Technische Hochschule Rosenheim
Technical University of Applied Sciences

Prof. Dr. Florian Künzner
Technical University of Applied Sciences Rosenheim, Computer Science

CA 8 – Memory 1

The lecture is based on the work and the documents of Prof. Dr. Theodor Tempelmeier

CAMPUS Rosenheim
Computer Science

Technische Hochschule Rosenheim
Technical University of Applied Sciences

Goal

CAMPUS Rosenheim
Computer Science

Technische Hochschule Rosenheim
Technical University of Applied Sciences

Goal

CA::Memory 1 - Hardware

- Memory types
- Memory chips
- Memory modules
- Modern memory modules

CAMPUS Rosenheim
Computer Science

Technische Hochschule Rosenheim
Technical University of Applied Sciences

Memory types

RAM vs ROM

RAM - Random access memory

- For **read and write** access
- Usage: **programs and data**
- It is (usually) a **volatile** memory (data are lost when power is switched off)
- **Very fast** access time
- **High power** consumption
- **Expensive**

ROM - Read only memory

- For **read only** memory access
- Usage: **firmware** (BIOS, UEFI)
- It is a **non-volatile** memory (remembers the data even if power is switched off)
- Usually **slower than RAM**
- **Low power** consumption
- **Cheaper** than RAM
- Example: EPROM, EEPROM

CAMPUS Rosenheim
Computer Science

Technische Hochschule Rosenheim
Technical University of Applied Sciences

Memory types

It's all about RAM!

CAMPUS Rosenheim
Computer Science

Technische Hochschule Rosenheim
Technical University of Applied Sciences

Memory types

SRAM vs DRAM

Property	SRAM - Static RAM	DRAM - Dynamic RAM
Construction	– Complex	+ Simple
Realisation of a bit	– 4..6 transistors	+ 1 transistor + 1 capacitor
Speed	+ Faster	– Slower
Size (capacity)	– Small	+ Large
Cost	– Expensive	+ Cheap
Used for	**Cache memory**	**Main memory**
Density	– Less dense	+ Highly dense
Charge leakage	+ Not present	– Present: refresh required
Power consumption	+ Low	– High

CAMPUS Rosenheim
Computer Science

Technische Hochschule Rosenheim
Technical University of Applied Sciences

Orders of magnitudes for bits and bytes

Bits:

Bits (decimal)			
Symbol	**Power**	**Num bits**	**Name**
1 kbit	10^3	1.000	kilobit
1 Mbit	10^6	1.000.000	megabit
1 Gbit	10^9	1.000.000.000	gigabit
1 Tbit	10^{12}	1.000.000.000.000	terabit
...			

Bits (**binary**)			
Symbol	**Power**	**Num bits**	**Name**
1 Kibit	2^{10}	1.024	kibibit
1 Mibit	2^{20}	1.048.576	mebibit
1 Gibit	2^{30}	1.073.741.824	gibibit
1 Tibit	2^{40}	1.099.511.627.776	tebibit
...			

Bytes:

Bytes (decimal)			
Symbol	**Power**	**Num bytes**	**Name**
1 kB	10^3	1.000	Kilobyte
1 MB	10^6	1.000.000	Megabyte
1 GB	10^9	1.000.000.000	Gigabyte
1 TB	10^{12}	1.000.000.000.000	Terabyte
...			

Bytes (**binary**)			
Symbol	**Power**	**Num bytes**	**Name**
1 KiB	2^{10}	1.024	Kibibyte
1 MiB	2^{20}	1.048.576	Mebibyte
1 GiB	2^{30}	1.073.741.824	Gibibyte
1 TiB	2^{40}	1.099.511.627.776	Tebibyte
...			

CAMPUS Rosenheim
Computer Science

Technische Hochschule Rosenheim
Technical University of Applied Sciences

Memory modules and chips - overview

[source: wikipedia.org]

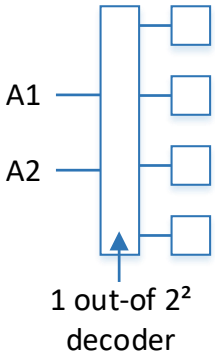
CAMPUS Rosenheim
Computer Science

Technische Hochschule Rosenheim
Technical University of Applied Sciences

Memory chips

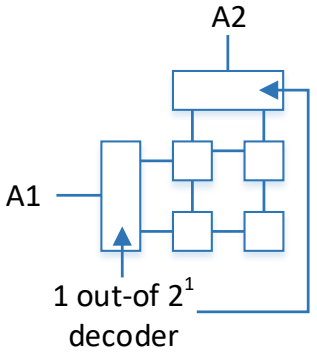
Arrangement of memory cells: (within a memory chip)

Linear Arrangement:



- To address 1-out-of-2^n memory cells, n address lanes are required.
- Problem: Address lanes are expensive (takes place on the chip)

Matrix Arrangement:



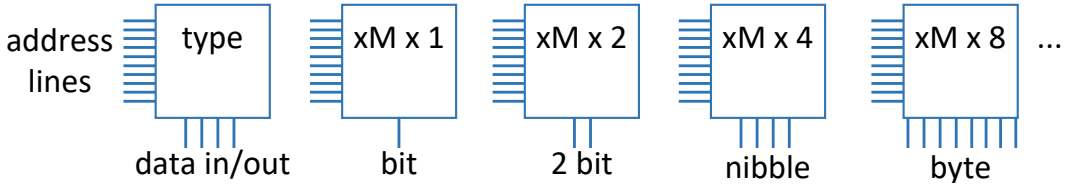
- To address 1-out-of-2^n memory cells, only $n/2$ address lanes are required.
- The address is usually transferred in two steps:
  - 1: Row address
  - 2: Column address
- Only half the address lanes are required

CAMPUS Rosenheim
Computer Science

Technische Hochschule Rosenheim
Technical University of Applied Sciences

Memory chips

Chip types:



Terminology:

Description	Type	Unit
Chip with x mega that provides 1 bit per address	xM x 1	bit
Chip with x mega that provides 2 bit per address	xM x 2	2 bit
Chip with x mega that provides 4 bit per address	xM x 4	nibble
Chip with x mega that provides 8 bit per address	xM x 8	byte

Memory chips

Chip capacity:

On these chips: K means Ki, M means Mi, G means Gi, ...
xK/xM/xG denotes the number of chip cell rows inside the chip

Chip capacity = xM x number of pins per chip

Examples:

- `16M x 1: 16Mi x 1 = 16 Mibit => 16Mi/8 =   2 MiB`
- `16M x 2: 16Mi x 2 = 32 Mibit => 32Mi/8 =   4 MiB`
- `1G x 4:  1Gi x 4 =  4 Gibit => 4Gi/8  = 512 MiB`
- `1G x 8:  1Gi x 8 =  8 Gibit => 8Gi/8  =   1 GiB`

Memory modules

The memory chips on a memory module are usually arranged in a matrix layout.

[image source: samsung.com]

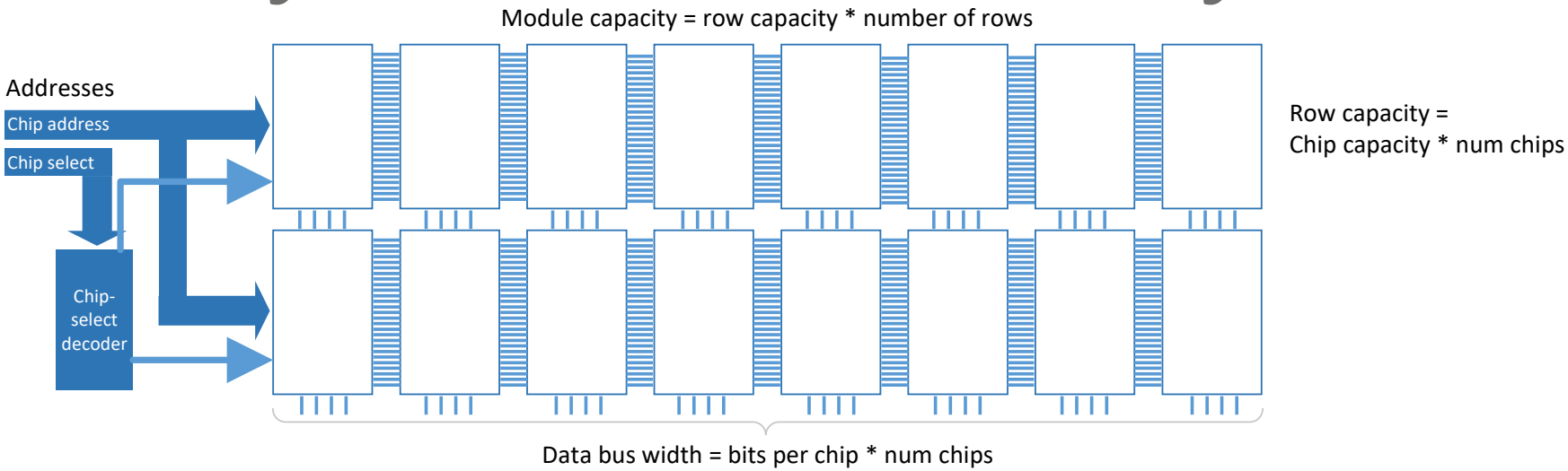
Memory address is divided into:

- Chip select address: for row (rank)
- Chip address (inside the chip)

CAMPUS Rosenheim
Computer Science

Technische Hochschule Rosenheim
Technical University of Applied Sciences

Memory modules - hardware layout



Address calculations:

Nr.	Descriptions	Calc	Results
(1)	Number of bits to address the module capacity	$log_2(module\ capacity)$	number of bits
(2)	Number of address lanes/bits for chip select	$log_2(num.\ rows)$	number of bits/address lanes
(3)	Number of address lanes/bits for chip address	$log_2(num.\ chip\ cell\ rows)$	number of bits/address lanes
(4)	Number of bits to address the bytes inside the word	$log_2(num.\ bytes\ per\ word)$	number of bits

Address calculation relationship:

Number of bits to address the module capacity: $(1) = \sum_{i=2}^{4}(i) = (2)+(3)+(4)$

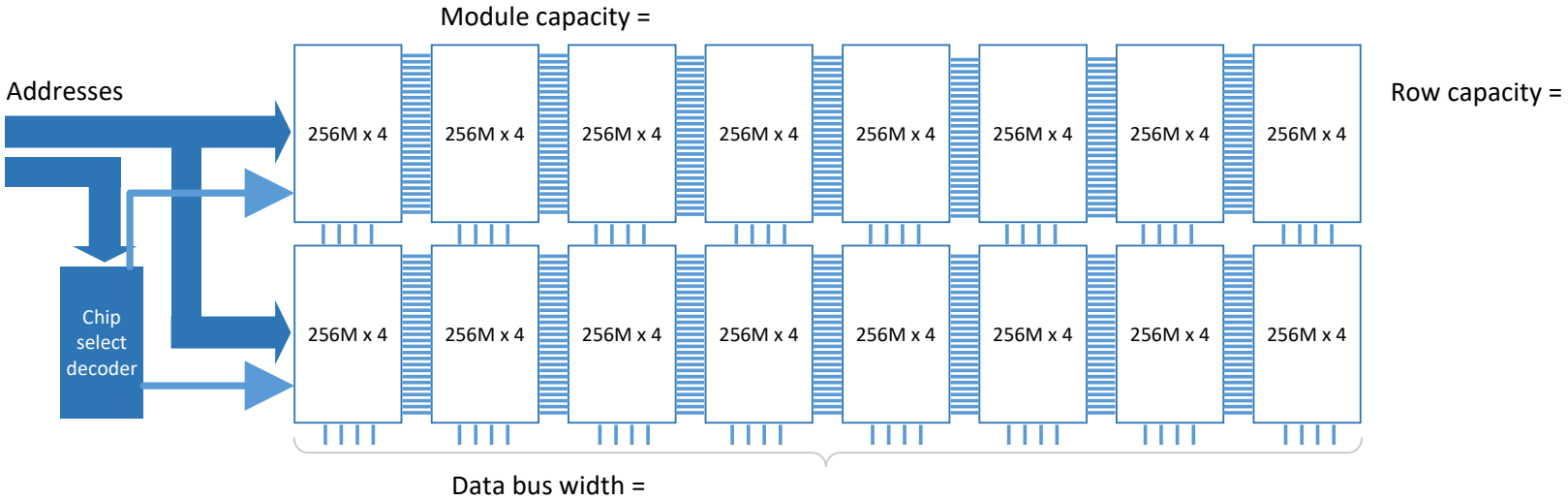
CAMPUS Rosenheim
Computer Science

Technische Hochschule Rosenheim
Technical University of Applied Sciences

Memory modules - example

Example:

- Design a memory module with: 2 GiB capacity and a 32 bit data bus width
- Use chips of type 256M x 4



Address calculations:

Nr.	Descriptions	Calc	Result
(1)	Number of bits to address the module capacity		
(2)	Number of address lanes/bits for chip select		
(3)	Number of address lanes/bits for chip address		
(4)	Number of bits to address the bytes inside the word		

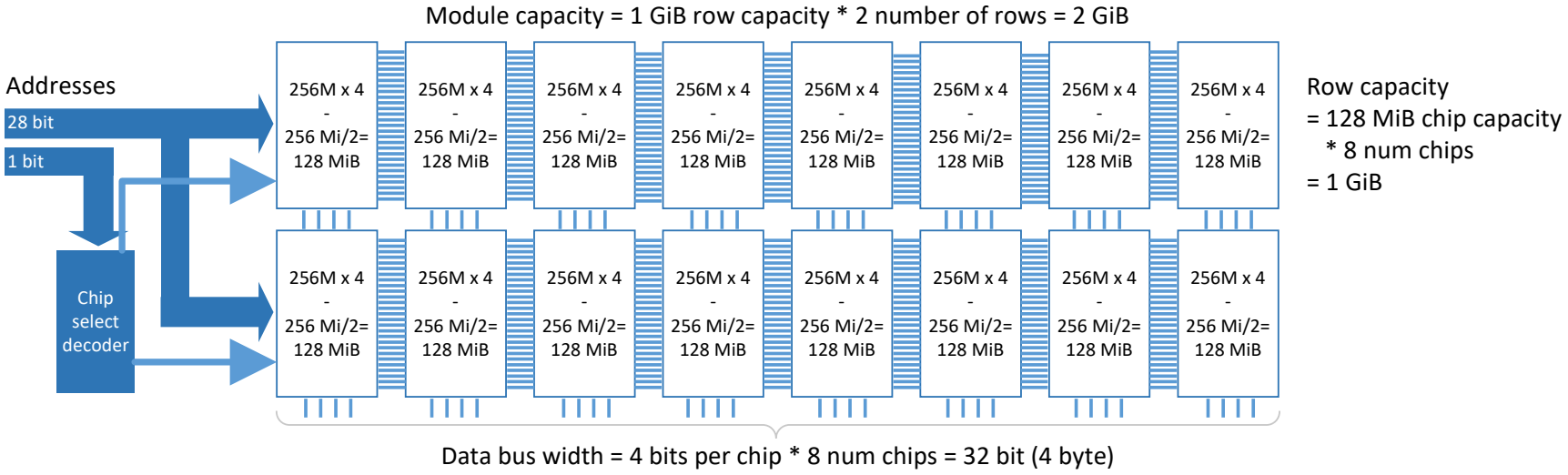
CAMPUS Rosenheim
Computer Science

Technische Hochschule Rosenheim
Technical University of Applied Sciences

Memory modules - example (solution)

Example:

- Design a memory module with: 2 GiB capacity and a 32 bit data bus width
- Use chips of type 256M x 4



Address calculations:

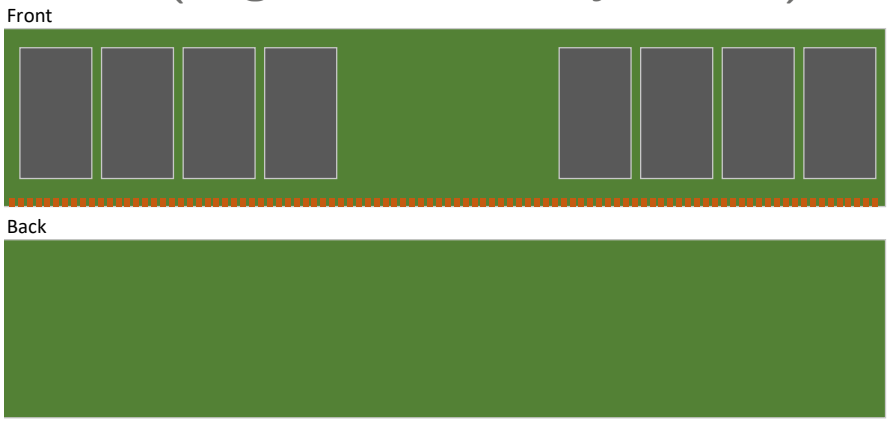
Nr.	Descriptions	Calc	Result
(1)	Number of bits to address the module capacity	$log_2(2\ GiB = 2^{31})$	31 bits
(2)	Number of address lanes/bits for chip select	$log_2(2)$	1 lanes/bits
(3)	Number of address lanes/bits for chip address	$log_2(256\ Mi = 2^{28})$	28 lanes/bits
(4)	Number of bits to address the bytes inside the word	$log_2(4bytes)$	2 bits

CAMPUS Rosenheim
Computer Science

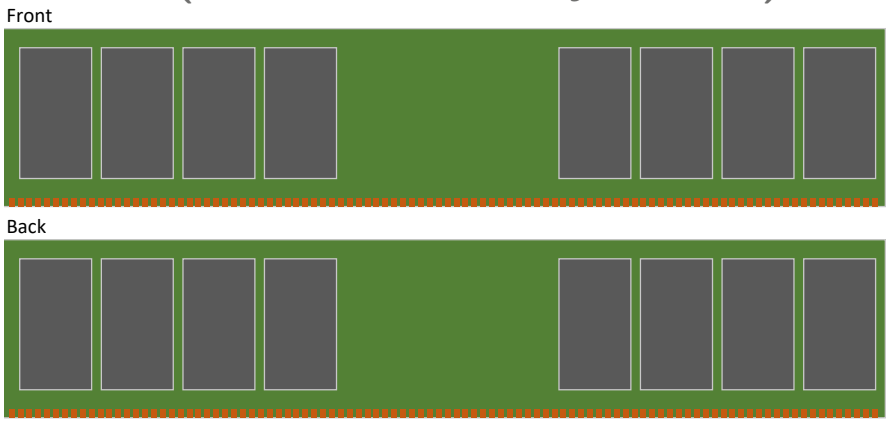
Technische Hochschule Rosenheim
Technical University of Applied Sciences

Memory modules - formats

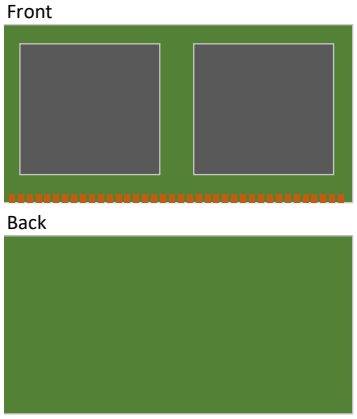
SIMM (single inline memory module):



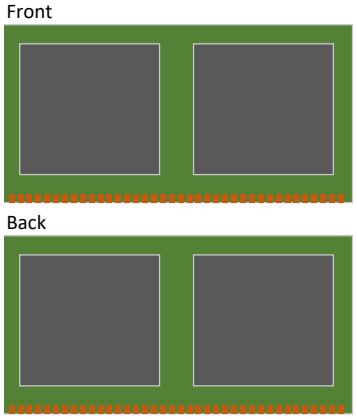
DIMM (dual inline memory module):



SO-SIMM (small outline SIMM):



SO-DIMM (small outline DIMM):



CAMPUS Rosenheim
Computer Science

Technische Hochschule Rosenheim
Technical University of Applied Sciences

Memory modules - interleaving

Problem:

After a **memory cell** is read in a DRAM, the cell **needs to be refreshed** and this takes some time.

Idea:

Distribute consecutive addresses evenly across the chip rows.

- **Reduces** the problem of **waiting** until the refresh is complete
- **Accelerates memory access** in an effect similar to pipelining
- But due to the increased capacities of the individual chips, a memory module has only one or two chip rows. -> Solution: **SDRAM**

CAMPUS Rosenheim
Computer Science

Technische Hochschule Rosenheim
Technical University of Applied Sciences

Modern memory modules

Overview of various terms in the memory area

SDRAM

SDRAM: synchronous DRAM

- **Synchronous** means there as a **clock** pulse
- **Dynamic** means there is a **refresh necessary**
- **Memory is divided** into several equally sized and independent **banks**: allows **interleaving** within chips
- Chips can accept new commands before finishing the previous one (for another bank).

ECC

ECC: error checking and correction

- ECC memory can **detect and correct** the most common kinds of internal **data corruption**
- Allows the **detection** and **correction** of **single bit errors**
- Some do also **detect double bit errors**
- Application area: Scientific and financial computing applications which operate on sensitive data

DDR-SDRAM

DDR-SDRAM: double data rate SDRAM

Idea:

- Transfers data at almost double the transfer rate
- **Data is transferred on rising and falling edges**
- DDR4-RAM is still available
- DDR5-RAM is now state of the art for new computers

For desktop/notebook systems:

	DDR3	DDR4	DDR5
Data transfer rate	17 GiB/s	25,6 GiB/s	51,2 GiB/s
Max module capacity	16 GiB	64 GiB	128 GiB

DDR RAM history (generations): `https://en.wikipedia.org/wiki/DDR_SDRAM#Generations`

CAMPUS Rosenheim
Computer Science

Technische Hochschule Rosenheim
Technical University of Applied Sciences

DDR-SDRAM labels

Example:

DDR5-6000 (PC5-48000) CL36-36-36-96 (links: shop, producer)

Details:

Standard	Speed (MT/s)	Bandwidth	CL	RCD	RP	RAS
DDR5	6000 MT/s	48000 MB/s	36	36	36	96

Symbols:

- CL: CAS (column address strobe) latency: number of clock cycles between read command and the moment data is available
- RCD: RAS-to-CAS (row-to-column) delay: delay in cycles after activating a row/bank
- RP: RAS precharge: number of cycles between activating/deactivating a row/bank
- RAS :Active to precharge delay: number of cycles between two accesses

Relationship of speed an bandwidth:

$$\texttt{bandwidth} = \texttt{speed (MT/s)} \times \texttt{bus width (bytes)}$$

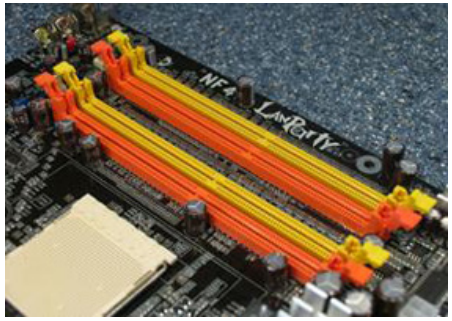
t_a: Access time t_a until the first word is available:

$$t_a \geq \frac{2 \times (\texttt{CL} + \texttt{RCD})}{\texttt{speed (MT/s)}}$$

CAMPUS Rosenheim
Computer Science

Technische Hochschule Rosenheim
Technical University of Applied Sciences

Multi-channel memory architecture



[source: wikipedia.com]

Idea:

- Adds **multiple channels** from the memory to the controller
- **Increase data transfer rate** of DRAM memory modules with the memory controller
- Dual/Triple/Quad-channels are possible
- **Dual-channel**: **theoretically doubles** the data transfer rate
- New Intel processors (like Intel Core i7-9800X) supports quad-channel memory architecture

CAMPUS Rosenheim
Computer Science

Technische Hochschule Rosenheim
Technical University of Applied Sciences

Summary and outlook

Summary

- Memory types
- Memory chips
- Memory modules
- Modern memory modules

Outlook

- MMU
- Virtual memory