## Schedule of Course Activities: Session 22

## *[Cloud 519: Introduction to Cloud Computing Online-Based]*

## *[Instructor: John C. Chan]*

|  |  |
| --- | --- |
| **Overview of Session** |  |
| We will answer the following questions: | 1. Memory Sub-system in a Computer System. 2. Peripheral IO, Graphics in a Computer System. 3. … |

# In the design of modern computer system, special considerations are made to its memory sub-system. The fundamental need include these:

# The CPU can execute 1-insreutction per clock cycle (Upton 5G Hz, 2ns per clock cycle).

# How can you avoid the CPU being starved of data input for it to process?

# How can you avoid CPU being idle, waiting for its data output (processed result), to be deposited?

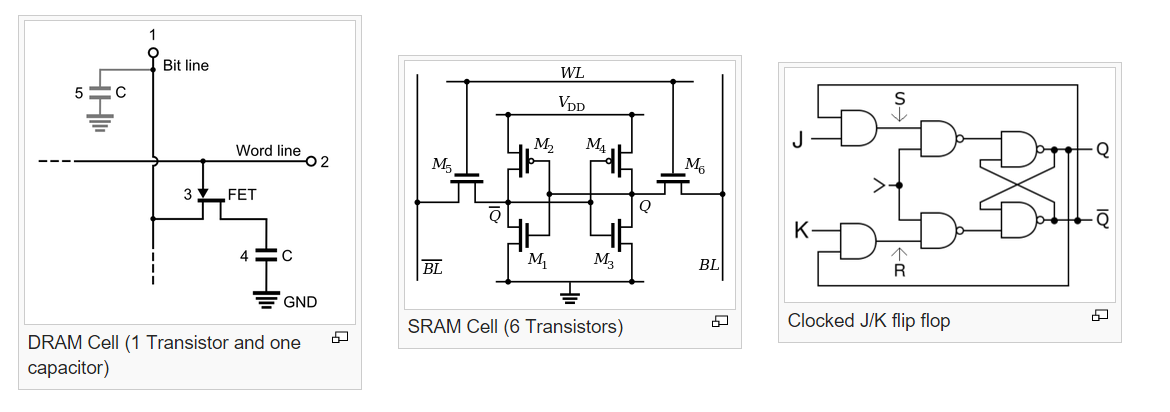
# How can you avoid memory access delay, due to bus contentions, data access locks etc.?

# How large should your memory buffering pool be?

(Detail answers to each of the questions above are beyond the scope of the class. Instead, let’s start by learning the basics of the memory/IO element in the computer system next).

# **Memory Technology: DRAM/DIMM in Finer Details.**

(The figure show below, is clearly beyond the scope of the class. It is at the transistor-level! But then, it conveys the most fundamental concept. This is how memory is design by electrical engineer. From left to right, you can see the complexity of the circuitry increases [DRAM->SRAM->FF]. These are the building blocks of the memory used in all computing [Your PC, or Cloud] included! This section is tailored for the hardware geeks, the game players).



**This Video summarizes many subjects covered computer memories from user’s perspectives (enjoy):**

<https://www.youtube.com/watch?v=dZcszUj5szA>

**Key Take Ways:**

* + DRAM, DDR3, volatile memory.
  + SRAM, Cache.
  + Flash Memory, EPROM. Non-volatile memory.
  + SSD.
  + HDD.

**Question:** Can we ever get rid of HDD?

Memory Technology: DRAM, Data access speeds:

* DDR: 3.2GB/s;
* DDR2: 8.5GB/s’
* DDR3: 1.5V; 12.8B/s; Up to 64GB on a single DIMM (240 Pin). ; 2 or 3 modules; Sockets is specific for the DIMM type. 30% less power.
  + The primary benefit of DDR3 SDRAM over its immediate predecessor, DDR2 SDRAM, is its ability to transfer data at twice the rate (eight times the speed of its internal memory arrays), enabling higher bandwidth or peak data rates.
  + With data being transferred 64 bits at a time per memory module, DDR3 SDRAM gives a transfer rate of (memory clock rate) × 4 (for bus clock multiplier) × 2 (for data rate) × 64 (number of bits transferred) / 8 (number of bits/byte). Thus with a memory clock frequency of 100 MHz, DDR3 SDRAM gives a maximum transfer rate of 6400 MB/s.
  + DDR3 prefetch buffer is 8-burst deep (vs. DDR2 of 4).
  + DDR3 transfer data on both the rising edge and falling edges of the IO clock. (e.g.400MHz).
  + Actual Performance is much lower, e.g.1200MB to 1900MB on memory bus.
* DDR4: 1.2V; 32GB/s; Upto 256GB on a single DIMM (288 Pin)
* DIMM: Replace SIMM, has 64-bit data path. ECC modules usually carry a multiple of 9 instead of a multiple of 8 chips.
* Mobile DRAM: Lower Power Consumptions, than PC DRAM.

Memory Technology: Graphic memory/Video-RAM:

<https://www.youtube.com/watch?v=Utv144XeHag>

Key Take-Away:

* + Like DRAM feeds CPU, VRAM feeds data to GPU e.g. Texture, Frame Buffer, Shadow Maps, requires memory; Graphics need to move massive amount of data into/out of the frame buffer; Whereas system memory focus on low latency on smaller transfers, high bandwidth, low power. 8GB is the typical requirement, rendering takes huge amount of memory.
  + GDDR5 is current Graphic memory. Similar to DDR3 but optimized for graphics application. Most graphic cards integrate the VRAM onboard.
  + Memory bandwidth in GForce: 92GB/s to 128GB/sec.
    - Screen Resolutions typical 1080x860 pixel, 65 frames/s, 32-bit depth resolution, a single frame means 32MB of data. This demands memory!
    - Anti-aliasing, takes a lots of memory; This demands memory.
    - Case run out of memory, grab from system memory…performance degrades.

**Thought provoking questions:**

* + **Will DRAM and SSD merge? (Hint: Intel and Micron just made such a technological breakthrough announcement).**
  + **…**

# **PCI Express: the Peripheral IO Interface**

**PCI Express** (**Peripheral Component Interconnect Express**), officially abbreviated as **PCIe**, is a high-speed serial computer expansion bus standard designed to replace the older PCI, PCI-X, and AGP bus standards. PCIe has numerous improvements over the older standards, including higher maximum system bus throughput, lower I/O pin count and smaller physical footprint, better performance scaling for bus devices, a more detailed error detection and reporting mechanism (Advanced Error Reporting, AER), and native hot-plug functionality. More recent revisions of the PCIe standard provide hardware support for I/O virtualization.

The PCI Express electrical interface is also used in a variety of other standards, most notably in ExpressCard as a laptop expansion card interface, and in SATA Express as a computer storage interface.

Format specifications are maintained and developed by the PCI-SIG (PCISpecial Interest Group), a group of more than 900 companies that also maintain the conventional PCI specifications. PCIe 3.0 is the latest standard for expansion cards that is in production and available on main stream personal computers.

PCI Express in the backbone of peripheral IO interfaces found on the motherboard of a typical computer server.

Form Factors:

A PCIe card fits into a slot of its physical size or larger (with ×16 as the largest used), but may not fit into a smaller PCIe slot; for example, a ×16 card may not fit into a ×4 or ×8 slot. Some slots use open-ended sockets to permit physically longer cards and negotiate the best available electrical and logical connection.

The next figure shows the various flavors of PCI Express.

Various slots on a computer motherboard, from top to bottom:

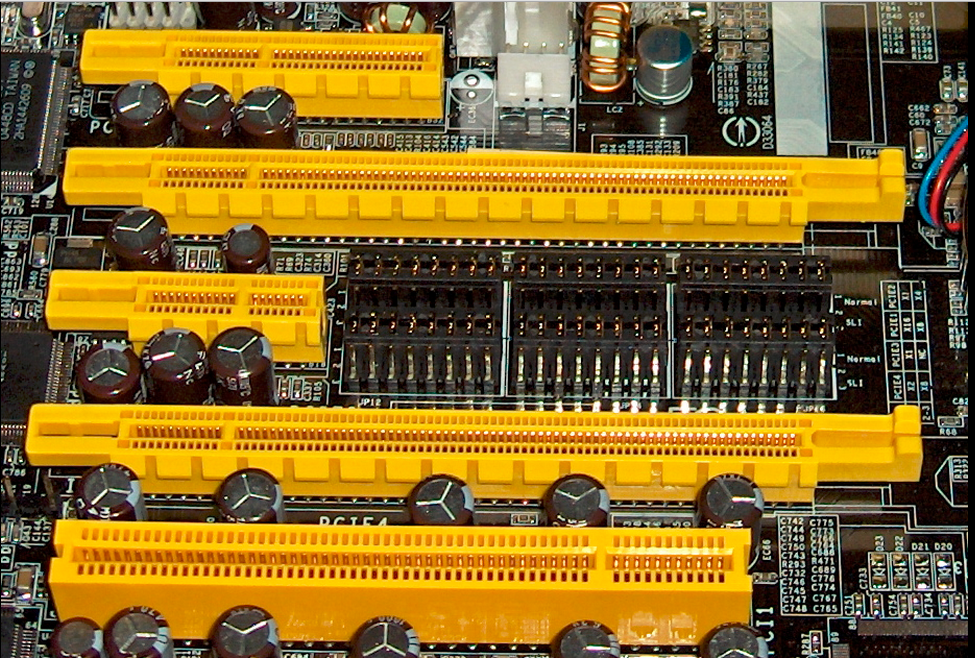
PCI Express ×4

PCI Express ×16

PCI Express ×1

PCI Express ×16

Legacy PCI (32-bit, 5 V)



Let learn about PCIe from this introductory video:

<https://www.youtube.com/watch?v=LSSHuMHbCWo>

Key Take-Away:

* IO peripheral plug-in.
* Compatibility, bandwidth compromise.
* Your Network Interface Card, plugs into here also.
* …

Class Assignment: Name 3 IO devices that are supported by the PCIe Bus:

1. USB
2. NIC: Network Interface Card.
3. Disk Drives
4. Wi-Fi.
5. CD-ROM
6. None of above

End-of-Class Module.

Questions? Please email to me, or post it on Blackboard.

Thank you.