

Chapter 4

INTRODUCTION TO REAL-TIME OPERATING SYSTEM (RTOS).

- Desired Service within a given period of time.
- RTOS are time driven system and deadline driven.
↳ Bounded response.

General purpose OS (GPOS)

- * time insensitive
- * non deterministic
(non predictable)
- * tasks will be non prioritized
- * may not be preemptive
- * Virtual memory
- * General purpose applicatns
- * unbounded response.

Real time OS (RTOS).

- * time sensitive
- * deterministic
* (predictable)
- * they are prioritized
- * mostly preemptive
- * no virtual memory
- * Specific applicatn.
- * Bounded response

Types of RTOS:

- ↳ soft RTOS
- ↳ hard RTOS.

soft RTOS: Meeting deadline will not lead to "duriable" or catastrophic changes but quality reduction may occurs such type of RTOS are called as soft RTOS.

Ex: Telecommunications applications, media applications.

Hard RTOS: It has to definitely meet deadline

Ex: Nuclear power plant

Key characteristics of RTOS:

- Reliability: long time service it should give
- Predictability: deterministic
- Compactness: as small as possible so that it consumes less memory.
- Performance: Throughput and response time should be good.
- Scalable: RTOS should allow to add some other component if required like any network protocol if needed.

Functions of RTOS:

* Task management: managing all tasks

* Scheduling

* Resource management / Resource allocation.

* Interrupt handling

Applications:

Medical, Telecommunication, defence, security, air traffic, Automobile.

RTOS

Application

RTOS

Networking
protocols

File
System

Other
components

C/C++ support
libraries

Kernel

POSIX
Support

Device
drivers

Debugging
facilities

Device
drivers

BSP (Board Support package)

Target hardware

BSP is used to connect some peripherals to hardware.

→ When we connect peripheral we will be in need of
Some drivers which will be installed in BSP (Board

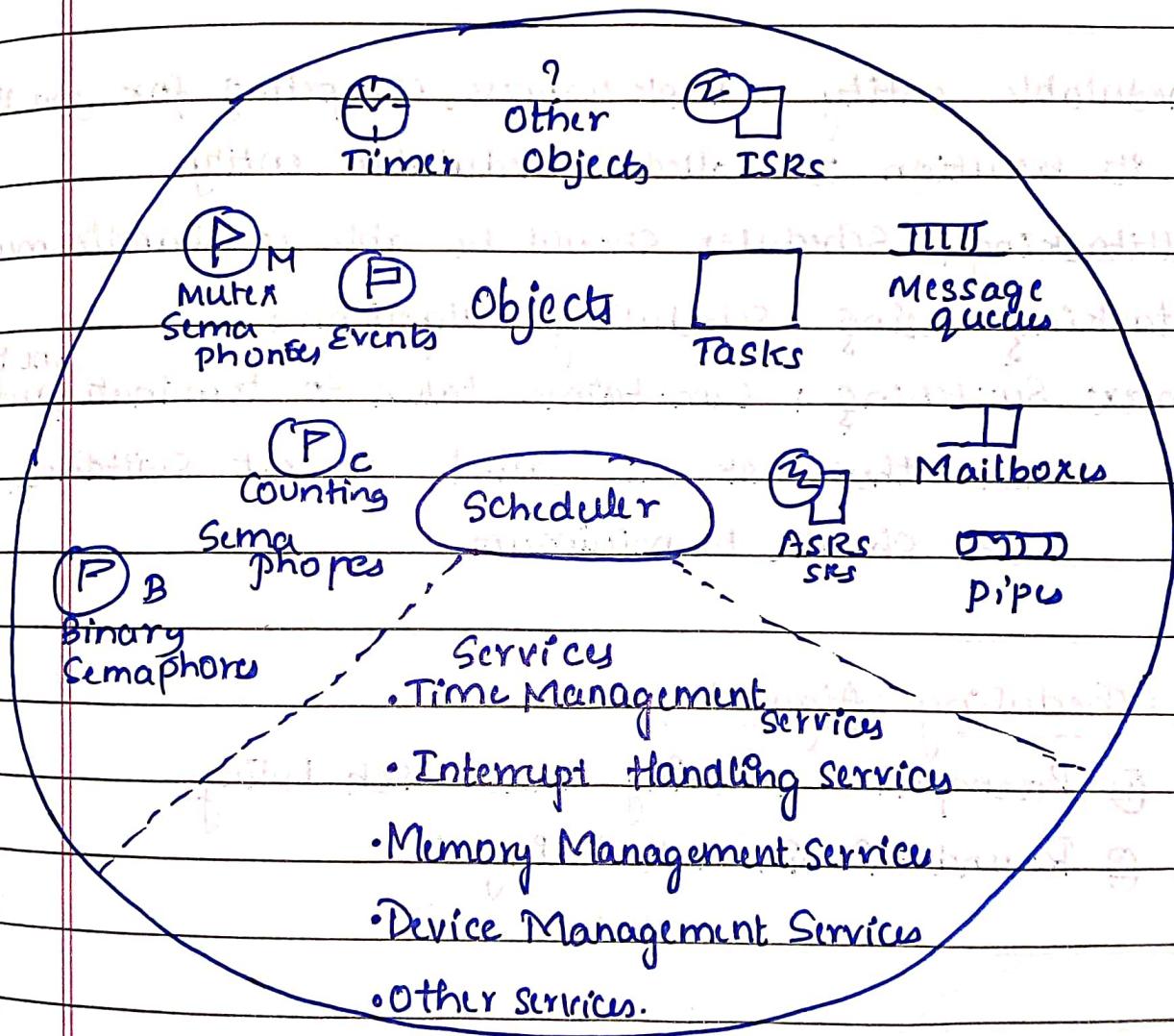
Support package software

Kernel Structure

Kernel is made up of mainly 3 components

- * Scheduler
- * Services
- * Objects

→ Scheduler make use of objects and gives services



Scheduler : heart of every kernel.

- * Schedulable entity
- * provides multitasking
- * context Switching
- * Dispatching
- * make use of Scheduling algorithms

- Schedulable entity : task / process competing for CPU time for its execution is called Schedulable entity.
- multitasking : Scheduler should be able to handle multitasking using scheduling algorithms.
- Context Switching : same time taken to terminate one task and load other task is called context switching.
It should be minimum.

Scheduling Algorithms

- ① Preemptive priority-based scheduling
- ② Round-Robin scheduling.

RTOS KERNEL OBJECT - TASK

- By use of Objects RTOS gives desired level of services
- Kernel Objects and
- ① Task / process
- ② Timer
- ③ ISR's → Interrupt service routines will help us to handle Interrupts.
- ④ Semaphores → for synchronization during multitasking.
- ⑤ Events / signals
- ⑥ Queues / pipes
- ⑦ Mail boxes

TASK

Task: A piece of program which is meant for realizing a specific operation is called as task.

Task is entity that competes for CPU time for its execution

Every task will have predefined algorithm / Scheduling algorithm.

Structure of task:

task name / ID

```
int my_task
{ while(1)
```

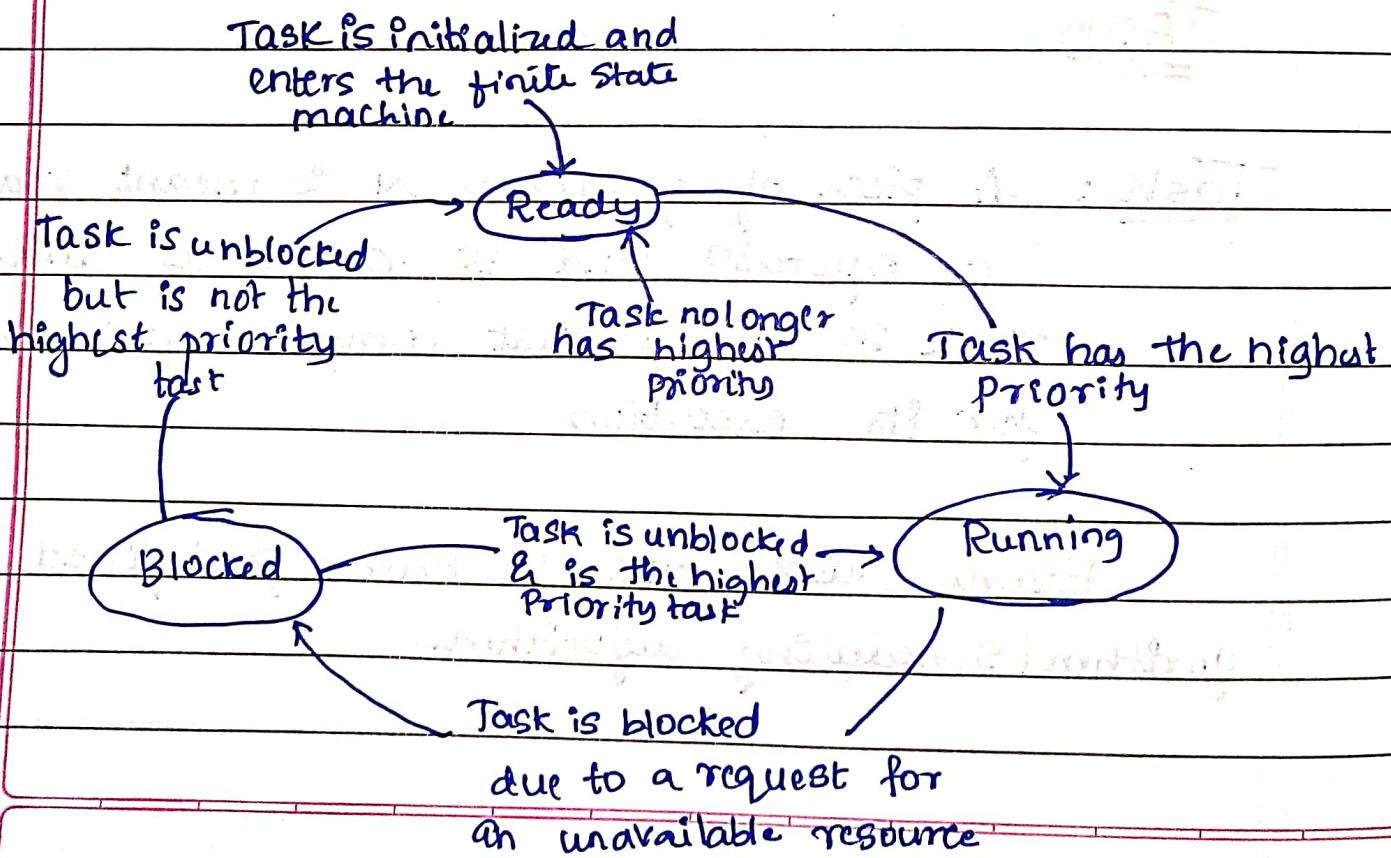
{
 }
 y
 y

→ We will have a TCB
(task control block)

which is going to
store the info about
task like Taskname
(ID) etc.
(parameters)

→ To store memory required to execute a task
will be having a stack memory.

TASK STATE DIAGRAM



TYPICAL TASK OPERATIONS

- * Creating & deleting of task

- * Scheduling

- * Task info

Operations

Description

Suspend

Suspend a task

Resume

Resumes a task

Delay

Delay a task

Restart

Restarts a task

Get priority

Get the current tasks priority

Set priority

Dynamically sets a tasks priority.

Preemption lock

Locks out higher pri tasks from

Preemption unlock

Unlocks preemption lock.

Operat's

Descript'

Get ID

Get current tasks ID

Get TCB

Get the current tasks TCB

RTOS KERNEL OBJECT - BASICS OF SEMAPHORE

→ Semaphore : for achieving synchronization, communication and Resource management.

→ act as key token for hardware resources whichever task will have this key they will be able to access CPU (resource).

→ 3 types of Semaphore

① Binary Semaphore

② Counting Semaphore

③ Mutex

I] **Binary Semaphore** : It comes in two values →

If it takes 0 : means that resource is unavailable

If it takes 1 : " " " " " "



→ To store the attributes related to semaphore
(Info like name, ID, tasks)

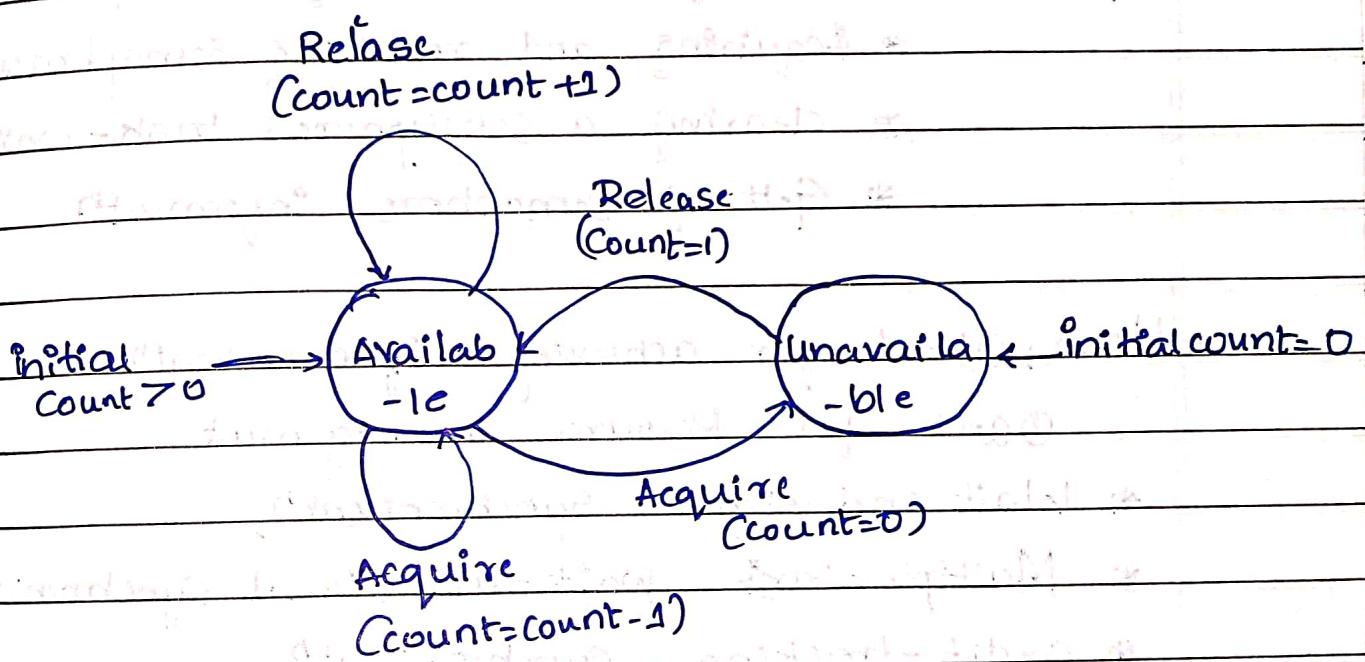
We have SCB (Semaphore control block).

It will have Id(name), task weight, list of tasks, value

problem of Binary S : no ownership means when someone access resource any other can release it by deleting task w may lead to memory leakage.

2) Counting Semaphore : having a count value (integer)
 How many times Particular Semaphore is being accessed that count value can be maintained (recorded) by this Semaphore.

→ Here also we will have 2 states (available and unavailable).



3) Mutex : To avoid problems in previous semaphores

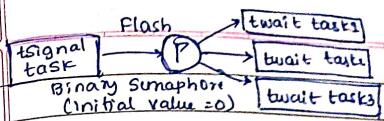
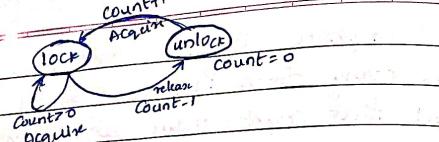
↳ 3 main features :

- * Ownership is provided

- * Task deletion safety

- * provides Recursive access

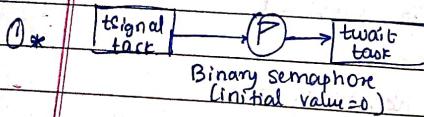
→ have 2 states : defined as lock and unlock.



SEMAPHORE OPERATION & USES.

- Operations:
 - * Creating and deleting Semaphores
 - * Acquiring and releasing Semaphores
 - * clearing a Semaphores task-waiting list
 - * Getting semaphore information

- USES:
 - ① Used for achieving Synchronization
 - ② and for Resource management.
 - * Wait-and-signal Synchronization,
 - * Multiple-task Wait-and-Signal Synchronization
 - * Credit-tracking synchronization
 - * Single Shared-resource-access synchronization
 - * Recursive Shared-resource-access synchronization
 - * Multiple Shared-resource-access "



OBJECT - PRIORITY INVERSION.

→ let us say we have 3 task

Low Priority

Med Pri

high priority



@ one instance of time when let us assume T1 is
in running mode

T1 utilizes resource called R1.

Now when T2 comes tries to preempt T1 as
it medium priority and when T3 comes it
also tries to preempt.

When T2 preempt T1 then R1 enters to
blocked mode now when T3 comes it want to
execute Though priority is high R1 has to be
released by T1 so T3 has to wait due to
poor design. This leads to convert high priority T3 to low priority
inversion.

→ Solving priority inversion problem.

We use mutex to solve this problem

Mutex solves problem by incorporating
priority inheritance and ceiling priority protocol

1. Priority inheritance protocol: means someone's priority being inherited

We make T_1 as highest priority so that T_2 cannot preempt T_1 and T_3 can do that its executing

2. Ceiling priority protocol:

Highest possible priority among all task will be assigned to running low priority task.

RTOS Kernel Object : Basics of Message Queue

Message Queues :

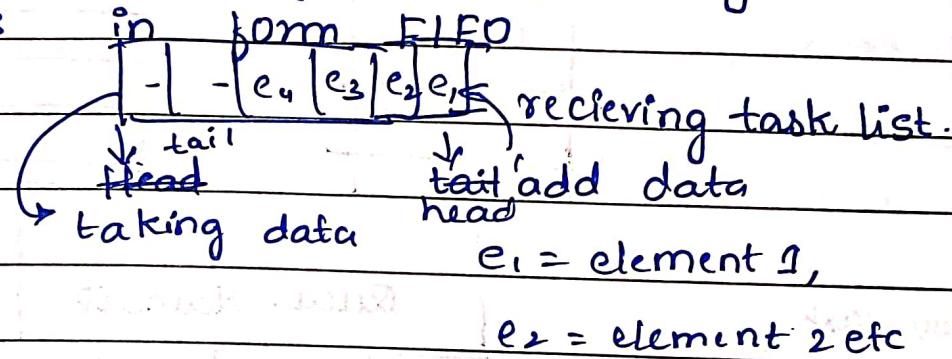
- Inter task processes (communication) (IPC) :
- Message Queues are achieving IPC
- Along with message queue we can also use pipes, mail boxes for IPC
- Message queue is going to carry message.
- This queue is in form FIFO
- 

Diagram illustrating a FIFO queue structure:

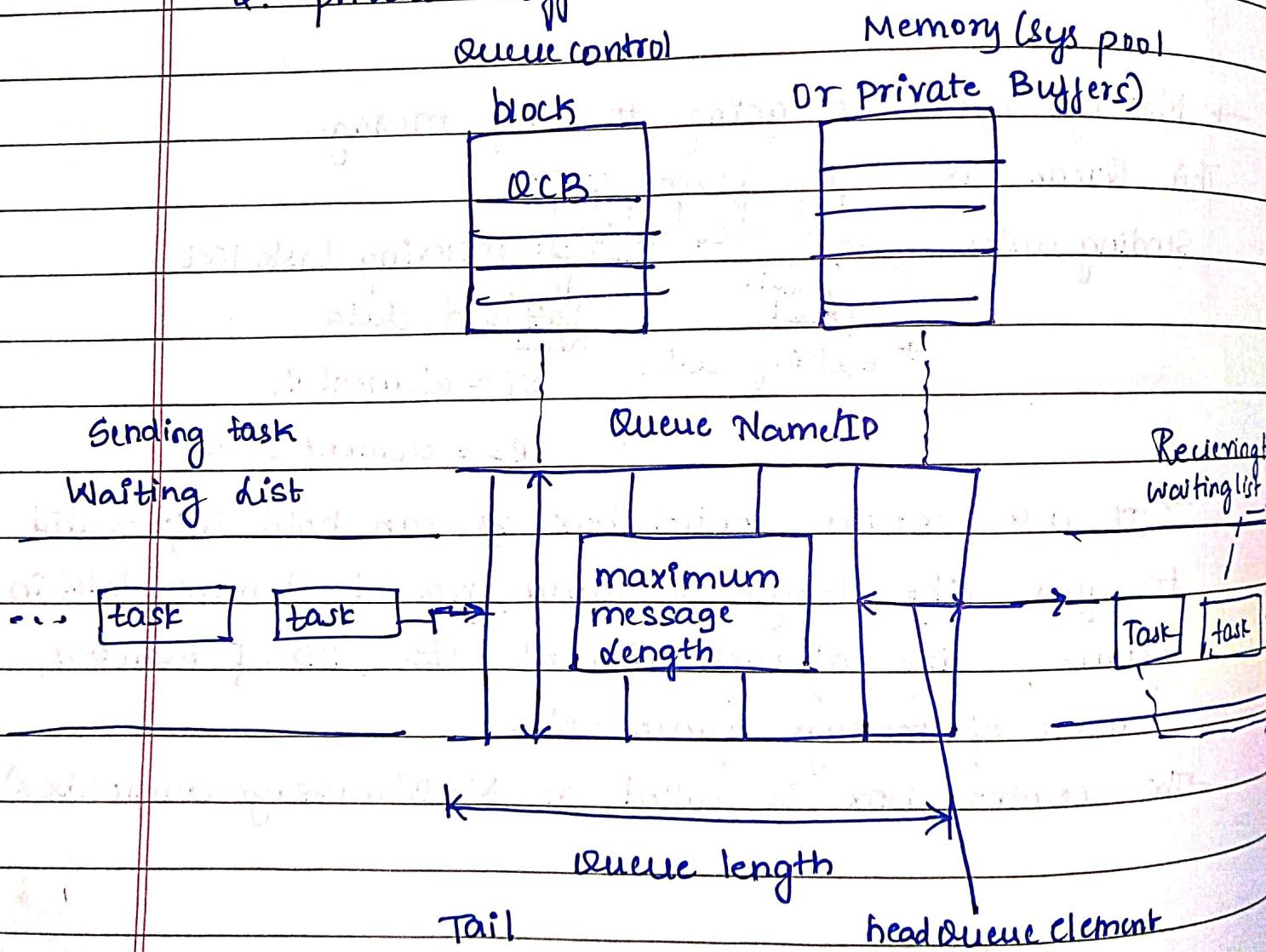
 - The queue is represented as a list of elements: $e_1 | e_2 | e_3 | e_4 | e_5$.
 - The **Head** pointer points to e_1 .
 - The **tail** pointer points to e_5 .
 - An arrow labeled "add data" points to the position after e_5 , indicating where new data is added.
 - A curved arrow labeled "taking data" points to e_1 , indicating where data is removed.
 - Annotations below the list define e_1 as "element 1" and e_2 as "element 2 etc".
- It also creates control box which can hold info related to queue like length of queue, no. of elements(data) in queues, size of each element etc., ID of mailbox, name of message queue etc.
- This control box is called as MCB(message control box).

→ A message queue is a buffer-like object through which tasks and TSRs send and receive messages. It communicates and synchronizes with data. A message queue is like a pipeline.

→ 2 means of allocating memory to queue:

1. System pool

2. private buffer.



* System pools : Using a system pool can be advantageous if it is certain that all message queues will never be filled to capacity @ the same time. The advantage occurs because system pools typically save on memory use.

The downside is that a message queue with large messages can easily use most of the pooled memory, not leaving enough memory for other message queues.

*) Private Buffers: Using private buffers, on the other hand, requires enough reserved memory area for the full capacity of every message queue that will be created.

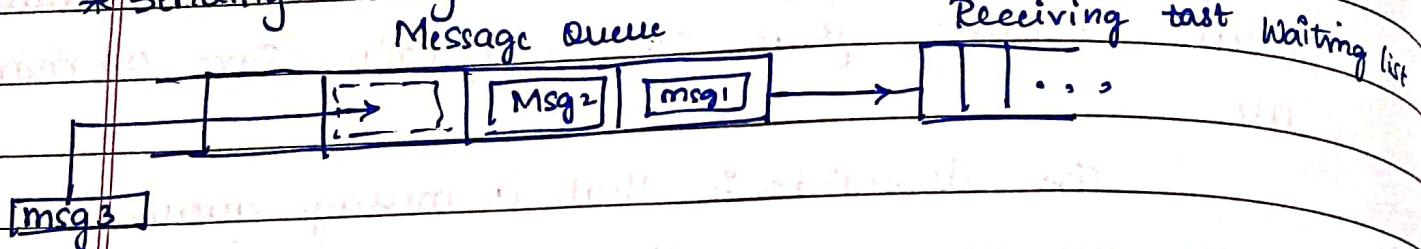
This approach clearly uses up more memory;

OPERATIONS ON MESSAGE QUEUES

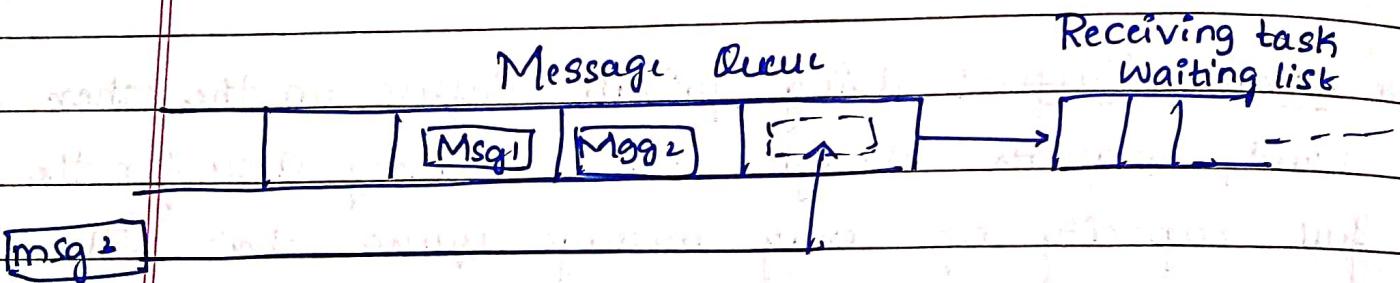
- ① Create queues:
- ② Deleting queue
- ③ Sending & Receiving messages:
- ④ At Get info abt message queue.

Sending & Receiving messages:

* Sending message - First in - First-out (FIFO) order



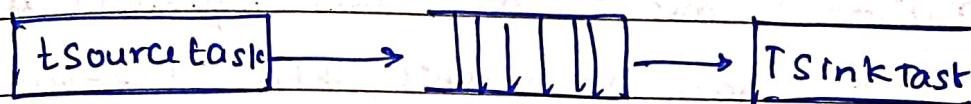
* Sending message - last-in, first-out (LIFO) order



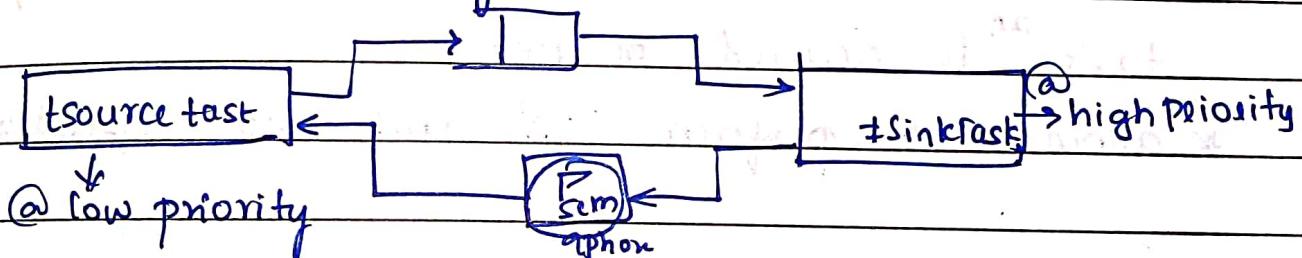
* for Task Waiting list also we have FIFO & priority based order.

Uses of Message Queue

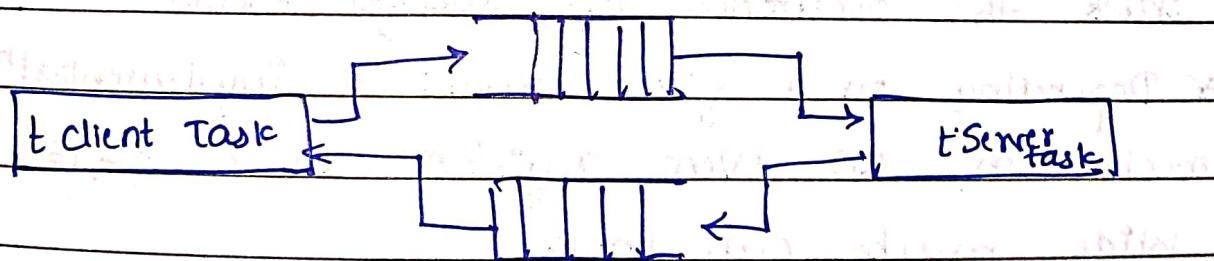
- > Non-interlocked, one-way data communication.
- > Interlocked, one-way data communicatn.
- > Interlocked, two-way " "
- > Broadcast communicatn.
- > Non-interlocked, one way



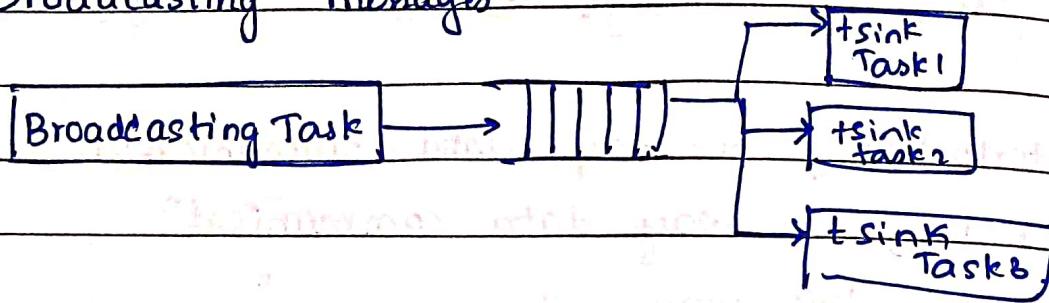
- > Interlocked, one way



- > Interlocked, 2-way communicatn.



Broadcasting messages

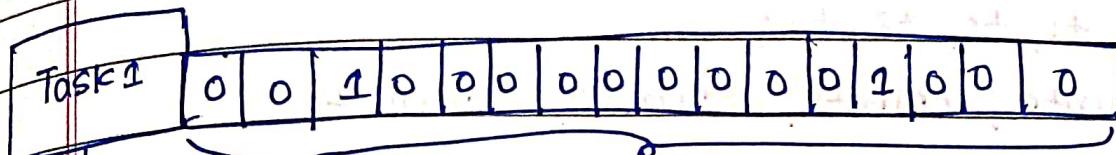
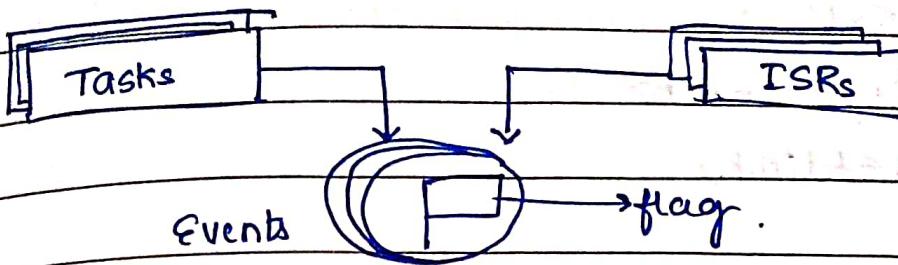


KERNEL OBJECTS - EVENT REGISTERS.

- * Event register is meant for task synchronization.
- * " " may be of 8bit, 16bit or 32bit.
- * Each bit will be acting as flag.
- * Each bit can be set or reset (0 or 1) to say a task is occurred or not.
- * allows us to perform 'OR' and 'AND' operations.

- * Event register, is an object belonging to a task and consists of a group of binary event flags used to track the occurrence of specific event.
- * Depending on a given kernel's implementation of this mechanism, an event register can be 8-, 16-, or 32-bit wide, maybe even more.
- * Each bit in this register is treated like a binary flag (also called an event flag) & can be either set or cleared.

Setting Events



Checking Events

- No wait
- Wait v.
- Wait with time out