Memory Devices

Paolo PRINETTO
Paolo.Prinetto@polito.it

Elena Ioana VATAJELU
Elena.Vatajelu@imag.fr

License Information

This presentation is licensed under the Creative Commons BY-NC License

To view a copy of the license, visit:
http://creativecommons.org/licenses/by-nc/3.0/legalcode

Disclaimer

- **We disclaim any warranties or representations as to the accuracy or completeness of this material.**
- **Materials are provided “as is” without warranty of any kind, either express or implied, including without limitation, warranties of merchantability, fitness for a particular purpose, and non-infringement.**
- **Under no circumstances shall we be liable for any loss, damage, liability or expense incurred or suffered which is claimed to have resulted from use of this material.**

Outline

- Introduction
- Memory Organization
- Memory Characteristics
- Access mechanisms
- Where we are coming from
- External vs Internal Architectures
- Memory Hierarchy
- Memories with ECCs

Outline

- Introduction
- **Memory Organization**
- Memory Characteristics
- Access mechanisms
- Where we are coming from
- External vs Internal Architectures
- Memory Hierarchy
- Memories with ECCs

Computer Memories

Physical devices used to store programs and data on a temporary or permanent basis in digital systems

Memory organization

- Memories are composed of **cells** (each cell stores 1 bit)
- Cells are organized in **words**
- From the user viewpoint, a *word is the information quantum that can be exchanged with the memory*

Memory operations

- On each word, data can be
  - Written
  - Read

Rel. 31.03.2017

Reset

- Memories are the only sequential devices without a reset signal

Outline

- Introduction
- Memory Organization
- **Memory Characteristics**
- Access mechanisms
- Where we are coming from
- External vs Internal Architectures
- Memory Hierarchy
- Memories with ECCs

Computer Memories characteristics

- Cell Size (F^2)

F = feature size = half the distance between like points in an array of minimum width and minimum spaced metal1

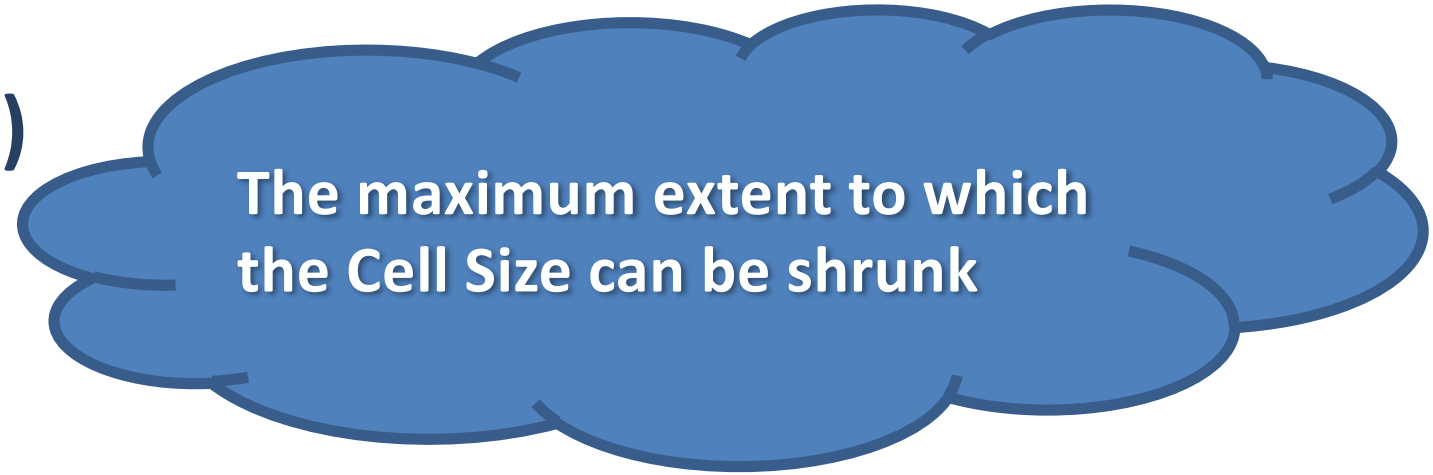
Memories Today

	SRAM	DRAM	NOR-Flash	NAND-Flash
Cell Size	$120F^2$	$4\text{-}6F^2$	$10\ F^2$	$4\text{-}5\ F^2$
Read Latency	<1ns	20ns	5,000ns	25,000ns
Write Latency	<1ns	20ns	1,000,000ns	200,000ns
Static power	YES	YES	NO	NO
Endurance	$>10^{15}$	$>10^{15}$	10^4	10^4
Non-volatility	NO	NO	YES	YES

Rel. 31.03.2017

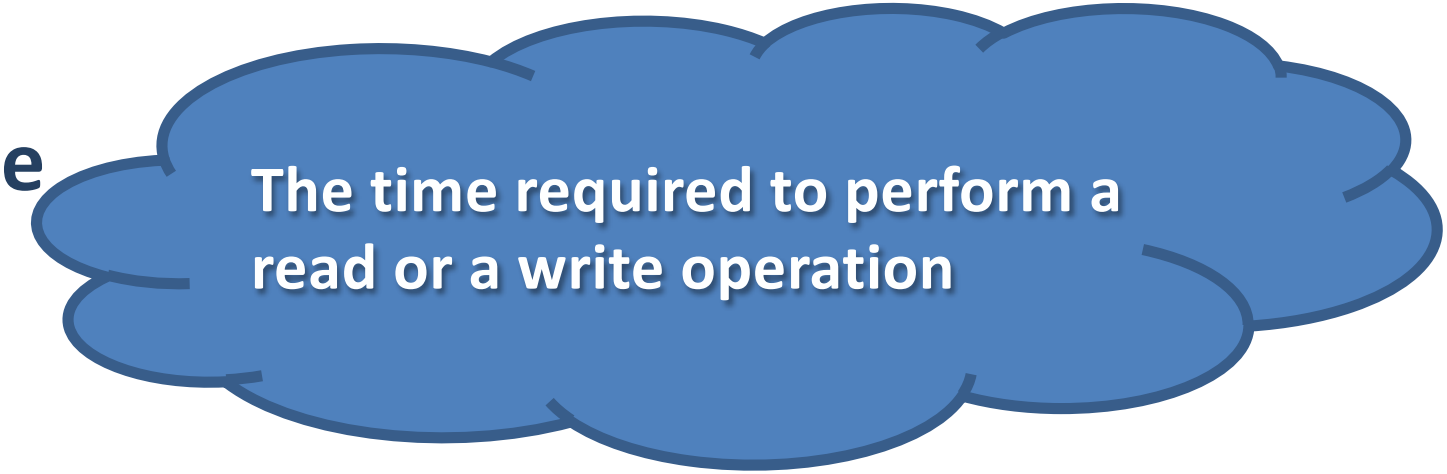
Computer Memories characteristics

- Cell Size (F^2)
- Scalability



Computer Memories characteristics

- Cell Size (F^2)
- Scalability
- Access Time



Memories Today

	SRAM	DRAM	NOR-Flash	NAND-Flash
Cell Size	$120F^2$	$4\text{-}6F^2$	$10\ F^2$	$4\text{-}5\ F^2$
Read Latency	<1ns	20ns	5,000ns	25,000ns
Write Latency	<1ns	20ns	1,000,000ns	200,000ns
Static power	YES	YES	NO	NO
Endurance	$>10^{15}$	$>10^{15}$	10^4	10^4
Non-volatility	NO	NO	YES	YES

Rel. 31.03.2017

Computer Memories characteristics

- Cell Size (F^2)
- Scalability
- Access Time
- Power Consumption

Rel. 31.03.2017

Computer Memories characteristics

- Cell Size (F^2)
- Scalability
- Access Time
- Power Consumption
- Endurance

The number of cycles the memory operates correctly

Memories Today

	SRAM	DRAM	NOR-Flash	NAND-Flash
Cell Size	120F^2	4-6F^2	10 F^2	4-5 F^2
Read Latency	<1ns	20ns	5,000ns	25,000ns
Write Latency	<1ns	20ns	1,000,000ns	200,000ns
Static power	YES	YES	NO	NO
Endurance	>10^{15}	>10^{15}	10^4	10^4
Non-volatility	NO	NO	YES	YES

Rel. 31.03.2017

Computer Memories characteristics

- Cell Size (F^2)
- Scalability
- Access Time
- Power Consumption
- Endurance
- Data Reliability

The ability to preserve data under disturbances

Computer Memories characteristics

- Cell Size (F^2)
- Scalability
- Access Time
- Power Consumption
- Endurance
- Data Reliability
- Volatility

The ability to preserve data when power is removed

Memories Today

	SRAM	DRAM	NOR-Flash	NAND-Flash
Cell Size	$120F^2$	$4\text{-}6F^2$	$10\ F^2$	$4\text{-}5\ F^2$
Read Latency	<1ns	20ns	5,000ns	25,000ns
Write Latency	<1ns	20ns	1,000,000ns	200,000ns
Static power	YES	YES	NO	NO
Endurance	$>10^{15}$	$>10^{15}$	10^4	10^4
Non-volatility	NO	NO	YES	YES

Rel. 31.03.2017

Computer Memories characteristics

- Cell Size (F^2)
- Scalability
- Access Time
- Power Consumption
- Endurance
- Data Reliability
- Volatility
- Fabrication Cost

Rel. 31.03.2017

Computer Memories characteristics

- Cell Size (F^2)
- Scalability
- Access Time
- Power Consumption
- Endurance
- Data Reliability
- Volatility
- Fabrication Cost
- Access mechanism

Computer Memories characteristics

- Cell Size (F^2)
- Scalability
- Access Time
- Power Consumption
- Endurance
- Data Reliability
- Volatility
- Fabrication Cost
- Access mechanism

- By *Address*
- By *Content*
- By *Writing order*

Rel. 31.03.2017

Outline

- Introduction
- Memory Organization
- Memory Characteristics
- **Access mechanisms**
- Where we are coming from
- External vs Internal Architectures
- Memory Hierarchy
- Memories with ECCs

Access Mechanisms

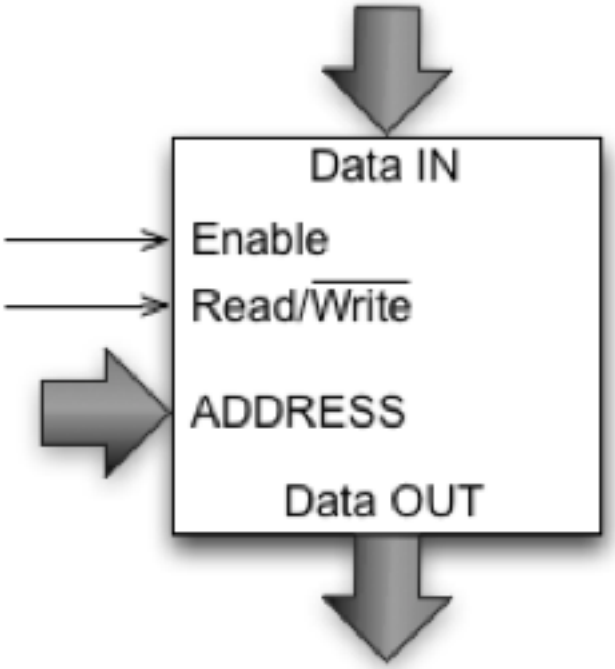
- *Data Access:*
  - By address
  - By content:
    - CAM (Content Addressable Memories)
    - Associative Memories
  - By writing order: FIFO (First-In First-Out)

Access Mechanisms

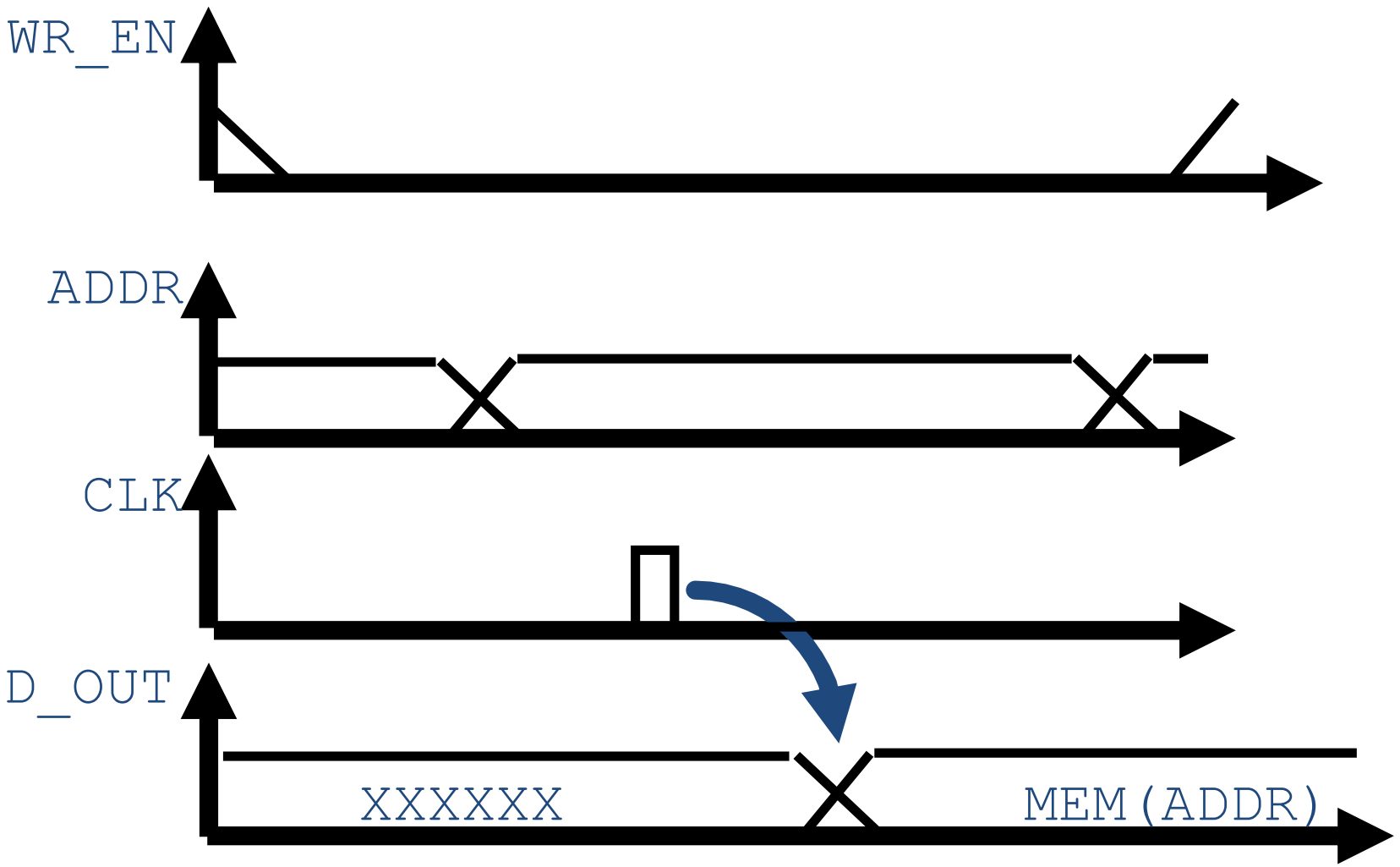
- *Data Access:*
  - By address
  - By content:
    - CAM (Content Addressable Memories)
    - Associative Memories
  - By writing order: FIFO (First-In First-Out)

Access by Address

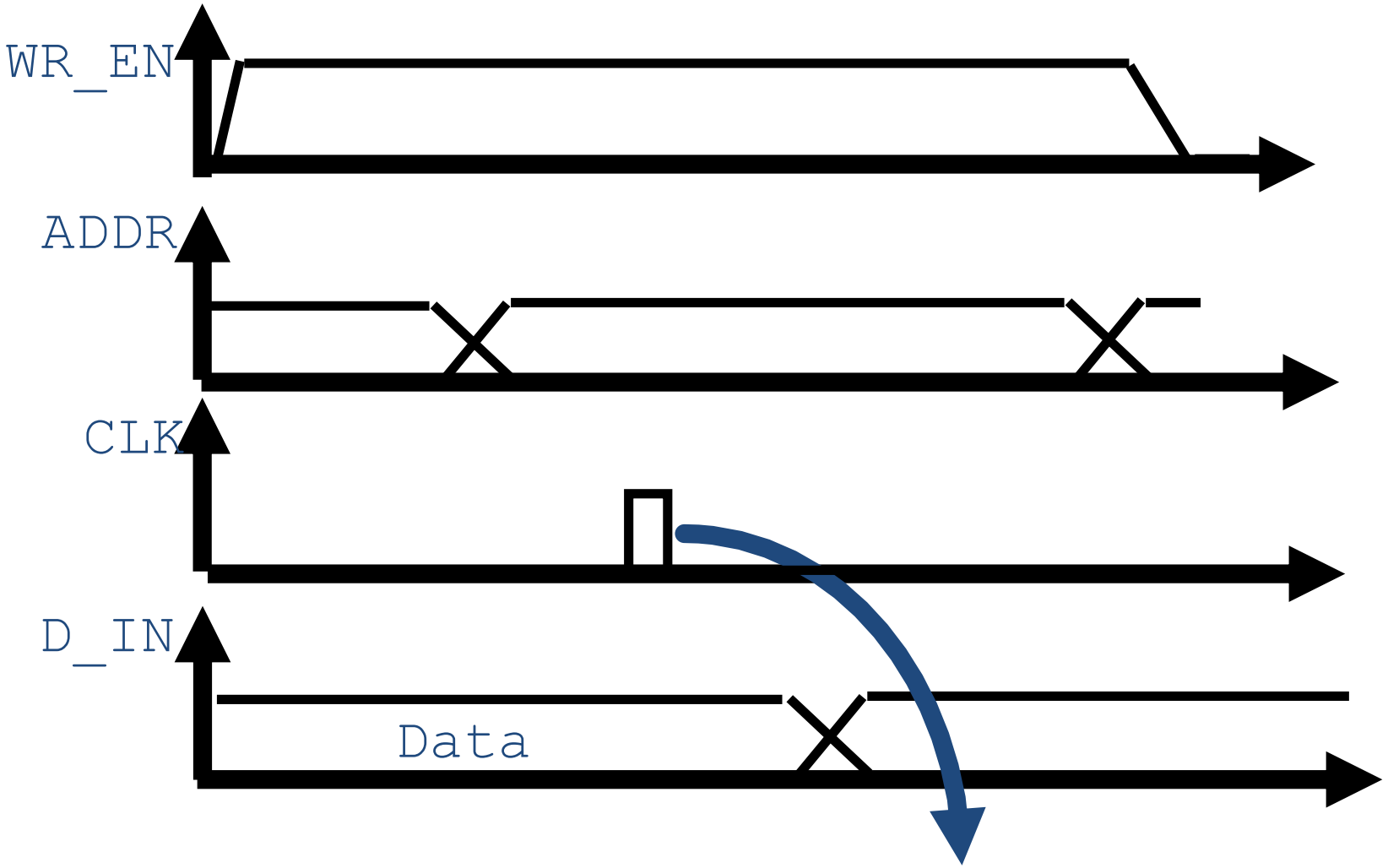
- Each word can be accessed independently and is univocally identified by an *Address*
- On each word, data can be
  - Written
  - Read



Timing diagram for a read operation



Timing diagram for a write operation



D_out *during write operation*

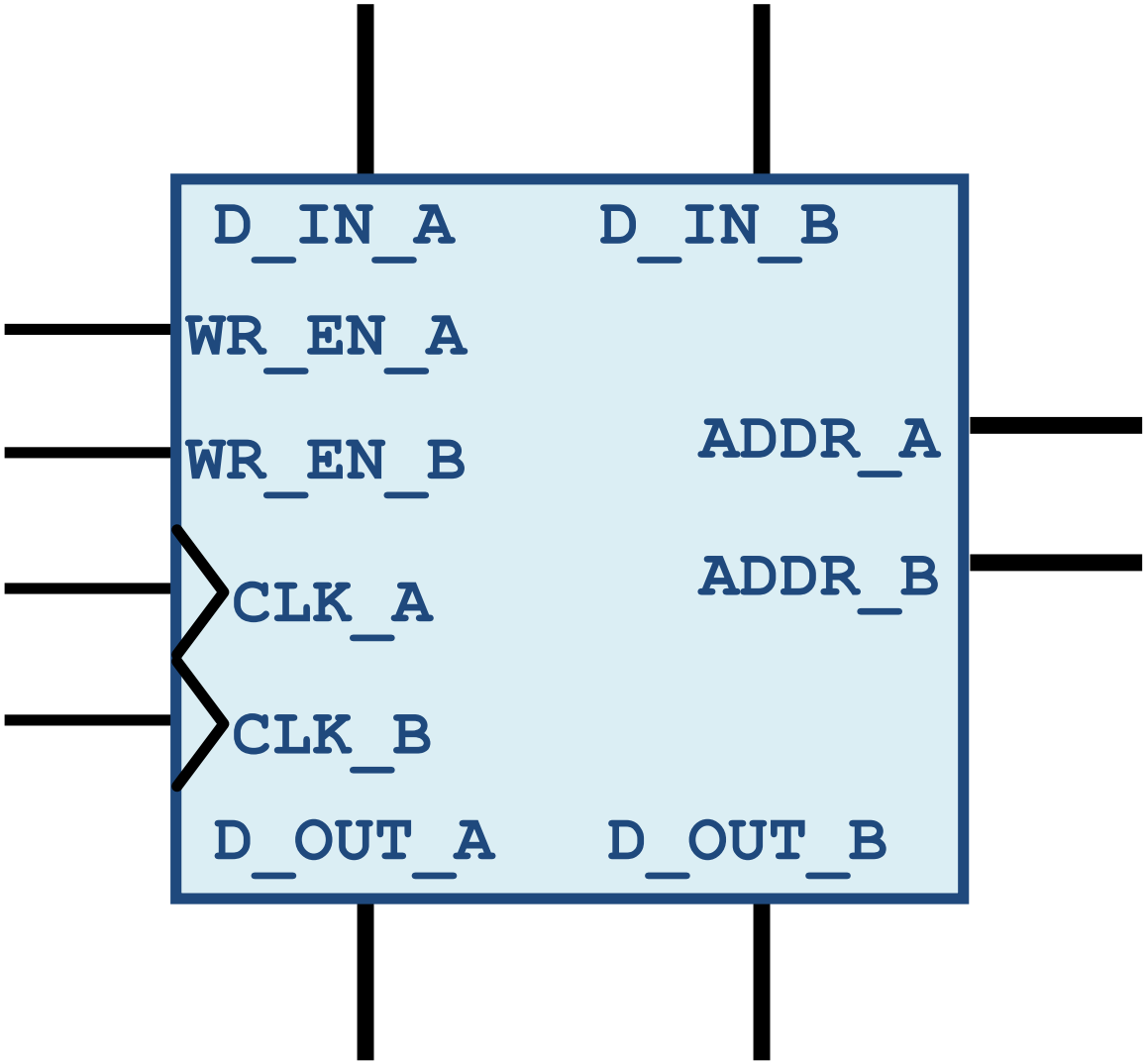
- Advanced memory supports three forms of write behavior:
  - **Normal** – Data on the output appears only during the read cycle. During a write cycle, the data (at the current address) does not appear on the output.
  - **Write Through** – A copy of the input data appears at the output of the same port.
  - **Read-Before-Write** – When new data is being written, the old contents of the address appears at the output.

Dual port RAM

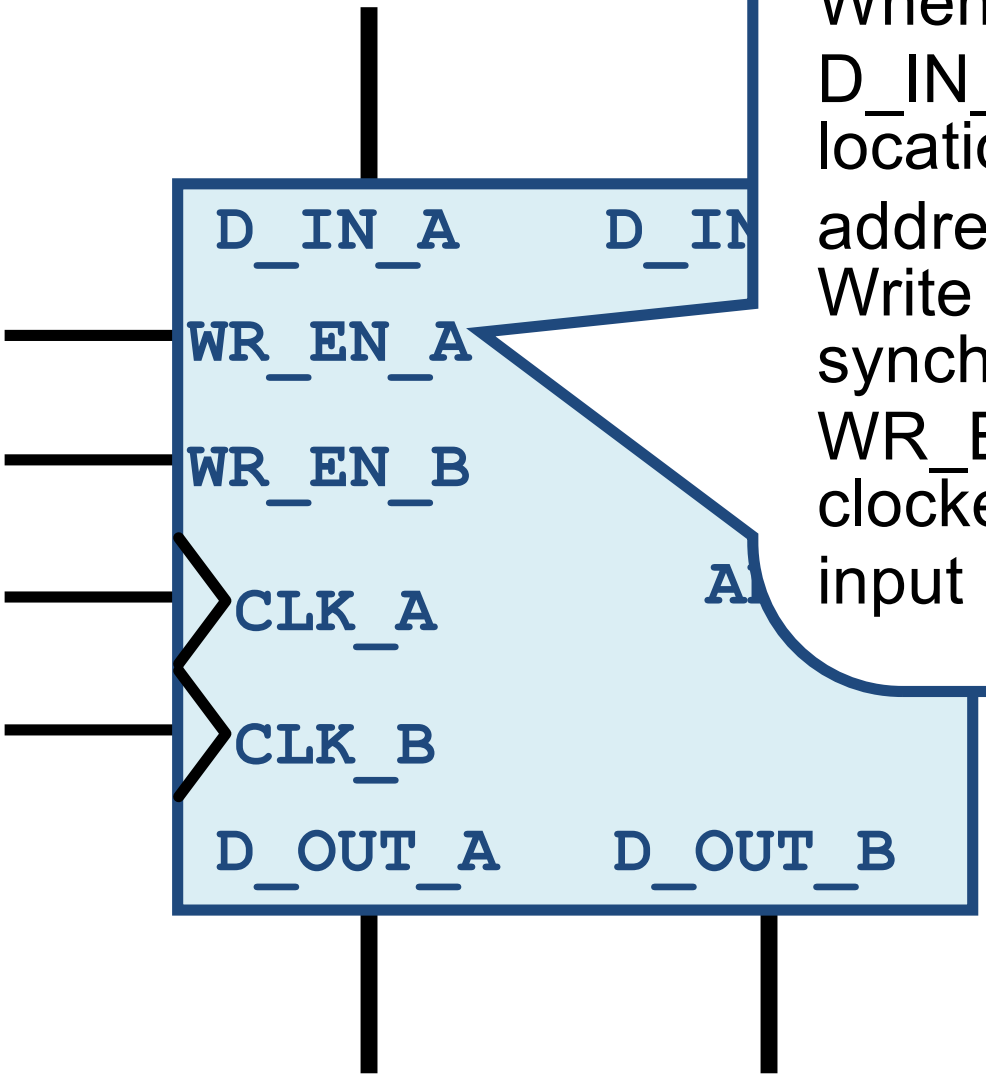
- Synchronous Read and Write Access
- Separate clock control of A and B ports
- Word Write
- Fully Synchronous independent operation from each port

Rel. 31.03.2017

Dual port RAM: symbol



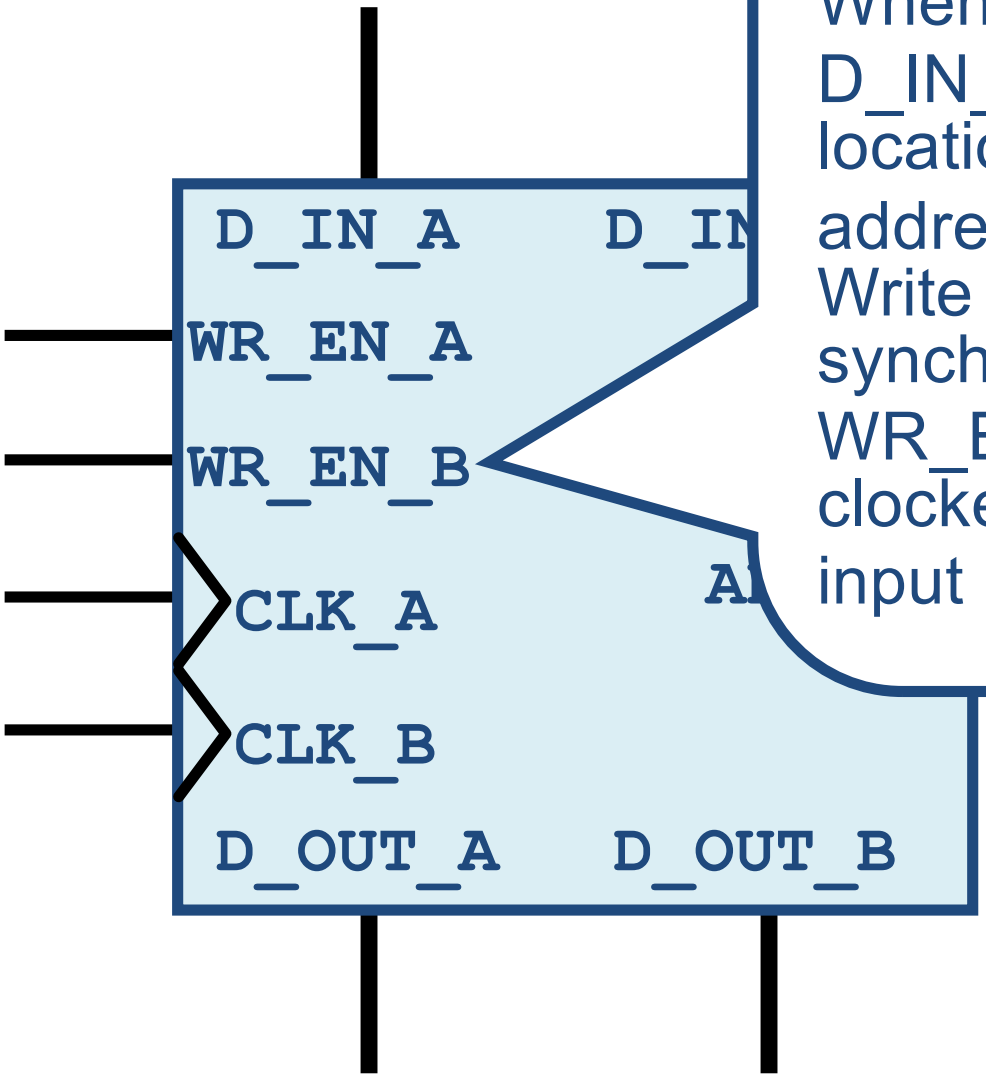
Dual port RAM: sym



Write Enable for port A

When high, writes D_IN_A into the memory location selected by address ADDR_A. Write operation is synchronous, and WR_EN_A signal is clocked in by CLK_A input

Dual port RAM: sym



Write Enable for port B

When high, writes D_IN_B into the memory location selected by address ADDR_B. Write operation is synchronous, and WR_EN_B signal is clocked in by CLK_B input

Dual port RAM: Principle of operation

- Simultaneously accessing the same memory location with same clock rate may result in the following:
  - Both ports writing: unknown data will be written
  - One port writing, other port reading: Write is completed, while the data read out is technology dependent.

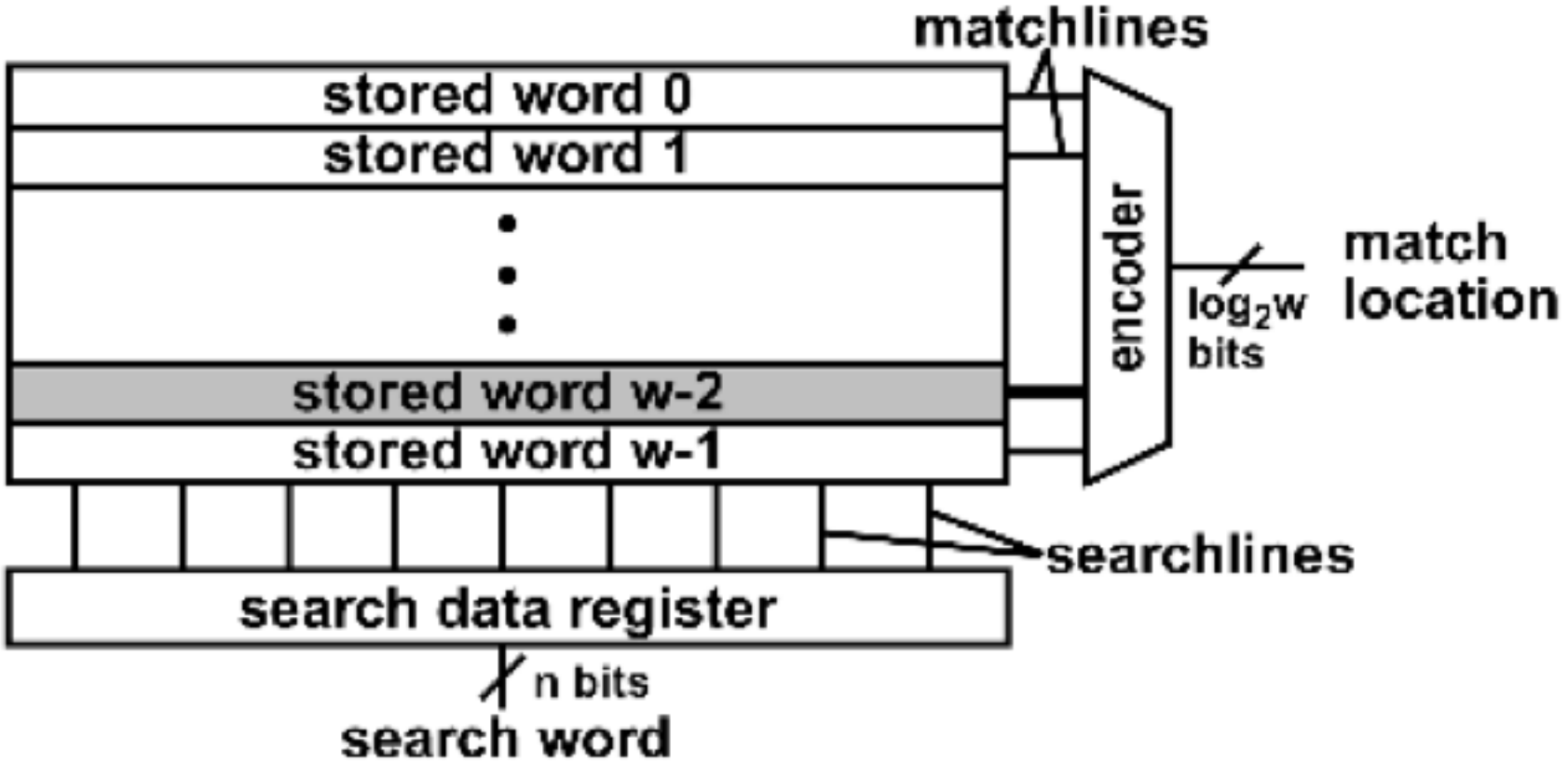
Access Mechanisms

- *Data Access:*
  - By address
  - By content:
    - CAM (Content Addressable Memories)
    - Ternary CAM
    - Associative Memories
  - By writing order: FIFO (First-In First-Out

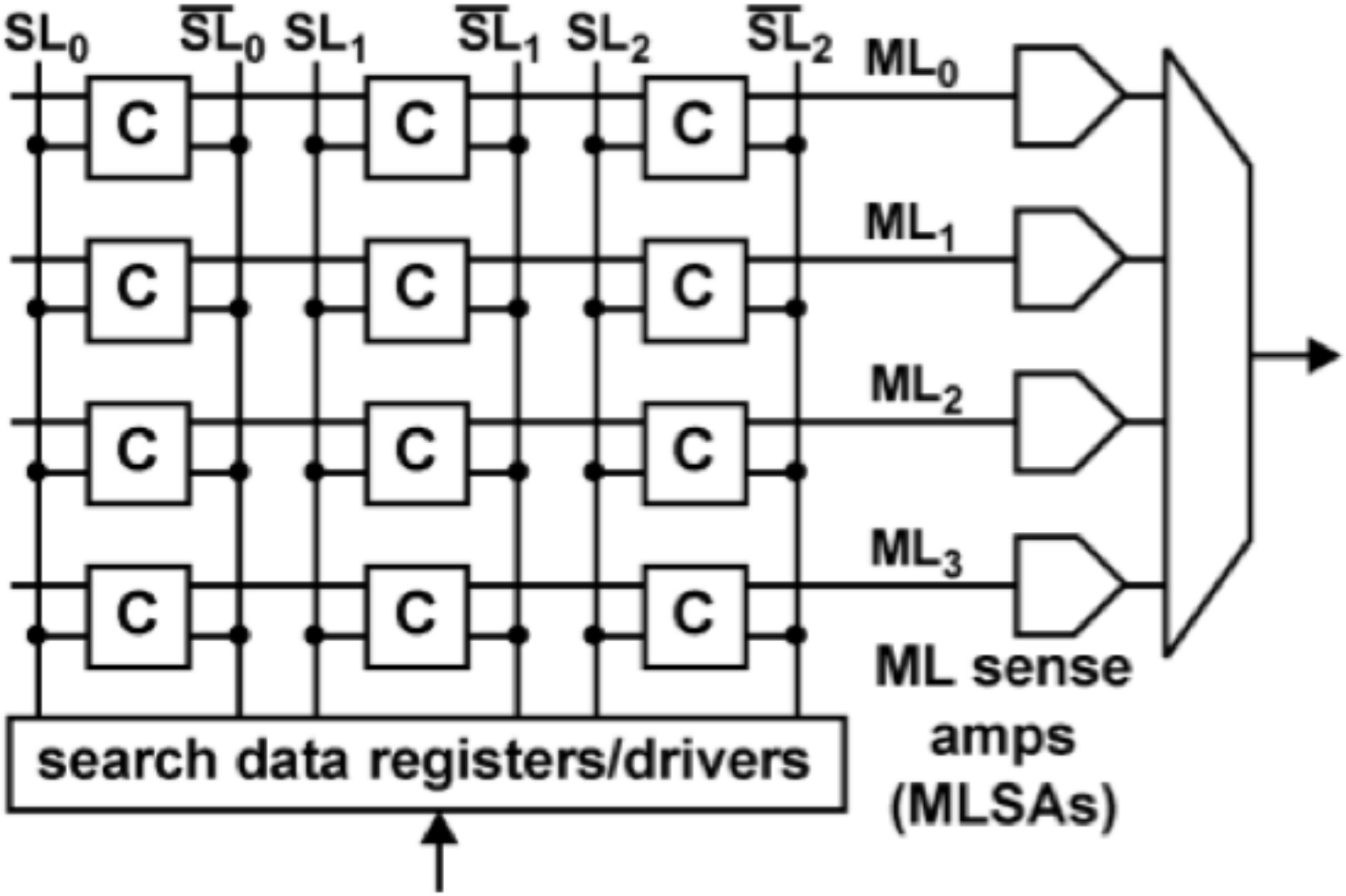
Content Addressable Memories

- A CAM gets in input a data (of the same size of the CAM words) and provides in output the address of the word whose content equals the input data
  - In case of multiple matches, the minum (or maximum) address is provided
  - In case of missmatch (the input data is not stored in the CAM), a status signal is asserted

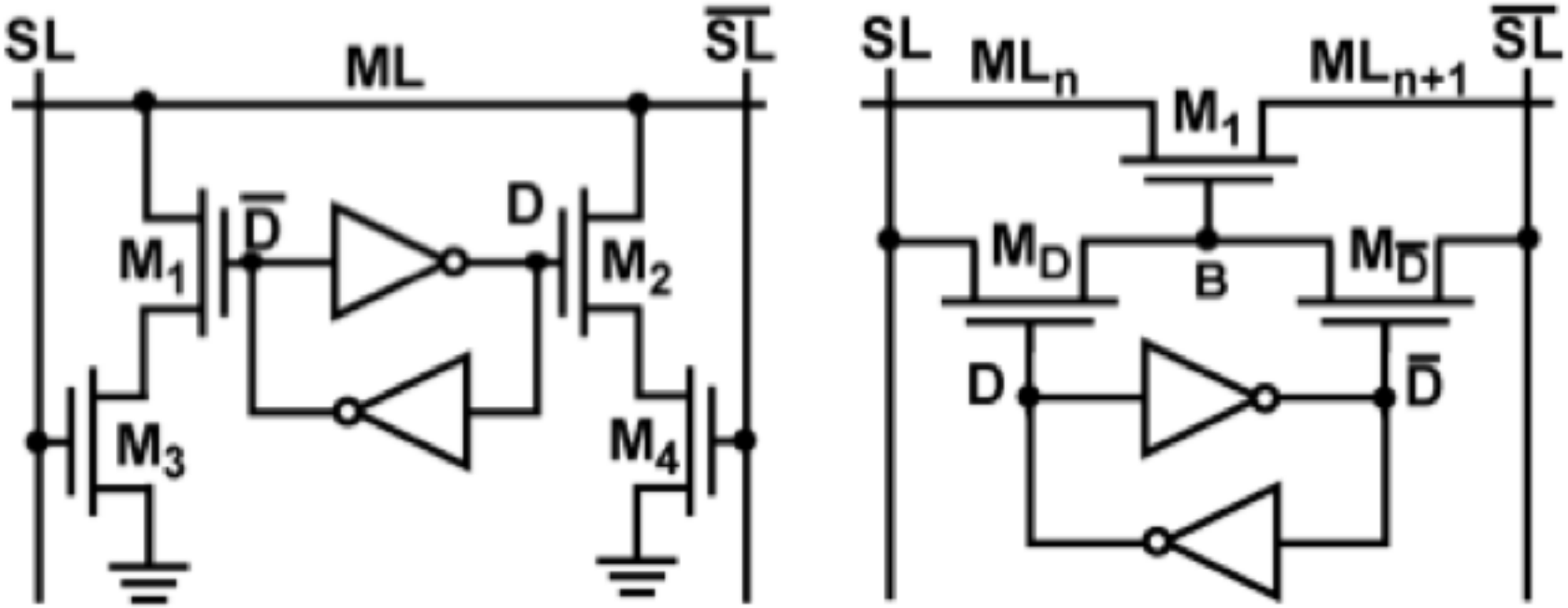
CAM conceptual view



CAM Array



CAM Cell



NOR

NAND

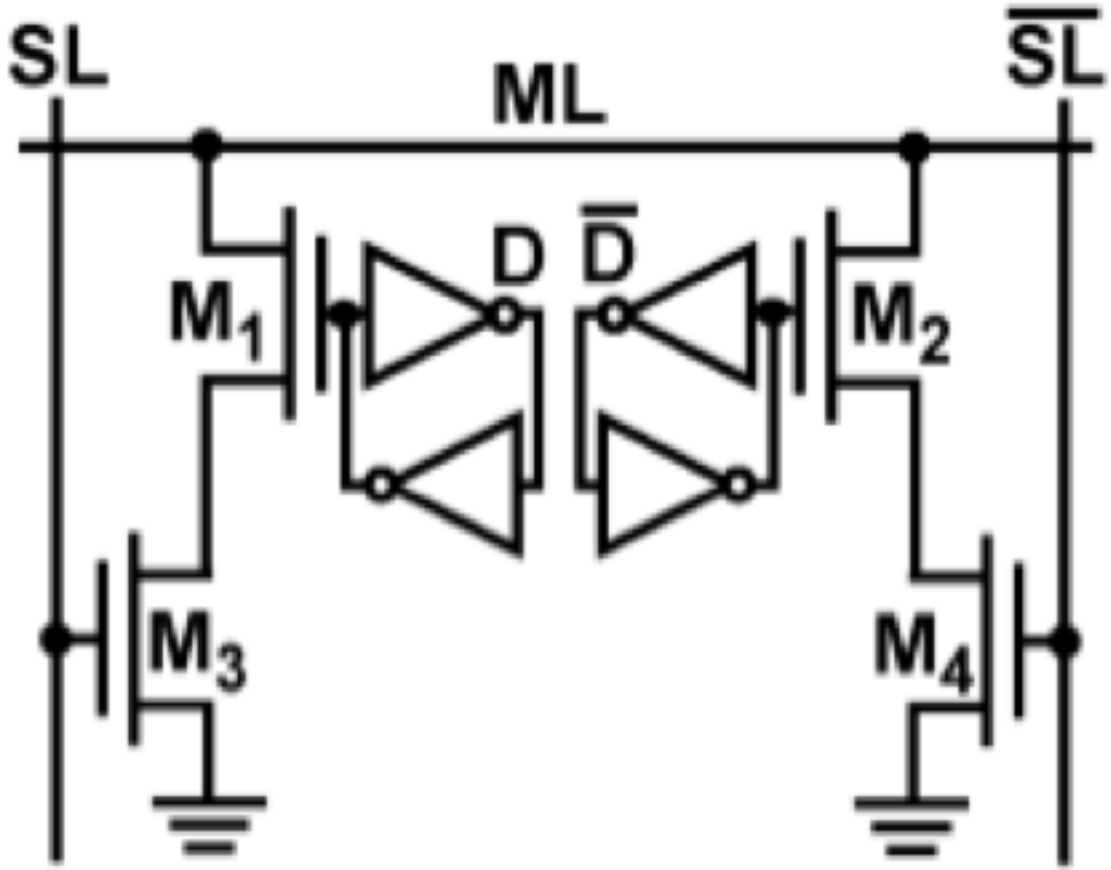
Access Mechanisms

- *Data Access:*
  - By address
  - By content:
    - CAM (Content Addressable Memories)
    - Ternary CAM
    - Associative Memories
  - By writing order: FIFO (First-In First-Out)

TCAM (Ternary CAM)

- A variation of CAM in which don't care values can be used during the search
- Conceptually, in a TCAM cell, one can store the value “-”, that during search will match with both “0” and “1”

TCAM Cell



Access Mechanisms

- *Data Access:*
  - By address
  - By content:
    - CAM (Content Addressable Memories)
    - Ternary CAM
    - Associative Memories
  - By writing order: FIFO (First-In First-Out)

Associative memories

- Sistemi principalmente costituiti da memorie di tipologia CAM più della logica sparsa e talvolta reti neurali
- Utilizzati per ricerche nelle quali il matching avvenga anche in caso di un certo grado di "similitudine" tra la stringa cercata e quella memorizzata nella memoria (e.g., Christophoro Columbus potrebbe essere riconosciuto come Cristoforo Colombo).

Access Mechanisms

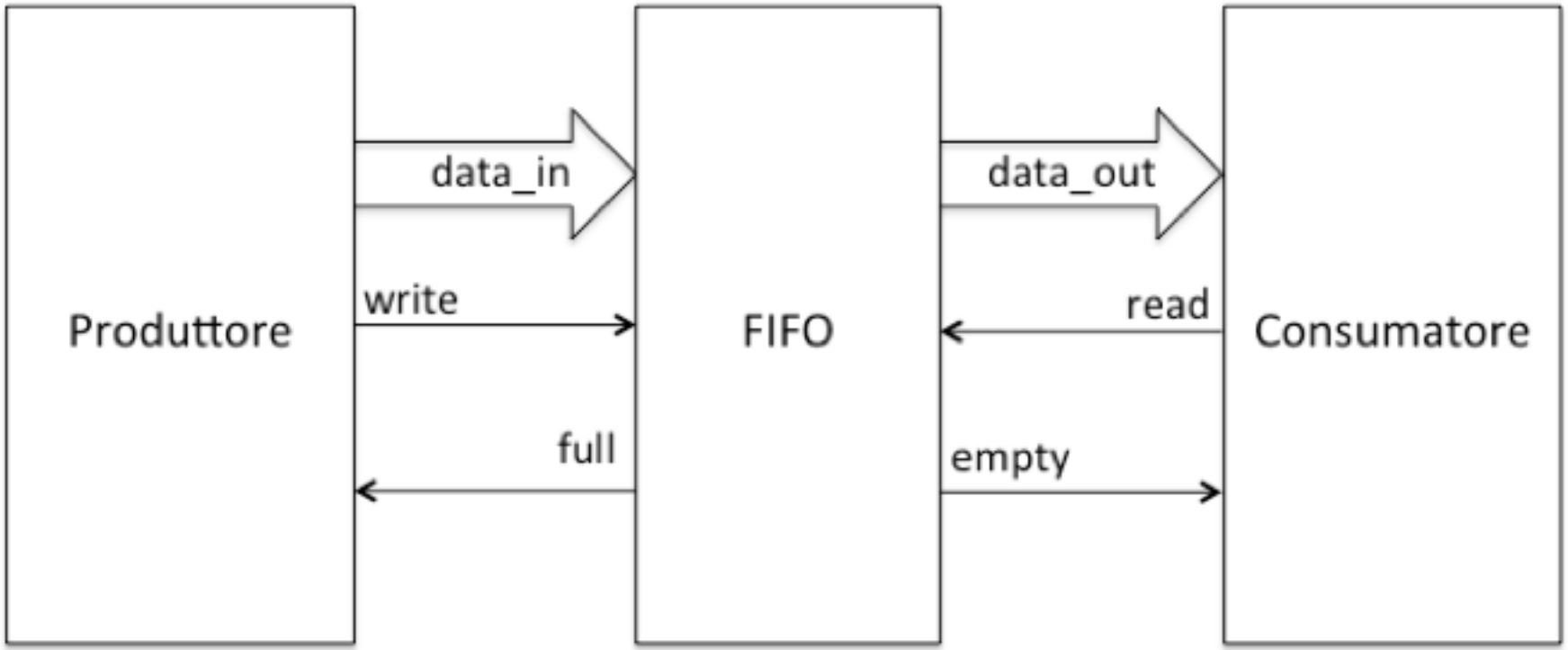
- *Data Access:*
  - By address
  - By content:
    - CAM (Content Addressable Memories)
    - Ternary CAM
    - Associative Memories
  - By writing order: FIFO (First-In First-Out)

FIFO memories

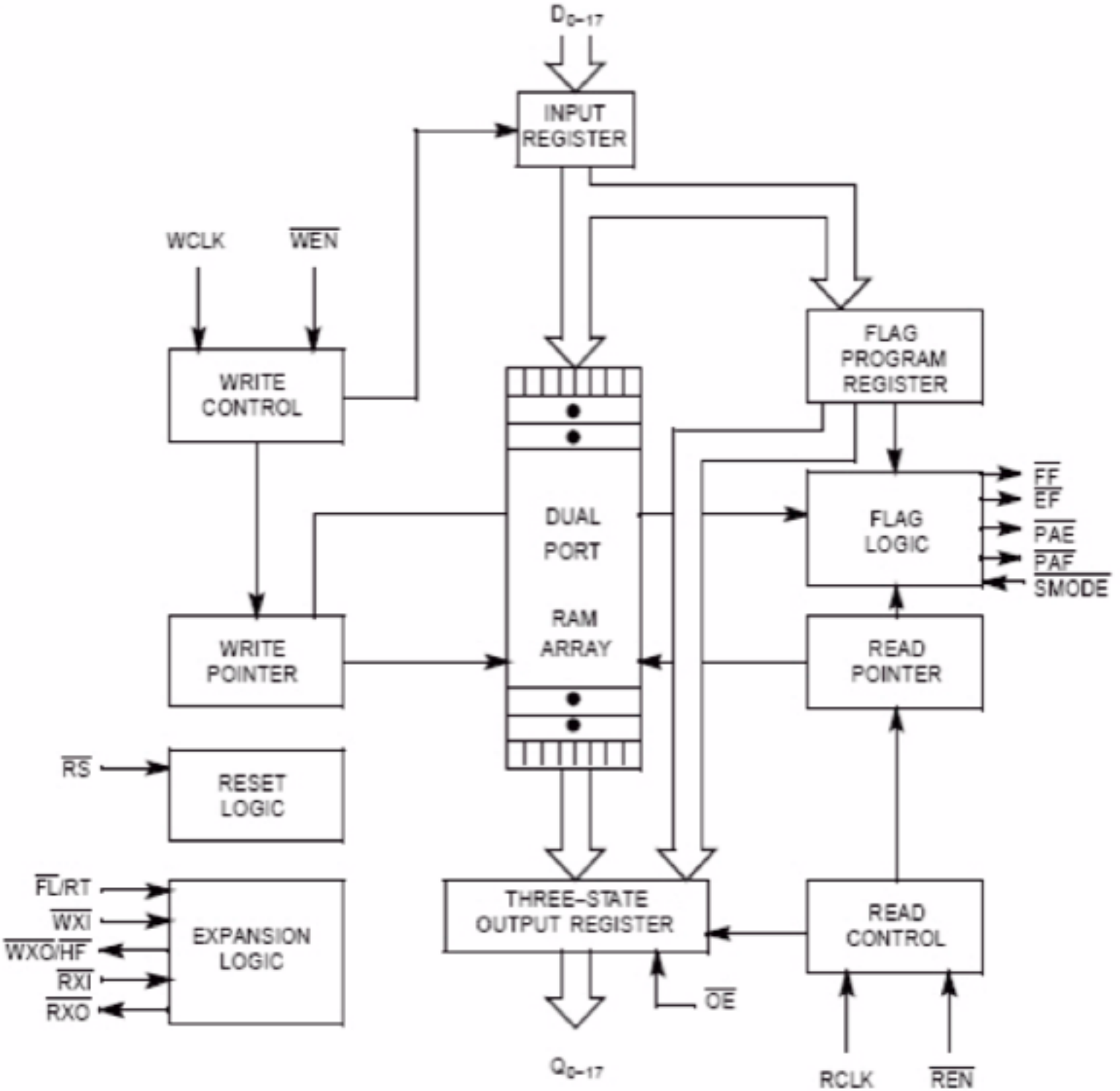
- Functionally equivalent to the homonymous well-known Abstract Data Type defined in CS
- Usually implemented resorting to dual port RAMs

FIFO memories

- Mostly used as interface buffers between producers and consumers with different throughputs



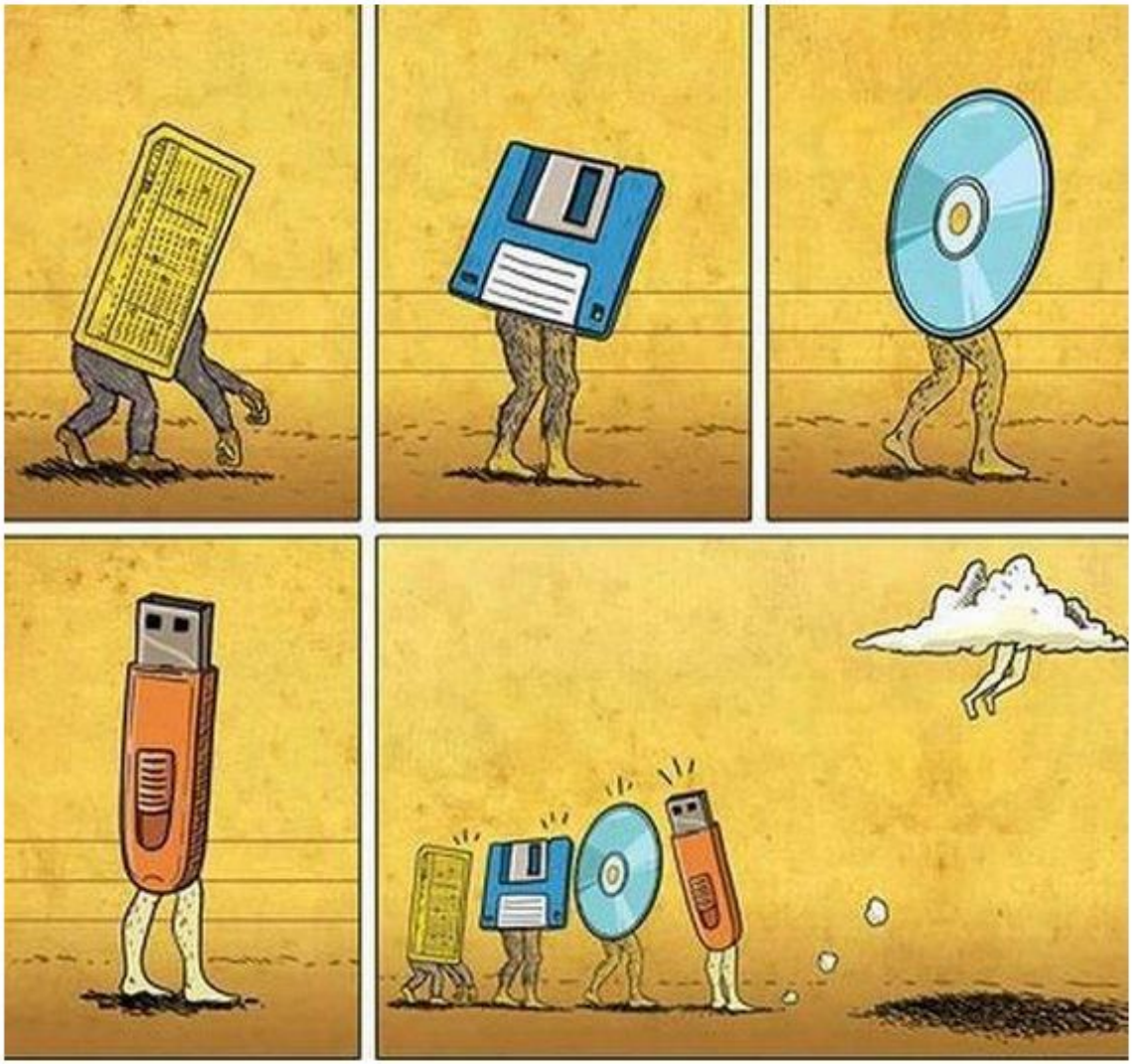
FIFO
Cypress
CY7C42x5



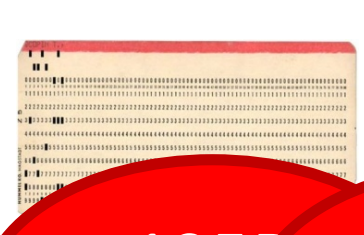
Outline

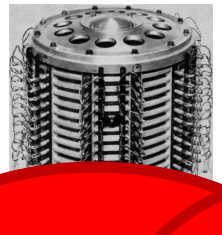
- Introduction
- Memory Organization
- Memory Characteristics
- Access mechanisms
- **Where we are coming from**
- External vs Internal Architectures
- Memory Hierarchy
- Memories with ECCs

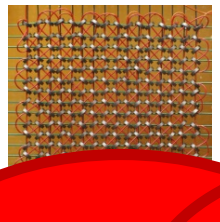
Data Storage Evolution



Memory Evolution







125[illegible]
2000[illegible]

62k[illegible]
2.5m[illegible]

156[illegible]
17µ[illegible]

1TB
1ms

1TB
~100µs

Punch Car[illegible]

Drum Me[illegible]

Magnetic

Core Mem[illegible]

H[illegible]

S[illegible]

D[illegible]

F[illegible]

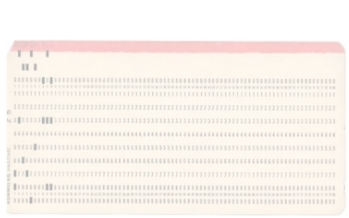
'Just' Technology Scaling

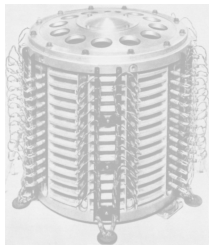
Emerging Memory Technologies

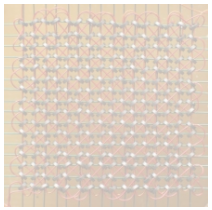
1837 1932 1960 1970 1980 1990 2000 2010

Rel. 31.03.2017

Memory Evolution







Punch Cards

Drum Memory

Magnetic Core Memory

Hard Drive

SRAM

DRAM

Flash Memory

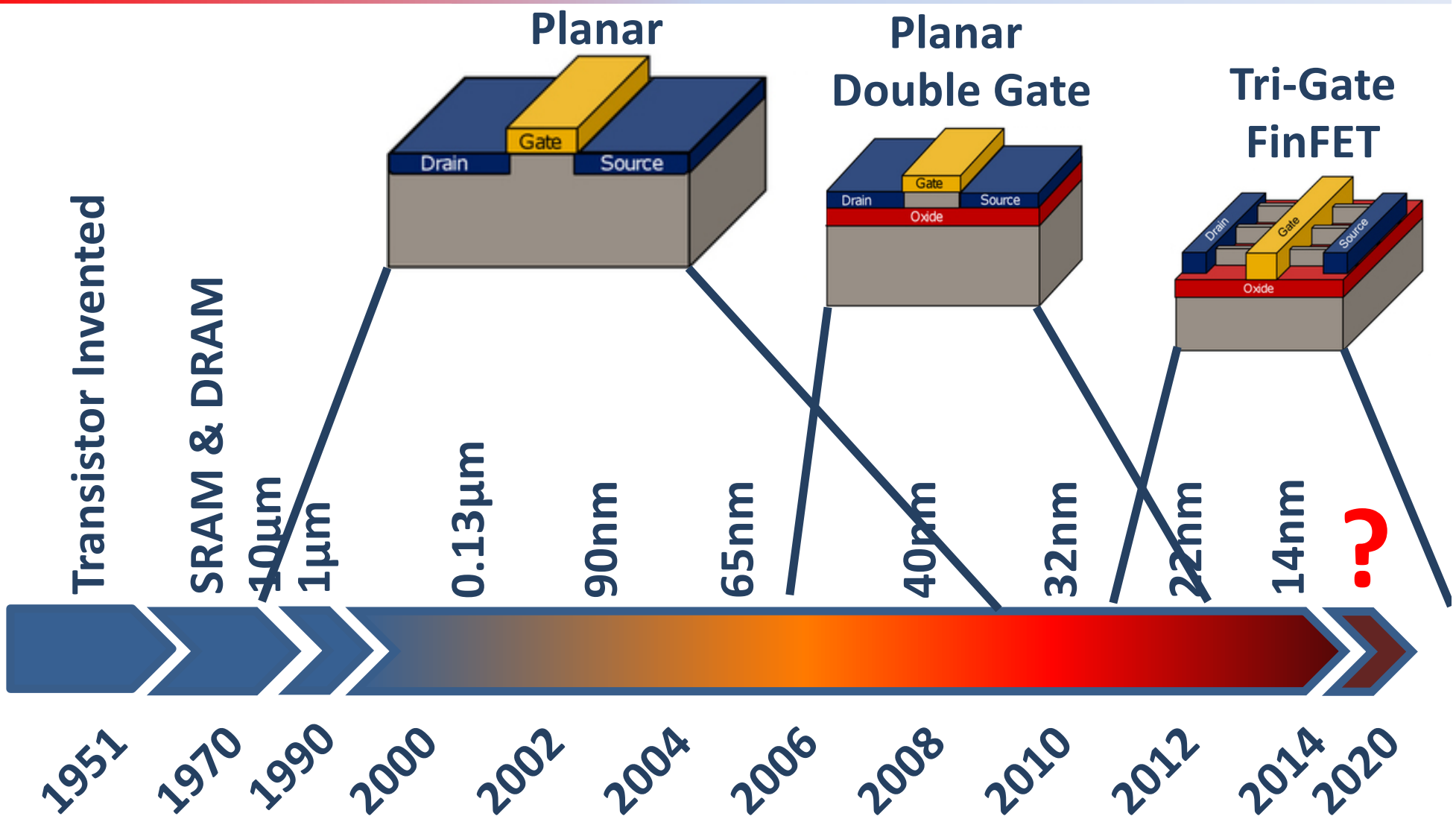
'Just' Technology Scaling

Emerging Memory Technologies

1837 1932 1960 1970 1980 1990 2000 2010

Rel. 31.03.2017

CMOS Technology Evolution

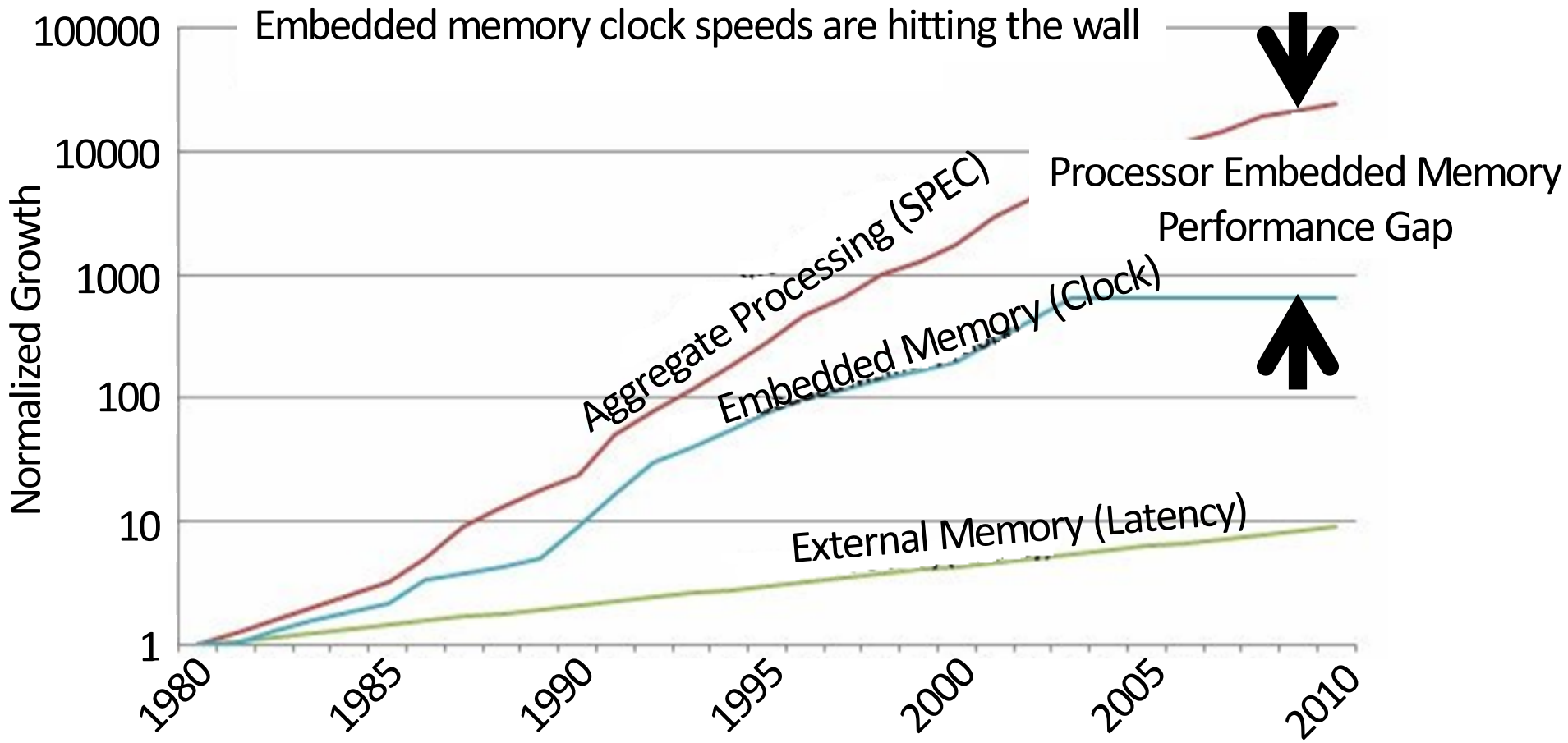


Rel. 31.03.2017

Technology Scaling Effects

ADVANTAGES	ISSUES
Higher integration	Higher variability
Higher speed	Decreased dependability
Lower capacitance	Increased leakage power
Lower switching energy	CPU vs. Memory

CPU vs. Memory



[Hennessy and Patterson, 5th Edition]

Outline

- Introduction
- Memory Organization
- Memory Characteristics
- Access mechanisms
- Where we are coming from
- **External vs Internal Architectures**
- Memory Hierarchy
- Memories with ECCs

External vs Internal Architectures

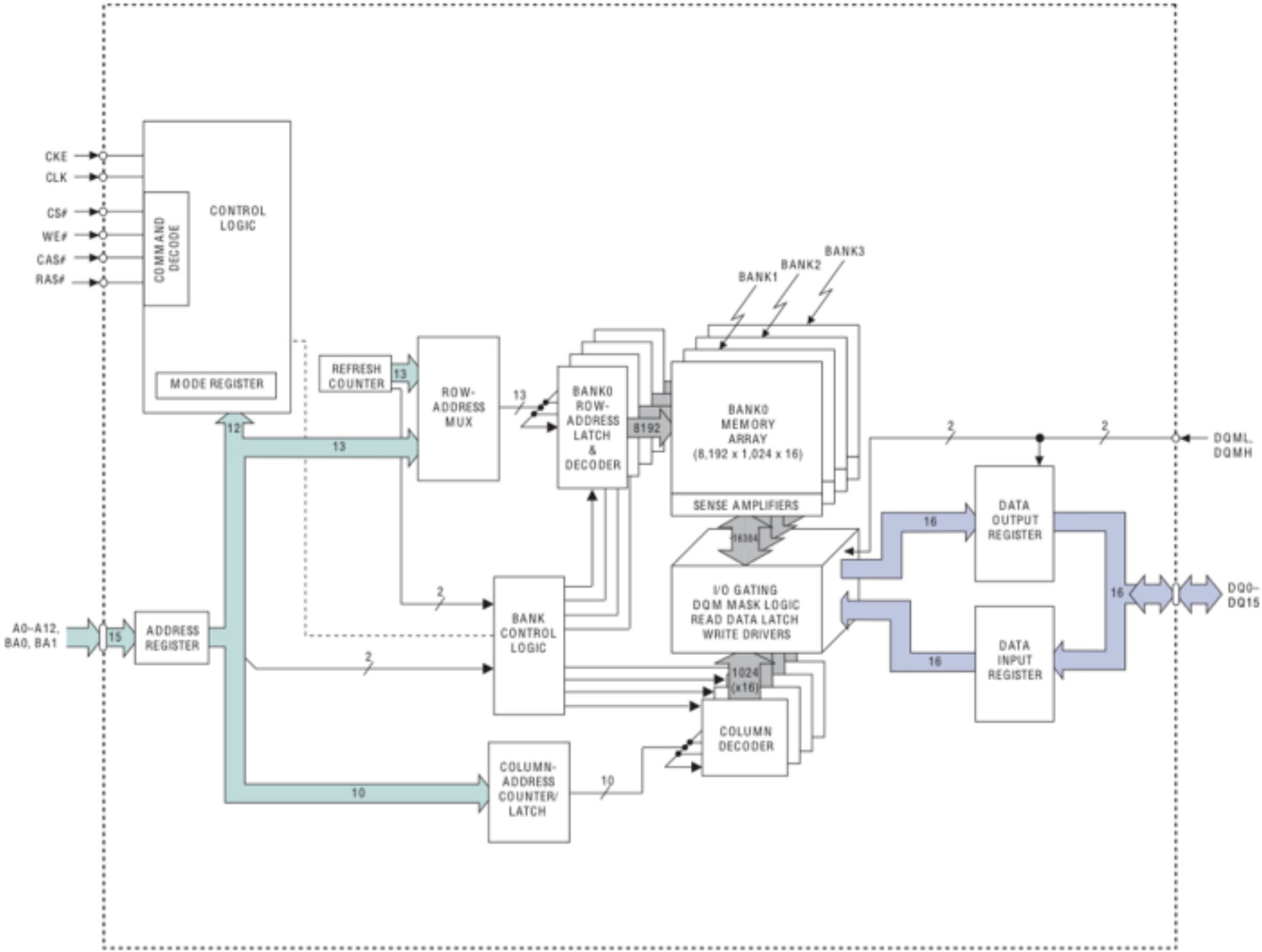
When dealing with a memory component, one must distinguish between

- ***External architecture*** : the way a user "sees" the components and can use it
- ***Internal architecture*** : how cells are actually organized within the component.

An example: the Micron MT48LC32M16A2 Memory

A Synchronous Memory of 512 Mb, accessible from a user as a memory of 32 M words of 16 bits, each.

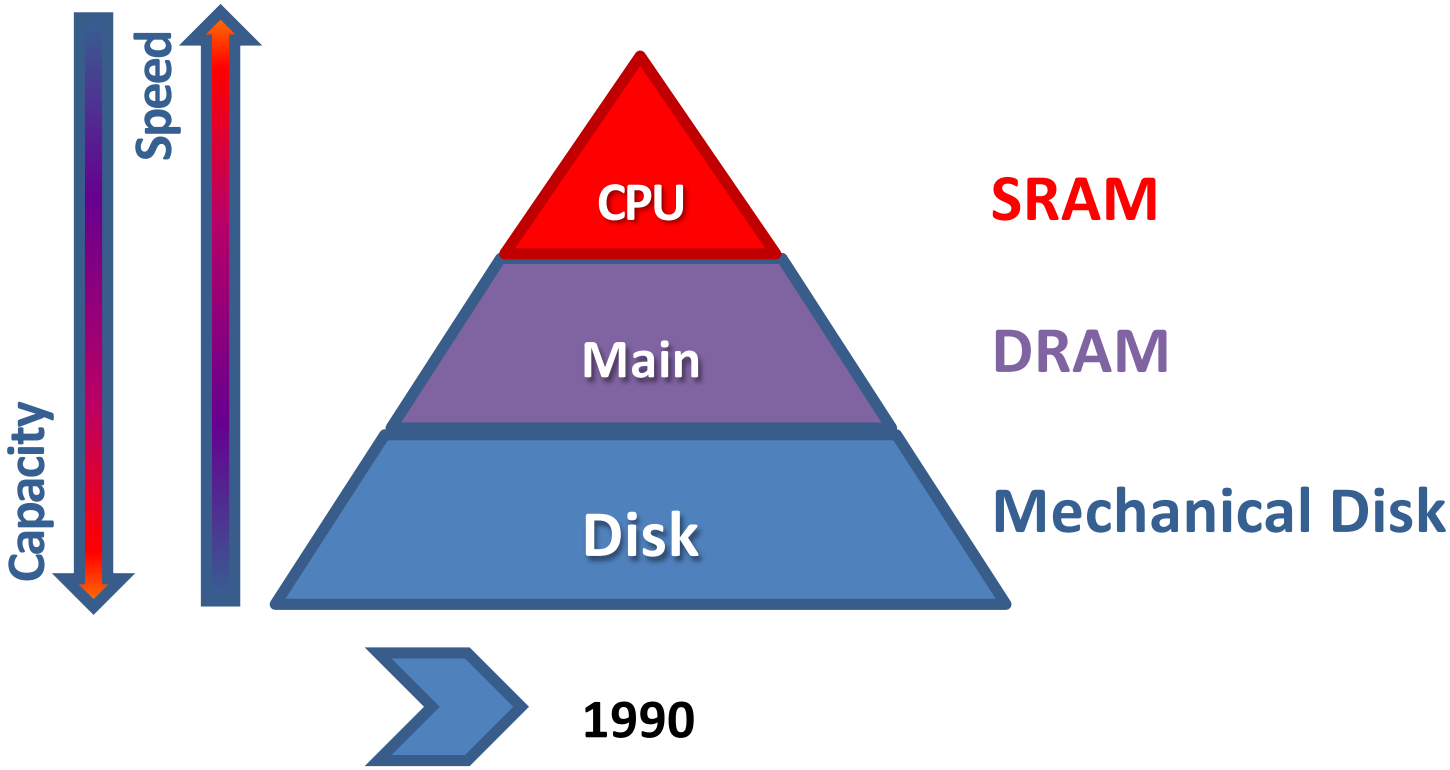
- *External architecture* :
  - Address bus of 15 bits to access each of the 32 M ($=2^{15}$) words
  - Data bus of 16 bit
  - Several control and status signals
- *Internal architecture* :
  - 4 banks (or matrices, or arrays), each composed of 8 k rows and 1 k columns, each column including 16 bits.


CKE
CLK
CS#
WE#
CAS#
RAS#
COMMAND DECODE
CONTROL LOGIC
MODE REGISTER
REFRESH COUNTER
13
ROW-ADDRESS MUX
13
BANK0 ROW-ADDRESS LATCH & DECODER
8192
BANK1
BANK2
BANK3
BANK0 MEMORY ARRAY (8,192 x 1,024 x 16)
SENSE AMPLIFIERS
16384
I/O GATING DQM MASK LOGIC READ DATA LATCH WRITE DRIVERS
1024 (x16)
COLUMN DECODER
12
13
2
BANK CONTROL LOGIC
2
A0–A12, BA0, BA1
15
ADDRESS REGISTER
10
COLUMN-ADDRESS COUNTER/ LATCH
10
2
2
DQML, DQMH
16
DATA OUTPUT REGISTER
16
DATA INPUT REGISTER
16
DQ0–DQ15

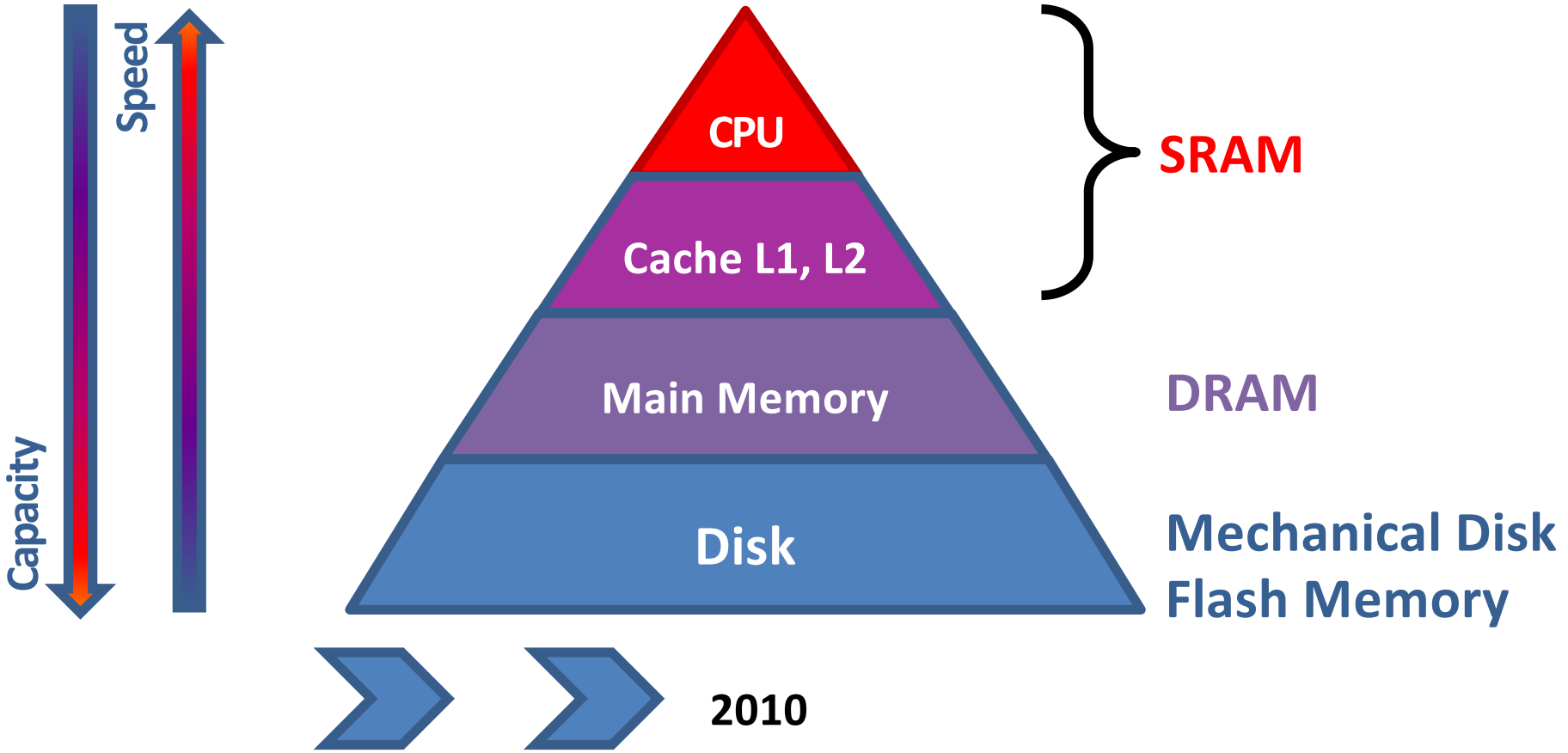
Outline

- Introduction
- Memory Organization
- Memory Characteristics
- Access mechanisms
- Where we are coming from
- External vs Internal Architectures
- **Memory Hierarchy**
- Memories with ECCs

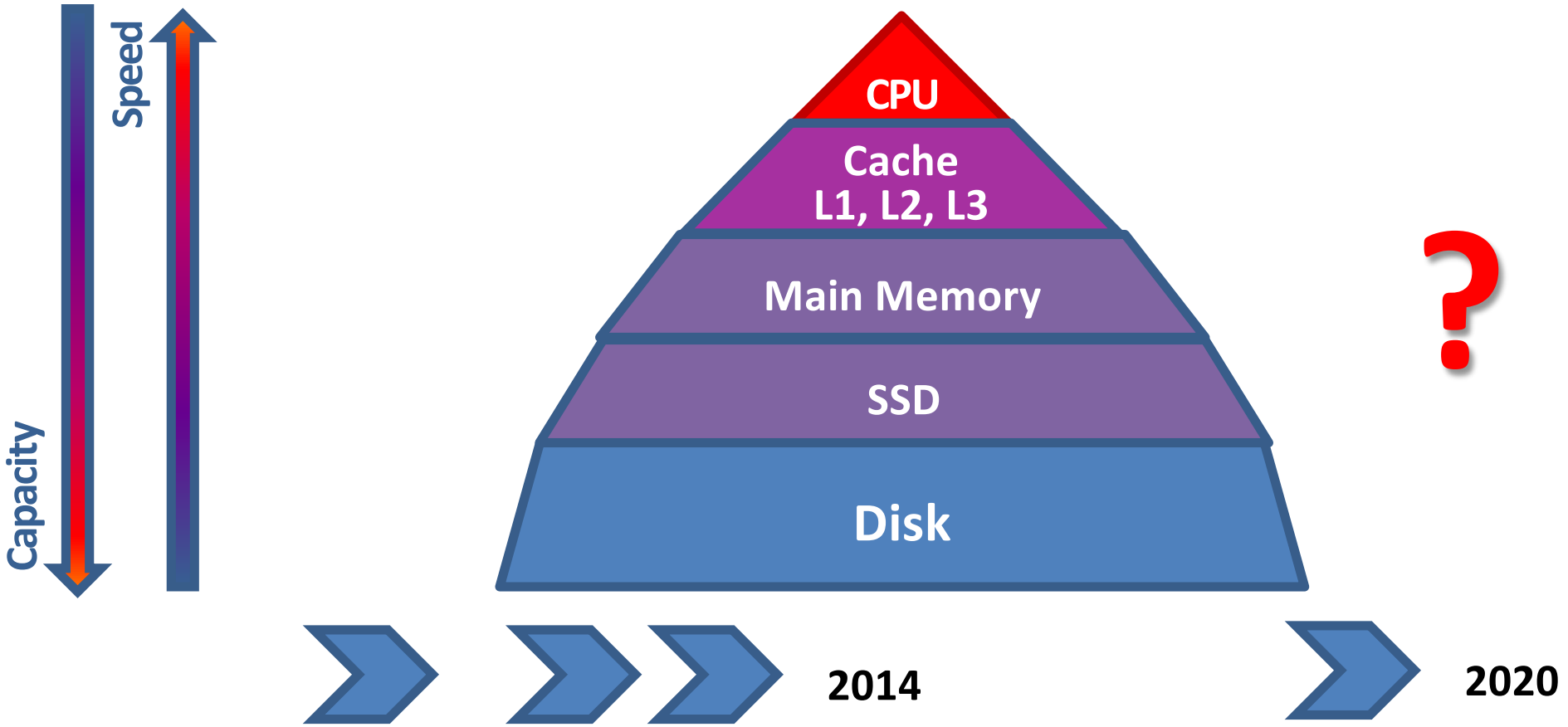
Memory Hierarchy



Memory Hierarchy

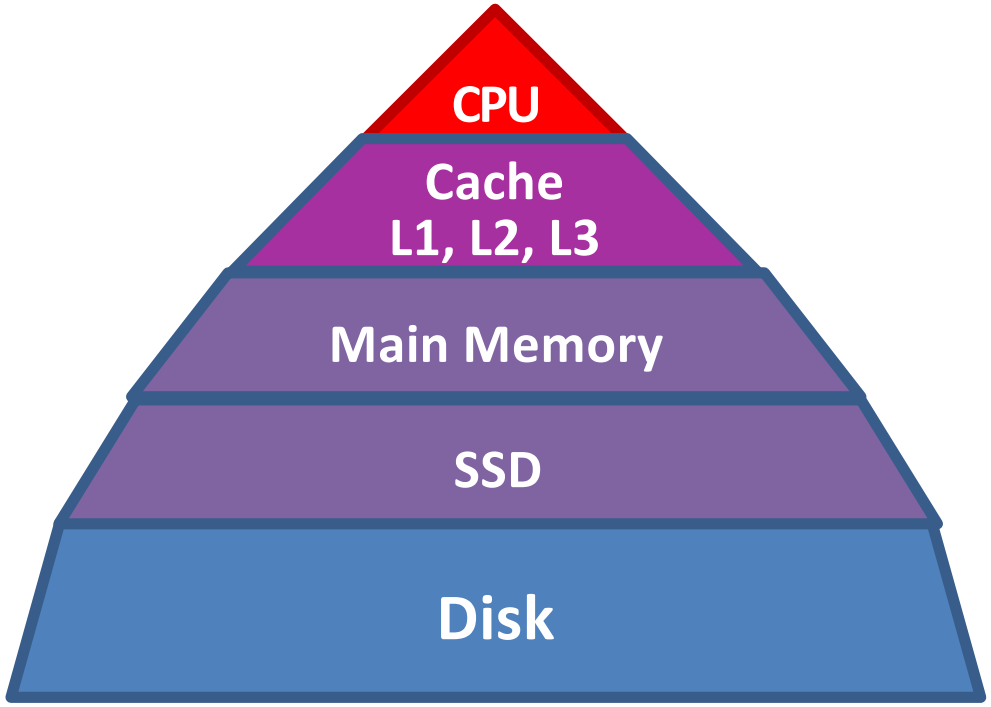


Memory Hierarchy

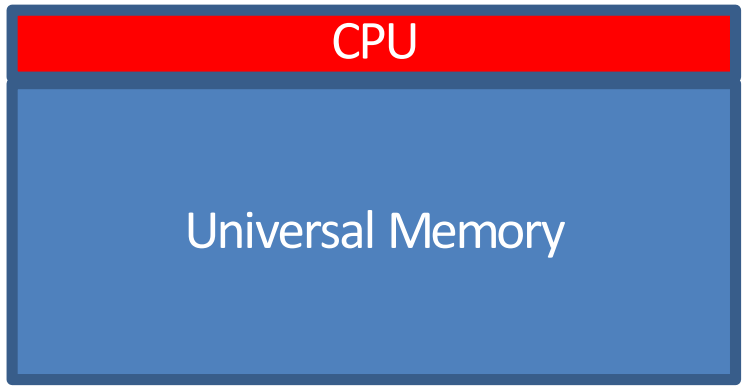


Rel. 31.03.2017

Where are we going
Architecture



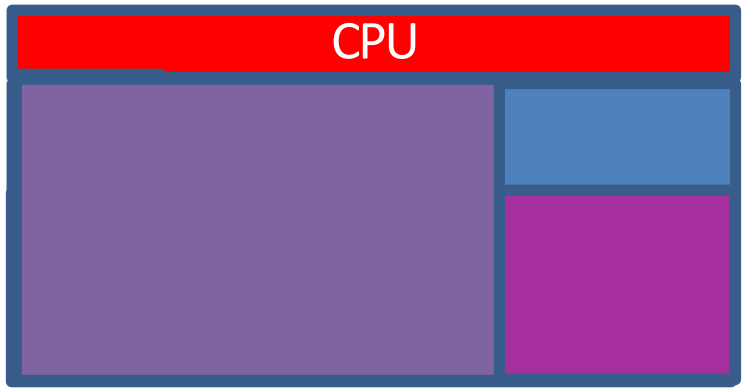
2014



2020

Where are we going
Architecture

- Reconfiguration:
  - Service-Oriented:
    - User Selectable
    - Application Driven
  - Aging-Dependent



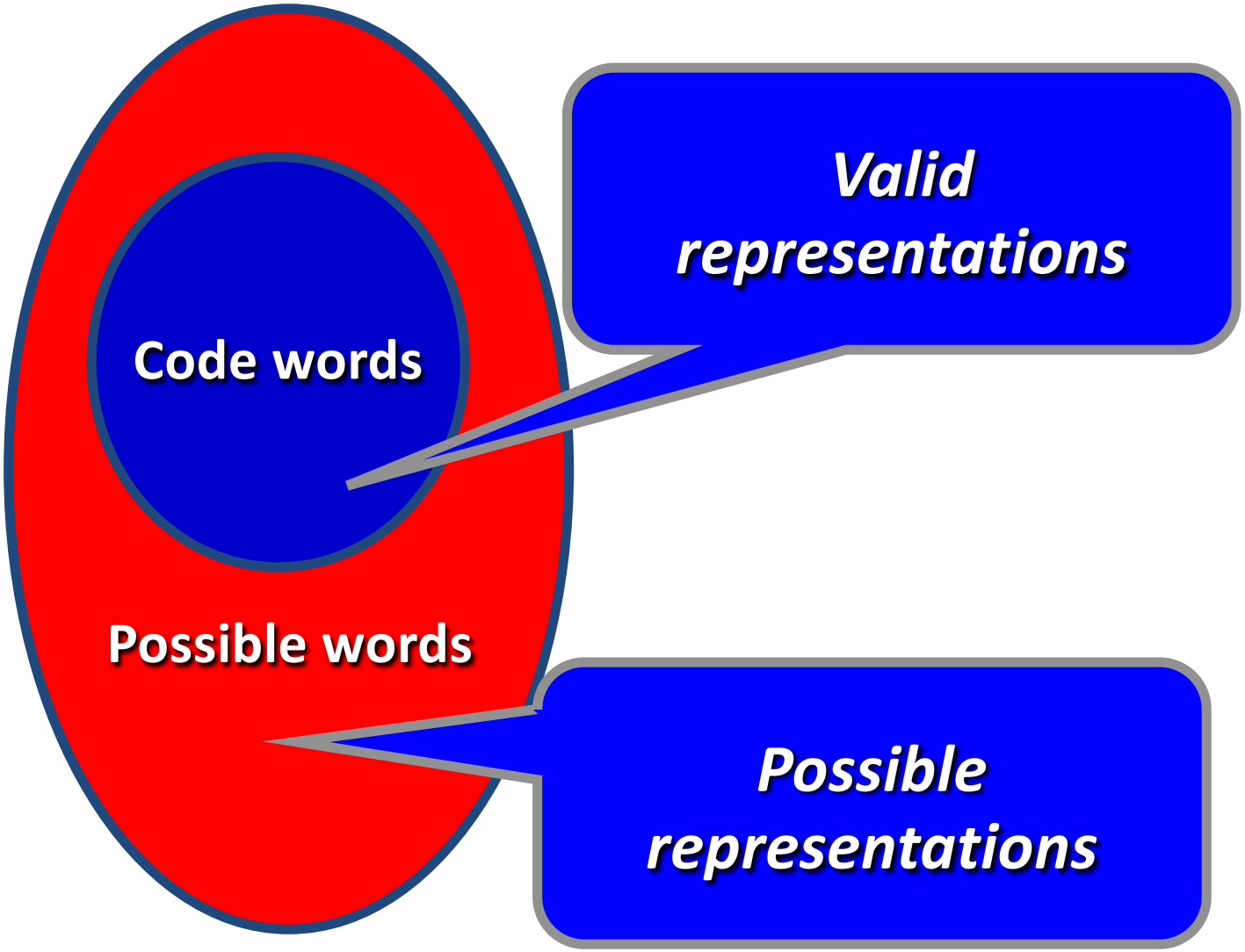
Outline

- Introduction
- Memory Organization
- Memory Characteristics
- Access mechanisms
- Where we are coming from
- External vs Internal Architectures
- Memory Hierarchy
- **Memories with ECCs**

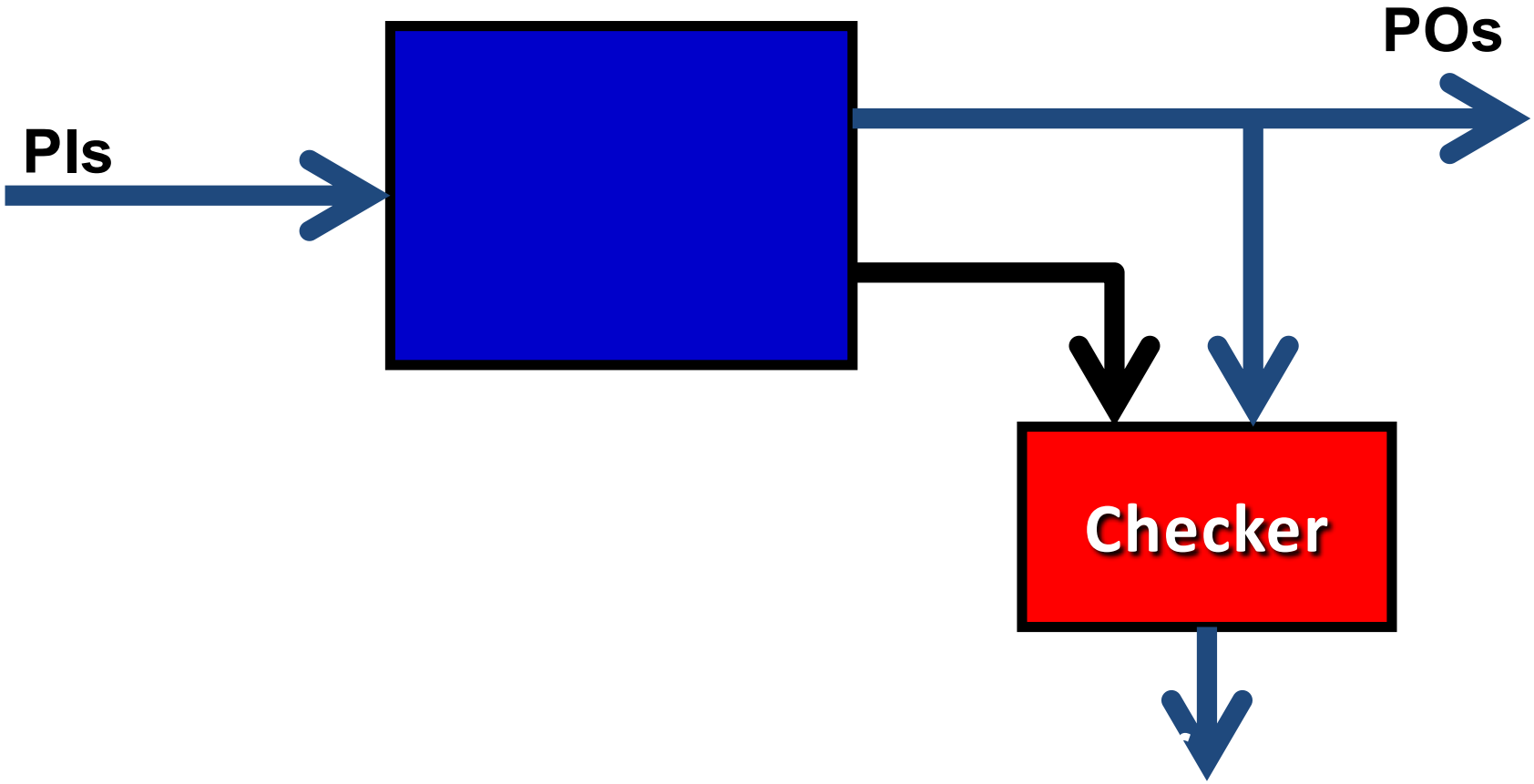
Memories with Error Correction Capability

- Memories are mostly equipped with EDACs (Error Detectors and Correctors), all exploiting ECCs (Error Correcting Codes)

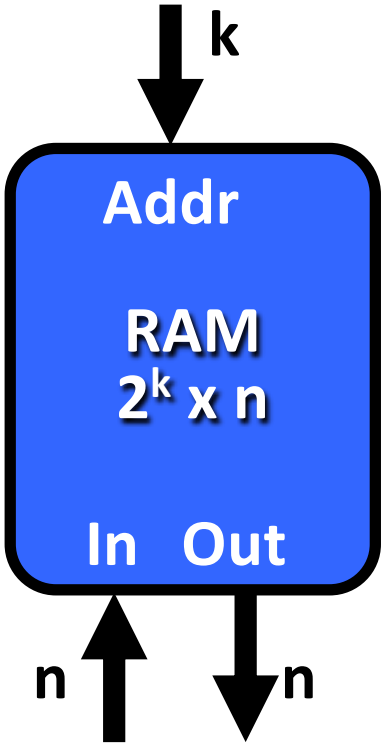
An example of information redundancy



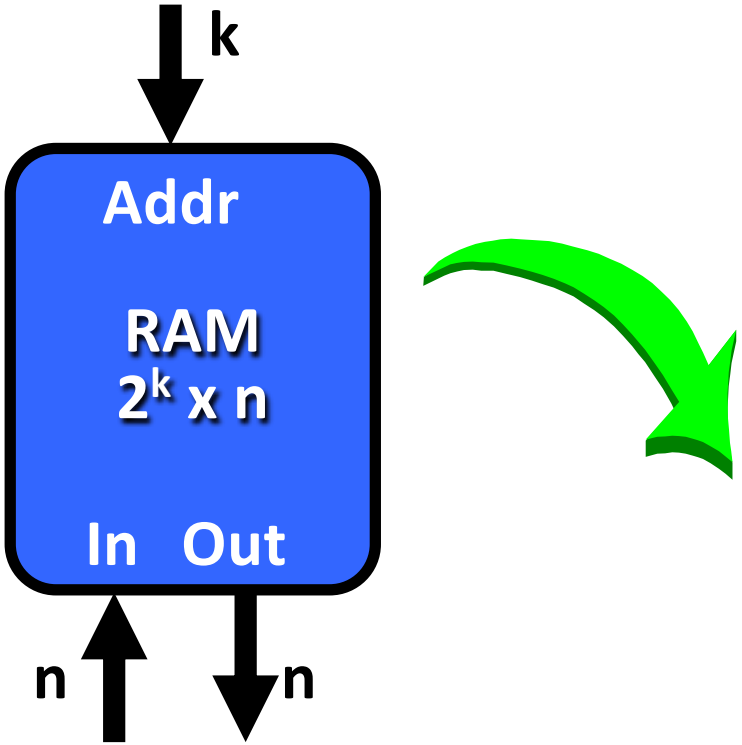
Typical structure

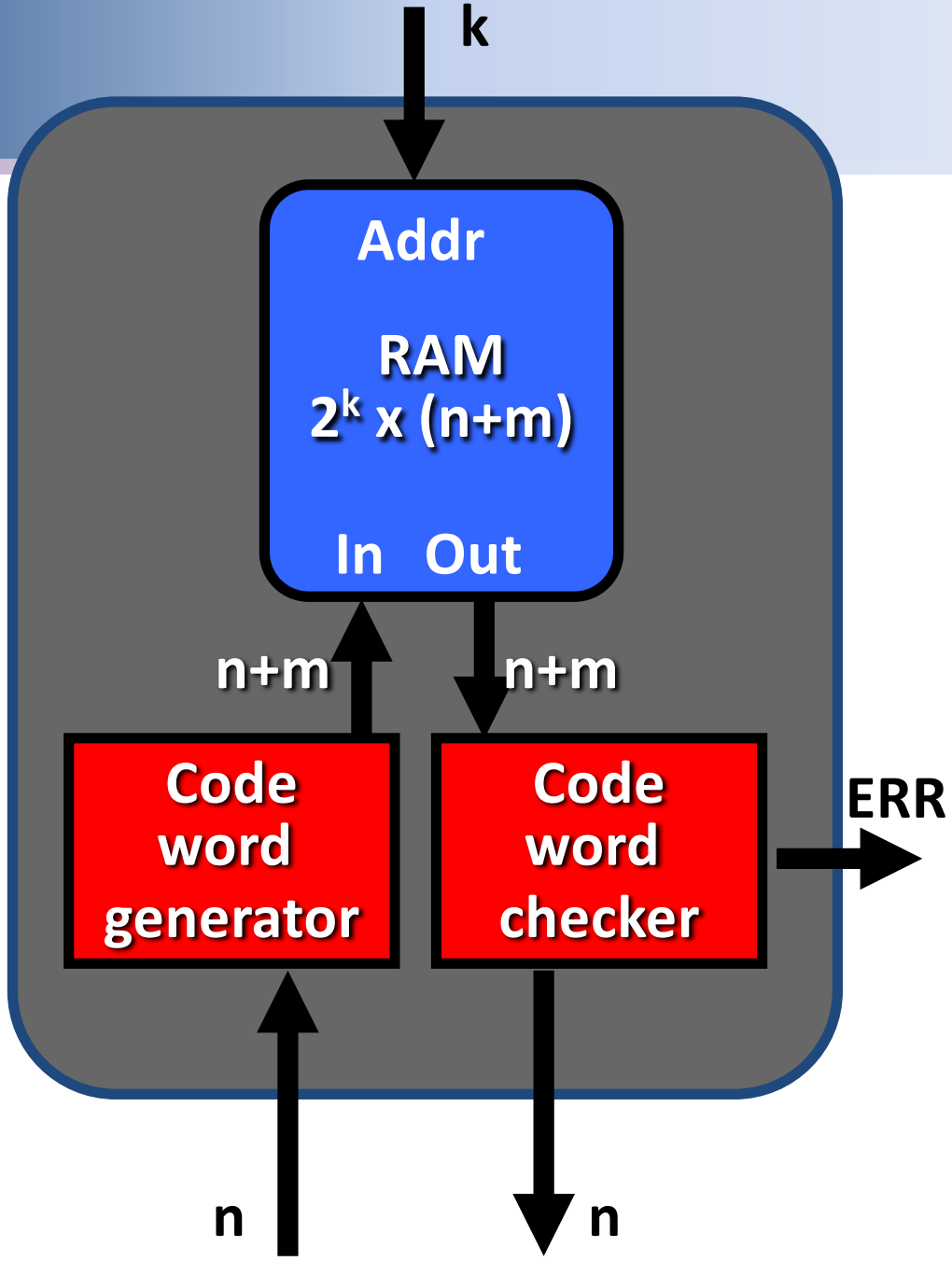


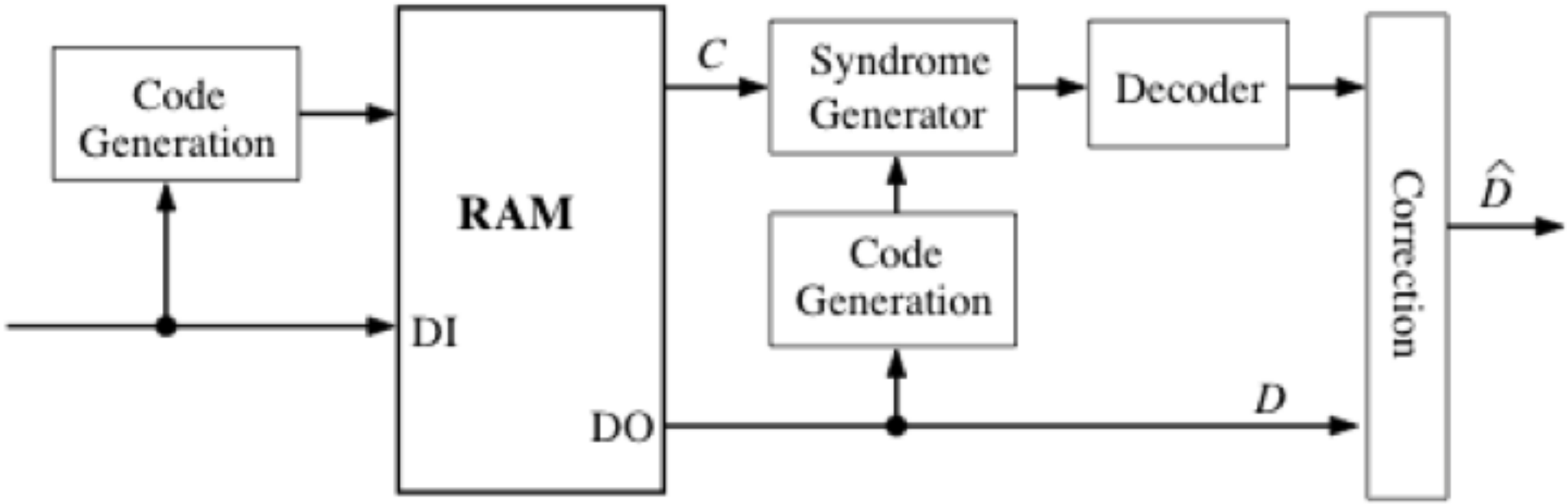
Example



Example







Codes

– Codes (parity, Hamming, Berger, Solomon, cyclic, arithmetic, …) are classified according to their ability to detect and correct errors affecting the code words:

- k-error detecting code: detect the presence of an error that modifies up to k bits
- k-error correcting code: correct an error that modifies up to k bits.

Hamming Codes

- Being N the word size, they require $\log_2 N$ bits of code and are able to:
  - Correct ALL Single Errors
  - Detect ALL Double Errors
- Thus, are often referred to as SEC-DED Code (Single Error Correction Double Error Detection).


Малые Автюхи, Калинковичский район, Республики Беларусь