OPERATING SYSTEMS

Unit1_Unit2_Unit3:Revision Class #6 Typical Concepts and QnAs related to Operating Systems

Nitin V Pujari
Faculty, Computer Science
Dean , IQAC, PES University

OPERATING SYSTEMS

Course Syllabus => Unit 1

1	Introduction: What Operating Systems Do, Computer-System Organization	1.1, 1.2	21.4
2	Computer-System Architecture, Operating-System Structure & Operations	1.3,1.4,1.5	
3	Kernel Data Structures, Computing Environments	1.10, 1.11	
4	Operating-System Services, Operating-System Design and Implementation	2.1, 2.6	
5	Process concept: Process in memory, Process State, Process Control Block, Context switch, Process Creation & Termination,	3.1 - 3.3	
6	CPU Scheduling - Preemptive and Non-Preemptive, Scheduling Criteria, FIFO Algrorithm	5.1-5.2	
7	Scheduling Algorithms:SJF, Round-Robin and Priority Scheduling	5.3	
8	Multi-Level Queue, Multi-Level Feedback Queue	5.3	
9	Multiprocessor and Real Time Scheduling	5.5, 5.6	
10	Case Study: Linux/ Windows Scheduling Policies.	5.7	
11	Inter Process Communication - Shared Memory, Messages	3.4	
12.	Named and unnamed pipes (+Review)	3.6.3	

OPERATING SYSTEMS

Course Syllabus => Unit 2

13	Introduction to Threads, types of threads, Multicore Programming, Multithreading Models	4.1 - 4.3	42.8
14	Thread creation, Thread Scheduling	5.4	
15	Pthreads and Windows Threads	4.4	
16	Mutual Exclusion and Synchronization: software approaches,	6.1-6.2	
17	principles of concurrency, hardware support	6.3-6.4	
18	Mutex Locks, Semaphores	6.5, 6.6	
19	Classic problems of Synchronization: Bounded-Buffer Problem, Readers-Writers problem	6.7-6.8	
20	Dining-Philosophers Problem	6.8	
21	Synchronization Examples: Synchronisation mechanisms provided by Linux/Windows/Pthreads.	6.9	
22	Deadlocks: principles of deadlock, Deadlock Characterization	7.1-7.3	
23	Deadlock Prevention, Deadlock example	7.4-7.5	
24	Deadlock Detection, Algorithm	7.6	

OPERATING SYSTEMS

Course Syllabus => Unit 3

25	Main Memory: Hardware and control structures, OS support, Address translation	8.1	64.2
26	Dynamic linking, Swapping	8.2	
27	Memory Allocation (Partitioning, relocation), Fragmentation	8.3	
28	Segmentation	8.4	
29	Paging: OS Support, TLBs, Address Translation	8.5	
30	Structure of the Page Table	8.6	
31	Design Alternatives - Inverted Page Tables, Bigger Pages	8.7-8.8	
32	Virtual Memory: Demand Paging, Copy-OnWrite	9.1-9.3	
33	Page replacement policy - LRU etc. (In comparison with FIFO and Optimal)	9.4	
34	Page Replacement (contd.), Frame allocation	9.4,9.5	
35	Thrashing	9.6	
36	Case Study: Linux/ Windows Memory Management	9.10	

OPERATING SYSTEMS

Do you know the meaning of the following w.r.t Operating Systems ?

- Process heap
  - Part of a process that stores dynamically allocated data.
- Process text.
  - Part of process that stores the instructions (the code).
- PCB (Process Control Block)
  - A data structure in the operating system that stores information about a process needed for scheduling.
- PCB information
  - Process state, Process number, program counter, stack pointer, registers with more.
- PCB Process state
  - Part of the PCB that stores information about the state of a process. (new, ready, running, waiting, dead)
- PCB Process number
  - Part of the PCB that stores the process identification (PID).

OPERATING SYSTEMS

Do you know the meaning of the following w.r.t Operating Systems ?

- PCB Program counter
  - Part of the PCB that stores the address for the next instruction in a process.
- PCB Registers
  - Part of the PCB that stores the values in the CPU-register.
- Context switch
  - The CPU switches to a different process and must first store the previous process state in the PCB.
- Zombie
  - A process that has terminated, but whose parent has not called wait().
- Orphan
  - A child whose parent is terminated.
- Long-term scheduler
  - Selects which processes should be brought from a storage device to the ready queue.

OPERATING SYSTEMS

Do you know the meaning of the following w.r.t Operating Systems ?

- Short-term scheduler
  - Selects which process from the ready queue that should be executed next.
- I/O-bound process
  - Spends more time doing I/O than computations, many short CPU bursts.
- CPU-bound process
  - Spends more time doing computation and generates I/O requests infrequently.
- Thread
  - A single execution stream within a process. with it's own program counter, stack, and registers.
- Kernel thread
  - The thread is controlled by the OS. The OS sees the different threads as different processes.
- Thread library
  - An API for creating and managing threads.

Do you know the meaning of the following w.r.t Operating Systems ?

- many-to-one
  - Many user-level threads mapped to single kernel thread.
- Asynchronous cancellation
  - Terminates the target thread immediately.
- Deferred cancellation
  - Allows the target thread to periodically check if it should be cancelled.
- one-to-one
  - Some user-level thread mapped to their own kernel thread.
- User-level thread
  - The thread is controlled by a thread library. The OS does not know about the thread.
- Thread-local storage
  - Allows each thread to have its own copy of data

Do you know the meaning of the following w.r.t Operating Systems ?

- Thread pool
  - Creating a number of threads and place them in a queue they will then wait for work. When a thread completes its service it returns to a queue and wait for more work.
- fork()
  - Creates a new process, with the same data as the process that called fork(). (code, pointers etc.)
- exec(program p)
  - Replaces the current process with a process with the code and data from the program p.
- sleep()
  - Suspends execution for at least the specified time.
- waitpid(int p)
  - Waits for the process with pid p to finish execution.
- wait()
  - Waits for the child process to finish execution.

OPERATING SYSTEMS

Do you know the meaning of the following w.r.t Operating Systems ?

- kill(int p)
  - A process kills the process with pid p.
- exit()
  - A process kills itself.
- Nonpreemptive
  - Once a process is in the running state, it will continue until it terminates or blocks itself for I/O
- Preemptive
  - Once a process is in the running state it can be interrupted for different purposes. (such as switching to a process with higher priority)
- CPU utilization
  - Keeping the CPU as busy as possible
- Waiting time
  - The amount of time a process has been waiting in the ready queue.

OPERATING SYSTEMS

Do you know the meaning of the following w.r.t Operating Systems ?

- Throughput
  - The amount of work performed by a system during a given period of time.
- Turnaround time
  - Amount of time to execute a particular process including the waiting time.
- Average turnaround time
  - Execution time plus wait time for every process divided by total amount of processes.
- FCFS (First Come First Served)
  - Gives the CPU to the first process in the ready queue.
- SJF (nonpreemptive)
  - Gives the CPU to the process that has the shortest burst time in the ready queue.
- SJF (preemptive)
  - Gives the CPU to the process that has the shortest burst time and switches to another process in the ready queue if it has a shorter burst time

Do you know the meaning of the following w.r.t Operating Systems ?

- Priority Scheduling (nonpreemptive)
  - Gives the CPU to the process that has the highest priority.
- Priority Scheduling (preemptive)
  - Gives the CPU to the process that has the highest priority and switches to another process in the ready queue if it has a higher priority.
- Aging
  - Giving a process with low priority higher priority if it is has been waiting in the ready queue too long without being executed.
- Round Robin (FCFS, Preemptive)
  - Gives the CPU to the first process and gives it one time quantum. When the timer runs out the process gets places last in the ready queue.
- Multilevel Queue Scheduling
  - Partitioning a ready queue into multiple queues where each queue could have its own scheduling and the queues have different priority over each other.

Do you know the meaning of the following w.r.t Operating Systems ?

- Multilevel Feedback Queue Scheduling
  - Multilevel queues that allow processes to move between them, to separate processes by the characteristics of their CPU bursts (form of ageing preventing starvation).
- Semaphore
  - A variable or data type used to control access to a common resource by multiple processes.
- Binary Semaphore
  - Integer value can range only between 0 and 1. (same as mutex lock)
- Counting Semaphore
  - Integer value can range over an unrestricted domain.
- wait(semaphore S)
  - Waits for the semaphore S to give the signal to continue.
- signal(semaphore S)
  - Gives a signal to the semaphore S.

Do you know the meaning of the following w.r.t Operating Systems ?

- Deadlock
  - A number of processes that are all waiting for a resource provided by some other process making it so that the processes stand still.
- Critical Section
  - A section of code within a process that requires access to shared resources and that must not be executed while another process is in a corresponding section of code.
- Mutual Exclusion
  - If process P is executing in its critical section, then no other processes can be executing in their critical sections.
- Circular wait condition
  - A number of processes that are all waiting for the next process is a circular queue to signal their semaphore.
- No preemption condition
  - A resource can only be provided by the process that holds it. Another thread can not force the thread to release the resource.

Do you know the meaning of the following w.r.t Operating Systems ?

- Hold and wait condition
  - A process holds at least one resource and is waiting for resources from other processes.
- The four conditions that if met can cause a deadlock.
  - Mutual exclusion, hold and wait, no preemption, circular wait.
- Deadlock detection
  - Finds instances of deadlock when threads stop making progress and tries to recover.
- Deadlock prevention
  - Provides methods for ensuring at least one of necessary conditions for deadlocks cannot hold.
- Deadlock avoidance
  - Requires that the operating system be given additional information in advance concerning which resources a process will request and use during its lifetime.

Do you know the meaning of the following w.r.t Operating Systems ?

- Memory Management Unit (MMU)
  - Hardware device that at run time maps virtual to physical address.
- Virtual address
  - An address that corresponds to a location in virtual space and is translated by address mapping to a physical address when memory is accessed.
- Physical address
  - An actual address in the main memory device seen by the MMU.
- First-fit memory allocation
  - A main memory allocation scheme that searches from the beginning of the free block list and selects for allocation the first block of memory large enough to fulfill the request.
- Best-fit memory allocation
  - A main memory allocation scheme that considers all free blocks and selects for allocation the one that will result in the least amount of wasted space

OPERATING SYSTEMS

Do you know the meaning of the following w.r.t Operating Systems ?

- Worst fit memory allocation
  - A main memory allocation scheme that considers all free blocks and selects for allocation the smallest one.
- External Fragmentation
  - As processes are loaded and terminated, the memory holes between the processes get broken down into smaller and smaller holes resulting in wasted memory.
- Internal Fragmentation
  - Allocated memory may be slightly larger than the process requested resulting in wasted memory.
- Virtual memory
  - A memory management scheme which a computer stores and retrieves data from secondary storage for use in main memory creating the illusion that the computer has more memory.

Do you know the meaning of the following w.r.t Operating Systems ?

- Paging
  - A memory management scheme which follows virtual memory by retrieving data from secondary storage in same-size blocks called pages.
- Demand Paging
  - An extension to paged memory management in which pages are brought into memory only when referenced (on demand).
- Page (Paging)
  - A fixed-length contiguous block of virtual memory.
- Frame (Paging)
  - A fixed-length contiguous block of physical memory into which memory pages are mapped by the operating system.
- TLB (translation look-aside buffer)
  - A fast memory that stores page numbers and frames.

OPERATING SYSTEMS

Do you know the meaning of the following w.r.t Operating Systems ?

- Swapping (Process)
  - Roll out a process to disk, releasing the memory it holds. (Can be used when a process is doing I/O)
- Thrashing
  - A condition of excessive paging in which the operating system becomes sluggish.
- File Control Block (FCB)
  - A file system structure in which the state of an open file is maintained.
- File Metadata
  - Metadata that describes interesting properties about the files that are not related to the main content of the file
- inode
  - Short for "information node". In UNIX/Linux, a system for storing key information about files.

Do you know the following w.r.t Operating Systems ?

- If processes P and Q wish to communicate, they need to establish a communication link between them exchange messages via send/receive
- IPC can be implemented in 2 possible ways Physical, Logical
- Physical IPC implementations => shared memory, hardware bus, network
- Logical IPC implementations => direct/indirect, synchronous/asynchronous, automatic or explicit buffering
- Properties of direct communication link
  - links are established automatically
  - a link is associated with exactly one pair of communicating processes
  - between each pair there exists exactly one link.
  - the link may be unidirectional, but is usually bi-directional
- messages are directed and received from this in indirect communication => mailboxes or ports

Do you know the following w.r.t Operating Systems ?

- Properties of indirect communication link
  - established only if processes share a common mailbox.
  - a link may be associated with many processes.
  - each pair of processes may share several communication links.
  - link may be unidirectional or bi-directional.
- message passing may be either blocking or non-blocking
- Blocking is considered as synchronous
- Non-Blocking is considered as asynchronous
- Message Buffering means queue of messages attached to the link
- Implementations of message buffering
  - zero capacity
  - bounded capacity
  - unbounded capacity

Do you know the following w.r.t Operating Systems ?

- Zero Capacity Buffering => no messages are queued on a link, sender must wait for receiver
- Bounded Capacity Buffering => finite length of n messages, sender must wait if link full
- unbounded capacity buffering
- Socket => an endpoint for communication
- one To one multithreading model
    - each user-level thread maps to kernel thread.
    - creating a user-level thread creates a kernel thread.
    - more concurrency than many-to-one.
    - number of threads per process sometimes restricted due to overhead
- many to one multithreading model
    - many user-level threads mapped to single kernel thread.
    - one thread blocking causes all to block.
    - multiple threads may not run in parallel on multicore system because only one may be in kernel at a time.
    - few systems currently use this model.

Do you know the following w.r.t Operating Systems ?

- many-to-many multithreading model
  - allows many user level threads to be mapped to many kernel threads.
  - allows the OS to create a sufficient number of kernel threads.
- Parallel regions => blocks of code that can run in parallel
- Default type of cancellation => deferred cancellation
- Light Weight Process => an intermediate data structure between user and kernel threads
- atomic => non interruptible
- block (waiting queue operation) => place the process invoking the operation on the appropriate waiting queue
- wakeup (waiting queue operation) => remove one of the processes in the waiting queue and place it in the ready queue
- starvation => indefinite blocking

Do you know the following w.r.t Operating Systems ?

- priority inversion => Scheduling problem when lower-priority process holds a lock needed by higher-priority process
- Dispatcher => module gives control of the CPU to the process selected by the short-term scheduler
- Dispatch latency => time it takes for the dispatcher to stop one process and start another running CPU utilization
- Lazy swapper => Pager that never brings in a page into memory unless it is needed.
- Zero-fill-on-demand => Pages are zeroed out before being allocated, for security reasons
- Belady's anomaly => The page fault rate may increase for some page replacement algorithms, as the number of page frames increases.
- Stack algorithms => Class of algorithms for page-replacement, that never suffer from Belady's anomaly (e.g. LRU).

Do you know the following w.r.t Operating Systems ?

- Locality model =>A process moves from locality (set of actually used pages) to locality during execution time.
- Working-set model => Based on the assumption of locality, using a moving working-set window. Prevents trashing, while allowing the highest degree of multiprogramming.
- Working set window => Set of recently moving pages, is a moving set. Based on locality.
- The Dining Philosopher's Problem is a standard test case for evaluating approaches in implementing Synchronization
- The type of memory that allows for very effective multiprogramming and relieves the user of memory size constraints is referred to as => Virtual memory
- Paging, which is transparent to the programmer, eliminates external fragmentation, provides efficient use of main memory, and has pieces of fixed, equal size.
- Segmentation, which is visible to the programmer, has the ability to handle growing data structures, modularity, and support for sharing and protection.

Do you know the following w.r.t Operating Systems ?

- Each entry in a Page Table contains control bits and the corresponding frame number if the page is resident in memory.

- In a segmentation system, each entry in a segment table contains control bits and the starting address and the length of the segment.

- Use a multiprocessor
    - To support multiprogramming. Large numbers of independent processes. To support parallel programming
    - Job consists of multiple threads and/or processes. Cooperating/communicating threads/processes. Not independent!

- The scheduling function should perform these five tasks:
    - 1. Share time fairly among processes
    - 2. Prevent starvation of a process
    - 3. Use the processor efficiently
    - 4. Have low overhead
    - 5. Prioritise processes when necessary (e.g. real time deadlines)

Do you know the following w.r.t Operating Systems ?

- Long-term scheduling is performed when a new process is created. This is a decision whether to add a new process to the set of processes that are currently active. (its stages include new & exit)
- Medium-term scheduling is a part of the swapping function. This is a decision whether to add a process to those that are at least partially in main memory and therefore available for execution. (its stages include Blocked Suspend & Ready Suspend)
- Short-term scheduling: the actual decision of which ready process to execute next. Its stages include running & ready & blocked
- Real-Time Scheduling
    - Correctness of the system depends not only on the logical result of the computation but also on the time at which the results are produced
    - Tasks or processes attempt to control or react to events that take place in the outside world
    - These events occur in "real time" and tasks must be able to keep up with them

OPERATING SYSTEMS

Do you know the following w.r.t Operating Systems ?

- Memory Hierarchy
  - The concept in which computers have a few megabytes of very fast, expensive, volatile cache memory, and a few gigabytes of medium-speed, medium-priced, volatile main memory, and a few terabytes of slow, cheap, nonvolatile magnetic or solid-state disk storage, not to mention removable storage, such as DVDs and USB sticks.

- Memory Manager
  - Used to keep track of which parts of memory are in use, allocate memory to process when they need it, and deallocate when they are done.

- Static Relocation
  - Used when two programs both reference absolute physical memory. Fixed by modifying the second program on the fly as it is loaded into memory.

- Address Space
  - The set of addresses that a process can use to address memory.

- Dynamic Relocation
  - Maps each process' address space onto a different part of physical memory in a simple way.

OPERATING SYSTEMS

Do you know the following w.r.t Operating Systems ?

- Base
  - The register that is loaded with the physical address where its program begins in memory.
- Limit
  - The register that is loaded with the length of the program.
- Swapping
  - Used to deal with memory overload. Brings in a process, runs it, then puts it back. This is like running processes one at a time.
- Virtual Memory
  - Used to deal with memory overload. Allows programs to run even when they are only partially in main memory.
- Memory Compaction
  - When swapping creates multiple holes in memory, this technique moves all processes downward / upward as the case may be as far as possible.

Do you know the following w.r.t Operating Systems ?

- Associative Memory
  - Usually inside the MMU and consists of small number of entries. Each entry contains information about one page, including the virutal page number, a bit that is set when the page is modified, the protection code, and the phycial page frame in which the page is located.
- Soft Miss
  - Occurs when a page referenced is not in the TLB, but is in memory.
- Hard Miss
  - Occurs when the page itself is not in memory, and also not in the TLB.
- Page Table Walk
  - Looking up the mapping in the page table hierarchy.
- Minor Page Fault
  - The data is in the page frame/memory, but not in the TLB, so it's still easy to access.
- Major Page Fault
  - The data is not in the page frame/memory or in the TLB, so the program needs to go to the disc, which takes time.

Do you know the following w.r.t Operating Systems ?

- Segmentation Fault
  - The data is not in the page frame/memory, not in the TLB, and not in the disc. In this case that the operating system typically kills the program.
- Multilevel Page Table
  - A partitioned page table that looks like a tree. It is used in computers with large amounts of memory so that it can be split up.
- Page Directory
  - A page table of page tables that contain page frames. This gives more flexibility in memory, allowing for more addressable memory.
- Page Directory Pointer Table
  - Extended the size of each page table from 32 bit to 64 bit, allowing for the total memory to exceed the 4GB boundary.
- Inverted Page Table
  - In this design, there is one entry per page frame in real memory, rather than one entry per page of virtual address space. Th entry keeps track of which (process, virtual page) is located in the page frame.

Do you know the following w.r.t Operating Systems ?

- Local page replacement
  - When a page fault occurs, a page from the same process should be replaced. These algorithms effectively correspond to allocating every process a fixed fraction of the memory.

- Global page replacement
  - When a page fault occurs, a page from any of the given processes may be replaced even between competing programs. These algorithms for page replacement dynamically allocate page frames among the runnable processes.

- PFF (Page Fault Frequency)
  - An algorithm that tells when to increase or decrease a process' page allocation but has nothing about which page to replace on a fault. It just controls the size of the allocation set.

- Paging Daemon
  - A background process that sleeps most of the time, but awakens periodically to inspect the state of memory. It is used to ensure a plentiful supply of free page frames in the paging system.

THANK YOU

Nitin V Pujari
Faculty, Computer Science
Dean, IQAC, PES University

nitin.pujari@pes.edu

For Course Deliverables by the Anchor Faculty click on www.pesuacademy.com