Chapter 8: Deadlocks
Chapter 8: Deadlocks

- System Model
- Deadlock in Multi-threaded Applications
- Deadlock Characterization
- Methods for Handling Deadlocks
- Deadlock Prevention
- Deadlock Avoidance
- Deadlock Detection
- Recovery from Deadlock
Chapter Objectives

- Illustrate how deadlock can occur when mutex locks are used
- Define the four necessary conditions that characterize deadlock
- Identify a deadlock situation in a resource allocation graph
- Evaluate the four different approaches for preventing deadlocks
- Apply the banker’s algorithm for deadlock avoidance
- Apply the deadlock detection algorithm
- Evaluate approaches for recovering from deadlock
System Model

- System consists of a collection of $m$ resource types and $n$ processes
  - Resource types $R_1, R_2, \ldots, R_m$
    - CPU cycles, memory space, I/O devices
- Each resource type $R_i$ has $W_i$ instances.
- Each process utilizes a resource by executing a sequence of three actions:
  - request it
  - use it
  - release it
Deadlock in Multithreaded Application

- Two mutex locks are created and initialized:

```c
pthread_mutex_t first_mutex;
pthread_mutex_t second_mutex;

pthread_mutex_init(&first_mutex, NULL);
pthread_mutex_init(&second_mutex, NULL);
```
Deadlock in Multithreaded Application

/* thread_one runs in this function */
void *do_work_one(void *param)
{
    pthread_mutex_lock(&first_mutex);
    pthread_mutex_lock(&second_mutex);
    /**
     * Do some work
     */
    pthread_mutex_unlock(&second_mutex);
    pthread_mutex_unlock(&first_mutex);

    pthread_exit(0);
}

/* thread_two runs in this function */
void *do_work_two(void *param)
{
    pthread_mutex_lock(&second_mutex);
    pthread_mutex_lock(&first_mutex);
    /**
     * Do some work
     */
    pthread_mutex_unlock(&first_mutex);
    pthread_mutex_unlock(&second_mutex);

    pthread_exit(0);
}
Deadlock in Multithreaded Application

- Deadlock is possible if thread 1 acquires `first_mutex` and thread 2 acquires `second_mutex`. Thread 1 then waits for `second_mutex` and thread 2 waits for `first_mutex`.
- Can be illustrated with a resource allocation graph:
Deadlock Characterization

Deadlock can arise if four conditions hold simultaneously.

- **Mutual exclusion**: only one process at a time can use a resource
- **Hold and wait**: a process holding at least one resource is waiting to acquire additional resources held by other processes
- **No preemption**: a resource can be released only voluntarily by the process holding it, after that process has completed its task
- **Circular wait**: there exists a set \( \{P_0, P_1, \ldots, P_n\} \) of waiting processes such that \( P_0 \) is waiting for a resource that is held by \( P_1 \), \( P_1 \) is waiting for a resource that is held by \( P_2 \), \ldots, \( P_{n-1} \) is waiting for a resource that is held by \( P_n \), and \( P_n \) is waiting for a resource that is held by \( P_0 \).
A resource allocation graph is a set of vertices $V$ and a set of edges $E$ such that:

- $V$ is partitioned into two types:
  - $P = \{P_1, P_2, \ldots, P_n\}$, the set consisting of all the processes in the system
  - $R = \{R_1, R_2, \ldots, R_m\}$, the set consisting of all resource types in the system

- **request edge** – directed edge $P_i \rightarrow R_j$ indicates $P_i$ has requested a unit of $R_j$

- **assignment edge** – directed edge $R_j \rightarrow P_i$
Resource Allocation Graph Example

- One instance of R1
- Two instances of R2
- One instance of R3
- Three instance of R4
- T1 holds one instance of R2 and is waiting for an instance of R1
- T2 holds one instance of R1, one instance of R2, and is waiting for an instance of R3
- T3 is holds one instance of R3
Resource Allocation Graph With A Deadlock

T₃ requests one unit of R₂ resulting in deadlock
Although this graph has a cycle, it is not the graph of a deadlock state.

\( T_2 \) holds a unit of \( R_1 \) and is not waiting for another resource, so it can release it when it is finished with it and \( T_1 \) can acquire the unit.
Resource Allocation Graph Facts

- If graph contains no cycles ⇒ no deadlock
- If graph contains a cycle ⇒
  - If each resource type has just a single instance of the resource, then a cycle implies deadlock
  - If there is at least one resource type that has multiple instances, a cycle does not necessarily imply deadlock.
Methods for Handling Deadlocks

- Ensure that the system will *never* enter a deadlock state:
  - Deadlock prevention
  - Deadlock avoidance
- Allow the system to enter a deadlock state and then recover
- Ignore the problem and pretend that deadlocks never occur in the system.
Deadlock Prevention

Invalidate one of the four necessary conditions for deadlock:

- **Mutual Exclusion** – not required for shareable resources (e.g., read-only files); must hold for non-shareable resources. Cannot remove mutual exclusion for reusable resources so this is not an option.

- **Hold and Wait** – must guarantee that whenever a process requests a resource, it does not hold any other resources. Two choices:
  1. Require process to request and be allocated **all of its resources before it begins execution**.
  2. Allow process to request resources **only when the process has none allocated to it**.

  - Poor resource utilization; starvation possible
Deadlock Prevention (Cont.)

- **No Preemption** – forcibly taking resources away from processes.
  - If a process that is holding some resources requests another resource that cannot be allocated immediately to it, then all resources currently being held by that process are released.
  - Preempted resources are added to the list of resources for which the process is waiting.
  - Process will be restarted only when it can regain its old resources, as well as the new ones that it is requesting.
  - On request, if unavailable, we preempt resources from a process that holds them if that process is in a waiting state; if the requested resources are not available or held only by ready processes, the requesting process waits. Process may lose resources while it waits. Process gets restarted only when it is allocated new resources and recovers the preempted ones.
Deadlock Prevention (Cont.)

- **Circular Wait** – impose a total ordering of all resource types, and require that each process requests resources in an increasing order of enumeration.
  - Proposed by Havender:
    - A process can request resources only in increasing order of its resource number.
    - If process holds $R_i$ it can only request $R_j$ if $j > i$.
    - If process wants $R_j$, it must first release all $R_i$ such that $i \geq j$.
    - If a process violates these rules, it is terminated.
  - This makes circular waiting impossible, so deadlock impossible.
  - But results in poor resource utilization.
Circular Wait

- Invalidating the circular wait condition is most common.
- Simply assign each resource (i.e. mutex locks) a unique number.
- Resources must be acquired in order.
- If:

```
first_mutex = 1
second_mutex = 5
```

code for `thread_two` could not be written as follows:

```c
/* thread_one runs in this function */
void *do_work_one(void *param)
{
    pthread_mutex_lock(&first_mutex);
    pthread_mutex_lock(&second_mutex);
    /**
    * Do some work
    */
    pthread_mutex_unlock(&second_mutex);
    pthread_mutex_unlock(&first_mutex);
    pthread.exit(0);
}
/* thread_two runs in this function */
void *do_work_two(void *param)
{
    pthread_mutex_lock(&second_mutex);
    pthread_mutex_lock(&first_mutex);
    /**
    * Do some work
    */
    pthread_mutex_unlock(&first_mutex);
    pthread_mutex_unlock(&second_mutex);
    pthread.exit(0);
}
Deadlock Avoidance

- In deadlock prevention schemes, requests are constrained resulting in poor utilization and throughput.
- Alternative is to use additional information to allow systems to decide when and whether to grant requests. This is called **deadlock avoidance**.
- Avoidance is a dynamic strategy.
- Requires that each process’s **maximum claim for each resource type is known in advance**. A request is granted only if the resulting system state is a **safe state**, meaning that there is a way to avoid deadlock in this state while still continuing to allocate resources to each process.
Deadlock Avoidance

Requires that the system has some additional *a priori* information available

- Simplest and most useful model requires that each process declare the **maximum number** of resources of each type that it may need
- The deadlock-avoidance algorithm dynamically examines the resource-allocation state to ensure that there can never be a circular-wait condition
- Resource-allocation *state* is defined by the number of available and allocated resources, and the maximum demands of the processes
Safe State

- When a process requests an available resource, system must decide if immediate allocation leaves the system in a safe state.

- System is in **safe state** if there exists a sequence \(<P_1, P_2, \ldots, P_n>\) of ALL the processes in the systems such that for each \(P_i\), the resources that \(P_i\) can still request can be satisfied by currently available resources + resources held by all the \(P_j\), with \(j < i\).

- That is:
  - If \(P_i\)'s resource needs are not immediately available, then \(P_i\) can wait until all \(P_j\) have finished.
  - When \(P_j\) is finished, \(P_i\) can obtain needed resources, execute, return allocated resources, and terminate.
  - When \(P_i\) terminates, \(P_{i+1}\) can obtain its needed resources, and so on.
Safe States and Unsafe States

- If a system is in safe state ⇒ it cannot be a deadlock state, i.e., there is no deadlock in the current state.

- If a system is in unsafe state ⇒ possibility of deadlock – it may not be possible to avoid deadlock.

- Avoidance algorithm ⇒ ensure that a system will never enter an unsafe state.
Safe, Unsafe, Deadlock State
Avoidance Algorithms

- Single instance of a resource type
  - Use a resource-allocation graph

- Multiple instances of a resource type
  - Use the Banker’s Algorithm
Resource-Allocation Graph Scheme

- **Claim edge** $P_i \rightarrow R_j$ indicated that process $P_j$ may request resource $R_j$; represented by a dashed line
- Claim edge converts to request edge when a process requests a resource
- Request edge converted to an assignment edge when the resource is allocated to the process
- When a resource is released by a process, assignment edge reconverts to a claim edge
- Resources must be claimed *a priori* in the system
Resource-Allocation Graph

- **Diagram**: A graph with two tasks, $T_1$ and $T_2$, each requiring resources $R_1$ and $R_2$, respectively. The arrows indicate the dependencies or resource allocations between the tasks and resources.
Unsafe State In Resource-Allocation Graph
Resource-Allocation Graph Algorithm

- Suppose that process $P_i$ requests a resource $R_j$
- The request can be granted only if converting the request edge to an assignment edge does not result in the formation of a cycle in the resource allocation graph
Banker’s Algorithm

- Multiple instances of resources
- Each process must a priori claim maximum use
- When a process requests a resource it may have to wait
- When a process gets all its resources it must return them in a finite amount of time
Let $n =$ number of processes, and $m =$ number of resources types.

- **Available**: Vector of length $m$. If available $[j] = k$, there are $k$ instances of resource type $R_j$ available

- **Max**: $n \times m$ matrix. If Max $[i,j] = k$, then process $P_i$ may hold at most $k$ instances of resource type $R_j$

- **Allocation**: $n \times m$ matrix. If Allocation$[i,j] = k$ then $P_i$ is currently allocated $k$ instances of $R_j$

- **Need**: $n \times m$ matrix. If Need$[i,j] = k$, then $P_i$ may need $k$ more instances of $R_j$ to complete its task

\[ \text{Need} \ [i,j] = \text{Max}[i,j] - \text{Allocation} \ [i,j] \]
Safety Algorithm

1. Let $Work$ and $Finish$ be vectors of length $m$ and $n$, respectively. Initialize:

   $Work = Available$

   $Finish[i] = false$ for $i = 0, 1, ..., n-1$

2. Find an $i$ such that both:
   (a) $Finish[i] = false$
   (b) $Need_i \leq Work$

   If no such $i$ exists, go to step 4

3. $Work = Work + Allocation_i$

   $Finish[i] = true$

   go to step 2

4. If $Finish[i] == true$ for all $i$, then the system is in a safe state
Resource-Request Algorithm for Process $P_i$

$\text{Request}_i = \text{request vector for process } P_i$. If $\text{Request}_i[j] = k$ then process $P_i$ wants $k$ instances of resource type $R_j$

1. If $\text{Request}_i \leq \text{Need}_i$ go to step 2. Otherwise, raise error condition, since process has exceeded its maximum claim

2. If $\text{Request}_i \leq \text{Available}$, go to step 3. Otherwise $P_i$ must wait, since resources are not available

3. Pretend to allocate requested resources to $P_i$ by modifying the state as follows:

   $$\text{Available} = \text{Available} - \text{Request}_i;$$
   $$\text{Allocation}_i = \text{Allocation}_i + \text{Request}_i;$$
   $$\text{Need}_i = \text{Need}_i - \text{Request}_i;$$

   - If safe $\Rightarrow$ the resources are allocated to $P_i$
   - If unsafe $\Rightarrow P_i$ must wait, and the old resource-allocation state is restored
Example of Banker’s Algorithm

- 5 processes $P_0$ through $P_4$;
- 3 resource types: $A$ (10 instances), $B$ (5 instances), and $C$ (7 instances)
- Snapshot at time $T_0$:

<table>
<thead>
<tr>
<th></th>
<th>Allocation</th>
<th>Max</th>
<th>Available</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$A$</td>
<td>$B$</td>
<td>$C$</td>
</tr>
<tr>
<td>$P_0$</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$P_1$</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$P_2$</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>$P_3$</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$P_4$</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>
Example (Cont.)

The content of the matrix *Need* is defined to be *Max – Allocation*

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Need</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P₀</td>
<td>7</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>P₁</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>P₂</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P₃</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>P₄</td>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

The system is in a safe state since the sequence < *P*₁, *P*₃, *P*₄, *P*₂, *P*₀> satisfies safety criteria.
Example: $P_1$ Request (1,0,2)

- Check that Request ≤ Available (that is, (1,0,2) ≤ (3,3,2) ⇒ true

<table>
<thead>
<tr>
<th>Allocation</th>
<th>Need</th>
<th>Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>$B$</td>
<td>$C$</td>
</tr>
<tr>
<td>$P_0$</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$P_1$</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>$P_2$</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>$P_3$</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>$P_4$</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

- Executing safety algorithm shows that sequence $< P_1, P_3, P_4, P_0, P_2 >$ satisfies safety requirement

- Can request for (3,3,0) by $P_4$ be granted?

- Can request for (0,2,0) by $P_0$ be granted?
Deadlock Detection

- Allow system to enter deadlock state
- Detection algorithm
- Recovery scheme
Single Instance of Each Resource Type

- Maintain **wait-for** graph
  - Nodes are processes
  - \( P_i \rightarrow P_j \) if \( P_i \) is waiting for \( P_j \)

- Periodically invoke an algorithm that searches for a cycle in the graph. If there is a cycle, there exists a deadlock

- An algorithm to detect a cycle in a graph requires an order of \( n^2 \) operations, where \( n \) is the number of vertices in the graph
Resource-Allocation Graph and Wait-for Graph

Resource-Allocation Graph

Corresponding wait-for graph
Several Instances of a Resource Type

- **Available**: A vector of length $m$ indicates the number of available resources of each type
- **Allocation**: An $n \times m$ matrix defines the number of resources of each type currently allocated to each process
- **Request**: An $n \times m$ matrix indicates the current request of each process. If $Request[i][j] = k$, then process $P_i$ is requesting $k$ more instances of resource type $R_j$. 
Detection Algorithm

1. Let $Work$ and $Finish$ be vectors of length $m$ and $n$, respectively. Initialize:
   (a) $Work = Available$
   (b) For $i = 1, 2, \ldots, n$, if $Allocation_i \neq 0$, then $Finish[i] = false$; otherwise, $Finish[i] = true$

2. Find an index $i$ such that both:
   (a) $Finish[i] == false$
   (b) $Request_i \leq Work$

   If no such $i$ exists, go to step 4
Detection Algorithm (Cont.)

3. \( \textit{Work} = \textit{Work} + \textit{Allocation}_i \)
   \( \textit{Finish}[i] = \text{true} \)
   go to step 2

4. If \( \textit{Finish}[i] == \text{false} \), for some \( 1 \leq i \leq n \), then the system is in deadlock state. Moreover, if \( \textit{Finish}[i] == \text{false} \), then \( \textit{P}_i \) is deadlocked

Algorithm requires an order of \( O(m \times n^2) \) operations to detect whether the system is in deadlocked state
Example of Detection Algorithm

- Five processes $P_0$ through $P_4$; three resource types $A$ (7 instances), $B$ (2 instances), and $C$ (6 instances)

- Snapshot at time $T_0$:

<table>
<thead>
<tr>
<th>Allocation</th>
<th>Request</th>
<th>Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$ $B$ $C$</td>
<td>$A$ $B$ $C$</td>
<td>$A$ $B$ $C$</td>
</tr>
<tr>
<td>$P_0$</td>
<td>0 1 0</td>
<td>0 0 0</td>
</tr>
<tr>
<td>$P_1$</td>
<td>2 0 0</td>
<td>2 0 2</td>
</tr>
<tr>
<td>$P_2$</td>
<td>3 0 3</td>
<td>0 0 0</td>
</tr>
<tr>
<td>$P_3$</td>
<td>2 1 1</td>
<td>1 0 0</td>
</tr>
<tr>
<td>$P_4$</td>
<td>0 0 2</td>
<td>0 0 2</td>
</tr>
</tbody>
</table>

- Sequence $<P_0, P_2, P_3, P_1, P_4>$ will result in $Finish[i] = true$ for all $i$
Example (Cont.)

- $P_2$ requests an additional instance of type $C$

<table>
<thead>
<tr>
<th>Request</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_0$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$P_1$</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>$P_2$</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$P_3$</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$P_4$</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

- State of system?
  - Can reclaim resources held by process $P_0$, but insufficient resources to fulfill other processes; requests
  - Deadlock exists, consisting of processes $P_1$, $P_2$, $P_3$, and $P_4$
Detection-Algorithm Usage

- When, and how often, to invoke depends on:
  - How often a deadlock is likely to occur?
  - How many processes will need to be rolled back?
    - one for each disjoint cycle

- If detection algorithm is invoked arbitrarily, there may be many cycles in the resource graph and so we would not be able to tell which of the many deadlocked processes “caused” the deadlock.
Recovery from Deadlock: Process Termination

- Abort all deadlocked processes
- Abort one process at a time until the deadlock cycle is eliminated

In which order should we choose to abort?
1. Priority of the process
2. How long process has computed, and how much longer to completion
3. Resources the process has used
4. Resources process needs to complete
5. How many processes will need to be terminated
6. Is process interactive or batch?
Recovery from Deadlock: Resource Preemption

- **Selecting a victim** – minimize cost

- **Rollback** – return to some safe state, restart process for that state

- **Starvation** – same process may always be picked as victim, include number of rollback in cost factor
End of Chapter 8