SYNERGY SCHOOL OF ENGINEERING, DHENKANAL

LECTURENOTES

ON

OPERATING SYSTEM

Compiled by

Ms.Aparna Priyadarshini

Lecturer, Department of Computer Science & Engineering,
Synergy School of Engineering, Dhenkanal

CONTENTS

UNIT-1
INTRODUCTION

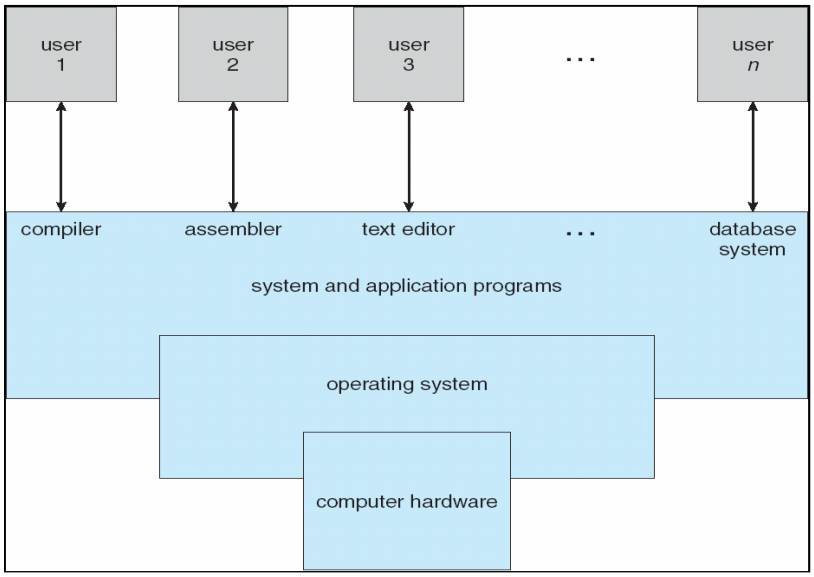
INTRODUCTION:

- Operating system is a system software that acts as an intermediary between the user of a computer and computer hardware.
- Itisconsideredasthebrainofthecomputer.
- Itcontrolstheinternalactivitiesofthecomp.hardwareandprovidestheuserinterface.
- Thisinterfaceenablesausertoutilizethehardwareresourcesveryefficiently.
- Itisthefirstprogramthatgetsloadedintothecomputermemorythroughtheprocess called "booting".

COMPONENTSOFCOMPUTERSYSTEM:

Ingeneral,wecandivideacomputersystemintothefollowingfourcomponents.

- Hardware
- Operatingsystem
- Applicationprograms
- Users



- Aswecanseeinthefiguretheuserinteractswiththeapplicationprograms.
- Theapplicationprogramdonotaccessthehardwareresourcesdirectly.
- HARDWARE resources include I/O devices, primary memory, secondary memory(hard disk, floppy disk etc.)and the microprocessor.
- SotheoperatingSystemisrequired to accessandusetheseresources.
- Theapplicationprogramsareprogrammedinsuchawaythattheycaneasily communicate with the resources.
- An operatingSystem is the firstprogram thatis loaded intointo the computers main memory,when acomputer is switched on.
- Somepopular operatingSystemarewindows9x(95,98),Linux,Unix, Windowsxp. vistaetc.

OBJECTIVESOFOPERATINGSYSTEM

Operatingsystemhasthreemain objectives

- **Convenience**:Anoperatingsystemmakesacomputersystemconvenientandeasytouse,fortheuser.
- **Efficiency:** Anoperatingsystemallowsthecomputerto usethecomputer hardwareinan efficient way, by handling the details of the operations of the hardware.
- **Ability to Evolve:** Anoperating systemshould be constructed in such a wayas to permit theeffectivedevelopment,testing,andintroductionofnewsystemfunctionswithoutatthe same time interfering with service.

NEEDSANDSERVICESOFOPERATINGSYSTEM/FUNCTIONSOFOPERATINGSYSTEM

OperatingSystemperformsanumberoffunctionsforthecomputersystemthatarefollows:

1) ***It actsasanCommandInterpreter:***
   - GenerallytheCPUcannotunderstandthecommandsgivenbyuser.Itisthefunction of operating System to translate this command (human understandable) into m/c understandable instructions that the system (CPU) can understand.
   - AftertheexecutionofinstructionsbyCPU,itretranslatestheo/pbackintoahuman understandable language.
   - Toexecutetheuserjobs,theOperarting Systeminteracts withthecomputerhardware.
2) ***ItactsastheResourceManager:***
   - AnOperating Systemactsasaresourcemanagerintwoways
     - Timemultiplexing
     - Spacemultiplexing
   - Intimemultiplexingthedifferent resources(hardwareorsoftware)canbeshared among different users for a optimal or fixed time slot.
   - e.g.the Operating Systemallocatesaresourcessuchas CPU to programAfor a fixed time slot. When the time slot ofprocess A is over, the CPU is allocated to another program
     B. If program A needs more CPU attention, then the CPU again allocated to program A after the time slice period allocated to program B is over.
   - In space multiplexing, different resources are shared at the same time among differentprograms.e.g.sharingofharddiskandmainmemorybydifferentusers at the same time.
3) ***MemoryManagement:***
   - Itkeepstrack of theresources(memory),whatpartof memory isinuseandby whom, which part of the memory is not in use.
   - Decideswhichprocessesaretobeloadedwhenmemoryspaceisavailable.
   - Allocationand de allocationofmemory
4) ***ProcessManagement:***
   - Aprocess(task)isaninstanceofaprograminexecution.Aprogramisjusta passive entity, but a process is an active entity.
   - To accomplish its task ,a process needs certain resources like CPU time, memory,files and I/O devices.
   - Theseresources areallocatedto process either atthetimeofcreationorwhenit is executing.

- Theoperatingsystemisresponsible forthefollowingfunctionsrelatedtoprocess management.
    i. Process creation (loading the prog. From secondary storage to main memory)
    ii. Processscheduling
    iii. Providemechanismforprocesssynchronization
    iv. Providemechanismfordeadlockhandling
    v. Processtermination

5) ***PeripheralorI/OdeviceManagement:***
    - Keeptrackofresources(device,channels,controlunits)attached tothe system.
    - CommunicationbetweenthesedevicesandCPUisobservedbyoperatingsystem.
    - An operating system will have device drivers to facilitateI/O functions involving devicelikekeyboard, mouse,monitor,disk,FDD,CD-ROM,printer etc.
    - Allocationand Deallocationofresourcestoinitiate I/Ooperation.
    - Othermanagementactivitiesare
        i. Spooling
        ii. Caching
        iii. Buffering
        iv. Devicedriverinterface

6) ***FileManagement:***
    - Afileisacollectionofrelatedinformationorrecorddefinedbythe user.
    - Theoperatingsystemisresponsibleforvariousactivitiesoffile managementare
        i. Creationanddeletionoffiles
        ii. Creationanddeletionofdirectories
        iii. Manipulationoffilesand directories
        iv. Mapping filesontosecondarystorage

7) ***SecondarystorageManagement:***
    - Itisalargermemoryusedtostorehugeamountofdata.Itscapacityismuch larger than primary memory. E.g. floppydisk, hard disk etc.
    - Theoperatingsystemisresponsible forhandlingallthedevicesthat canbedone by thesecondary storage management.
    - Thevariousactivities are:
        i. Freespacemanagement
        ii. Storage allocation(allocation of storage space when new files have to bewritten).
        iii. Diskscheduling(schedulingtherequestformemoryaccess)

8) ***Protection/SecurityManagement:***
    - If a computer system has a multiple processor, then the various processes must beprotected of one another's activities.
    - Protection refers to mechanism for controlling user access of programs orprocesses or user to resources defined by the computer system.

9) ***ErrordetectionandRecovery:***
    - Errormayoccurduringexecutionlikedividebyzerobyaprocess,memory access violation, deadlock, I/O device error or a connection failure.
    - Theoperatingsystemshoulddetectsucherrorsandhandles them.

CLASSIFICATION/TYPESOFOPERATING SYSTEM

All operatingSystem consists of similar component and can perform all most similar function but the methodand procedure for performing these functions are different.

OPERATING SYSTEM are classified into different categories according to their differentfeatures. Thefollowingsectionwilldiscusstheclassificationofoperatingsystem.

SingleuserOPERATING SYSTEM:

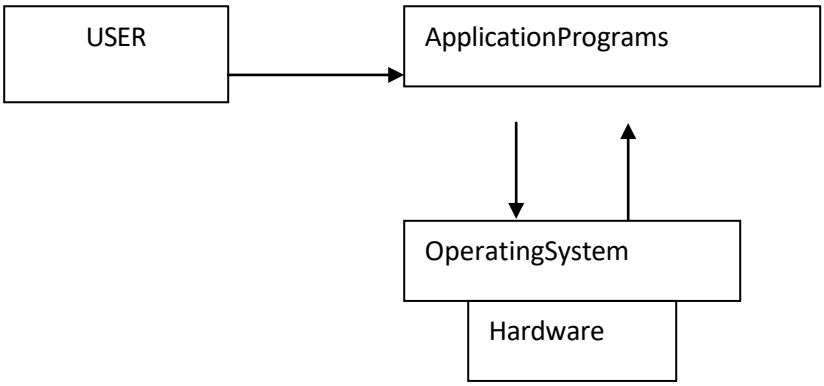
- Inasingleuseroperatingsystemasingleuser canaccessthecomputerataparticulartime.
- Thissystemprovidesallthe resourcestothe userat allthe time.
- ThesingleuseroperatingSystemisdividedintothefollowingtypes.
    - Singleuser,singletaskingoperatingSystem
    - Singleuser,multitaskingoperatingSystem

Singleuser,singletaskingoperatingSystem:

- Inasingleuser,singletaskingoperatingsystem,Thereisasingleusertoexecutea programataparticular system.
- Example–MS-DOS

Singleuser, multitaskingoperating System

- Insingleuser,multitaskingOPERATINGSYSTEMasingleusercanexecutemultiple programs.
- Example–Ausercanprogramdifferentprogramssuchasmakingcalculationsinexcelsheet, printingaworddocument&downloadingintothefilefrominternetatthesametime.



Advantage:

- TheCPU hasto handle onlyone application programat atime sothat process management is easy in this environment.

Disadvantage:

- Astheoperatingsystemishandlingoneapplicationatatimemostofthe CPUtimeis wasted.

MultiuserOPERATINGSYSTEM:

- Ina multi-user operatingsystem, multiple numberofuserscanaccessdifferent resources of a computerat a time.
- This system provides access with the help of a network. Network generally consists of various personal computers that can and receive information to multi user mainframe computer system.
- Hence, the mainframe computer acts as the server and other personalcomputer act asthe client for that server.
- Ex:UNIX,Window 2000

Advantage:

Sharingofdataandinformationamongdifferentuser.

Disadvantage:

Useofexpensivehardwareforthemainframecomputer.

BatchOperatingSystem

- Inabatchprocessingoperatingsysteminteractionbetweentheuserandprocessoris limited or there is no interaction at all during the execution of work.
- Dataandprogramsthatneed tobeprocessedarebundled andcollected asa 'batch'.
- Thesejobsaresubmittedtothecomputerthroughthepunchedcard.thenthejobwith similar needs executed simultaneously.

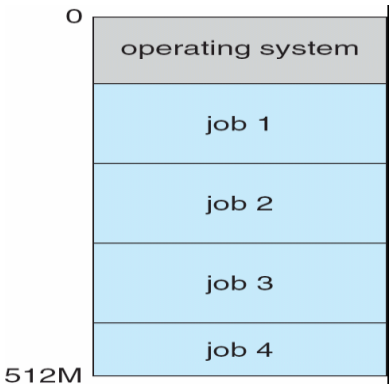
Advantage:

Itissimpletoimplement.

Disadvantage:

Lack ofinteractionbetweenuser and theprogram.

MultiprogrammingOPERATINGSYSTEM:



- Inamultiprogramming operatingSystemseveraluser canexecutemultiple jobsbyusinga singleCPUat the same time.
- The operatingSystem keeps several program or job in the mainmemory.
- When ajobissubmitted to thesystem in amagnetic disk or job pool.
- Then some of the jobs are transferred to the main memoryaccordingtothesizeofthemainmemory.
- The CPU execute only one job which is selected by theoperating System.
- WhenthejobrequiresanyI/Ooperation,thenCPU switchesto next jobinthe main memoryi.eCPU do nothaveto wait forthecompletion ofI/O operation of that job.
- WhentheI/OoperationofthatjobiscompletedthentheCPUswitchestothenext jobafter the the execution of the current job.
- E,g.UNIX,Windows95etc

Advantage:

CPUutilizationismorei.ethemostofthetimetheCPUisbusy.

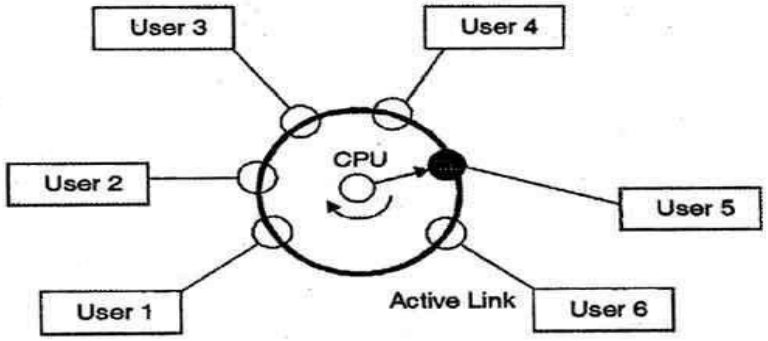
Disadvantage:

Theusercan't directlyinteractwiththesystem.

TimesharingOperatingSystem:

- ThisistheLogicalextensionofmultiprogramming system.
- TheCPUismultiplexedamongseveraljobs thatarekeptinmemory andon disk(the CPU is allocated to a job only if the job is in memory).

- Here the CPU can execute more than one job simultaneously by switching among themselves.
- The switching processisvery fastso that the usercan directly interactwith the system during the execution of the program.
- This system storesmultiplejobsinthemainmemory andCPUexecuteall thejobsin asequence.
- GenerallyCPU time isdivided into no.ofsmall intervalknownas **time slice period**.
- Every process has to execute for the time slice period; then the CPU switch over to nextprocess.
- The switching process is veryfast,so it seems that several processes are executed simultaneously.



Inabovefiguretheuser5isactivebutuser1,user2,user3,anduser4areinwaitingstatewhereas user 6 is in ready status.

As soon as the time slice of user 5 is completed, the control moves on to the next ready user i.e. user6.Inthisstateuser2,user3,user4,anduser5areinwaitingstateanduser1isinreadystate. The process continues in the same wayand so on.

Advantage:

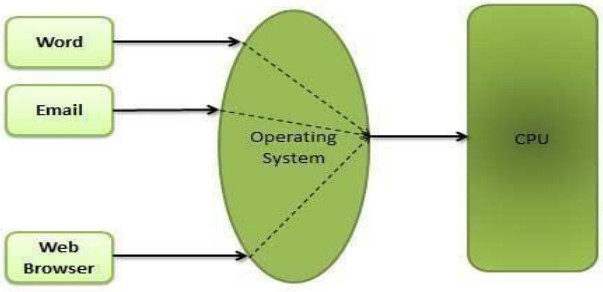
- CPUutilizationismorei.ethemostofthetimetheCPUisbusy.

Disadvantage:

- Theoperatingsystemis morecomplexduetomemorymanagement,Diskmanagementetc.

MultitaskingOperating System:

- A multi-tasking operating system allows more than one program to be running at the same time.
- E.g.-one user can open the word document and can simultaneouslyaccess the internet.
- While the processor handles only one application at a particular time it is capable of switching between the applicationseffectivelytoapparentlysimultaneouslyexecuteeachapplication.



- This type of operating system is seen everywhere today and is the most common type of operating system, the Windows operating system would be an example.

MultiprocessingOperatingSystem:

- When a system contains more than one processor in close communication, sharing the computer bus, the clock and sometimes memory and peripheral devices is known as **multiprocessing operating System.**

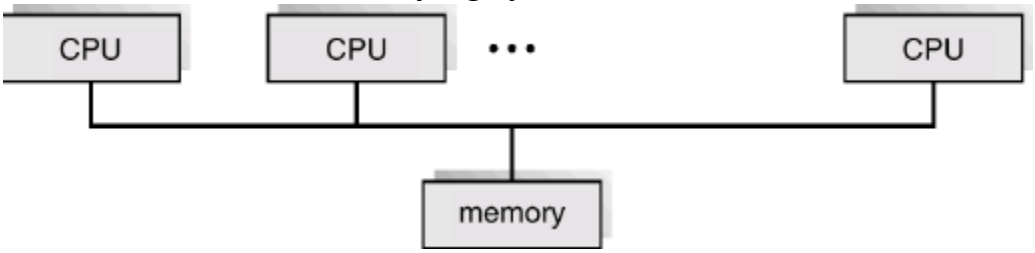
- This isdividedinto2types:
- Symmetricmultiprocessingsystem
- Asymmetricmultiprocessingsystem

Symmetricmultiprocessing(SMP)

- Eachprocessorruns a sharedcopyoftheoperatingsystem.
- Different processor can communicate with each other and are able to execute this copy at the same time.
- These processor are executed bya single operatingSystem and have equalright to access all theI/O devices connected to the system.

Asymmetricmultiprocessing(ASMP)

- Itisbaseduponthepprincipleofmasterslave relationship.
- In this system one processor runs the operatingSystem and other processor run the userprocesses.
- Theprocessorwhichrunstheoperating Systemisknownasmaster processortheprocessorwhichrunstheuserprocessesknownastheslave processor.
- Itisusedinlarge system.
- Each processor is assigned a specific task; master processor schedules and allocated work to slave processors.
- Morecommoninextremelylargesystems



Advantage:

- **Improved Reliability**:-As the system consists of multiple processor, failure of one processor does not disturb the computer system. The other processor in the system continues the task.
- **Improved throughput**:-throughput is defined asthe totalno of jobs which are executed bytheCPU inonesecond. Asthissystemuse multipleprocessor alltheworkloaddivided between the different processor.
- **Economica**l:-in this system different processor share the clock, bus, peripheral and memorybetweenthem. Duetothisreasonthesystemare moreeconomicalthanmultiple single processor system.

Real timeOperatingSystem:

- In a real-time operating system a job is to be completed within the right time constraint otherwise job loses its meaning.
- Thesesystemcompeteaparticularjobinthefixedtimeslotinordertorespondtoan event quickly.
- Real time introduces for correct operation and it required to produce result within a non negotiable time period.
- Real-timesystemsareusuallyusedtocontrolcomplexsystemsthatrequirealotof processing like machinery and industrial systems.

- Thisisof2types:
  - Hardrealtimeoperatingsystem
  - Softrealtimeoperatingsystem

Hardreal-timesystem:

- Thissystemcompletesthecriticaltaskswithinthedefinite intervaloftimeconstraint. If the
- critical task is not completed within the time constraint, then the system fails.
- Thissystemhastocompleteallthe processeswithinthedefinitedeadlineandasinglemiss leads to critical failure.
- E.g.:-Pacemaker, flightcontrolsystem(anymissindeadlineleadstocrash).

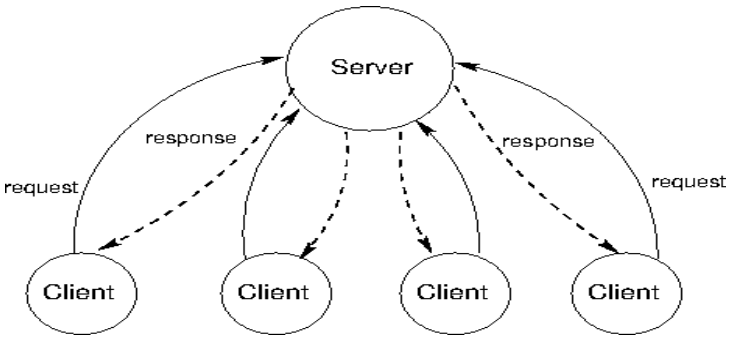
Softreal-timesystem:

- These systems are not affected bythe lapse of time interval and do not cause anycritical failure.
- E.g.:-Livevideostreaming.

DistributedOperating Systems:

- Indistributedoperating system, the users access remoteresources inthe same wayas the local resources are accessed.
- Distributethe computationamongseveralphysicalprocessors.
- Loosely coupled system – each processor has its own local memory; processors communicatewithoneanotherthroughvariouscommunicationslines,suchashigh-speed buses or telephone lines.
- Thesesystemsprovidefeaturessuchasdataandprocessmigration. This
- operating system based on two models.
  - Client-servermodel
  - Peer–to-peermodel

Client-server model:-In this model, the client sendsa request for a resourcetothe server and the server, in turn provides the requested resource as a response back to the client.



Peer –to-peer model: In a peer-to-peer model,all the computers behave
as peers as well as clients. These peers communicate with each other for exchange of their resources.

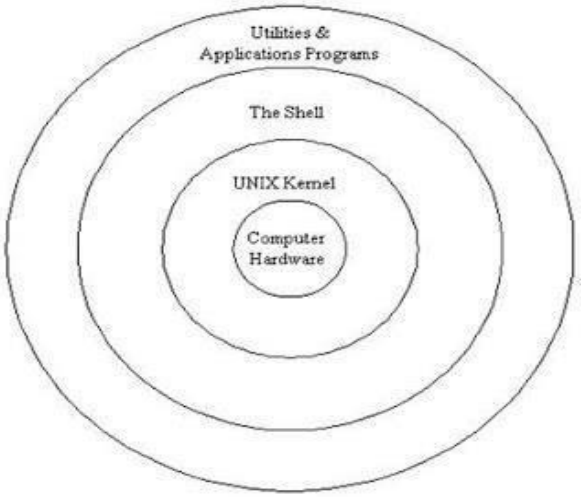
Advantages*:*

- Itfacilitatesthesharingofhardwareand softwareresourcesbetweendifferentprocessors.
- Itincreasesreliabilityasfailureofonenodedoesnotaffect theentirenetwork.
- It increases the computational speed of computer by sharing the workload into different nodes.
- Itenabledifferentuserstocommunicatewitheachotherusing email.

StructureofOperating System

ThestructureofOperatingSystemcomprises of4layers.

- Hardware
- Kernel
- Systemcallinterface(shell)
- Applicationprograms



Kernel:

- It is the vital part of the operatingsystem. It interacts directlywith the hardware of a computer.
- Programs interact with thekernelthrough 10 system calls.
- **System call:-** The *system call*provides an interface to the operating system services.
- System call tells the kernel to carry outvarious asks for the program such as opening afile,writing to a file,obtaining informationabouta file,executingaprogram,terminatingaprocessetc.
- Themain functionofthe**kernel**are
  - Tomanagecomputermemory
  - Tomaintainfilesystem
  - Allocationofresources
  - Controlaccesstothecomputer
  - Handleinterrupts

SystemCallInterface(Shell):

- Shelliscommandlineinterpreter whichinterpretsthecommandsgivenbythe user.
- Itissoftwareorprogramwhichactsasa mediatorbetweenkernelanduser.

- The shell read the commands, whatyou typed at command line and interprets them and sends request to execute a program. That's why shell is called as command line interpreter.

Hardware:

- Computer hardware refers to the physical parts or components of a computer such as the monitor,mouse,keyboard, computer data storage,hard drive disk (HDD), system unit(graphiccards,soundcards,memory,motherboardandchips),etc.allofwhichare physical objects that can be touched

Utilityandapplication programs:

- Utility programs help **manage, maintain and control computer resources**. These programsareavailabletohelpyouwiththeday-to-daychoresassociatedwithpersonal computing and to keep your system running at **peak performance**.
- Application software is all the computer software that causes a computer to perform useful tasks beyond the running of the computer itself.
- Examples of application programs include word processors; database programs; Web browsers; development tools; drawing, paint, and image editing programs; and communication programs.

EvolutionofOperatingsystem:

1. Serialoperatingsystem.
2. Batchoperatingsystem.
3. Multiprogrammingoperatingsystem.
4. Time-Sharingoperatingsystem.
5. Real-Timeoperatingsystem.
6. Multiprocessingoperatingsystem
7. Distributedoperatingsystem

UNIT-2

PROCESSMANAGEMENT

PROCESS:

- Process isaprograminexecution
- Processisacurrentlyexecutabletask.
- Processexecutionmustprogressinasequentialmanner.

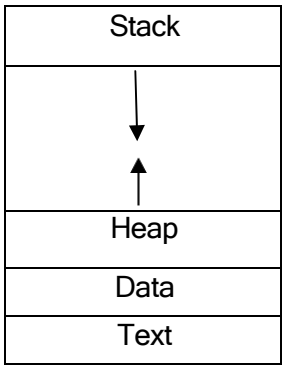
Process	Program
i)Aprocess isthesetofexecutable instruction,thosearethemachine code.	i)It isasetofinstructionwrittenin programming language.
ii)Processisdynamicinnature.	ii)Programisstaticinnature.
iii)Processisanactiveentity.	iii)Programisapassiveentity.
iv)Aprocessresidesinmain memory.	iv)Aprogramresides insecondarystorage.
v) A process is expressed in assemblylanguageormachine level language.	v)Aprogramisexpressedthrougha programmable language.
vi)Thetimeperiodlimited.	vi)Spantimeperiodisunlimited.

ProcessinMemory:-

⇒Aprocessresidesinmemorythroughfollowingsectioni.e.

1) Stack
2) Heap
3) Data
4) Text

- Stacksectioncontainslocalvariable
- Datasectioncontainsglobalvariable
- Textsectioncontains codeor instruction.
- Heapsectioncontainsmemorywhichwillbe dynamicallyallocated during runtime.



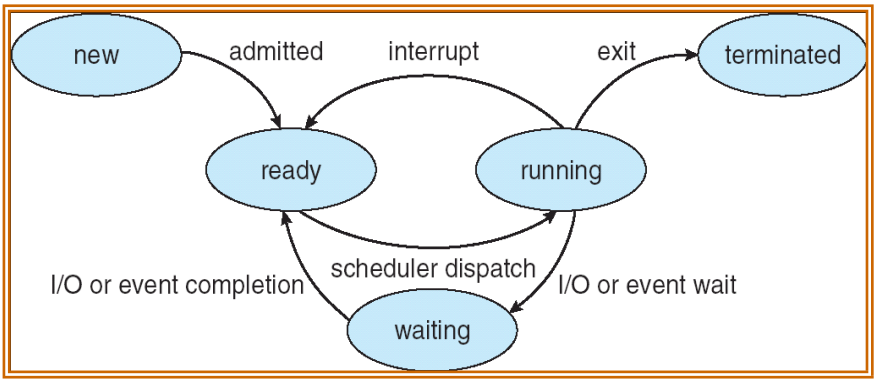
PROCESSSTATE:

Whenaprocess isexecuted, it changes itsstate.Thecurrent activityofthat processisknownasProcess state.A process has different states. They may be

- **Newstate**:

- Whentherequestismadebytheuser,theprocessis created.

- Thenewlycreatedprocessmovesintoanewstatement.
- Theprocess residesinsecondarymemorythrougha queuenamedasjobqueueorjobpool.



⇒.

Diagramofprocessstate

- **Readystate:-**
  - Aprocessissaid tobereadyifitneedstheCPUto execute.
  - Outoftotalnewlycreated processes, specified processesareselected andcopied to temporary memory or main memory.
  - Inmain memorytheyresidesinaqueuenamed asreadyqueue.
- **Running:-**

Aprocessissaidtoberunningifitmovesfromreadyqueueand startsexecutionusingCPU.

- **Waitingstate/blockedstate:-**
  - Aprocessmaymovein to thewaitingstateduetothefollowingreasons.
    - If a process needs anevent to occuror an input or output device and the operatingsystemdoesnotprovideI/Odeviceorevent immediately, thenthe process moved into a waiting state.
    - Ifa higher priorityprocessarrivesat theCPUduring theexecutionofan ongoing process,thenthe processorswitchestothe newprocessand current process enter into the waiting state.
- **Terminatedstate:-**
  - After completionofexecutiontheprocessmoves intotheterminatedstateby exiting thesystem. Theterminatedstateconvertstheprocess into aprogram.
  - Sometimesoperatingsystemterminatestheprocessduetothefollowingreasons.
    - Exceedingthetimelimit
    - Input/outputfailure
    - Unavailabilityof memory
    - Protectionerror
    -

PROCESSCONTROLBLOCK(PCB)/TASKCONTROLBLOCK(TCB)

- Torepresentaprocesstheoperating Systemneedstogroupalltheinformationofaprocess insideadatastructure. This data structure is known as **process control block(PCB).**
- In other words operating System represents each process by a PCB. An operating System considers a processas afundamentalunit forResourceAllocation.Following resourcescouldbe allocated to a process.

TheinformationstoredinsidethePCBincludes

i. **Pointer-**Itstoresthe startingaddressofthe process.
ii. **ProcessState-**Thisfieldstoresorrepresentthecurrentstateoftheprocesswhetheritsin ready/running/new/waiting/terminating.
iii. **ProcessID/Number-**Eachprocess has uniqueIDor serialno. Eachprocess is shownanunique no. known as its Process ID or ProcessNumber.
iv. **ProgramCounter-**Itstoresthe addressofthe nextinstructionto beexecuted.
v. **Register**-This field contains the name of the registers which are currently used by the processor.
vi. **SchedulingInformation-**Thisfieldstorestheinformationaboutthe scheduling algo.used by operating System for scheduling that process.
vii. **MemoryManagementInformation-**Thisfield containsthe valueof thebase table, segment table and page table.
viii. **Account Information-**This field contains the total no. processes, time slice period it used.
ix. **FileManagementInformation-**It storesvariousinformationaboutthe files used by the process.
x. **I/O StatusInformation-**It storesthe information about various allocated I/O devicesto theprocess, a list ofopenfiles &so on.

PROCESS SCHEDULING

- The objective of multiprogrammingis tohave some process running atall times,tomaximize CPU utilization.
- The objective of time sharing is to switchthe CPU among processes so frequentlythat users can interact with each program while it is running.
- ThispurposecanbeachievedbykeepingtheCPU busyatallthetimes.
- So,whentwoormoreprocessescompetefortheCPUatthesametime, achoicehastobe made.
- This procedure of determining the next process to be executed on the CPU is called as**Process Scheduling.**
- Themoduleoftheoperatingsystemthatmakesthisdecision iscalledas**Schedule**r.
- Processschedulingconsistsofthree sub functions:
  - **I.** SchedulingQueue
  - **II.** Scheduler
  - **III.** ContextSwitching

I. Scheduling Queue

Theoperatingsystemmaintainsseveralqueues forefficientmanagementofprocesses. Theseareasfollows:

1. JobQueue:

- Whentheprocessentersintothesystem,theyareput intoajobqueue.
- Thisqueueconsistsofallprocessesinthe systemonamassstorage device suchasharddisk.

2.Ready Queue:

- Fromthe jobqueue,theprocesseswhicharereadyforexecutionareshiftedtothemain memory. In the
- main memorythe processes are kept in the **ready queue**.
- Inotherwords,the readyqueue containsallthoseprocessesthatarewaitingfortheCPU.

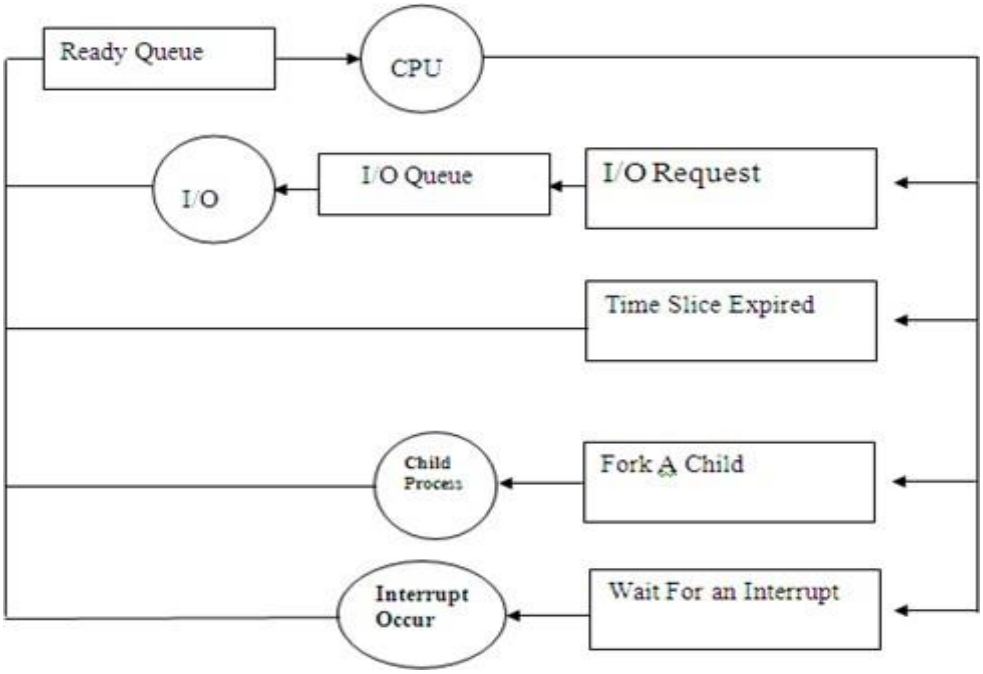
3.DeviceQueue:

- Devicequeue isaqueue forwhichalist ofprocesseswaiting foraparticularI/Odevice. Each
- device has its own device queue.
- WhenaprocessrequiredsomeI/Ooperation,itisthentakenoutof theready queueandkept under the device queue.

4.**SuspendedQueue:**Itstoresthelistofsuspended process.

QueuingDiagram:

- Theprocesscould issueanI/Orequestandthenbe placed inanI/Oqueue. The
- process could create a new subprocess and wait for its termination.
- The process couldbe removedforcibly from the CPUas a resultof aninterrupt, andagain put back in the ready queue.



II. Scheduler:

- Themoduleoftheoperatingsystemthatmakesthedecisionofprocessschedulingisknownas **Scheduler**.
- Their main task is to select the jobs to be submitted into the systemand to decide which process to run.

- Schedulersare ofthreetypes.
  1. LongTermScheduler
  2. Short TermScheduler
  3. MediumTermScheduler

LongTermScheduler(LTS)

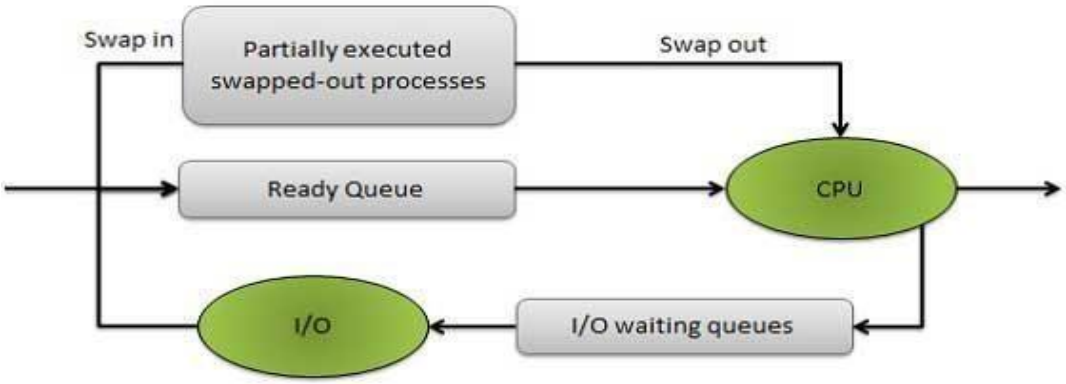
- Itisalsocalled**jobscheduler**;itworkswiththejobqueue.
- Jobschedulerselectsprocessesfromthejobqueueandloadsthemintothemainmemoryfor execution.
- Itexecutesmuchlessfrequently,astheremaybelongtimegapbetweenthecreationofnew process in the system.
- The**primaryobjective**ofthejobscheduleristo controlthe degreeof multiprogramming.
- If the degree of multiprogrammingis stable, then the average rate of process creation mustbe equal to the average departure rate of processes leaving the system.
- Whenprocesschangesthe statefromnewtoready,thenthere isalongtermscheduler.
- LTSselectsabalancedmixofCPUboundand I/Oboundprocesses.

ShortTermScheduler(STS):

- Itisalsocalled**CPUscheduler**or**processscheduler**.
- Itselectstheprocessfromreadyqueue andallocates CPUtoit.
- Mainobjectiveistoincreasethesystemperformance.
- Thisscheduler is frequentlyinvokedascomparedtoLongtermscheduler.
- It isthechangeofreadystatetorunningstateoftheprocess.
- This is faster one because the process executes for short time period before waitingfor an I/Orequest.

MediumTermScheduler(MTS):

- Itisalsoknownas Swapper.
- Sometimes the processes are removedfrom the memory and from CPU to reduce the degree of multiprogramming.
- Thenaftersometimetheprocessescanbereintroducedintomemoryandexecutioncanbe continued where it is left off. This scheme is known as swapping.
- TheMediumTermSchedulerselectsaprocessamongthepartiallyexecutedorunexecuted swapped out processes and swaps it in the main memory.



III. ContextSwitching

- Transferring the control of the CPU from one process to other requires saving the context of currently running process and loading the context of another ready process. This mechanism of saving and restoring the context is known as **context switch**.
- The portion of the PCB including the process state, memory management information and CPU scheduling information together constitutesthe **Context** or **State** of the process.
- Theswitchingperiodsdependsuponthe memoryspeed and thenumberofregistersused.

CPUSCHEDULING

BasicConcept:

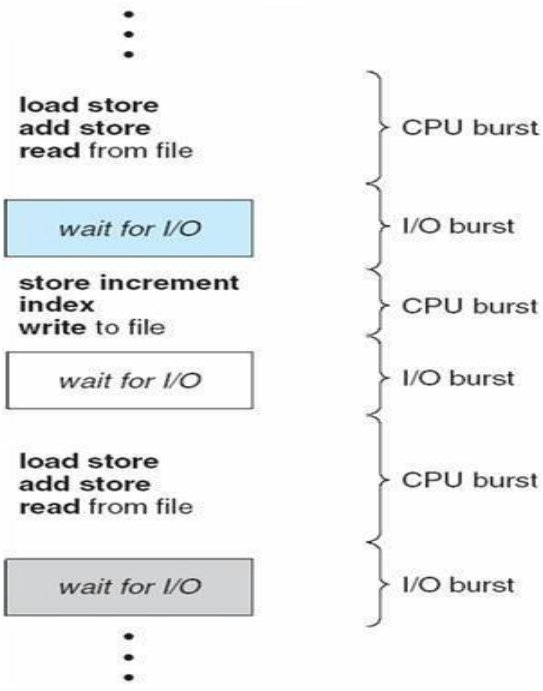
The objective of multiprogramming is to improve the productivity of the computer. It can be done by maximizing the CPU utilization. That means some process running at all times and this is happened by switching the CPU among processor.

Butinaunipolarsystemonlyoneprocessmayrunatatimeandotherprocessesmust wait until the CPU is free and can be rescheduled.

Schedulingisafundamentaloperatingsystemfunction.Almostallcomputerresourcesare scheduledbeforeuse.TheCPUisoneoftheprimarycomputerresource.Thusitsscheduling istocontrol the operating System design.

CPU-I/OBurstCycle

- ThesuccessofCPUschedulingdependsonanobservedpropertyofprocesses. Process
- execution consists of a cycle of CPU execution and I/O wait.
- Processesalternatebetweenthesetwostates.ProcessexecutionbeginswithaCPUburst.
- ThatisfollowedbyanI/Oburst,whichisfollowedbyanotherCPUburst,thenanotherI/O burst, and so on.
- ThefinalCPU burstendswithasystemrequesttoterminate execution.



CPUScheduler(ShortTermSchedular):

- Whenever the CPU becomes idle, the operating systemmust select one ofthe processes inthe ready queue to be executed.
- Theselectionprocessiscarriedoutbytheshort-termscheduler (orCPUscheduler). Scheduling can
- be of 2 types:
    - Non-PreemptiveScheduling
    - PreemptiveScheduling

Non-PreemptiveScheduling:

Inthiscaseoncethe CPUhas beenallocatedtoaprocess,theprocesskeepsthe CPUuntilit releasesthe CPUbyterminatingorbyswitchingtothewaitingstate.Thatiswhenitis(process)iscomputedorrequired any I/O operation.

PreemptiveScheduling:

In this case CPU can be released forcefully**.** Under this scheduling the process has to leave the CPU forcefullyonthebasisofcriteria likerunningtoreadyandwaitingstatetoreadystate(i.e. when interrupt occur or due to completion of time slice period).

DISPATCHER:

- Dispatcheristhemodulethatgivescontrol of theCPUtotheprocessselectedbytheshort- term scheduler.
- The time it takes for the dispatcher to stop one process and start another running is known as the **dispatch latency.**
- Thisfunctioninvolvesthefollowing:
    - Switchingcontext
    - Switchingtouser mode
    - JumpingtotheproperlocationintheuserprogramtorestartthatprogramIn.

SchedulingCriteria:

- There are severalCPUscheduling algorithm. But we have to select one which is suitable forour system.
- Therearesomecriteriabasedonwhich CPU schedulingalgorithm selectthenextprocess to execute.
    - **CPU utilization**. We want to keep the CPU as busy as possible. Conceptually, CPU utilizationcanrangefrom0to100percent.Inarealsystem,itshouldrangefrom40percent (for a lightly loaded system) to 90 percent (for a heavily used system).
    - **Throughput:**itcanbedefinedfor asystemas"theno.ofjobscompletedperunit time".
    - **Turnaround time:** The interval of time from submission of a process to the time of completion. It is the total time spent by a process within the system.
      Turnaround time=timespent inthe readyqueue +time spent inexecution + time spentin I/O operation.

      OR

      TurnAroundtime=Completiontime–Arrivaltime

- **Waiting time**: It is the sum of the periods spent waiting in the ready queue .[that means CPUschedulingaffectsonlytheamountoftimethataprocessspendswaitinginthe ready queue,butdoesnotaffecttheamountoftimeduringwhichprocessexecutesor doesI/O].

  Waitingtime=TurnAroundtime–Bursttime

  Itistheamount oftime duringwhichprocessisinthereadyqueue.
- **Responsetime:** It is theamount oftime,processtakestostartresponding(first response after submission)

  ResponseTime=TimeatwhichprocessfirstgetstheCPU–Arrivaltime

SCHEDULINGALGORITHM

TheScheduling algorithmdecideswhichoftheprocess inreadyqueue isto be attending theCPU.There are various scheduling algorithms:

1. FirstComeFirstServescheduling(FCFS)
2. Shortest JobFirst(SJF)
3. Priorityscheduling
4. RoundRobinScheduling
5. MultilevelQueuescheduling

FirstComeFirstServescheduling(FCFS)

- Thisisthesimplestandeasiestschedulingalgorithm.
- Inthisscheme, theprocessthatrequeststheCPUfirstisallocatedtheCPUfirst. The
- first process is stored in the first position of the readyqueue.
- HerethedatastructureofthereadyqueueisFIFOqueue.
- FCFS is non preemptive. when CPU is free, it is allocated to other process i.e the CPU has been allocated to process, that process keeps the CPU untilit release the CPU either byterminating or by requesting I/O.

- Letthe processarrivesintheorderp1,p2,P3,p4,p5.

Process	ArrivalTime	CPUBurst
P1	**0**	**20**
P2	**4**	**2**
P3	**6**	**40**
P4	**8**	**8**
P5	**10**	**4**

FindouttheAverageTurnAroundTime(ATAT)andAvg. WaitingTime(AWT).

Solution:

Theresult ofexecutionshowninGANTT CHART:

P1	P2	P3	P4	P5
0 20	22	62	70	74

Waiting time:

P1=0
P2=20-4=16
P3=22-6=16
P4=62-8=54
P5=70-
10=60

HencetheAWT(AverageWaitingTime)=(0+16+16+54+60)/5=29.2

TurnAround Time(TAT):

P1=20-0=20
P2=22-4=18
P3=62-6=56
P4=70-8=62
P5=74-
10=64

Hence**AverageTAT=(20+18+56+62+64)/5=44**

Disadvantage:

- Theuser havingsmalljobhastowaitforalongtime.
- Thisalgorithmisparticularlytroublesome fortiesharing systembecauseeachuser needsto geta share of the CPU at regular time intervals.

Advantage:

- FCFSschedulingisverysimpletoimplementandunderstand.

ShortestJobFirstScheduling(SJF)

- In this type of scheduling when the CPU is available , itis assigned to the process that has thesmallest next CPU burst.
- Iftwoprocesseshavethesame lengthnextCPUburst,FCFSscheduling isusedtobreakthetie. It is also
- known as shortest next CPU burst.
- SJFalgorithmmaybeeitherpreemptiveornonpreemptive.
    - The choice arises whena new process arrives atthe readyqueue while a previous process is executing.
    - ThenewprocessmayhaveashortestnextCPUburst thanthecurrentlyexecutingprocess.
    - A preemptive SJF algorithm will pre-emptthe currently executing process where as anon premptive SJF algorithm will allow the currently running process to finish its CPU burst.

- ❖ Preemptive SJF scheduling is sometimes called “shortest remaining time first scheduling”.
- ❖ Largerjobswillneverbeexecutedifsmallestjobsarrives.

Process	ArrivalTime	CPUBurst
P1	00	15
P2	05	10
P3	07	05
P4	10	08

Non preemptive

GanttChart:

P1	P3	P4	P2

0 15 20 28 38

WaitingTime:

P1=0

P3=15-7=8

P4=20-10=10

P2=28-5=23

AWT=0+8+10+23=41/4=10.25

TurnAroundTime:

P1=15

P2=38-5=33

P3=20-7=13

P4=28-10=18

ATAT=(15+33+13+18)/4=19.75

Priorityscheduling:

Incaseofpriorityscheduling theprocesshavinghighest priorityvaluewill beexecuted first.

Problem:

processs	AT	BT	Priority
P1	00	15	3
P2	04	10	2
P3	06	05	1
P4	08	08	4

SOLUTION(NON-PREEMPTIVEGANTTCHART):

P1	P3	P2	P4

0 15 20 30 38

W.T.

P1=0

P2=20-4=16

P3=15-6=9

P4=30-8=22
A.W.T=0+16+9+22=47/4=11.75

T.A.T

P1=15-0=15
P2=30-4=26
P3=20-6=14
P4=38-8=30
A.T.A.T=15+26+14+30=85/4=21.25

InternalPriority:

InPriorityScheduling apriorityvalue isassignedto eachoftheprocess inthereadyqueue. Thepriority value can be assigned either internallyor externally. The factor for assigning internalpriority is:

- Bursttime
- MemoryRequirement
- I/Odevices
- No.Offiles

ExternalPriority:

Thefactorforassigningtheexternalpriorityvalueare:

- Importantofprocess
- Amountoffundgiven
- Politicalpressure

Thepriorityschedulingmaybepreemptiveornon-preemptive.themajorproblemwiththepriority scheduling is indefinite blocking or starvation. The solution to the problem is **aging**. Aging is a techniquewhichgraduallyincreasestheprioritvalueoftheprocessthatwaitsinthe readyqueue for a long time.

Problem:

Process	**B.T**	**Priority**
P1	10	3
P2	1	1
P3	2	4
P4	1	2

GanttChart:

P2	P4	P1	P3

0 1 2 12 14

W.T

P1=2 P2=0 P3=12 P4=1

A.W.T=(2+0+12+1)/4=3.75

T.A.T

P1=12 P2=1 P3=14 P4=2

A.T.A.T=(12+1+14+2)/4=7.25

RoundRobin Scheduling:

- ThisisdesignedforTimesharingsystem.ItissimilartoFCFS scheduling.
- ButtheCPUpre-emptsamongtheready processin every timesliceperiod,whicharein the ready queue.
- IncaseofFCFSschedulingthereadyqueueisaFIFOqueue.ButinRRschedulingthe readyqueue is a circular queue.
- Round Robin Scheduling is a purelypreemptive scheduling algorithm. Because after everytime slice period CPU will switch over to the next process in the ready queue.

Process	A.T.	B.T.
P1	00	20
P2	10	10
P3	15	15
P4	15	10

CPUTime=5ms

P1	P1	P2	P3	P4	P1	P2	P3	P4	P1	P3

0 5 10 15 20 25 30 35 40 45 50 55

W.T:

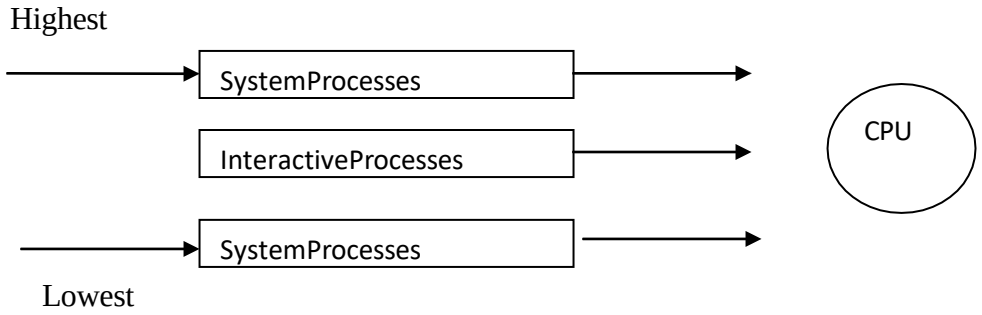
P1=(25-10)+(45-30)=30
P2=30-15=15
P3=(35-20)+(50-40)=20
P4=(20-15)+(40-25)=20
A.W.T=(30+15+20+20)/4=21.25

T.A.T

P1=50
P2=(35-10)=25
P3=(55-15)=40
P4=(45-15)=30
A.T.A.T=(50+25+40+30)/4=36.25

MultilevelQueuescheduling

- Thisalgorithmpartitionsthereadyqueueinto severalseparatequeues.
- Processes are permanentlyassigned to onequeue based onsome criteriasuchas memorysize, process priority.
- Eachqueuehasitsownschedulingalgorithm.
- ForegroundqueuemaybeinRRandbackgroundqueuemaybein FCFS.
- Againthereisaschedulingalgorithmtoselectaqueueamongmanyqueues.
- If priority queueisapplied.thennoprocessin thelowestpriority queuecanbeexecutedtill there is a process in the higher priority queue.
- If Round Robin scheduling is applied, then each queue gets CPU for a certain amount of time.Again that time will be divided among the processes in that queue.



InterprocessCommunication(IPC)

Overview:

Processesareclassifiedinto2 categories.

Theyare:i)Independent process

ii)CooperatingprocessIndepende

nt process:-

It is defined as a process thatdoes not share anydata and does not communicate withother process. Inotherwordswecansaythatmodificationmadetoanindependentprocessdoesnotaffectthe functioning of other processes.

Co-operatingprocess:-

Itisdefined asaprocess,whichgetsaffectedbyanyother process.

Theseprocessesareusedfor resourcesharingandtospeedupacomputationprocedure.

InterprocessCommunication(IPC)

Interprocesscommunicationisthe mechanismprovided bytheoperatingsystemthat allowsprocessesto communicate with each other.

Processesareclassifiedinto2categories.Theyare:

Independentprocess: Anindependentprocessisnotaffectedbytheexecutionofotherprocesses.

Cooperatingprocess:aco-operatingprocesscanbeaffectedbyotherexecutingprocesses.

Advantagesofprocesscooperation

Informationsharing: Sinceseveralusersmaybe interested inthesamepieceofinformation (forinstance,ashared file), we must provide an environment to allow concurrent access to these types ofresources.

Computationspeedup: Ifwewant aparticulartaskto runfaster,wemust break it into subtasks,eachofwhichwill be executing in parallelwith the others. Such a speedup can be achieved only if the computer has multiple processing elements (such as CPUS or I/O channels).

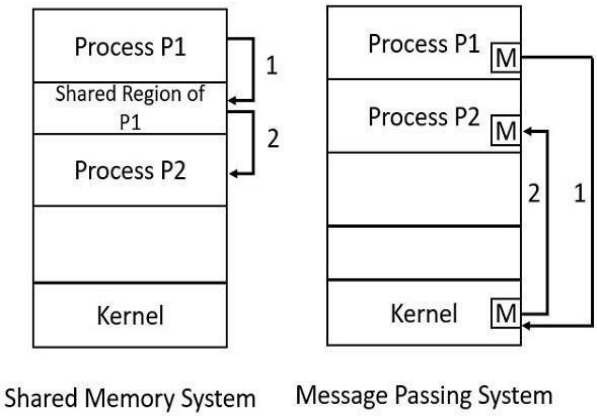
Modularity: We maywant toconstructthesystemina modularfashion,dividingthesystem functions intoseparate processes or threads.

Convenience: Evenanindividualuser mayhave manytasksonwhichto workatonetime. For instance, auser may be editing, printing, and compiling in parallel.

WaystoImplementIPC

1.SharedMemory: Multipleprocessescanaccessacommonshared memory. Multiple processes communicate by shared memory,where one process makes changes at a time and then others view thechange.Sharedmemorydoesnotusekernel.

2. MessagePassing: Messagepassingprovidesamechanismtoallow processes to communicate and to synchronize their actions without sharing the same address space. It is very useful in case where the tasks or processes reside on different computers and are connected byanetwork.Messagescanbeoffixedor variablesize.



UNIT-

3MEMORYMANAGEMEN

T

Oneofthemajorfunctionsofoperatingsystemismemorymanagement. Itcontrolsthe

- Allocationandde-allocationofphysicalmemory.
- Whichpartofthe memoryis currentlyusedbywhichprocess.
- Decidewhichprocessesaretobeloadedintomemory.
- Freespacemanagement.
- Dynamicallocation/de-allocationofmemorytoexecutingprocessesetc.

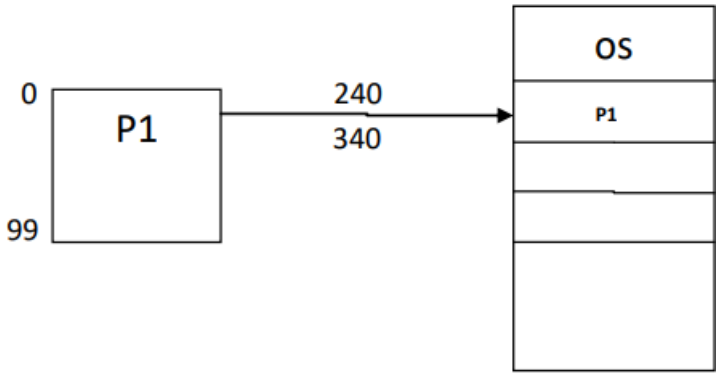
LogicalAddress&PhysicalAddress:-

LogicalAddress:AddressgeneratedbyaCPUiscalledas logicaladdress.

PhysicalAddress:-Addressgeneratedbymemorymanagement unit iscalledas physicaladdress.The logical address is known as "virtual address".

Thesetofalllogicalunit oraddressgeneratedbyprogramsreferredas"logicaladdressspace". The set of physical address corresponding to logical address is referred as "physical address space".Suppose the program size= 100 KB

Butitisloadedinthemainmemoryfrom240to340KB.



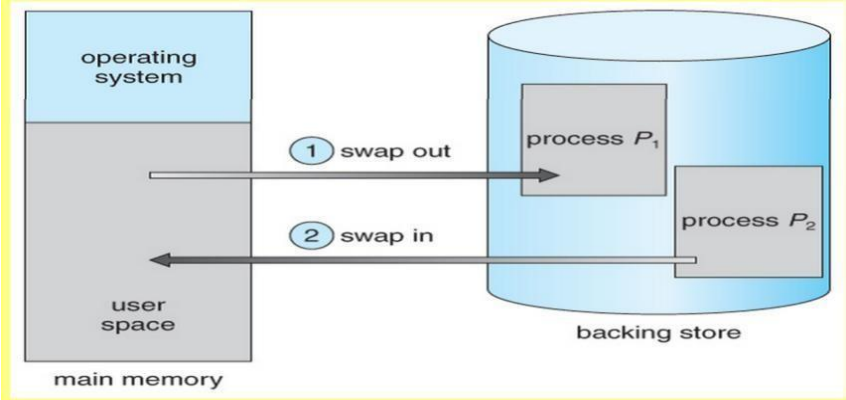
- So0to99KBisthelogicaladdressspace but240to 340KBisthephysicaladdressspace.
- Physicaladdressspace=logicaladdressspace +contentofrelocation register.
- Themappingbetweenlogicalandphysicaladdressesaredoneatrun-timebythe memorymanagement unit (MMU).

SWAPPING:-

- Swappingisthemethodtoimprovemainmemoryutilization.
- Whenaprocessisexecuteditmustbe inthemain memory.
- A process can be swapped out temporarily to secondary memoryor hard disk or backing memoryand thenagain brought back to secondarymemoryfor execution. This technique is known as "**Swapping".**
- Thebasicoperationofswappingis
    - Swap-out(roll-out)

- Swap-in(roll-in)

Swap-out:-The mechanism to transfer the process from main memory to secondary memory. Swap-in:- Themechanismthat shiftstheprocessfromsecondarymemoryto primary memory.



MEMORYALLOCATIONMETHODS

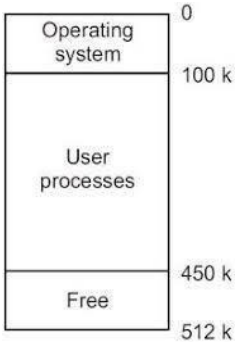
- Themainmemorymust accommodatebothoperatingsystemand varioususerprocesses.
- Generally,themainmemoryisdividedinto2partitions.
  - Operatingsystem.
  - Applicationprogram/userprocesses.
- Theoperatingsystemplaceineitherlowmemoryorhigh memory.
- Commonlytheoperatingsystemis loadedinlowmemory.
- Generally,therearetwomethodsusedforpartitioningthememoryallocation.
  - Contiguousmemoryallocation
  - Non-Contiguous memory

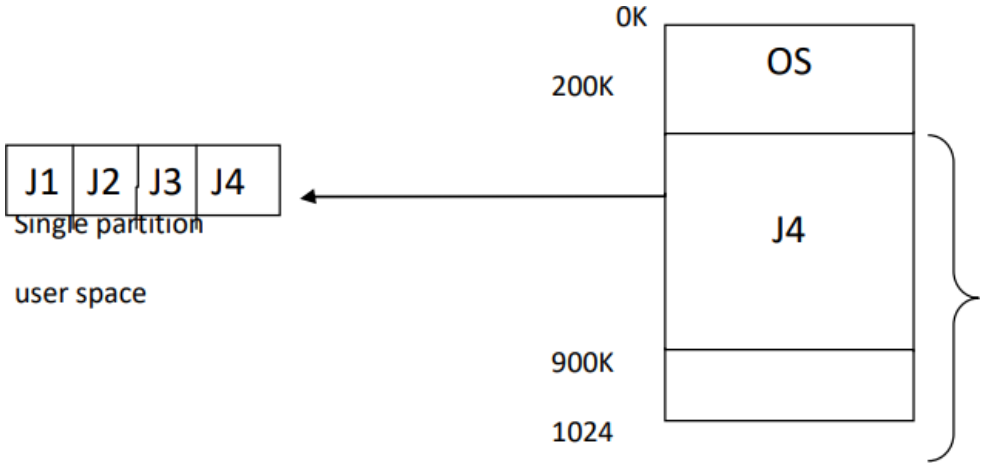
allocationContiguousMemoryAllocation:-

- Itisagaindividedintotwo parts.
  - Singlepartitionallocation.
  - Multiplepartitionallocation.

SinglePartitionAllocation:-

- Inthismemoryallocationmethod,theoperatingsystemresideinthelow memory.





- Andtheremainingpart/spacewillbetreatedasasinglepartition.
- Thissinglepartitionisavailableforuserspace/applicationprogram.
- Onlyone jobcanbe loaded inthisuser space isthe mainmemoryconsisting ofonlyone processat a time, because the user space treated as a single partition.

Advantage:-

i. It isverysimple.
ii. Itdoesnot requireexpertiseto understand.

Disadvantage:-

i. Memoryisnotutilized property.
ii. PoorutilizationofprocessorforI/O).

Multiple Partition Allocation:-

This methodcanbeimplementedin3ways. Theseare:

- Fixedequalmultiplepartition.
- Fixedvariablemultiplepartition.
- Dynamicmultiplepartition.

Fixedequalmultiplepartition:-

i. Inthis memorymanagement scheme the operating systemoccupies the low memoryand rest of mainmemory is available for user space.
ii. Theuserspaceisdividedintofixedpartitions.Thepartitionsizedependsupontheoperating system.
iii. A partition of main memory is wasted within a partition is said to be “Internal Fragmentation”and the wastage of an entire partition is said to be ”External Fragmentation”.
iv. Thereisoneproblemwiththismethodismemoryutilizationisnotefficientwhich causesinternal and external fragmentation.

Advantages:-

- Thisscheme supportsmultiprogramming.
- EfficientutilizationofCPU &I/O devices.
- Simple andeasytoimplement.

Disadvantages:-

- Thisschemesuffersfrominternalaswellasexternal fragmentation.
- Since,thesizeofpartitionsarefixed,thedegreeofmultiprogrammingisalsofixed.

Fixedvariablepartition:-(unequalsizepartition)

- Inthisschemetheuserspaceofmainmemoryisdividedintonumberof partitions, butthepartitions sizes are different length.
- The operating system keepa table whichindicates,which partition of memory are availableandwhichareoccurred.Thistable isknownas"PartitionDescriptionTable" (PDT).
- Whenaprocessarrivesandneedsallocationormemory,wesearchforpartitionwhich isbigenough to allocate this process. If find one allocation, then allocate the partition to that process.

Advantage:-

i. Supportsmultiprogramming.
ii. Smallermemoryloss(expected).
iii. Simple&easyto implement.Disadvantage:-

i. Suffersfrominternalaswellasexternalfragmentation.

Figure 4: Fixed Partitioning

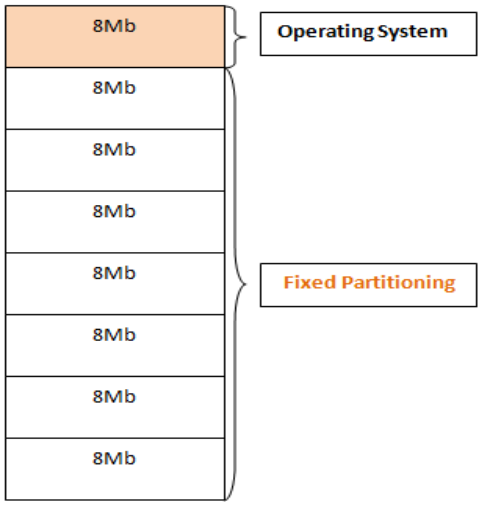
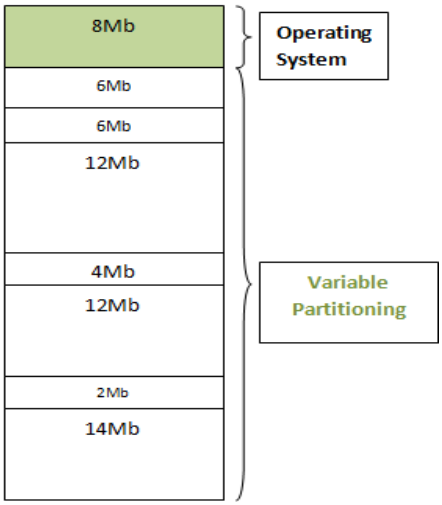


Figure 5: Variable Partitioning

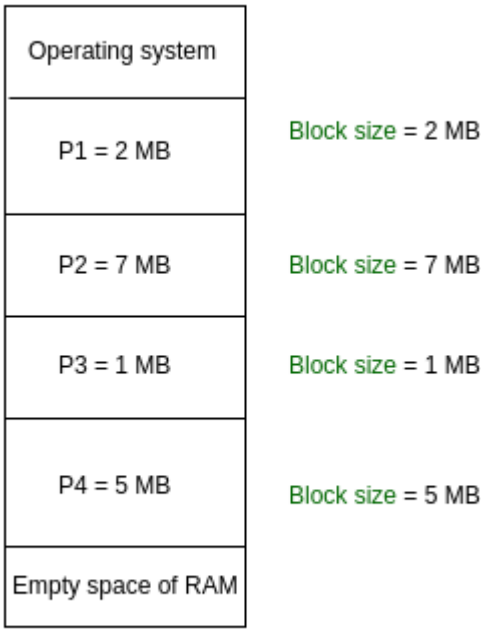


DynamicMultiplePartitionMemoryManagement:- (Variablepartition)

- Toovercome/eliminatesomeoftheproblemswithfixedpartition,anothermethodis developedknown as "Dynamic Partitioning".
- In this technique, the amount of memory allocated is exactly the amount of memory a processrequires.

- Inthis methodthepartitionaremadedynamically.
- Initially when there is no process in the memory, the whole memory is available for allocationand it is treated as a single large partition of available memory (a hole).
- Wheneveraprocessrequest formemory,theholelargeenoughto accommodatethat process isallocated.
- Therestofthememoryisavailabletootherprocess.
- As soon as the process terminates, the memory occupied by it is de-allocated and can be usedby other process.

Dynamic partitioning



Partition size = process size
So, no internal Fragmentation

Advantage:-

1. Partitionchangeddynamically.Sonointernalfragmentation.
2. EfficientmemoryandCPU

utilization.Disadvantage:-

1. Suffers from external

fragmentation.PartitionSelection Algorithms:-

Whenever aprocessarrivesandthere arevariousholes largeenoughto accommodate it,the operatingsystemmayuseone ofthefollowing algorithmtoselectapartitionforthe process.

- First fit:- Inthis algorithm, the operatingsystemscans the free storage list and allocates the first partition thatis large enough for that process.

  Advantage:-

  1. Thisalgorithmisfastbecauseverylittlesearchis

  involved.Disadvantage:-

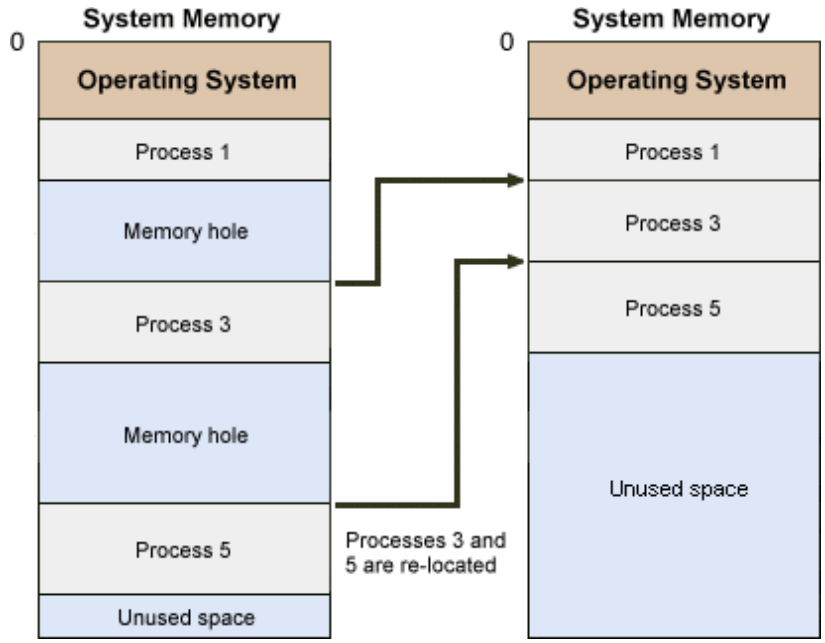
  1. Memorylossmaybehigh.

- Best fit:- In this algorithm the operatingsystem scans the free storage list and allocate the smallestpartitionthat is big enough for the process.
  Advantage:-
  1. Memorylosswillbesmaller thanthefirstfit.

  Disadvantage:-Searchtimewillbelarger ascomparedtofirstfit.
- Worst-fit:-In thisalgorithm theoperatingsystemscanstheentirefreestoragelistand allocatethelargestpartition to the process.
  Disadvantage:-Maximumintervalfragmentation.

Compaction:-

- Compaction is a technique of collecting all the free spaces together in one block, so that other process can use this block or partition.
- There are large no. of small chunks of free memory that may be scattered all over the physicalmemoryandindividualeachofchunksmaynotbigenoughtoaccommodate even a small program.
- So,compactionisatechniquebywhichthesmallchunkoffreespacesaremadecontiguous to each other into a single free partition, that may be big enough to accommodate some other processes.
  Ex-Collectallthefragmentationtogetherinoneblockandnowthefigureis:-



Noncontiguousmemorypartition:-

- As one program terminates, the memory partition occupied by it becomes available to be usedby another program.
- Let the size of the freed memory be S, the next program to be run onthe memory may need aspace which is larger or smaller than S.
- Ifitislargerthenitcannotbeloaded,ifitissmaller,thanapartofthepartition remainsunutilized.
- Thisunutilized memoryisknownasfragment..Thisconceptisknownasfragmentation.
- Fragmentation is of 2 types:-
  - ✓ Externalfragmentation
  - ✓ Internalfragmentation.

Externalfragmentation

Whenthefragmentis toosmallforarunningprogram tobeload,thenthereafragmentor portionremains unutilized.

Internalfragmentation

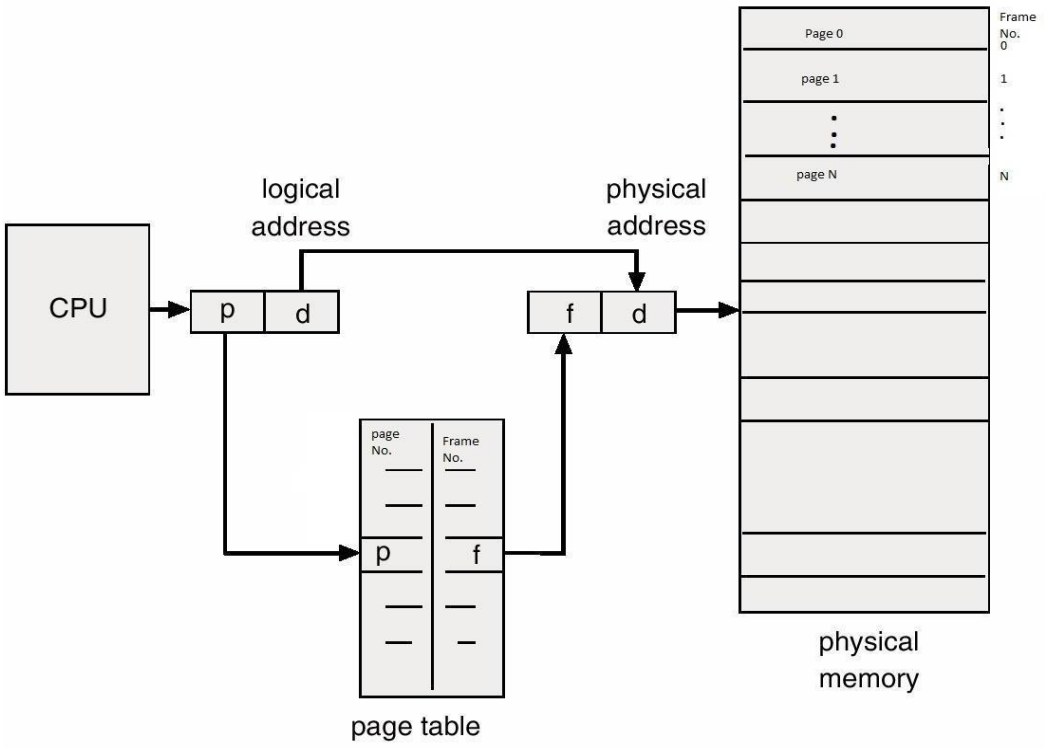
Whenthefragmentremainsunutilizedinsidealarger memorypartitionalreadyallocatedtoaprogram.

- Bothlead topoormemoryutilization.
- To overcome this problem the memory is allocated in such a way that parts of a single processmaybeplacedinnon-contiguousareasofphysicalmemory.Thistypeofallocation is known as Non-contiguous allocation.
- Thetwopopularschemesin Non-contiguousallocationarepaging&segmentation.

Paging

- Pagingisanefficientmemorymanagement schemebecauseitisNon-contiguousmemory allocation method.
- The partition methodsupports the contiguous memory allocation i.e the entire process loaded in partition but in paging the process is divided into small parts, these are loaded into elsewhere in main memory.
- Thebasicideaofpagingisphysicalmemory/ mainmemoryisdividedintofixedsize blocks calledas frames.
- Logicalmemory(oruserjob) isdivided intofixedsizeblockcalledpages. Page
- size and frame size should be equal.
- Backingstoreisalso divided intofixedsize blockthatareofsamesizeasmemoryframes. When
- a process is tobe executed its pages are loadedinto the main memory in any availablememory frame.
- EverylogicaladdressgeneratedbyCPU isdividedintotwo parts:-

1. Pagenumber (P)2.Pageoffset(d)



Structureofpagingscheme

- Pagenumberisusedasanindex intothe pagetable.
- Pagetable isadatastructuremaintained byoperatingsystem.It isused formappingpurpose. The
- page table specifies-
  - Whichframesareallocated
  - Whichframesareavailable
  - Howmanytotalframesarethereandsoon.
- Thepagetableconsistsof2fields-1)**pagenumber2)frame number**
- pagetable containsthe base addressofeachpageinphysical memory.
- Thebaseaddressiscombinedwiththepageoffset todefinethephysicalmemoryaddress.
- Thepagesizeorframesizeisdependinguponoperatingsystem.But it is generally a power of 2,such as4MB,8MB,16MB etc
- Thepagemaptablespecifieswhichpageisloadedinwhichframe,but displacement or offset iscommon.
- Thepaginghasnoexternalfragmentation,but theremaybeinternalfragmentation.In paging itis called as **pagebreak**.

Advantage

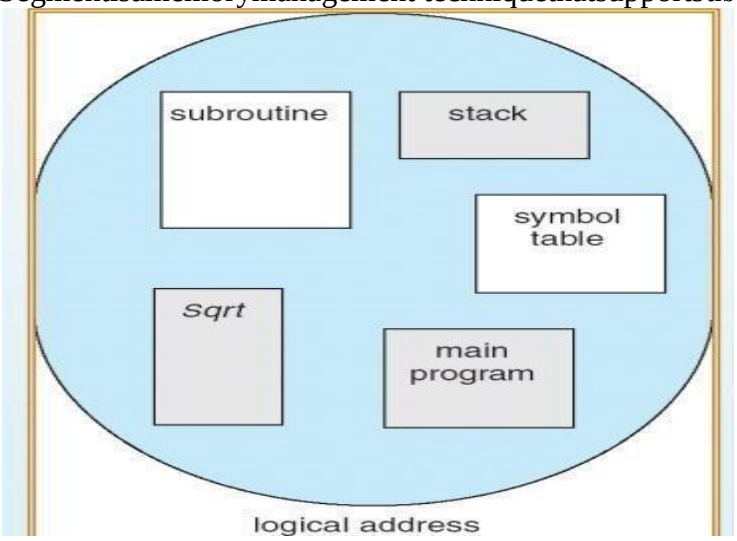
- Itsupportstimesharing system
- Itdoesnoteffectfromfragmentation It
- support virtual memory.

Disadvantage

- Theschememaysuffer“pagebreak”.
- Ifthenumberofpagesarehigh,itisdifficultto maintainpagetable.

Segmentation

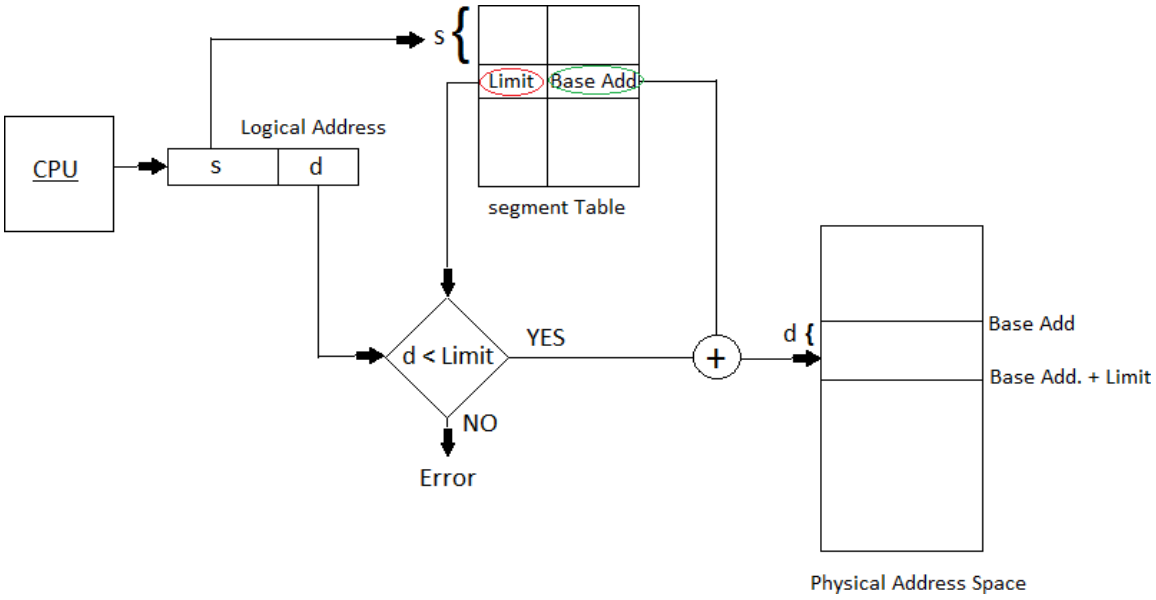
- Incaseofpagingtheuser’sviewofmemoryisdifferentfromphysicalmemory.
- User don’t think thatmemory is alinear array of byte, some containinginstruction and somecontaing data.
- But he view the memory as a collection of variable sized segments and there is no ordering ofsegments.
- Segmentisamemorymanagement techniquethatsupports**user’sviewofmemory.**



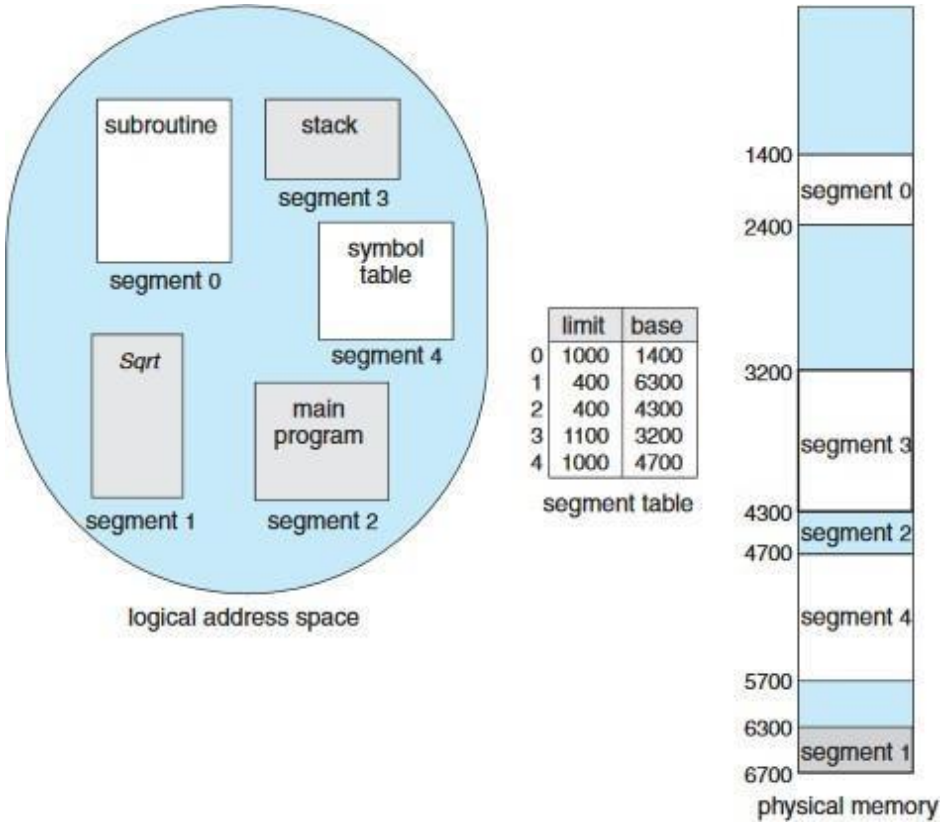
- Asegment canbedefinedasa logicalgroupingofinstruction,suchassubroutine, array or a dataarea.

- “Everyprogramisacollectionofthesesegments”
- Here the logicaladdress is a collectionofsegment
- Each segment has a name and length.
- Segmentation is a technique for managing these segments.
- Eachsegmentarenumberedandreferredbysegmentnumber.
- The logicaladdressisconsistingoftwotuples<segmentno,offset>
- Ex-the lengthofa segment main is 100K, here ‘main’ is the name ofthe segment and theoffsetvalue is100K.theoperatingsystemsearchestheentire mainmemoryforfree space to load a segment. Thismapping is done by segment table.
- Thesegmenttableisatable,eachentryofit hasa segment“Base”andasegment “limit”. Logical
- address consist of two parts.
    1. SegmentNumber(s)
    2. Offsetintothatsegment(d)
- Thesegmentnumberisusedasanindexintothesegmenttable.
- Theoffset iscomparedwithsegment limit.offset shouldbe lessthanorequalto limit,else there is an error. If the offset is valid then “d” will be added with base value to get the actual physical address.

Diagramofsegmentationscheme



Example:-



PAGE FAULT

- Whentheprocessorneedtoexecuteaparticularpage,thatpageisnotavailablein mainmemory then an interrupt occurs to the operating system called as **page fault.**
- When the page fault is happened, the page replacement will be needed. The word page replacement means to select a victim page in the main memory.
- Replacethe pagewiththerequired pagefrombacking storeorsecondarymemory.

STEPSFORHANDLINGPAGEFAULT

- Toaccessapagetheoperatingsystemfirst checkthepagetabletoknowwhetherthereference is valid or not.
- Ifinvalid,aninterruptoccurtooperatingsystemcalledpage fault. Then
- operating system search for free frame in the memory.
- Thenthedesirepageisloadedfromdisktoallocatefreeframe.
- Whenthediskreadiscompletethepagetableentryis modifiedbysettingthevalidbit. The the
- execution ofthe process starts where it was left.

Difference.BetweenPagingandsegmentation

Paging	Segmentation
Themainmemorypartitionedintoframes or blocks	Themainmemorypartitionedinto segments.
Thelogicaladdressspacedividedintopages by compiler or memory management unit.	Thelogicaladdressspaceisdivideinto segmentsspecifiedbytheprogrammer.
Itmaysuffer frompagebreakorinternal fragmentation	Thisschemesufferfromexternal fragmentation
Theoperatingsystemmaintainpagemap table for mapping between frames and pages.	Segmentmaptableisusedfor mapping.
Itdoesn’tsupportuserviewofmemory	Itsupportuser viewofmemory.
Theprocessorusespageno.andoffsetto calculate absolute address	Theprocessorusesthesegmentno.and displacement to calculate the absolute address.

VIRTUALMEMORY

Virtual memory is a technique which allows the execution of a process, even the logical address space is a greater than the physical memory.

Ex:letthe programsizeorlogicaladdressspaceis15MB.,buttheavailablememoryis12MB. So, the 12MB is loaded in main memory and remaining 3MB is loaded in thesecondary memory. When the 3MB is needed for execution then swap out the 3MB from main memory tosecondary memory and swap in 3MB from secondary memory to main memory.

Advantages:

Largeprogramscanbewritten,asvirtualspaceavailableishugecomparedtophysicalmemory. Less I/O

required, leads to faster and easy swapping of processes.

Morephysicalmemory available,asprogramsarestored onvirtualmemory,sothey occupy very less space on actual physical memory.

DEMANDPAGING

Demandpagingistheapplicationofvirtualmemory. It is

the combination of paging and swapping.

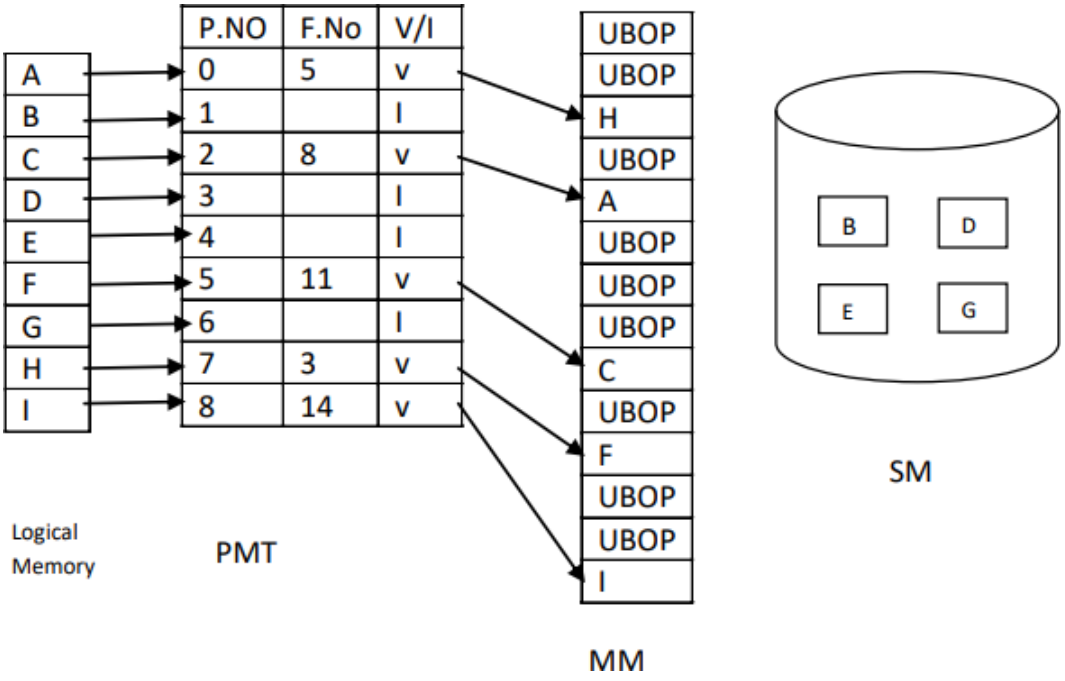
Thecriteriaofthisschemeis“apageisnotloadedintothemainmemoryfromsecondary memory,until it is needed”.

So,apageisloadedintothemainmemory by demand,sothisschemeiscalledas“Demand Paging”.

For ex: Assume that the logical address space is 72 KB. The page and frame size is 8KB. So thelogical address space is divided into 9 pages, i.e numbered from 0to 8.

Theavailablemainmemoryis40KB.i.e.5framesareavailable.Theremaining4pages areloaded in the secondary storage.

Wheneverthosepagesarerequired,theoperatingSystemswap-inthosepagesintomain memory.



Intheabovefigurethemappingbetweenpagesandframesdonebypagemap table.

Indemand paging the PMT consisting of3 fields i.e. page no., frame no. and valid/invalid bit. Ifa page resides in the main memory the v/I bit set to valid.

OtherwisethepageresidesinthesecondarystorageandthebitsettoInvalid.

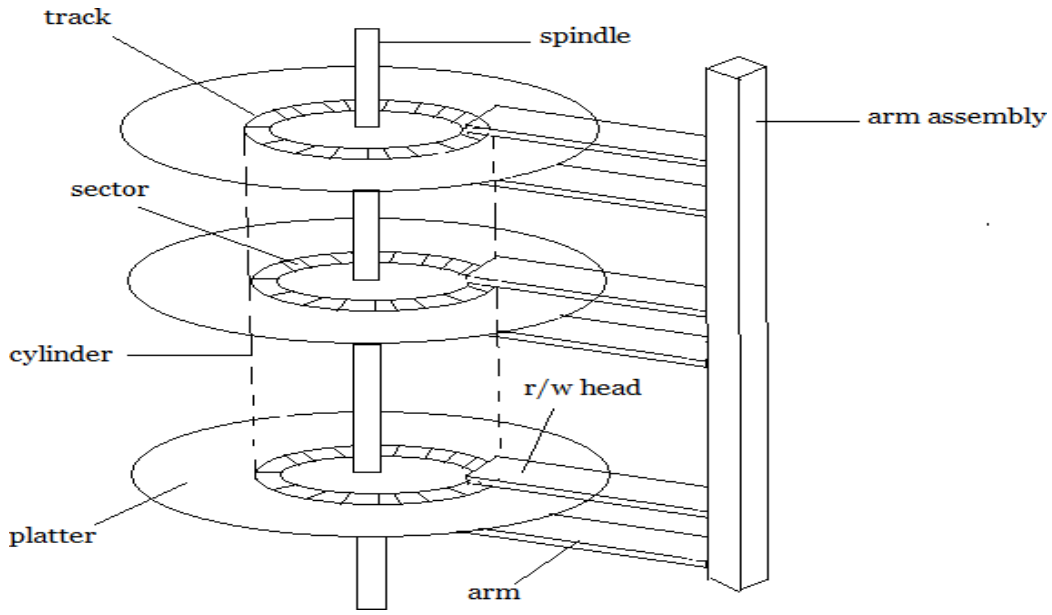
Thepage numbers 1,3,4,6are loaded inthe secondarymemory. So thosebits are setto invalid. Remaining pages resides in the main memory, so those bits are set to valid.

Theavailablefreeframesinmainmemoryis5,so5pagesareloaded,remainingframesare used by other process(UBOP).

UNIT-4
DEVICEMANAGEMENT

MagneticDisk Structure:

Inmoderncomputers,mostofthesecondarystorage isinthe formofmagneticdisks.Hence,knowing the structure of a magnetic disk is necessary to understand how the data in thediskisaccessedbythecomputer.



Physicalstructure ofa magneticdisc:-

Platter:-Eachdisc hasaflat circularshapenamed asplatter.Thediameterofplatteris fromrange $1.8^{11}$to 2.25^{11} inches.

TheinformationArestoredmagneticallyonplatter.

Tracks:- Surface of the platter are logically divided into circular

tracks.**Sector:-**Tracksaresubdivided into no.ofsectionstermedas

sector. Each track may contain hundreds of sectors.

Cylinder:-Setoftracksthat areatonecompositionmakesupacylinder.There may be

thousands of concentric cylinders in disc drive.

Read/writehead:- Thisispresentjustaboveeachsurfaceofeveryplatter.

Discarm:-Theheadsareattached to adiscarmthatmovesalltheheadsasaunit.

Note:-Thestorage capacityofanormaldiskdriveismeasured byGB.

Seektime:- Thetimerequiredtoreachthedesired track bythereadandwritehead isaseektime. There are 2 components to calculate seek time.

i) Internalstartuptime
ii) Thetimetakentotraversethecylinder Ts
= m + n + s
WhereTs=Estimatedseektime n
= no. of tracks traverses
m=constant
s=startuptime.

Rotationaldelays time:-

Thetimerequiredtoreachthedesiredsectorbywrite/readhead iscalledrotationaldelaytime. Generally the average of rotational delay is between 100 to 200 ms.

Diskschedulingalgorithm:-

Whenever aprocessrequest ani/p, o/poperationsfromthebasic it willissueasystemcalltotheOS. The request have been processed according to some sequence. This sequence has been made by various algorithm named as disk scheduling algorithm
DiskschedulingalgorithmSchedulesalltherequest properlyand insomeorder. This
algorithm. Is implemented in a multi programming system.
Followingarethediskschedulingalgorithm
i) FCFS
ii) SSTF(Shortestseektime first)
iii) SCAN
iv) C-SCAN(Circular scan)
v) Look
vi) C-Look

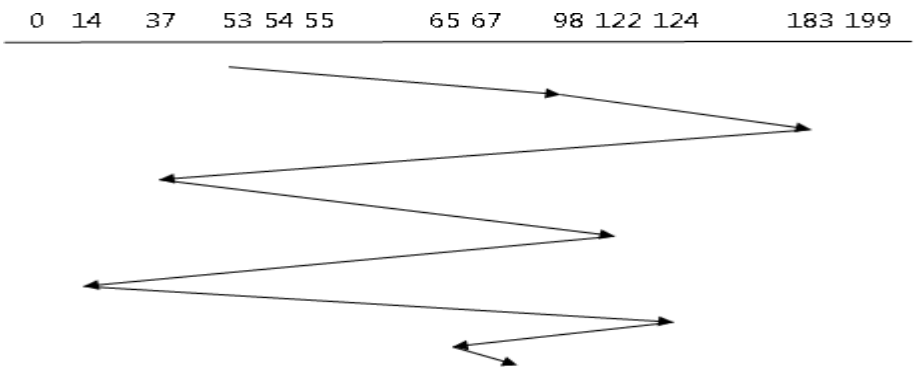
FIRSTCOME,FIRSTSERVE

Itisnotanefficientalgorithmbutitisfairinschedulingthedisk access.

Forexample, wearegivena listofrequest fordisk I/Otoblocksonacylinder. 98, 183,

37, 122, 14, 124, 65, 67

Ifthestarting pointis53thenthe accesswould belikebelow



Thebig jumpfrom183to37couldbeavoided ifsomehow14,37and122,124areservedtogether. This indicates the problem with the FCFS algorithm which is larger head movement.

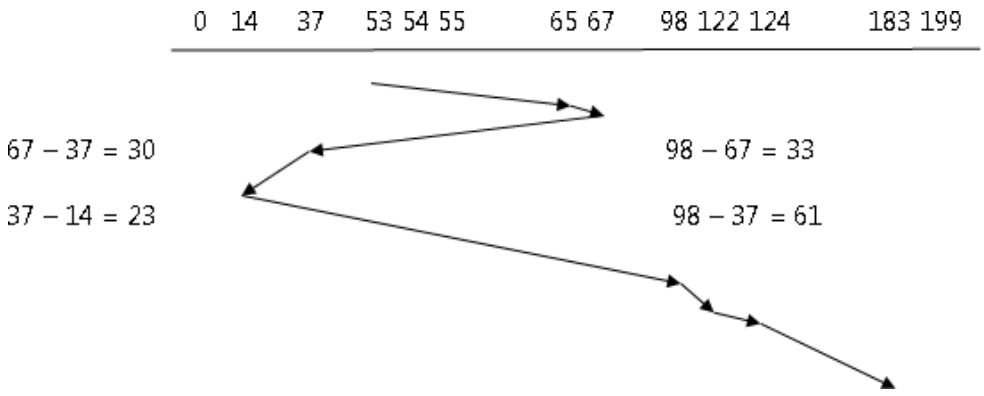
SSTFSCHEDULING

The main ideaoftheShortest-Seek-Time-First algorithmistoservicealltherequestscloseto the current position of the head before moving far awayto service other requests.

Example:

Consideringourprevioussequenceofdiskblocksaccess. Queue

= 98, 183, 37, 122, 14, 124, 65, 67



ThereisasubstantialimprovementcomparedtoFCFSalgorithm.Thetotalheadmovement is as follows.

65–53=1237–14=23124–122=2

67–65=298–14=84183–124=59

67–37=30122–98=24

TotalHead Movement=236cylinders.

But suppose 14 and 183 are a queue and a request near 14 came,it will be served and next one isalso closeto14came, it willbe served firstandthiswillleadto starvationof183 inthe queue.

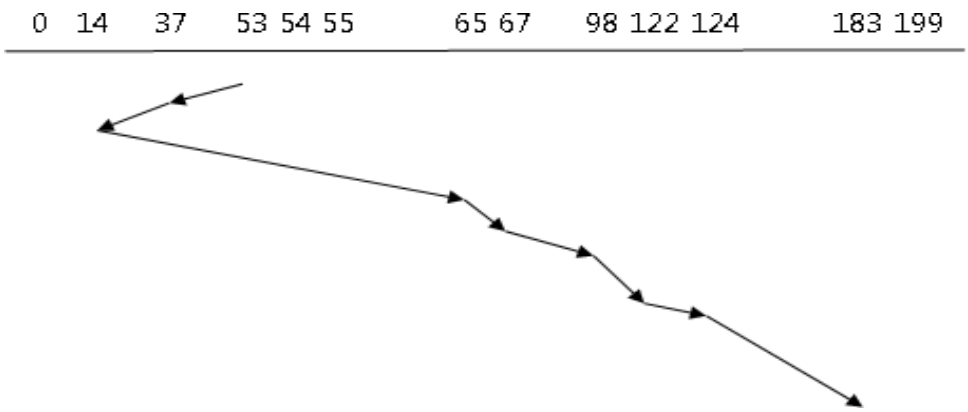
SSTFisanimprovementbutnottheoptimalalgorithm.

SCANSCHEDULING

Inthisalgorithm, the disk armworkslikeanelevatorstarting at oneend servicing allthe wayup totheotherendandthenstartfromtheotherendinreverseorder.TousetheSCANalgorithm,we need to know two information.

1. DirectionofScan
2. Startingpoint

Let'sconsiderourexampleandsupposethediskstart at53andmoveinthedirectionof0.

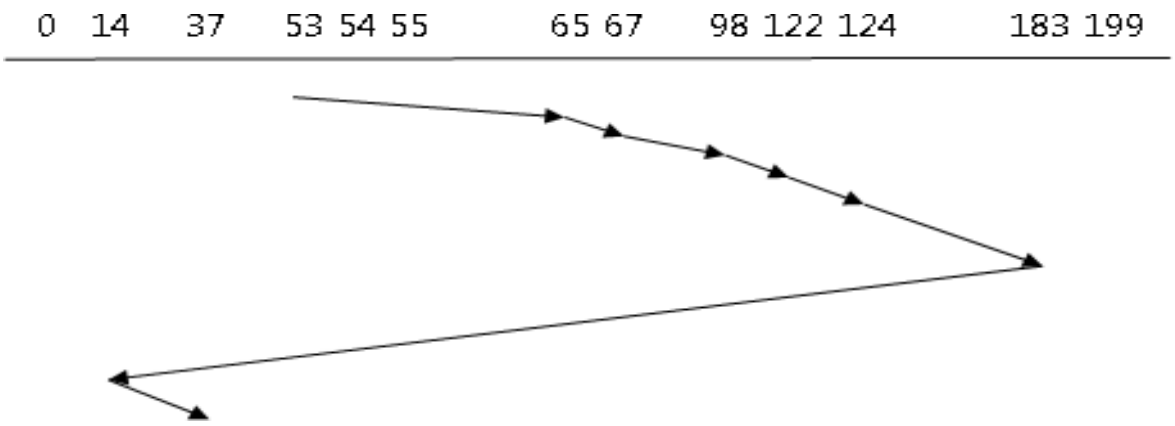


TotalHeadMovement

53–37=1667–65=2124–122=2
37–14=2398–67=31183–124=59
65–14=51122–98=24
Totalheadmovement=158
TheSCAN move inonedirectionandservicealltherequest immediately,but whilereturning inreverse direction it does not serve anyrequest since theyhave been serviced recently. Moreoftherequest isattheoppositeend, wewill seeanalgorithmthat wantstogotheother end directly.

C-SCANALGORITHM

Inthisalgorithm,theheadfromoneendtotheotherservicingrequestalongtheway, however,it does not do a reverse trip and go to the beginning directlyas if it is a circular queue.



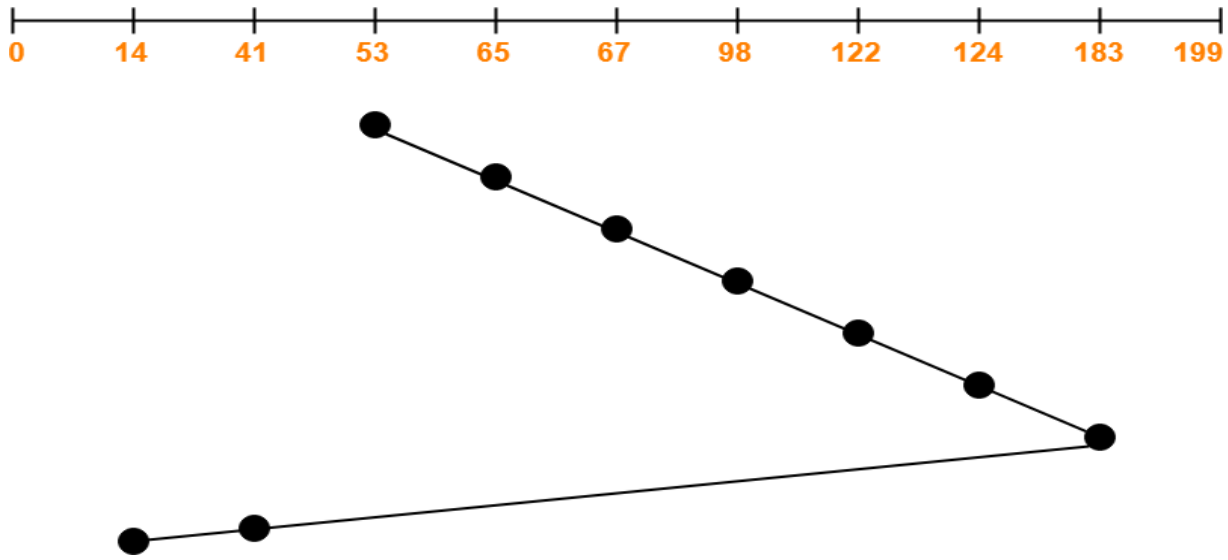
TotalHeadmovement

183-124= 5998–67=31183–14= LookforRequest
124–122=267–65=237–14=23
122–98=2465–53=12
TotalHeadMovement=153

LOOKDiskSchedulingAlgorithm-

- LOOKAlgorithmisanimprovedversionofthe SCAN Algorithm**.**
- Headstartsfromthe first requestatoneendofthediskand movestowardsthe lastrequestat the other end servicing all the requests in between.
- Afterreachingthelastrequestattheotherend, headreversesitsdirection.
- Itthenreturnstothefirst requestatthestarting endservicingalltherequestsinbetween.
- Thesameprocessrepeats.

Consider adiskqueuewithrequestsforI/Oto blocksoncylinders98,183,41,122,14,124,65, 67.TheLOOKschedulingalgorithmisused.Thehead isinitiallyat cylindernumber53moving towards larger cylinder numbers on its servicing pass. The cylinders are numbered from0 to 199.



Totalheadmovementsincurredwhileservicingthese requests

= (65–53)+(67–65)+(98–67)+(122–98)+(124–122)+(183–124)+(183–41)+(41–14)

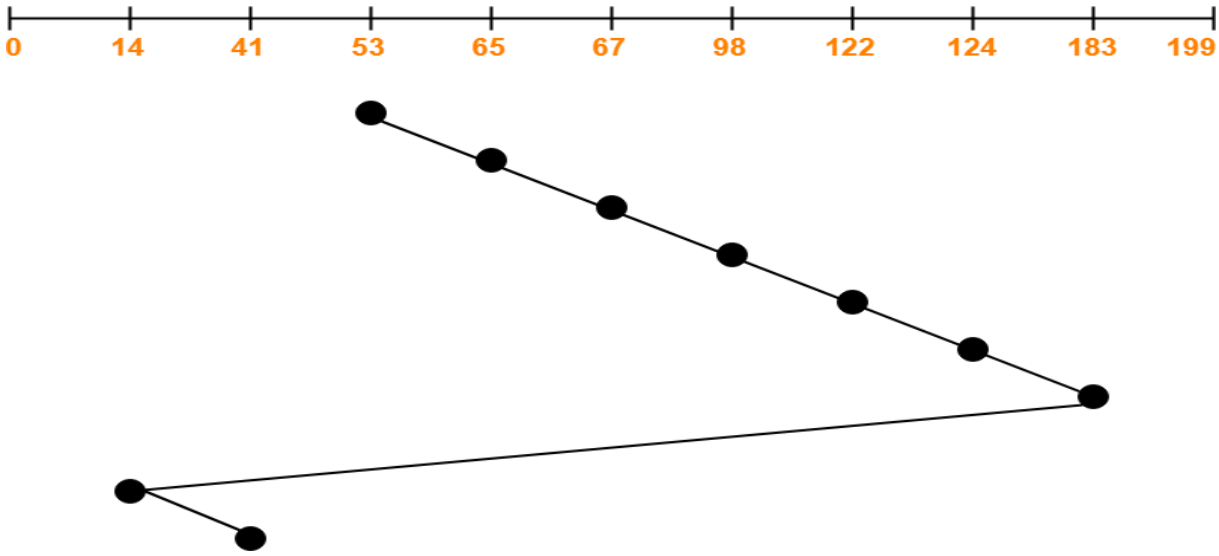
=12+2+31+24+2+59+142+27

=299

C-LOOKDiskSchedulingAlgorithm-

- Circular-LOOKAlgorithmisan improved versionoftheLOOKAlgorithm.
- Headstartsfromthe first requestatoneendofthediskand movestowardsthe last request at the other end servicing all the requests in between.
- Afterreachingthelastrequestattheotherend,headreversesitsdirection.
- Itthenreturnstothe first request atthestartingendwithout servicinganyrequest in between.
- Thesameprocessrepeats.

Consider a disk queue with requests for I/O to blocks on cylinders 98, 183, 41, 122, 14, 124, 65, 67. TheC-LOOKschedulingalgorithmisused.Thehead isinitiallyat cylindernumber53movingtowards larger cylinder numbers on its servicing pass. The cylinders are numbered from 0 to 199.



Totalheadmovementsincurredwhileservicingthese requests

= (65–53)+(67–65)+(98–67)+(122–98)+(124–122)+(183–124)+(183–14)+(41–14)

=12+2+31+24+2+59+169+27

=326

Devicemanagement

Themainfunctionsofthedevicemanager are:

1. Monitorthestatusofalldevices,including storagedrives,printersandotherperipherals
2. Enforcepre-setpoliciesonwhichprocessgetswhichdevicefor how long
3. Dealwiththeallocationofdevices toprocesses
4. Dealwiththede-allocationofdevicestoprocesses, bothat atemporarybasis(e.g. whenthe process is interrupted) and on a permanent basis (e.g. when the process is completed).

Devicemanagementtechnique:-

Thereare3techniquesfordevicemanagement,i.e.

i) Dedicated.
ii) Shared
iii) Virtual

Dedicated:-

- Thesearedevicesthat areassignedtooneprocessat atime, andtheprocessonlyreleasesthedevice once it is completed.
- Theproblemwiththisisthat it meansonlyoneuserisusing it at atime,and it might be inefficient ifthe device isn't being used 100% ofthe time that it is being locked bythe user.
- Ex.:-Printer,cardreaderanddisk.

Shared:-

- Thesearedevicesthatcanbesharedbetweenseveralprocesses.
- Consideringanexample likeaharddisk, itisshared,butinterleavingbetweendifferent processes'requests.
- Allconflictsfordeviceneedtoberesolvedbut predeterminedpoliciesto decidewhichrequest is handled first.

Virtual:-

- ThesearedevicesareacombinationofDedicatedand SharedDevices.
- Soaprinterisadedicateddevice,butusingthespooling(queues)meansitcanbeshared.
- Aprint job isn't sent straight to the printer, instead it goesto the disk (spool) until it is fully preparedwithallthenecessaryprintersequencesand formatting, thenitgoestotheprinter,ensuring that the printers (and all I/O devices) are used efficiently.

I/Otrafficcontroller:-

I/Otrafficcontroller, controlallthedevicetrackorchannel.The traffic controller maintain all the status information.

Thetrafficcontroller, attend3questions,i.e.

i) Isthereapath,availabletoserverorI/Orequest?
ii) Ifmore,than,onepathavailable.
iii) Ifnopathcurrentlyavailable,when, onewillbefree.

Inordertoanswerthesequestion,I/Otrafficcontrolleruse oneofthefollowingdatabase,i.e.,

i) Unitcontrolblock(UCB)
ii) CentralUnitControlBlock(CUCB)
iii) ChannelcontrolBlock(CCB)

UCB	CUCB	CCB
⇒ Device unit identification. ⇒ Status of device. ⇒ List of control unit, connected to the device. ⇒ List of processes, waiting for this device.	⇒ It is control unit identification ⇒ List of device connected to the control unit. ⇒ List of channel connected to the control unit ⇒ Waiting for the control.	⇒ Channel identification. ⇒ Status of the channel. ⇒ List of control unit connected to the channel. ⇒ List of the process, waiting for the channel

I/O scheduler:-

Ifthereare moreI/Orequest,pending, then,availablepathisnecessarytochoose,which,I/O requestis satisfied, first. Then, the process ofscheduling, is applied here, and it is known as I/Oscheduler.

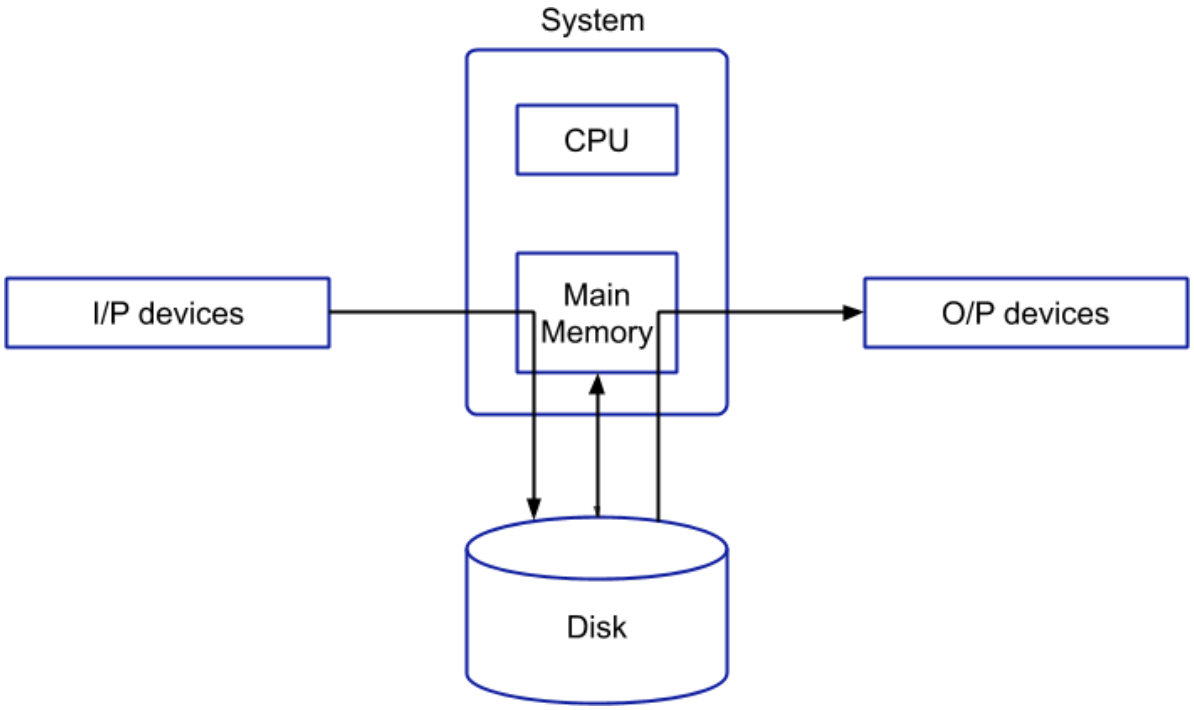
I/O device handler:-

- I/OprocessestheI/Ointerrupts,handleserrorcondition,and providesdetailed scheduling algorithms, which are extremely device dependent.

- EachtypeofI/O devicehasowndevicehandler algorithmlikeFCFS,SSTF,SCAN.

Spooling:

- SPOOLisanacronymfor**simultaneousperipheraloperations on-line**.
- It is a kind of buffering mechanism or a process in which data is temporarily held to beusedand executed by a device, program or the system.
- Dataissenttoandstoredinmemoryorothervolatilestorageuntiltheprogramor computer requests it for execution.



- Spoolingusesbuffer tomanagefilestobeprinted.
- Fileswhicharespooledarequeuedandcopiedtoprinterone at atime.
- TomanageI/Orequests,operatingsystemhasacomponent thatiscalled spooler.
- Spooler managesI/Oreqeststoaprinter. Spooler operatesinthebackgroundandcreatesa printing schedule.

Racecondition:-

Raceconditionisasituation,where,severalprocess,accessand manipulate,samedata,concurrentlyandthe execution depend on a particular order.

UNIT-5

DEADLOCKS

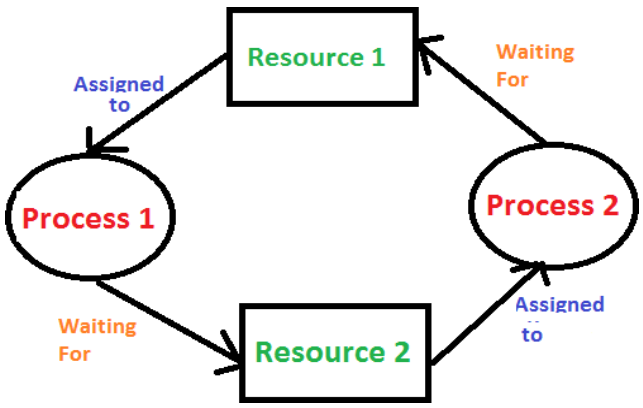
SystemModel:
Asystemconsistsofafinite number ofresourcestobedistributed amonganumber ofcompeting processes.The resources arepartitioned into severaltypes eachof whichconsists ofa number of identicalinstances. Aprocess may utilized a resources in the following sequence

1) **Request:-**processrequest foraresourcethroughasystemcall.Ifthe resourceisnotavailableitwillwait.
   Example:systemcallsopen(), malloc(),new( ),andrequest().
2) **Use:-**Aftergettingtheresource,theprocesscanmakeuseof itbyperforming the execution.
   Example:printstotheprinterorreadsfromthefile.
3) **Release:-**Afterthecompletionofthetasktheresource isnotrequiredbythat process, inthat it shouldbe released.
   Example:close(),free(),delete(), andrelease().

Resourcessuchas->CPU, memory, I/O devicesetc.Whenaprocessrequestsforaresourcesthenif it is freethen it will be allocated to that process. But if theresource is busywithother process thenthe previous processhas to wait till that resource is free.

Deadlock:*Deadlock*isasituationwherea set ofprocessesareblocked becauseeachprocess isholding aresource and waiting for another resource acquired by some other process.

Forexample, inthe belowdiagram, Process1isholding Resource1andwaiting forresource2whichisacquired by process 2, and process 2 is waiting for resource 1.



REASONS/NECESSARYCONDITIONSFORARISINGDEADLOCK:--

Adeadlocksituationcanariseifthefollowingfour conditionholdsimultaneouslyinthesystem.

1. Mutualexclusion
2. Holdand wait
3. Nopre-emption
4. Circularwait

1) MUTUALEXCLUSION:-At leastoneresourcesmust beheldinanon-sharablemode.That meansonly one process can use that resource at a time.
EX:printerisnon-sharableButHDDisasharableresource.
Sothat iftheresourceisnot freethentherequestingprocess hastowait tilltheresourceisreleasedbythe other process.

2)HOLD AND WAIT:-There must be aprocess which is already holding using one resource andrequesting(waiting) foranotherresourcewhichiscurrentlyheld byanotherwaiting process.

3)NOPREEMPTION:-Resourcescannotbepre-empted.That meansaresourcecan't releasedbythe process unless until it has completed itstask. I.e. printer will be released only when printing work is finished.

4)CIRCULARWAIT:-Supposetherearen-processes{P0,P1,P2,……Pn-1}theyallare waitingprocesses.

P0iswaitingfortheresourceholdbyP1.

P1iswaitingfortheresourceholdbyP2.

P2iswaitingfor theresourceholdbyP3.

_____

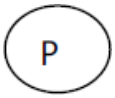
– –

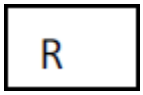
Pn-1iswaiting fortheresourceholdbyP0.

Ifallabove4 conditionsaresatisfied inasystem, thendeadlock mayoccur but ifanyoneofthecondition (criteria) is not satisfied then deadlock will never occur.

Resourceallocationgraph (RAG):-

- Adiagrammatic representto det.Theexistenceofdeadlock inthesystembyusing agraph named as RAG.
- It isa directedgraph.

- RAGconsistsofseveral no. ofnodesandedges.
- Itcontains i)Process 

nodeCircleii)Resourcenode 

Square.
- Thebulletsymbolwithinthe resourceisknownasinstancesofthat resource.
- Instanceofresourcesmeansidenticalresourcesofsametype.
- Thereexist2kindsof edges.

i) Request edge.

ii) Allocation/assignment edge

REQUESTEDGE:-Wheneveraprocessrequestforresourcesthenitiscalledarequestedge.

- Arequestedge isdrawnfromtheprocesstoresource.

ASSIGNMENT EDGE:- Whenever a resource is allocated to a process the request edge is converted to an assignment edge from the instance of the resource to the process.

NOTE:-IftheRAGcontainsNO CYCLE,thenthereisNODEADLOCKinthe system.

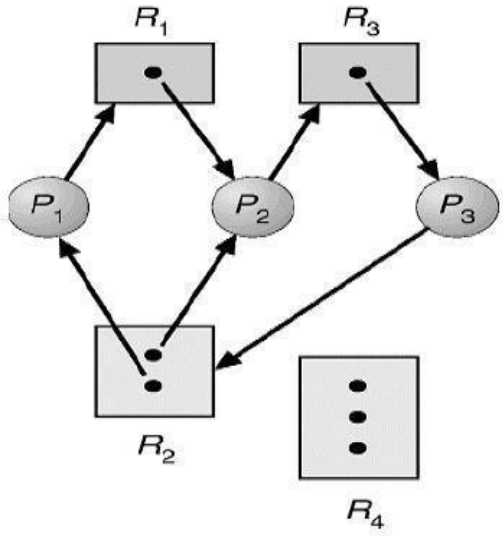
- Ifthe RAGcontainsa CYCLE,thenthere MAYBEADEADLOCKinthesystem.
- Iftheresourceshave exactlyoneinstance thena cycleindicatesadeadlock.
- If the resources havemultipleinstanceperresourcethen acycleindicates that"there maybe deadlock".
- Process waitforgraph(PWFG) :- Aprocess waitforgraph can be obtained by removing/collapsing resource symbols in the RAG.

Theresourceallocationgraphshownin figurehasthe following situation.

- ThesetsP,R,E
- P= {P1,P2,P3}
- R= {R1,R2,R3,R4}
- E={P1→R1,P2→R3,R1→P2,R2→P2,R2 →P1,R3→P3}

Theresourceinstancesare

- ResourceR1hasoneinstance
- Resource R2 hastwoinstances.
- ResourceR3hasoneinstance
- Resource R4hasthreeinstances.



Theprocess states are:

- ProcessP1isholdinganinstanceofR2andwaiting for an instance of R1.
- ProcessP2is holdinganinstanceofR1andR2andwaitingforaninstanceR3.
- ProcessP3isholdingan instanceofR3.

Thefollowing exampleshowstheresourceallocationgraphwithadeadlock.

- P1->R1->P2->R3->P3->R2->P1
- P2->R3->P3->R2->P1

MethodsforHandlingDeadlocks

Deadlockscan behandledbyanofthefollowingmethods.

- Deadlockprevention
- Deadlockavoidance
- Deadlockdetectionandrecovery

DEADLOCKPREVENTION:-Deadlockscanbepreventedfromoccurringbypreventingoneofthe necessary four conditions. I.e. mutual exclusion, hold and wait, no pre-emption and circular wait.
Ifone oftheconditioncanbepreventedfromoccurringthendeadlockwillnotoccur.

EliminateMutualExclusion
Itis not possibleto dis-satisfythe mutualexclusionbecausesomeresources, suchasthetapedriveand printer, are inherently non-shareable.

EliminateHold andwait

1. Allocate allrequired resourcesto the process before the start of its execution, this way hold and wait condition is eliminated but it will lead to low device utilization. for example, if a process requiresprinterat alatertimeandwe haveallocatedprinterbeforethestartofitsexecutionprinter will remain blocked till it has completed its execution.
2. Theprocesswillmakeanewrequest for resourcesafter releasingthecurrent setofresources. This solution may lead to starvation.

EliminateNo Preemption
Preemptresourcesfromtheprocess whenresourcesrequiredbyotherhighpriorityprocesses.

EliminateCircularWait
Eachresourcewill beassignedwitha numericalnumber. Aprocesscanrequest theresources increasing/decreasing. order of numbering.
For Example, ifP1process isallocatedR5resources, nownexttime ifP1ask for R4,R3lesser thanR5such request will not be granted, only request for resources more than R5 will be granted.

DEADLOCKAVOIDANCE

-> Toavoiddeadlockitisrequiredtokeep theadditionalinformationof theprocessi.e.the operating System willbegiven prior information about the process such that

- Whichprocesswillrequestforwhichresource.

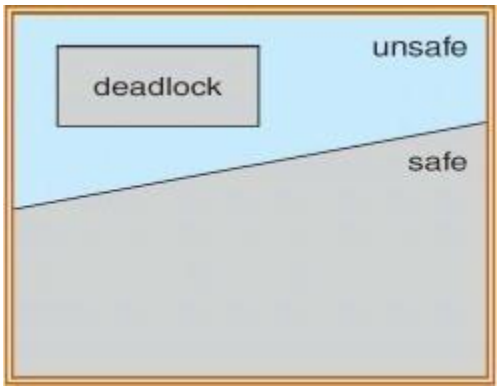
- Atwhattimeandforhowlongtime.

->So thattheoperating Systemfind outa sequenceof execution.

->Ifalltheprocessisexecuteinthe sequence then systemwillneverenterinto adeadlockstate.

SAFESTATE:-Astateissafeifthesystemcanallocatetheavailableresourcestoeachprocessin same order and still avoid deadlock.

*SAFESEQUENCE:-Asequenceofprocesses(P1,P2, and P3----------Pn) is a safe sequence if the available resources can be allowed to the processes in that sequence and avoid deadlock.



Ifnosafesequenceexiststhesystemissaidtobe

i

nunsafestate.Anunsafe state mayleadtoadeadlock.

Ex.Supposea systemcontains 12tapedrives

Process	max.	Allocation	need
P0	10	05	5
P1	04	02	2
P2	09	02	7

Free= 12-9=3

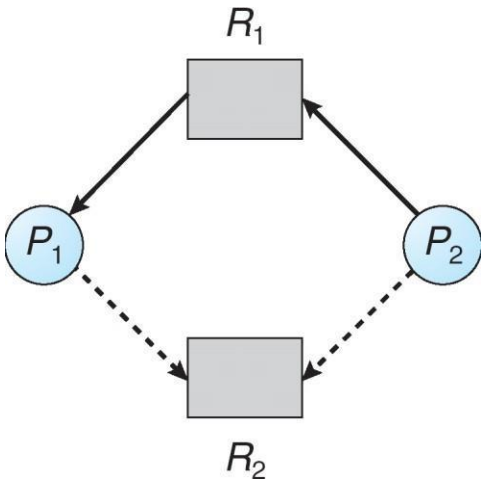
So,thesafesequenceis<P1,P0,andP2>.Ifthesystemwillalwaysremaininsafestatethendeadlock will never occur. So, whena process requests for aresourcethat is currentlyavailable the systemmust decide whether that resource will be allocated or the process will wait. The request will be granted only if the allocation leaves the system in a safe state.

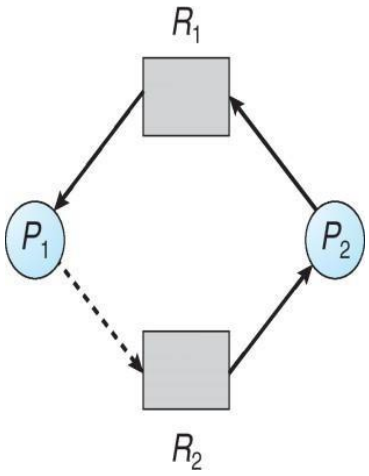
This methodisused onlywhenthesystemcontain onetypeofresourcehavingmultipleinstances.

Resourceallocationgraph:-(MultipleResourcesHavingsingleinstance)

We can use this algorithm for deadlock avoidance if the system contains different types of resources but each is having single instances.

- Inthisgraph, besideassignment edgeandrequest edge,thirdedgeknownas "claimedge" isadded. A
- claim edge from process Pi to Ri indicates that the process Pi may request for resourceRj in FUTURE. Claim edge is similar to request edge but it is represented as dashed line,
- Ifaprocessrequest for resourceRjthenthat request onlybegranted if byconverting the requesting edge Pi->Rj to the assignment edge Rj->Pi does not form a cycle.





- If there aretwo claimedges for the same resource then it canavoid. If onlyone of the processesis allocated the resource R1 then a deadlock can arise.

Banker'salgorithm:-(MultipleResourceshavingmultipleinstances)

Theresourceallocationgraphallocationisnotapplicabletoaresourcehavingmultipleinstancesof each resource type.

- Thisalgorithmusedforsystemhavingmultipleresourcesalongwithmultipleinstances.
- Banker'salgorithmislessefficienttheRAG.
- This name was choose because it could be used in a banking system to ensure that bank never allocates its available cash such that it can no longer satisfy the needs of all its customers.
- Whenanew processentersintothe system

  - Itmustdeclaremaximumnumberofinstancesofeachresourcetypethatitmay need.
  - Thisnumber shouldbelessthanthetotalnumberofresources.

- Whena processrequests a setofresources, thesystemmustdetermine:-
  - Whether theallocationofthese resourceswillleave thesysteminasafe state.
  - IfYES,theresourcesareallocatedto chatprocess.
- IsNO,theprocesshavetowaituntilsomeotherprocessreleasesenoughresources.
- Banker'salgorithmconsistsoftwo parts.
  - Safetyalgorithm
  - Resourcerequestalgorithm.
- Thesafetyalgorithmisused todeterminewhetherasystemisin safestateornot.
- The resource request algorithm is used determine whether or not a request generated bya process for a resource, would lead the system to an unsafe state.
- The algorithm uses several data structures such as vector & matrixes. Ex. Let in a system there are 'n' processes and 'm' resources.

1) **AVAILABLE**:-Itisanarray/vectoroflength'm'indicatesthenumberofavailableresourceof each type.

Ifavailable[J]=kmeansthere arekinstancesofresourcetypeRjavailable.

2) **MAX**:-Itisammatrixdefinesthemaximumdemandofeach process.

If max[i,j]=k,thentheprocessPimayrequestatmostkinstancesofresourcetypeRj.

3) **ALLOCATION:** - It is an n*m matrix definesthe number of resources of each type currently allocatedto eachprocess.Ex.Ifallocation[i,j]=k,thenprocessPiiscurrentlyallocatedk instances of resource to Rj.

4) **NEED**:-Itisann*mmatrixindicatestheremainingresourcesneedeachprocessifneed [i,j]=k, then the process Pi may need k more instances of resource type Rj to complete its task.

Need[i,j]=max[i,j] – allocation[i,j].

Safetyalgorithm

Thealgorithmforfindingoutwhetherornotasystemisinsafestate.Itcanbedescribedas follow.

STEP1:-workisa vectoroflength
m.Finish is a vector of length n.

Work=
available.
Finish[i]=false.
For i=1, 2,3----------n.

STEP2:-findan isuch that

Finish[i]=false.
Need i <=work.
Ifnosuchithengotostep4.

STEP3:-work=work

allocationFinish[i] = true.

Gotostep2.

STEP4:-Iffinish[i]=trueforallI,thensystemisin safestate.

- Thecomplexityofthealgorithmis0(m*n^2)i.e.thealgorithmmayrequireonorderofm*n^2 operation to decide whether a state is safe.

Resourcerequestalgorithm

Requestisavectoroflengthm.

1. Ifrequest i<= need igo to step-2.
   Else raise an error condition.
2. Ifrequest <availablegoto step-3.
   Otherwise Pi must wait.
3. Available=Available-Request(i)
   Allocation = Allocation + RequestiNeed i= need i-Request i

DEADLOCKDETECTION

Ifa systemdoes not employeither a deadlock prevention or a deadlock aavoidance algorithm, thena deadlock situation may occur.

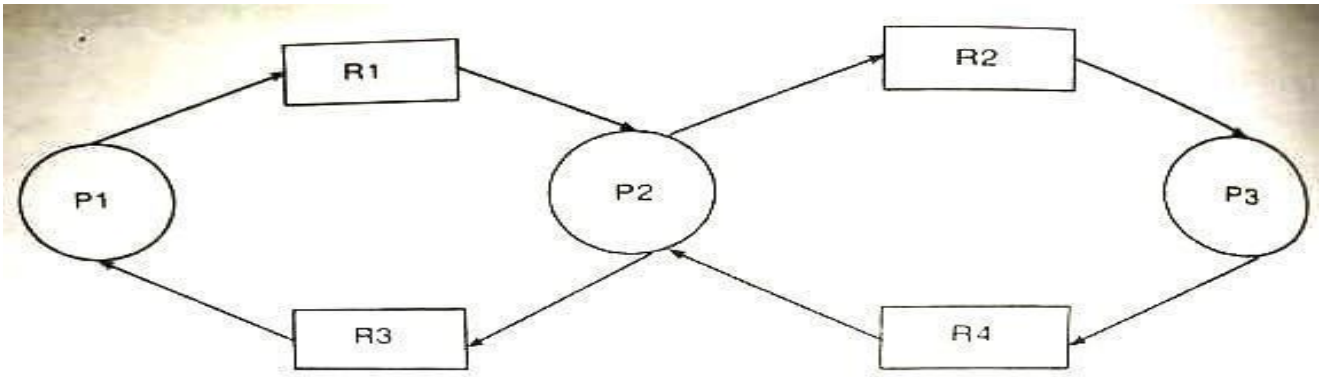
- Soatthetimeorenvironmentthesystemshouldprovide:-
- Analgorithmthat willcheckwhether thedeadlock hasoccurredintothesystem(deadlock detection).
- Analgorithmtorecoverthesystemfromdeadlock.State(deadlockRecovery)

Singleinstanceofeachresourcetype:-

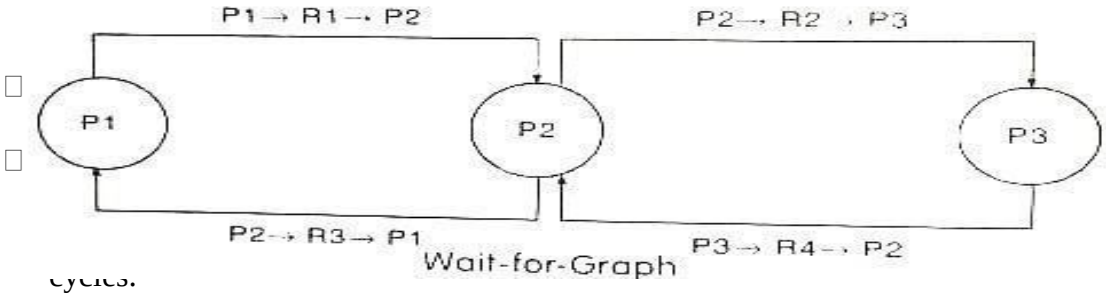
Ifallthe resources have onlysingle instance, thenwe candetect adeadlock state byusing “wait-for- graph” (WFG).

Itissimilar toRAGbutonlydifferenceisthatheretheverticesareonlyprocesses.

- ThereisanedgefromPitoPjifthereis anedgefromPitoR andalsoonedge fromRtoPj.



Resource allocation Graph



Wait-for-Graph

with cycles.

- In thisfig.thereistwocyclesoneisP1 to P2 to P1.SecondareP2 to P3 toP2.Sothesystem consisting of two deadlocks.

Multiple/severalinstancesofResourceType:-

Thewaitforgraphmethodisnotapplicableto severalinstanceofaresourcetype.

- So, we need another method to resolve this problem. The algorithm used is known as “deadlock detection algo”.
- Thisalgorithmlike“banker’salgo”anditusesseveraldatastructures.
  -Available: -Avectoroflength‘m’indicatesthenumberofavailableresourcesofeachtype.
  -Allocation: - An n*m matrix defines the no. of resources of each type currently allocated to each process.

-Request: - An n*m matrix indicates the currentrequestof each process. If request[I,j]=k, then the process Pi is requesting k more instances of resource type Rj.

Detectionalgorithm:-

STEP 1:-

(work=available)

AvailableI,fori=1,2,3,4--------- If

Allocation i! = 0

Thenfinish[i]

=false

Otherwisefinish[i]=true.

STEP2:-Findanindexsuchthat

Bfinish[i]=false

Ijrequest[i]<=availableor

workFinish [i] = true

Gotostep-2

STEP4:-Iffinish[i]=false,forsomei,thenthesystemisindeadlockstate.I.e.processPiis deadlocked.

RECOVERYFROM DEADLOCK

Whenthedetectionalgorithmdetectsthatdeadlockexistsinthesystemtherearetwo methods for breaking a deadlock.

- Onesolutionissimplytoabort onebyoneprocessto Breakthecircularwait.
- Secondsolution istopre-emptsomeresourcesfromoneor moreofthedeadlock process.

⇨ **PROCESSTERMINATION:-**

Thismethodusedtorecoverfromdeadlock.Weuseoneoftwomethodsforprocesstermination.

- Abortalldeadlockedprocess
- Abortonebyone processuntilallcycleis eliminating.

i. **Abortalldeadlockprocesses: -** Itmeansthatreleaseallthe processes inthedeadlockedstate and start the allocation from the starting point.

- Itisanexpensivemethod.

ii. **Abort one byoneprocessuntilthe deadlockcycle is broken**:- Inthis method first abortthe one ofthe processes inthe deadlocked state and allocatedthe resources(resourcesfromabort process) to some other process in the DL state.

- Thencheck whetherthe deadlockbrakedornot.
- If YES, then it is ok i.e. deadlock is eliminated. If NO, abort the process from the deadlock state then check.
- Continuethisprocessuntilwerecoverfromdeadlock.
- Thisalsoexpensivemethodbutbetterthanfirstone.

- In thismethod thereisan overheadbecauseevery timetheDLdetection algorithmisinvoked after each process is aborted.
- Ex.Endtaskinwindows.
- Therearesomefeatureswhichdeterminesthe whichprocessto be aborted::
    I. Priorityoftheprocess.
    II. Howlongtheprocess iscomputed andhowlongtimeitisneededtocompletion.
    III. Howmanyresourcesthe processhas currently used.
    IV. Howmanymore resourcesitneedsfor completion.

⇨ **RESOURCE PREEMPTION:-** Thereare three methods toeliminatedeadlocksusing resourcepre-emption. They are-

- Selectiona victim.
- Rollback
- Starvation

Selectingavictim:-SelectavictimresourcefromtheDLstate,andpre-emptthatone.

- Selectionofvictimdonesothatthecostwillbeminimum.

Rollback:-Whenaresourcewillbepre-emptedfromaprocess,thennaturallytheprocesswillgo into the waitingstate.Sowemustrollbacktheprocesstosomesafestatesothatitwillbestarted fromsame state againor not fromthe beginning. I.e. rollback the processes and resources upto some safe state, and restart it from that state.

- This methodrequiresthesystemto keepmoreinformationaboutthe stateofallthe runningprocesses.

Starvation:-Howtoensurethatstarvationwillnotoccur?Itshouldbekeptinmindthatresourcesshould not be pre-empted from etc. same process again and again; otherwise that process will not be completed for a long period of time.

- That is a process cana resources can be picked as a victimonlya finite number oftimes,not morethan that, otherwise it create a starvation.

Unit-6

File Management

File is a primary resource in whichwe can store information and can retrieve theinformation when it is required.

Therecanbeanumericdatafile,analphabeticdatafileoralphanumericandbinarydatafiles. In general terms a file is sequence of bits, bytes, lines or records.

All computer applications need to store and retrieve information. As computers canstore information on various storage media, in the same way, the operating System provides a logical view of information storage on various secondary storage media like magnetic disks, magnetic tapes and optical disks etc.

Thisuniformlogicalstorageunit iscalledasfile. Soafile isthecollectionofrelated information, which is stored on secondary storage.

FILEATTRIBUTES

Afilehasdifferentattributes.Theattributesmayvaryfromoneoperating Systemtoother.

*Name- A name is usually a string of characters. A symbolic name which is in human readable form.

*Identifier-itisusuallyanumberandisauniquetagthatidentifiesthefilewithinthefilesystem.it is a unique identification of a file which is internal to the system.

*Type-Normallyexpressed asanextensiontothefilename.Itindicatesthe typeoffile.

Ex: .exe-executablefile, .src-sourcefile,.obj-objectfile

*Size-thecurrentsizeofthefile(in bytes)

*Location-itisapointertothelocationwhereafileisstoredinsecondarymemory.

*Protection-Itspecifiestheaccesscontrolinformation.itcontrolswhocandoreading ,writing,executingandsoon.

*TimeandDate-itspecifiestimeofcreationand filecreateddate.

*User identification-this isusefulforprotectionandsecurityandlastusagemonitoring.

FileSystem

Thefilesystemconsistsof2distinctsubcomponents.A

.collectionoffiles,eachstoringrelateddata

B. Directorystructure, whichorganizes and provides information about all the file in thefile system.

FILEORGANIZATION

Fileorganizationreferstothemannerinwhichtherecordsofafileareorganizedonthesecondary storage.

Basicallyfile isasetoflogicalrecords.It isallocatedadiskspace intermsofphysical blocks. The most

common file organization schemes are:

- Sequential
- Direct
- Indexed
- Partitioned

Sequential:-Inthismethod,informationorrecordthat isstoredinafileisprocessedinasequence .i.etherecordsarestored strictlyin thesameorderastheyoccurphysicallyin thefile.

Direct:-Therecordsarestoredinanyorderassuitedforapplication.The system

supports random or direct access of any record in the file.

Indexed:-Inthismethod,anindexiscreatedforthefile.Thisindexcontainsthepointers(physicaladdress) for various blocks or records.

Partitioned:-Inthismethod,file ispartitioned intosequentialsubfiles.Each sequential

sub file is called a member of the partitioned file.

FILEOPERATION

Todefineafileinapropermanner ,there aredifferentoperationsareperformedonfiles.

Toallowstorageandretrievalof information fromafiledifferent systemprovide different operations.

Themostcommonoperationsthatcanbeperformedonafileasfollows:

*__create-__2stepsareneededtocreateafile.

-checkwhether thespaceisavailableor not.

-ifyes,2ndoneismadeanentryfornew file.

*__write-__Towriteafile,wehavetoknow2things

i) Name ofthefile.

ii) Informationordatatobewrittenonthefile.

Thesystemfirst searchestheentiregivenlocationsforthefile.ifthe file is found,thesystemmust keep a pointer to the location in the file where the next write is to take place.

*__Read:-__Toreadafile,firstofallwesearchthe directoriesforthefile.

If file is found, the system needs to keep a read pointer to the location in thefile where the next read is to tae place.

Oncethereadhastakenplace,thereadpointeris updated

*__Seek:-__Torepositionthefilepointertospecifiedlocation. Thisis done to

read or write a record at a specified position.

*__Delete:-__Whenthefileisnolongerrequired,itisneededtodeletetheoccupieddiskspace.

To delete a file ,we searchthe directoryfor the named file and whenthe file is found ,werelease all file space so that other files can reuse this space and erase the directory entry.

__Truncate:-__Whenthe user wants to erasethe contents ofa file, but wantsto retain it's attribute. Itisnot necessaryto deletethefileandthenrecreateit.It ispossiblebydoingtruncateoperation. i.etotruncateafile,removethecontentsonly,buttheattributesareasitis.

__Open:-__Aprocessopenmustopenafilebefore usingit.

__Close:-__ Whenalltheaccessesarefinished, theattributesanddiskaddressesarenolonger needed, you need to close the file in order to release the internal table space.

__Append:-__

- Thisoperation isrestrictedformofwrite.
- Itonlyallowsyoutoadddatatothe endoffile.

__Rename:-__

- Itfrequentlyhappensthatauser needstochangethenameofanexistingfile.
- Thisoperationallowsyou torenameanexisting file.

FILESTYPES

When designing a file system, we need to consider whether or not the operating system wouldrecognizeandsupport filetypes.Acommontechniqueforimplementingfiletypesis to include the types as the part of the file names.

- Generally,thenameofthefilesplit intotwoparts:-1-name2-extension(whichis usually separated by 0).
- Thefiletypeisdepending onextensionofthefile.
- The followingsectiondescribesdifferenttypesoffileswiththeirextensionand function

Filetype	Extension	Purpose/function
executable	.exe .com .bin .none	Readyto runorreadyto run m/c language
Object files	.obj	Instructionsareintheform of m/c language. A linked uses this information and coverts itinto executable format.
baten	.bat .sh	Commands to the commandinterpreter.
Sourcecode	.c .cc .cpp .java .pas .asm .f77	Sourcecodeinvarious language.
text	.txt .doc	Usedtocreatetext documents.
Wordprocessor	.wp .rtf .etc	Thesefileallowsvarious word processor formats.
library	.lib .a .so .dll	Explain the entire library functionsinanyprogram.
Print orviewfiles	.ps .pdf .jpg .dvi .sif	ASCII or binary files in a formatprintingorviewing.

archive	.arc .zip .tar	Groupedfiles,compressed file archivingor storage
multimedia	.mov.mpeg .mp3 .ym .mp4 .avi	Thesearebinaryfilesthat contains audio/alu information.

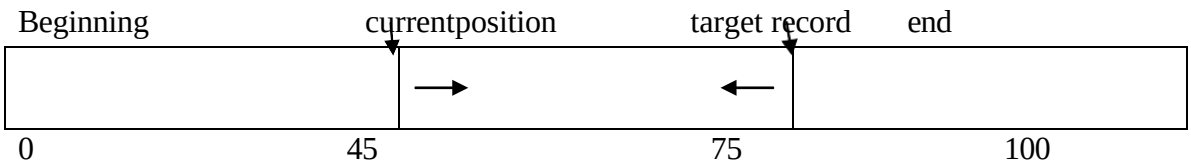
FILEACCESSINGMETHODS

Filesareusedto storedata.The informationpresent inthe fileanbeaccessedbyvarious access methods. Different system uses different access methods .Followingarethemostcommonlyusedaccess methods:

- Sequentialaccess
- Directaccess
- Indexedsequentialaccess.

Sequentialaccessmethod:-

- Thismethodissimplestamongallmethods.Informationinthefileis processedin order, one record after the other.
- Magnetictapesaresupporting thistypeoffileaccessing
- Ex-a fileconsistingof100records,thecurrent positionofread/writehead is45threcord, suppose we wants to read the 75th record, then it access sequentially from 45,46 70,71,72,73,74,75.
- So,the read/writeheadtransverseallthe recordsbetween45to75.



- Sequentialfilesaretypicallyused inbatchapplicationandparallel processing.

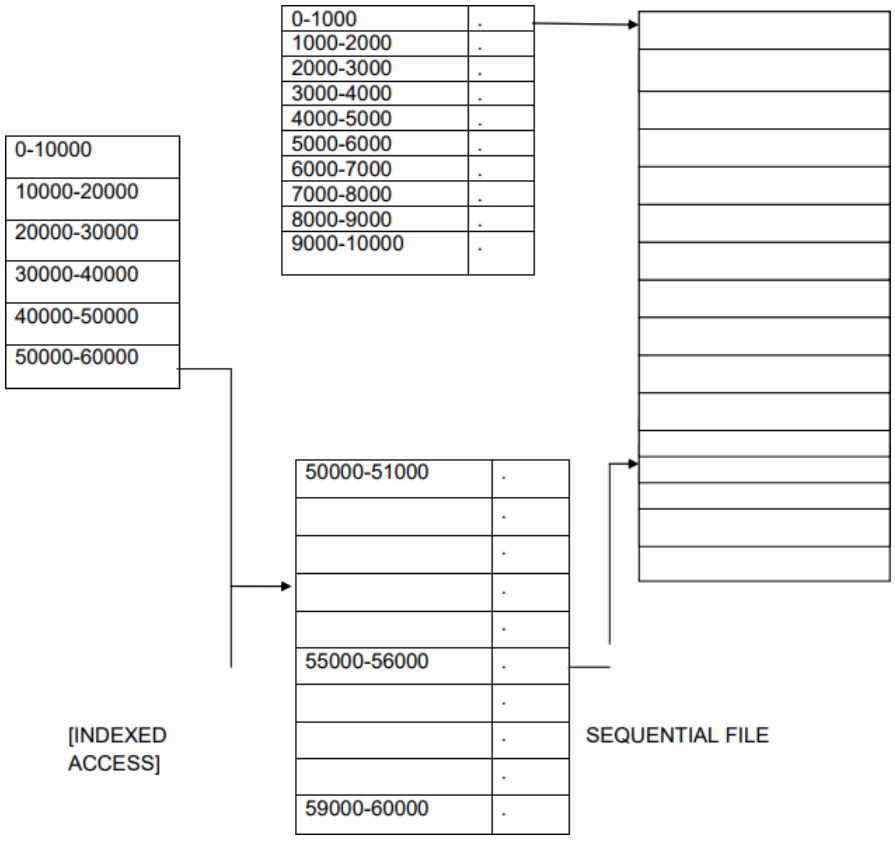
Direct access:-

- Direct accessisalso calledrelativeaccess.Inthismethodrecordscan read/write randomly without any order.
- Thedirect accessmethod isbasedondisk modeloffilesbecause,diskallows random access to any file block.
- Ex:- a disk containing of 256 blocks. The current position of r/w head is 55th block, suppose we want 200th block. Then we can access 200th block directly without anyrestrictions. Anotherexample issuppose aCDcontaining10songs, at the present we are listening the song no.3 and wewant to listen song no. 7, then we can shift from song no.3 to 7 without any restrictions.

IndexedsequentialAccess:-

The maindisadvantage inthesequentialfile is, ittakesmoretimetoaccessarecord.To overcome this problem, we can use this method.

- Inthismethod(ISF),therecordsarestoredsequentiallyforefficientprocessing.But,they can be accessed directly using Index or key field. Keys are the pointerwhich contains adder of various blocks.
- Recordsareorganizedinsequencebased onakeyfield.



- Suppose a file consisting of 60,000 records, the master index divided thetotal index into 6 blocks.
- Eachblockconsistingofa pointertosecondaryindex.
- Thesecondaryindexdivide the 10000records into 10indexes.
- Eachindexconsistingofapointerto itsoriginallocation.1-A key field
  2-Apointerfield

- Suppose we want to access the 55,550threcord, thenthe file managementsystem(FMS) access the index that is 50000 to 60000.
- This block (50000 to 60000) consisting of a pointer, this pointer points to the 6thindexof the secondary index.
- Thisindexpointstotheoriginallocationoftherecordsfrom55000to56000.
- Fromthis itfollowsthesequentialmethod
- That'swhyismethodissaidtobeindexedsequentialfile.sothismethodisneither purely sequential nor purely direct access.
- Generallyindexedfilesareusedinairlinereservationsystemandpayrollsystem.

File directories:-

- The directory contains information about the files including attributeslocations and ownership.
- Sometimesthedirectoriesconsistingofsub-directoriesalso.
- Thedirectoryisitselfafileanditisownedbytheoperatingsystem.
- Itisaccessiblebyvariousfilemanagement units.

Directorystructure:-

Sometimes the file system consisting a millions of files when no. Of filesincreases, then it is very difficult to manage the files.

- Tomanagethesefiles:-
  - Firstgroupthesefiles.
  - Then loadone groupoffileinto onepartition.

- Thiseachpartitioniscalled"directory".
- A directory structure providesamechanism fororganizingmany filesin thefile system.

  Different operationsonthefile directories:-

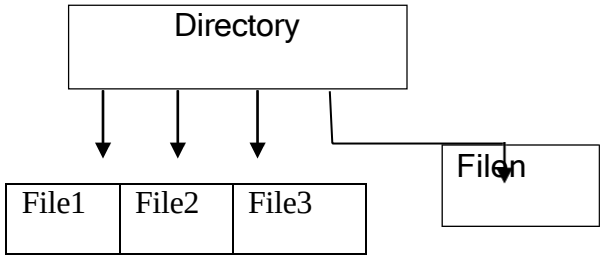
- Searchforafile:-searchthedirectorystructure forrequiredfile.
- Createafile:-wheneverwecreateafilethenweshouldmakeanentryinthe directory.
- Deleteafile:- whenfilenolongerneeded,thenweremoveis fromthedirectory.
- List adirectory:-wecanseethelistoffilesinthe directory.

- Renameafile:-wheneverwewanttochangethenameoffilethenwecanchange.
- Traverseafile:-ifweneedtoaccesseverydirectoryandeveryfilewithinthe directory structure we can traverse the file system.

Therearedifferenttypes ofdirectorystructures areavailabletheyare:-

- Singleleveldirectory
- Twoleveldirectory
- Treestructured directory
- Acyclicdirectory
- Generalgraphdirectory

- Singleleveldirectory:-
  - Itis simplestofalldirectorystructure.
  - Inthisdirectorysystem,onlyonedirectoryisthereandisconsistofallfiles
  - Alfilescontainedinthesamedirectoryname.



Advantage:-Thisschemeisverysimpleandabilitytolocatefileseasily.

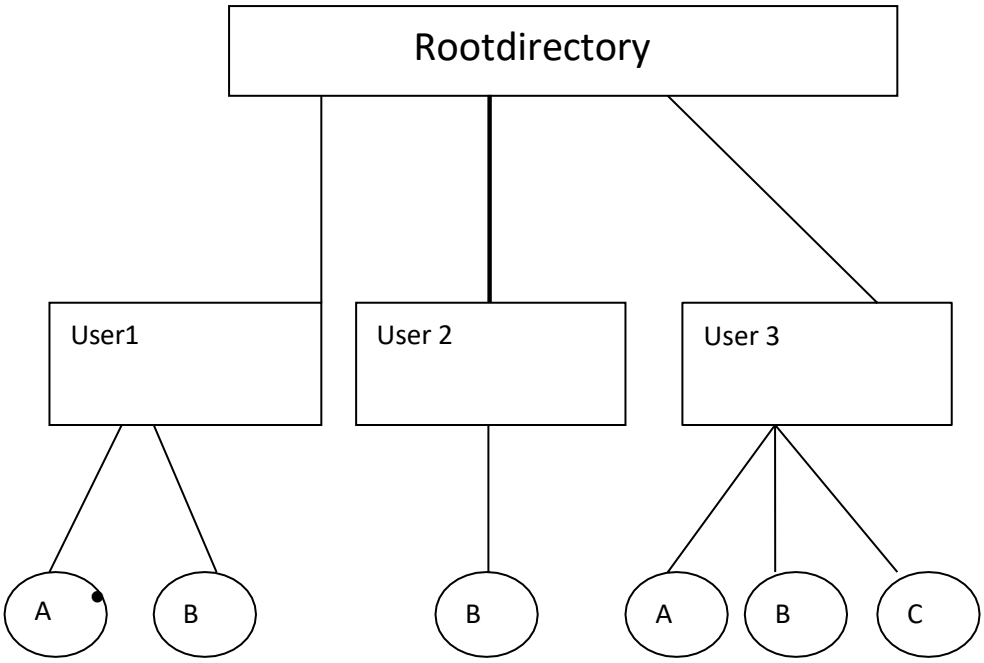
Disadvantage:-

- This structure have some significant limitations even for a single user, because if the no. Of files increases, then it is difficult to keep track of the file and also quite difficult to remember the names of all the files.
- Asthesefilesareinsamedirectory,thereforethesefileswillhaveunique names.

Two-leveldirectory:-

- The problemis singlelevel directory is differentusers may be accidentally usethe same name for their files.
- To avoid this problem each user need a private directory,so that name chosenbyoneuser don't interface with the name chosen by different user.
- Two-levelstructureisdivided intotwolevelsofdirectories:-1- master directory (root directory)
  2-sub-directory(userdirectory)

- Considerthefollowingfigureforbetterunderstanding:-

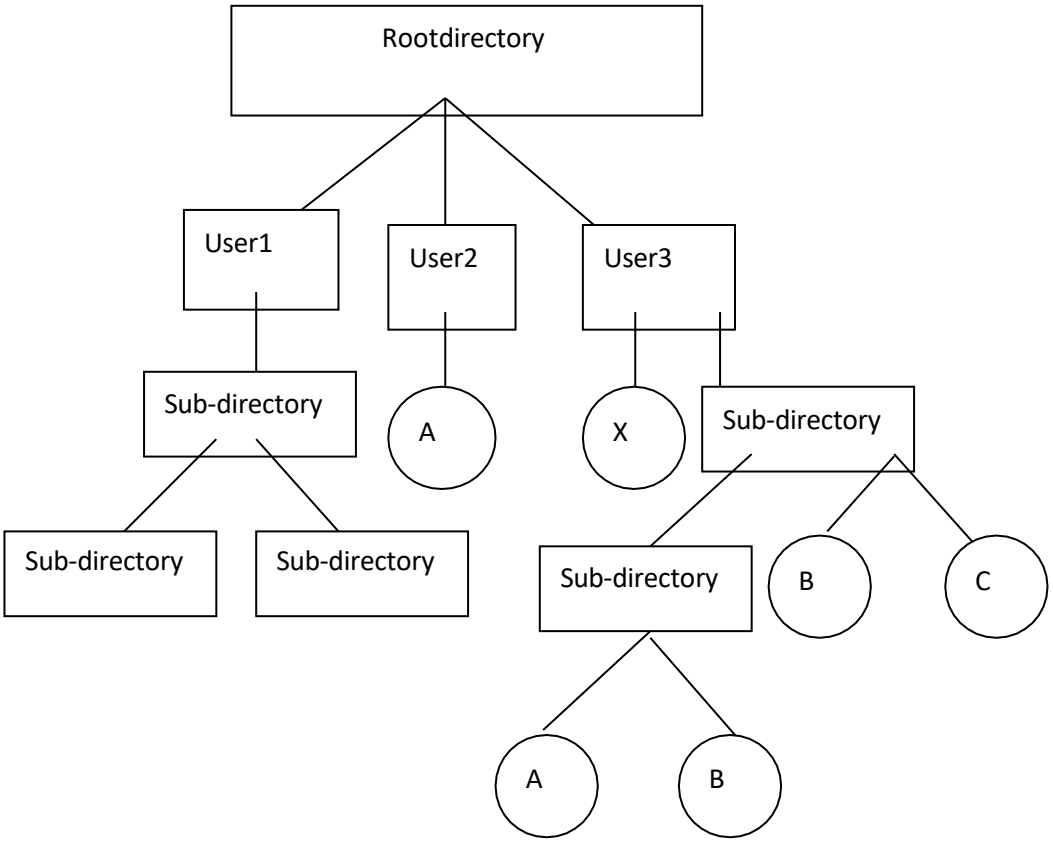


- Hererootdirectoryis firstleveldirectory.Itconsistsofentriesof"user directory".
- Userleveldirectoriesareuser1,user2,user3anditcontainsA, B, Cfiles.

Treestructureddirectory/hierarchicaldirectorysystem:-

- This structure allows user to create their own sub-directories and then organizethe files into it.
- MS-DOSoperatingsystemusesthistreestructured directories.
- Onedirectorymaycontainanotherdirectoryaswellasfiles.

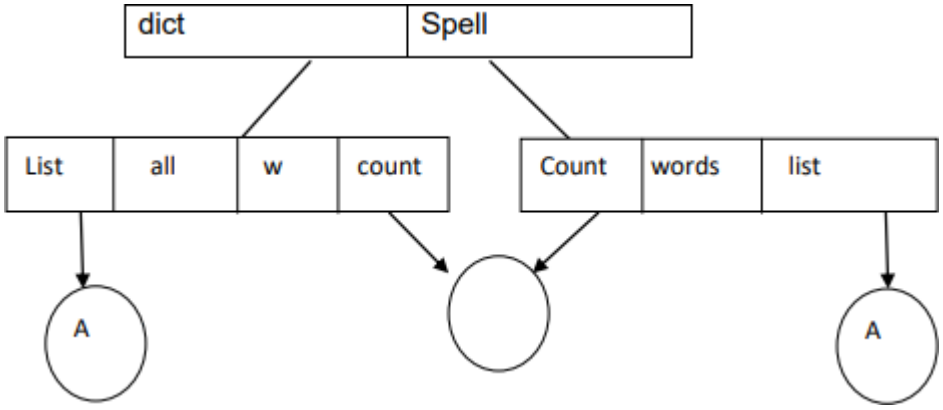
- Considerthefollowingfigure:-



[atreestructureddirectory]

Acyclicgraphdirectory:-

- Thisgraphallowsdirectoriesto havesharedsub-directoriesandfiles.
- Same filemaybe intwodifferentdirectories.
- Acyclic graph is a generalization of the tree structured directoryscheme but herecycle informed.

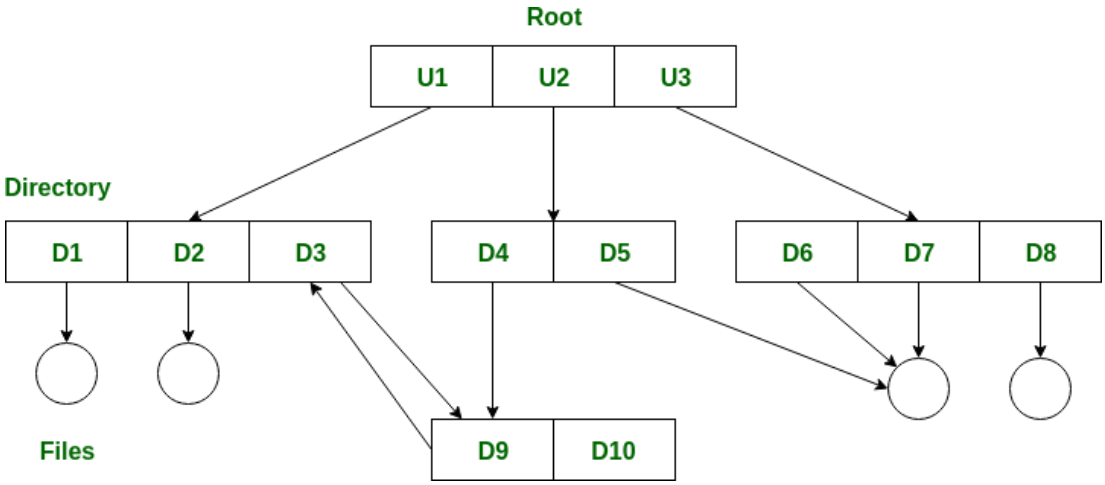


Fileimplementation:-

- Sharedfiles/sub-directoriescanbe implementedintwoways
  - Symboliclink:-
    - Apointertoanotherfileor directory.Ex-ln-s/spell/count/dict/count.
  - Hardlink:-
    - Duplicatealllinksinformationabouttheminbothsharingdirectory.

Generalgraphdirectorystructure:-

- Whenweaddlinkstoanexistingtreestructureddirectory thetreestructureis destroyed, resulting a simple graph structure.
- The primaryadvantage ofthis structure is traversing is easyand file sharing also possible.



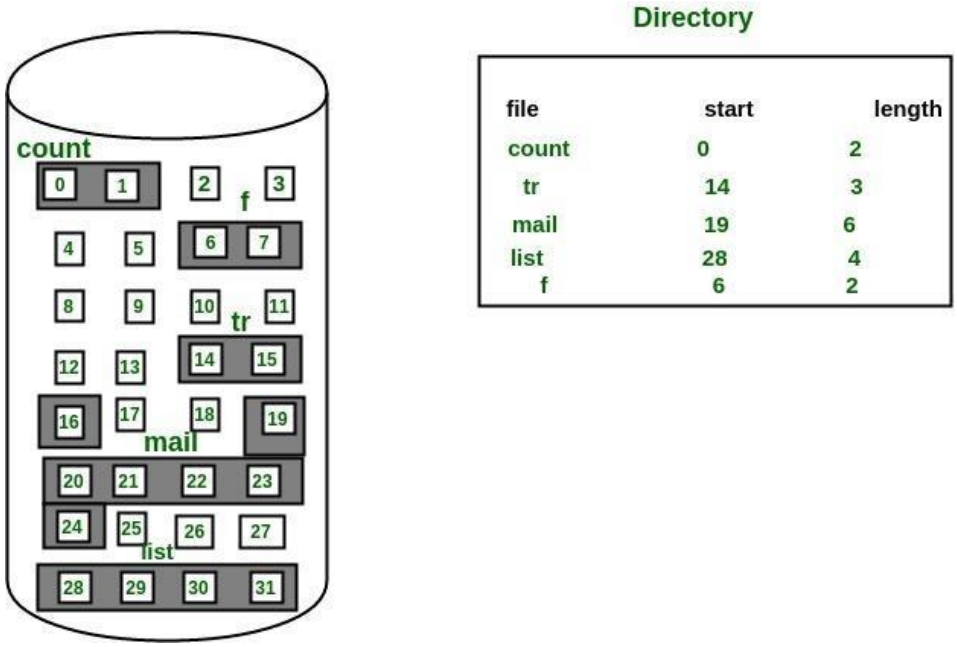
Fileallocationmethod:-

Filesarenormally storedin thediskssothemainproblemishow toallocatetothesefilesso that disk space is utilized effectively and files can be accessed quickly.

- Three major methodsofallocatingdiskspaceare inwideuse
- Theyare:-
  1-contiguousallocation2-
  linked allocation
  3-groupedallocationorindexed allocation

Contiguousallocation:-

- Inthismethodeach filesoccupiesasetofcontiguousblocksonthedisks.
- Ex:-adisksconsistingof1Kbblocks.A100kbfilewouldbeallocatedto100 consecutive blocks. With 2kbblocks, it would be allocated 50 consecutive blocks.



The*file'mail'*intheabovefigurestartsfromtheblock 19with length=6blocks.Therefore,itoccupies *19, 20,21, 22,23,24*blocks.

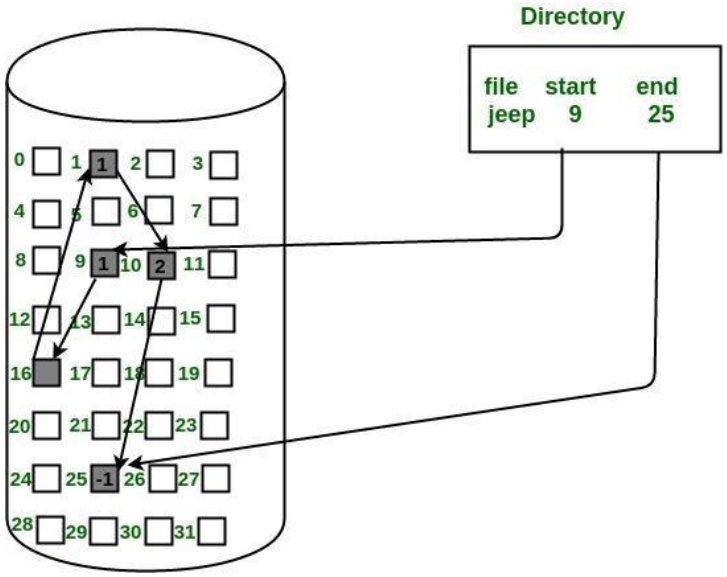
- Inthisfiguretherighthandsidepartisthefileallocatingtable.
- It is consisting of a single entryfor each file. It shows the file name starting blockofthe file and size of the file.
- Thismethodisbestsuitedforsequentialfiles.

Disadvantage:

- Itisdifficulttofindthecontiguousfreeblocksinthedisks.
- External fragmentation occurs (i.e:- some of the free blocks may leftbetweentwo files).

Linked allocation:-

- Inthismethod, everyfileisalinkedlistofdiskblocks.
- Itiseasytolocatethefiles,becauseallocationisonan individualblockbasis.
- Thesediskblocksarepresentallover the disks.
- Everyblock containsapointerforthenextfreebocks.



- Thesepointersareavailableto users.
- Ex:- there is a file “sort”, which is consist of 7 blocks. It starts from block 8and continues to block 15 and from block 15 to block 22 and so onfinally itis ended with the blocks.

Advantages:

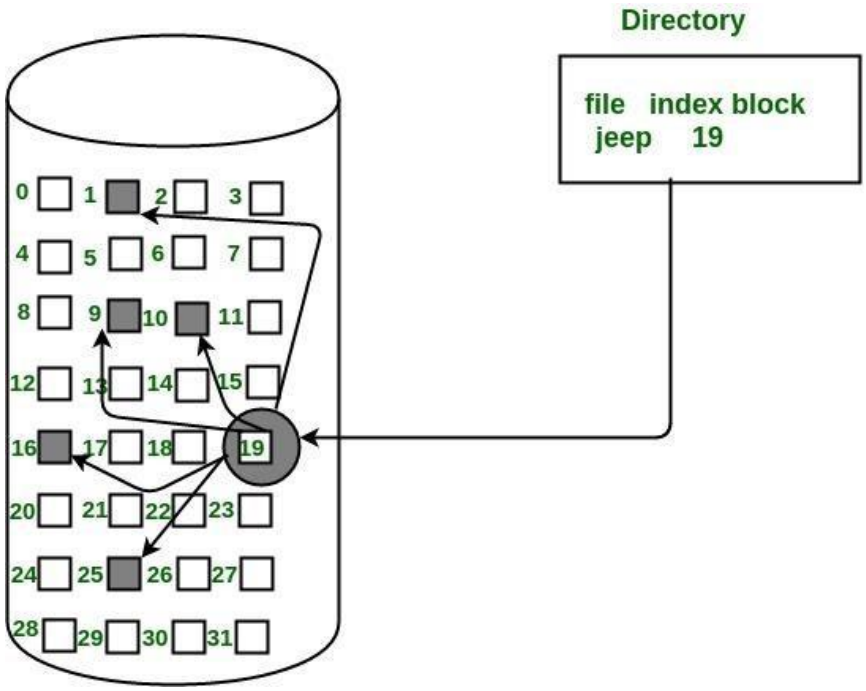
- Avoidexternalfragmentation.
- Suitedforsequentialfiles.

Disadvantages:

- Thepointeritselfoccupiessomememorywithitintheblock.Solessspace available for storing information.
- Takesmuchaccessing time.

Groupedallocation orindexedallocation:-

- Thismethodsolvesalltheproblemofthelinkedallocationmethod.
- Itsolvestheproblembybringingallthepointersataparticularplace,whichisknown as index value.
- Anindividualblockhavingthepointerstotheotherblocks.
- An individualindexblock isprovided to everyfile and it containsallthedisk block addresses.
- Whencreating afile,allthepointersare set toit.
- Thefig. Showstheindexedallocationofdiskspace.



Advantages:

- Indexedallocationsupportsbothsequentialanddirectaccessing offiles.
- Thefileindexed arenot physicallystoredaspartoffileallocationtable.
- Whenthefilesizeincreases,wecaneasilyaddsomemoreblockstotheindex.
- Noexternalfragmentation.

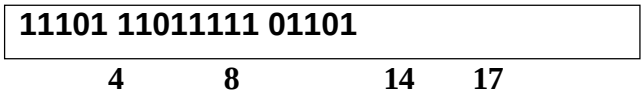
Freespacemanagement (or)diskspacemanagement:-

Generallythe filesare storedondisk so management ofthe disk space isa major problemtothe designer,ifuserwantstoallocatethespaceforthefileswehavetoknow what stocksonthe disk available.

- Thusweneedadiskallocationtableinadditiontothe**file**allocationtable.
- Tokeeptrackoffreediskspace,thefilesystemmaintainsa freespace list.The freespace list records all the disks bocks which are free i.e not allocated some other files.
- Tocreateafile,wesearchthefreespacelist.Whenafileisdeleteditsdiskspaceisadded to the free space list.
- These are the number oftechniques used for disk space management:- 1:-bit sector or bit value.

  2:-chain free points or linked free space list.3:-
  index block list.

  Bitsectororbittable
  - A bitvectorisacollectionofbits,inwhicheachblockisrepresentedbyone bit.
  - Iftheblockisfree, thebitis0.
  - Ifthebockisallocatedthebitis1.
  - Ex:-consideradisk whereblocks4,8,14,17arefree



Chain freepointsorlinkedfreespacelist:-

- Another approach is to link allthe free space blocks together keeping a pointertothe first free block.
- Thisblockcontainsapointertothenextfreediskblockandsoon.
- In an example we keep a pointer to the block 4, as the first free block 4 would contain a pointer to block 8, which would point to block 14, which would block topoint to 17 and so on.

Indexedblocklist:-

- Thechain freepointsarenotveryefficienttotraverse thelist.
- In index block list technique it would store the address of n free blocks in thefirst free block.
- Then-1ofthese areactuallyfree.
- Thelastoneisthediskaddressofanotherblockcontainingtheaddressofanother ‘n’ free blocks.

Advantages:

Theaddressoflargenumber offreeofblockscanbefound quickly.

Fileprotection/sharingoffiles:-

Any information present in the computer system must be protected from physicaldamageand improper access.

- Files can be damaged due to h/w problems such as temperature and validationand maybe deleted accidently.
- So,thereisneedtoprotectthesefiles.Therearemanymethodsforproviding protection to various files.
- Fileprotectionisdepending onthe system
  -inasingleusersystemwecanprovideprotectionbysimplyremovingfloppydisks and storing them at a safe place.
  -butinmulti-usersystem,therearevariousmechanismusedtoprovideprotection. They are:-
  1-typeofaccess2-
  accesscontrol
  3-otherprotectionapproaches(suchas password).
  Typeofaccess:-
- Wecaneasilyprovideprotectionbyprohibitingaccess**.**
- Controlled access is provided byprotection mechanism. These mechanisms canaccept or reject an access depending on the type of access.
- Thevariousoperationsthat canbecontrolled are:-
  - Read:-helpsto readfromthefile
  - Write:-helpsto writeorrewrite thefile
  - Execute:-helpstoexecutea stored file
  - Append:-helpstowritenewinformationattheendoffile.
  - Delete:- helpstodeletethefile
  - List:-helpstolistthe nameandattribute ofthefile

- Variousoperationssuchasrenaming,copying,editing ofafilecanbecontrolled.
- Different protectionmechanismsareintroducedforvarioussystemsandeverymechanismhasits own advantages and disadvantages.

Accesscontrol:-

- Inthis approachofprotection,accessdependsontheidentityoftheuser
- Everyuserusesadifferenttypetoaccessafileordirectory
- Themostcommonmethodtomakealistwithidentityofeachuserandtheiraccess control.
- Whenauserrequestsanaccessforfile,thenfirstitcheckstheaccesslistrelatedto that file.
- Ifthatparticular userislisted,thentheoperatingsystemallowstoaccessthatuser.
- Ifnot,thenitleadstoprotectionviolationandoperatingsystemdeniestherequest.
- Advantage:-itcanhandlecomplexmethodologies.
- Disadvantage:-the list becomes very large, when the no. Ofuser increases. So itisvery difficult to maintain and construct a list.
- Inorderto solve thisproblem, access controlis introduced.The systemclassifiestheuser into three different categories related to each file.
  - Owner:-itistheuserwhocreatesthefile.
  - Group:-itisthesetofuserwhosharesthefileandrequiresthesametimetoaccess.
  - Universe:-referstoallotheruserofsystemthatconstitutesa universe.

Otherprotectionapproaches (password):-

- Anotherprotectionapproachisto use apassword foreveryfile.
- Soaccessingafile iscontrolledbypassword.

Disadvantages:

- Userhastorememberalargeno. Ofpassword.
- If a single password is used for all the files then if it gets discovered,then it makesallthefiles accessible.

Unit-

7Systemprogrammin

g

Systemprogramming:-systemprogrammingistheactivityofprogrammingsystemsoftware.
The primary distinguishing characteristics of the systemprogramming when compared to that application programming aims to produce software which provides services to the user directly whereas systems programming aims to produce software and software platforms which provide services to other software.

Systemprogrammingrequiresagreaterdegreeofhardwareawareness.

Applicationprogram:-

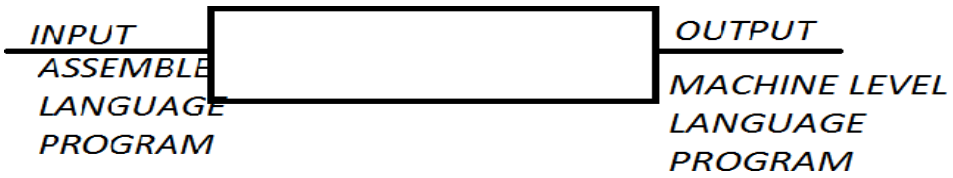
1. Applicationsoftwareisasetofoneormorethanoneprogramswhicharedesignedtocarry out operation for a specified application.
2. For example payroll packages are designed to produce pay slip as the majorproduct. An application package for processing examination result produced mark sheet as the major product.
3. Thepersonwhopreparestheapplicationprogramisknownasapplicationprogrammer.
4. Nowadaysapplicationpackagesareusedforapplicationsuchasbanking, administration, insurance, publishing, manufacturing science and engineering.

Systemsoftware:-

1. Systemsoftwarealsoknownassystempackages.Thesearethesetofoneormorethanone programwhicharedesigned not to performspecific application. Butthesearedesignedto operate computer system properly.
2. The system programs helps or assist human for performing several application such as input and output data to the system.
3. Italsoexecutestheapplicationprogram.
4. It manages and monitors the activities ofall hardware suchas memory, printer, keyboard etc.
5. These are verycomplex to design. So rarely it is designed in houses. These are designed by system programmers.

Assemblers:

1. A computer program which translates an assembly language program toitsmachine language equivalent is known as assembler.
2. Theassemblerisasystemprogramwhichissuppliedbythe manufacture.
3. A symbolic program written by a programmer is assembly language is called a source program. After the source program has been converted into machine language by an assembler it is referred to as an object program.

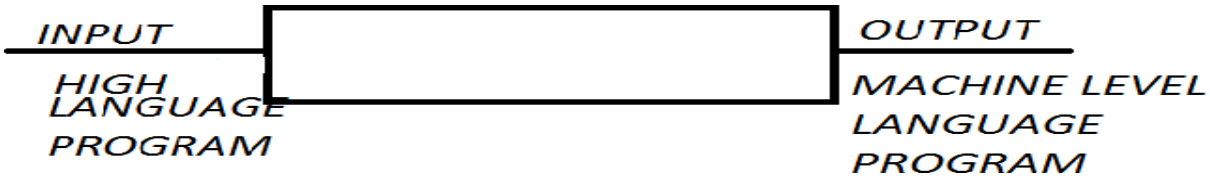


4. Theinputoftheassembler istheassemblylanguageprogramandoutputfromthe assembler is the machine language program.

Complier:

1.Acomplier isaprogramthattranslatesthehighlevellanguage intomachine level language by reading the entire source code.

2-Aprogramwritten by a programmer in high level language is called source program thathasbeenconvertedintomachinelanguagebyacomplier isreferredtoasobject program.



3. So input to a complier is known as source program and output from acomplier is known as object program.

4. Asinglecompliercannottranslateallthehighlevellanguageintomachinelevel language should have a dedicated complier for itscompilation.

5. Complier is a large program whichresides insecondary. When it isrequired it is copied into main memory.

Interpreter

Aninterpreterisalso atranslatorthattranslatesthehighlevellanguage intomachinelevel language by reading one by one.

Heretranslationandexecutionalternateforeachstatement encounteredinhighlevellanguage program.

Aninterpretertranslatesthe instructionandthecontrolunit executestheresulting machine code so on.

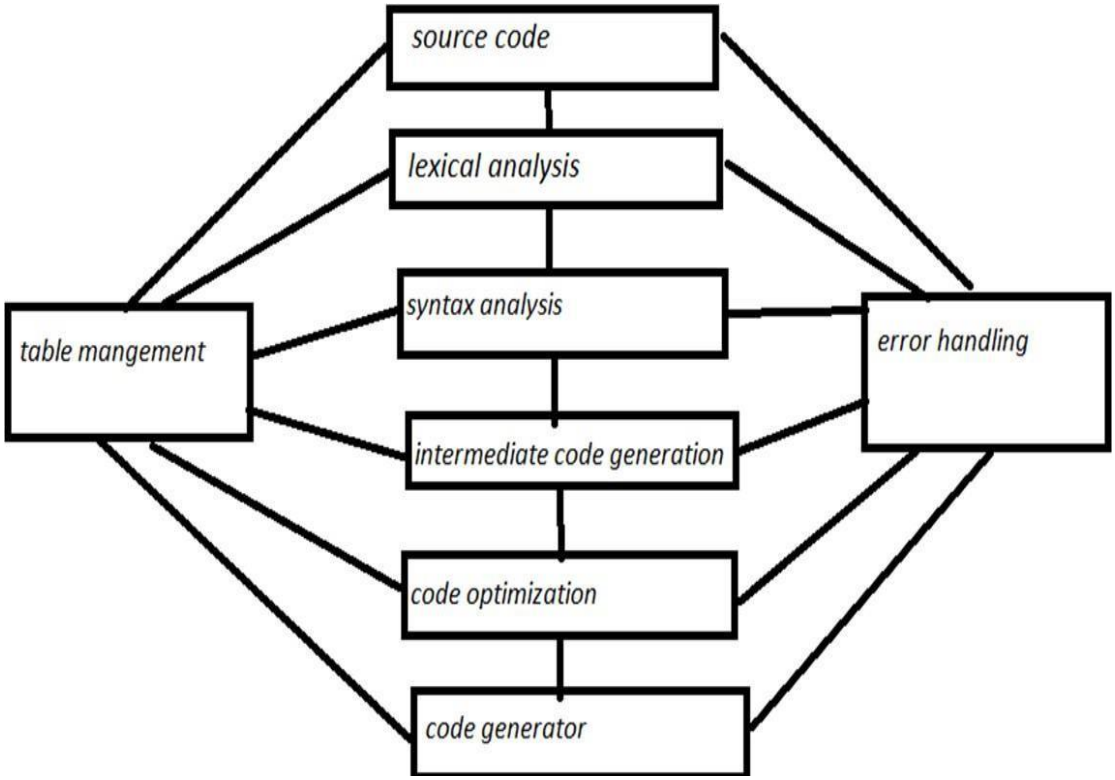
Itissimpletowriteandrequired lessspace inmainmemoryforstorage.Asoneby one line is

translated so it is slower.

Complier	Interpreter
1.Complier is software that translates the high level language into machine language by readingtheentirecode at a time.	1.Interpreterisasoftwarethattranslatesthehigh level language into machine language by reading one statement at a time.
2. Repeated compilation is notnecessaryforrepeatedexecutionofa program.	2. Repeated interpretation isnecessary.
3.slowresponsetothechangeinsource code.	3.Fastresponsetothechangeinsource code.
4. it is a complex program as compared to interpreter.	4.itisasimple program.
5. itrequireslargememoryspacein computer.	5. easytowriteanddonotrequire large memory space in computer.
6.it isfaster.	6.it isslower.

Stagesofcomplier

Thecomplier takesaninput asourceprogramand producesasoutput anequivalentsequenceof machine instruction. The complier does this transition by some sequence of gates or phase.



Lexicalanalyzer/scanner:-

- Lexicalanalysisisthefirstphaseofcompiler whichisalsotermedasscanning.
- Sourceprogramisscannedtoreadthestreamofcharactersandthosecharactersaregroupedto form a sequence called lexemes which produces token as output.

Token: Token is a sequence of characters that represent lexical unit, which matches with the pattern, such as keywords, operators, identifiers etc. This separatorscharacterofsources languageintogroups thatlogicallybelong together, these groups are called tokens. The usual tokens are keyword operator symbol.

Syntax analyzer/parser:- The output of the lexical analyzer is passed tothis syntax analyzer. Thesyntaxanalyzercheckwhetherthestatementisvalidornot everylanguagehasitsproduction.

Ifsentencefollowsthisrulesthisrulesthenthesentenceisvalid.

Tocheckthe validationofasentencetwotechniquesareused:-

- Topdownapproach
- Bottomup approach

Intermediatecodegeneration:-Thisphaseusesthestructureproducesbythesyntaxanalyzerto create a stream of simple instruction. These instructions are similar to assembly language.

Codeoptimization:-This isanoptionalphase, whosejobisto improvethe intermediatecode. So that the ultimate object program can run faster.

Codegeneration:-Thisphaseproducestheobjectcodebydecidingwherethe memoryspacewill be allocated to the variables, literals and constant.

Table management:- Thisportionofthecomplier keepstrackofthe namesused bytheprogram and records essential information. The data structure used to record this information is called a symbol table.

Error handler:-The error handler is involved when an error in the source programis detected. Generally the error occurs at the syntax analyzer phase.

Boththetablemanagementanderrorhandlerroutinesinteractwithallthephaseofthe complier.

References

1. “OperatingSystemConcepts”byAviSilberschatz,Peter BaerGalvinand Greg Gagne
2. “OperatingSystem”byEr.RajivChopra
3. “OperatingSystemandSystemProgramming”byP.Balkrishna Prasad
4. “OperatingSystems” byVijayShukla
5. “OperatingSystem”byStuartMadnick andJohn Donovan.
6. https://www.geeksforgeeks.org
7. https://nptel.ac.in
8. https://en.wikipedia.org