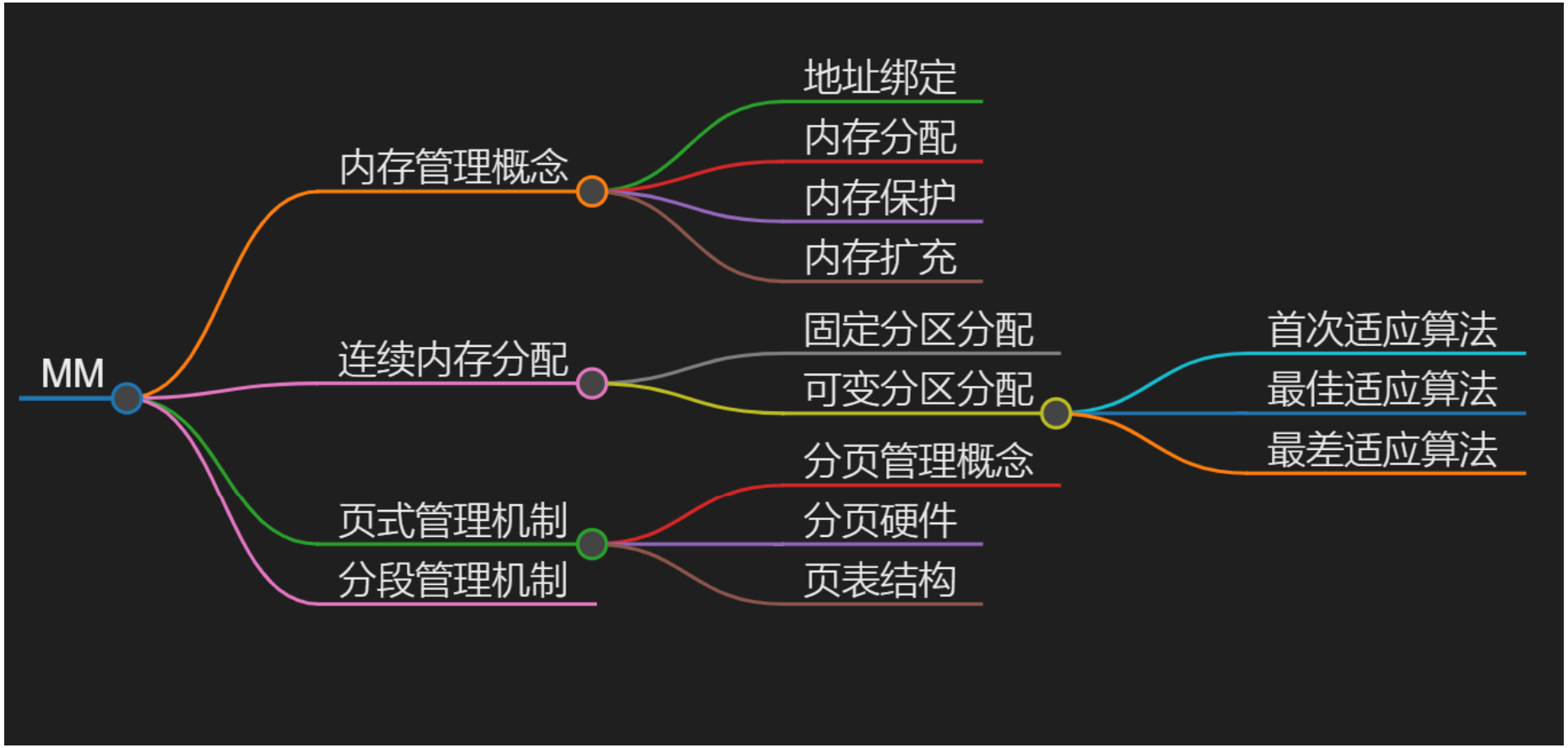
操作系统

L15 虚存（上）

胡燕
大连理工大学 软件学院

0-主存管理回顾

主体内容



单选题 5分

设置

下列关于TLB（快表）的描述中，错误的是（ ）。

A 快表的内容是页表的子集

B 快表保存在内存固定位置

C 对快表的查找是按内容并行完成

D 引入快表可以加快地址转换速度

提交

单选题 5分

设置

已知系统为32位物理地址，采用48位虚拟地址，页面大小为4KB，页表项大小为8字节。假设系统使用纯页式存储，则要采用几级页表，页内偏移多少位?

A 3，12

B 4，12

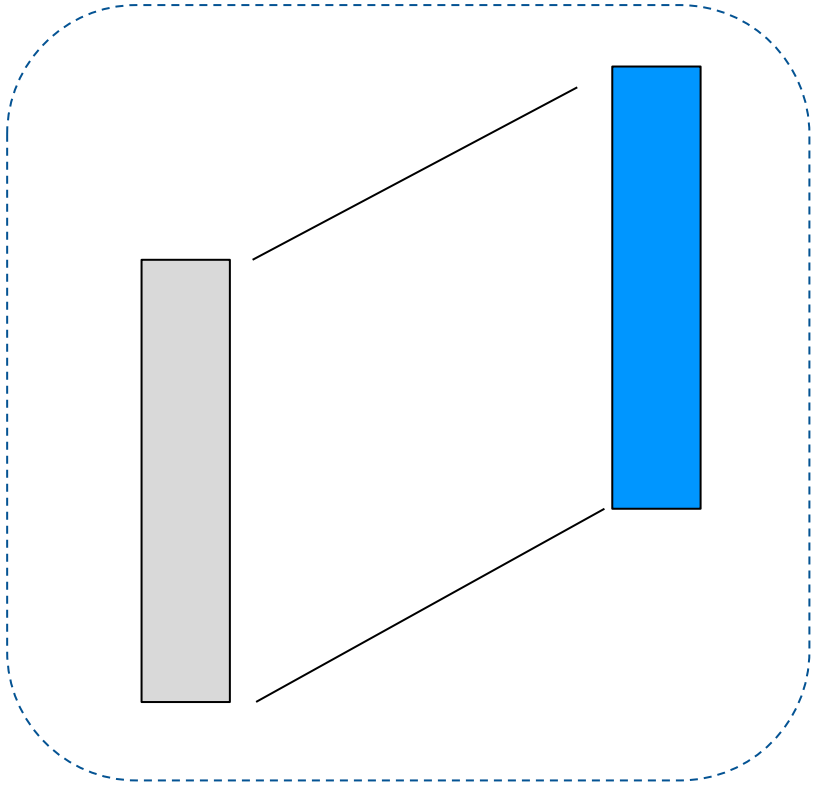
C 3，14

D 4，14

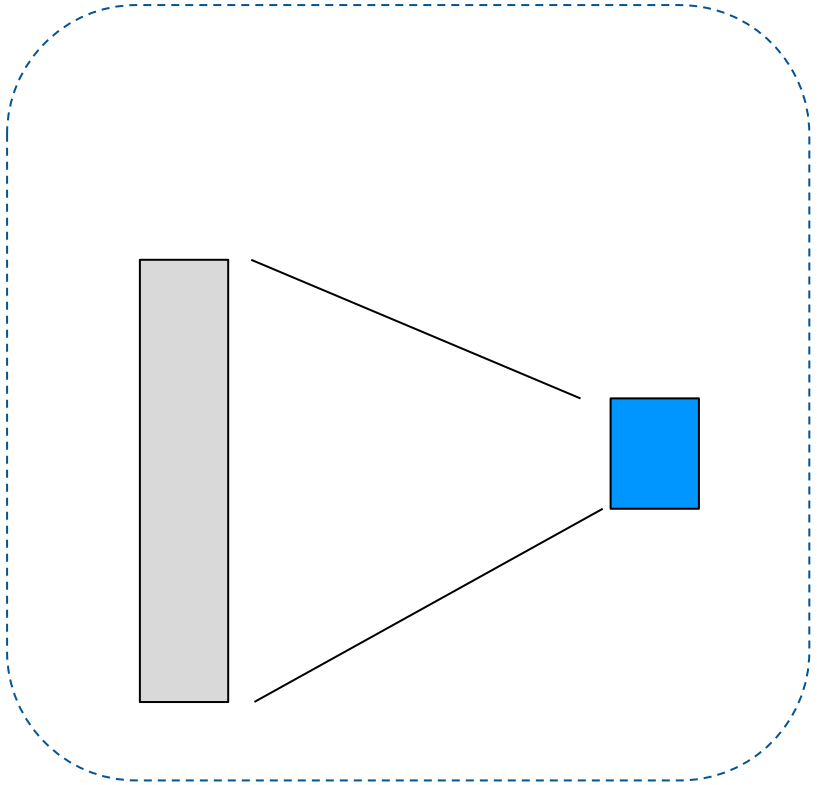
提交

主体内容

Q: 为系统中的每个进程如何进行物理内存分配?

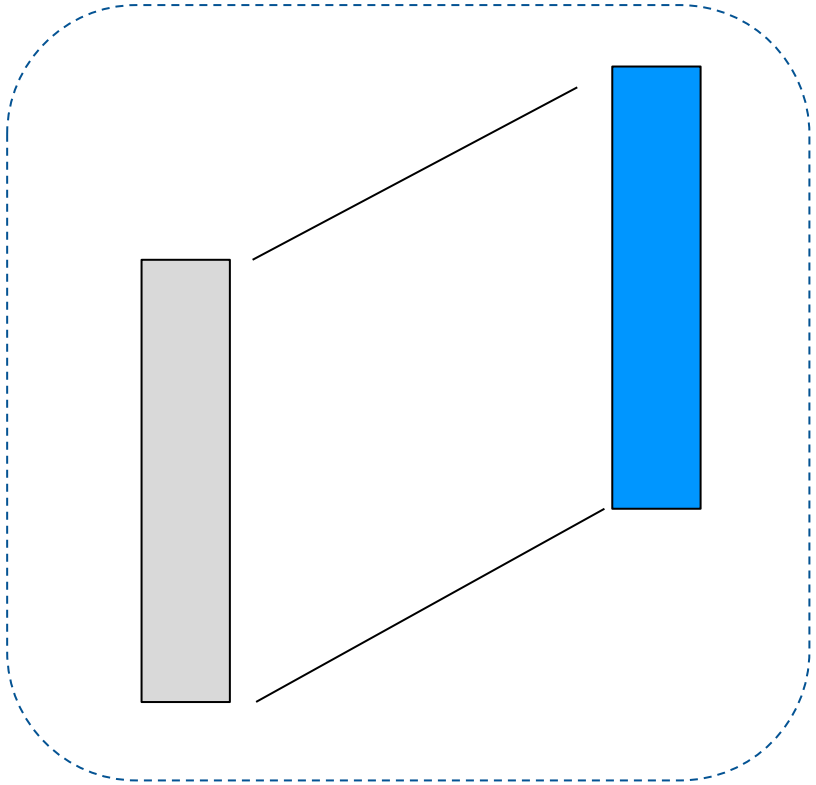


or

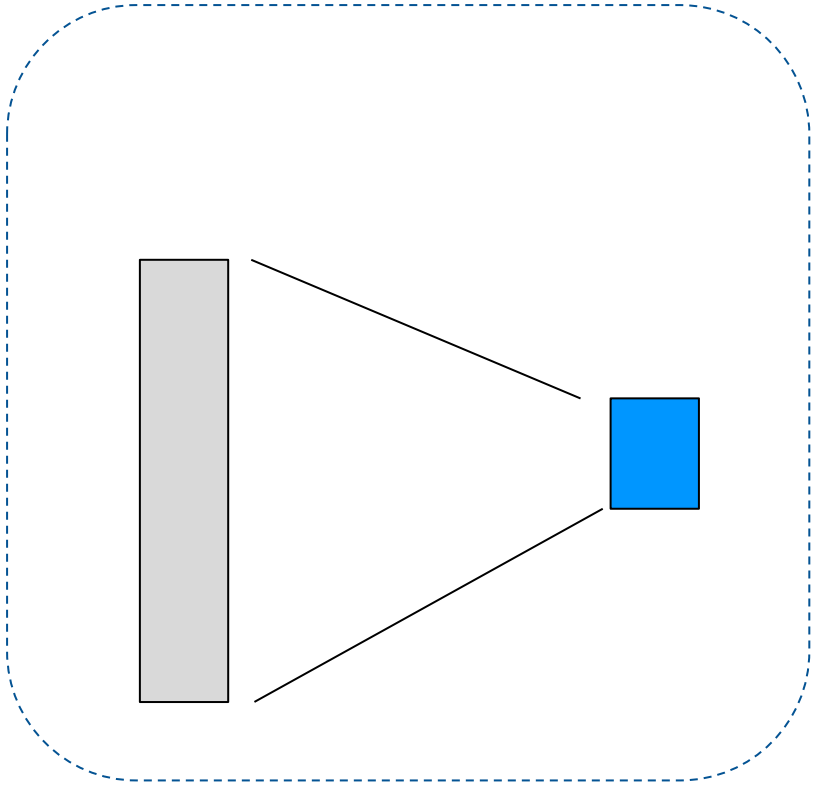


主体内容

Q: 为系统中的每个进程如何进行物理内存分配?



or



因为僧多粥少，所以内存资源的使用要精打细算

虚存概念

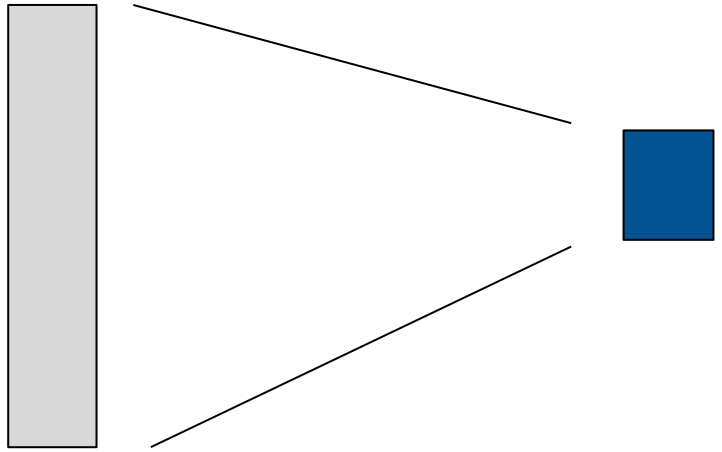
Virtual Memory

01

虚存思想

虚存的核心思想

用较少的物理内存，支撑较大的逻辑地址空间



如何践行虚存的核心思想

实存思想下的任务执行方案

每个程序执行时，将对应程序的所有代码和数据全部加载进物理内存，然后，再开始执行

虚存思想下的任务执行方案

进程初始状态时，并不一次性将逻辑地址空间的所有代码和数据加载入物理内存（若不考虑性能因素，甚至可以初始零加载）

进程执行过程中，根据当前执行的局部，根据需要加载所需的逻辑空间内容到物理内存

允许进程根据当前工作的实际需要，进行适量物理内存分配

虚存带来的优势

- 不再需要一次性加载程序的所有代码和一次运行所需的所有数据，可以节省大量物理内存
- 使得进程的创建更有效率
- 减少IO操作
- 支持更多的并发进程
- 方便实施在进程间共享物理内存

虚存思想

虚存机制的理论依据

大逻辑空间

小物理空间

支点：局部性原理

程序的局部性原理：程序在执行时呈现出局部性规律，即在一段时间内，整个程序的执行仅限于程序中的某一小部分

实现虚存的关键技术：

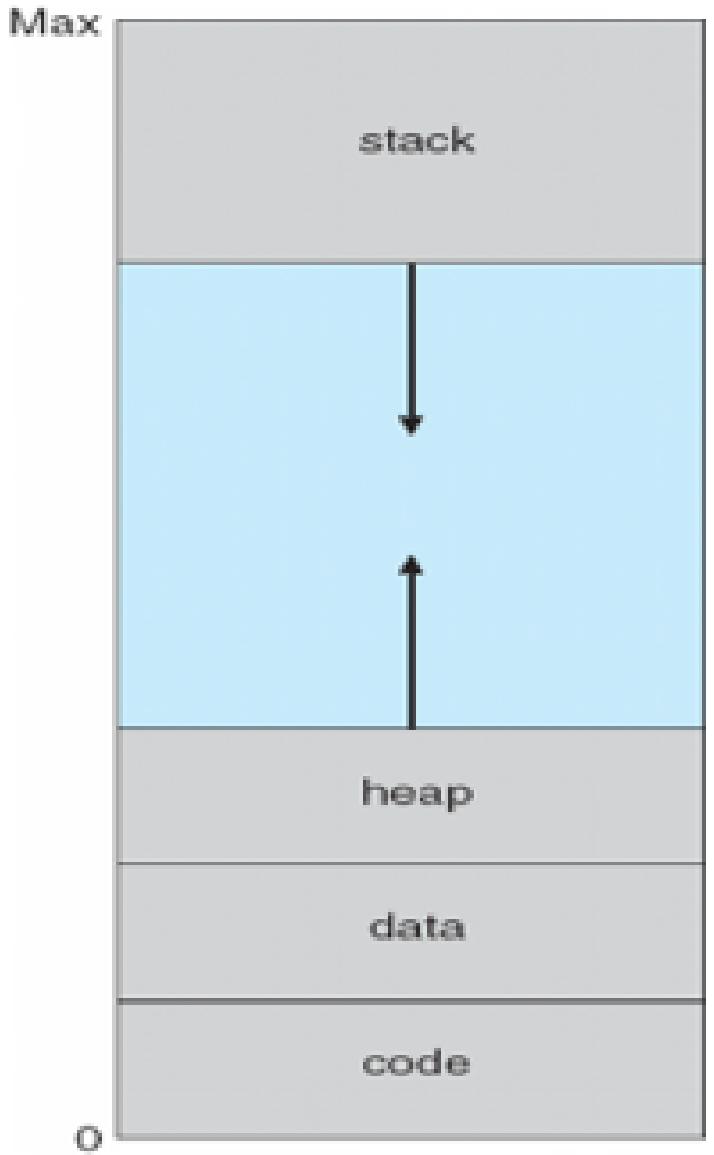
按需分配 demand driven

惰性加载 lazy loading

引入虚存的意义

形成较大的虚存空间

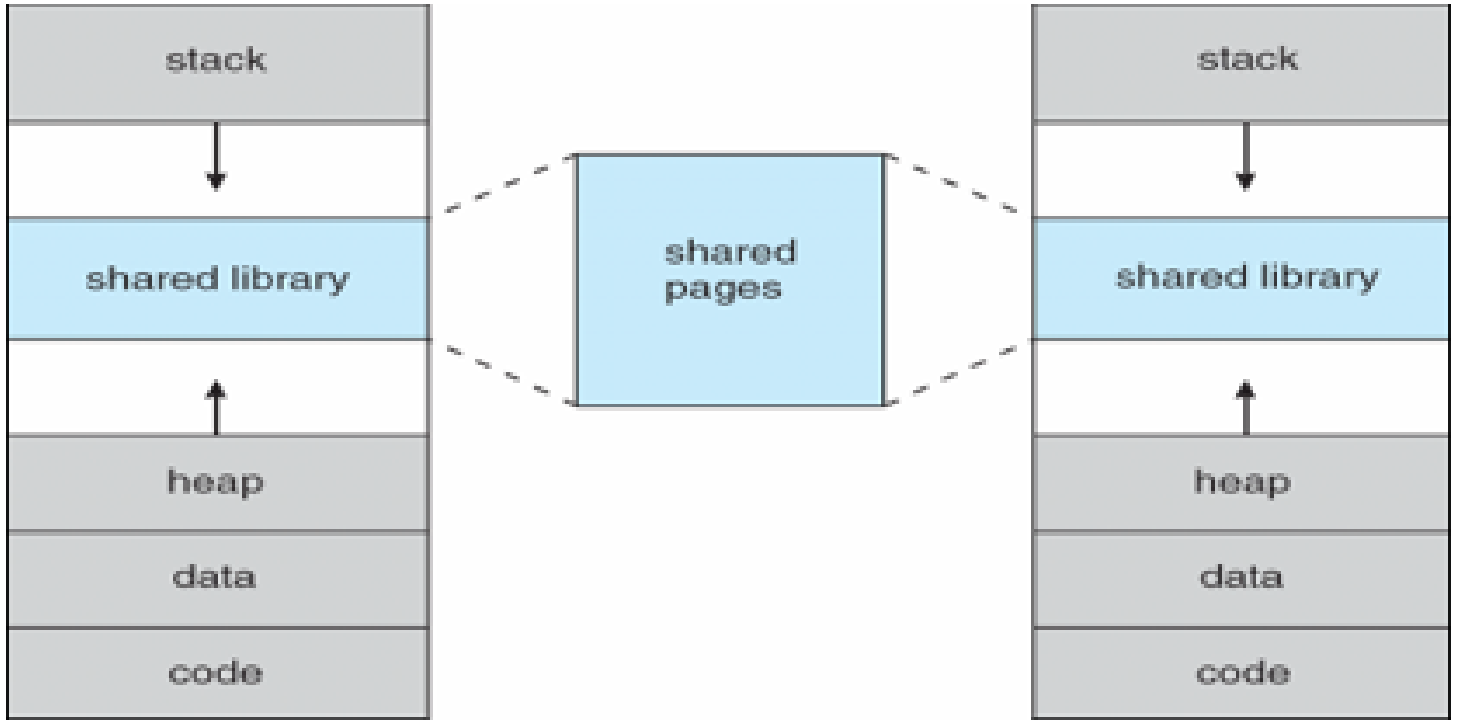
开发人员进行应用开发时，不需要太多考虑对逻辑地址空间大小进行限制的问题
通常，包含代码、数据、运行时堆和栈的逻辑地址空间都会存在较大富余



引入虚存的意义

便于进程间进行内存共享

可以将相同一块物理内存区域映射到不同进程的逻辑地址空间：仅需将需共享的物理页内容映射到进程的虚地址空间即可



按需调页

Pure Demand Paging

进程创建初始，系统仅为其划定虚存空间，而不实际分配任何物理页

执行期间，每一次访存，都会面临两种可能：

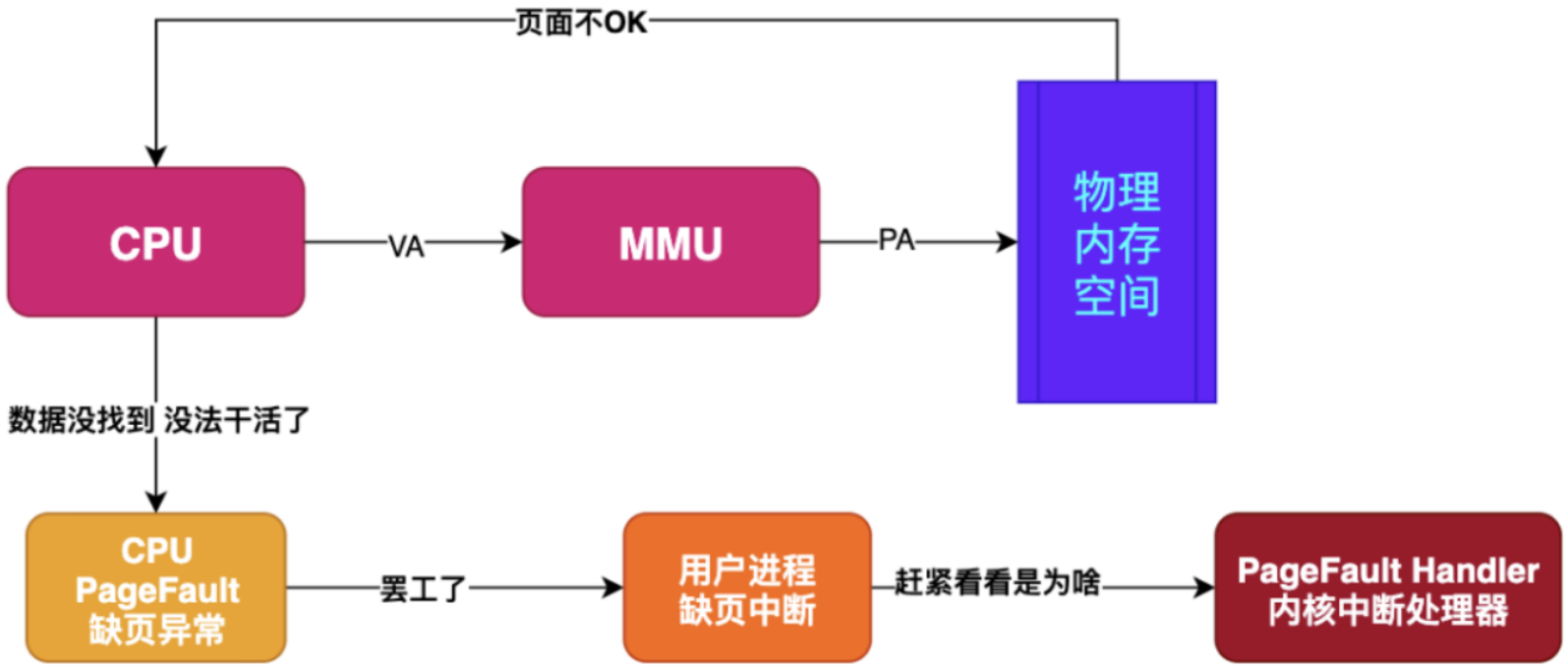
(1) 若访问的逻辑页未在内存，导致缺页异常

怎么办：在处理缺页异常的过程中，将该逻辑页调入某物理页

(2) 若访问的逻辑页已在内存，则正常访问

按需调页

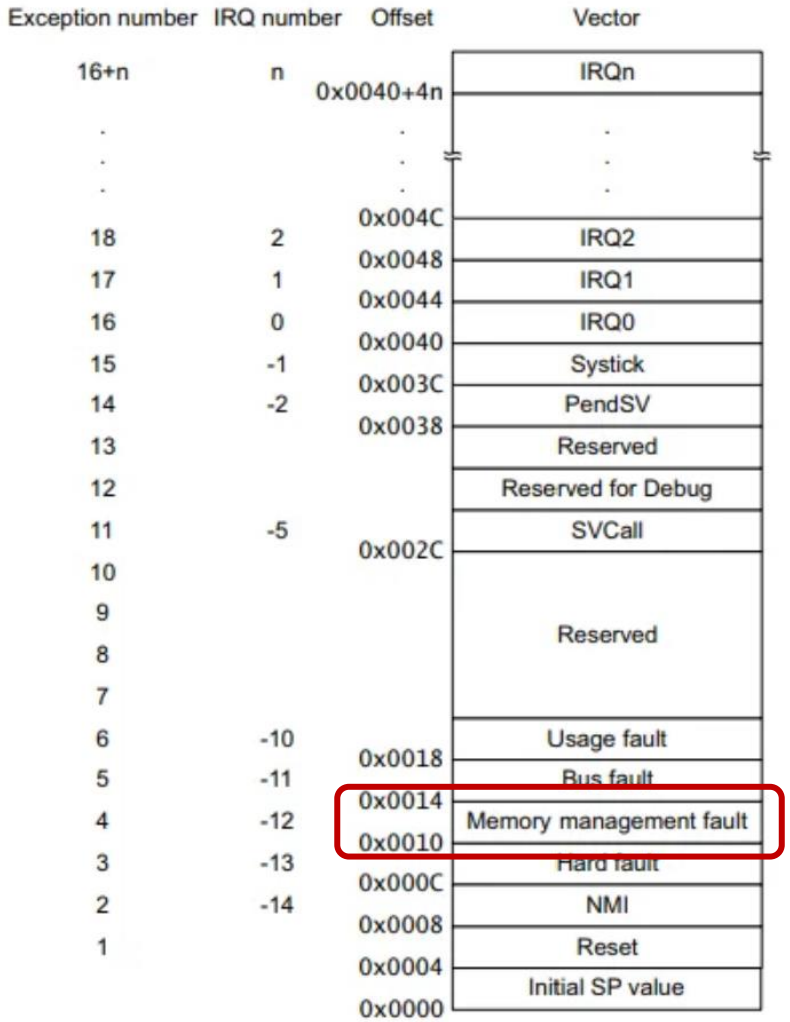
缺页异常



虚拟地址经过一番寻址后，没有找到其对应的物理地址，会发生缺页异常（Page Fault，页故障）

https://blog.csdn.net/g6U8W7p06dCO99fQ3/article/details/107240849

按需调页



ARM Contex M CPU的中断向量示意图

MMU fault checking

During the processing of a section or page, the MMU behaves differently because it is checking for faults. The MMU can generate these faults:

- *Alignment fault*
- *Translation fault*
- *Access bit fault*
- *Domain fault*
- *Permission fault*.

https://developer.arm.com/documentation/ddi0301/h/memory-management-unit/mmu-fault-checking

https://microcontrollerslab.com/what-is-interrupt-vector-table/

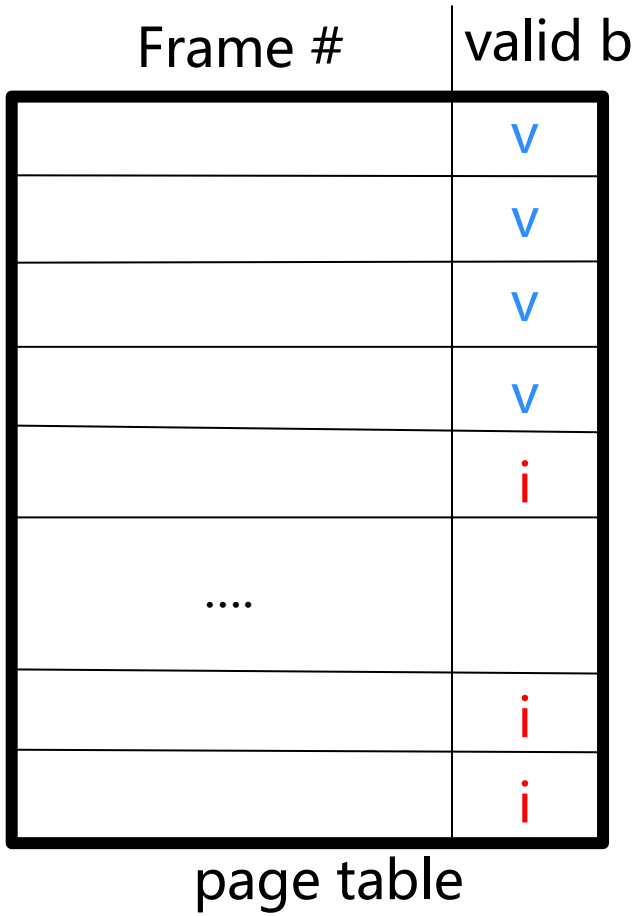
Q:

在基于按需调页的虚存机制中，如何使CPU能够判断发生缺页，从而通知其处理呢?

Demand Paging

按需调页：页表有效位

若要实施按需调页，MMU需要能够判定访问的逻辑页面是否在内存



方法：为每个页表项设定一个有效位（Valid bit）

有效位=v ➡ 对应逻辑页已经在物理内存

有效位=i ➡ 对应逻辑页未在物理内存

Q1: 有效位为i的逻辑页在哪里?

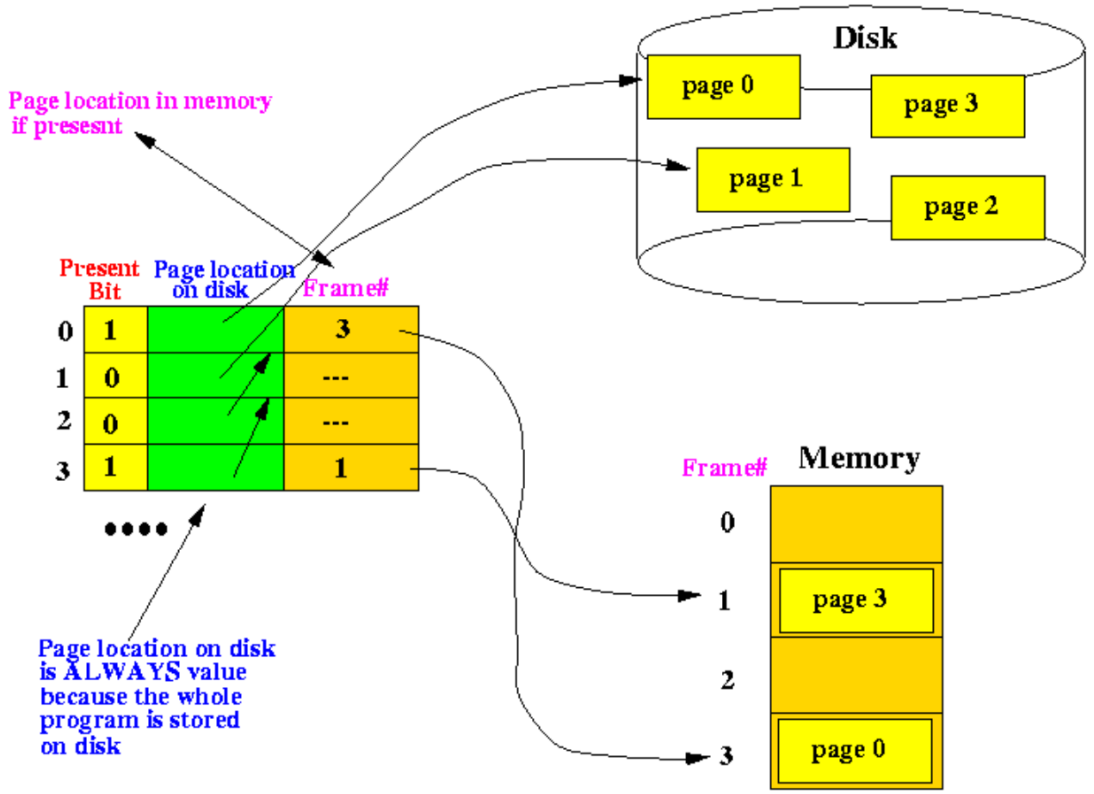
在磁盘上，可执行文件内

Demand Paging

发生缺页时，通过lazy page swapper（惰性页交换）进行加载

需要对页表项的结构做相应调整：增加有效位(Valid bit)

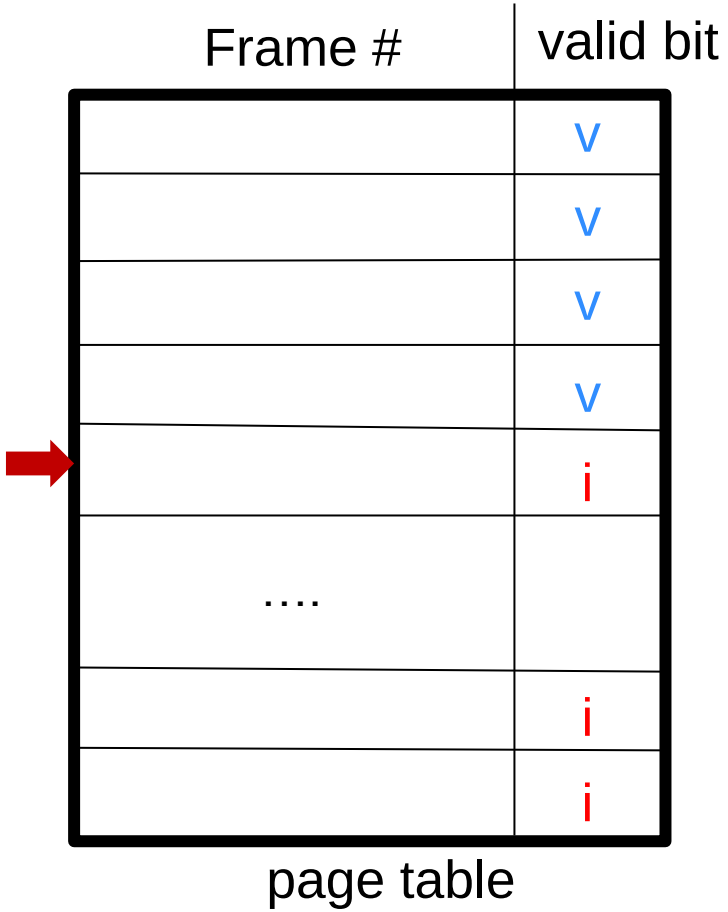
需要在访问未在物理内存的信息时，做出必要的处理（页面的惰性加载）



https://www.cs.emory.edu/~cheung/Courses/355/Syllabus/9-virtual-mem/page-fault.html

Demand Paging

按需调页：页表有效位



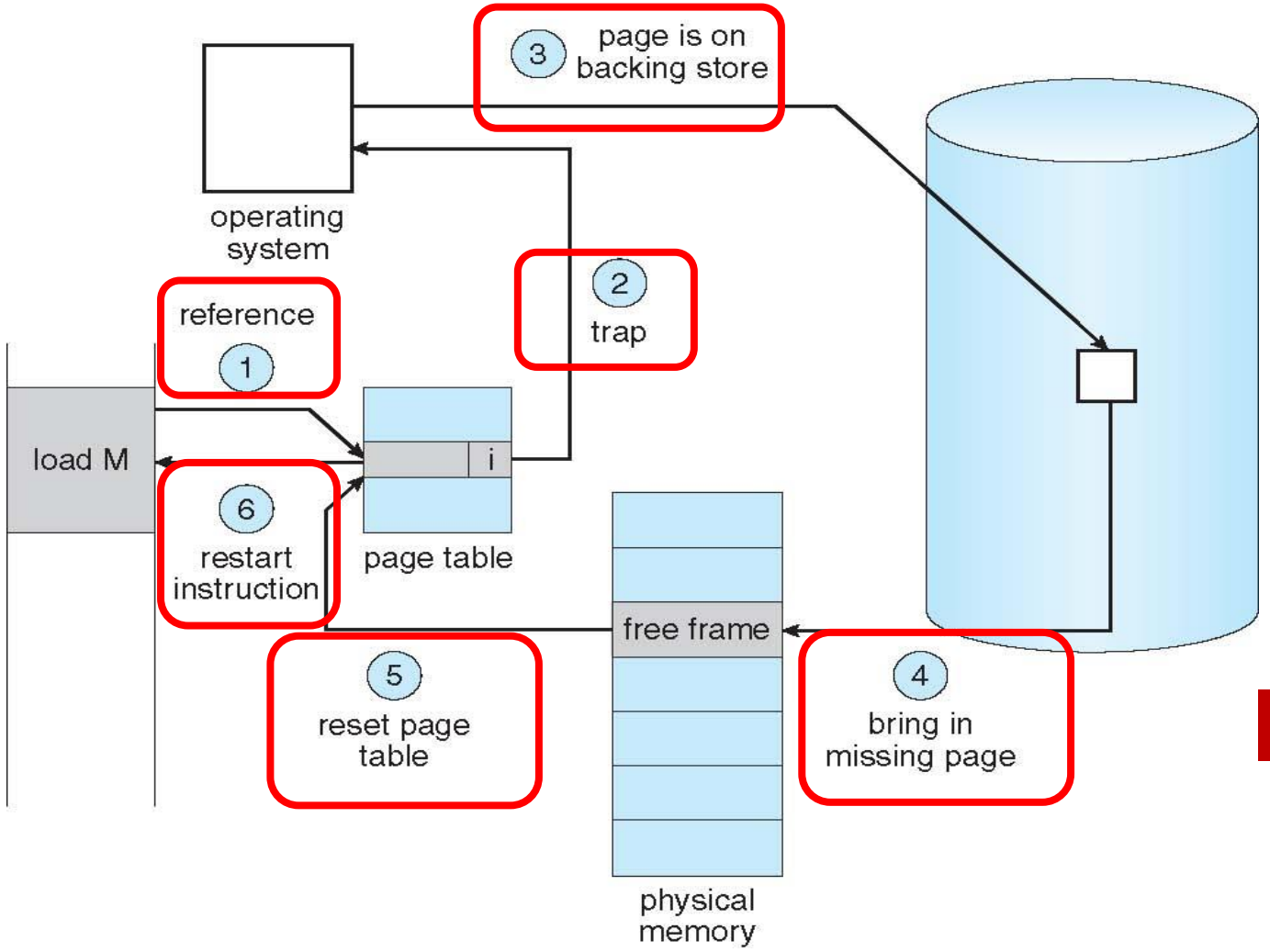
Q2: 访问有效位为i的逻辑页时，会发生什么?

Page Fault （页故障，或称缺页异常、缺页中断）

页故障，指的是当软件试图访问已映射在虚拟地址空间中，但是目前尚未加载在物理内存中的一个页时，由CPU的内存管理的单元发出的一种特殊中断

Q3: 发生页故障时，操作系统需要怎么应对?

按需调页：页故障处理流程



1-进程内存访问触及无效页表项

2-导致页故障

3-惰性加载器开始工作

4-惰性加载器将页面调入物理内存

5-更新相应页表项

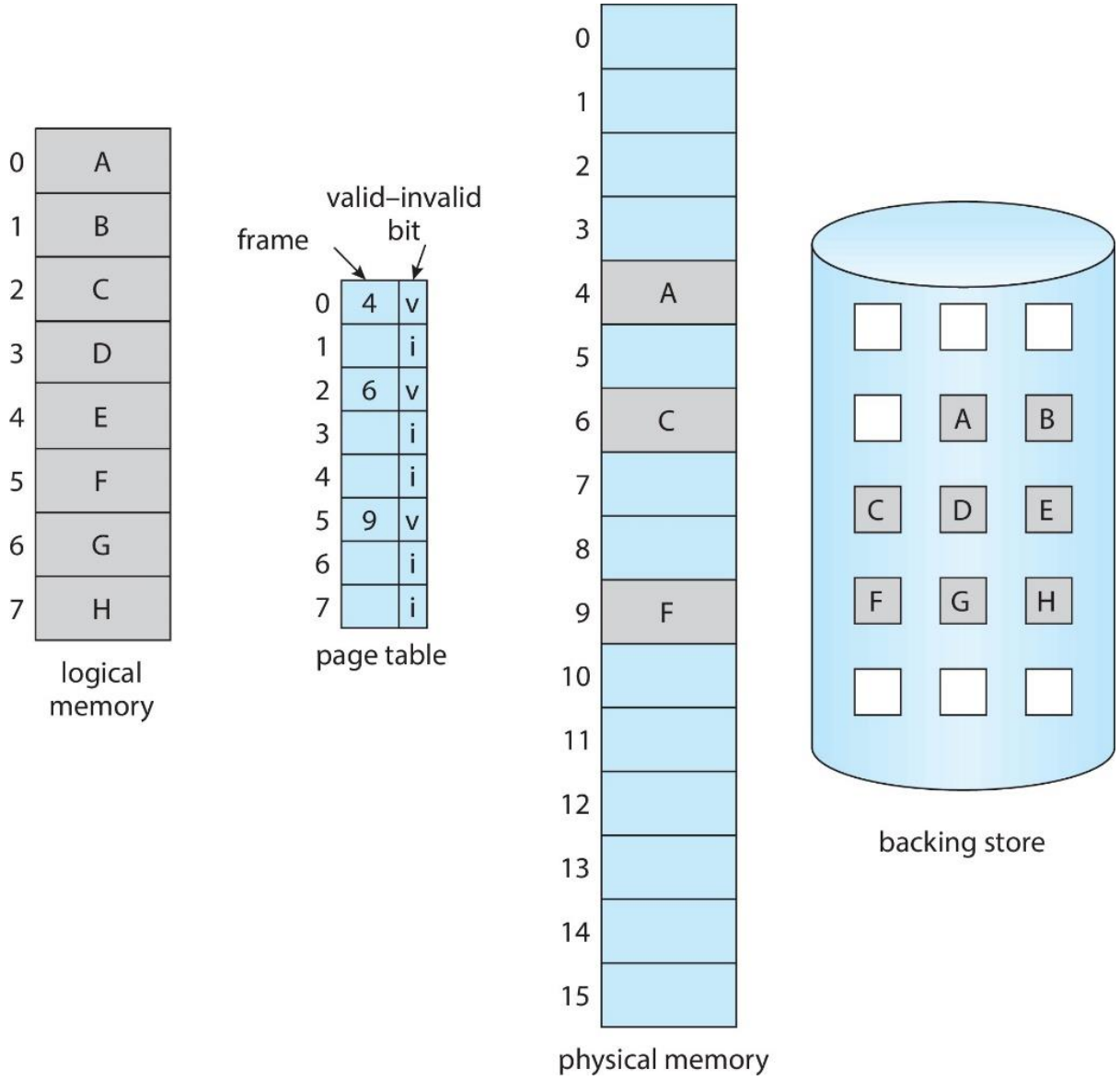
6-重启触发页故障的指令

Q:为何要重启触发页故障的指令?

因为触发页故障的指令未能成功执行，
而经过页故障处理后，具备了执行条件

主观题 4分

设置



如图所示页面状态下，访问逻辑页面3，请简述按需调页的处理流程。

作答

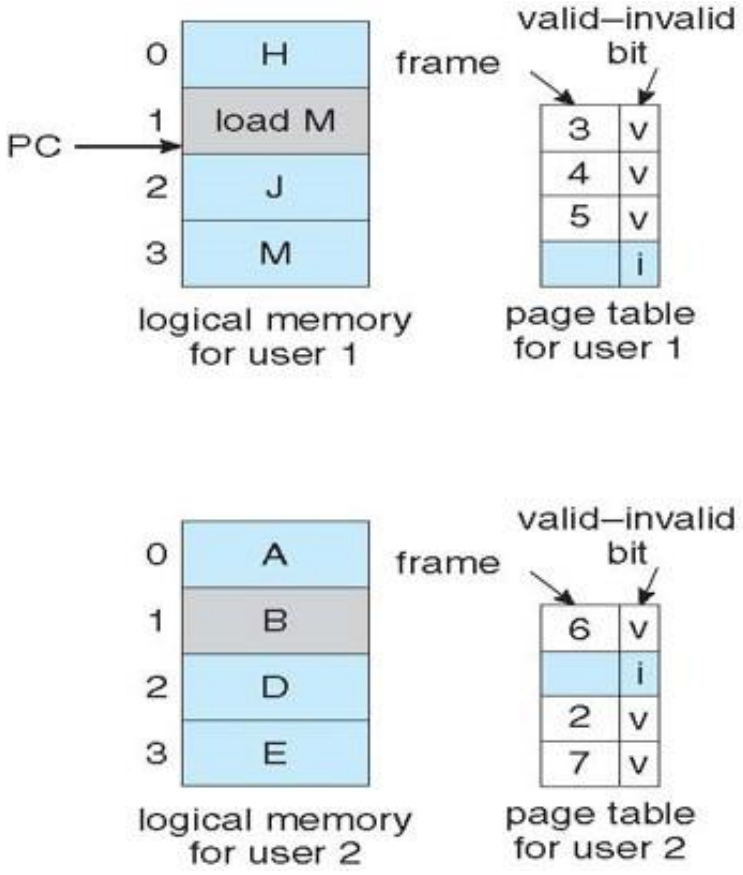
页置换算法

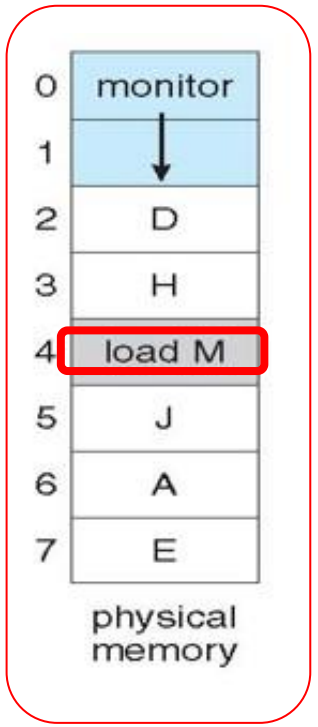
Page Replacement Algorithms

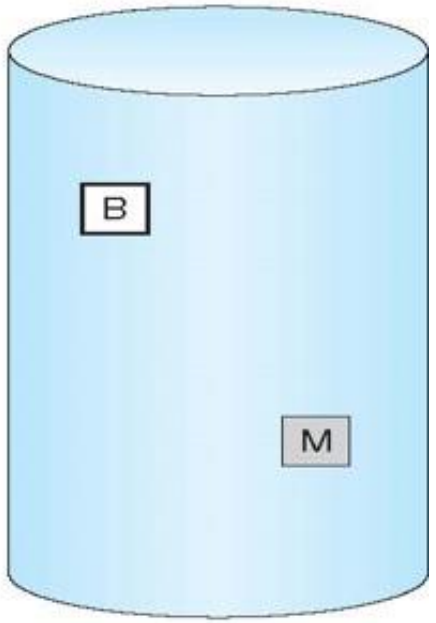
02

简介

- 何时需要进行页置换?







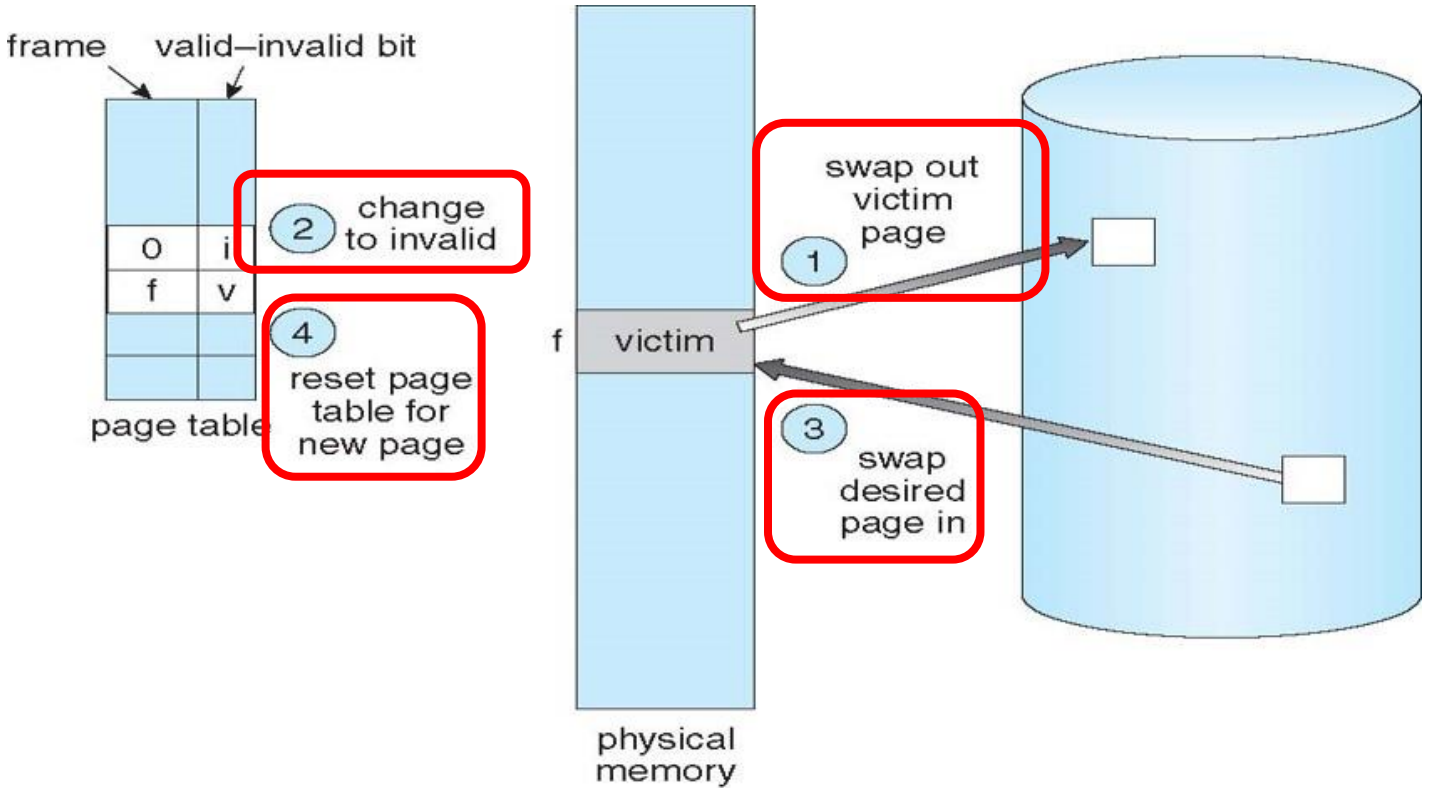
当前状态：物理内存已满

要执行一条指令load M，
需要调入逻辑页M，
但是物理内存满了，怎么办?

需要腾出一个已被占用的物理页

关键问题：牺牲哪个物理页

页置换详细步骤



1. 选择牺牲页帧f，将其中的逻辑页换出
2. 将被换出的被牺牲的逻辑页对应的页表项置为invalid
3. 将需要的页面调入页框f
4. 更新被调入页的页表项

简介

考虑了页置换的按需调页完整流程

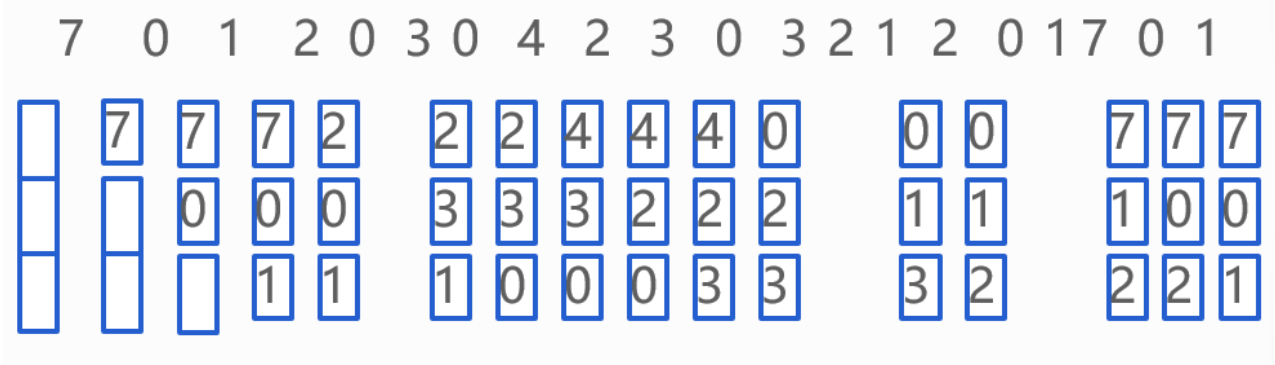
1.从磁盘上获取所需的逻辑页
2.在分配页框时，如果没有空闲页框可用了，则选择一个牺牲页。如果选中的牺牲页内有数据被修改过，那么必须将页框内的数据写回磁盘
3.分配页帧，将需要的页面调入该页帧，并同时更新页表
4.重启引发页故障的指令

典型的置换算法

1. FIFO
2. OPT
3. LRU

(1)FIFO算法

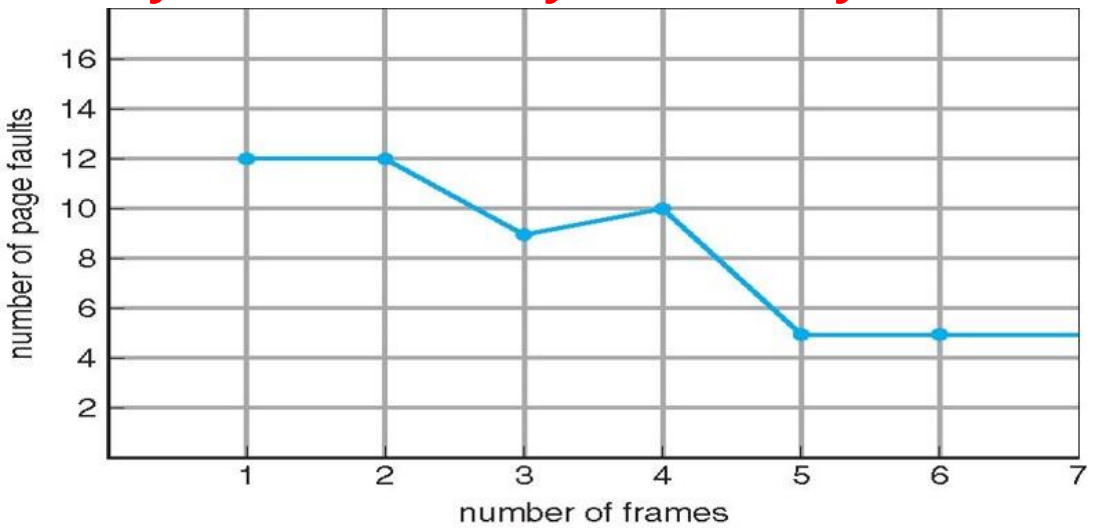
- 每次选择在页帧内停留最久的页面将其换出



- 进程拥有3个物理页框
- 20次内存访问发生了15次页故障，其中包含12次页置换

(1)FIFO算法

- Belady异象（Belady Anomaly）



- Reference String： 1,2,3,4,1,2,5,1,2,3,4,5
- FIFO算法不仅导致页故障率偏高，且不够稳定
  - 该示例中，从3个页帧增加到4个页帧，页故障率反而增加

主观题 2分

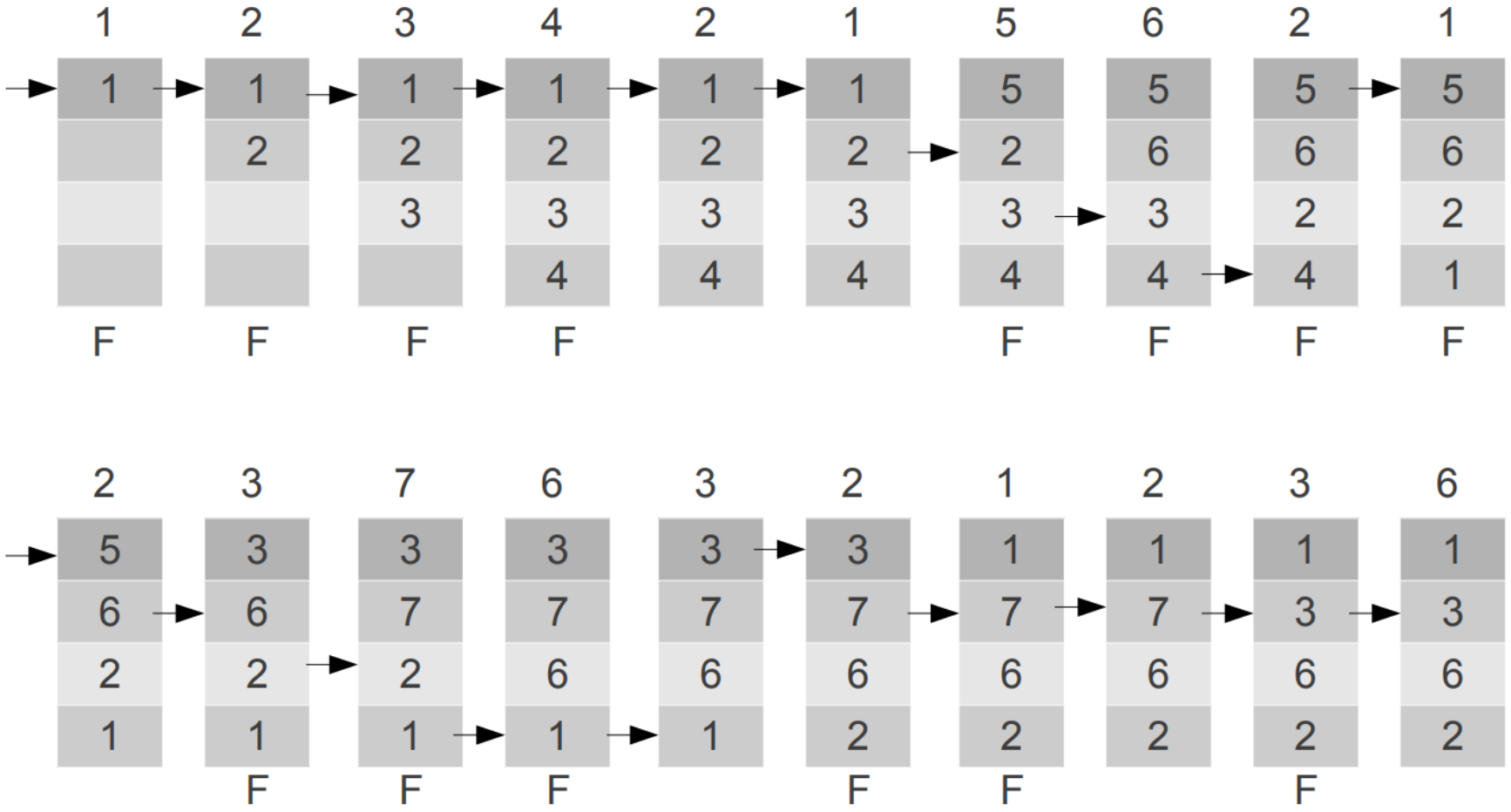
设置

页面引用串：1,2,3,4,2,1,5,6,2,1,2,3,7,6,3,2,1,2,3,6
分配的物理页帧数=4
请给出FIFO置换算法下页面命中情况，并计算页故障次数。

作答

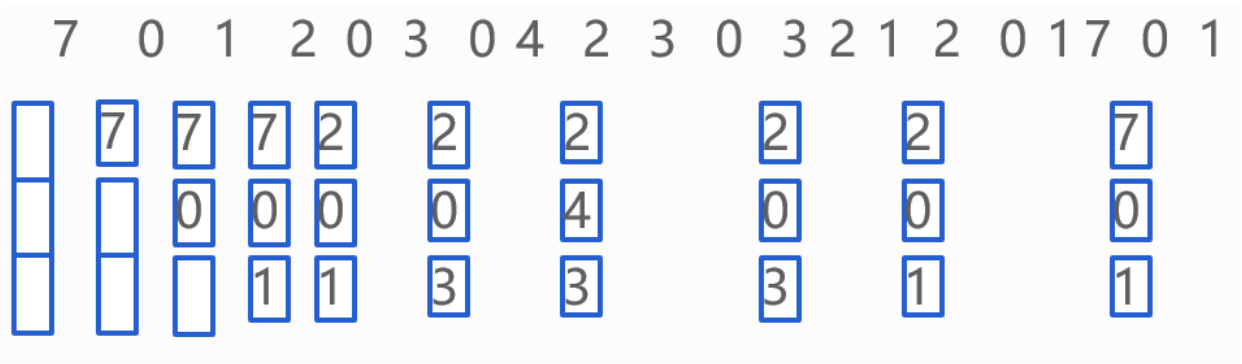
(1)FIFO算法

- 练习答案



(2) OPT算法

- 展望未来，最久未被使用的页被选为牺牲页



- 进程拥有3个页帧
- 发生9次页故障，其中包含6次页置换
- 问题：实际在进程运行时，未来不可精确预知

主观题 2分

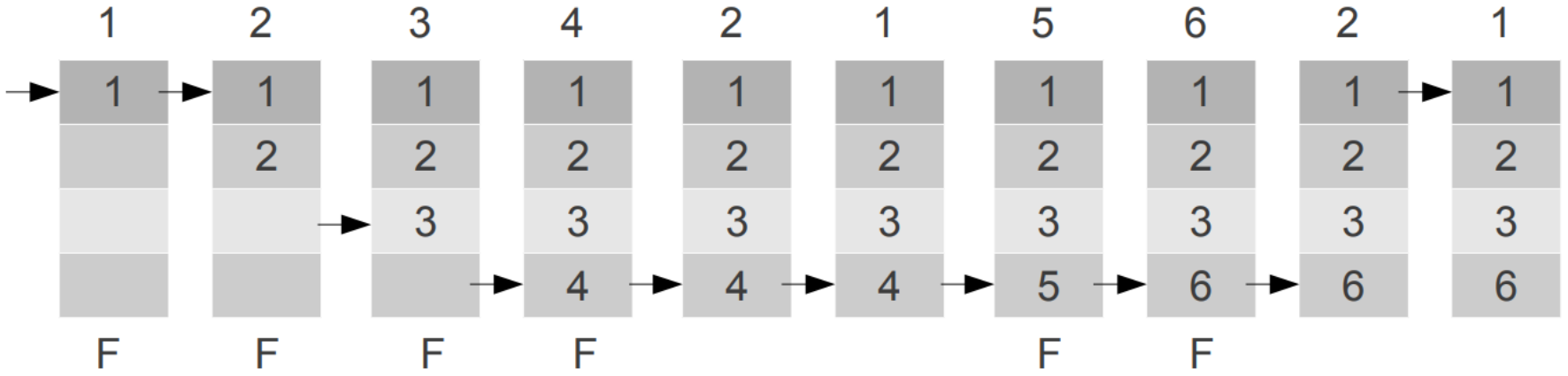
设置

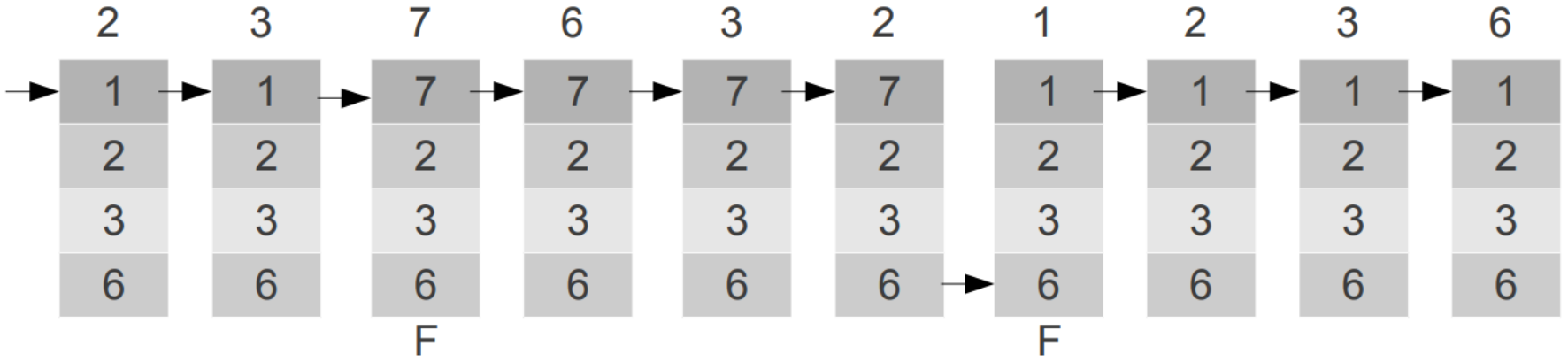
页面引用串：1,2,3,4,2,1,5,6,2,1,2,3,7,6,3,2,1,2,3,6
分配的物理页帧数=4
请给出OPT置换算法下页面命中情况，并计算页故障次数。

作答

(2) OPT算法

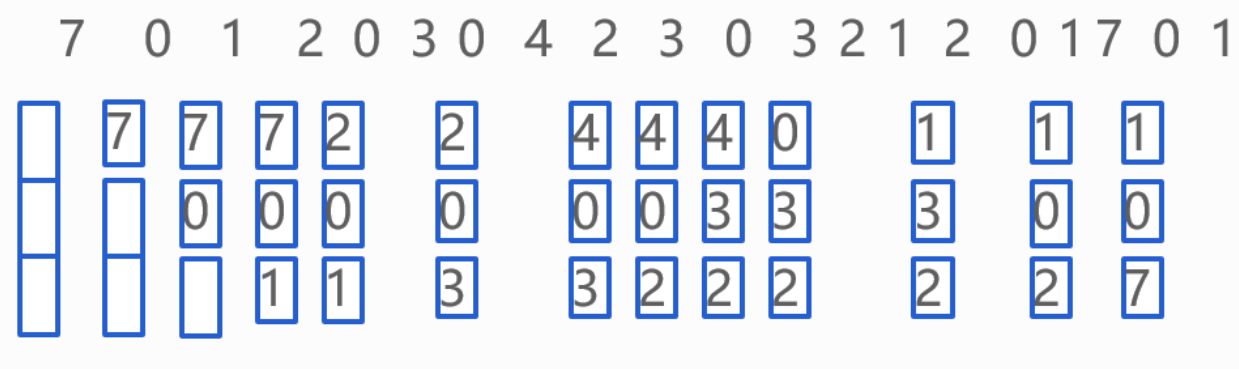
- 练习答案


1 2 3 4 2 1 5 6 2 1
1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2
3 3 3 3 3 3 3 3 3 3
4 4 4 5 6 6 6
F F F F F F


2 3 7 6 3 2 1 2 3 6
1 1 7 7 7 7 1 1 1 1
2 2 2 2 2 2 2 2 2 2
3 3 3 3 3 3 3 3 3 3
6 6 6 6 6 6 6 6 6 6
F F

(3) LRU算法

- 以史为鉴，最近最久未被使用的页被选中置换



  - 进程拥有3个页帧
  - 发生12次页故障，其中包含9次页置换

主观题 2分

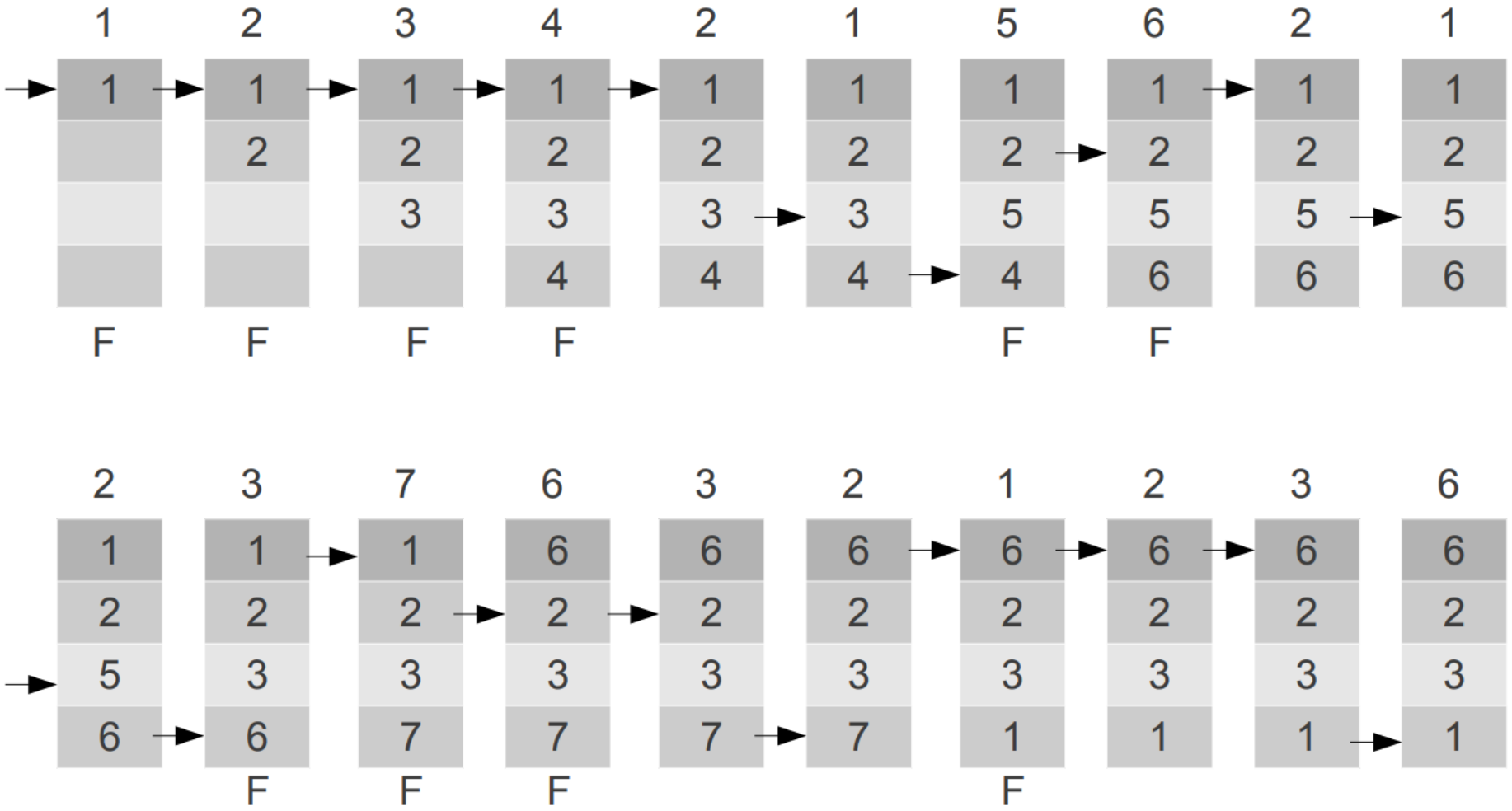
设置

页面引用串：1,2,3,4,2,1,5,6,2,1,2,3,7,6,3,2,1,2,3,6
分配的物理页帧数=4
请给出LRU置换算法下页面命中情况，并计算页故障次数。

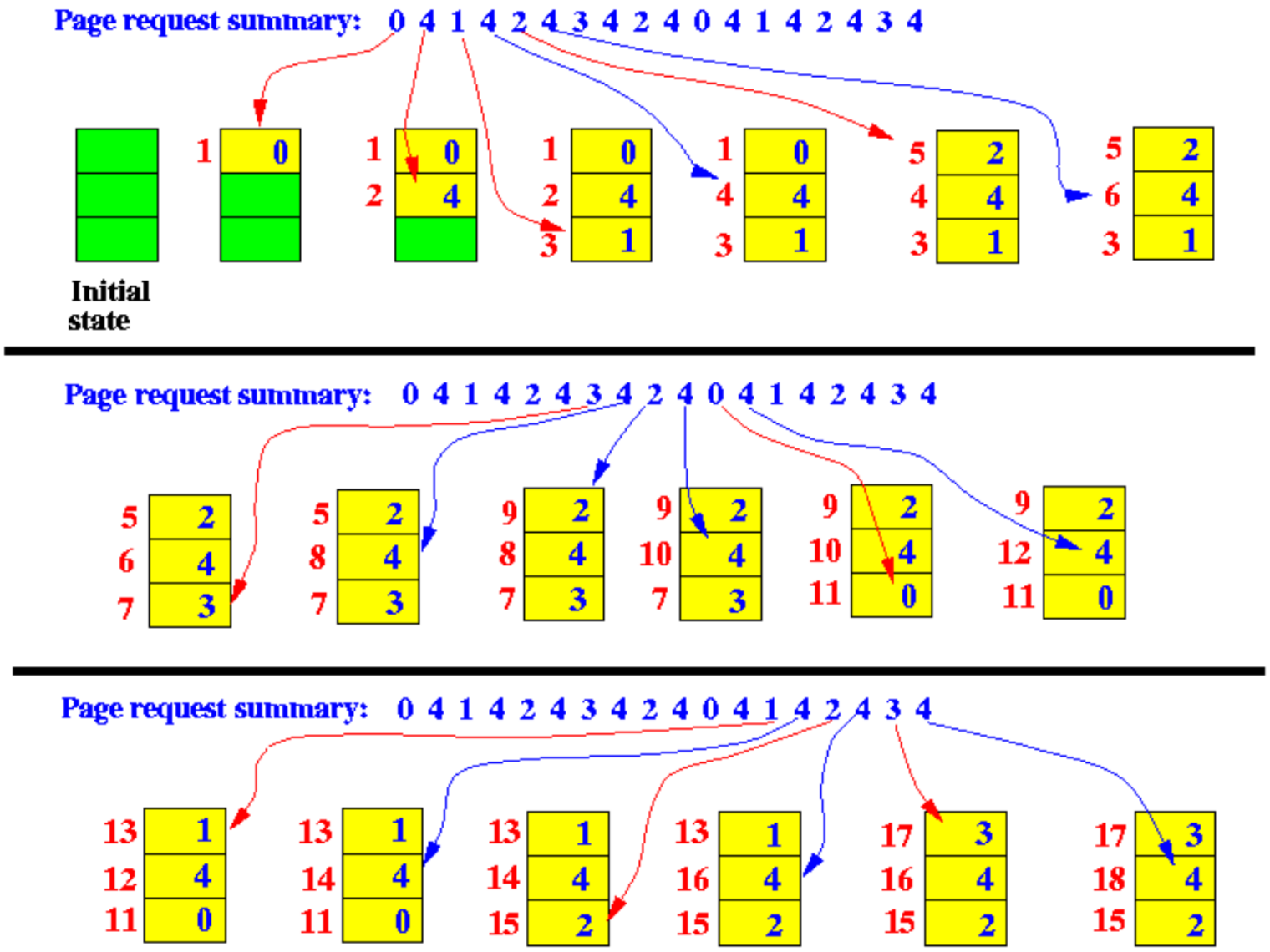
作答

(3) LRU算法

- 练习答案


1 2 3 4 2 1 5 6 2 1
1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2
3 3 3 3 5 5 5 5
4 4 4 4 6 6 6
F F F F F F
2 3 7 6 3 2 1 2 3 6
1 1 1 6 6 6 6 6 6 6
2 2 2 2 2 2 2 2 2 2
5 3 3 3 3 3 3 3 3 3
6 6 7 7 7 7 1 1 1 1
F F F F

(3) LRU算法



- 为每个页帧增加一个时间戳计数
- 每次优先置换时间戳值最小的

Summary

小结：

- 虚存概念
- 页置换算法

练习

习题1

内存保护需要(　)，以保证整个内存空间不被非法访问。

A. 由操作系统和内存硬件协同完成

B. 由操作系统或内存硬件独立完成

C. 由内存硬件独立完成

D. 由操作系统独立完成

练习

习题1

内存保护需要(　)，以保证整个内存空间不被非法访问。

A. 由操作系统和内存硬件协同完成

B. 由操作系统或内存硬件独立完成

C. 由内存硬件独立完成

D. 由操作系统独立完成

练习

习题2

在可变分区内存管理方案中，某一进程执行结束后，系统回收其主存空间并与相邻空闲分区合并，为此需修改空闲分区表，造成空闲分区数减1的情况是该回收分区（ ）。

A. 前后均无邻接空闲分区

B. 前后均有邻接空闲分区

C. 前有邻接空闲分区，但后无邻接空闲分区

D. 前无邻接空闲分区，但后有邻接空闲分区

练习

习题2

在可变分区内存管理方案中，某一进程执行结束后，系统回收其主存空间并与相邻空闲分区合并，为此需修改空闲分区表，造成空闲分区数减1的情况是该回收分区（ ）。

A. 前后均无邻接空闲分区

B. 前后均有邻接空闲分区

C. 前有邻接空闲分区，但后无邻接空闲分区

D. 前无邻接空闲分区，但后有邻接空闲分区

练习

习题3

动态重定位是在程序的（ ）过程中进行的。

A. 链接

B. 装入

C. 执行

D. 编译

练习

习题3

动态重定位是在程序的（ ）过程中进行的。

A. 链接

B. 装入

C. 执行

D. 编译

练习

习题4

下面关于内存管理的叙述，正确的是（ ）。

A. 存储保护的目的是限制内存的分配

B. 在内存大小为M、有N个用户的分时系统中，每个用户占用M/N大小的内存空间

C. 在虚拟内存系统中，只要磁盘空间无限大，进程就能拥有任意大的编址空间

D. 实现虚拟内存管理必须要有相应的硬件支持

练习

习题4

下面关于内存管理的叙述，正确的是（ ）。

A. 存储保护的目的是限制内存的分配

B. 在内存大小为M、有N个用户的分时系统中，每个用户占用M/N大小的内存空间

C. 在虚拟内存系统中，只要磁盘空间无限大，进程就能拥有任意大的编址空间

D. 实现虚拟内存管理必须要有相应的硬件支持

练习

习题5

下面的内存管理方案中，（ ）内存管理方式适宜采用静态重定位。

A. 固定分区

B. 分页

C. 分段

D. 动态重定位分区

练习

习题5

下面的内存管理方案中，（ ）内存管理方式适宜采用静态重定位。

A. 固定分区

B. 分页

C. 分段

D. 动态重定位分区

练习

习题6

连续分区分配和分页管理的主要区别是（ ）。

A. 连续分区分配中的块比分页管理中的页要小

B. 连续分区管理有地址映射而分区管理没有

C. 连续分区管理有存储保护而分区管理没有

D. 连续分区管理要求分配连续的空间，而分页管理没有这种要求

练习

习题6

连续分区分配和分页管理的主要区别是（ ）。

A. 连续分区分配中的块比分页管理中的页要小

B. 连续分区管理有地址映射而分区管理没有

C. 连续分区管理有存储保护而分区管理没有

D. 连续分区管理要求分配连续的空间，而分页管理没有这种要求

练习

习题7

采用两级页表的页式存储管理中，进程执行中CPU首次访问某逻辑地址时，需访问主存的次数是（ ）。

A. 1次

B. 2次

C. 3次

D. 4次

练习

习题7

采用两级页表的页式存储管理中，进程执行中CPU首次访问某逻辑地址时，需访问主存的次数是（ ）。

A. 1次

B. 2次

C. 3次

D. 4次

练习

习题8

段页式存储管理汲取了页式管理和段式管理的长处，其实现原理结合了页式和段式管理的基本思想，即（ ）。

A. 用分段方法来分配和管理物理存储空间，用分页方法来管理用户地址空间

B. 用分段方法来分配和管理用户地址空间，用分页方法来管理物理存储空间

C. 用分段方法来分配和管理主存空间，用分页方法来管理辅存空间

D. 用分段方法来分配和管理辅存空间，用分页方法来管理主存空间

练习

习题8

段页式存储管理汲取了页式管理和段式管理的长处，其实现原理结合了页式和段式管理的基本思想，即（ ）。

A. 用分段方法来分配和管理物理存储空间，用分页方法来管理用户地址空间

B. 用分段方法来分配和管理用户地址空间，用分页方法来管理物理存储空间

C. 用分段方法来分配和管理主存空间，用分页方法来管理辅存空间

D. 用分段方法来分配和管理辅存空间，用分页方法来管理主存空间

按需调页系统的设计考虑

预调页的设计。

谢 谢!
Thank you!