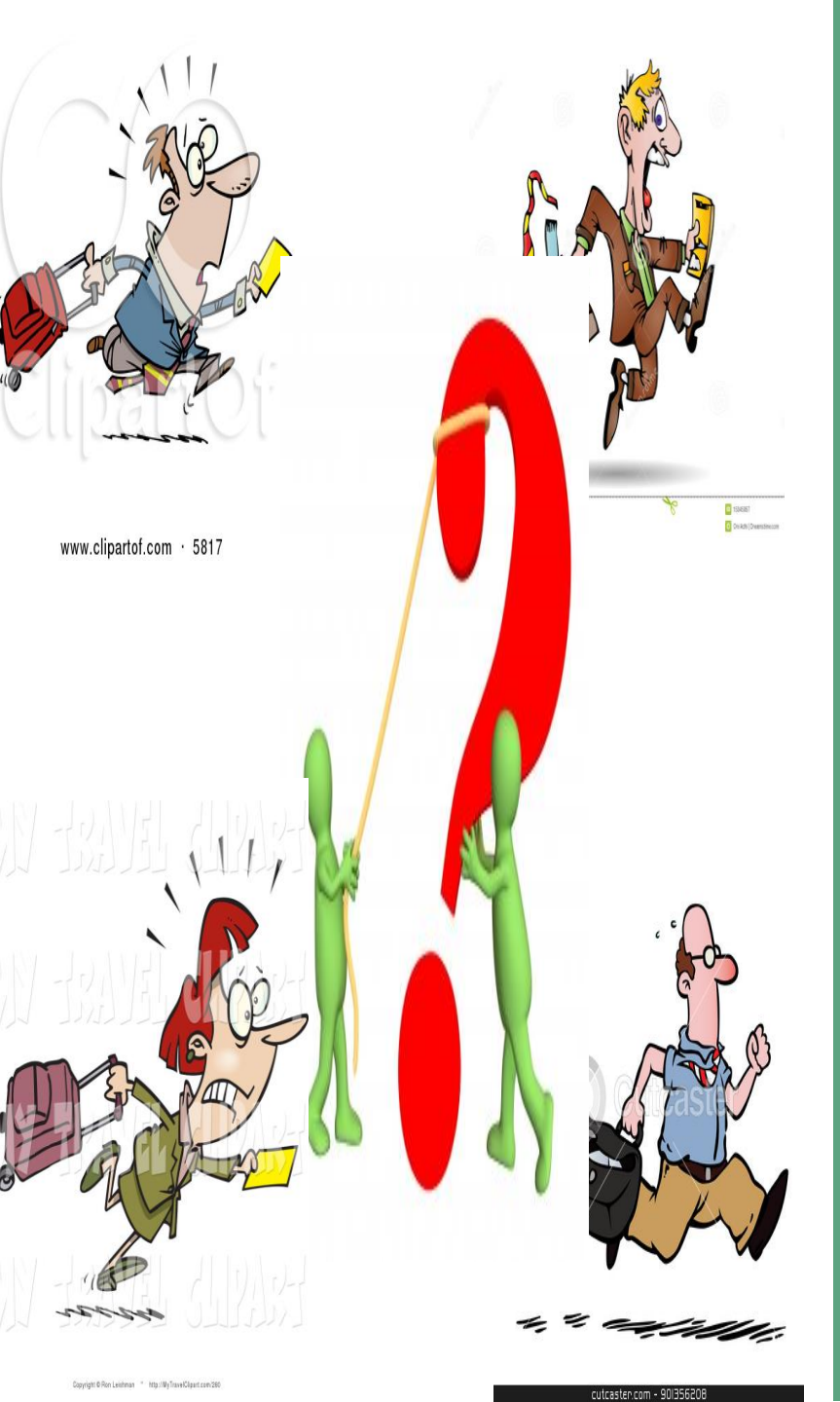


FR-FCFS BW

RESULTS

Main Lessons:

1. We investigate the major sources of memory interference in COTS architectures upon deploying mixed criticality systems.
2. We identified two main contributors to this interference:
   - Write batching and FR-FCFS arbitration.
3. Although theoretically increasing the amount of reordering done helps increasing the system peak BW, after certain point there are diminishing returns in practice
4. such increase also significantly increases worst-case delays
5. Reaching a compromise point between BW and WCD to meet all requirements of MCS is the potential direction to explore. Such point will be application-dependent.
6. We provide analytical equations as well as experimental evaluation to help designers deciding this point



Main Lessons:

1. We investigate the major sources of memory interference in COTS architectures upon deploying mixed criticality systems.
2. We identified two main contributors to this interference:
   - Write batching and FR-FCFS arbitration.
3. Although theoretically increasing the amount of reordering done helps increasing the system peak BW, after certain point there are diminishing returns in practice
4. such increase also significantly increases worst-case delays
5. Reaching a compromise point between BW and WCD to meet all requirements of MCS is the potential direction to explore. Such point will be application-dependent.
6. We provide analytical equations as well as experimental evaluation to help designers deciding this point