MADALGO
CENTER FOR MASSIVE DATA ALGORITHMICS


Andrei Negoescu
Goethe University

GOETHE UNIVERSITÄT FRANKFURT AM MAIN

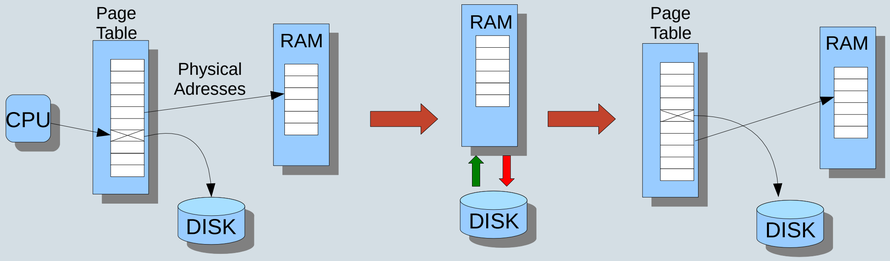
# Online Paging for Flash Memory Devices

## Paging

Paging is a memory-management scheme used by operating systems when the required memory space does not fit into the fast memory (RAM) and thus parts of it reside on the secondary device (disk).

**Page fault:**
A memory page accessed by an active process which resides on the disk, has to be loaded into RAM involving slow I/O operations including:

- Loading the requested page from disk
- Evicting one or more pages to disk if the fast memory is full

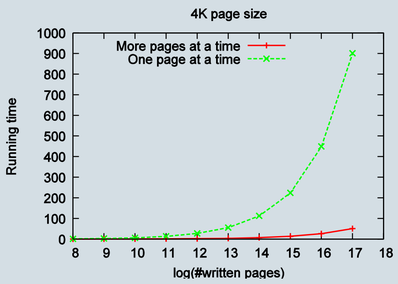


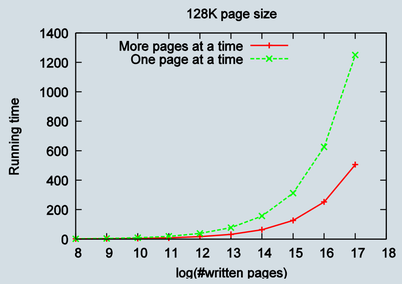

## Secondary Device: Flash Disk (SSD)

Efficient I/O in the context of paging:

- Read on a page basis
- Write in blocks of $\alpha$ pages ($\alpha>1$)
- Write operations are slow

Writing pages from random positions of the memory to a SSD using a translation layer:



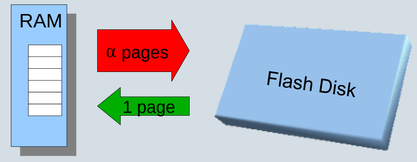

## α-Paging Model

**Input** consists of an online sequence of page requests.

**Online Algorithm**
If the memory is full and a page fault occurs, the algorithm has to decide which pages to swap to disk during an eviction.

**Eviction**
Up to **α** pages can be swapped to disk during **one** eviction.

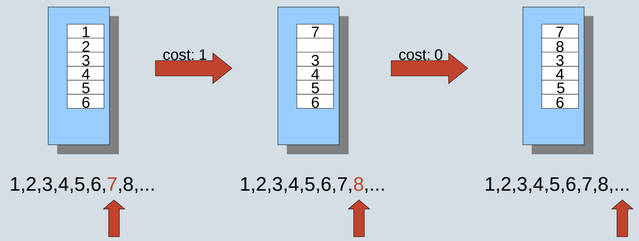
**Goal**: Minimize the number of evictions.




## Algorithms

- **α-LRU** evicts the α least recently used pages**.**

  Example for $\alpha=2$:




- **α-MARK** is a randomized algorithm which keeps track of an interval splitting**.** Only pages not requested in the current interval can be evicted. These are chosen uniformly at random.
- **α-MIN** is an optimal offline algorithm, which knows all future requests in advance. It evicts the α pages whose first requests occur farthest in the future.

## Results

**Competitive Ratio** is a performance measure for online algorithms. It provides the worst case ratio between the cost of the online algorithm and an optimal offline algorithm.

The variable ***k*** denotes the number of pages fitting in RAM.

| Setting | Algorithm | Competitive Ratio | Lower Bound |
|---|---|---|---|
| **Deterministic** | α-LRU | $k/\alpha$ | $k/\alpha$ |
| **Randomized** | α-MARK | $2.466\cdot\ln(k/a)$ | $1.44\cdot\ln(k/a)$ |

Known competitive ratio in the classical model of the adapted algorithms:

| Setting | Algorithm | Competitive Ratio | Lower Bound |
|---|---|---|---|
| **Deterministic** | LRU | $k$ | $k$ |
| **Randomized** | MARK | $2\ln(k)$-$1$ | $\ln(k)$ |

## Future Work

- Strongly competitive (optimal) randomized online algorithm
- Distinguish between clean and dirty pages