Computer Structure

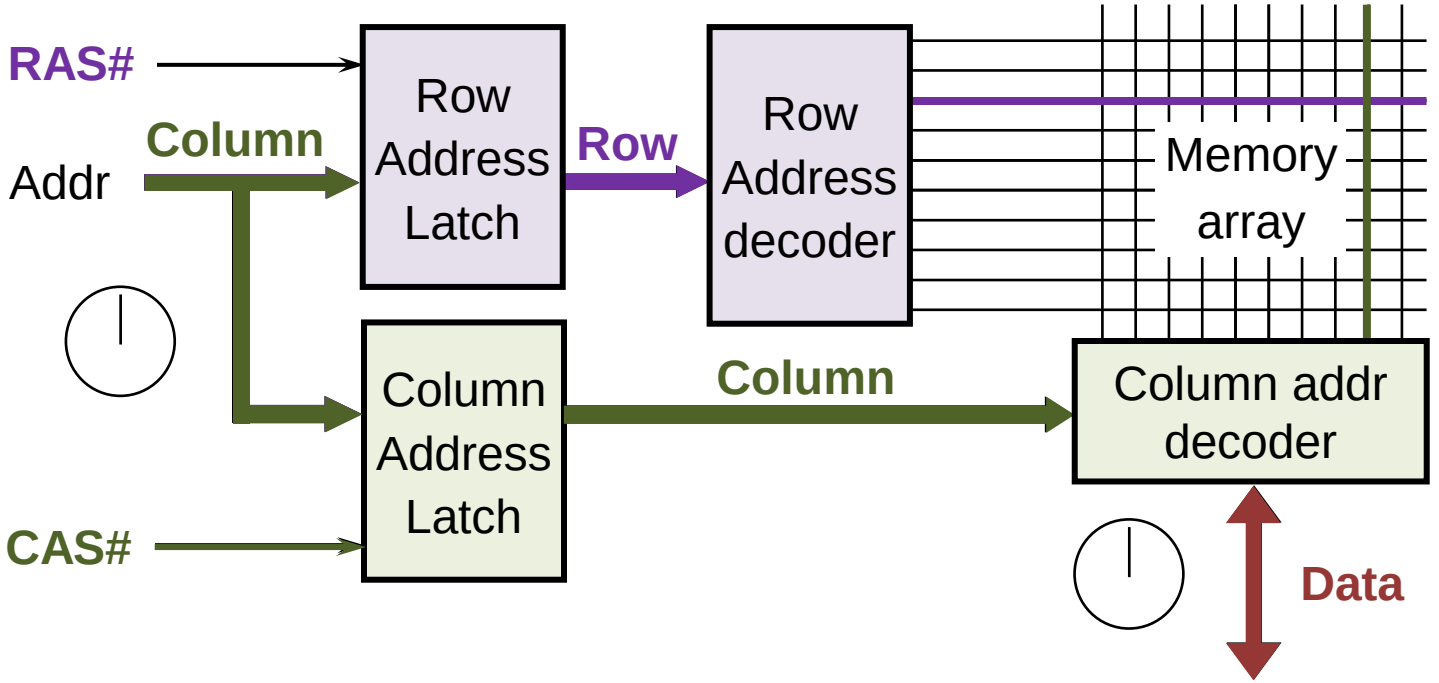
DRAM

Lihu Rappoport And Adi Yoaz

DRAM – Dynamic Random Access Memory

- **Radom Access Memory**
  - The data access time is independent of the access sequence
  - E.g., unlike a tape (serial access) or disk (direct access)
  - We will see that DRAM is not 100% a RAM
- **Dynamic – requires periodic refresh**
  - A DRAM cell consists of transistor + capacitor
    - The capacitor keeps the state, the transistor guards access to the state
    - Charging and draining a capacitor takes time
    - Leakage current from the transistor drains the capacitor even when transistor is closed
    - The DRAM cell must be periodically refreshed (every 64ms)
  - SRAM (static RAM) is faster and does not require refersh
    - But each cell is comprised of a few transistors – it is bigger and requires higher power
    - For main memory, memory capacity and low power are the prime factors, making DRAM the right choice

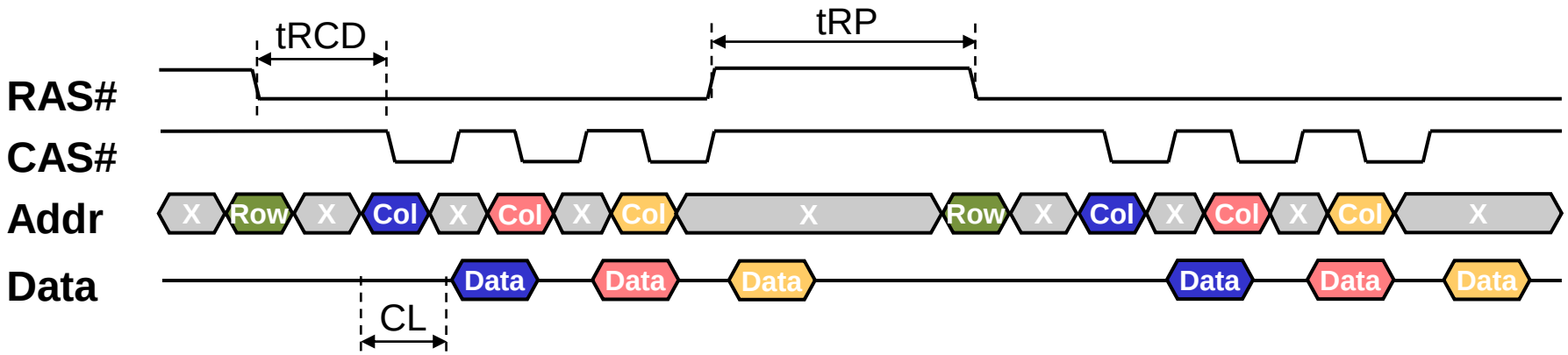
Basic DRAM chip



- **DRAM access sequence**
  - Put Row on addr. bus
  - Assert RAS# (Row Addr. Strobe) to latch Row
  - Put Column on addr. bus
  - Wait RAS# to CAS# delay and assert CAS# (Column Addr. Strobe) to latch Col
  - Get data on address bus after CL (CAS latency)

Page Mode DRAM

- **Allows Multiple accesses to different columns within the same row**
  - Saves RAS, RAS to CAS delay, and Row pre-charge



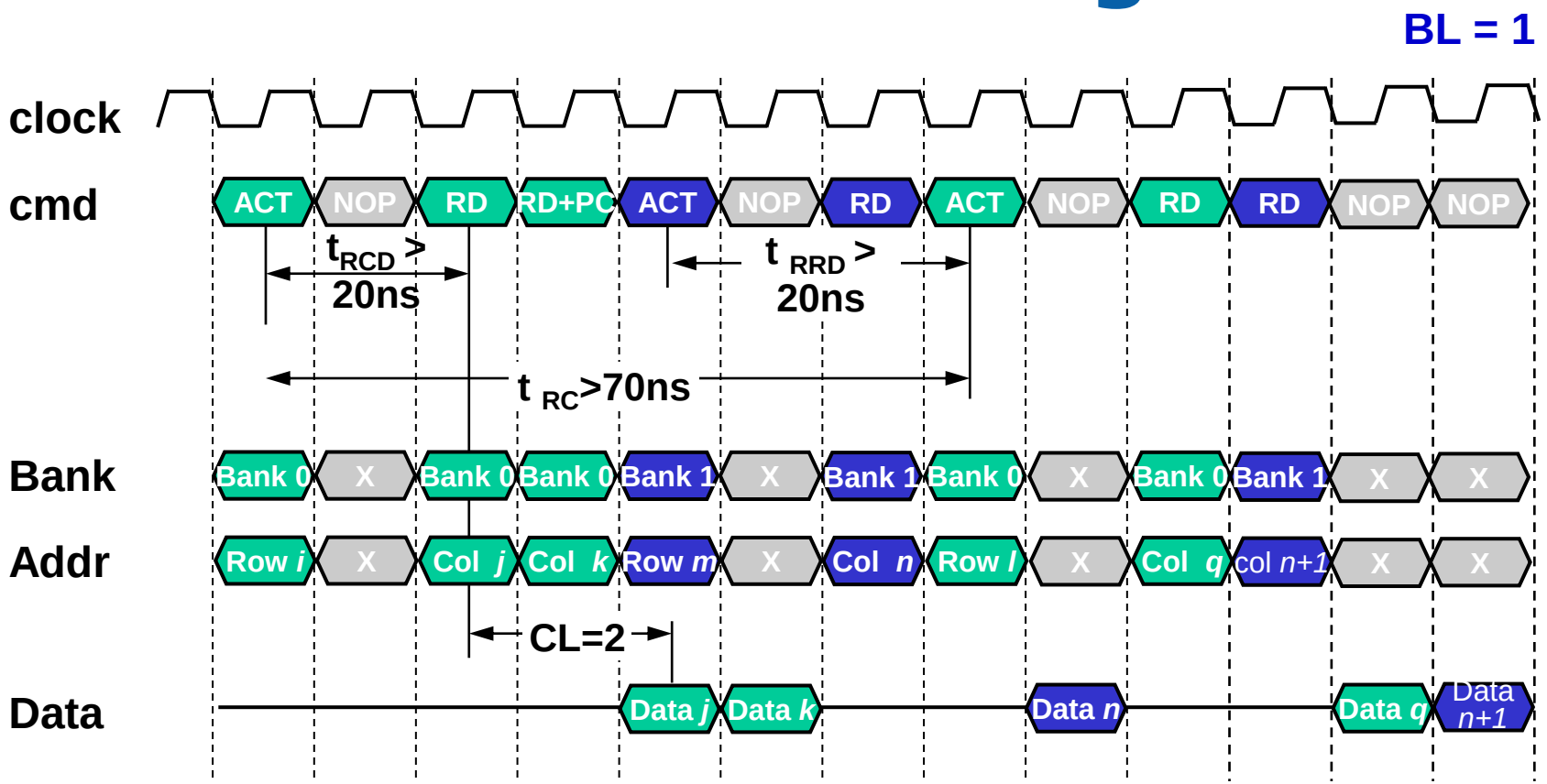
  - CAS-to-CAS delay: delay between two CASs within the same row
  - tRP: Row pre-charge time: the time to close current row, and open a new row

- **DRAM is not true random access memory**
  - A random access memory provides a constant access time regardless of the sequence of memory locations accessed
  - Accessing multiple columns within the same row is much faster than accessing addresses from different rows

Synchronous DRAM – SDRAM

- **All signals are referenced to an external clock (100MHz-200MHz)**
  - Makes timing more precise with other system devices
- **4 banks – multiple pages (rows) open simultaneously (one per bank)**
  - Accessing columns within each one of the 4 open rows is fast
- **Command driven functionality instead of signal driven**
  - ACTIVE: selects both the bank and the row to be activated
    - ACTIVE to a new bank can be issued while accessing current bank
  - READ/WRITE: select column
- **Burst oriented read and write accesses**
  - Successive column locations accessed in the given row
  - Burst length is programmable: 1, 2, 4, 8, full-page (may end by burst terminate)
- **A user programmable Mode Register**
  - CAS latency, burst length, burst type
- **Auto pre-charge: may close row at last read/write in burst**
- **Auto refresh: internal counters generate refresh address**

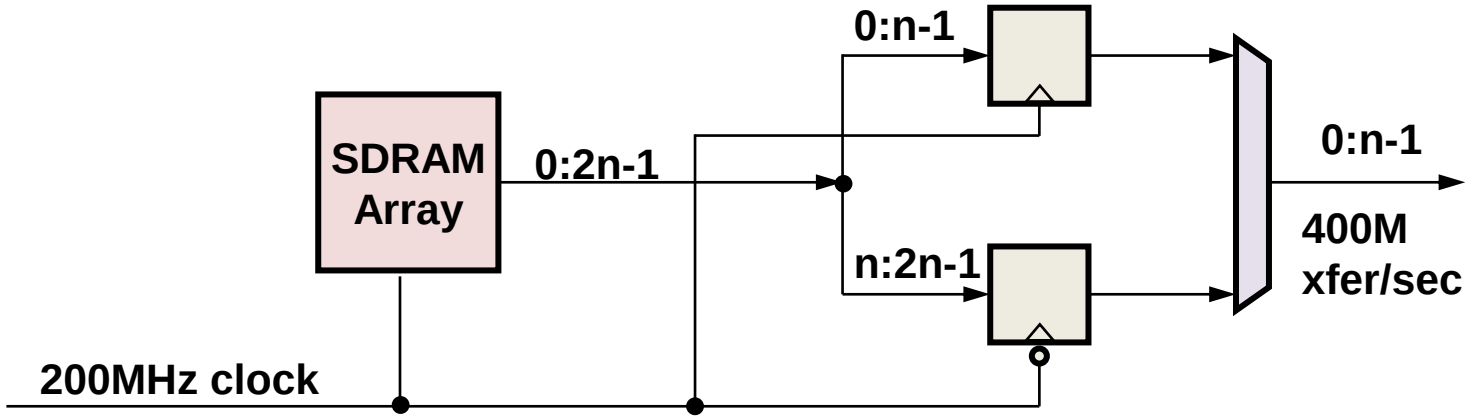
SDRAM Timing



- t_{RCD}: ACTIVE to READ/WRITE gap = $\lceil t_{RCD}(MIN) / \text{clock period} \rceil$
- t_{RC}: successive ACTIVE to a different row in the same bank
- t_{RRD}: successive ACTIVE commands to different banks

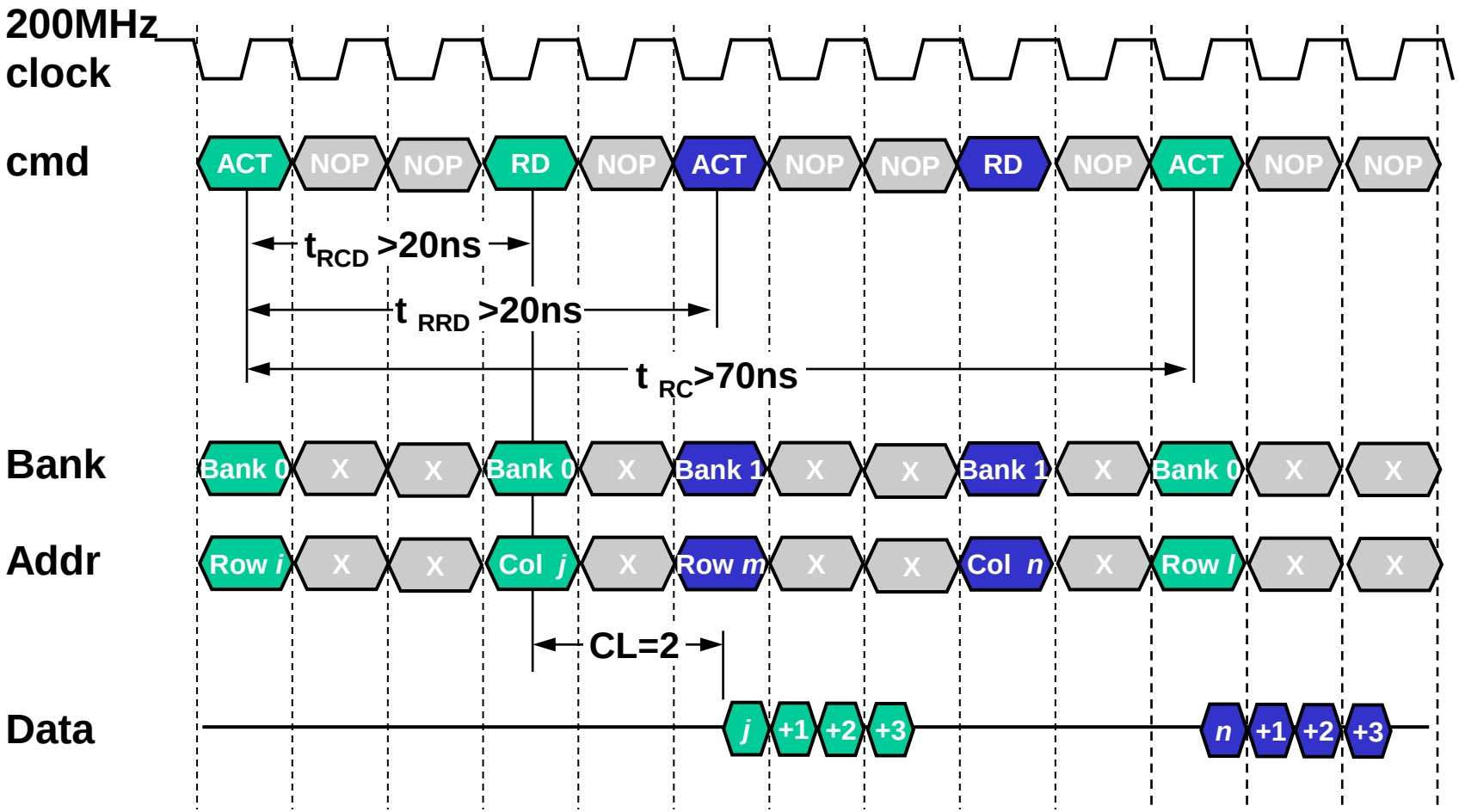
DDR-SDRAM

- **2n-prefetch architecture**
  - DRAM cells are clocked at the same speed as SDR SDRAM cells
  - Internal data bus is twice the width of the external data bus
  - Data capture occurs twice per clock cycle
    - Lower half of the bus sampled at clock rise
    - Upper half of the bus sampled at clock fall



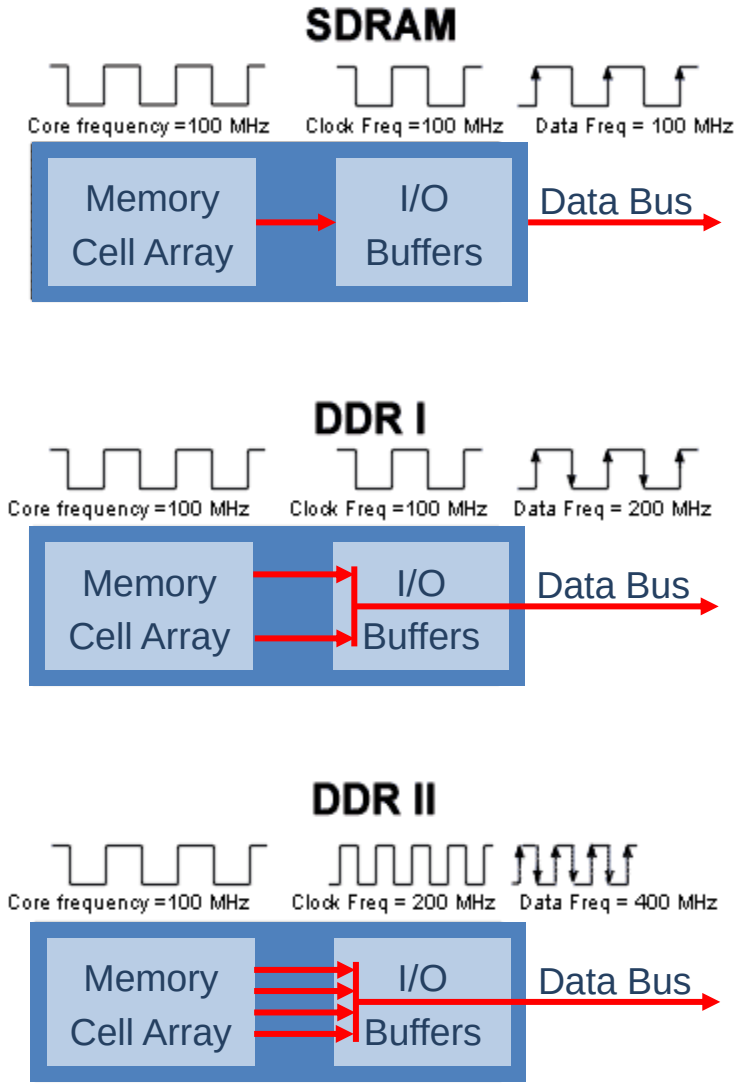
- **Uses 2.5V (vs. 3.3V in SDRAM)**
  - Reduced power consumption

DDR SDRAM Timing



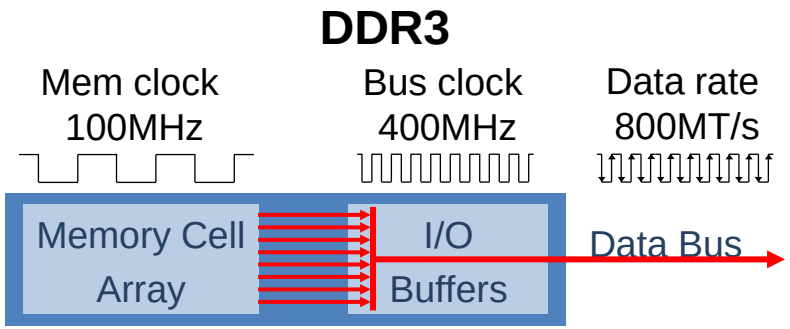
DDR2

- **DDR2 doubles the bandwidth**
  - 4n pre-fetch: internally read/write 4× the amount of data as the external bus
  - DDR2-533 cell works at the same freq. as a DDR266 cell or a PC133 cell
  - Prefetching increases latency
- **Smaller page size: 1KB vs. 2KB**
  - Reduces activation power – ACTIVATE command reads all bits in the page
- **8 banks in 1Gb densities and above**
  - Increases random accesses
- **1.8V (vs 2.5V) operation voltage**
  - Significantly lower power



DDR3

- **30% power consumption reduction compared to DDR2**
  - 1.5V supply voltage (vs. 1.8V in DDR2)
  - 90 nanometer fabrication technology
- **Higher bandwidth**
  - 8 bit deep prefetch buffer (vs. 4 bit in DDR2)



- **2× transfer data rate vs DDR2**
  - Effective clock rate of 800–1600 MHz
    - using both rising and falling edges of a 400–800 MHz I/O clock
    - DDR2: 400–800 MHz using a 200–400 MHz I/O clock
- **DDR3 DIMMs**
  - 240 pins, the same number as DDR2, and are the same size
  - Electrically incompatible, and have a different key notch location

DDR4 – 4th Gen. DDR SDRAM

- **Released to the market in 2014**
  - Higher module density: up to 64GB DIMMs vs. 16GB in DDR3
  - Lower voltage requirements: 1.2V and 1.4V vs. 1.5V and 1.65V in DDR3
  - Higher data rate transfer speeds
- **Prefetch has not been increased above the 8n used in DDR3**
  - The basic burst size is eight words
  - Higher bandwidth achieved by more read/write commands per sec
    - To allow this, DRAM banks are divided into two or four selectable bank groups, where transfers to different bank groups may be done more rapidly
- **The reduced power allows higher operation speeds**
  - Without going to expensive power and cooling requirements
  - Power consumption increases with speed
- **Frequency: 800 to 2133 MHz (DDR4-1600 through DDR4-4167)**
  - compared 400 to 1067 MHz in DDR3
- **Speeds are typically advertised as doubles of these numbers**
  - DDR3-1600 and DDR4-2400 are common
  - DDR3-3200 and DDR4-4800 available at high cost

DDR4 Standards

- **Standard name vs. module name**
  - DDR4-xxxx denotes per-bit data transfer rate in MT/sec
  - PC4-xxxxx denotes overall module (DIMM) transfer rate in MByte/sec
    - DDR4 modules transfer data on a 8byte data bus
      $\Rightarrow$ module peak transfer rate in MByte/sec = transfer rate × 8
- **Memory clock frequency, I/O bus clock frequency, Data rate**
  - 8 wide prefetch $\Rightarrow$ Data rate = Memory clock freq × 8
  - Data transferred on both clock edges $\Rightarrow$ Data rate = I/O bus clock freq × 2
- **DRAM timing, measure in I/O bus cycles, specifies 3 numbers**
  - CAS Latency – RAS-to-CAS Delay – RAS Pre-charge Time
- **CAS latency (latency to get data in an open page) in nsec**
  - CAS Latency × I/O bus cycle time in the example: $[11/(800\times10^6)]\times10^9 = 13.75$

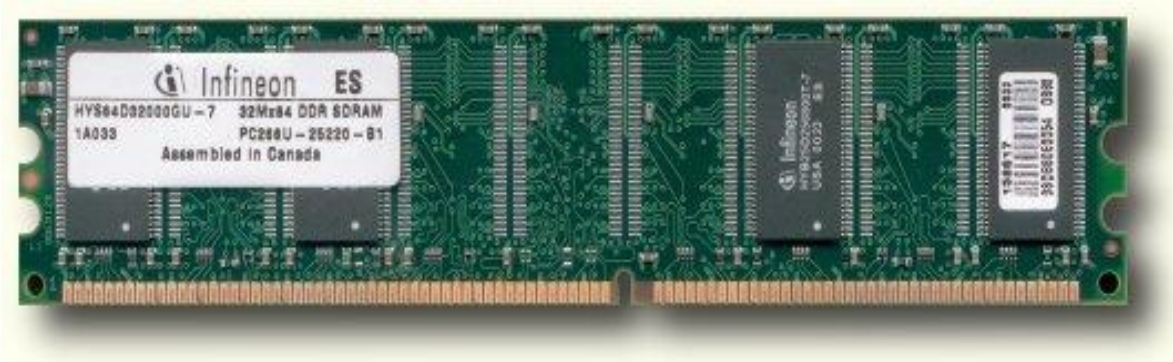
Standard name	Memory clock (MHz)	I/O bus clock (MHz)	Data rate (MT/s)	Module name	Peak transfer rate (MB/s)	Timings CL-tRCD-tRP	CAS latency (ns)
DDR4-1600K	200	800	1600	PC4-12800	12800	11-11-11	13.75

DDR4 Standards

Standard name	Memory clock (MHz)	I/O bus clock (MHz)	Data rate (MT/s)	Module name	Peak transfer rate (MB/s)	Timings CL-tRCD-tRP	CAS latency (ns)
DDR4-1600J* DDR4-1600K DDR4-1600L	200	800	1600	PC4-12800	12800	10-10-10 11-11-11 12-12-12	12.5 13.75 15
DDR4-1866L* DDR4-1866M DDR4-1866N	233.33	933.33	1866.67	PC4-14900	14933.33	12-12-12 13-13-13 14-14-14	12.857 13.929 15
DDR4-2133N* DDR4-2133P DDR4-2133R	266.67	1066.67	2133.33	PC4-17000	17066.67	14-14-14 15-15-15 16-16-16	13.125 14.063 15
DDR4-2400P* DDR4-2400R DDR4-2400T DDR4-2400U	300	1200	2400	PC4-19200	19200	15-15-15 16-16-16 17-17-17 18-18-18	12.5 13.32 14.16 15
DDR4-2666T DDR4-2666U DDR4-2666V DDR4-2666W	325	1333	2666	PC4-21333	21333	17-17-17 18-18-18 19-19-19 20-20-20	12.75 13.50 14.25 15
DDR4-2933V DDR4-2933W DDR4-2933Y DDR4-2933AA	366.6	1466.5	2933	PC4-23466	23466	19-19-19 20-20-20 21-21-21 22-22-22	12.96 13.64 14.32 15
DDR4-3200W DDR4-3200AA DDR4-3200AC	400	1600	3200	PC4-25600	25600	20-20-20 22-22-22 24-24-24	12.50 13.75 15

DIMMs

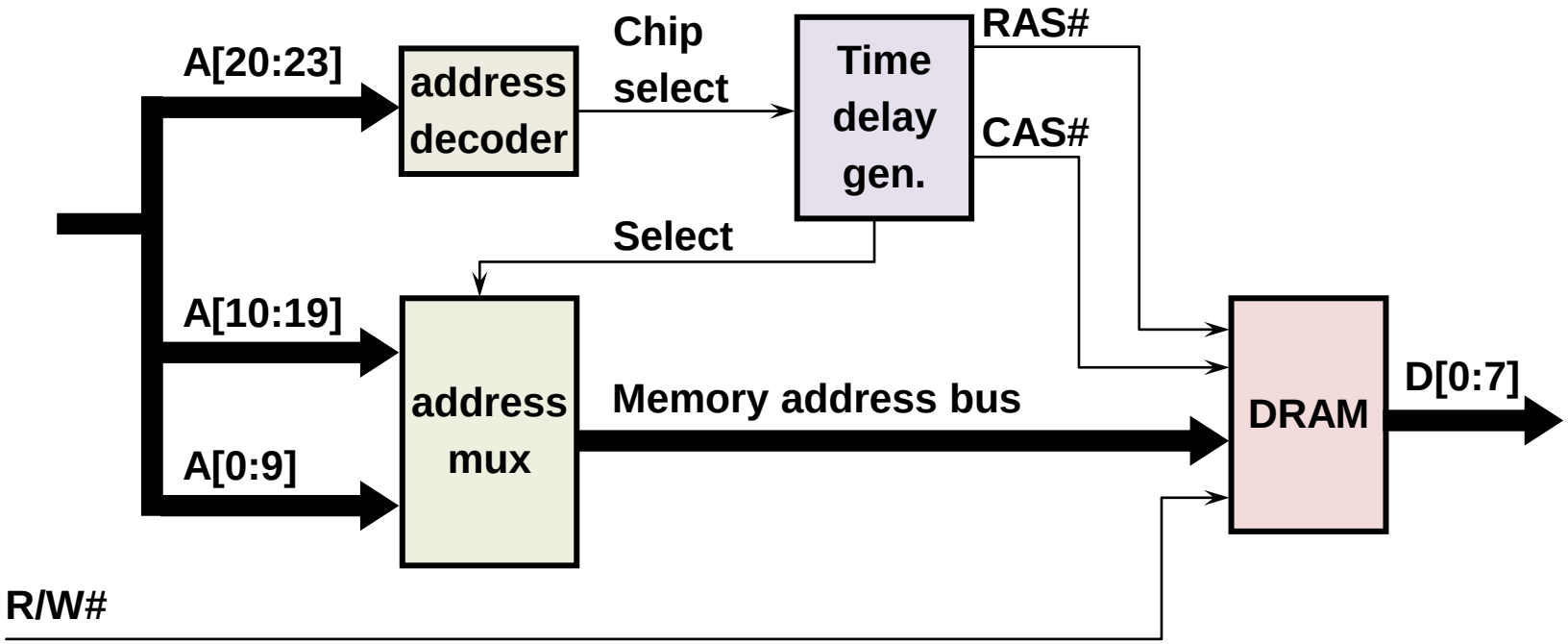
- **DIMM: Dual In-line Memory Module**
  - A small circuit board that holds memory chips



- **64-bit wide data path (72 bit with parity)**
  - Single sided: 9 chips, each with 8 bit data bus
  - Dual sided: 18 chips, each with 4 bit data bus
  - Data BW: 64 bits on each rising and falling edge of the clock
- **Other pins**
  - Address – 14, RAS, CAS, chip select – 4, VDC – 17, Gnd – 18, clock – 4, serial address – 3, …

The DRAM Controller

- **DRAM controller gets address and command**
  - Splits address to Row and Column
  - Generates DRAM control signals at the proper timing



- **DRAM data must be periodically refreshed**
  - The DRAM controller performs DRAM refresh, using refresh counter

The DRAM Controller (cont.)

- **A typical DRAM controller supports two channels of DDR4**
- **Each channel has its own resources**
  - Handles memory requests independently
  - Contains a 32 cache-line write-data-buffer
  - Supports 8 bytes per cycle

- **A hash function distributes addresses between channels**
  - Attempts to balance the load between the channels in order to achieve maximum bandwidth and minimum hotspot collisions
- **An out-of-order scheduler maximizes BW and minimize latency**
  - Writes to the memory controller are considered completed when they are written to the write-data-buffer
  - The write-data-buffer is flushed out to main memory at a later time, not impacting write latency

Partial Memory Writes

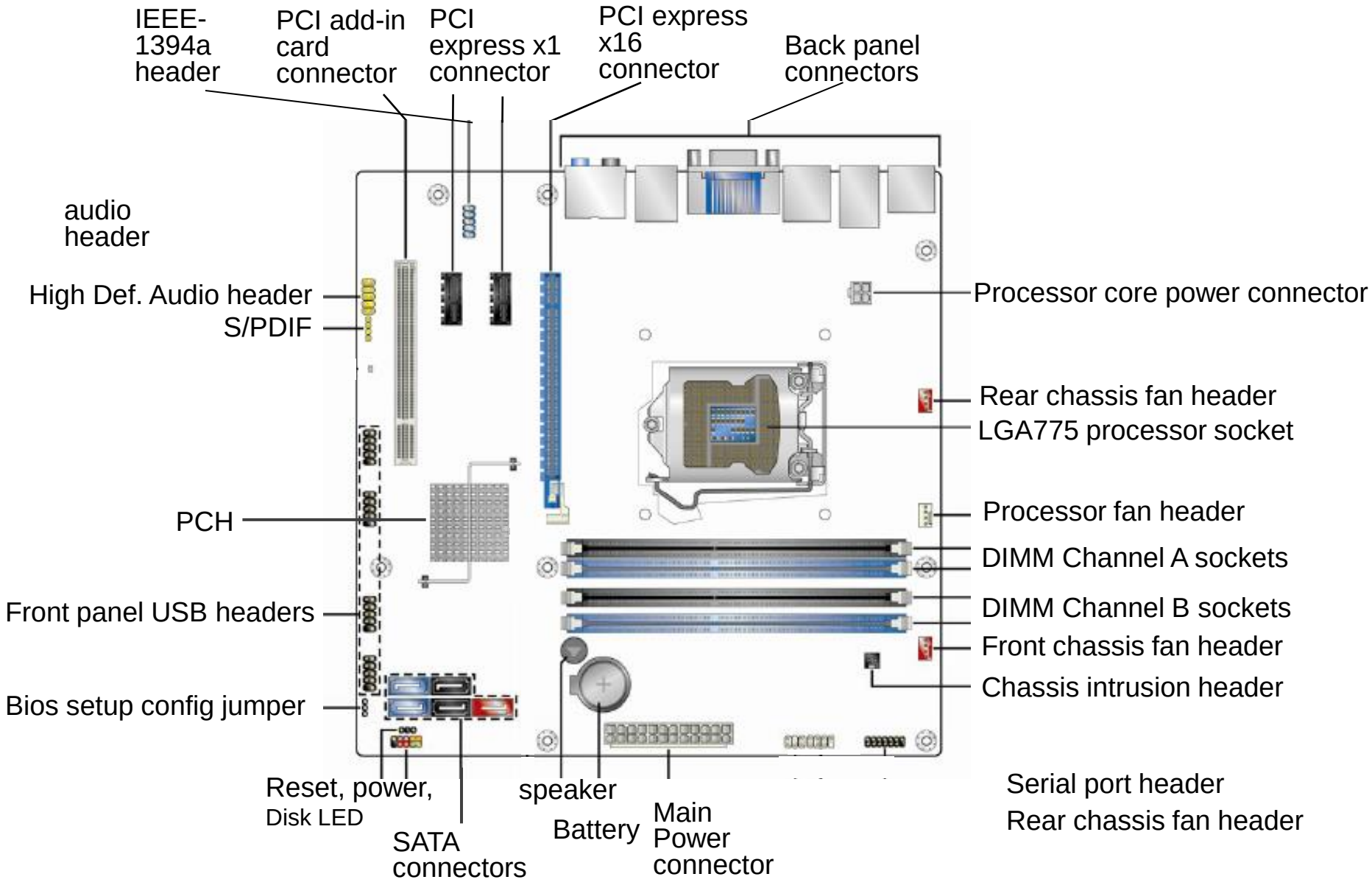
- **Memory reads and writes are done at full Cache Line granularity**
- ***Partial Write* transactions are writes to a subset of a Cache Line**
  - Non-cacheable writes, e.g., write requests driven by IO devices
  - Would require adding more pins to indicate to memory which bytes to write
  - Not supported with ECC
- **Partial writes are not common enough to justify the extra cost**
  - The MC supports a partial write, by a RMW operation
    - Reads the current value of the full CL
    - Replaces the bytes that have to be modified
    - Writes back the full CL to memory
- **Software should avoid creating partial write transactions whenever possible**
  - E.g., buffer the partial writes into full cache line writes

How to get the most of Memory ?

- **For best performance**
  - Populate both channels with equal amounts of memory
    - Preferably the exact same types of DIMMs
  - Use highest supported speed DRAM, with the best DRAM timings

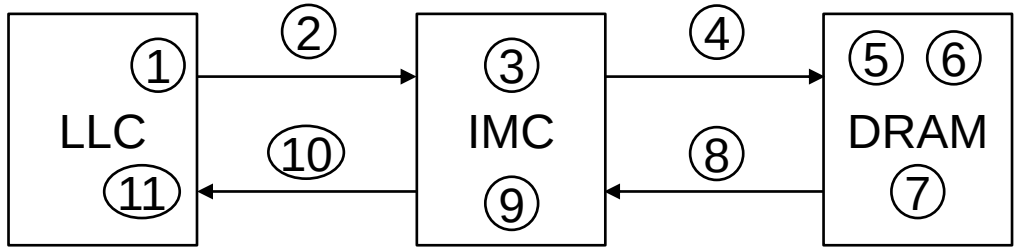
- **Each DIMM supports 4 open pages simultaneously**
  - The more open pages, the more random access
  - It is better to have more DIMMs: *n* DIMMs $\Rightarrow$ 4*n* open pages
  - Dual sided DIMMs may have separate CS of each side
    - Support 8 open pages
    - Dual sided DIMMs may also have a common CS

Motherboard Layout



DRAM Access Latency

1. Delay in LLC request buffer
2. Request sent from LLC to IMC
3. IMC delay
4. IMC sends command to DRAM
5. Row pre-charge
6. RAS-to-CAS
7. CAS latency
8. Data returns to IMC
9. IMC delay
10. Data sent from IMC to LLC
11. Delay in LLC fill buffer



DRAM Access Latency

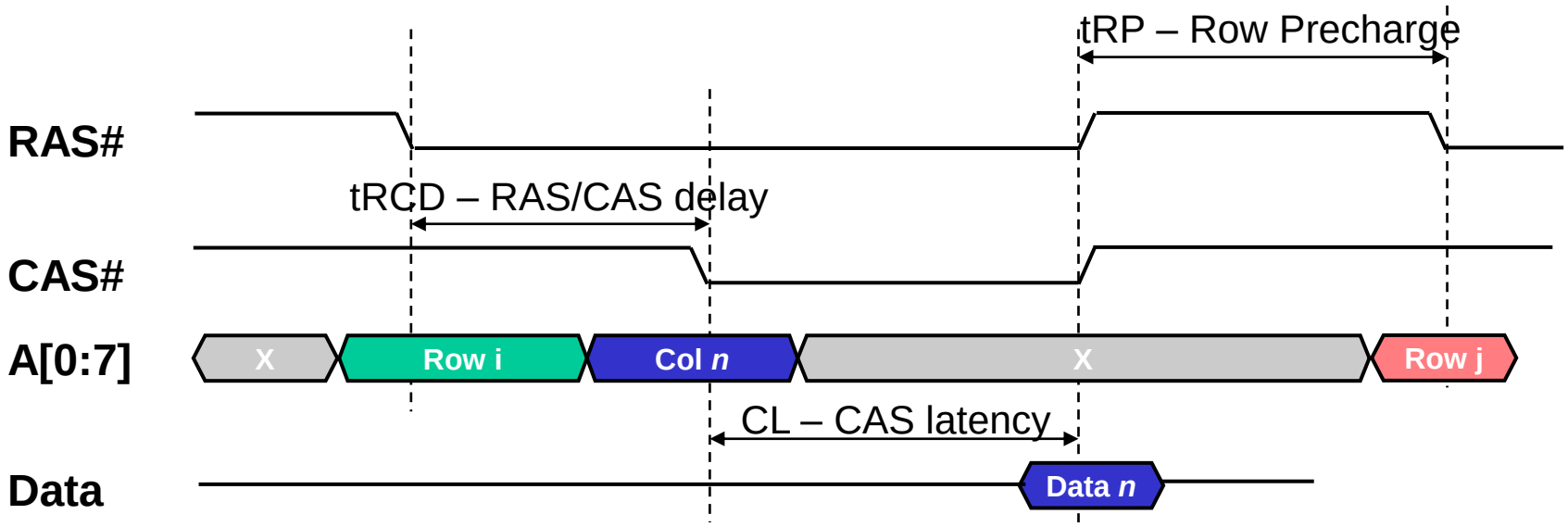
- **Assume CPU clock cycle of 4GHz = 250ps = 0.25ns**
- **Single access to a new page**
  - In the DRAM
    - Row pre-charge 13ns
    - RAS-to-CAS 13ns
    - CAS latency 13ns
  - Rest of the system 25ns
  - Total 64ns = 256 CPU clock cycles
- **Single access to a an open page**
  - In the DRAM
    - CAS-to-CAS 6ns
    - CAS latency 13ns
  - Rest of the system 25ns
  - Total 44ns = 176 CPU clock cycles

DDR5 vs. DDR4

	DDR4	DDR5	DDR5 Advantage
Speed	1.6 – 3.2 GT/s 0.8 – 1.6 GHz clock	4.8 – 8.4 GT/s 1.6 – 4.2 GHz clock	Higher bandwidth
Voltage	1.2V	1.1V	Lower power
Power management	On motherboard	On DIMM PMIC	Better power efficiency Better scalability
Channels	64bit data + 8bit ECC 1 channels / DIMM	32bit data + 8bit ECC 2 channels / DIMM	Higher memory efficiency Lower latency
Burst Length	BC4, BL8	BC8, BL16	Higher memory efficiency
Max memory / chip	16 Gbit	64 Gbit	Higher capacity DIMMs

Backup

DRAM Access Sequence Timing



- Put row address on address bus and assert RAS#
- Wait for RAS# to CAS# delay (tRCD) between asserting RAS and CAS
- Put column address on address bus and assert CAS#
- Wait for CAS latency (CL) between time CAS# asserted and data ready
- Row pre-charge time: time to close current row, and open a new row

SRAM vs. DRAM

- **Random Access: access time is the same for all locations**

	DRAM – Dynamic RAM	**SRAM – Static RAM**
Refresh	Refresh needed	No refresh needed
Address	Address muxed: row+ column	Address not multiplexed
Access	Not true "Random Access"	True "Random Access"
density	High (1 Transistor/bit)	Low (6 Transistor/bit)
Power	low	high
Speed	slow	fast
Price/bit	low	high
Typical usage	Main memory	cache

DDR4 vs DDR3

eature/Option	DDR3	DDR4	DDR4 Advantage
Voltage (core and I/O)	1.5V	1.2V	Reduces memory power demand
V_{REF} inputs	2 – DQs and CMD/ADDR	1 – CMD/ADDR	V_{REFDQ} now internal
Low voltage standard	Yes (DDR3L at 1.35V)	No	Memory power reductions
Data rate (Mb/s)	800, 1066, 1333, 1600, 1866, 2133	1600, 1866, 2133, 2400, 2666, 3200	Migration to higher-speed I/O
Densities	512Mb–8Gb	2Gb–16Gb	Better enablement for large-capacity memory systems
Internal banks	8	16	More banks
Bank groups (BG)	0	4	Faster burst accesses
^{t}CK – DLL enabled	300 MHz to 800 MHz	667 MHz to 1.6 GHz	Higher data rates
^{t}CK – DLL disabled	10 MHz to 125 MHz (optional)	Undefined to 125 MHz	DLL-off now fully supported
Read latency	AL + CL	AL + CL	Expanded values
Write latency	AL + CWL	AL + CWL	Expanded values
DQ driver (ALT)	40Ω	48Ω	Optimized for PtP (point-to-point) applications
DQ bus	SSTL15	POD12	Mitigate I/O noise and power
R_{TT} values (in Ω)	120, 60, 40, 30, 20	240, 120, 80, 60, 48, 40, 34	Support higher data rates
R_{TT} not allowed	READ bursts	Disables during READ bursts	Ease-of-use
ODT modes	Nominal, dynamic	Nominal, dynamic, park	Additional control mode; supports OTF value change
ODT control	ODT signaling required	ODT signaling not required	Ease of ODT control, allows non-ODT routing on PtP applications
Multipurpose register (MPR)	Four registers – 1 defined, 3 RFU	Four registers – 3 defined, 1 RFU	Provides additional specialty readout

DDR Comparison

	DDR	DDR2	DDR3	DDR4
Prefetch depth	2	4	8	8
Memory Clock (MHz)	100-200	100-200	100-266⅔	200-400
I/O bus clock (MHz)	100-200	200-400	400-1066⅔	800-1600
Data rate (MT/s)	200-400	400-800	800-2133	1600-3200
Module rate (GB/s)	1.6-3.2	3.2-6.4	6.4-17.1	12.8-25.6
CAS latency (ns)	9.4-12.5	11.2-15	11-15	12.5-15
DRAM timing (lowest)	2.5-3-3	3-3-3	5-5-5	10-10-10
Voltage (standard/low)	2.5V / 1.8V	1.8V	1.5V / 1.35V	1.2V / 1.05V
Power (mW)	399	217		
Max DIMM size	1GB	4GB	16GB	64GB
Internal banks	4	4 / 8	8	16
Banks Groups	n.a.	n.a.	n.a.	4
Year released	2000	2003	2007	2014

DDR5

- Higher bandwidth, improved power consumption, higher capacity per memory module
- Supports up to 64-gigabit memory chips, or 128GB/DIMM (DDR4: 16-gigabit)
  - Initial DDR5 memory modules use 16-gigabit memory chips, i.e. 32GB per module
- Supports die stacking with up to eight dies on a chip – 2TB per module
- Data rate: 3200 – 6400 MT/s (DDR4: 1600 – 3200 MT/s)
  - Actual mainstream starts at 4800 MT/s (DDR4: 2133 MT/s)
- DDR5 memory modules support two independent 32-bit channels (40-bit with ECC)
  - DDR4: single 64-bit channel per module (72-bit with ECC)
  - For a dual DIMM: 4 × 32-bit (vs. 2 × 64-bit on DDR4)
- Double burst length: from eight bytes (BL8) to 16 bytes (BL16)
- 1.1V (vs 1.2V for DDR4)
- DDR5 memory modules include a power management IC (PMIC)
  - Reduces motherboard cost and design complexity, but adds cost to the memory modules
- 32-bank structure divided into eight groups (16-bank / 4 groups)
  - Enables more pages to be open consecutively