Page Replacement

Computer Operating Systems, Spring 2025

Instructors: Joel Ramirez Travis McGaha

Head TAs: Ash Fujiyama Emily Shen Maya Huizar

TAs:

Ahmed Abdellah	Bo Sun	Joy Liu	Susan Zhang	Zihao Zhou
Akash Kaukuntla	Connor Cummings	Khush Gupta	Vedansh Goenka	
Alexander Cho	Eric Zou	Kyrie Dowling	Vivi Li	
Alicia Sun	Haoyun Qin	Rafael Sakamoto	Yousef AlRabiah	
August Fu	Jonathan Hong	Sarah Zhang	Yu Cao	

pollev.com/cis5480

- What's your coffee order?

Administrivia

- PennOS Milestone 1 is due this week Apr 25
  - You have the first milestone, which should be done by tomorrow
  - Everyone should already have/had a meetup scheduled with your TAs.
  - Have a plan (a REAL plan) for how to complete the rest
  - Full Thing due ~Apr 25
    - You can technically turn it in late, but by then you're really just grasping at straws to finish it up…

Administrivia

- Check-in Out Tomorrow: due at end of Friday
  - Another one will be released this week, due sometime next week

Lecture Outline

- **Page Replacement: High Level**
  - **FIFO**
  - **Reference Strings**
  - **Beladys**
- LRU
- Thrashing
- FIFO w/ Reference bit

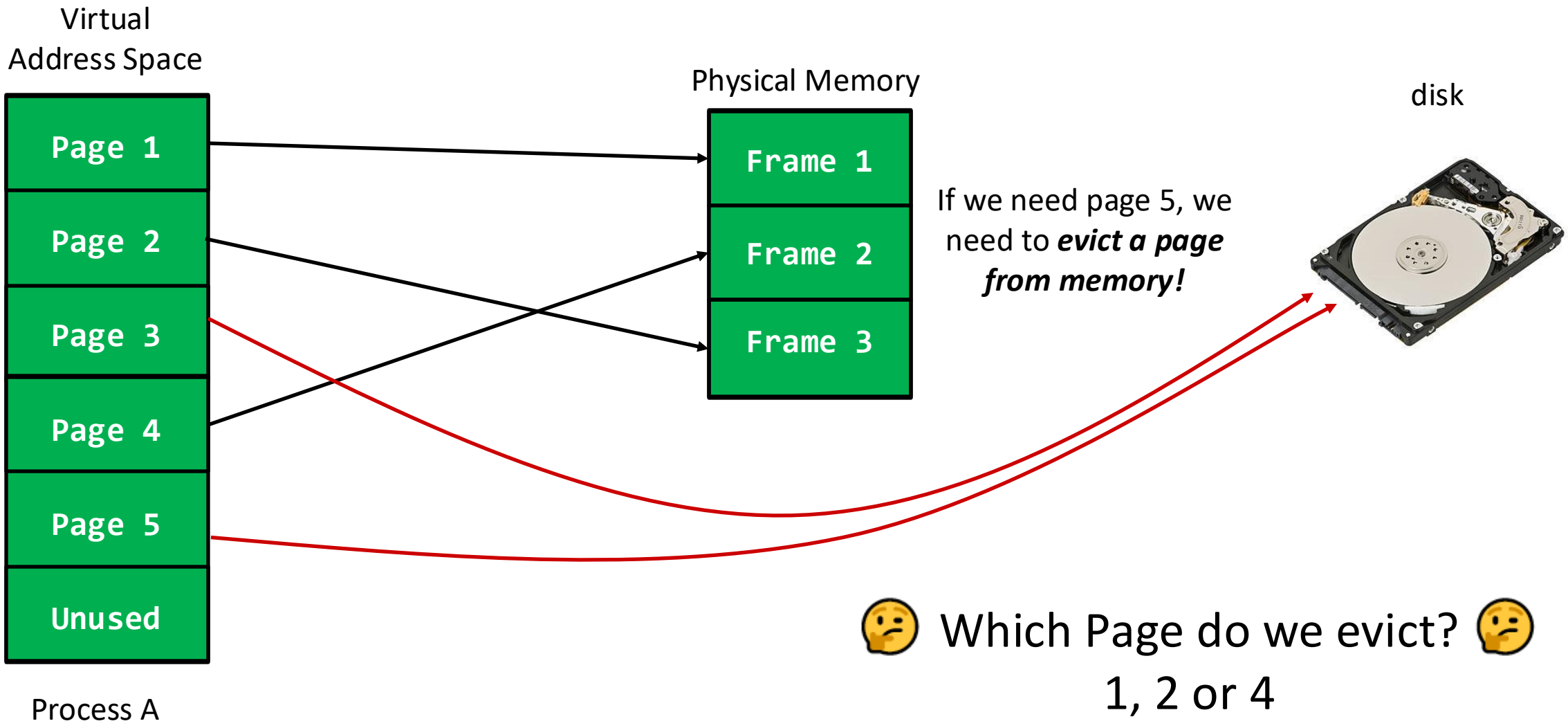
Page Replacement

- The operating system will sometimes have to evict a page from physical memory to make room for another page.

- If the evicted page is access again in the future, it will cause a page fault, and the Operating System will have to go to Disk to load the page into memory again

Evicting a Page

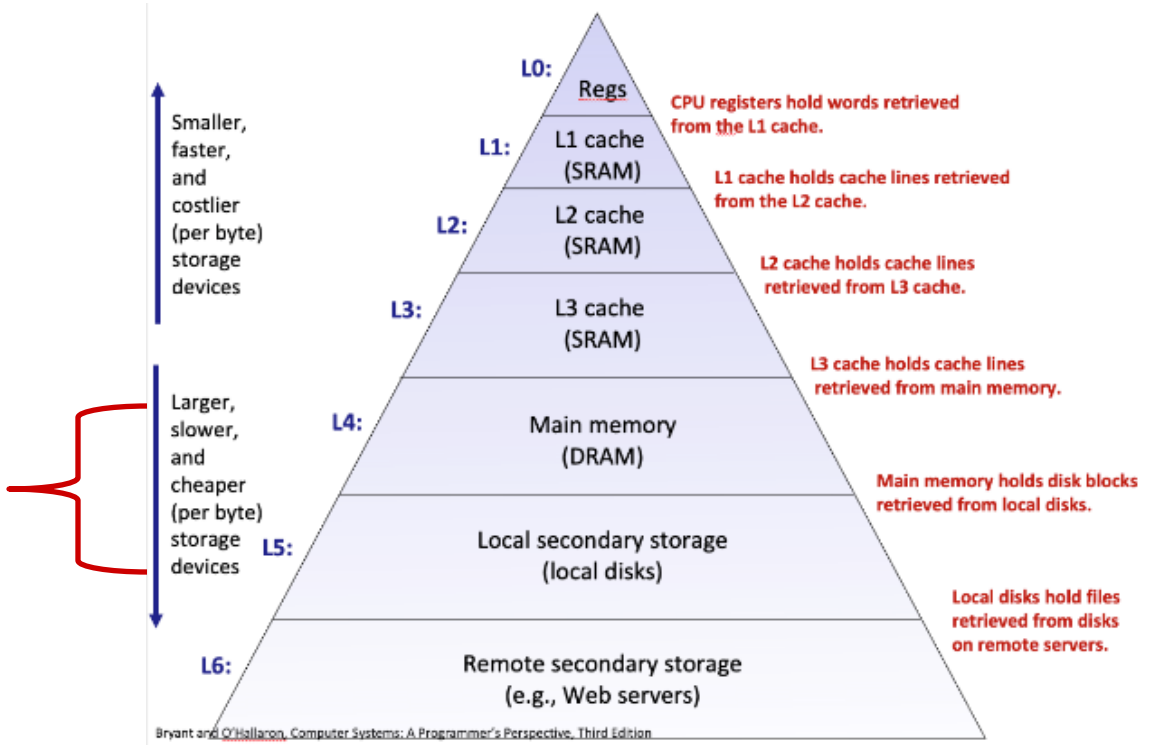
- Physical Memory is limited in size. Not all pages can exist in Memory.



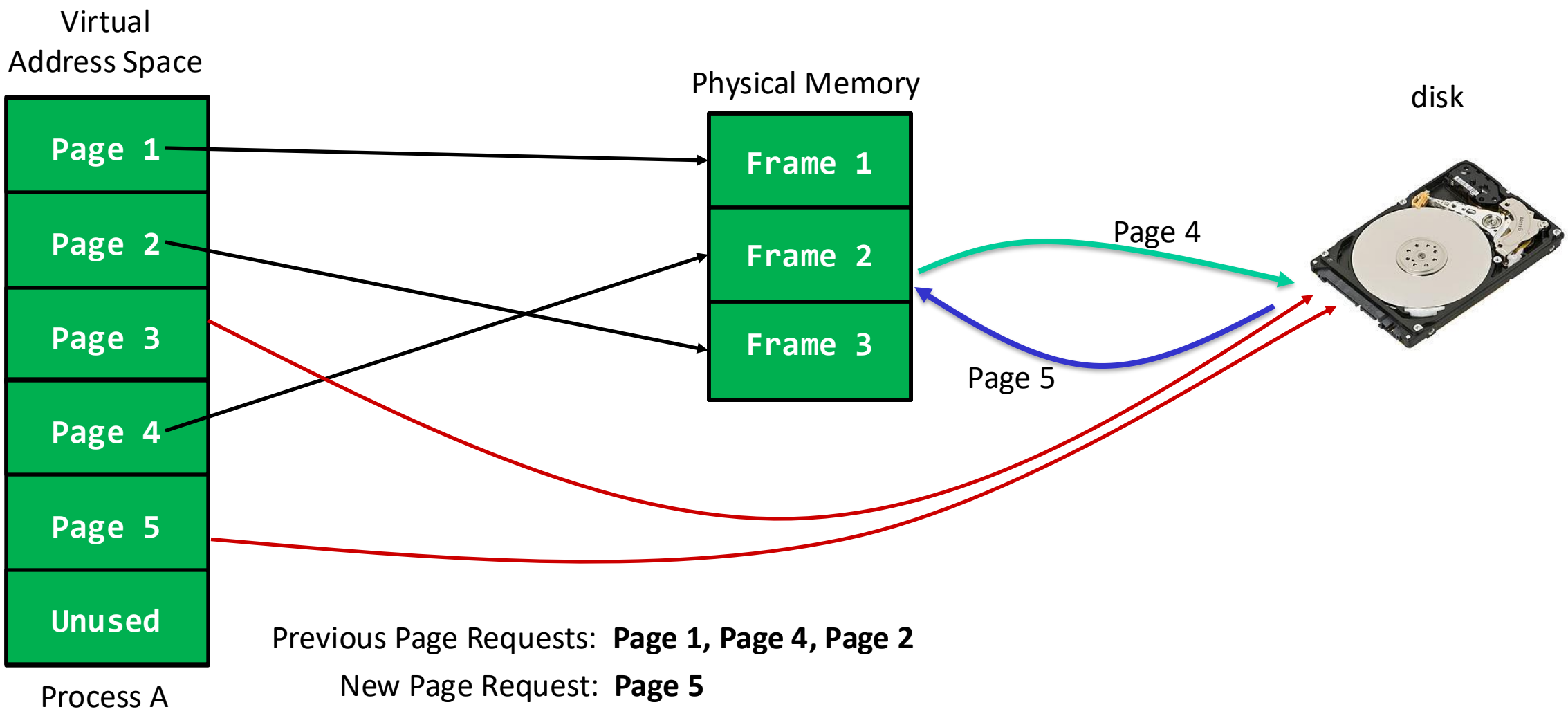
Page Replacement

- Remember this? Disk access is very very slow (relatively speaking).
  - How can we ***minimize disk accesses***?
  - How can we try to ensure the page we evict from memory is unlikely to be
  - used again in the future?

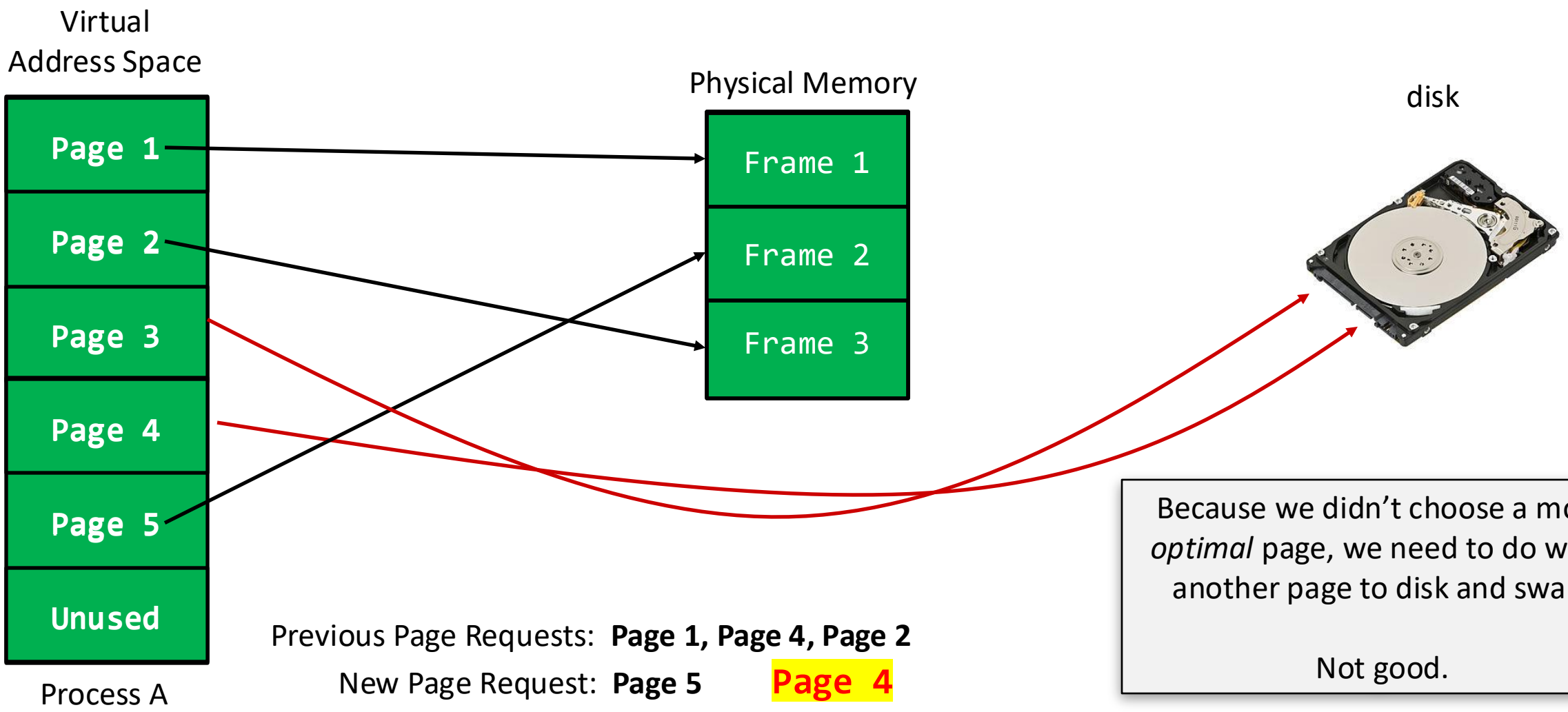
We are at this Level:



Page Replacement Scenario



Page Replacement Scenario



Previous Page Requests: **Page 1, Page 4, Page 2**

New Page Request: **Page 5** **Page 4**

We just evicted page four!!

Because we didn't choose a more *optimal* page, we need to do write another page to disk and swap.

Not good.

Reference String

- A reference string is a string representing a sequence of virtual page accesses. By a given process on some input.
  - E.g., `0 1 2 3 4 1 2 9 5 3 2 2 …`
  - Page 0 is accessed, then 1, then 2, then 3 …

- Having the page access history, we can now optimize for which page to select when evicting!

Review: Fault vs Eviction

- Page Fault
  - When a corresponding page is *not* in memory, we need to load the page from memory.
- Eviction
  - When there are too many pages in memory, and we need to evict one to make space for another.
  - Just because there is a page fault does not mean there is an eviction.

FIFO Replacement

- One way to decide which pages can be evicted is to use FIFO (First in First Out)
- If a page needs to be evicted from physical memory, then the page that has been in memory the longest can be evicted.

```
#define MAX_PAGES 4

typedef struct page_stack {
   short *stack; // array
   short size;
} page_stack;
```

```
short evict(page_stack *ps){
   // There's none to evict,
   //  -1 page doesn't exist
   if(ps->size == 0) return -1;

   ps->size--;
   short page_num = (ps->stack)[0];
   memmove(ps->stack,
                        ps->stack + 1, ps->size);
   return page_num;
}
```

*simplified code, doesn't work for all edge cases.

FIFO Replacement

- One way to decide which pages can be evicted is to use FIFO (First in First Out)
- If a page needs to be evicted from physical memory, then the page that has been in memory the longest can be evicted.

```
void add(page_stack *ps, short page_num){
      for(int i = 0; i < ps->size; i++){
            if(ps->stack[i] == page_num){
               return; // In Mem.
                              // No page fault
         }
   }
      if(ps->size == MAX_PAGES) evict(ps);
      //page fault!
      (ps->stack)[ps->size] = page_num;
      ps->size++; //increment size
}
```

```
short evict(page_stack *ps){
   // There's none to evict,
   //  -1 page doesn't exist
   if(ps->size == 0) return -1;

   ps->size--;
   short page_num = (ps->stack)[0];
   memmove(ps->stack,
                              ps->stack + 1, ps->size);
   return page_num;
}
```

*simplified code, doesn't work for all edge cases.

FIFO Replacement

- If we have 4 frames, and the reference string:
  4 1 1 2 3 4 **5**
  - Red numbers indicate that accessing the page caused a page fault. Accessing 5 also causes 4 to be evicted from physical memory

```
short page_str[] = {4, 1, 1, 2, 3, 4, 5};
page_stack ps = {0}; // No pages in memory at start.
ps.stack = (short *)malloc(MAX_PAGES * sizeof(short));

for(short x: page_str){ // for x in page_str
        add(&ps, x);
}
```

*simplified code, doesn't work for all edge cases.

FIFO Replacement

- If we have 4 frames, and the reference string:
  4 1 1 2 3 4 **5**
  - Red numbers indicate that accessing the page caused a page fault. Accessing 5 also causes 4 to be evicted from physical memory
- For those who like tables ☺

	Ref str:	4	1	1	2	3	4	5
Newest		4	1	1	2	3	3	5
			4	4	1	2	2	3
					4	1	1	2
Oldest						4	4	1

Poll Everywhere

pollev.com/cis5480

- Given the following reference string, how many page faults (not evictions) occur when using a FIFO algorithm given no pages are in memory at the start.
- 1 2 3 4 1 2 5 1 2 3 4 5

```
#define MAX_PAGES 3

typedef struct page_stack {
   short *stack; // init empty
   short size;
} page_stack;
```

- Part 2: If we didn't have to follow a strict policy, what is the "optimal" # of pages that could be evicted to minimize faults? How many less faults would we have?

Poll Everywhere

pollev.com/cis5480

	Ref str:	1	2	3	4	1	2	5	1	2	3	4	5
Newest													
Oldest													
Evicted													

Poll Everywhere

pollev.com/cis5480

- Given the following reference string, how many page faults occur when using a FIFO algorithm
- 1 2 3 4 1 2 5 1 2 3 4 5
- FIFO

	Ref str:	1	2	3	4	1	2	5	1	2	3	4	5
Newest		**1**	2	3	4	1	2	5	5	5	3	4	4
			1	2	3	4	1	2	2	2	5	3	3
Oldest				1	2	3	4	1	1	1	2	5	5
Evicted					**1**	2	3	4			1	2	

- 9 faults

Poll Everywhere

pollev.com/cis5480

- Given the following reference string, how many page faults occur when using a FIFO algorithm
- 1 2 3 4 1 2 5 1 2 3 4 5
- Theoretical optimal?

	Ref str:	1	2	3	4	1	2	5	1	2	3	4	5
		1	2	3	4	4	4	5	5	5	3	4	4
			1	2	2	2	2	2	2	2	5	3	3
				1	1	1	1	1	1	1	2	5	5
Evicted					**3**			4			1	2	

- 7 faults

"optimal" replacement

- If you knew the exact sequence of page accesses in advance, you could optimize for smallest number of page faults

- Always replace the page that is furthest away from being used again in the future
  - How do we predict the future??????
  - You can't, but you can make a "best guess" (later in lecture)

- Optimal replacement is still a handy metric. Used for testing replacement algorithms, see how an algorithm compares to various "optimal" possibilities.

Poll Everywhere

pollev.com/cis5480

- Given the following reference string, how many page faults occur when using a FIFO algorithm:
- 3 2 1 0 3 2 4 3 2 1 0 4

```
#define MAX_PAGES 3

typedef struct page_stack {
   short *stack; // init empty
   short size;
} page_stack;
```

- Part 2: What if we had **4** page frames, how many faults would we have?

Poll Everywhere

pollev.com/cis5480

	Ref str:	3	2	1	0	3	2	4	3	2	1	0	4
Newest													
Oldest													
Evicted													

Poll Everywhere

pollev.com/cis5480

	Ref str:	3	2	1	0	3	2	4	3	2	1	0	4
Newest													
Oldest													
Evicted													

Poll Everywhere

pollev.com/cis5480

- Given the following reference string, how many page faults occur when using a FIFO algorithm
- 3 2 1 0 3 2 4 3 2 1 0 4

- Three page frames

	Ref str:	3	2	1	0	3	2	4	3	2	1	0	4
Newest		3	2	1	0	3	2	4	4	4	1	0	0
			3	2	1	0	3	2	2	2	4	1	1
Oldest				3	2	1	0	3	3	3	2	4	4
Victim					3	2	1	0			3	2	

- 9 faults

Poll Everywhere

pollev.com/cis5480

- Given the following reference string, how many page faults occur when using a FIFO algorithm
- 3 2 1 0 3 2 4 3 2 1 0 4
- Four page frames

	Ref str:	3	2	1	0	3	2	4	3	2	1	0	4
Newest		3	2	1	0	0	0	4	3	2	1	0	4
			3	2	1	1	1	0	4	3	2	1	0
				3	2	2	2	1	0	4	3	2	1
Oldest					3	3	3	2	1	0	4	3	2
Victim								3	2	1	0	4	3

- 10 faults

Bélády's anomaly

- Sometimes increasing the number of page frames in the data structure results in an increase in the number of page faults ☹
- This behavior is something that we want to avoid/minimize the possibility of.
- Some algorithms avoid this *anomaly (LRU, LIFO, etc.)*

Lecture Outline

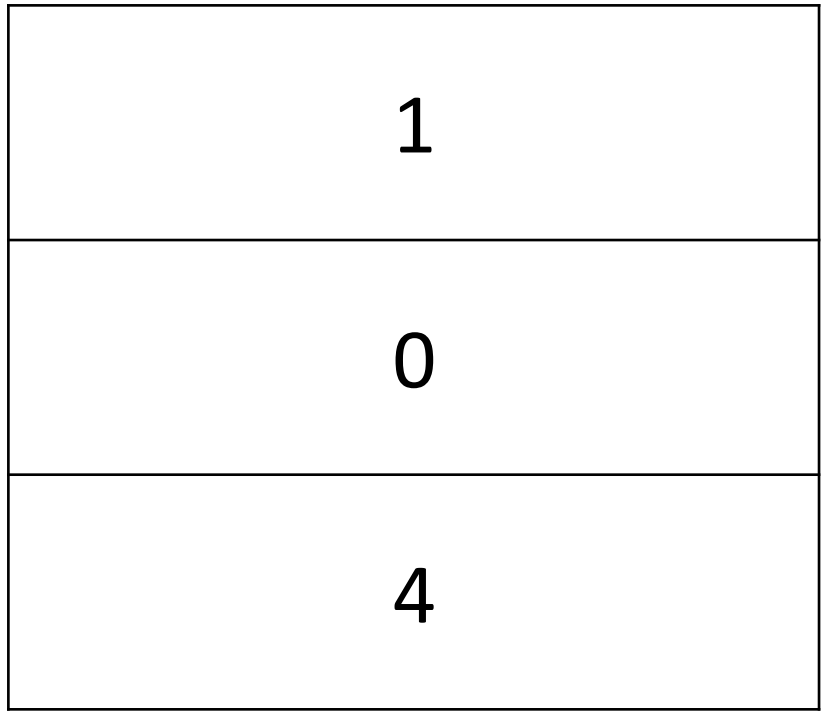
- Page Replacement: High Level
  - FIFO
  - Reference Strings
  - Beladys
- **LRU**
- Thrashing
- FIFO w/ Reference bit

LRU (Least Recently Used)

- Assumption:
  - If a page is used recently, it is likely to be used again in the future
- Use prior knowledge to predict the future (update posterior)
- Replace the page that has had the longest time since it was last used
- Sorta Reminiscent of a Priority Queue, where smaller time since last access indicates lower priority of eviction. (But too complicated)

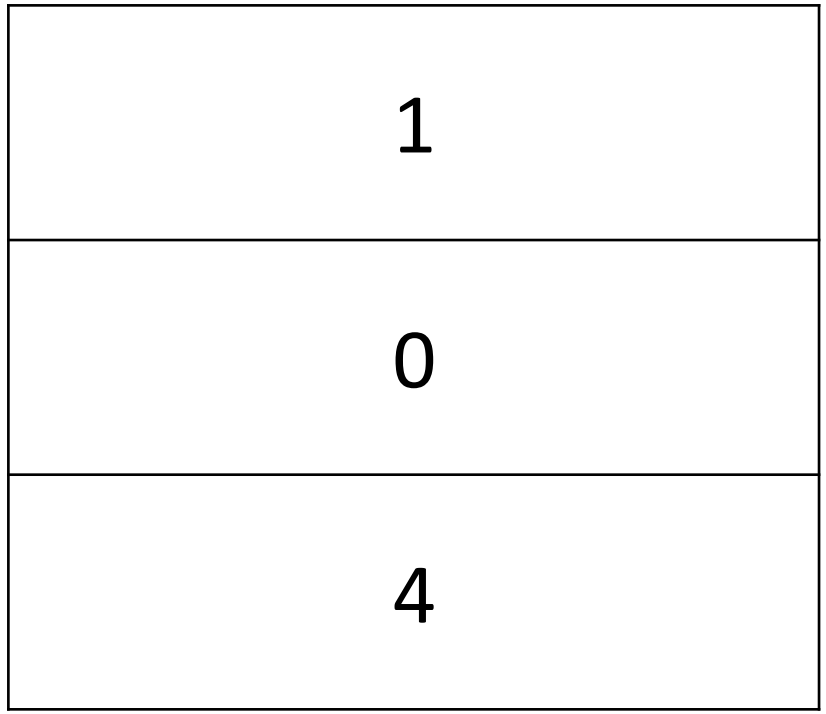
Small Example: 3 Pages of Space

4, 0, 1, 2, 0, 3, 0, 4, 2, 3, 0, 3



Small Example: 3 Pages of Space

4, 0, 1, 2, 0, 3, 0, 4, 2, 3, 0, 3

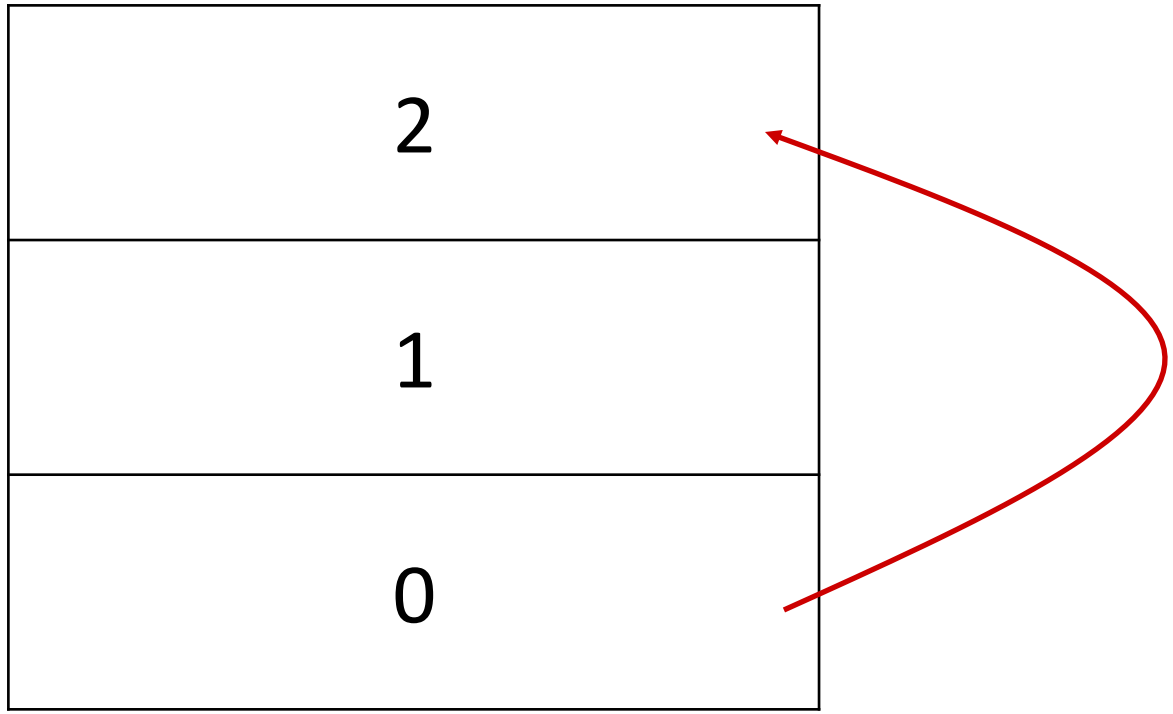


To Evict

To make space for Page 2 – we need to evict page 4.

Small Example: 3 Pages of Space

4, 0, 1, 2, 0, 3, 0, 4, 2, 3, 0, 3

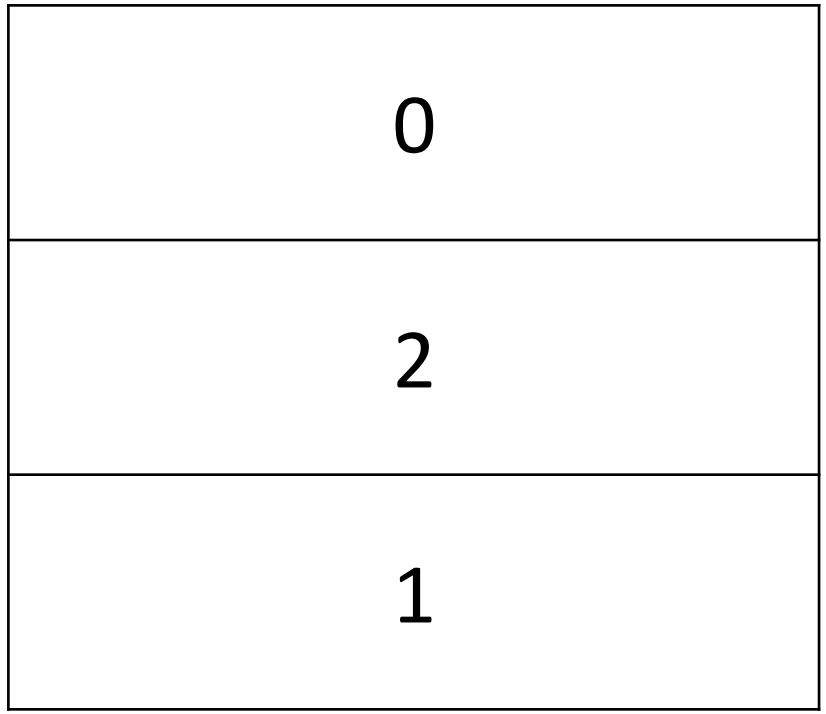


To Evict

As we access 0 again, we move it to the top of the

Small Example: 3 Pages of Space

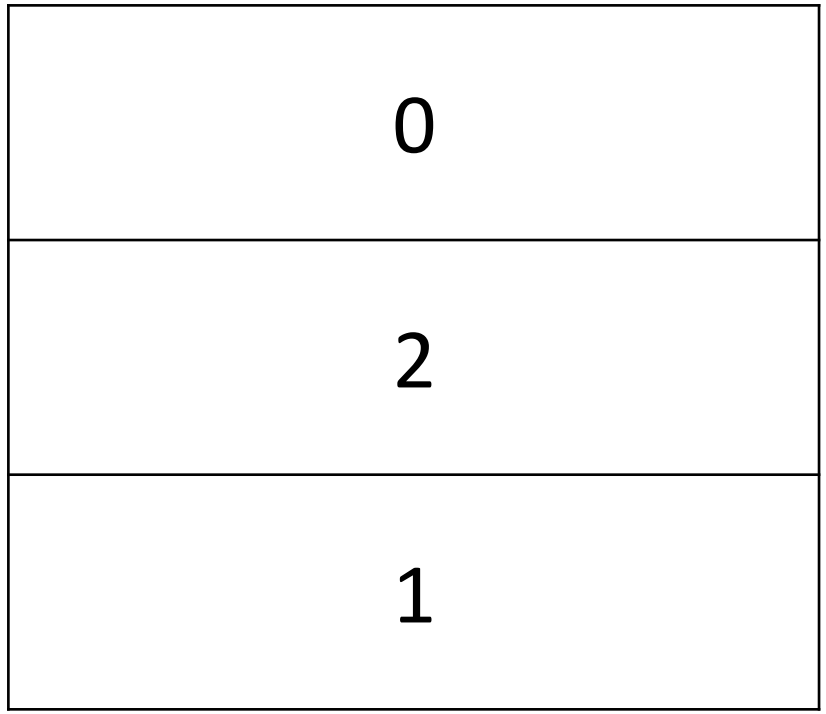
4, 0, 1, 2, 0, 3, 0, 4, 2, 3, 0, 3



Observation:
The ‘order’ of the values when using the LRU is always a subsequence of page accesses.

Small Example: 3 Pages of Space

4, 0, 1, 2, 0, 3, 0, 4, 2, 3, 0, 3



Observation:
The ‘order’ of the values when using the LRU is always a subsequence of page accesses.

Lots of ways to reason about this…try to find a way that makes more sense to you.

Poll Everywhere

pollev.com/cis5480

- Now, using the same Reference String with LRU, let's try to fill this table out…
- What if there are four frames instead of 3? How Many ***Page Faults***?

LRU	Ref str:	4	0	1	2	0	3	0	4	2	3	0	3
Recent													
To Evict													
Evicted													

Poll Everywhere

pollev.com/cis5480

- What if there are four frames instead of 3? How Many Page Faults?
  - 6 Faults
  - 2 Evictions

LRU	Ref str:	4	0	1	2	0	3	0	4	2	3	0	3
Recent		4	0	1	2	0	3	0	4	2	3	0	3
			4	0	1	2	0	3	0	4	2	3	0
				4	0	1	2	2	3	0	4	2	2
To Evict					4	4	1	1	2	3	0	4	4
Evicted							4		1				

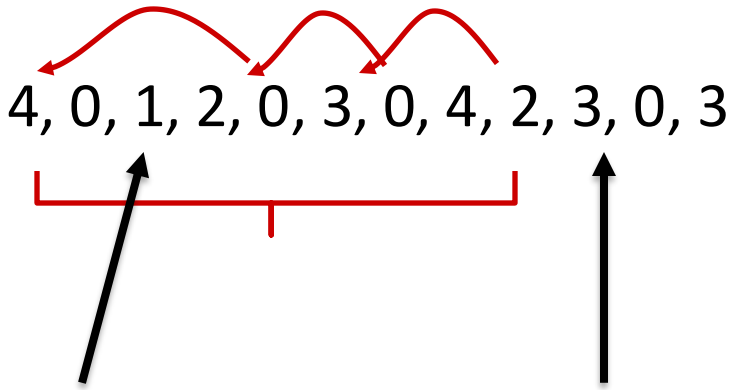
Poll Everywhere

pollev.com/cis5480

- What if there are four frames instead of 3? How Many Page Faults?

- Easier for me to think about this in terms of subsequences

4 Page Faults as we start off empty.



Not included in sequence here as it is already in the 'stack' at it's start.

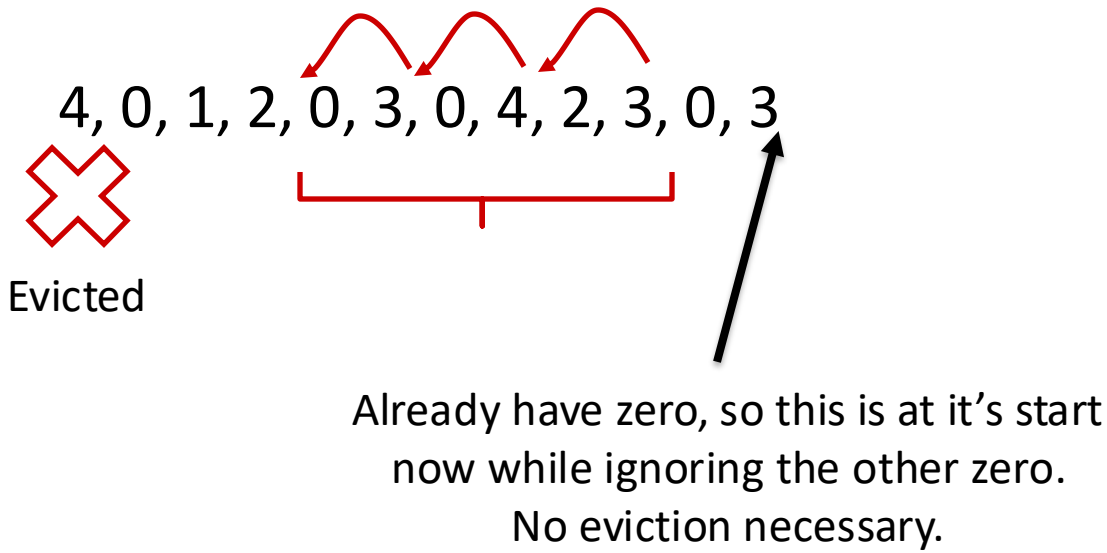
The newest page to add; we just remove the last value in the sequence. (Our first eviction!)

Poll Everywhere

pollev.com/cis5480

- What if there are four frames instead of 3? How Many Page Faults?

- Easier for me to think about this in terms of subsequences

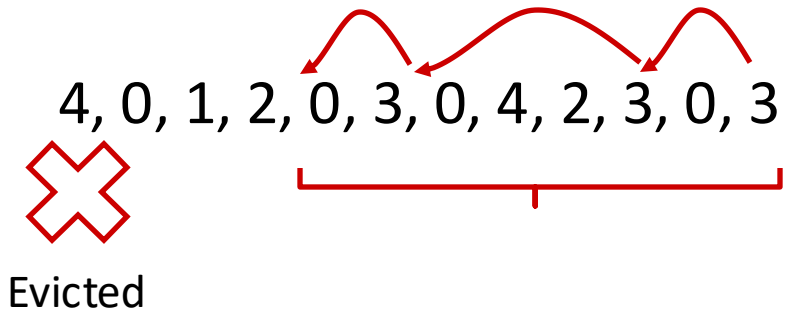


Poll Everywhere

pollev.com/cis5480

- What if there are four frames instead of 3? How Many Page Faults?

- Easier for me to think about this in terms of subsequences



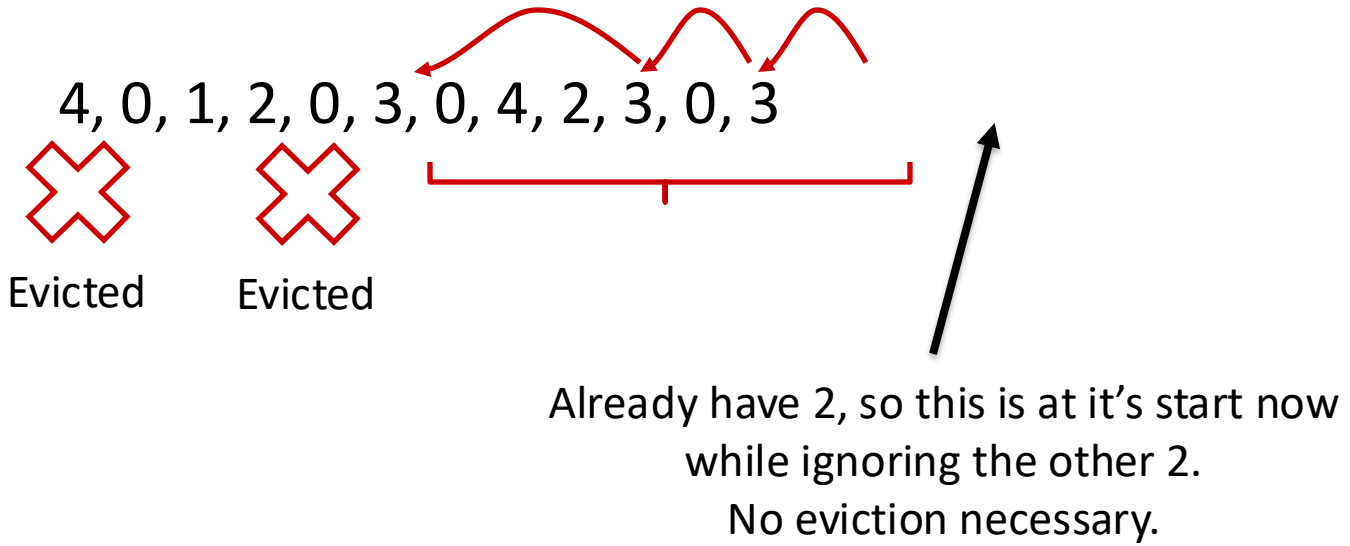
The newest page to add; we just remove the last value in the sequence. (Our Second eviction!)

Poll Everywhere

pollev.com/cis5480

- What if there are four frames instead of 3? How Many Page Faults?

- Easier for me to think about this in terms of subsequences

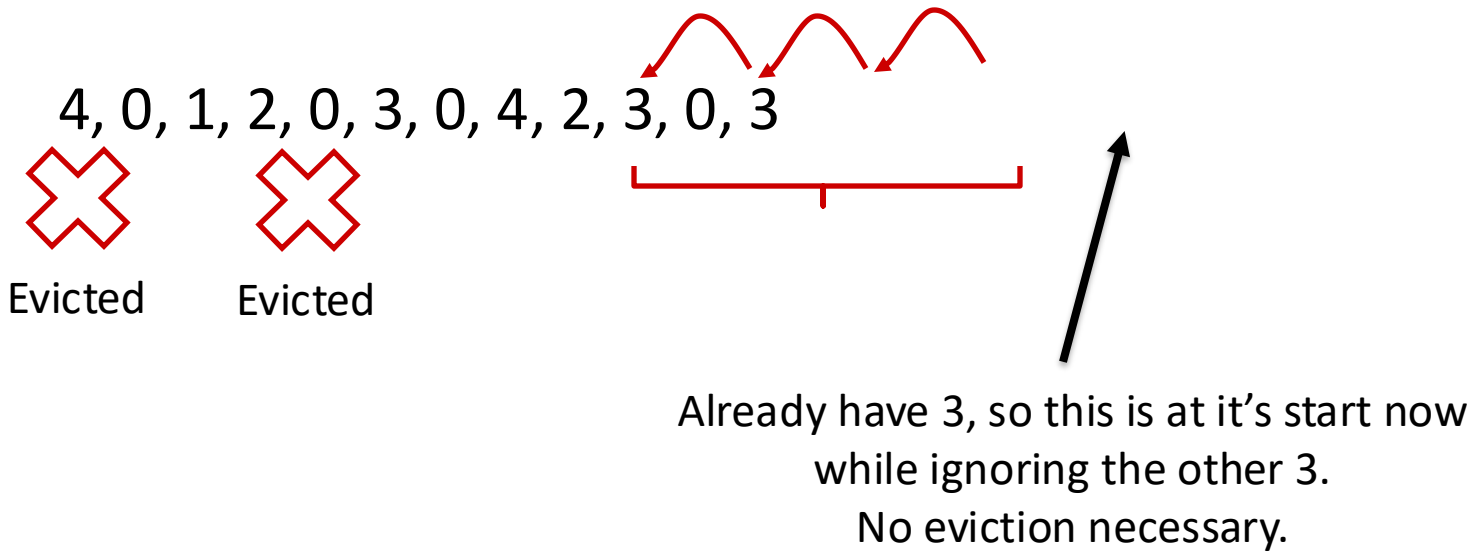


Poll Everywhere

pollev.com/cis5480

- What if there are four frames instead of 3? How Many Page Faults?

- Easier for me to think about this in terms of subsequences

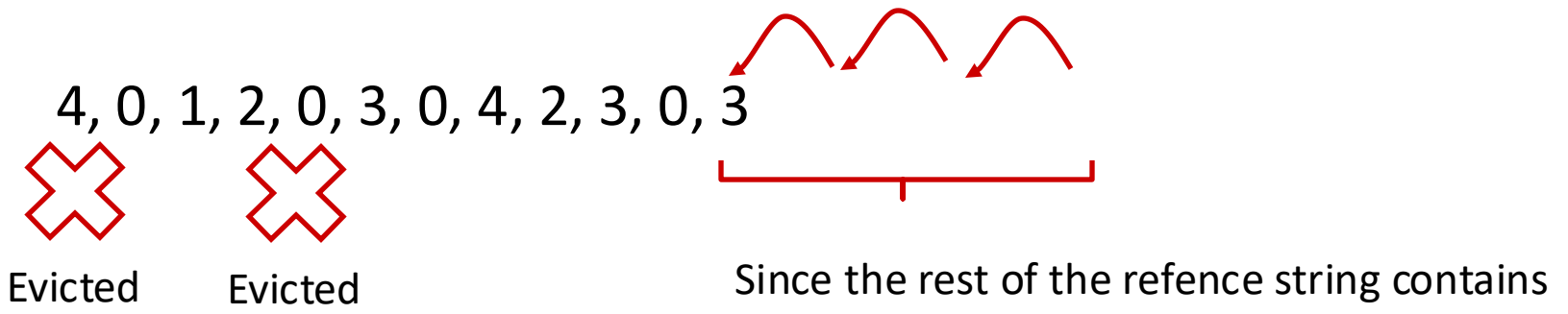


Poll Everywhere

pollev.com/cis5480

- What if there are four frames instead of 3? How Many Page Faults?

- Easier for me to think about this in terms of subsequences



Since the rest of the refence string contains pages already in the stack, there's no need to keep going.

There will be no more page evictions! (yay!)

Easier for me to think about it in this way … but do what you gotta do!

LRU Implementation

- Couple of Possibilities
  - we would need to timestamp each memory access and keep a sorted list of these pages
    - High overhead, timestamps can be tricky to manage :/
  - Keep a counter that is incremented for each memory access
    Look through the table to find the lowest counter value on eviction
    - Looking through the table can be slow
    - Should you weigh time of access more when it's more recent? (e.g. 3,3,3,3,3,3,3,3,3,1,1)
  - Whenever a page is accessed find it in the stack of active pages and move it to the bottom

LRU Approximation: Reference Bit & Clock

- It is expensive to do bookkeeping every time a page is accessed. Minimize the bookkeeping if possible

- When we access a page, we can update the reference bit for that PTE to show that it was accessed recently
  - This is done automatically by hardware, when accessing memory.
  - Setting a bit to 1 is much quicker than managing time stamps and re-organizing a stack

- We could check the reference bit at some clock interval to see if the page was used at all in the last interval period

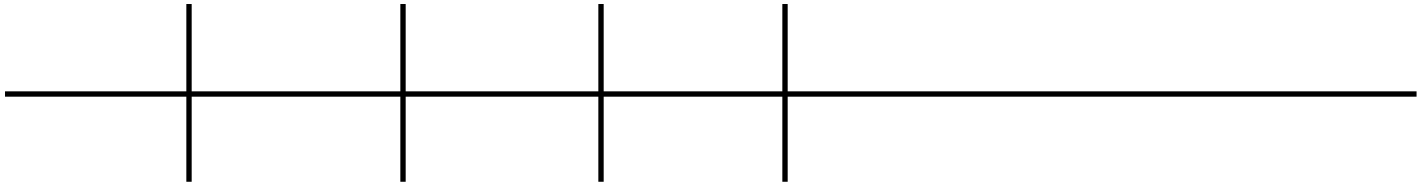
LRU Approximation: Aging

- Each page gets an 8-bit “counter”.
- On clock interval and for every page:
  - shift the counter to the right by 1 bit (**>> 1**)
  - write the reference bit into the MSB of the counter.
  - Current reference bit is reset to 0

- If we read the counter as an unsigned integer, then a larger value means the counter was accessed more recently
  - Right shifting allows us to take into consideration time since the last access as we essentially divide the value by 2 if there were no accesses.

Aging Illustration

- Timeline



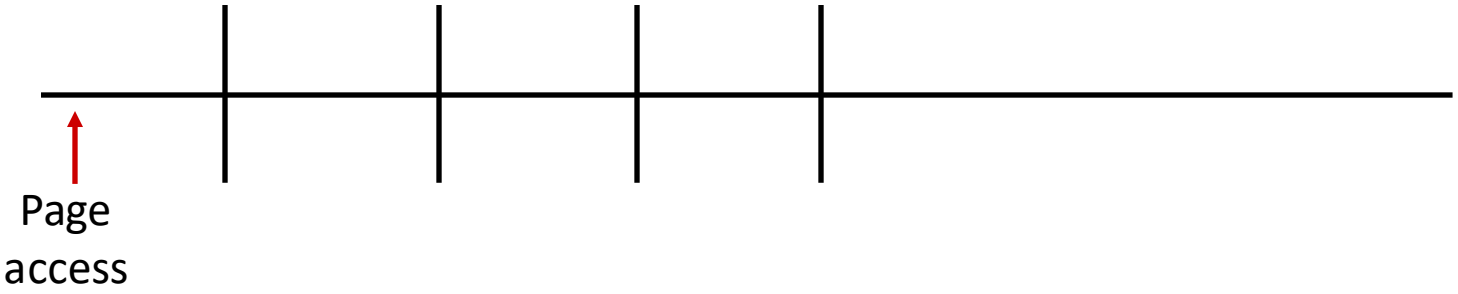
- Counter:

0	0	0	0	0	0	0	0

- Ref bit: 0

Aging Illustration

- Timeline



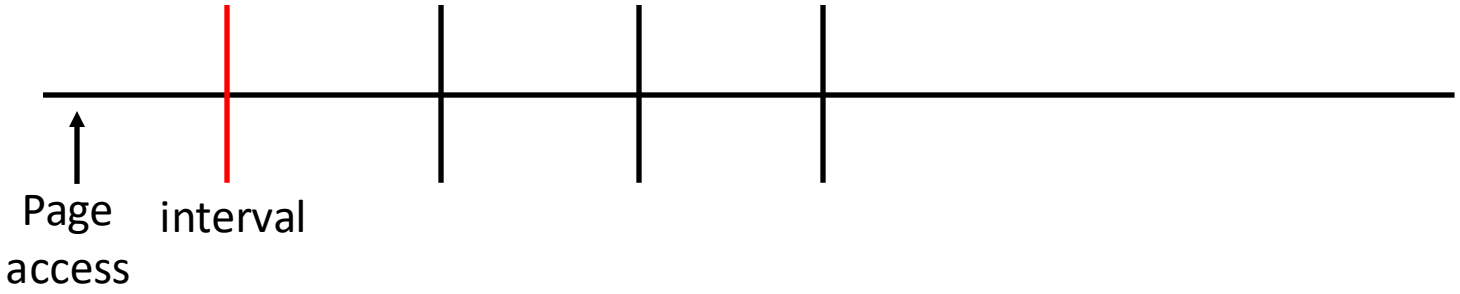
- Counter:

0	0	0	0	0	0	0	0

- Ref bit: 1

Aging Illustration

- Timeline



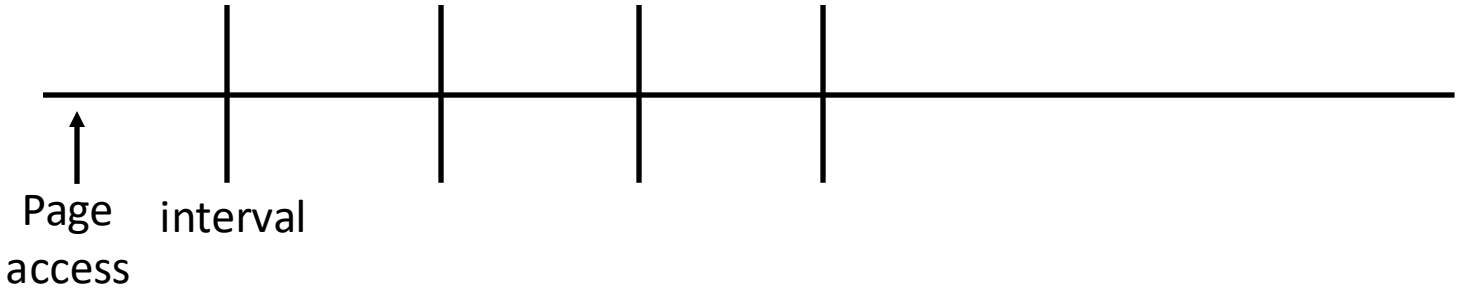
- Counter:

1	0	0	0	0	0	0	0

- Ref bit: 0

Aging Illustration

- Timeline



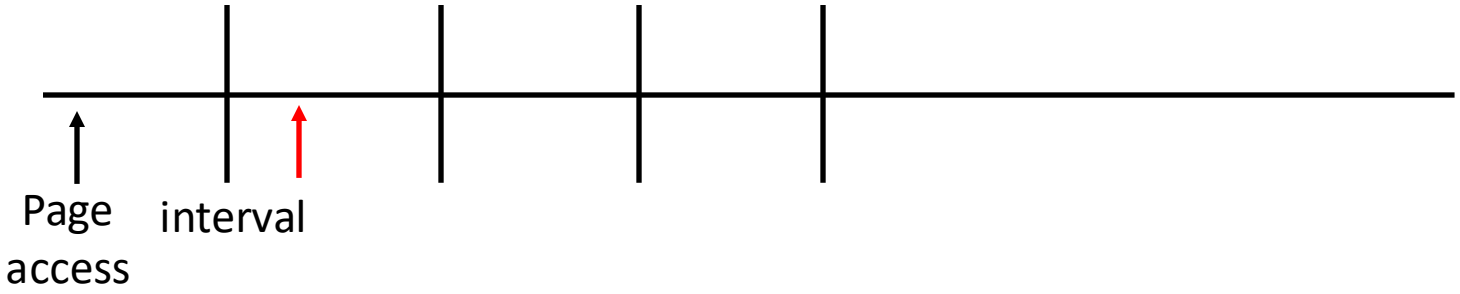
- Counter:

1	0	0	0	0	0	0	0

- Ref bit: 0

Aging Illustration

- Timeline



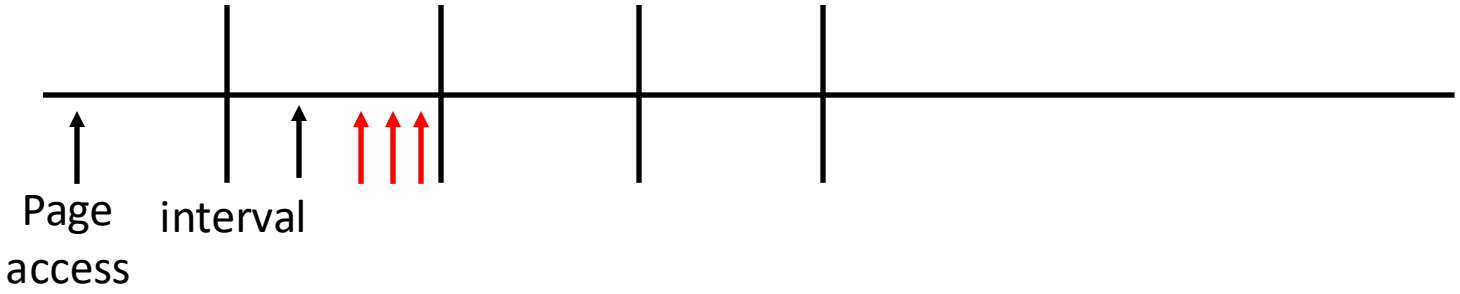
- Counter:

1	0	0	0	0	0	0	0

- Ref bit: 1

Aging Illustration

- Timeline



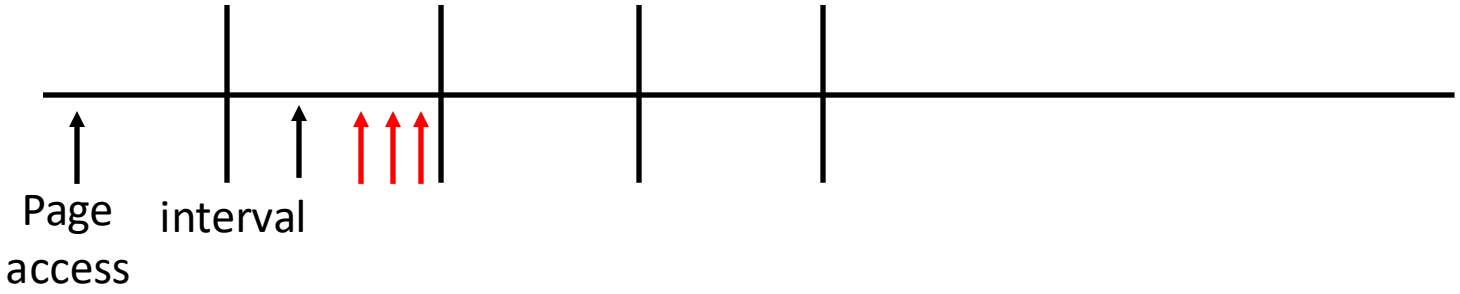
- Counter:

1	0	0	0	0	0	0	0

- Ref bit: 1

Aging Illustration

- Timeline



- Counter:

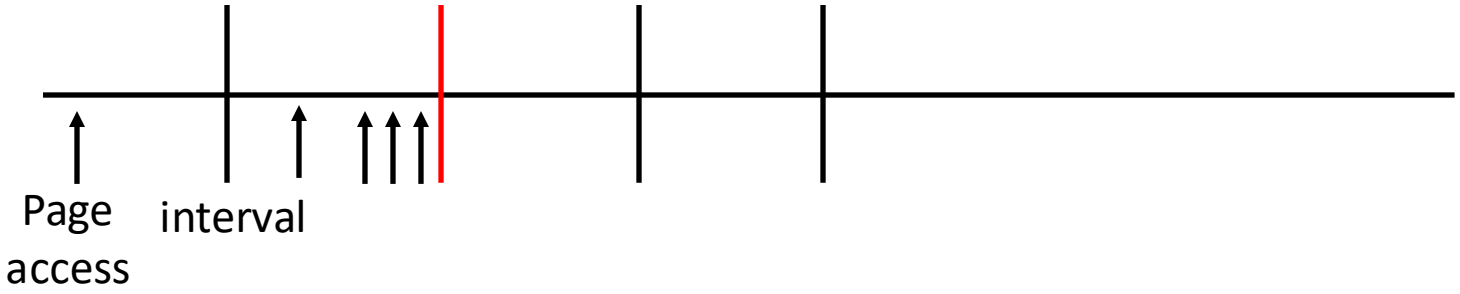
1	0	0	0	0	0	0	0

counter = (**uint8_t**)counter >> 1
counter = counter | (ref << 7)

- Ref bit: 1

Aging Illustration

- Timeline



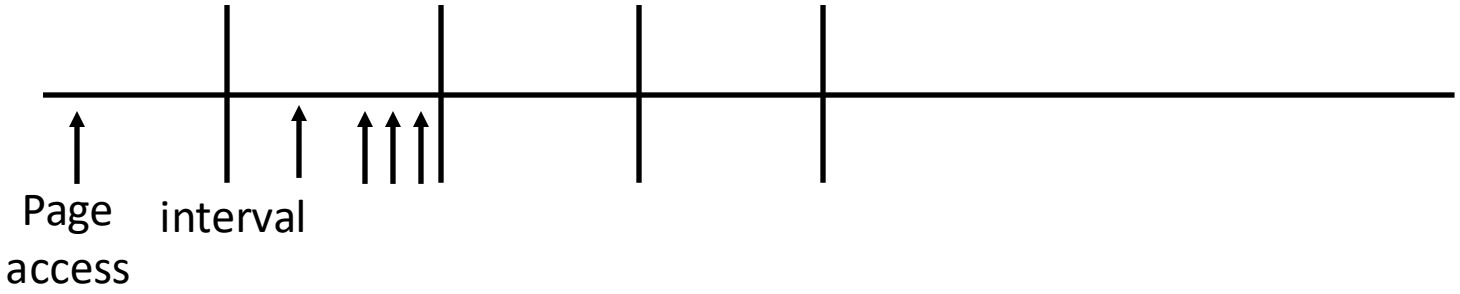
- Counter:

1	1	0	0	0	0	0	0

- Ref bit: 0

Aging Illustration

- Timeline



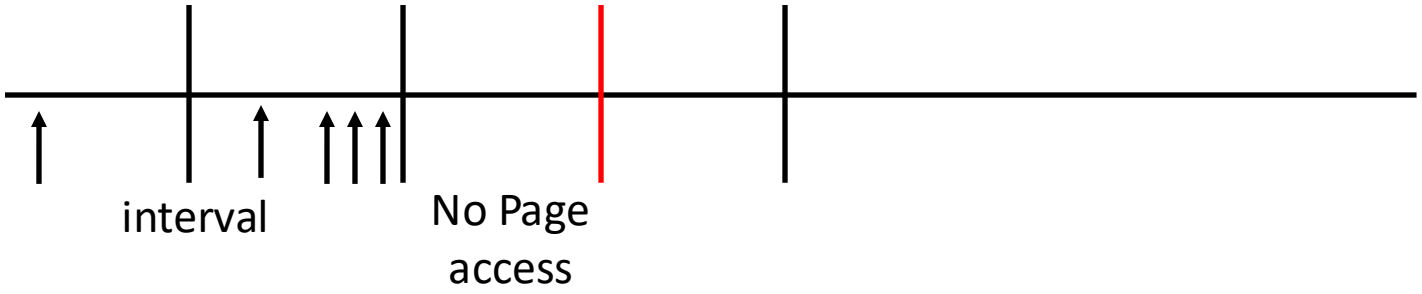
- Counter:

1	1	0	0	0	0	0	0

- Ref bit: 0

Aging Illustration

- Timeline



- Counter:

0	1	1	0	0	0	0	0

- Ref bit: 0

Aging: Analysis

- Analysis
  - Low overhead on clock tick and memory access
  - Still must search page table for entry to remove/update
  - Insufficient information to handle some ties
    - Only one bit information per clock cycle
    - Information past a certain clock cycle is lost

Lecture Outline

- Page Replacement: High Level
  - FIFO
  - Reference Strings
  - Beladys
- LRU
- **Thrashing**
- FIFO w/ Reference bit

Thrashing

- This is not specific to LRU, but it is easiest to demonstrate with LRU

- When the physical memory of a computer is overcommitted, causing almost constant page faults (which are slow)
  - Overcommitment most commonly happens when there are too many processes, and thus too much memory needed
  - Can also happen with a few processes, if the process needs too much memory

Thrashing: LRU Example

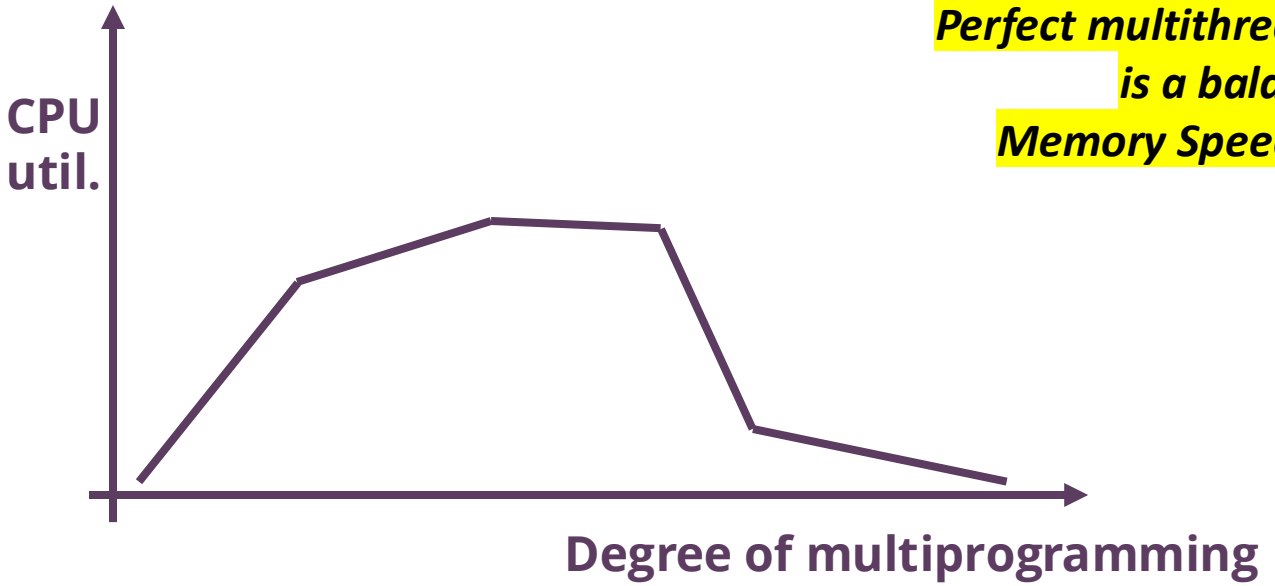
- Consider the following example with three page frames and LRU

LRU	Ref str:	0	1	2	3	0	1	2	3	0	1	2	3
Recent		0	1	1	2	0	1	2	3	0	1	2	3
			0	1	2	3	0	1	2	3	0	1	2
To Evict				0	1	2	3	0	1	2	3	0	1
Evicted					0	1	2	3	0	1	2	3	0

- Page fault on every memory access ☹

Thrashing: Multiprogramming

- It is good to have more processes running, then we can have better utilization of CPU.
  - While one process waits on something, another can run
  - More on CPU Utilization later
- As we use more processes running at once, more memory is needed, can cause thrashing ☹



Perfect multithreading/multiprocessing is a balance between Memory Speed/Size vs CPU Speed

Lecture Outline

- Page Replacement: High Level
  - FIFO
  - Reference Strings
  - Beladys
- LRU
- Thrashing
- **FIFO w/ Reference bit**

FIFO Analysis

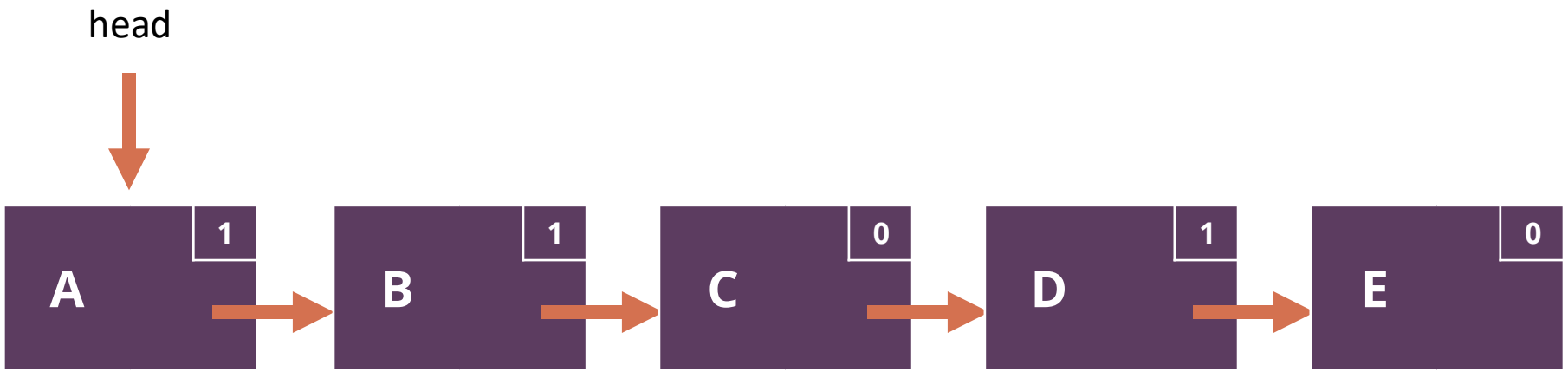
- Remember FIFO? The first page replacement algorithm we covered?
  - Evict the page that has been in physical memory the longest
- Analysis:
  - Low overhead. No need to do any work on each memory access, instead just need to do something when loading a new page into memory & evicting an existing page
  - Not the best at predicting which pages are used in the future ☹
- Could we modify FIFO to better suit our needs?

Second Chance

- Second chance algorithm is very similar to FIFO
  - Still have a FIFO queue
  - When we take the first page of the queue, instead of immediately evicting it, we instead check to see if the reference bit is 1 (was used in the last time interval)
  - If so, move it to the end of the queue
  - Repeat until we find a value that does not have the reference bit set (if all pages have reference bit as 1, then we eventually get back to the first page we looked at)

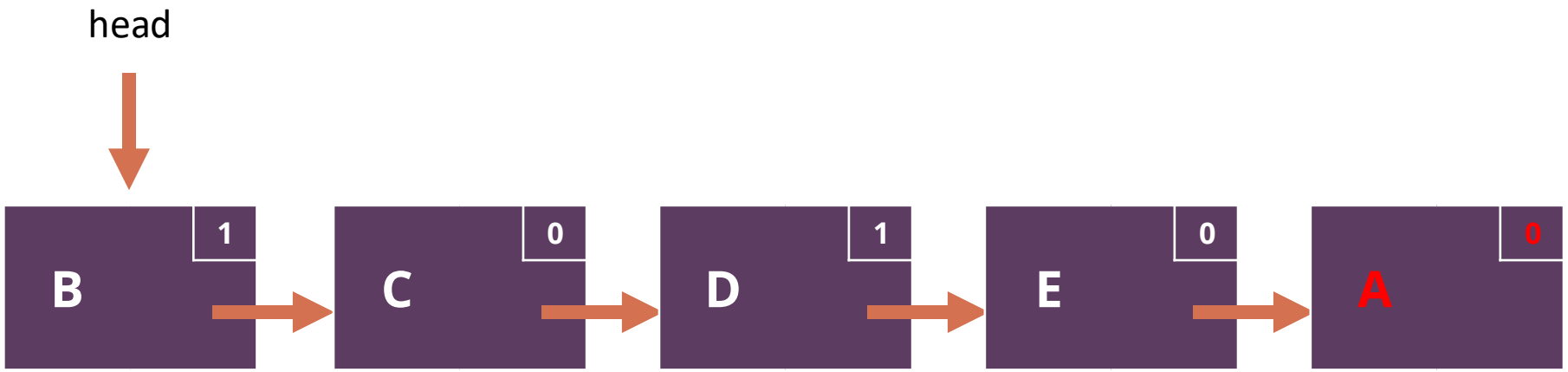
Second Chance Example

- If we need to evict a page: start at the front
- Reference bit is 1, so set to 0 and move to end



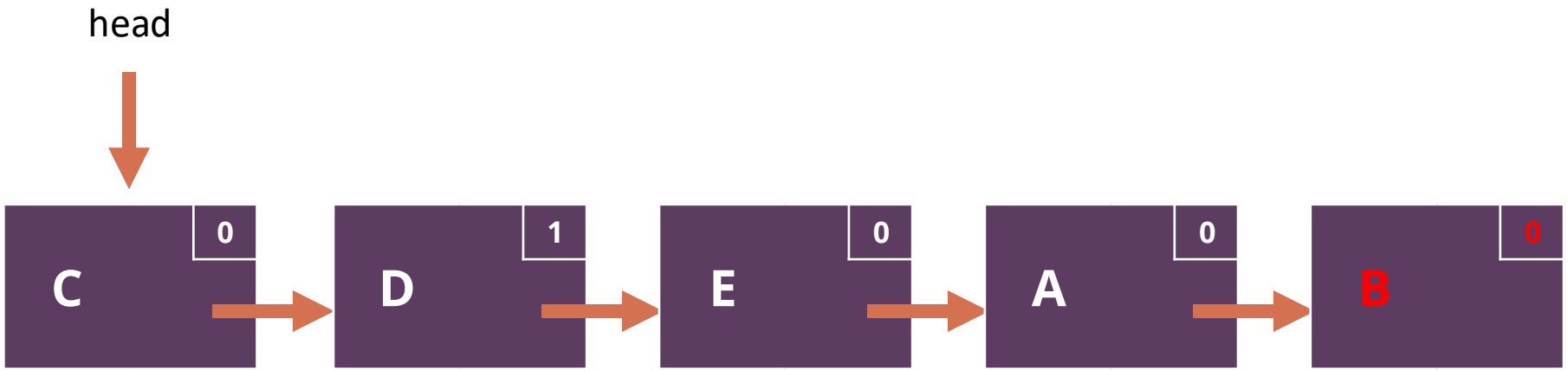
Second Chance Example

- If we need to evict a page: start at the front
- Reference bit is 1, so move to end



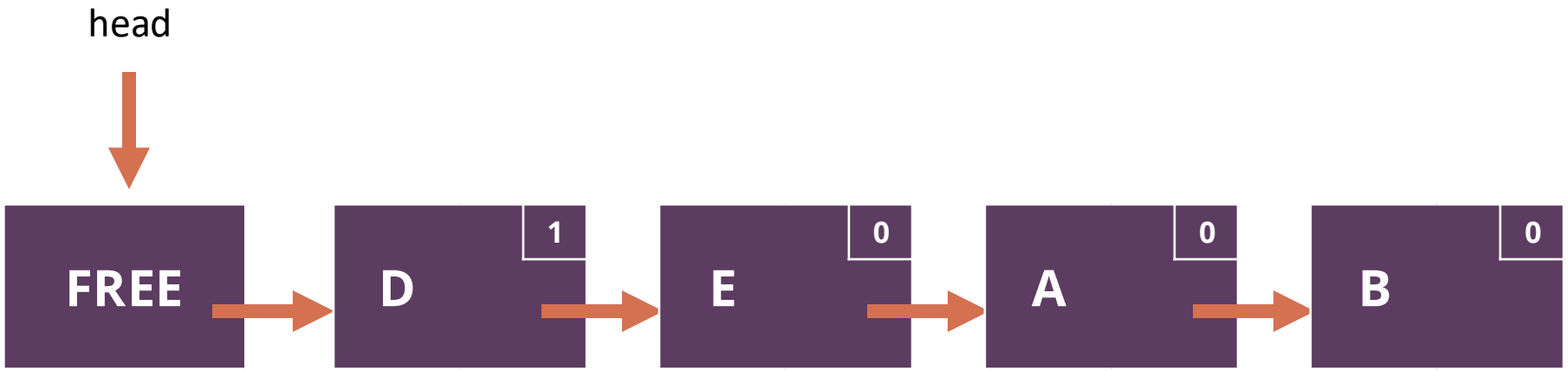
Second Chance Example

- If we need to evict a page: start at the front
- Reference bit is 1, so move to end



Second Chance Example

- If we need to evict a page: start at the front
- Found a page with reference bit = 0, evict Page C!

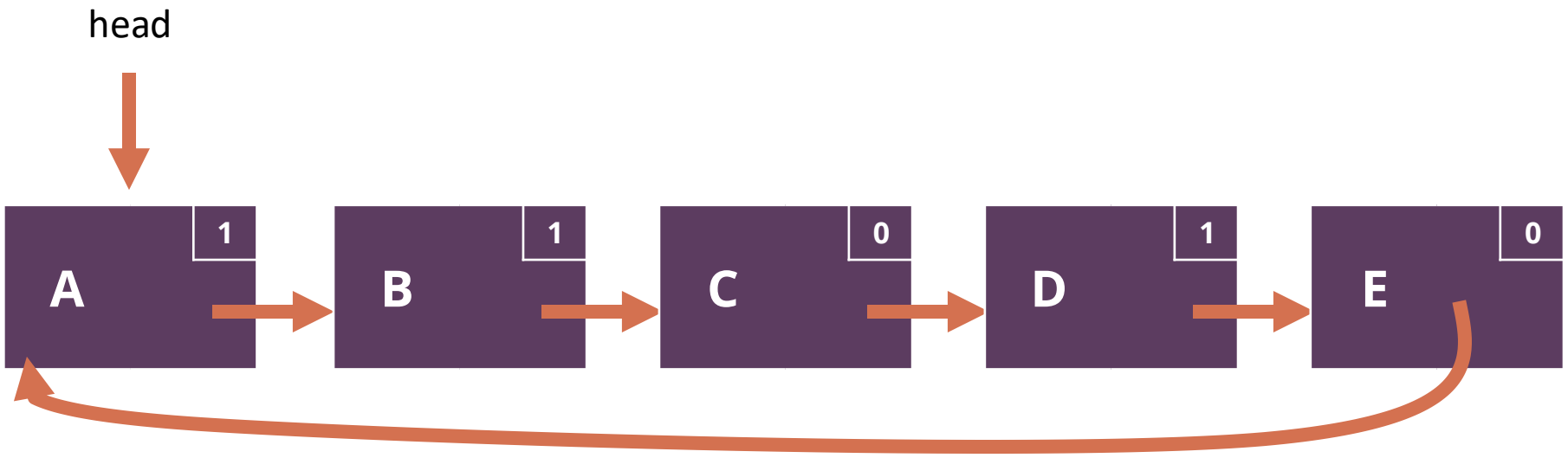


Clock

- Optimization on the second chance algorithm
- Have the queue be circular, thus the cost to moving something to the “end” is minimal

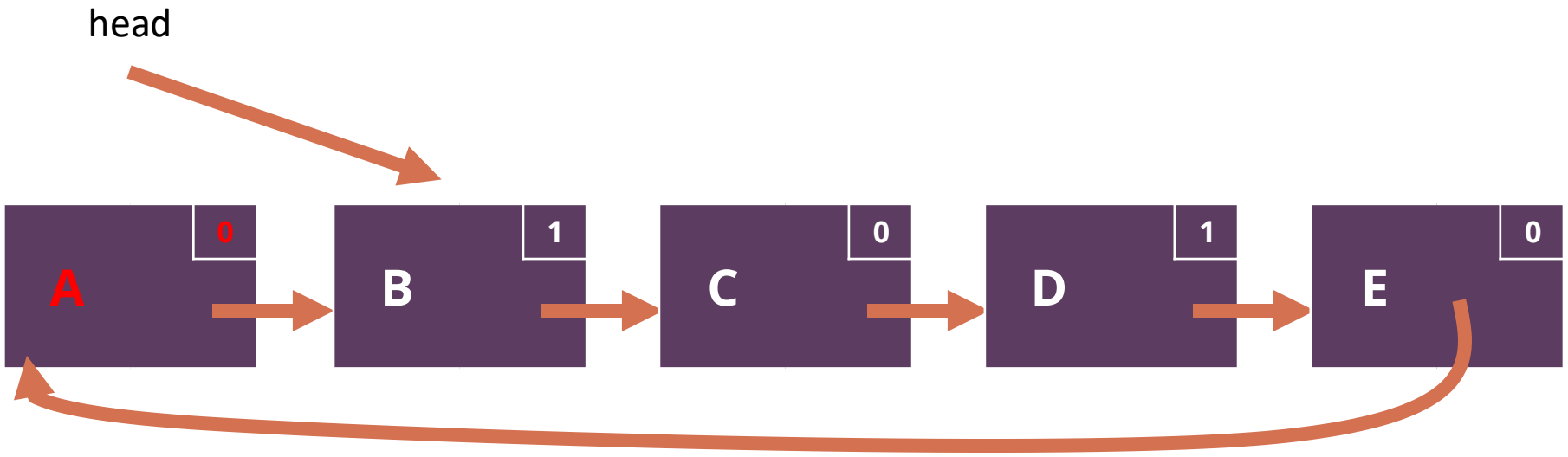
Clock Example

- If we need to evict a page: start at the front
- Reference bit is 1, so set to 0 and move to end



Clock Example

- If we need to evict a page: start at the front
- Reference bit is 1, so set to 0 and move to end

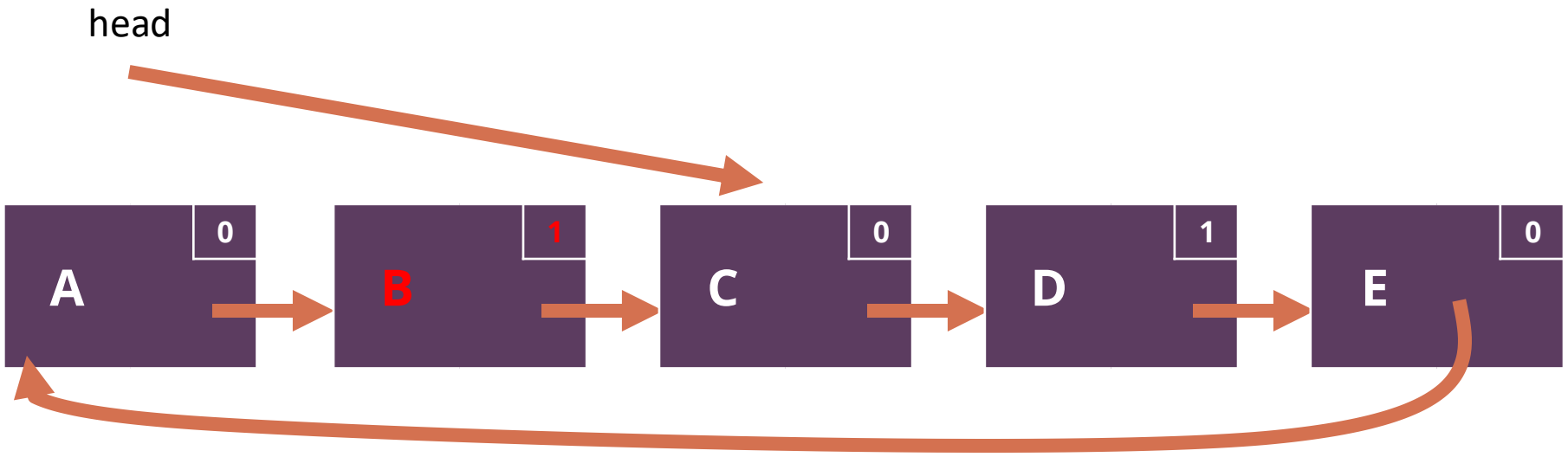


Clock Example

- If we need to evict a page: start at the front

- Reference bit is 1, so set to 0 and move to end

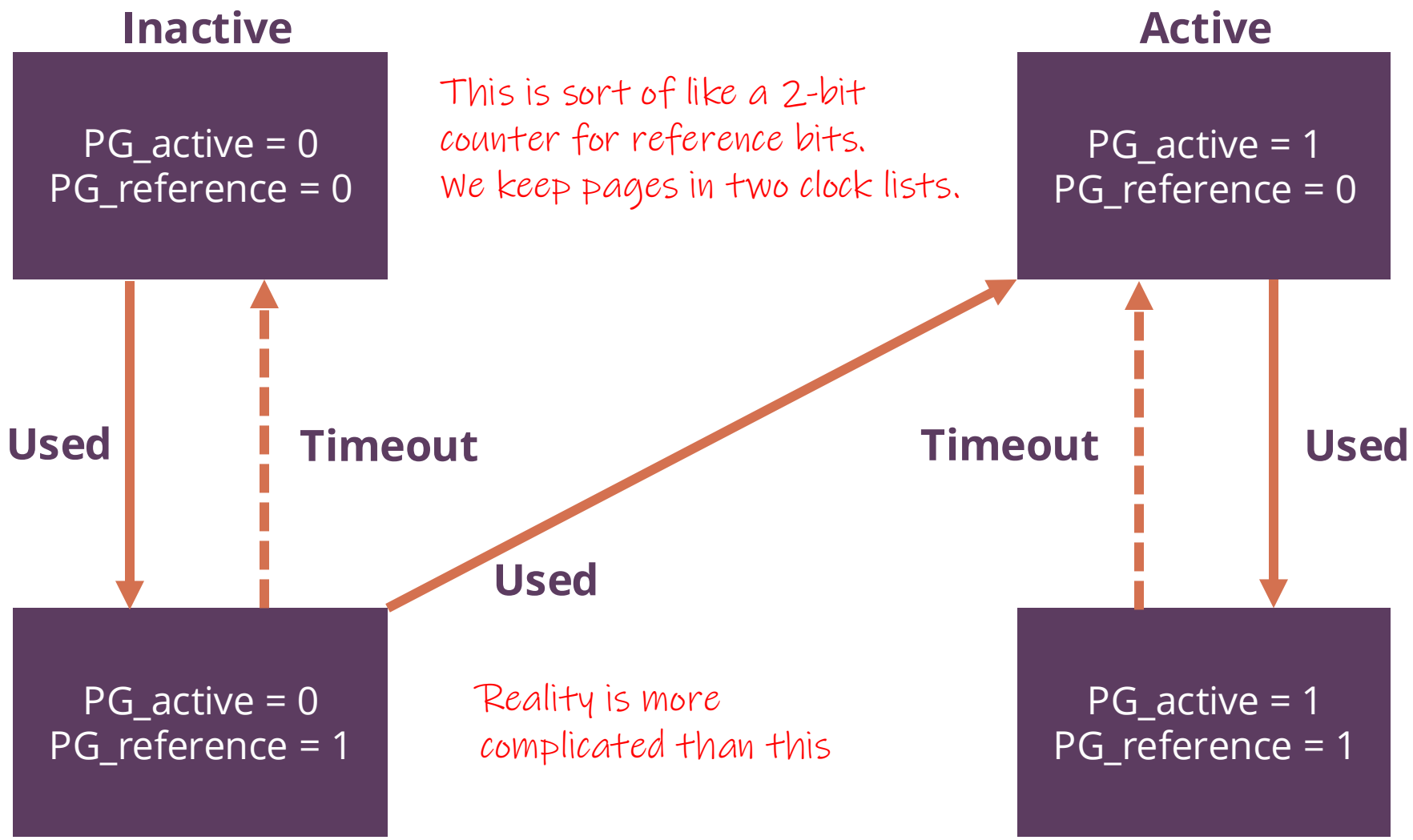
Can also be modified to prefer to evict clean pages instead of dirty pages



Linux Two-List Clock Page Replacement Algorithm

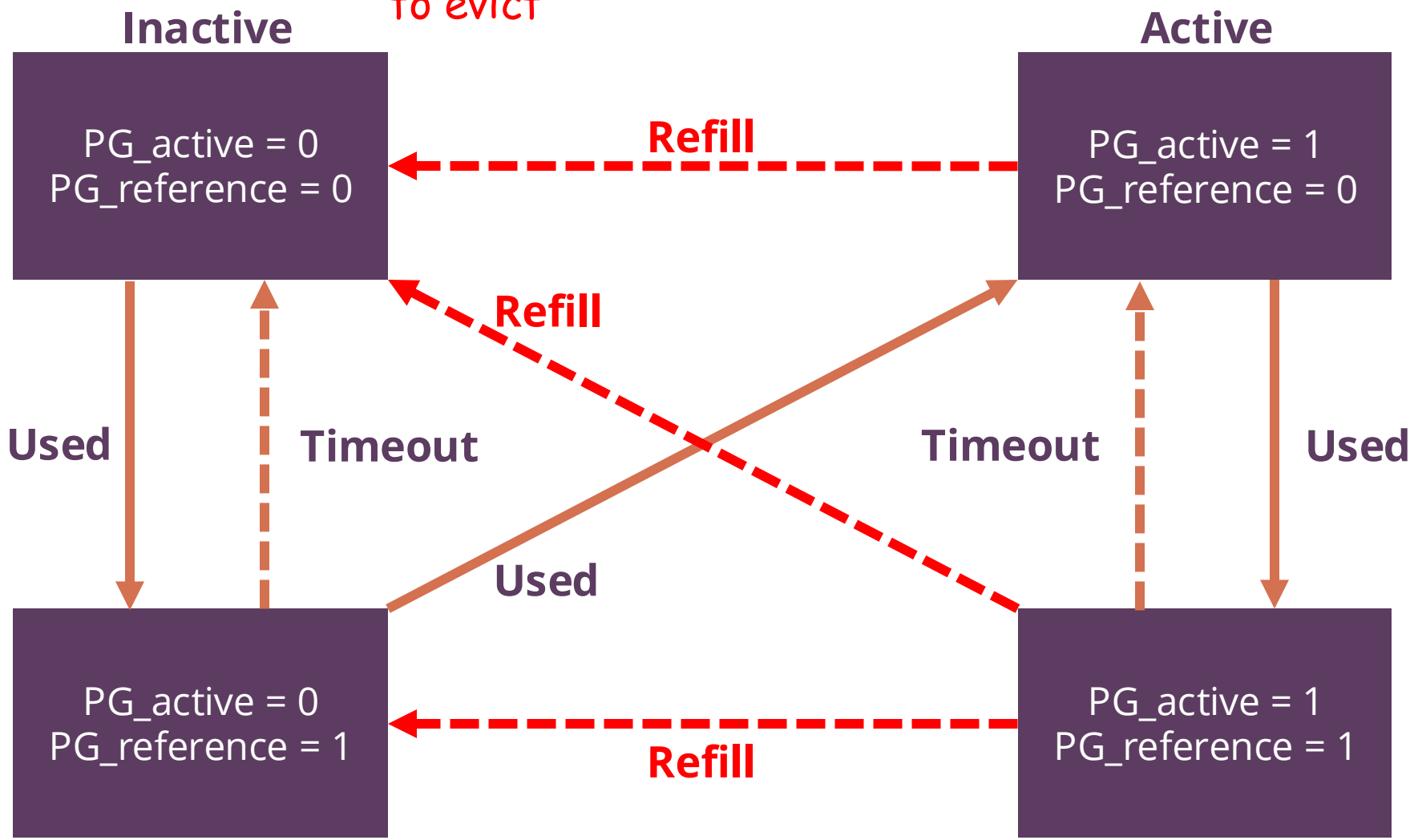
- Maintains two lists: Active list and Inactive list
- Eviction Priority:
  - Chose a page from the inactive list first
- Page Access Behavior:
  - If a page has not been referenced recently, move it to the inactive list
- If a page is referenced:
  - Set its reference flag to true
  - It will be moved to the active list on the next access
  - Two accesses are required for a page to become *active*
- Decay Mechanism:
  - If the second access doesn't happen, the reference flag is reset periodically
  - After two timeouts without activity, the page is moved to the inactive list

Linux diagram



Linux diagram

Linux will want to keep a good ratio of inactive to active, so that there are always some pages that are considered "more ok" to evict



Active should be ~2/3 of pages at most