Coeus: *Clustering (A)like Patterns For Practical Machine Intelligent Hybrid Memory Management*

Thaleia Dimitra Doudali (IMDEA Software Institute)*
Ada Gavrilovska (Georgia Tech)

@ CCGrid 2022

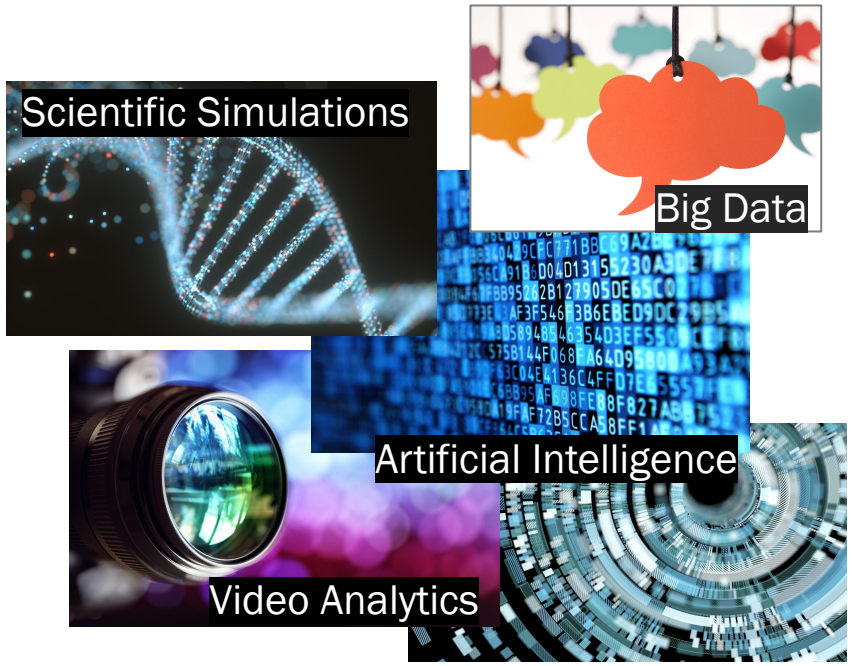
* Work done while at Georgia Tech (PhD).

The Era of Data

"More than 65 ZB of data will be created, captured, copied, and consumed in the world this year."

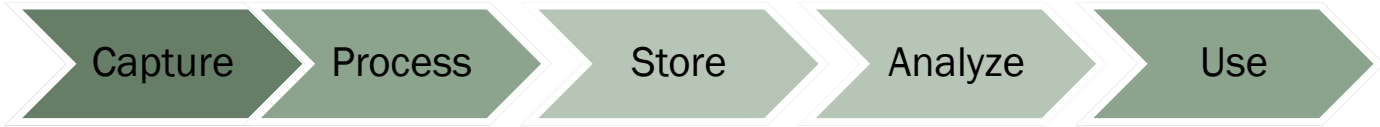
Source: International Data Corporation, March 2021.

Exploded Data Sizes



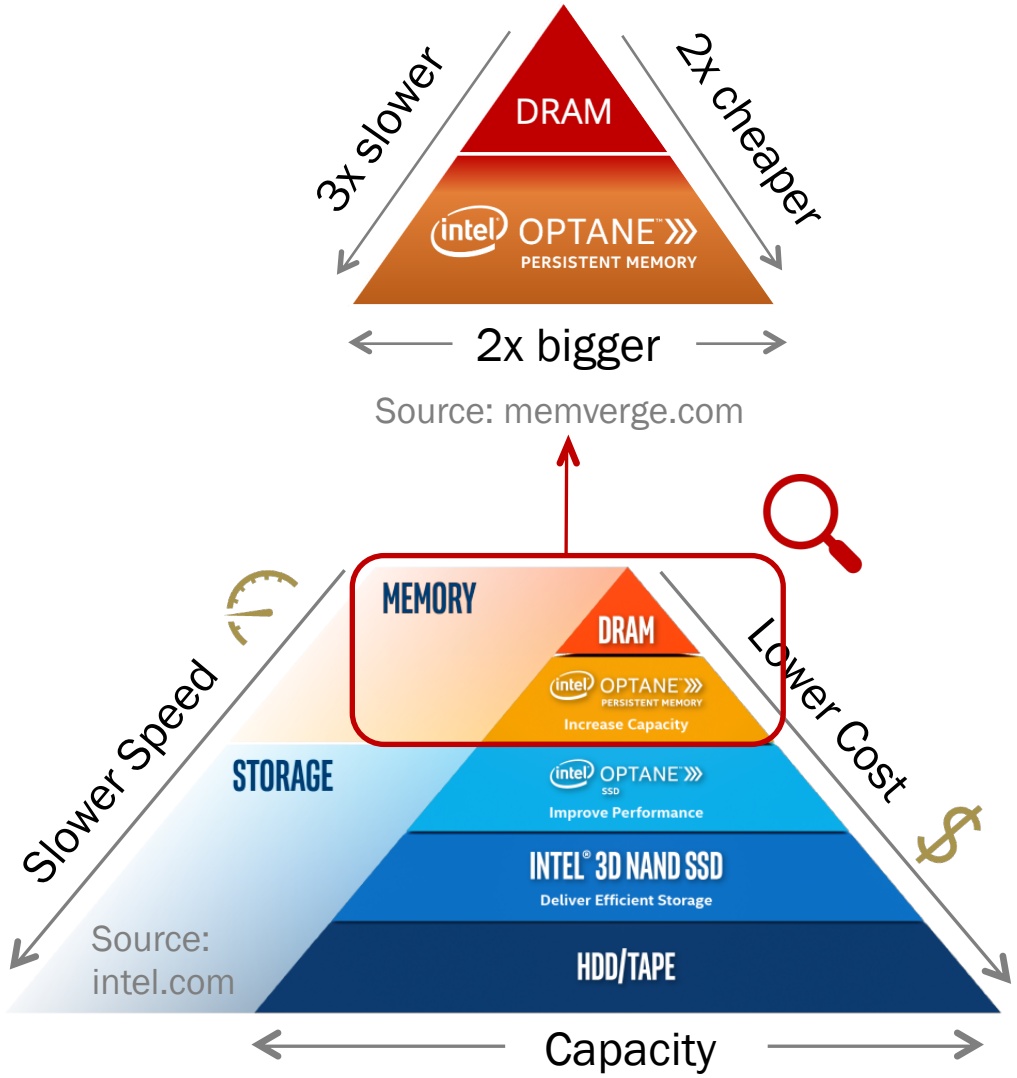
Data Analytics Pipeline

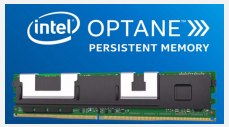
ZBs of data



Need for speed and massive storage capacities!

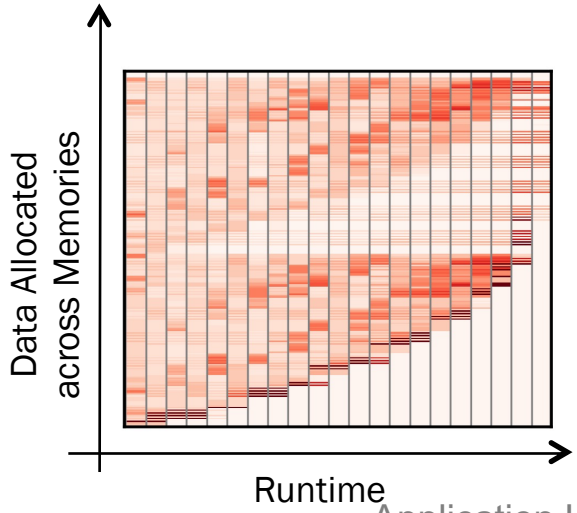
The Era of Heterogeneous Hardware



Characteristic	Technology	Hardware Vendors
Low Latency	MRAM	
High Bandwidth	HBM	
Persistence	PMEM	

Examples of other heterogeneous memory technologies.

Hybrid Memory Management is Complex

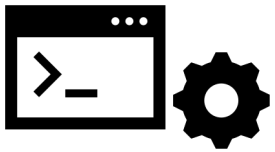


Dynamic Data Movements!

It is a **complex decision mix** to manage the data allocated across memories.

E.g., Which / How much / Where / When to move data?

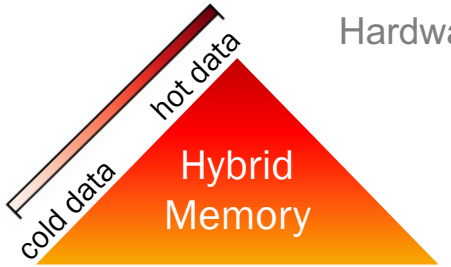
Application Level



Resource Management Software

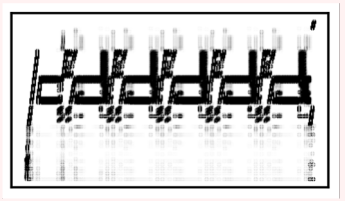
System Level

Hardware Level

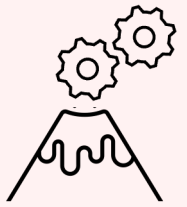


Why do we need more intelligent systems?

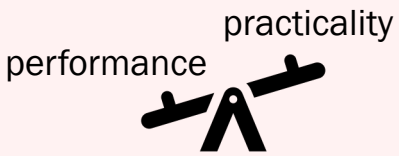
Application data sizes



Complex data access patterns



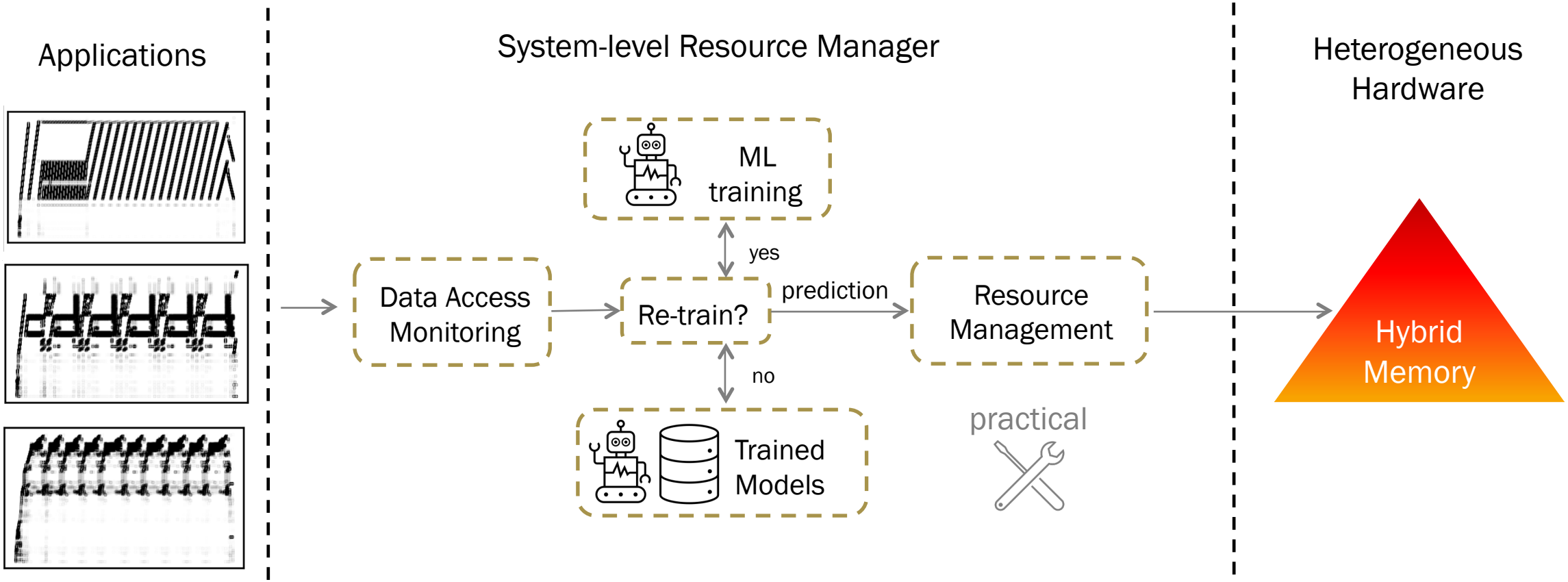
Exploded system parameter space



Hard to balance

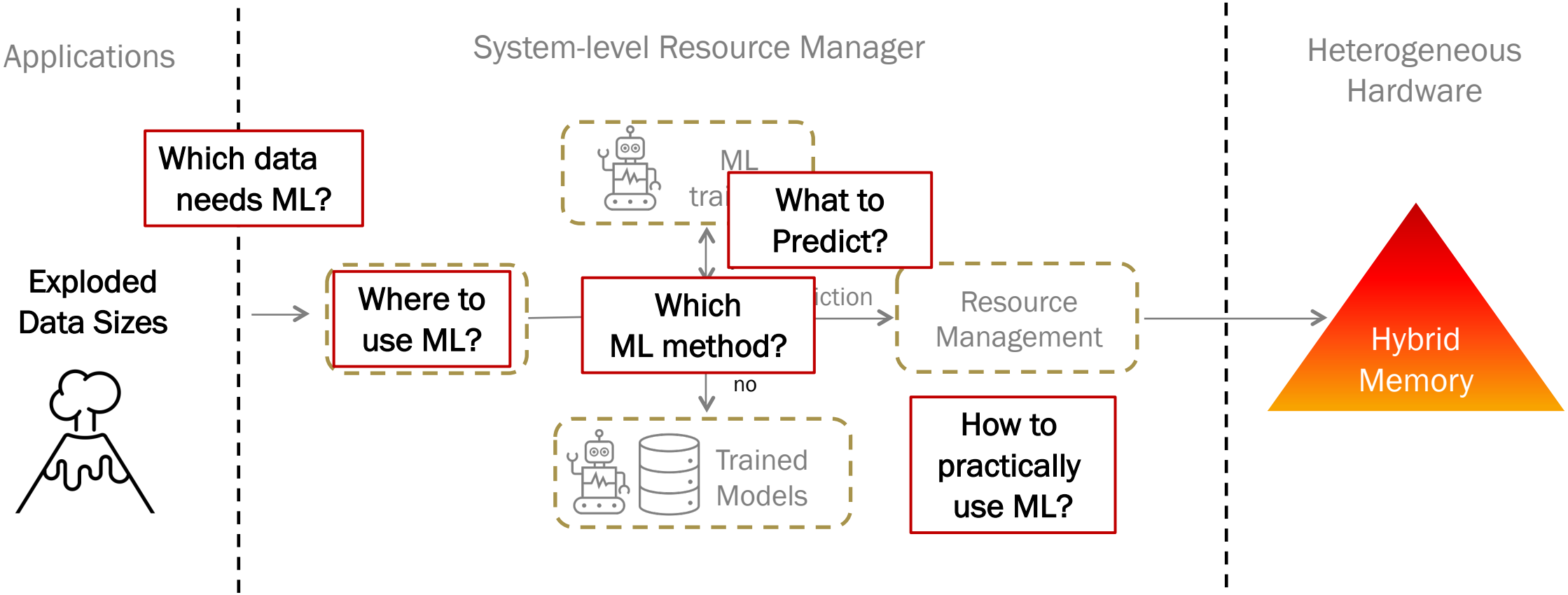
Machine Intelligent Hybrid Memory Management

The Vision.



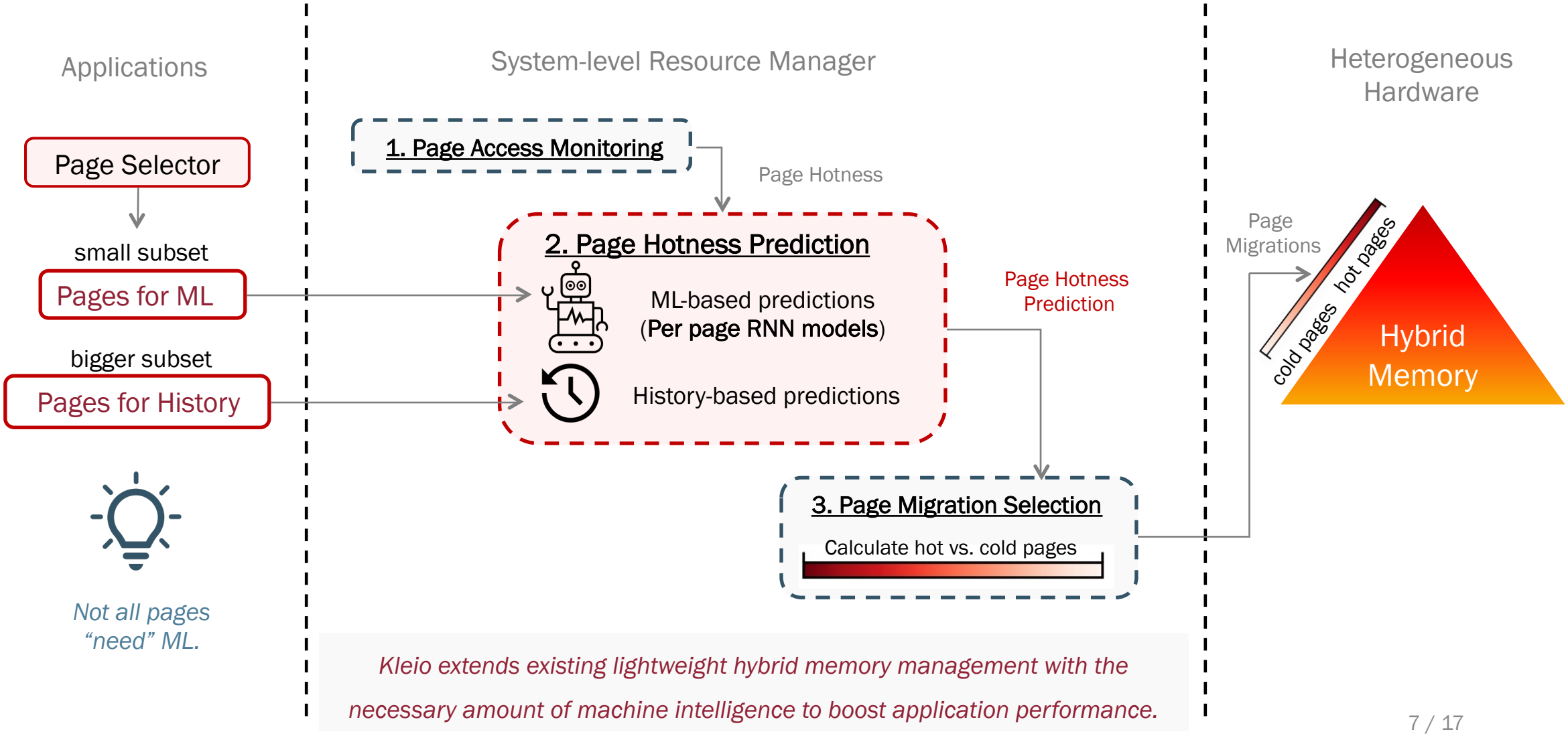
Machine Intelligent Hybrid Memory Management

Laying the grounds for the practical integration of ML.

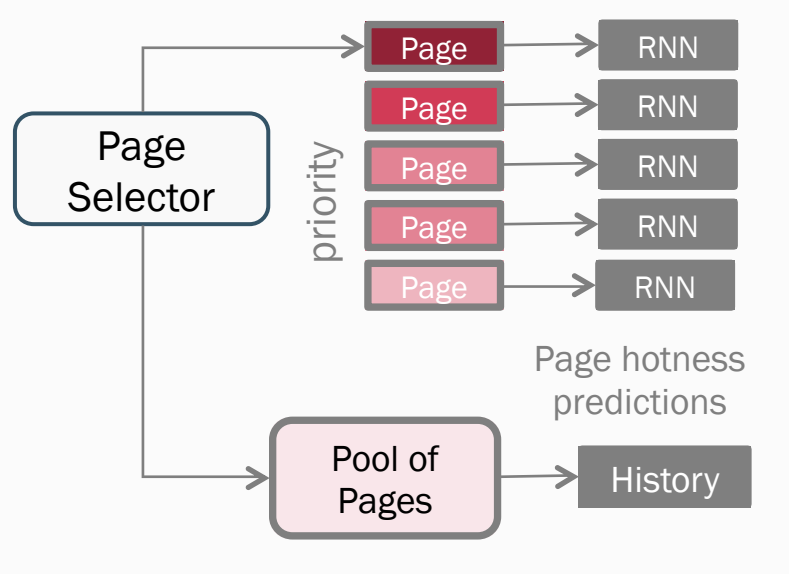


System Design of *Kleio*

Kleio is a hybrid memory page scheduler with machine intelligence. [Best Paper Award Finalist at HPDC 2019.]

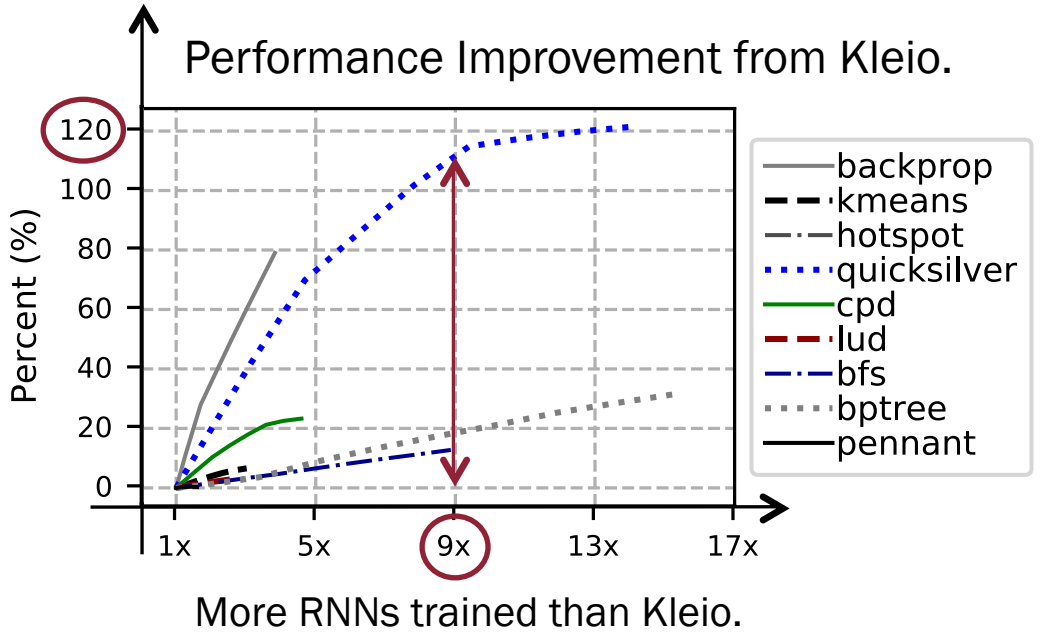


How “Small” is the Page Subset for ML?



Kleio’s evaluation trains is limited to training 100 RNNs.

The higher, the better.



The more, the worse.

*The current design of Kleio (1 model per page) allows for a **big variability** in the number of ML models (resource and time overheads) required to maximize application performance.*

Scaling ML-based Management

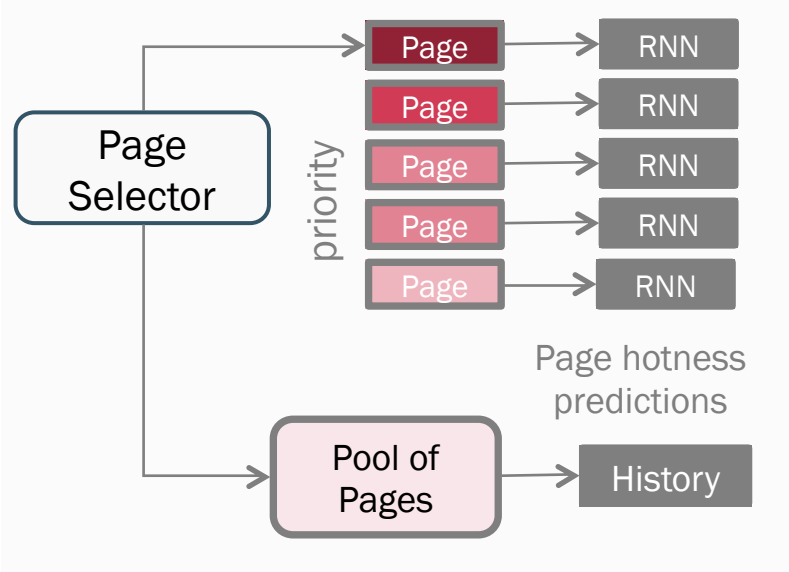
↑ Application Memory Footprints

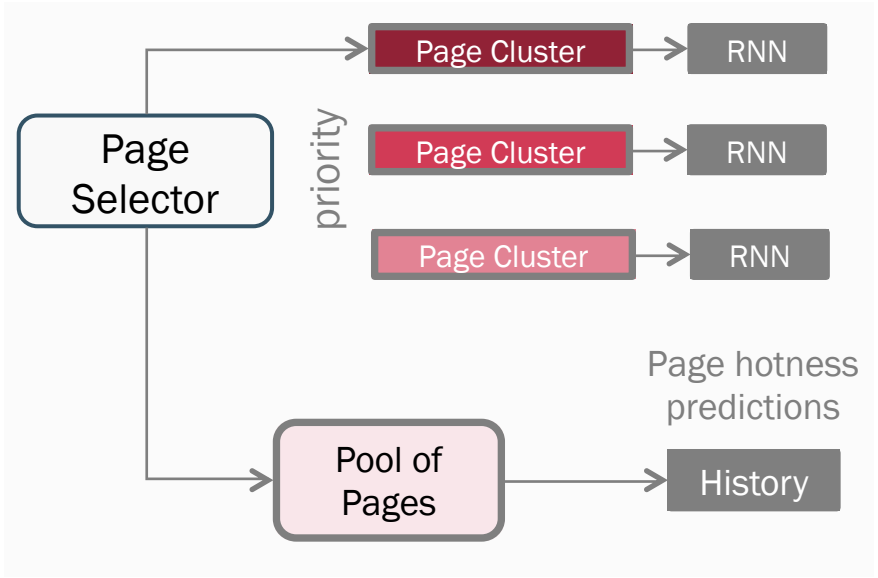
↑ Irregularity in data access patterns

↑ Absolute number of RNN models per page

↑ ML overheads

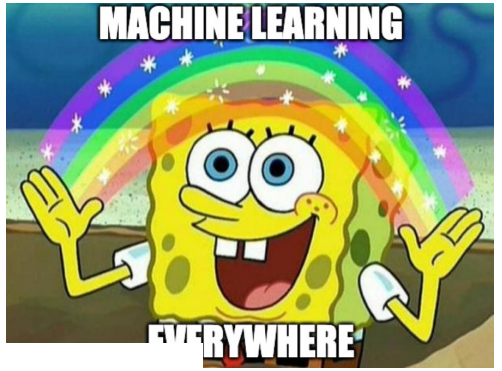
Can we scale the RNN models to learn patterns of more pages?



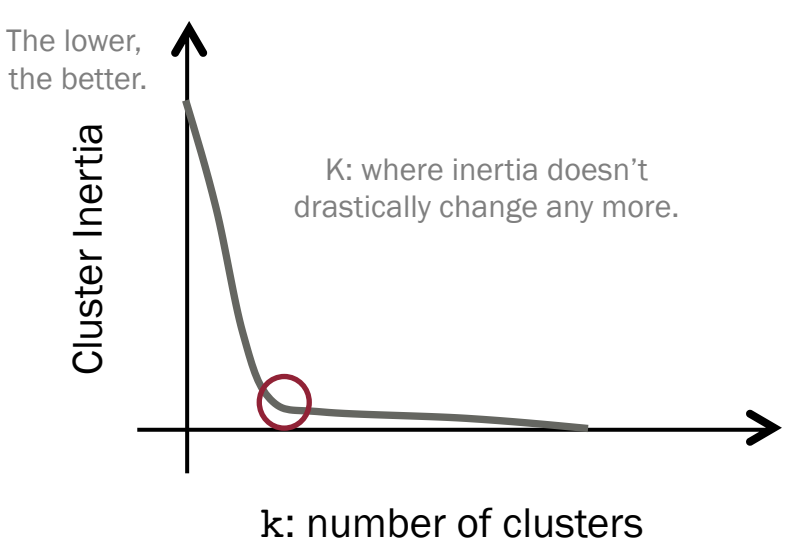


Clustering? Let's use Machine Learning!

K-means is a very popular unsupervised learning data clustering method.

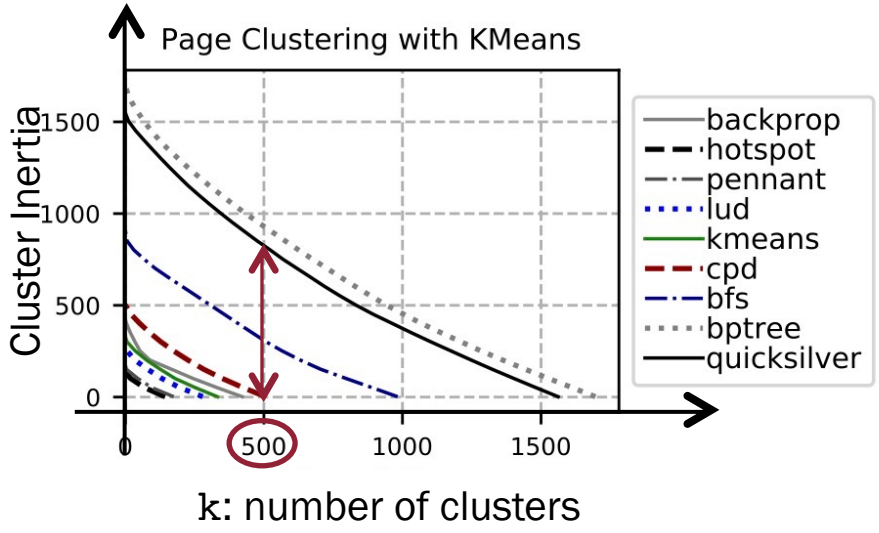


Challenge 1: Decide the number of clusters (k).



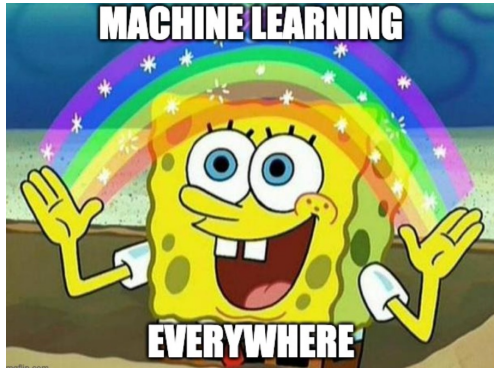
Lower inertia means higher similarity within a cluster.

Inertia = 0: data in cluster is identical.



- No single value of `k` works best for all.
- It is not obvious which value of `k` to choose.

Clustering? Let's use Machine Learning!



Let's assume we know how many clusters to create.

Challenge 2: What input to use for RNN training over a cluster of pages?

Option 1: input data from all pages.

↑ Number of Pages

↑ Learning Overheads

Option 2: input data from the "centroid" page.

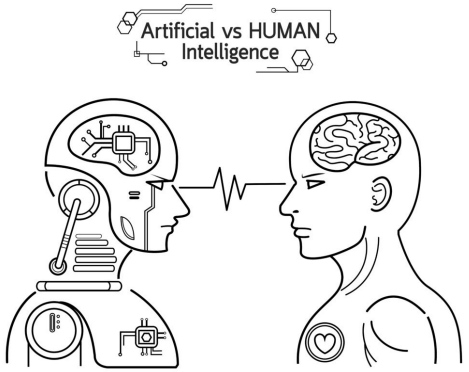
↑ Cluster Dissimilarity

↓ ML Prediction Accuracy

Ineffective clustering may lead to large learning overheads and reduced prediction accuracy levels.

☹ Clustering with ML methods introduces non trivial complexity and overheads.

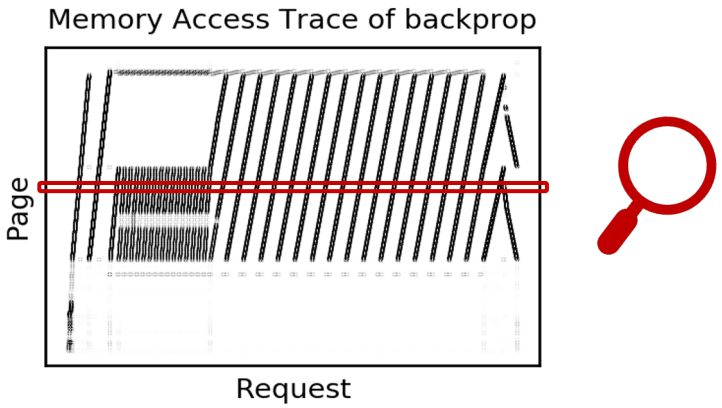
😎 How can we Keep it Smart but Simple ?

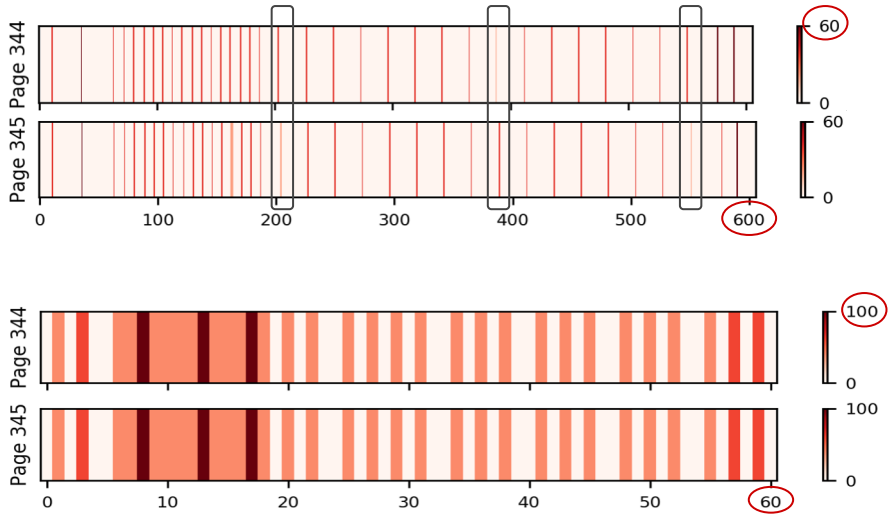


Insight on Similar vs. Identical Patterns

Let’s take a step back and observe the per page patterns.

Input to the ML model is the sequence of page hotness across periods of time.





Short Periods

Sequences are slightly different.

Long Periods

Sequences are **identical**.

It is all about the right granularity!

Which is the Right Granularity?

Longer periods result in more pages having identical patterns of page access hotness across time.

Longer periods may result in insufficient data movements and impact application performance.

Periods that align with the page reuse distance, maximize performance.

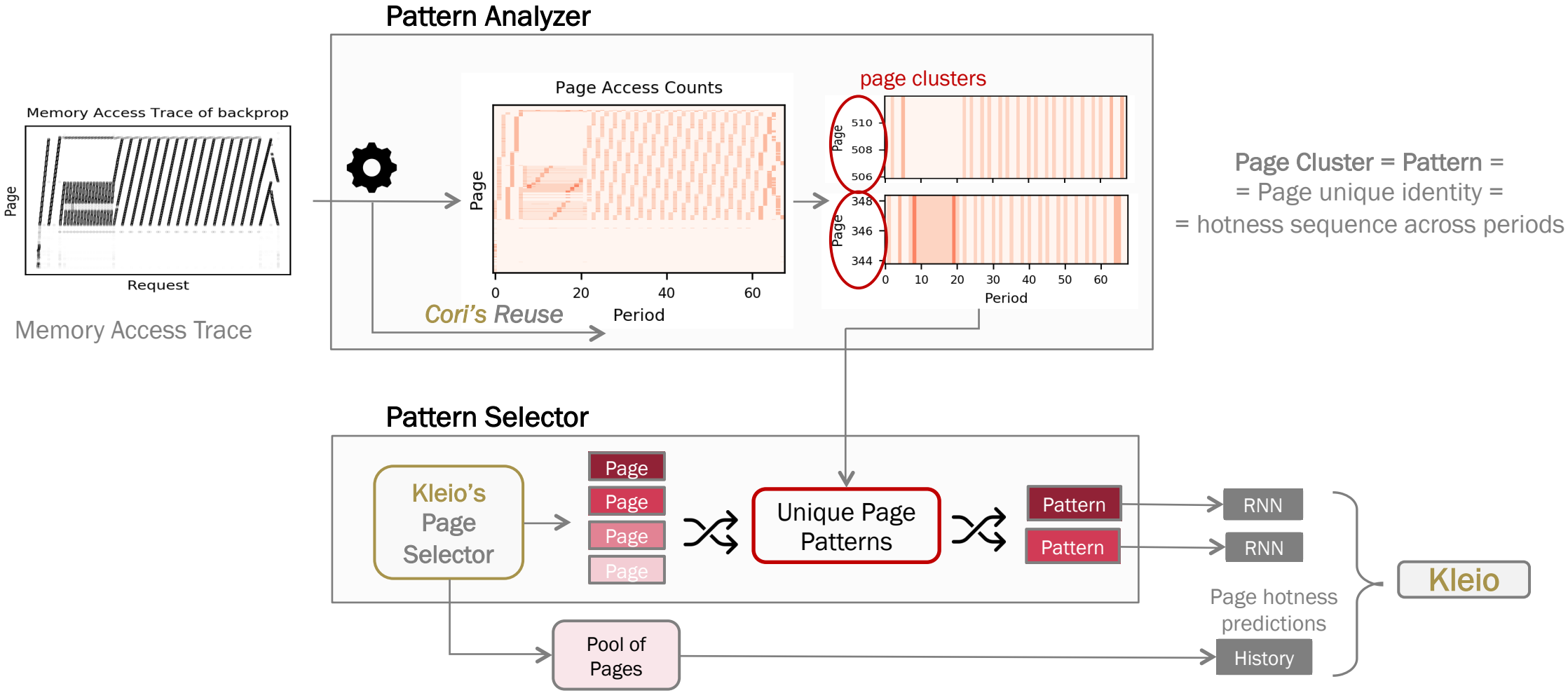
***Page Reuse Distance** = The time gap between two accesses to the same page.

Insight from ***Cori**: Dancing to the Right Beat of Periodic Data Movements over Hybrid Memory Systems.* [IPDPS 2021]

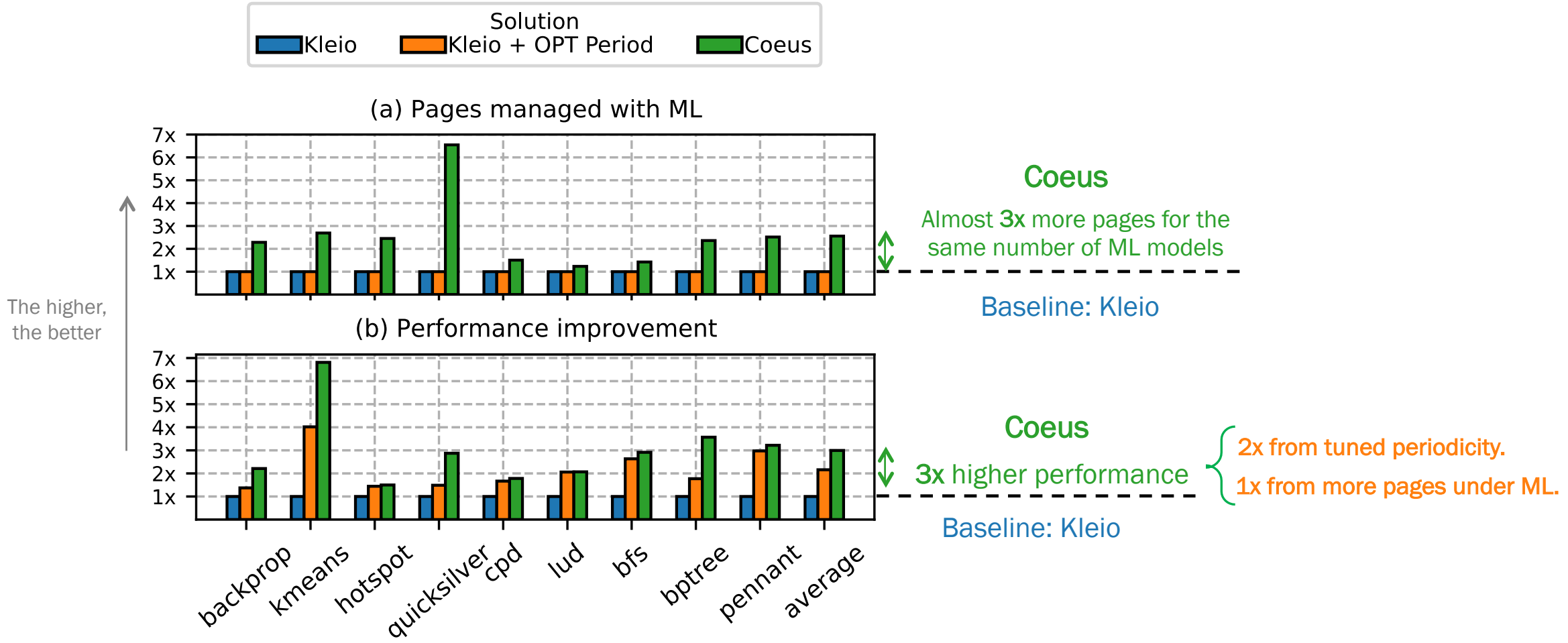
Cori is a lightweight tuning solution for hybrid memory page schedulers, that we will use to determine the “right granularity”.

System Design of *Coeus*

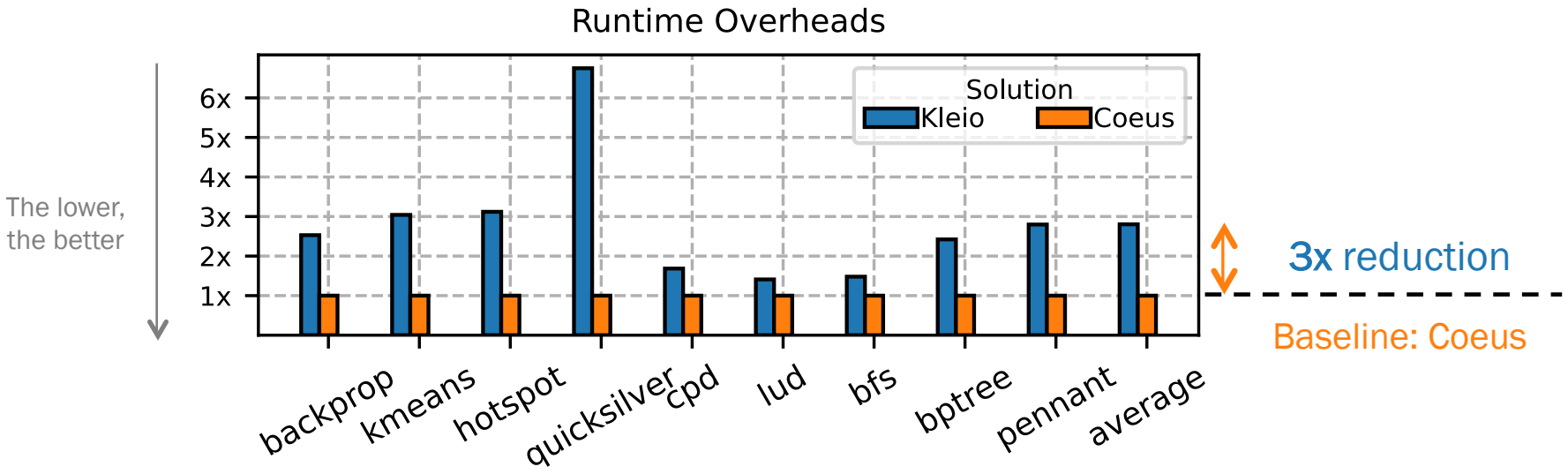
Coeus is a page grouping mechanism for machine intelligent page schedulers.



Scaling ML to More Pages and Improving Performance



Reducing Runtime Overheads of ML-based Management

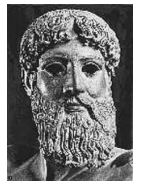


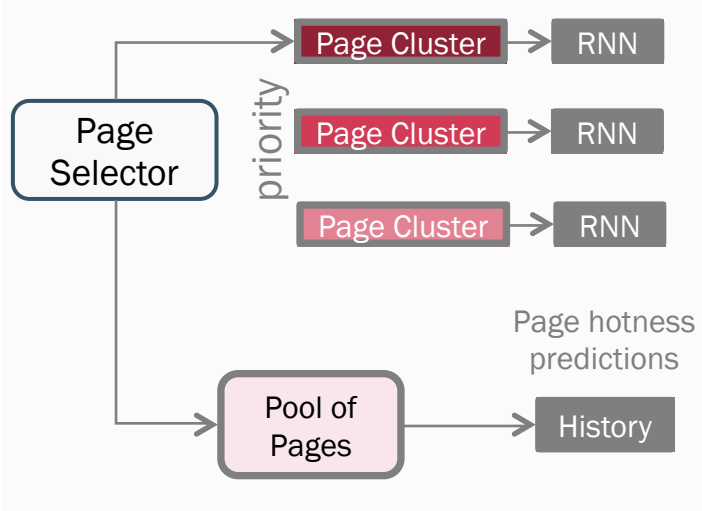
Coeus enables:

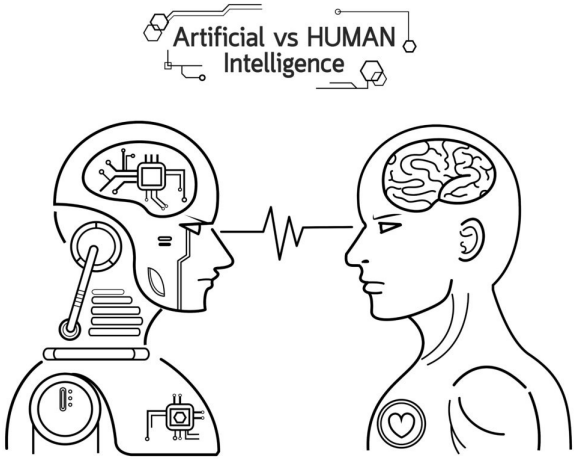
- ☑ Quick page clustering process.
- ☑ Fewer ML models deployed.

Summary of Coeus

Greek Trivia: According to the ancient Greek mythology, Coeus was the titan god of intelligence.







ML = Machine Learning.

Takeaways:

- Keep it Smart, but Simple.
  - ML is not always necessary.
- It is all about the right granularity.
  - For patterns and performance.
- Coeus reduces ML overheads by 3x.
  - Quick clustering.
  - Reduces total number of ML models deployed.
- Improves application performance by 3x.
  - From allowing more pages to be managed with ML.
  - From the tuned periodicity of the management.