DRAM CONTROLLER

Mahdi Nazm Bojnordi

Assistant Professor

School of Computing

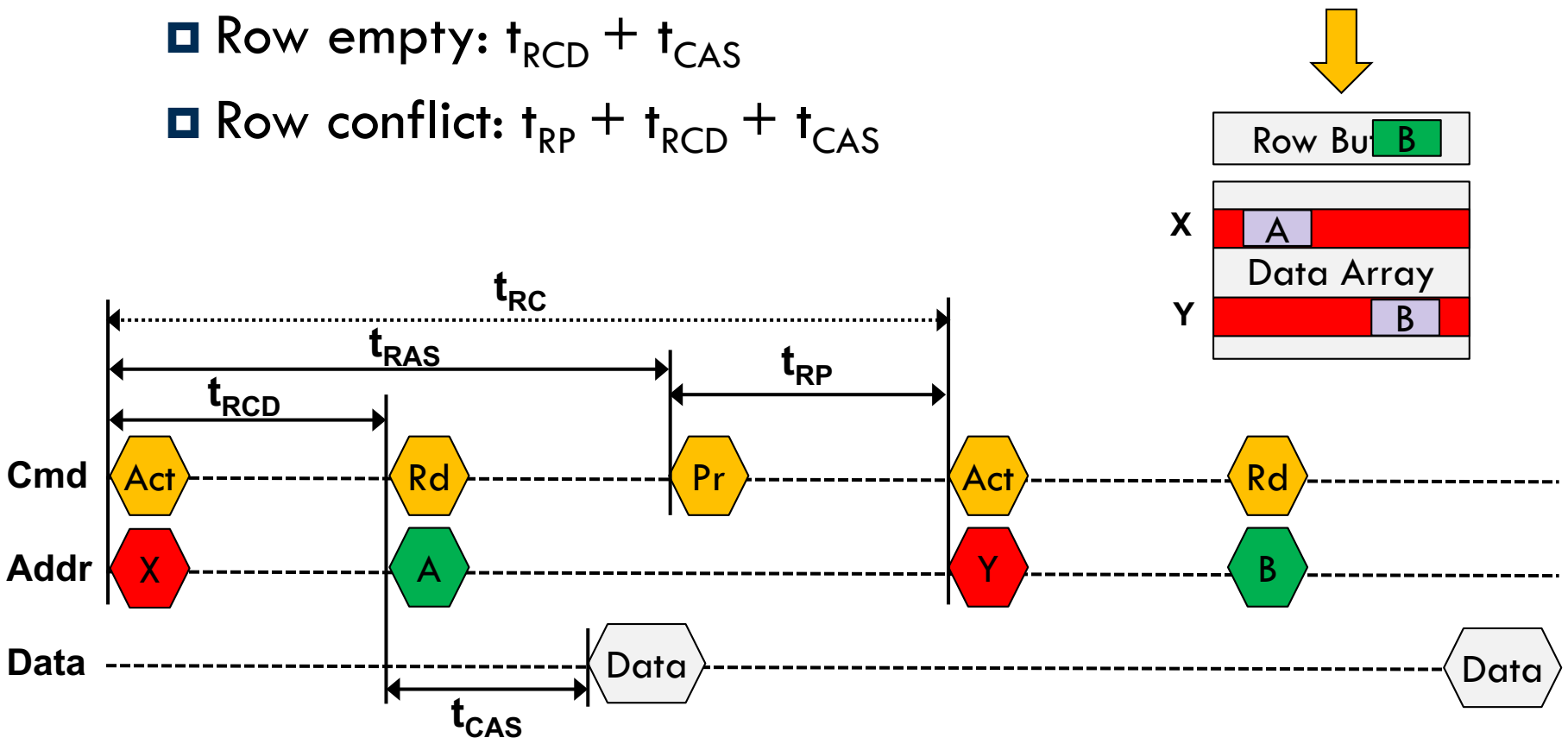
University of Utah

Overview

- Announcement
  - Homework 5 will be released on Nov. 13th
- This lecture
  - DRAM control
  - DRAM timing
  - DRAM hierarchy
    - Channel, bank

Recall: DRAM Timing Example

- Access time
  - Row hit: t_{CAS}
  - Row empty: $t_{RCD} + t_{CAS}$
  - Row conflict: $t_{RP} + t_{RCD} + t_{CAS}$

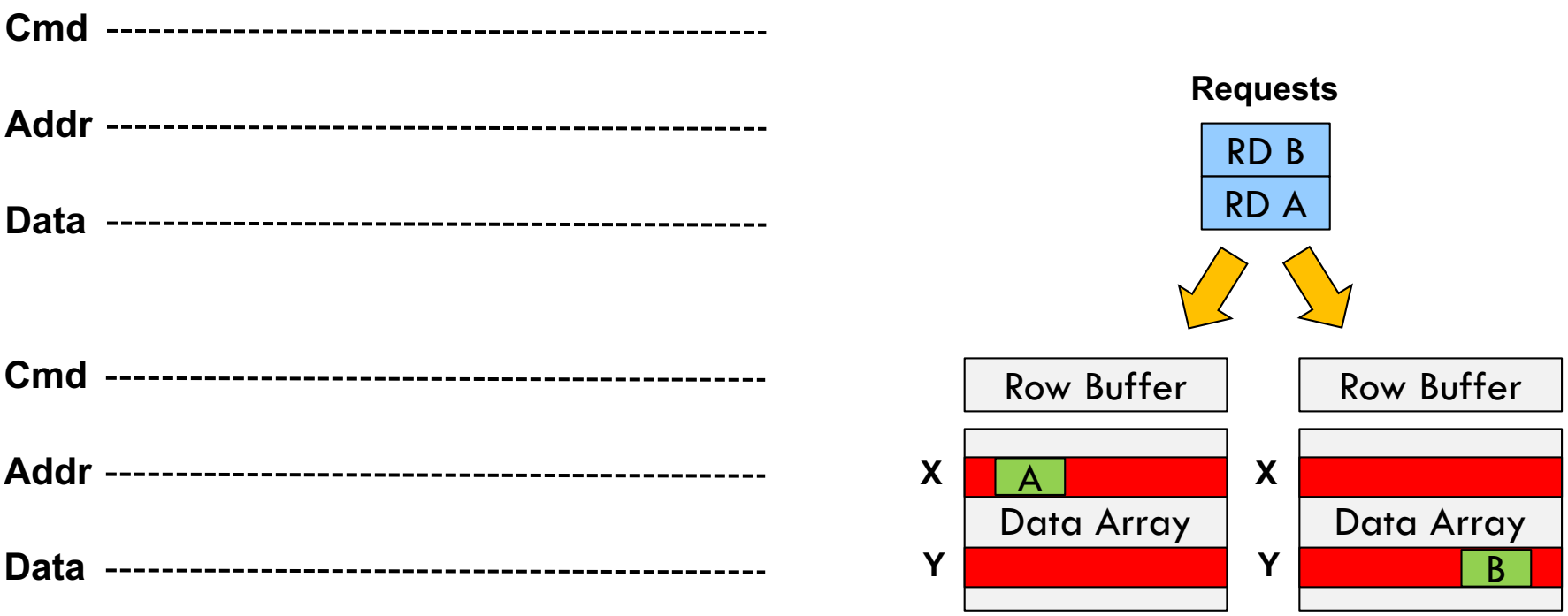


Improving Performance

DRAM Channels

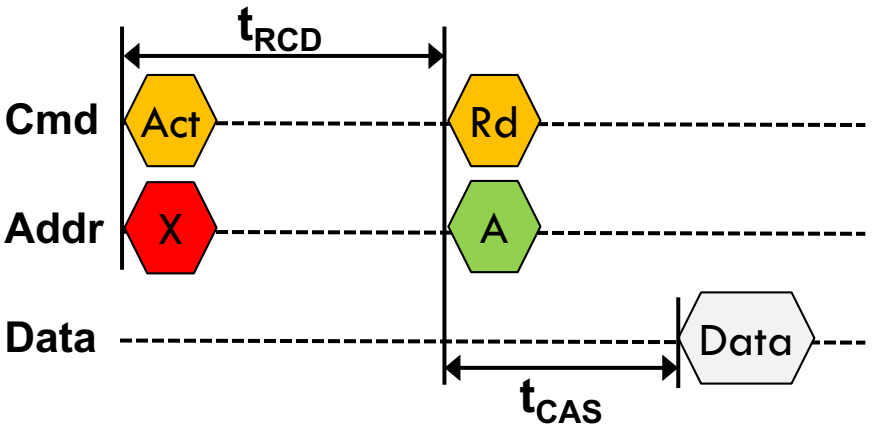
Memory Channels

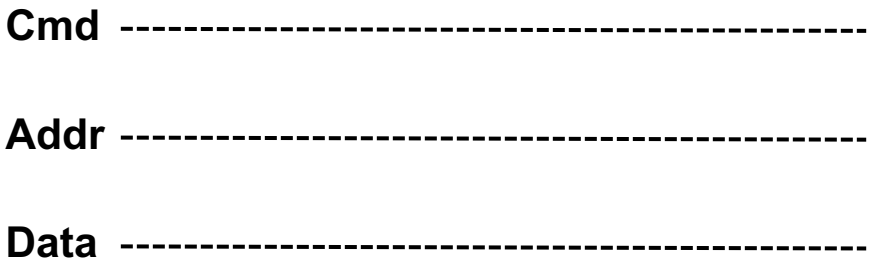
- Memory channels provide fully parallel accesses
  - Separate data, control, and address buses

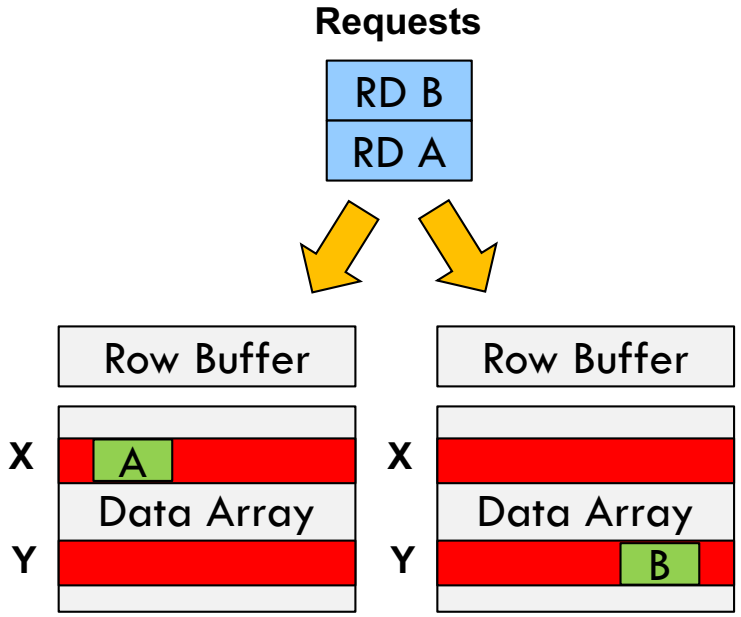


Memory Channels

- Memory channels provide fully parallel accesses
  - Separate data, control, and address buses

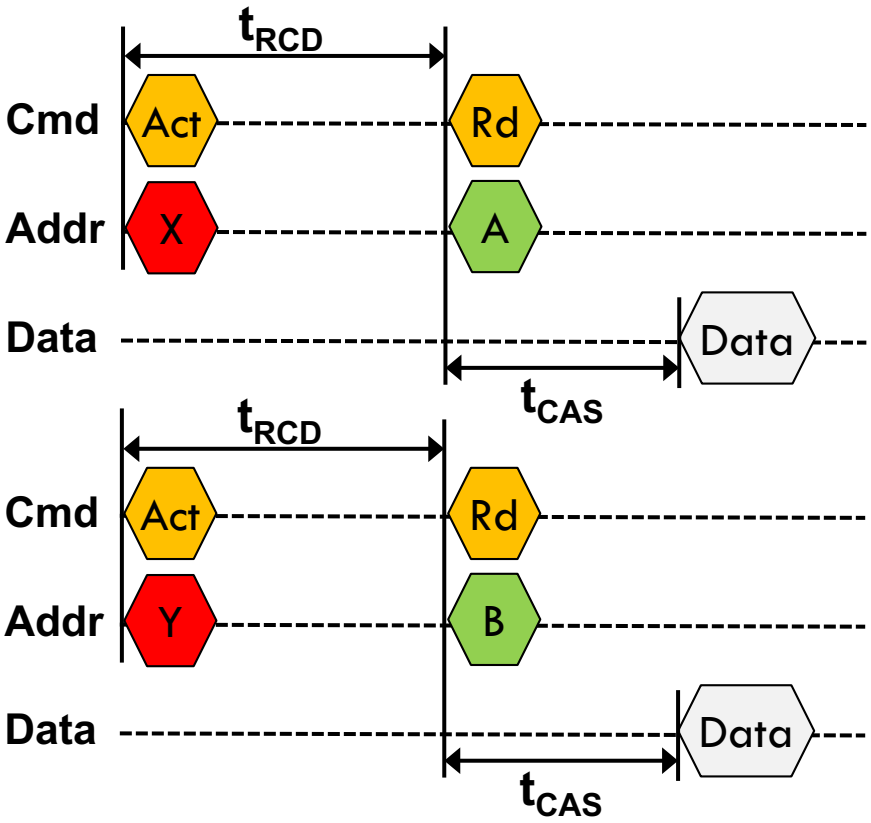


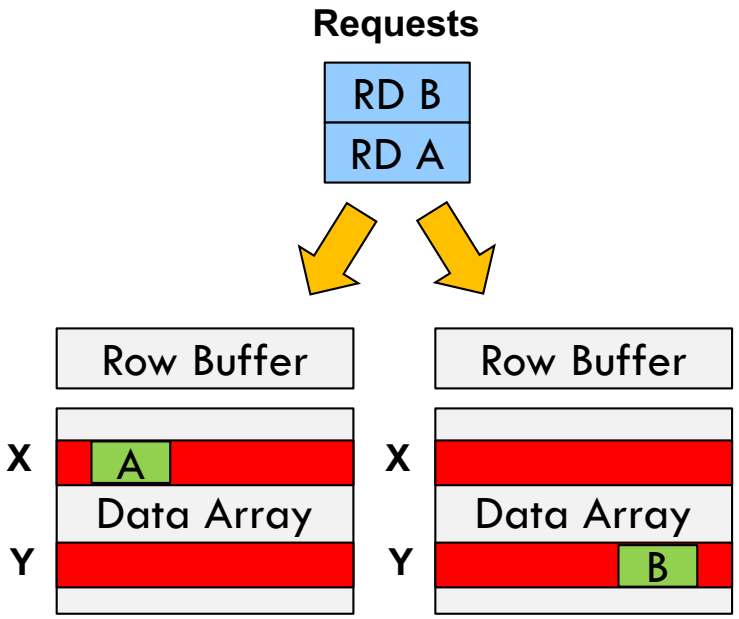




Memory Channels

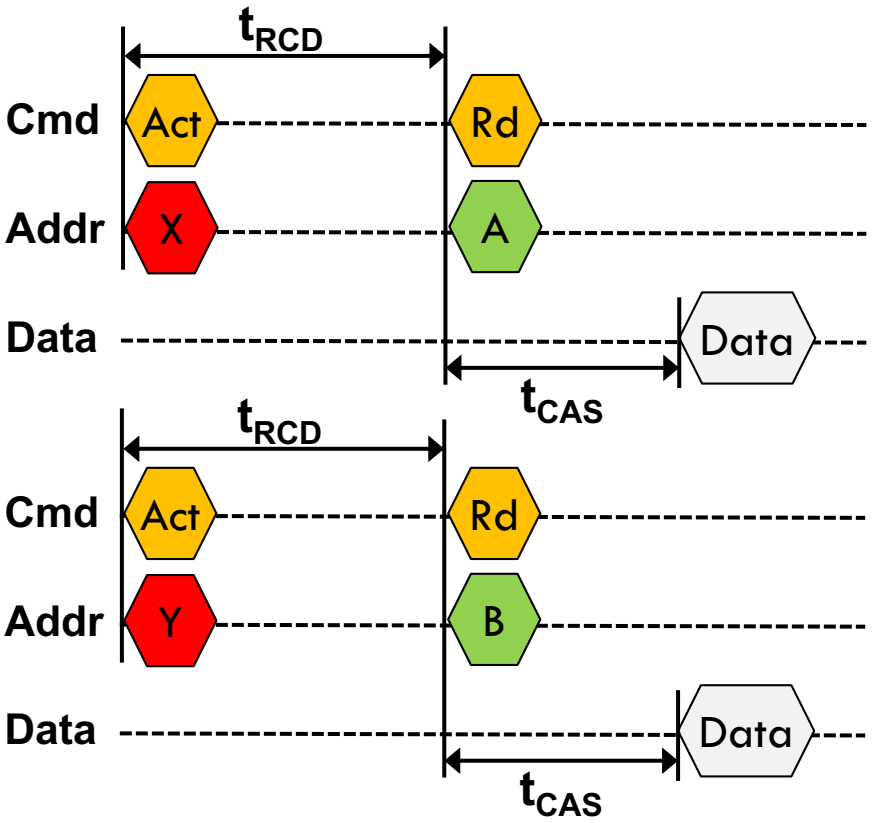
- Memory channels provide fully parallel accesses
  - Separate data, control, and address buses



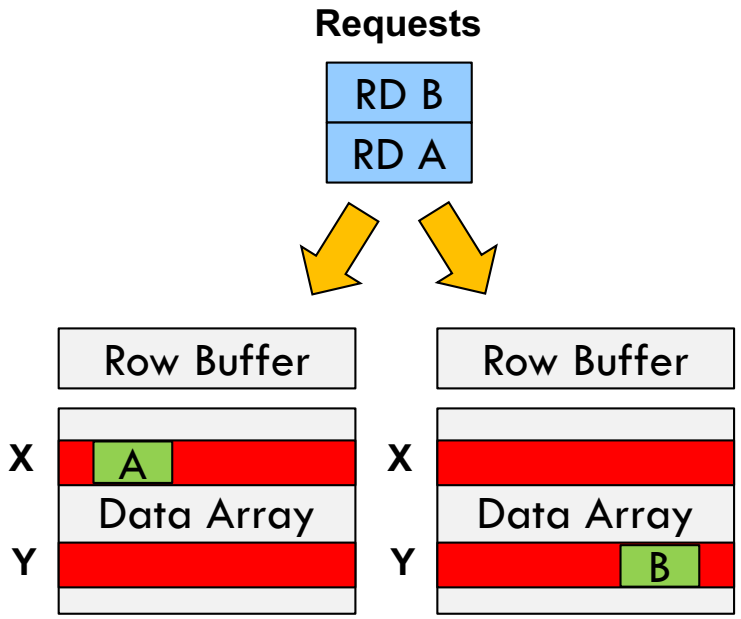


Memory Channels

- Memory channels provide fully parallel accesses
  - Separate data, control, and address buses



Not scalable due to pin overhead

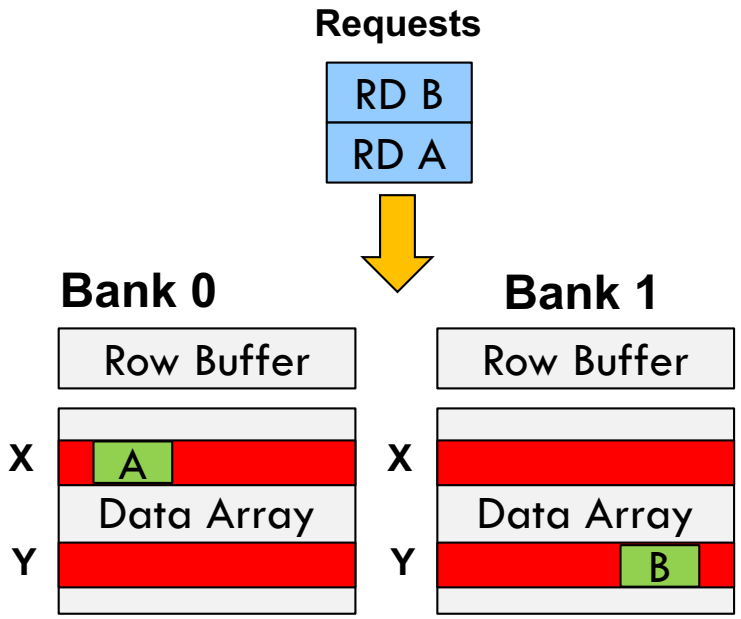


Improving Performance

DRAM Ranks

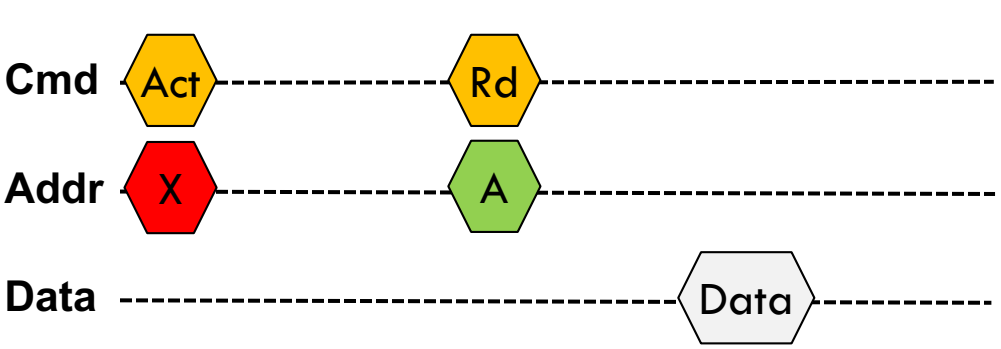
Memory Banks

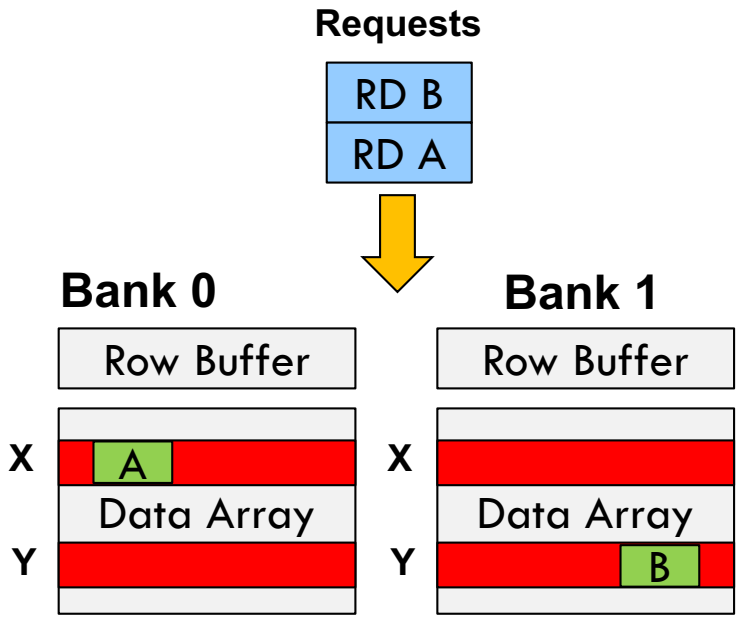
- Memory banks provide parallel operations
  - Shared data, control, and address buses
- The goal is to keep the data bus fully utilized



Memory Banks

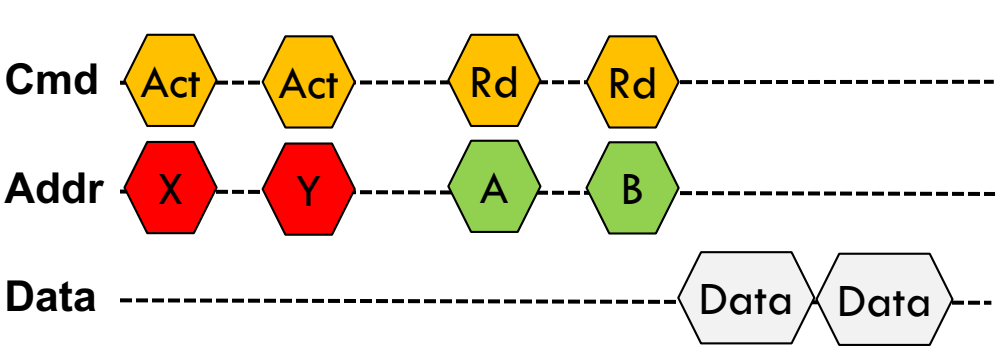
- Memory banks provide parallel operations
  - Shared data, control, and address buses
- The goal is to keep the data bus fully utilized

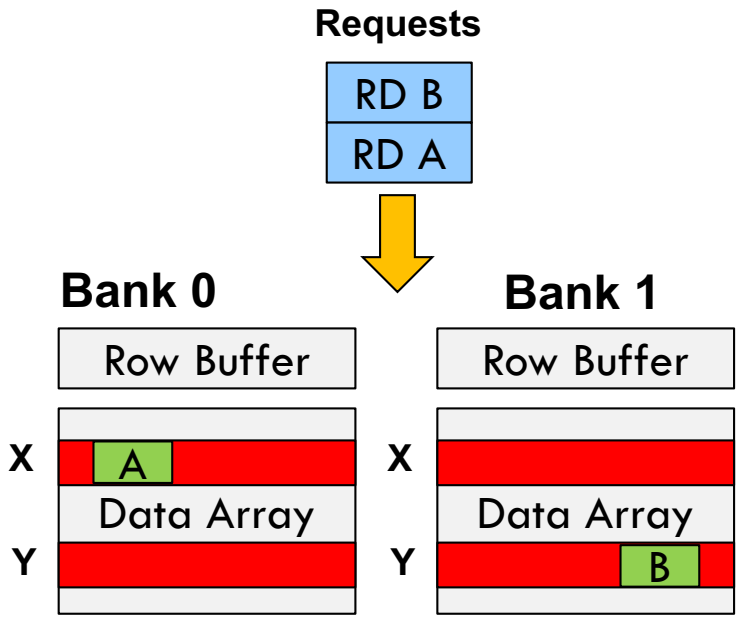




Memory Banks

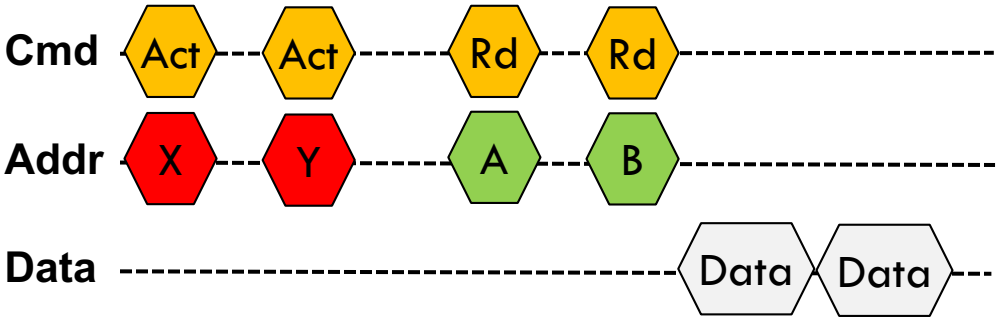
- Memory banks provide parallel operations
  - Shared data, control, and address buses
- The goal is to keep the data bus fully utilized

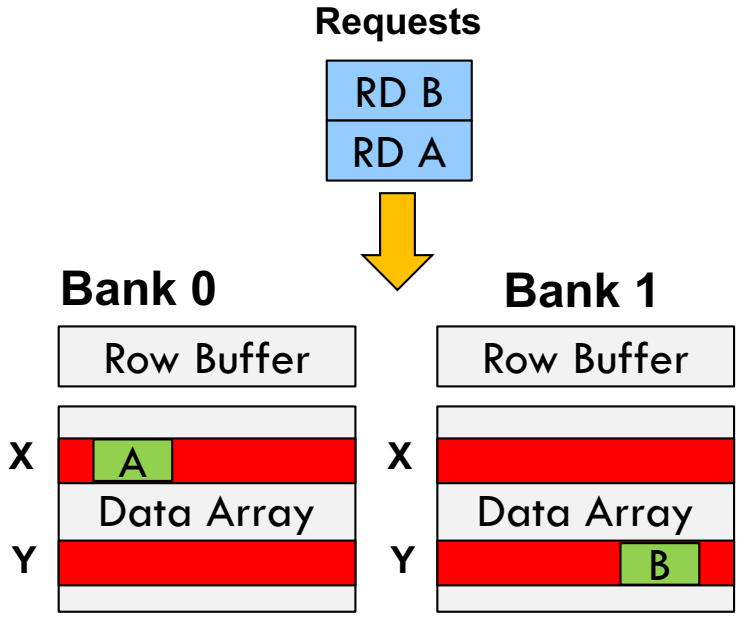




Memory Banks

- Memory banks provide parallel operations
  - Shared data, control, and address buses
- The goal is to keep the data bus fully utilized

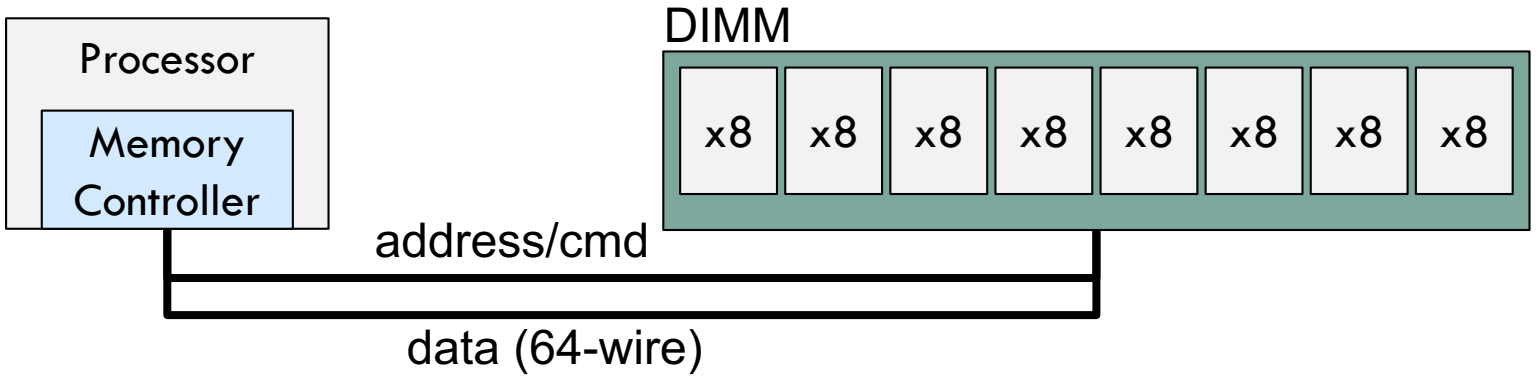




Shorter data transfer time to reduce bus conflicts
Double data rate vs. single rate

DRAM Organization

- DRAM channels are independently accessed through dedicated data, address, and command buses
  - Physically broken down into DIMMs (dual in-line memory modules)
  - Logically divided into ranks, which are a collection of DRAM chips responding to the same memory request



Memory Controller

- Memory controller connects CPU and DRAM
- Receives requests after cache misses in LLC
  - Possibly originating from multiple cores
- Complicated piece of hardware, handles:
  - DRAM refresh management
  - Command scheduling
  - Row-buffer management policies
  - Address mapping schemes

DRAM Control Tasks

- Refresh management
  - Periodically replenish the DRAM cells (burst vs. distributed)
- Address mapping
  - Distribute the requests to destination banks (load balancing)
- Request scheduling
  - Generate a sequence of commands for memory requests
    - Reduce overheads by eliminating unnecessary commands
- Power management
  - Keep the power consumption under a cap
- Error detection/correction
  - Detect and recover corrupted data

DRAM Addressing

Address Mapping

- A memory request

Type	Address	Data

- Address is used to find the location in memory
  - Channel, rank, bank, row, and column IDs
- Example physical address format

Row ID	Channel ID	Rank ID	Bank ID	Column ID

- A 4GB channel, 2 ranks, 4 banks/rank, 8KB page

Address Mapping

- A memory request

Type	Address	Data

- Address is used to find the location in memory
  - Channel, rank, bank, row, and column IDs
- Example physical address format

Row ID	Channel ID	Rank ID	Bank ID	Column ID
16	0	1	2	13

- A 4GB channel, 2 ranks, 4 banks/rank, 8KB page

Example Problem

- Start with empty row buffers, find the total number of commands if all the request are served in order
  - Address= row(12):channel(0):rank(1):bank(3):column(16)

addr

00000010

20000001

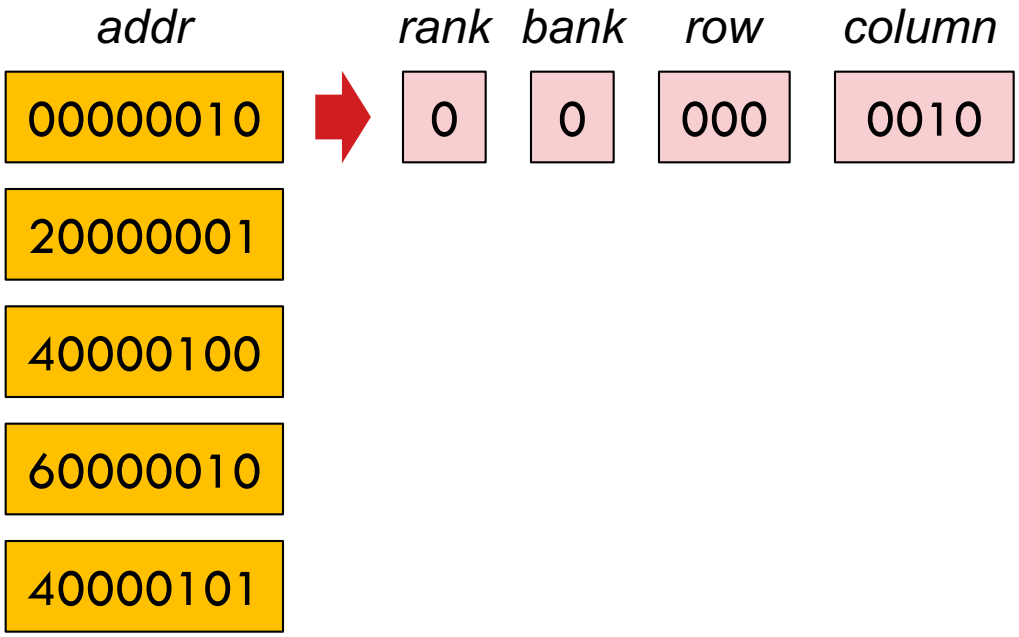
40000100

60000010

40000101

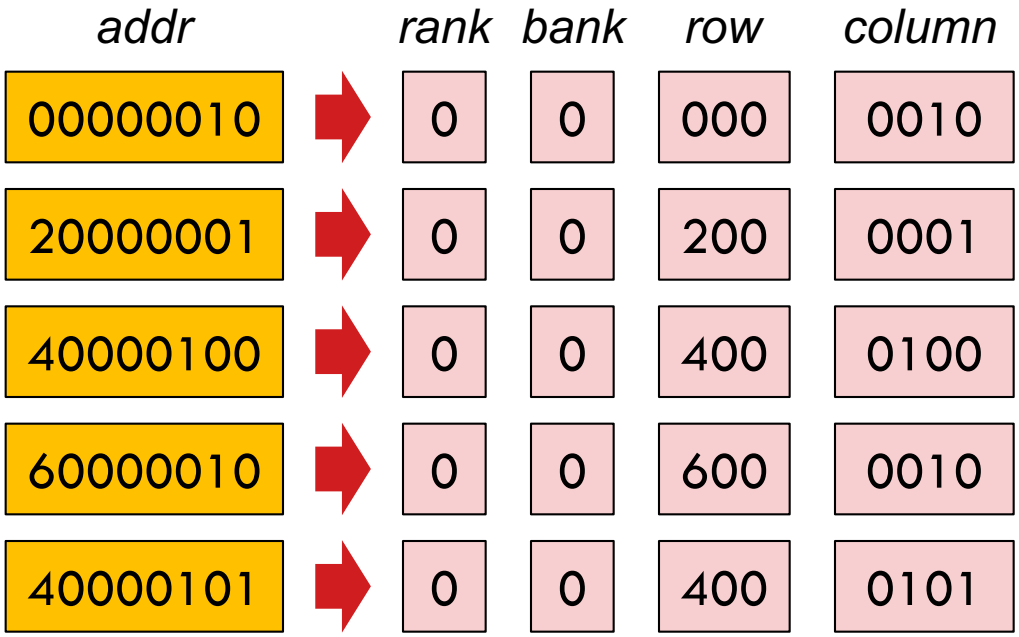
Example Problem

- Start with empty row buffers, find the total number of commands if all the request are served in order
  - Address= row(12):channel(0):rank(1):bank(3):column(16)



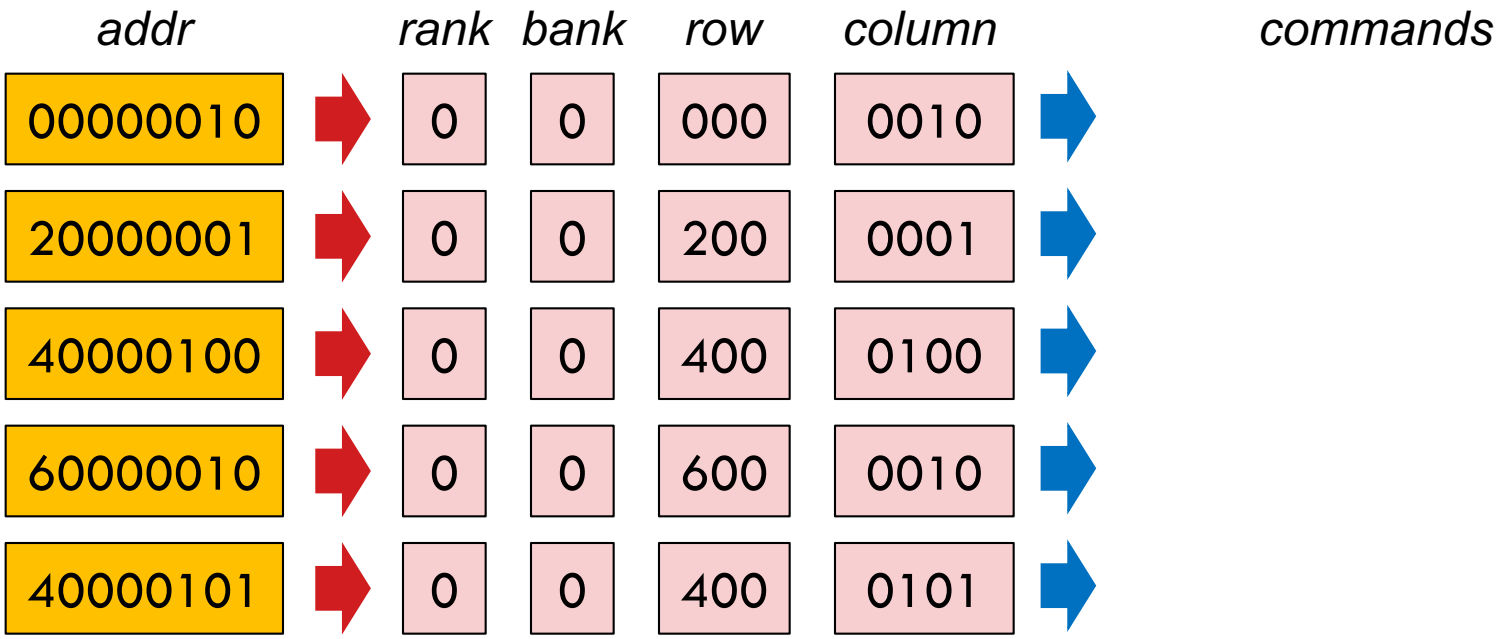
Example Problem

- Start with empty row buffers, find the total number of commands if all the request are served in order
  - Address= row(12):channel(0):rank(1):bank(3):column(16)



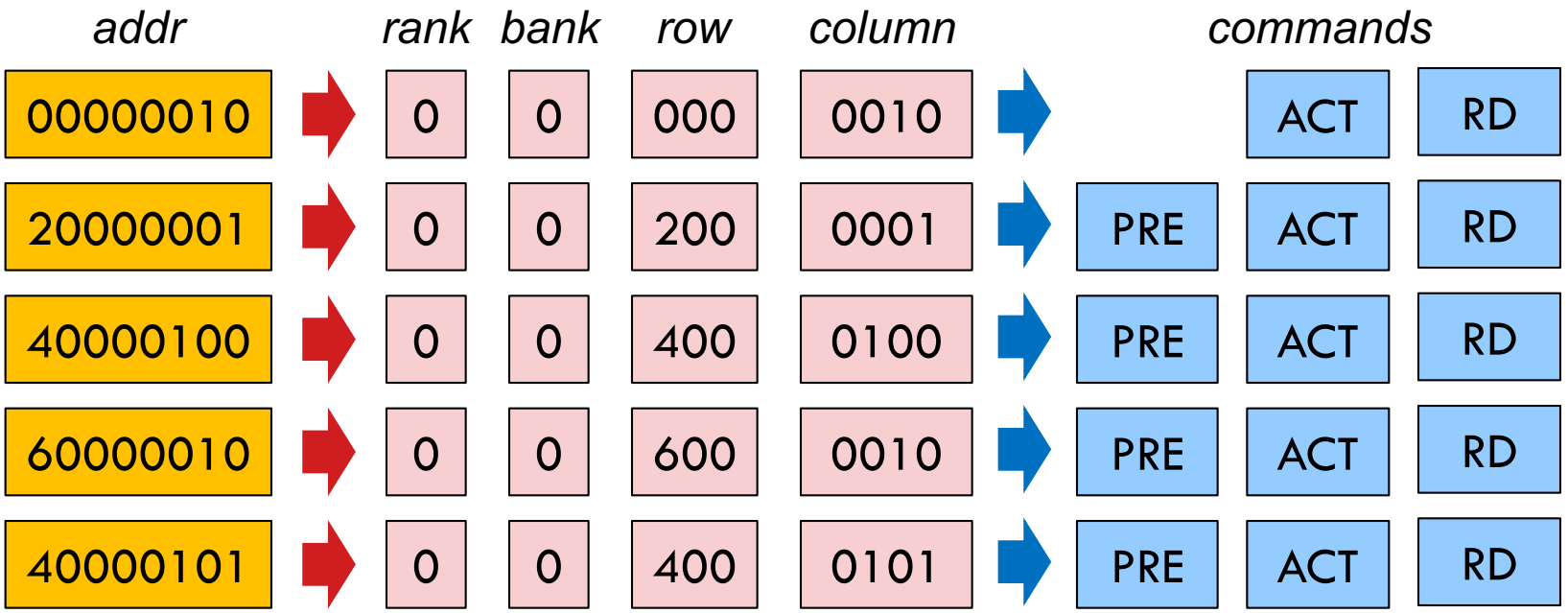
Example Problem

- Start with empty row buffers, find the total number of commands if all the request are served in order
  - Address= row(12):channel(0):rank(1):bank(3):column(16)



Example Problem

- Start with empty row buffers, find the total number of commands if all the request are served in order
  - Address= row(12):channel(0):rank(1):bank(3):column(16)



Example Problem

- Find the total number of commands using the following address mapping scheme
  - Address= bank(3):rank(1):channel(0):row(12):column(16)

addr

00000010

20000001

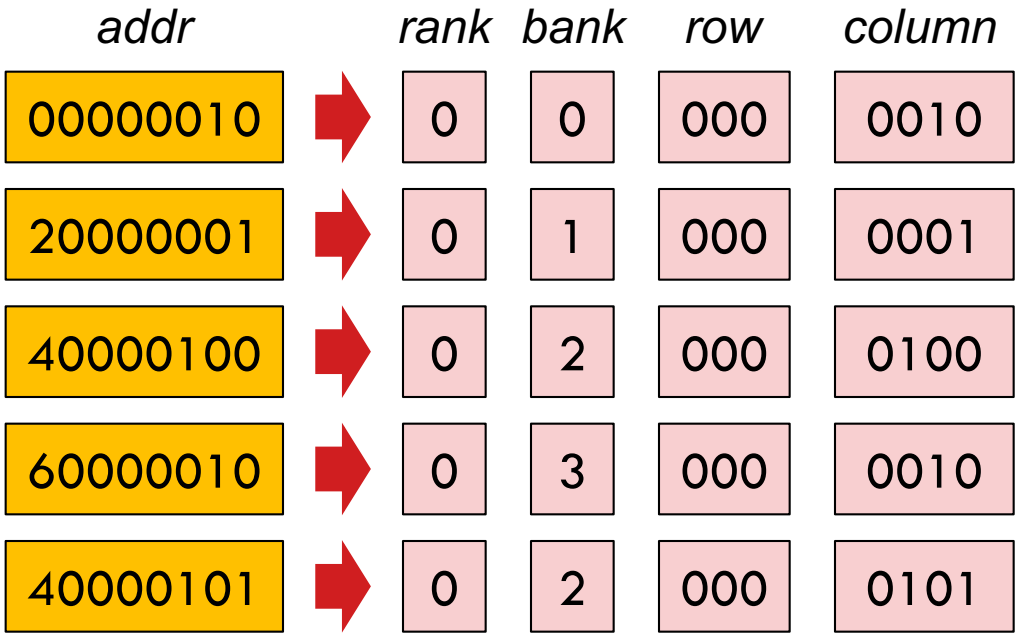
40000100

60000010

40000101

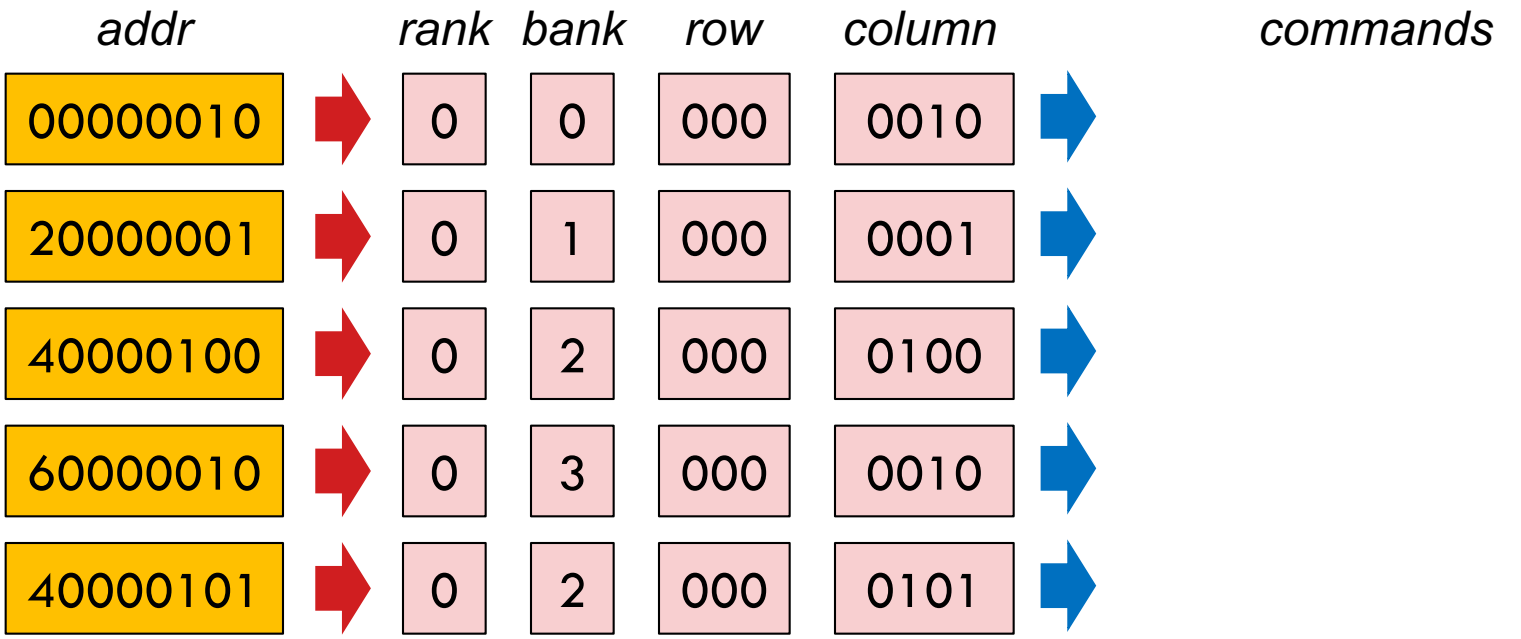
Example Problem

- Find the total number of commands using the following address mapping scheme
  - Address= bank(3):rank(1):channel(0):row(12):column(16)



Example Problem

- Find the total number of commands using the following address mapping scheme
  - Address= bank(3):rank(1):channel(0):row(12):column(16)



Example Problem

- Find the total number of commands using the following address mapping scheme
  - Address= bank(3):rank(1):channel(0):row(12):column(16)