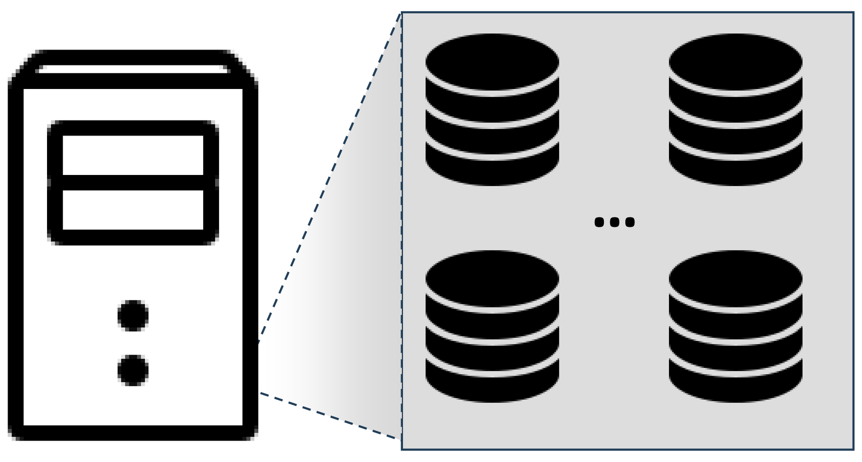
NOCAP: Near-Optimal Correlation-Aware Partitioning Joins

Zichen Zhu Xiao Hu Manos Athanassoulis

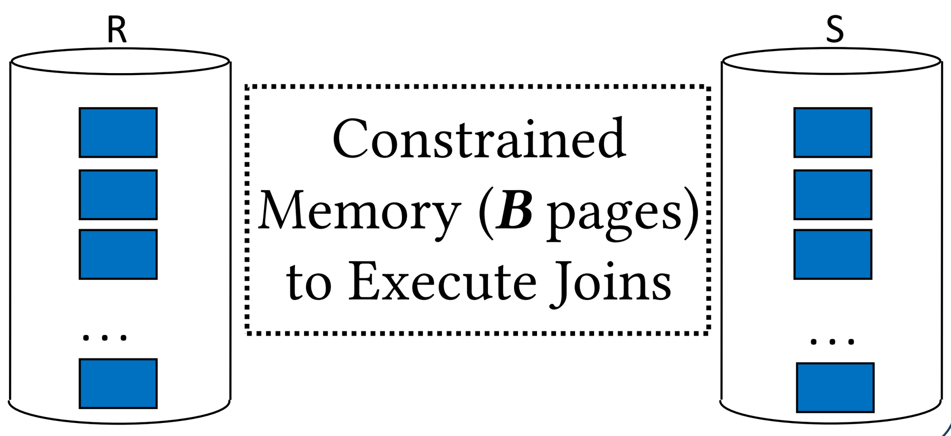
BOSTON UNIVERSITY

Storage-based Joins



Multiple DB instances in the cloud share the memory resources

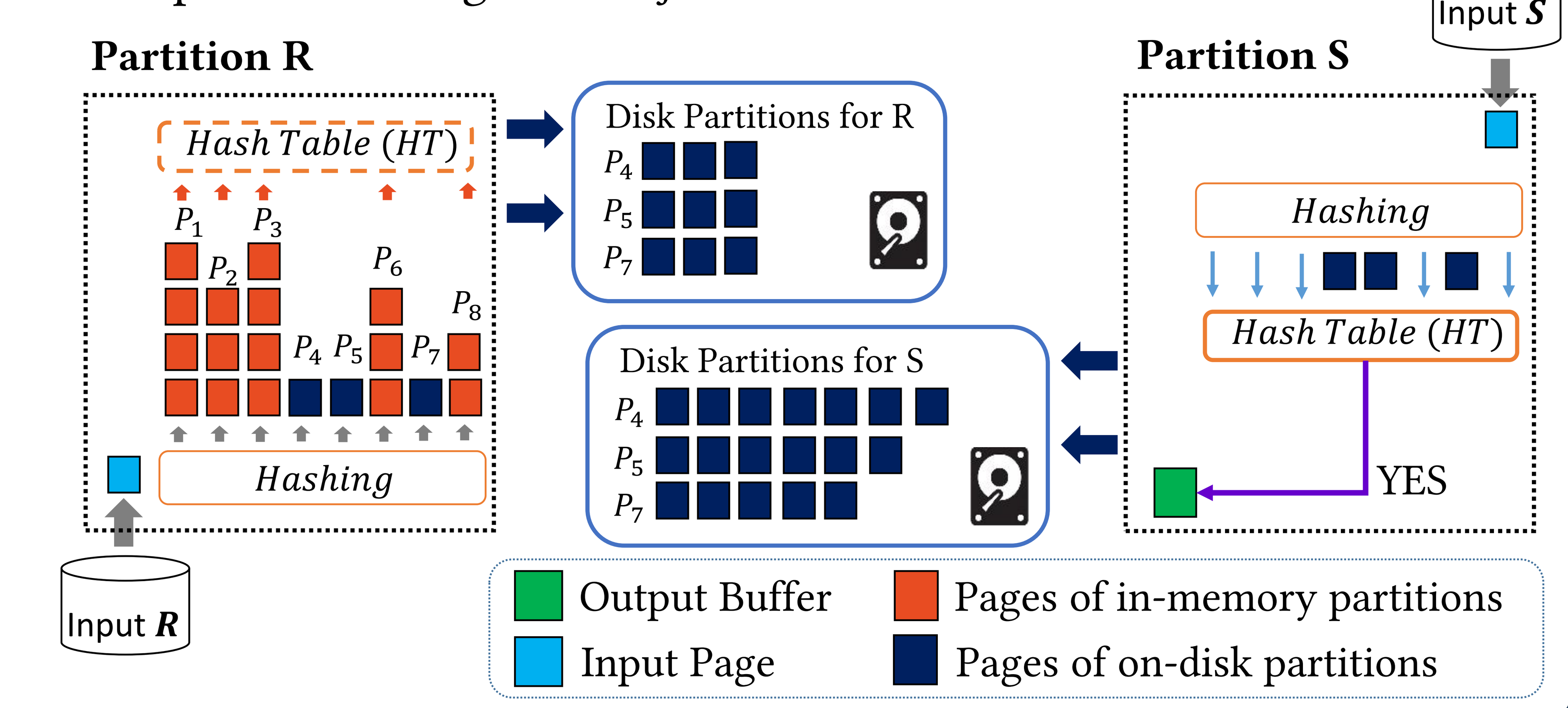
Tables may not fit in memory, which leads to storage-based joins.



Dynamic Hybrid Hash Join (DHH)

State-of-the-art DB engines use DHH to implement storage-based joins

AsterixDB PostgreSQL



Optimal Correlation-Aware Partitioning (OCAP)

An $n \times (m+1)$ *Boolean* matrix P that represents the partitioning assignment (***list-based partitioning***)

$$P = \begin{bmatrix} 0 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 0 \end{bmatrix}_{n\times(m+1)}$$

Notation	Meaning
n (n_R)	The number of tuples in relation R
m	The number of partitions on disk
R_1	A partition cached in memory
$Cost(\|\|R_j\|\|, \|\|S_j\|\|)$	The read and write cost of R_j, S_j plus the join cost between R_j, S_j

$$\arg\min_{P,m} \sum_{j=2}^{m+1} Cost(||R_j||, ||S_j||)$$

s.t. $\forall i \in [n],\ \sum_{j=1}^{m+1} P_{i,j} = 1$

$$||R_1|| + m + 2 \le B$$

$$P_{i,j} \in \{0,1\}, \forall i \in [n], \forall j \in [m+1]$$

Exponential searching space !

↓ *Consecutiveness* ↓ *Monotonicity* ↓ *Divisibility*

$\boldsymbol{O(n^2 \cdot \log B / B)}$ **for PK-FK joins**

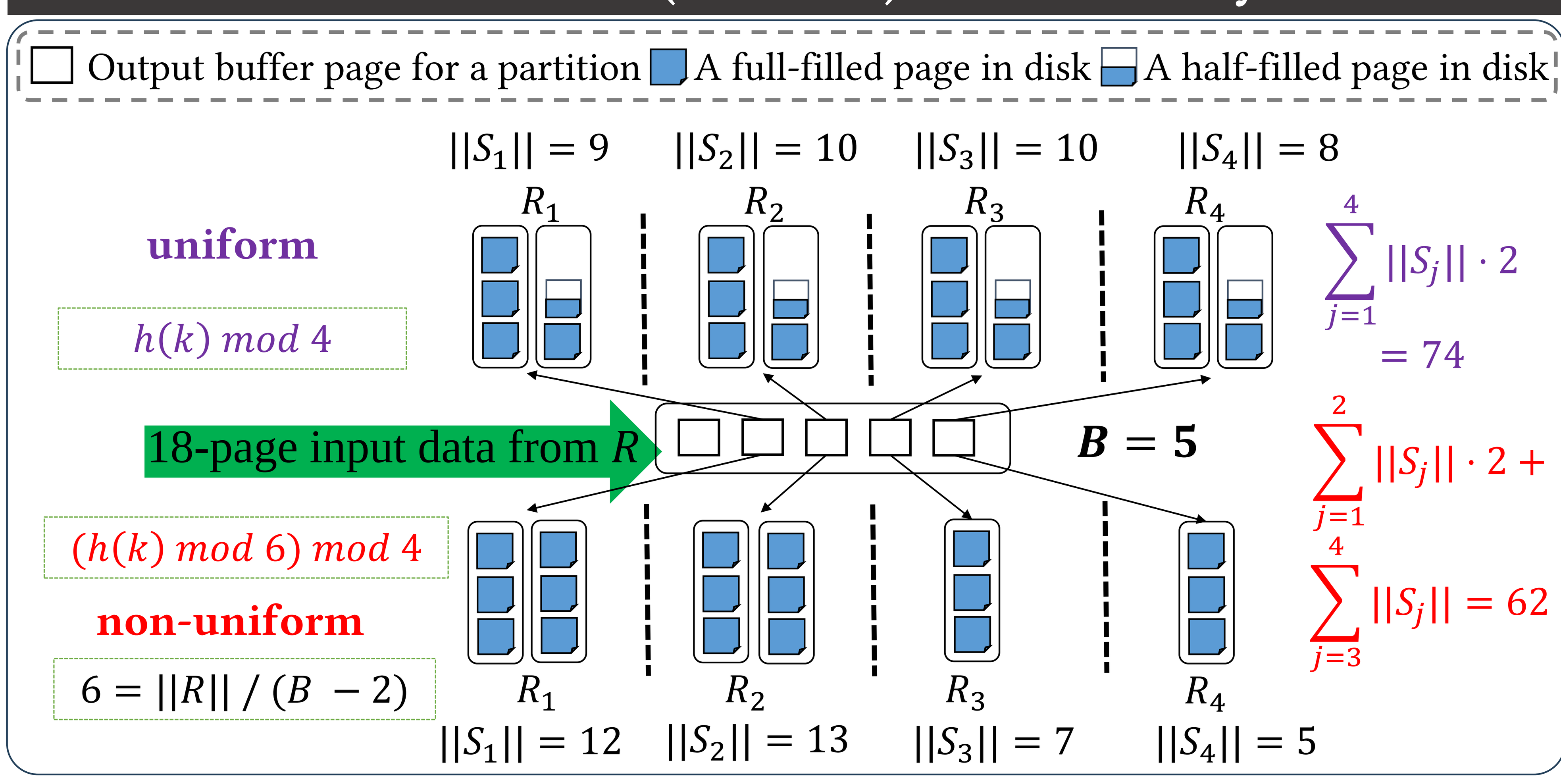
☹

Practical Challenges

(1) Limited matching info

(2) List-based partitioning also occupies memory

Near-OCAP (NOCAP) = Divisibility + Prioritizing MCVs with Constrained Memory



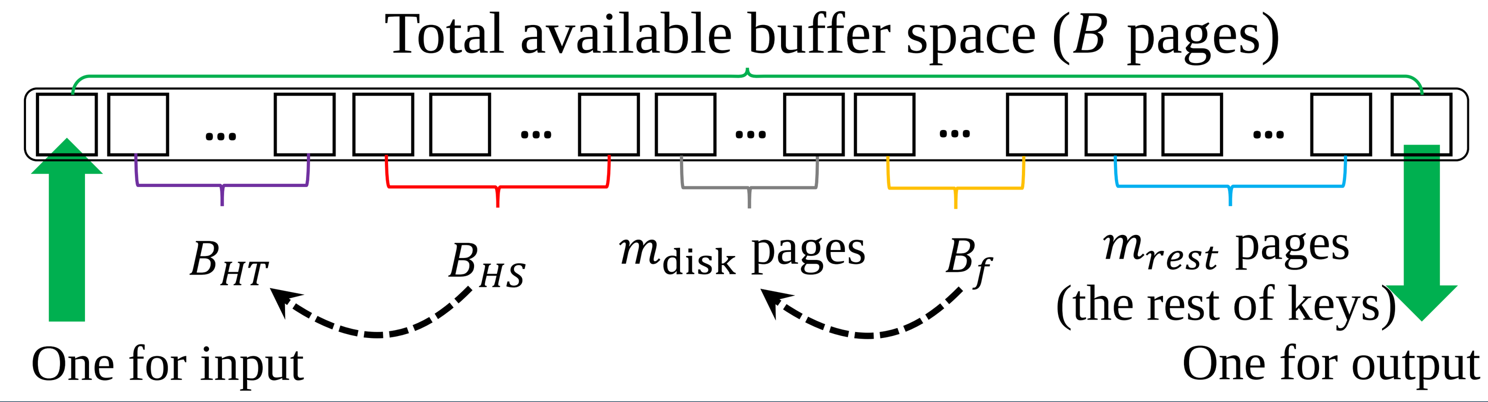
Index i	Frequency in S
1	1
...	...
n-k+1	73
...	...
n	100

We only have top-k frequent keys (MCVs)

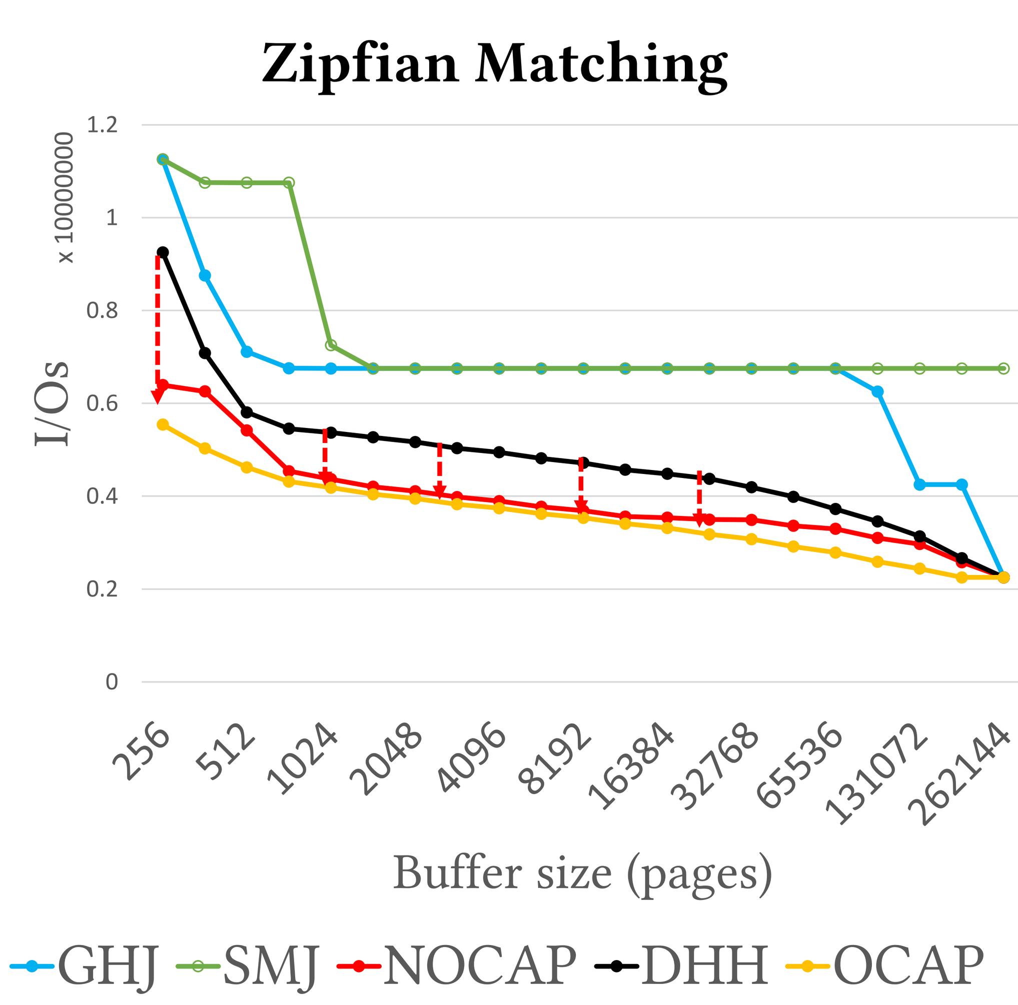
Run augmented DHH with memory budget m_{rest}

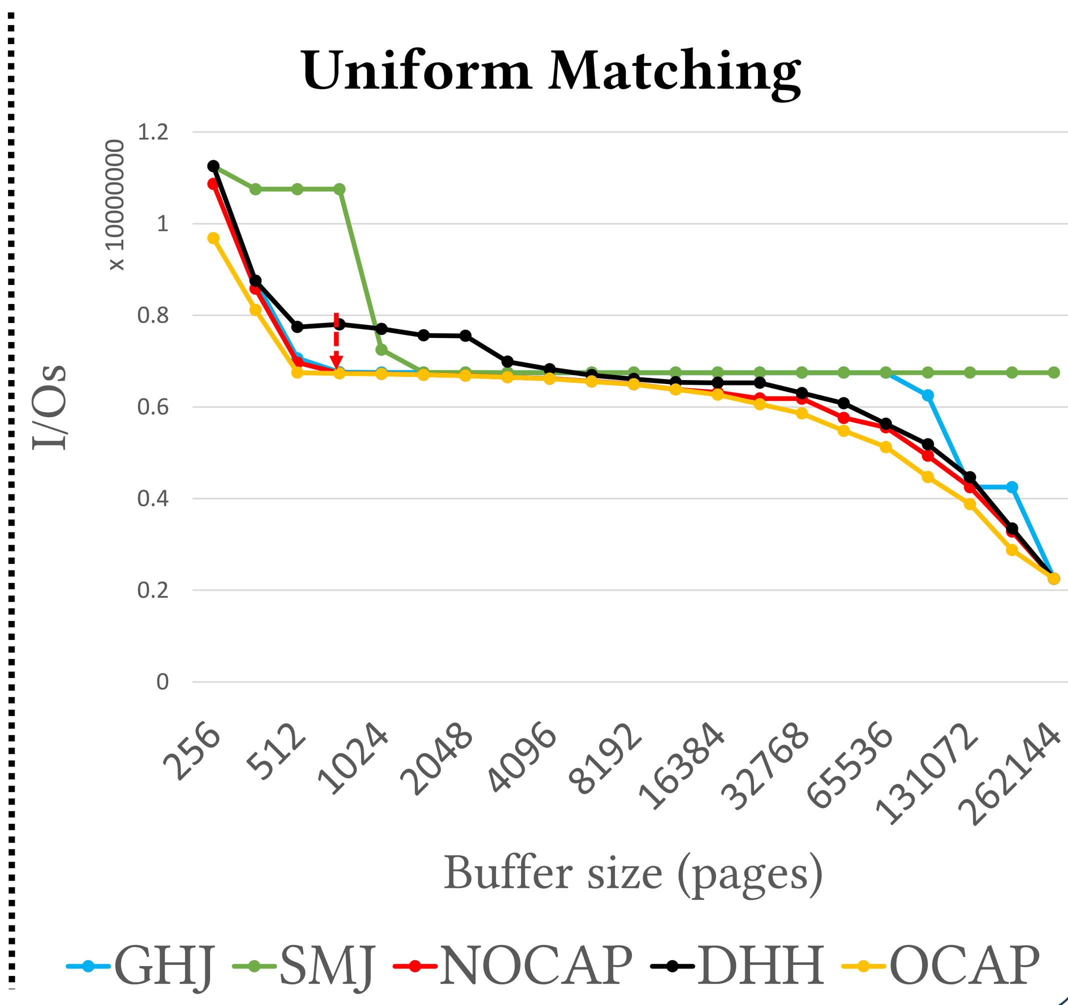
Run OCAP for top-k' MCVs with memory budget $m = B - 2 - m_{rest}$

Total available buffer space (B pages)



Experiments



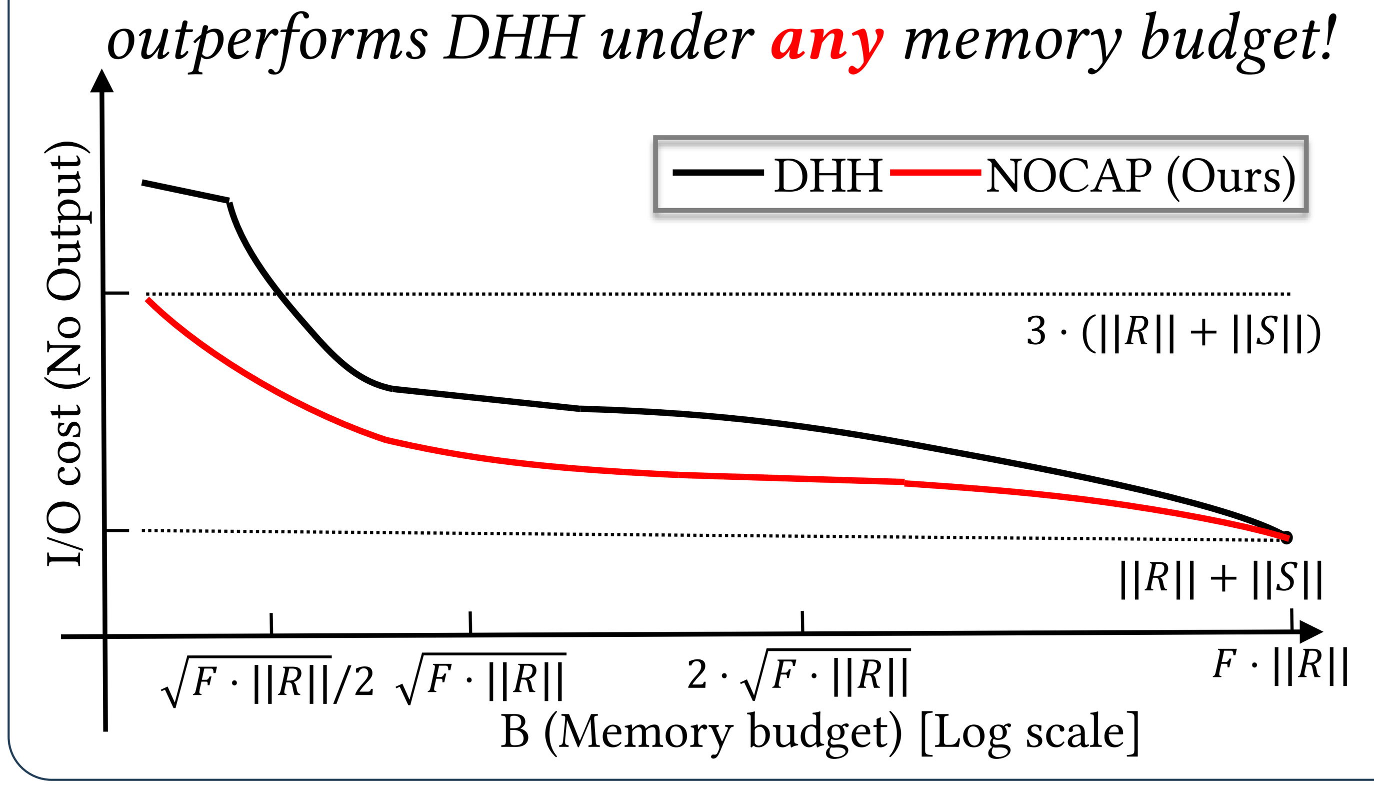


Summary of NOCAP

Up to 30% *improvement for skewed join correlation,*

Up to 10% *even for uniform case!*

outperforms DHH under ***any*** *memory budget!*



BOSTON UNIVERSITY · Meta · NSF · DiSC lab