Contents:

1. Slide1

Top

---

2. Slide2

MMU(Memory Management Unit)

- MMU is a class of computer hardware components responsible for handling memory accesses requested by the CPU.

- Among the functions of such devices are

  - the translation of virtual addreses to physical addresses (i.e., virtual memory management)

In computer terminology a **virtual address** is an address identifying a virtual (non-physical) entity. The term virtual address is most commonly used for virtual memory or virtual network address.

Top

---

3. Slide3

In computer science, a **physical address** is the address presented to a computer's main memory in a virtual memory system, in contrast to the virtual address which is the address generated by the CPU. Virtual addresses are translated into **physical addresses** by a memory management unit(abbreviated MMU).

- Memory protection,

Memory protection is a system that prevents one process from corrupting the memory of another process running on the same computer at the same time.

It usually employs hardware and system software to allocate distinct memory to different processes and to handle exceptions arising when a process tries to access memory outside its bounds.

There are different ways to achieve memory protection. These include:

Top

---

4. Slide4

Segmentation, Paging, Protection Keys.

Segmentation: means that the currently running process is restricted to using only certain parts of memory called segments.

A segment is defined by a pair of hardware registers which give the limits of allowed memory addresses.

Paging: is the method mostly used for memory protection. In paging, the memory address space is divided into small pieces, called pages.

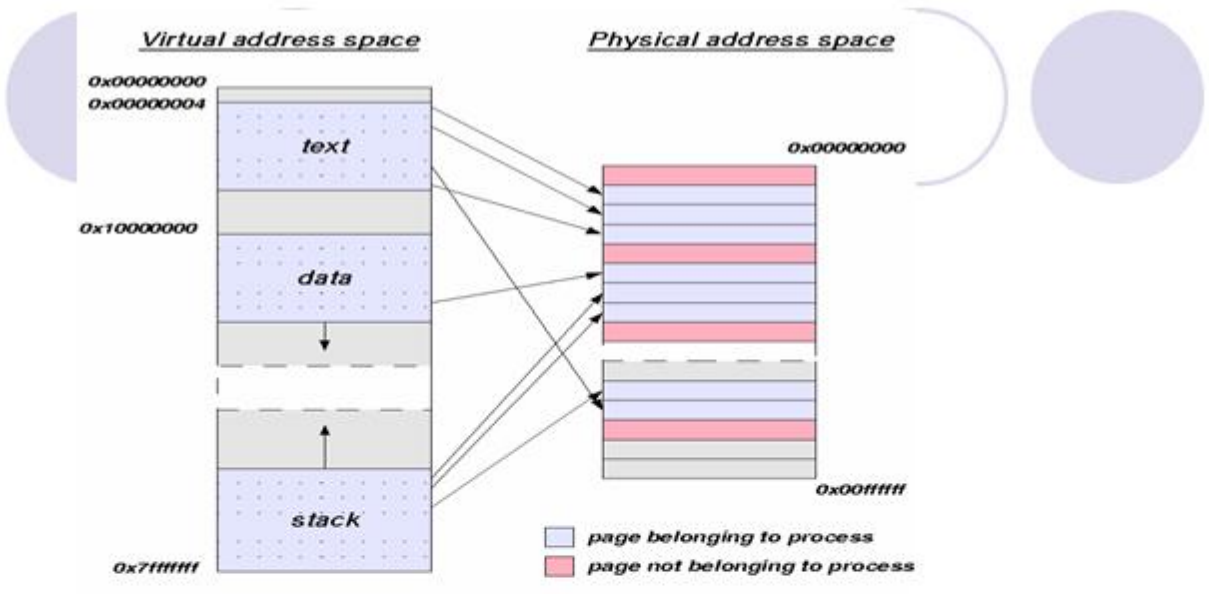
The memory access part of paging is done at the hardware level via page tables, and is handled by the memory management unit.

A page table is the data structure used by a virtual memory system in a computer operating system to store the mapping between virtual addresses and physical addresses. Virtual addresses are those unique to the accessing process. Physical addresses are those unique to the CPU, i.e., RAM.

Top

---

5. Slide5



Relationship between pages addressed by virtual addresses and the frames in physical memory, within a simple address space scheme. Physical memory can contain pages belonging to many processes. Pages can be swapped to disk if used infrequently, or if physical memory is full. Not all pages are in physical memory, here.

Top

---

6. Slide6

As mentioned earlier, physical memory is divided into small blocks called pages (typically 4 kilobytes or less) in size.
Each block is assigned a page number.

The operating system may keep a list of free pages in its memory, or may choose to probe the memory each time a memory request is made.

Whatever the case, when a program makes a request for memory, the operating system allocates a number of pages to the program, and keeps a list of allocated pages for that particular program in memory

When paging is used alongside with virtual memory. The operating system has to keep track of pages in use.Pages will not be used or have not been used for some time. Then, when the operating system deems fit, or when a program requests a page that has been swapped out, the operating system swaps out a page to disk, and brings another page into memory.

Top

---

7. Slide7

In this way, you can use more memory than your computer physically has.

Protection Keys:

A protection key mechanism divides physical memory up into blocks of a particular size (e.g. 2KB), each of which has an associated numerical value called a protection key.

- Cache control

A CPU cache is a cache used by the central processing unit of a computer to reduce the average time to access memory. The cache is a smaller, faster memory which stores copies of the data from the most frequently used main memory locations.

- Bus arbitration,
- Bank switching.

Top

---

8. Slide8

Bank switching (also known as "paging", but only loosely related to the ordinary meaning of "paging" in computing) was a technique common in 8-bit microcomputer systems, to increase the amount of addressable RAM and ROM without extending the address bus.

- Modern MMUs typically divide the virtual address space (the range of addresses used by the processor) into pages.
  - An **address space** defines a context in which a memory address makes sense. A memory address identifies a physical location in computer memory.

  Example address spaces:

  - Main memory (physical memory)
  - Virtual memory : memory created by using the hard disk to simulate additional

Top

---

9. Slide9

random-access memory; the addressable storage space available to the user of a computer system in which virtual addresses are mapped İnto real address.

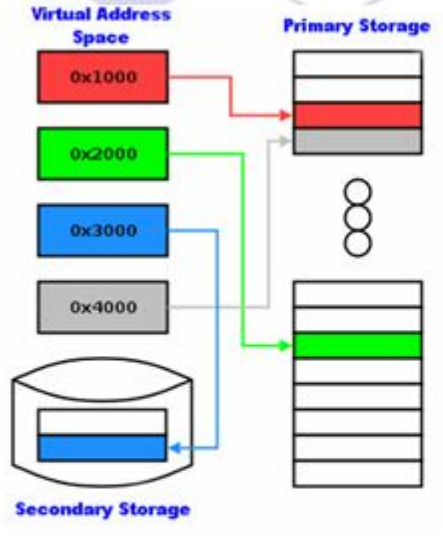
Virtual memory or virtual memory addressing is a memory management technique.

Virtual memory addressing is typically used in paged memory systems. This in turn is often combined with *memory swapping, whereby memory pages stored in primary storage are written to secondary storage (often to a swap file or swap partition), thus freeing faster primary storage for other processes to use.*

Top

---

10. Slide10



The memory pages of the virtual address space seen by the process, may reside non-contiguously in primary, or even secondary storage.

Top

---

11. Slide13

- I/O port space
- IP address

An **IP address (Internet Protocol address)** is a unique number that devices use in order to identify and communicate with each other on a computer network utilizing the Internet Protocol standard (**IP**).

- **Address space** size is 2N, usually a few kilobytes. The bottom ***n*** bits of the address (the offset within a page) are left unchanged. The upper address bits are the (virtual) page number.

The MMU normally translates virtual page numbers to physical page numbers via an associative cache called a Translation Lookaside Buffer (TLB).

Top

---

12. Slide11

A **Translation Lookaside Buffer (TLB)** is a buffer (or cache) in a CPU that contains parts of the page table which translate from virtual İnto real addresses.

- Translation Lookaside Buffer has a fixed number of entries and is used to improve the speed of virtual address translation.
- The buffer is typically a **content addressable memory** (CAM is a special type of computer memory used in certain very high speed searching applications. It is also known as *associative memory*) in which the search key is the virtual address and the search result is a real or physical address .

Top

---

13. Slide12

- When the TLB lacks a translation, a slower mechanism involving hardware-specific data structures or software assistance will be used.
- The data items found in such data structures are typically called *page table entries* (PTEs). The physical page number is combined with the page offset to give the complete physical address.
- A PTE or TLB entry may also include information about whether the page has been written to (the *dirty bit*), when it was last used , what kind of processes may read and write it, and whether it should be cached.

It is possible that TLB entry or PTE prohibits access to a virtual page, perhaps because no physical memory (RAM) has been allocated to that virtual page. In this case the MMU will signal a page fault to the CPU.

Top

---

14. Slide14

Page Fault:

In computer storage technology, a **page fault** is an interrupt(or exception) to the software raised by the hardware, when a program accesses a page that is not mapped in physical memory.

- The operating system will then handle the situation appropriately, perhaps by trying to find a spare page of RAM and set up a new PTE to map it to the requested virtual address.
- If no RAM is free it may be necessary to choose an existing page, using some replacement algorithm, and save it to disk (paging).
- In some cases a "page fault" may indicate a software bug. A key benefit of an MMU is memory protection.

Top

---

15. Slide15

An MMU also reduces the problem of fragmentation of memory. (**Fragmentation** is a term that occurs in several fields and describes a process of something breaking or being divided into pieces (fragments)).

- Blocks of memory have been allocated and freed.
- the free memory may become fragmented (discontinuous) so that the largest contiguous block of free memory may be much smaller than the total amount.
- With virtual memory, a contiguous range of virtual addresses can be mapped to several non-contiguous blocks of physical memory.

Top

---

16. Slide16

REFERENCES:

1) http://en.wikipedia.org/wiki/Memory_management_unit

2)http://www.memorymanagement.org/glossary/m.html#mmu

Top