Andrew Mackay

Comp 314

ID: 3441988

Operating systems assignment 3

1.

The logical address, otherwise known as virtual address, is generated by the CPU. The physical address is the one that is loaded into the memory address register or seen by the memory unit. Memory mapping can convert logical to physical addresses, as logical addresses are mapped to physical ones as well.

The advantages of allowing user programs to access the virtual address are that it allows the operating system to control the view of memory that is made available to the software. Virtual addresses are beneficial to software developers who don't know a systems exact memory address, as in this situation they don't need to know the physical memory. Virtual addresses are translated through mapping to physical addresses, where the mappings are stored in translation tables. Each application uses its own set of virtual addresses which are mapped to the corresponding physical locations. As the OS switches applications it reprograms the map, which are mapped to the correct physical location. (baeldung, 2022).

2.

Linking is to connect the modules of functions of a program for execution, whereas loading is the process of loading a program from secondary memory to main memory where it will be then executed. (developer.arm, 2022)
The dynamic loading of resources is a technique that helps utilize better memory space. Within dynamic loading the routine is only loaded when needed which is useful when sizeable amounts of code are used to handle infrequent cases.
Dynamic linking is similar to dynamic loading although in contrast linking is postponed until execution time, usually paired with C libraries and intended to reduce the shortcomings of static linking .

3.

With paging the programmer views memory as one single space, only containing one program. This is advantageous as all mapping of addresses is done by the OS out of view of the programmer.

A paging based memory scheme is based around the concept of the physical address space not needing to be contiguous, or in other words located near to or adjacent, within a process, as other memory management techniques use this as a requirement. Paging prevents memory fragmentation and the need for compaction which other contiguous methods are prone to.

Paging is done by breaking the memory into frames which are fixed sized blocks then breaking logical memory into blocks of the same size called pages. Pages are loaded into their memory frames when a process is to be executed. The backing store is then converted into the same size blocks as the frames. Paging is a form of dynamic relocation, every logical address is still bound to a physical address.

4.

The hardware cache of the page table in the simplest instance is implemented as a set of dedicated high speed hardware registers, which translates the page address efficiently. Although this increases context switch time and frequency.

The TLB otherwise known as the Translation Look-Aside Buffer is a hardware cache that is small but is used for fast look up of paged memory address translation which provides quick access to a subset of memory addresses. Functionally it has 2 parts for each entry, a value and a key tag. When associated memory is presented with an item it searches the memory for all corresponding keys at the same time for a match. When an item is found the corresponding value field is then returned, which is done very fast through the instruction pipeline.

5.

Hierarchical paging otherwise known as multilevel paging is used where some page tables may be too big so can separate into a hierarchy of several levels. 2 and 3 level page tables can be used for simplicity, as the logical address is broken into multiple page tables. There are situations where a 64 bit logical address space needs to be configured within the page table and in this instance the 2 page table will not suffice and will have to be converted to 3 due to containerizing the extra volume of bits.(studytonight, 2022).

A hashed page table is a a page table that has a hash value indicative of its page number which is hashed for quick access. The hashed page table is used to more effeciently solve the problem of managing addresses greater than 32 bits by linking a list of elements that hash to the same location. Each element contains a virtual page number, a value of the mapped page frame and a pointer to the next element of the list.

6.

Demand paging requires much hardware support by means of a TB/ address translation mechanism, page table entries plus disk addresses, ability to detect a page fault, restartable instructions .Instead of swapping all pages at once when a context switch occurs ,instead with demand paging loading or storing any page is done only

until it is absolutely necessary.  Handling page faults requires much planning and can take a serious toll on the systems performance if page demands cause many faults and the system is trying to constantly handle them. (cseweb, 2022) Pages that are never needed are not loaded into memory at all, which creates extra space compared to the amount of space the unneeded pages would take up comparatively.

7.

Swapping or replacement is to move a process or portion of a process temporarily out of memory to a backing store, and then later brought back into memory to continue execution. With swapping the processes that occupy the memory may exceed the total memory limit while still performing their duties which will provide greater overall performance.
Swapping with paging is not commonly used anymore as it is generally deemed inefficient due to the amount of time required to the backing store being prohibitive in combination with the extra time frame it takes to move entire processes between memory.

8.

The minimum amount of frames is defined by architecture, but the maximum is defined by physical space. When allocating frames this is something that needs to be considered.

9.

Adding paging to an existing architecture may prove to be problematic for the systems performance as much careful consideration is placed around implementation of paging within the system. Programming certain data structures can prevent problems with paging whereas not keeping them in mind will create much slowdown and possible page faults, where they need to be programatically dealt with. Page faults are never the less inevitable but they can be reduced significantly by precaution of careful selection of data structures which increases locality lowering the page fault rate, but also the number of pages to sort through in the local set. A hash table when dealing with paging in a pre designed system for example would prove to be counter intuitive as hash tables scatter references producing bad locality thus affecting the amount of page faults and time spent sifting through page tables.

10.

When a process is trying to access a page which isn't in the memory, it can accept a new page then allocate it a frame from the set of frames, and if that frame is currently allocated to a different process it can still be taken and accepted to a different process. This doesn't hinder performance of processes, and can be beneficial

to performance. (geeksforgeeks, 2018) Global replacement is compared with local replacement as they are 2 techniques to deal with allocation of frames but differ in the sense that local allocation can only accept new frames from only its own set of allocated frames.

11.

Disk scheduling is also known as I/O scheduling. This is implemented due to I/O requests arriving by different processes while only one can be served at a time by the disk controller, multiple requests may be far from eachother which results in greater disk arm movement, one of the slowest parts of the computer is the hard drive therefor has to be accessed efficiently.(geeksforgeeks,2015) For HDDS access time has two major components, the seek time, and rotational latency. There are multiple disk scheduling algorithms which can be implemented and they are broken down into FCFS - first come first serve , SSTF - shortest seek time first, SCAN, CSCAN, LOOK, CLOOK, RSS, LIFO, N-STEP LOOK, and FSCAN. With there being so many different algorithms that all have different approaches they can be assumed to have different results as well whereas the advantages of FCFS are every request gets a fair chance, and no indefinite postponement, the disadvantages are that it does not optimize seek time, and may not provide the best possible service so we have many other algorithms as listed above to sift through to try and decide which will best suit our needs.

12.

Error detection and correction are somewhat self explanatory by definition but by taking a deeper dive we can look at how errors are detected and how they are corrected.  This process is fundamental to various areas of computing such as networking and storage, and memory. By detecting an issue the system can halt an operation before further damage is done, report the error to the admin, or warn that a internal device may be failing. Parity uses a hash function to detect multiple bit errors and is a form of checksum functions which use modular arithmetic to store, compute and compare values. Parity may be used to detect errors within single (0-1) bits but can be faulty if encountering a double bit error as this will go undetected.

Error correction code seeks to both detect and correct the problem. Supplemental amounts of storage coupled with specified algorithms can achieve error correction effectively. The ECC is written with a value calculated from all the bytes in the data being written, this value is then recalculated after the page is read and compared to the original predicted value, if there is a difference than there was an error with the data storage. If only a few bits of data were corrupted during this comparison the ECC can identify and correct them accordingly.

13.

The directory structure is of a higher level of organizational hierarchy than files and folders in the way that the directory is thought of as a file with information

within it, while the directory structure is the overarching relational database of all directories and how they are organized. The directory structure keeps info on the entries of all these related files like type, filename, location and user.(binaryterms, 2019) Directories can be updated by the user in such categories as searching, creating, deleting, renaming, listing and can be organized in many different ways. The simplest directory structure is the single level directory, although 2 and 3 level directories also exist, along with acyclic graph directories, and general graph directories.

An example of a faulty directory structure is my own file system. Within my directory structure it uses cloud based storage involuntarily as I haven't purchased the full version of windows 10. So within the directory structure all things within the C: drive are actually stored like OneDrive/C: and can only be referenced like this, although I was able to manually override this through much frustration for certain programs, it largely has stayed like this even with an IT workers attempt at fixing it.

14.

Buffering is the transfer of data between a device and application, which can be thought of a type of loading. Buffering is done for 3 separate reasons which are the mismatch in speed between producer and consumer of data stream, providing adaptions to devices for different data transfer sizes, and to support the copying of semantics for application input and output.

Cache is a area of fast memory which holds copies of data, in which the data accessed from a cache is a copy of the data it represents. A buffer may also hold copy of an item but the cache has its copy on faster storage which is stored separately.

Spooling is partly a buffer but it differs in the sense that it holds output for a device that can't accept interleaved data streams. When multiple processes are trying to access a hardware device at once these processes are spooled in a sort of queue.

15.

The advantages of placing functionalities within a device controller vs in the kernal are that the kernal is freed of a lot of responsibilities and this will allow it to perform all other activities much better, the whole system will crash less often as there will be less bugs overall.

The disadvantages of placing functionalities within a device controller are that there would be a massive decrease in performance as embedded algorithms could conflict with a programs use of the device as well as improving said algorithms would need constant updating of the hardware, and the bugs would be harder to fix overall.

16.

Disk scheduling algorithms should be written as a separate module within the operating system, the file allocation method influences which technique to choose aka if contiguous files chosen then FCFS(first come first serve) is best suited as it is designed to handle this situation best. The location of directories and index blocks is important as well. The disk head has to move the entire width of the disk if the directory entry is on the first cylinder and the data on the final one. Main memory catching of index blocks can also help reduce disk arm movement.

Keeping in mind that every Disk-Scheduling algorithm has unique properties, advantages and disadvantages. Something like SCAN and C-SCAN have less of a starvation problem and are better at handling systems that have more read and write operations from the disk and are considered a heavier load.

References

Silberschatz, A., Gagne, G., & Galvin, P. B. (2018). Operating System Concepts (10th Edition). Wiley Global Education US. https://online.vitalsource.com/books/9781119320913

developer.arm.com. 2022. Virtual and physical addresses. [online] Available at: <https://developer.arm.com/documentation/101811/0102/Virtual-and-physical-

addresses#:~:text=The%20benefit%20of%20using%20virtual,are%20permitted%20to%20that%20memory.> [Accessed 1st August 2022].

baeldung.com. 2022. Dynamic Linking vs. Dynamic Loading. [online] Available at: <https://www.baeldung.com/cs/dynamic-linking-vs-dynamic-loading> [Accessed 1st August 2022].

Structure of Page Table in Operating Systems - Studytonight. (2022). Retrieved 1st August 2022, from https://www.studytonight.com/operating-system/structure-of-page-table-in-operating-systems

Demand Paging. (2022). Retrieved 1st August 2022, from https://cseweb.ucsd.edu/classes/sp17/cse120-a/applications/ln/lecture13.html

Allocation of frames in Operating System - GeeksforGeeks. (2018). Retrieved 2nd August 2022, from https://www.geeksforgeeks.org/operating-system-allocation-frames/

Disk Scheduling Algorithms - GeeksforGeeks. (2015). Retrieved 2nd August 2022, from https://www.geeksforgeeks.org/disk-scheduling-algorithms/

T, N. (2019). What is Directory Structure in Operating System (OS) ? its types with diagram - Binary Terms. Retrieved 22 August 2022, from https://binaryterms.com/directory-structure-in-os.html

Selection of a Disk-Scheduling Algorithm - Educate. (2020). Retrieved 20th August 2022, from https://educatech.in/selection-of-a-disk-scheduling-algorithm/