CS370 Operating Systems
Colorado State University
Yashwant K Malaiya
Spring 1019 L14
Deadlocks

Slides based on
• Text by Silberschatz, Galvin, Gagne
• Various sources
FAQ

- Creating threads without using pthreads
- Do java threads use pthreads?
- Why modern OSs do not actively prevent deadlocks?
  - mechanism by which locks are always acquired in a defined order

<table>
<thead>
<tr>
<th>thread one function</th>
<th>thread two function</th>
</tr>
</thead>
<tbody>
<tr>
<td>void *do_work_one(void *param)</td>
<td></td>
</tr>
</tbody>
</table>
| { | }
|   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); |
| } | 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); |
| } |
Methods for Handling Deadlocks

• Ensure that the system will never enter a deadlock state:
  – Deadlock prevention
    • ensuring that at least one of the 4 conditions cannot hold
  – Deadlock avoidance
    • Dynamically examines the resource-allocation state to ensure that there can never be a circular-wait condition

• Allow the system to enter a deadlock state
  – Detect and then recover. Hope is that it happens rarely.

• Ignore the problem and pretend that deadlocks never occur in the system; used by most operating systems, including UNIX.
For a deadlock to occur, each of the four necessary conditions must hold. By ensuring that at least one of these conditions cannot hold, we can prevent the occurrence of a deadlock.

**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 \( \{P_0, P_1, ..., P_n\} \) of waiting processes that are circularly waiting.
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
Deadlock Avoidance: Handling resource requests

• For each resource request:
  – Decide whether or not process should wait
    • To avoid possible future deadlock

• Predicated on:
  1. Currently available resources
  2. Currently allocated resources
  3. *Future requests and releases of each process*
Avoidance: amount and type of information needed

- **Resource allocation state**
  - Number of available and allocated resources
  - Maximum demands of processes

- **Dynamically** examine resource allocation state
  - Ensure circular-wait cannot exist

- Simplest model:
  - Declare maximum number of resources for each type
  - Use information to avoid deadlock
Safe Sequence

System must decide if immediate allocation leaves the system in a safe state

System is in **safe state** if there exists a sequence \(<P_1, P_2, ..., P_n>\) of ALL the processes 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 \(P_i\) resource needs are not immediately available, then \(P_i\) can wait until all \(P_j\) have finished and released resources
    - When \(P_i\) terminates, \(P_{i+1}\) can obtain its needed resources, and so on

- If no such sequence exists: system state is **unsafe**
Deadlock avoidance: Safe states

• If the system can:
  – Allocate resources to each process in some order
    • Up to the maximum for the process
  – Still avoid deadlock
  – Then it is in a safe state

• A system is safe ONLY IF there is a safe sequence

• A safe state is not a deadlocked state
  – Deadlocked state is an unsafe state
  – Not all unsafe states are deadlock
Safe, Unsafe, Deadlock State

Examples of safe and unsafe states in next 3 slides
Example A: Assume 12 Units in the system

<table>
<thead>
<tr>
<th></th>
<th>Max need</th>
<th>Current holding</th>
</tr>
</thead>
<tbody>
<tr>
<td>P0</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>P1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>P2</td>
<td>9</td>
<td>2</td>
</tr>
</tbody>
</table>

At T0:
9 units allocated
3 (12-9) units available

A unit could be a drive, a block of memory etc.

• **At time T0 is the system is in a safe state?**
  – **Try sequence** <P1, P0, P2>
  – P1 can be given 2 units
  – When P1 releases its resources; there are 5 units
  – P0 uses 5 and subsequently releases them (# 10 now)
  – P2 can then proceed.

• Thus <P1, P0, P2> is a safe sequence, and at T0 system was in a safe state
Example A: Assume 12 Units in the system (timing)

Is the state at T0 safe?

<table>
<thead>
<tr>
<th></th>
<th>Max need</th>
<th>Current holding</th>
<th>T0</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
</tr>
</thead>
<tbody>
<tr>
<td>av</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P0</td>
<td>10</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td>9</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Thus the state at T0 is safe.
Example B: 12 Units available in the system

<table>
<thead>
<tr>
<th>Max need</th>
<th>T0</th>
<th>T1 safe?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Av</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>P0</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>P1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>P2</td>
<td>9</td>
<td>3 Is that OK?</td>
</tr>
</tbody>
</table>

Before T1: 3 units available
At T1: 2 units available

- At time **T1**, P2 is allocated 1 more units. Is that a good decision?
  - Now only P1 can proceed.
  - When P1 releases its resources; there are 4 units
  - P0 needs 5 more, P2 needs 6 more. Deadlock.
    - **Mistake** in granting P2 the additional unit.
- The state at **T1** is not a safe state.
Avoidance Algorithms

• **Single instance** of a resource type
  – Use a resource-allocation graph scheme

• **Multiple instances** of a resource type
  – Use the **banker’s algorithm** (Dijkstra).
Resource-Allocation Graph Scheme

- **Claim edge** $P_i \rightarrow R_j$ indicated that process $P_i$ 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.
Suppose $P_2$ requests $R_2$. Although $R_2$ is currently free, we cannot allocate it to $P_2$, since this action will create a cycle getting system in an unsafe state. If $P_1$ requests $R_2$, and $P_2$ requests $R_1$, then a deadlock will occur.
• 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: examining a request

- Multiple instances of resources.
- Each process must a priori claim maximum use.
- When a process requests a resource:
  - it may have to wait (resource request algorithm)
  - Request not granted if the resulting system state is unsafe (safety algorithm)
- When a process gets all its resources it must return them in a finite amount of time.
- Modeled after a banker in a small town making loans.
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

**Processes vs resources:**

- **Max**: $n \times m$ matrix. If $Max [i,j] = k$, then process $P_i$ may request 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

\[
Need [i,j] = Max[i,j] – Allocation [i,j]
\]
1. Let Work and Finish be vectors of length $m$ and $n$, respectively. Initialize:
   \[ \text{Work} = \text{Initially Available resources} \]
   \[ \text{Finish} [i] = false \text{ for } i = 0, 1, ..., n-1 \]

2. Find a process $i$ such that both:
   (a) $\text{Finish} [i] = false$
   (b) $\text{Need}_i \leq \text{Work}$
   If no such $i$ exists, go to step 4

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

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

Notation: $Request_i$ = request vector for process $P_i$. If $Request_i[j] = k$ then process $P_i$ wants $k$ instances of resource type $R_j$

Algorithm: Should the allocation request be granted?

1. If $Request_i \leq Need_i$ go to step 2. Otherwise, raise error condition, since process has exceeded its maximum claim.
2. If $Request_i \leq Available$, go to step 3. Otherwise $P_i$ must wait, since resources are not available.
3. Is allocation safe?: Pretend to allocate requested resources to $P_i$ by modifying the state as follows:
   \[
   \begin{align*}
   Available &= Available - Request_i; \\
   Allocation_i &= Allocation_i + Request_i; \\
   Need_i &= Need_i - Request_i;
   \end{align*}
   \]
   - If safe $\Rightarrow$ the resources are allocated to $P_i$
   - If unsafe $\Rightarrow$ $P_i$ must wait, and the old resource-allocation state is preserved.

Safety Algorithm
Example A: Banker’s Algorithm

- 5 processes $P_0$ through $P_4$;
- 3 resource types: $A$ (10 instances), $B$ (5 instances), and $C$ (7 instances)
- Is it a safe state?

<table>
<thead>
<tr>
<th>Process</th>
<th>Max</th>
<th>Allocation</th>
<th>Need</th>
</tr>
</thead>
<tbody>
<tr>
<td>type</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>available</td>
<td></td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>P0</td>
<td>7</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>P1</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>P2</td>
<td>9</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P3</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>P4</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>
Example A: Banker’s Algorithm

- Is it a safe state?
- Yes, since the sequence < P1, P3, P4, P2, P0> satisfies safety criteria

<table>
<thead>
<tr>
<th>Process</th>
<th>Max</th>
<th>Allocation</th>
<th>Need</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>type</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>available</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>P0</td>
<td>7</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>P1</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>P2</td>
<td>9</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>P3</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>P4</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

P1 run to completion. Available becomes \([3 \, 3 \, 2]+[2 \, 0 \, 0] = [5 \, 3 \, 2]\)
P3 run to completion. Available becomes \([5 \, 3 \, 2]+[2 \, 1 \, 1] = [7 \, 4 \, 3]\)
P4 run to completion. Available becomes \([7 \, 4 \, 3]+[0 \, 0 \, 2] = [7 \, 4 \, 5]\)
P2 run to completion. Available becomes \([7 \, 4 \, 5]+[3 \, 0 \, 2] = [10 \, 4 \, 7]\)
P0 run to completion. Available becomes \([10 \, 4 \, 7]+[0 \, 1 \, 0] = [10 \, 5 \, 7]\)

Hence state above is safe.
Ex B: Assume now $P_1$ Requests (1,0,2)

- Check that Request ≤ Available. $(1,0,2) \leq (3,3,2) \rightarrow true$.
- Check for safety after pretend allocation. $P_1$ allocation would be $(2,0,0) + (1,0,2) = 302$

<table>
<thead>
<tr>
<th>Process</th>
<th>Max</th>
<th>Allocation</th>
<th>Need</th>
</tr>
</thead>
<tbody>
<tr>
<td>type A</td>
<td>B</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>available</td>
<td></td>
<td>2 3 0</td>
<td>A</td>
</tr>
<tr>
<td>P0</td>
<td>7</td>
<td>5 3 0</td>
<td>0</td>
</tr>
<tr>
<td>P1</td>
<td>3</td>
<td>2 2 2</td>
<td>3</td>
</tr>
<tr>
<td>P2</td>
<td>9</td>
<td>0 2 0</td>
<td>3</td>
</tr>
<tr>
<td>P3</td>
<td>2</td>
<td>2 2 2</td>
<td>2</td>
</tr>
<tr>
<td>P4</td>
<td>4</td>
<td>3 3 3</td>
<td>0</td>
</tr>
</tbody>
</table>

Sequence $< P_1, P_3, P_4, P_0, P_2 >$ satisfies safety requirement.

Hence state above is safe, thus the allocation would be safe.
Ex C,D: Additional Requests ..

- Given State is (same as previous slide)

<table>
<thead>
<tr>
<th>Process</th>
<th>Max</th>
<th>Allocation</th>
<th>Need</th>
</tr>
</thead>
<tbody>
<tr>
<td>type</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>available</td>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>P0</td>
<td>7</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>P1</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>P2</td>
<td>9</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>P3</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>P4</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

P4 request for (3,3,0): cannot be granted - resources are not available.
P0 request for (0,2,0): cannot be granted since the resulting state is unsafe.
Bankers Algorithm: Practical Issues

- Processes rarely know in advance about their maximum resource needs
- Number of processes is not fixed
  - Varies dynamically
- Resources thought to be available can disappear
- Few systems use this algorithm
Deadlock Detection

• Allow system to enter deadlock state
• Detection algorithm
  – Single instance of each resource:
    • wait-for graph
  – Multiple instances:
    • detection algorithm (based on Banker’s algorithm)
• Recovery scheme
• Maintain **wait-for graph** (based on resource allocation graph)
  – Nodes are processes
  – $P_i \rightarrow P_j$ if $P_i$ is waiting for $P_j$
    
  – **Deadlock if cycles**

• 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

(a) Resource-Allocation Graph

(b) Corresponding wait-for graph

Has cycles. Deadlock.
**Banker’s algorithm:** Can requests by all processes be satisfied?

- **Available:** A vector of length \(m\) indicates the number of available (currently free) 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 \(\text{Request}_{ij} = k\), then process \(P_i\) is requesting \(k\) more instances of resource type \(R_j\).
Detection Algorithm

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

2. Find an index \(i\) such that both:
   (a) \(\text{Finish}[i] = \text{false}\)
   (b) \(\text{Request}_i \leq \text{Work}\)
   If no such \(i\) exists, go to step 4

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

4. If \(\text{Finish}[i] = \text{false}\), for some \(i, 1 \leq i \leq n\), then the system is in deadlock state. Moreover, if \(\text{Finish}[i] = \text{false}\), then \(P_i\) 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)
- Sequence $<P_0, P_2, P_3, P_1, P_4>$ will result in $\text{Finish}[i] = true$ for all $i$. No deadlock

<table>
<thead>
<tr>
<th>Process</th>
<th>Allocation</th>
<th>Request</th>
</tr>
</thead>
<tbody>
<tr>
<td>type</td>
<td>A  B  C</td>
<td>A  B  C</td>
</tr>
<tr>
<td>available</td>
<td>0  0  0</td>
<td></td>
</tr>
<tr>
<td>P0</td>
<td>0  1  0</td>
<td>0  0  0</td>
</tr>
<tr>
<td>P1</td>
<td>2  0  0</td>
<td>2  0  2</td>
</tr>
<tr>
<td>P2</td>
<td>3  0  3</td>
<td>0  0  0</td>
</tr>
<tr>
<td>P3</td>
<td>2  1  1</td>
<td>1  0  0</td>
</tr>
<tr>
<td>P4</td>
<td>0  0  2</td>
<td>0  0  2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>After</th>
<th>available</th>
</tr>
</thead>
<tbody>
<tr>
<td>ini</td>
<td>0  0  0</td>
</tr>
<tr>
<td>P0</td>
<td>0  1  0</td>
</tr>
<tr>
<td>P2</td>
<td>3  1  3</td>
</tr>
<tr>
<td>P3</td>
<td>5  2  4</td>
</tr>
<tr>
<td>P1</td>
<td>7  2  4</td>
</tr>
<tr>
<td>P4</td>
<td>7  2  6</td>
</tr>
</tbody>
</table>
Example of Detection Algorithm (cont)

- \( P_2 \) requests an additional instance of type \( C \)

<table>
<thead>
<tr>
<th>Process</th>
<th>Allocation</th>
<th>Request</th>
</tr>
</thead>
<tbody>
<tr>
<td>type</td>
<td>A  B  C</td>
<td>A  B  C</td>
</tr>
<tr>
<td>available</td>
<td>0  0  0</td>
<td></td>
</tr>
<tr>
<td>P0</td>
<td>0  1  0</td>
<td>0  0  0</td>
</tr>
<tr>
<td>P1</td>
<td>2  0  0</td>
<td>2  0  2</td>
</tr>
<tr>
<td>P2</td>
<td>3  0  3</td>
<td>0  0  1</td>
</tr>
<tr>
<td>P3</td>
<td>2  1  1</td>
<td>1  0  0</td>
</tr>
<tr>
<td>P4</td>
<td>0  0  2</td>
<td>0  0  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

Choices

• 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?
• 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
Deadlock recovery through rollbacks

- **Checkpoint** process periodically
  - Contains memory image and resource state
- Deadlock detection tells us *which* resources are needed
- Process owning a needed resource
  - **Rolled back** to before it acquired needed resource
    - Work done since rolled back checkpoint discarded
  - **Assign** resource to deadlocked process
Livelocks

In a livelock two processes need each other’s resource

• Both run and make no progress, but neither process blocks
• Use CPU quantum over and over without making progress

Ex: If fork fails because process table is full

• Wait for some time and try again
• But there could be a collection of processes each trying to do the same thing
• Avoided by ensuring that only one process (chosen randomly or by priority) takes action
Main Memory

Slides based on
• Text by Silberschatz, Galvin, Gagne
• Various sources
Chapter 9: Main Memory

Objectives:
- Organizing memory for multiprogramming environment
  - Partitioned vs separate address spaces
- Memory-management techniques
  - Virtual vs physical addresses
  - Segmentation
  - Paging: page tables, caching (“TLBs”)
- Examples: the Intel (old/new) and ARM architectures
- Not considered here: Virtual memory: Main memory/disk interaction. Next topic.