**Module 4: Discussion**

In this module, you were introduced to memory concepts and management. Discuss the differences between physical and logical memory. What methods are used for allocating memory contiguously? In response to your peers, provide constructive feedback and at least one difference between internal and external fragmentation.

Alejandro Ricciardi

June 7, 2023

Hello class,

The discussion question refers to physical and logical memory and it is a little different concept than physical and logical addresses. In this post, I will define and discuss both. In simple terms, a memory is a physical place where the data is stored, and an address is the memory-address of the data stored in a register. Registers are commonly referred to as a type of computer memory built directly into the processor or Central Processing Unit (CPU) that is used to store and manipulate data during the execution of instructions (Hopkins, 2023). Furthermore, I will discuss fixed partitioning, dynamic partitioning, and the buddy system which are methods to allocate memory contiguously.

To begin, memory is an essential component of a computer system's functionality. Furthermore, "Computer memory is organized into at least two levels, referred to as main memory and secondary memory" (Stallings, 2018, p. 315). Main memory is also called system memory or Random-Access Memory (RAM). RAM is volatile, meaning that it does not provide permanent storage, i.e., if your turn a computer off, the data stored in the computer RAM will be lost. Moreover, the CPU utilizes the RAM to store and quickly access data necessary to execute processes. The secondary memory, on the other hand, is slower than the RAM and it is mainly utilized as long-term storage. In the context of virtual memory, second memory can also be utilized to store and access data necessary to execute processes.

* Physical memory refers to the actual memory hardware, that is, the RAM. While it can also refer to secondary memory, for the purpose of this discussion, I will define it as the RAM.
* The logical memory, in the context of this discussion, can be defined as the CPU address register, which is where data logical addresses are stored.

Second, logical and physical addresses are sets of finite bits pointing to data stored in the RAM.

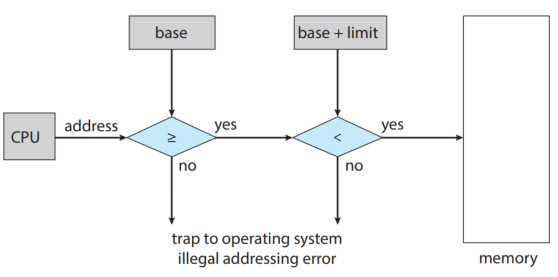
* The logical address, also called virtual address (not to be confused with virtual memory) is generated by the CPU during program execution. The set of all logical addresses generated for a program is referred to as the logical address space of the program.
* The physical address, on the other hand, points to a location in the RAM, and it is computed by the Memory Management Unit (MMU), which is a hardware component responsible for translating a logical address to a physical address. through the MMU. This safeguards or protects the physical memory from being directly accessed by a program. In other words, programs cannot directly access the data stored on the RAM, they access it indirectly The set of all physical addresses corresponding to logical addresses is referred to as the physical address space.

The figures below depict the relationship between the logical address, physical addresses, CPU, and MMU.

A diagram of a process

Description automatically generatedA diagram of a computer

Description automatically generated



(Silberschatz et al., 2018, Figure 9.5, Figure 9.2, Figure 9.1)

Different mapping schemes exist to map the logical address space to the physical address space. The example above illustrates a simple MMU scheme that is a generalization of the base register scheme. The base register or relocation register stores a base value that is used to compute the physical address from the logical address.

For example, in the first figure:

1. The CPU generated a logical address value of 346.
2. MMU generated a base address (Relocation Register) value of 14000 corresponding to the base value of the physical address space.
3. The MMU translates the logical address into the physical address utilizing the base address, here 346 + 14000 = 14346.

In summary, the CPU generates logical addresses and the MMU translates them into physical addresses to access the data stored in the physical memory.

Next, how is the RAM partitioned to store data? The two basic methods are fixed and dynamic partitioning, a third one is the buddy system which is a mix of the fixed and dynamic partitioning methods. These methods allocated memory (RAM) contiguously, that is a process is allocated a single block of memory that is contiguous or adjacent to another process block. This ensures that memory is efficiently utilized (Contiguous Memory Allocation, n.d.)

A diagram of process and process

Description automatically generatedA diagram of a process

Description automatically generatedContiguous blocks Not contiguous blocks

(Images from Stallings, 2018, p. 321)

Fixed partitioning characteristics:

* The size of the block is equal for all processes. The size is determined when the system is started up and remains the same until the system is shut down.
* Each process that is loaded into memory is assigned to a partition. If a process is larger than a single partition, it must be divided into smaller pieces that fit into separate partitions.
* One downside to fixed partitioning is the potential for memory waste, known as internal fragmentation. This happens when any program, no matter how small, occupies an entire partition. This can cause internal fragmentation.
* Since the number of partitions is fixed, the number of processes that can be loaded into memory at the same time is limited by the number of partitions.

Dynamic partitioning characteristics:

* The size of the partition isn't fixed. When a process needs to be loaded into memory, a partition just large enough to hold the process is allocated.
* Often results in more efficient use of memory because it minimizes internal fragmentation. Since each partition is just large enough to hold its process, there's less unused space within each partition.
* Despite reducing internal fragmentation, dynamic partitioning can lead to a problem called external fragmentation. This occurs when free memory is broken into small, non-contiguous blocks over time. Even if there is enough total free memory to accommodate a process, it may not be able to be allocated if the free memory isn't contiguous.
* To deal with external fragmentation, a process called compaction may be used. This involves shifting the processes in memory to make all the free memory contiguous. However, compaction can be a resource-intensive task.
* It requires efficient algorithms to allocate and de-allocate memory. The most common of these are the "first-fit" and "next-fit" algorithms, which each have their own trade-offs in terms of memory utilization and allocation speed.

The buddy system Characteristics:

* It is a mix of the fixed and dynamic partitioning methods
* Memory is divided into partitions of different sizes, but each size is a power of 2. For example, starting with a block of size 1024KB, it can be divided into smaller "buddy" blocks of sizes 512KB, 256KB, 128KB, 64KB, 32KB, 16KB, 8KB, 4KB, 2KB, and 1KB.
* When a process requests memory, the system checks for a free partition of the exact size requested. If a partition of the exact size doesn't exist, the system splits a larger partition in half (and continues splitting if necessary) until it has a partition of the correct size.
* When memory is freed, the system checks if the "buddy" of the freed partition (the adjacent partition of the same size) is also free. If it is, the system combines the two partitions into a larger partition of the next power of 2 size. This combining (or coalescing) continues until the system can't combine the freed partition with its buddy.
* The Buddy System offers a good compromise between efficient use of memory and reduction of fragmentation. It's more efficient than fixed partitioning because it can handle processes of different sizes more flexibly. And it reduces the external fragmentation that can be a problem with dynamic partitioning by merging free memory blocks.
* The buddy system, however, has its drawbacks. One of these is overheads is involved splitting and combining blocks of memory.

- Alex

I utilize my discussion posts as a base for my class notes, as a result, they can be quite lengthy.

**Reference:**

Contiguous Memory Allocation (n.d.). *Online Courses and eBooks Library*. https://www.tutorialspoint.com/contiguous-memory-allocation#:~:text=Contiguous%20memory%20allocation%20is%20a,or%20adjacent%20to%20each%20other.

Hopkins, J. (2023, May 24). What is a register in a CPU and how does it work? *Total Phase Blog*. https://www.totalphase.com/blog/2023/05/what-is-register-in-cpu-how-does-it-work/

Silberschatz, A., Galvin, P. B., & Gagne, G. (2018). *Operating System Concepts* [PDF]. Wiley. Retrieved from: https://os.ecci.ucr.ac.cr/slides/Abraham-Silberschatz-Operating-System-Concepts-10th-2018.pdf Links to an external site.

Stallings, W. (2018). *Operating Systems: Internals and design principles*. Pearson.

Afteracademy (2020, March 30). What is the difference between logical and Physical Address WRT operating system? *AfterAcademy*. https://afteracademy.com/blog/what-is-the-difference-between-logical-and-physical-address-wrt-operating-system/