Improving Query Execution on Object-based Computational Storage(OCS) System

Soon Hwang[1], Junhyeok Park[1], Junghyun Ryu[1], Jungahn Park[2], Jeongjin Lee[2], Jungki Noh[2], Soonyeal Yang[2], Woosuk Chung[2], Youngjae Kim[1]

1) Sogang University, 2) SK Hynix

The 9th International Parallel Data Systems Workshop

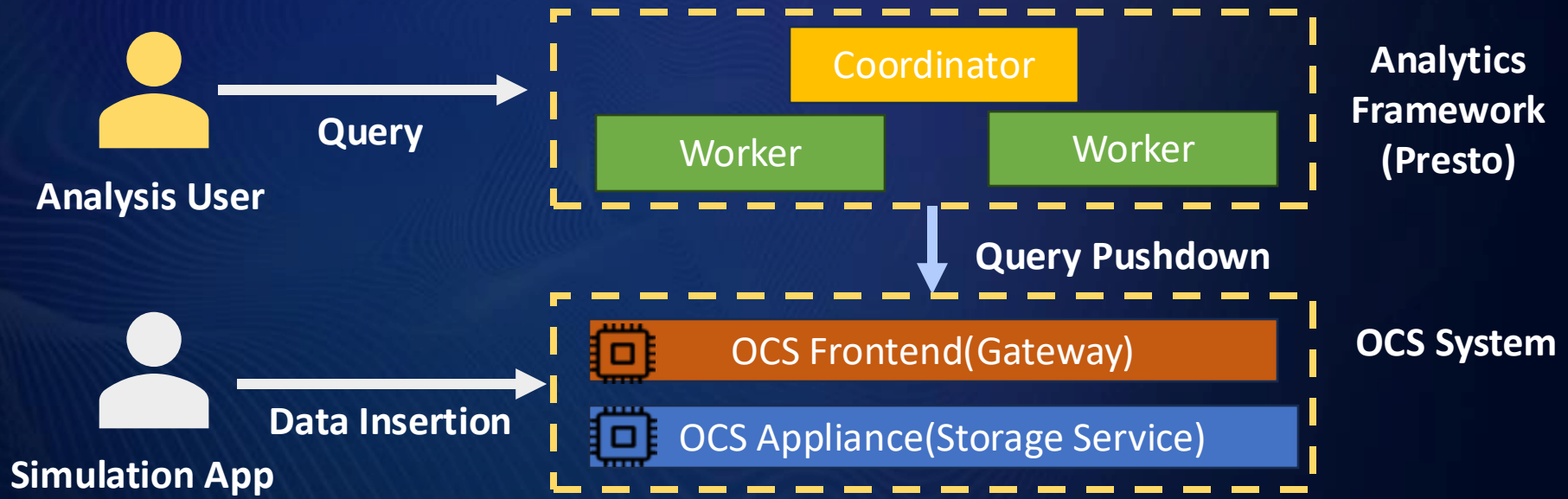
서강대학교
SOGANG UNIVERSITY

OCS System

Research Approach

Enabling vertical query pushdown in Object-based Computational Storage(OCS) system[1] within LANL analytics ecosystem[2]

- OCS system is composed of OCS Frontend(OCS-FE) and OCS Appliance(OCSA) with pushdown interface.
- Scientific query is vertically offloaded to the OCS system to reduce data movement.



[1] Kim, J. (2023). Accelerating Data Analytics Using Object Based Computational Storage in a HPC. Presented at the SC23 International Conference for High Performance Computing, Networking, Storage, and Analysis.
[2] Qing et al, Toward Standardized, Open Object-Based Computational Storage For Large-Scale Scientific Data Analytics, Work-In-Progress (WIP) Presenation in the 8th International Parallel Data Systems Workshop (PDSW) held in conjunction with SC23, Denver, CO, November 2023

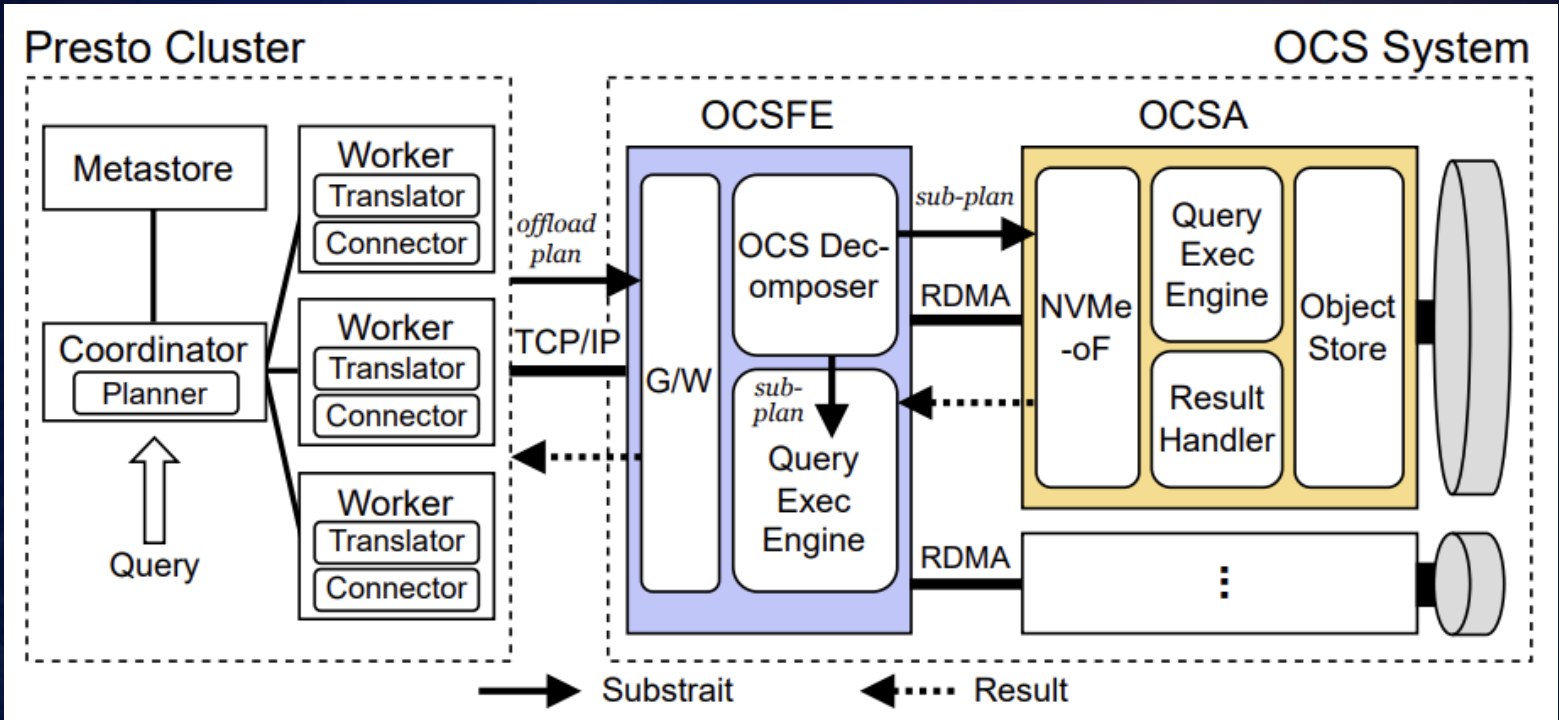
OCS System's Vertical Pushdown Objective

Objective

Utilize the rich resources in the OCS system with pushdown decision making, not simply pushdown query to the storage!

- Data transfer reduction is important since the size of data for analytics is getting bigger.
- The resources in the OCS system implies **the opportunity for vertical pushdown** not only **at the worker level to storage**, but also **inside the storage system level itself.**
- Q) Is the data movement the one and only important factor for additional pushdown point?
    1. Data transfer reduction is the first factor.
       ➔ OCS system does not allow round-trip of intermediate results between layers.
    2. Computation should be considered for pushdown point decision.

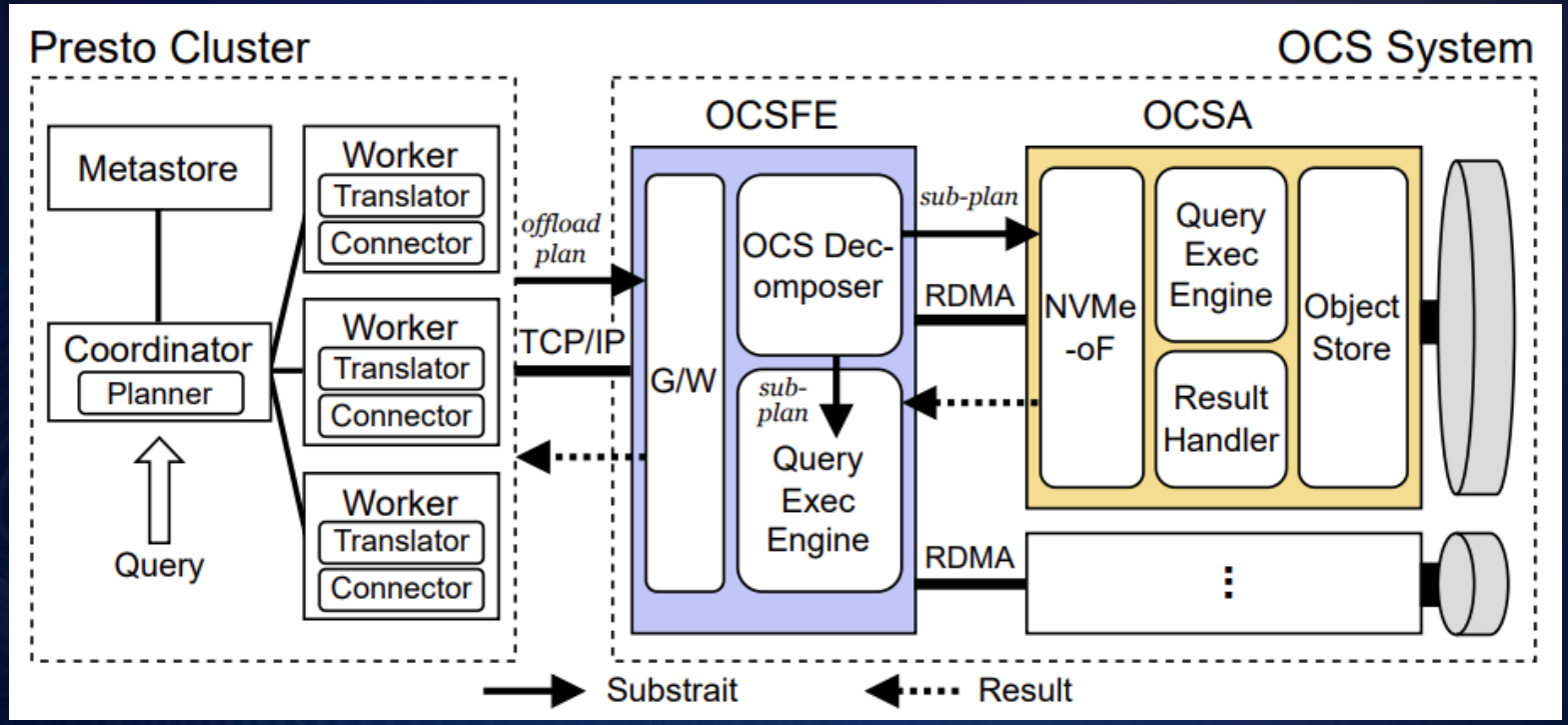
OCS System Design



Goal) Identify the optimal point for query pushdown **from the analytics framework to the OCS system**, and **within the OCS system itself**, taking into account both data reduction and computation

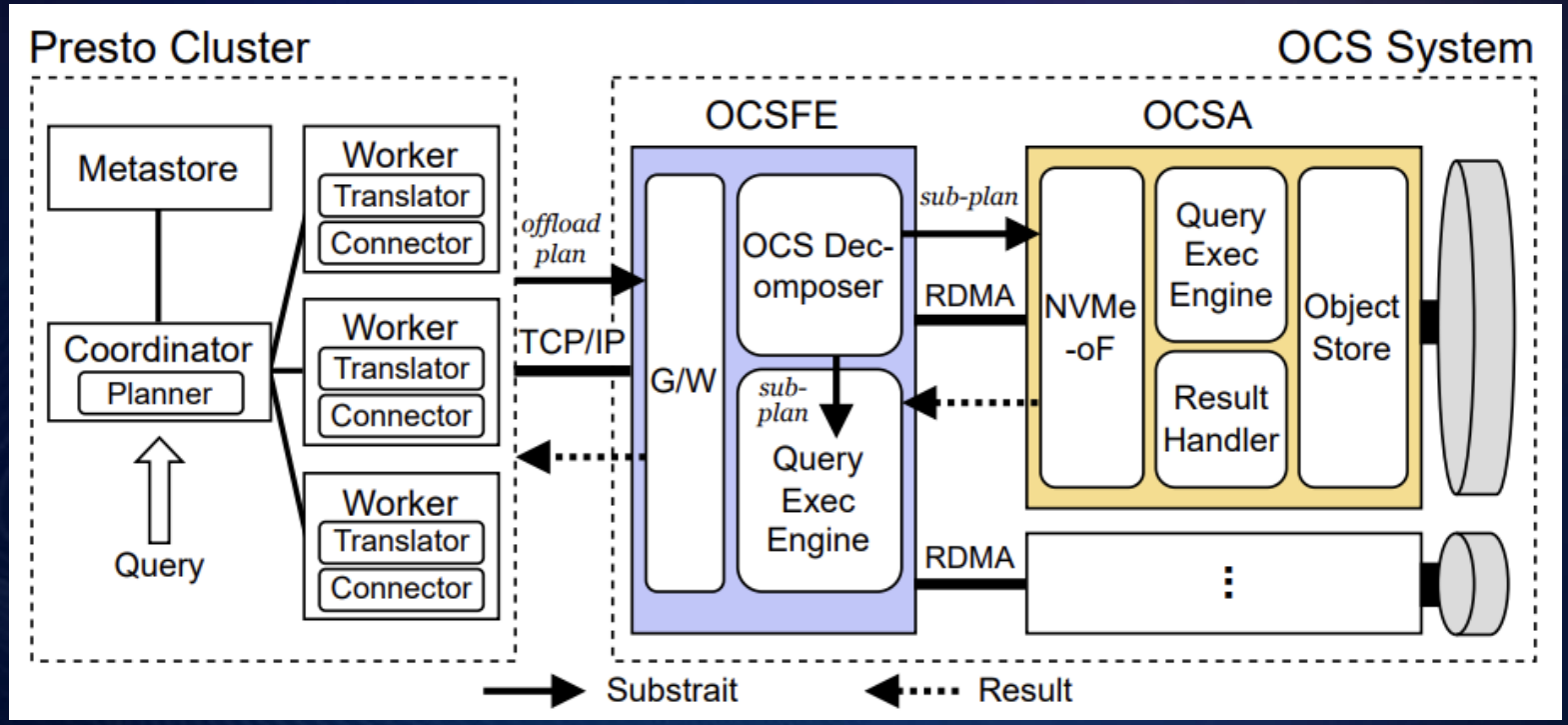
1. Vertical Pushdown Framework at OCS System level
2. Pushdown Decision Maker at Presto level

Vertical Pushdown Framework at OCS System



- Vertical Pushdown Framework at OCS System level
  1. Data Movement & Computation Cost-based Algorithm
  2. OCS Decomposer for decomposing plan into OCS-FE and OCSA sub-plan

Pushdown Decision Maker at Presto Level

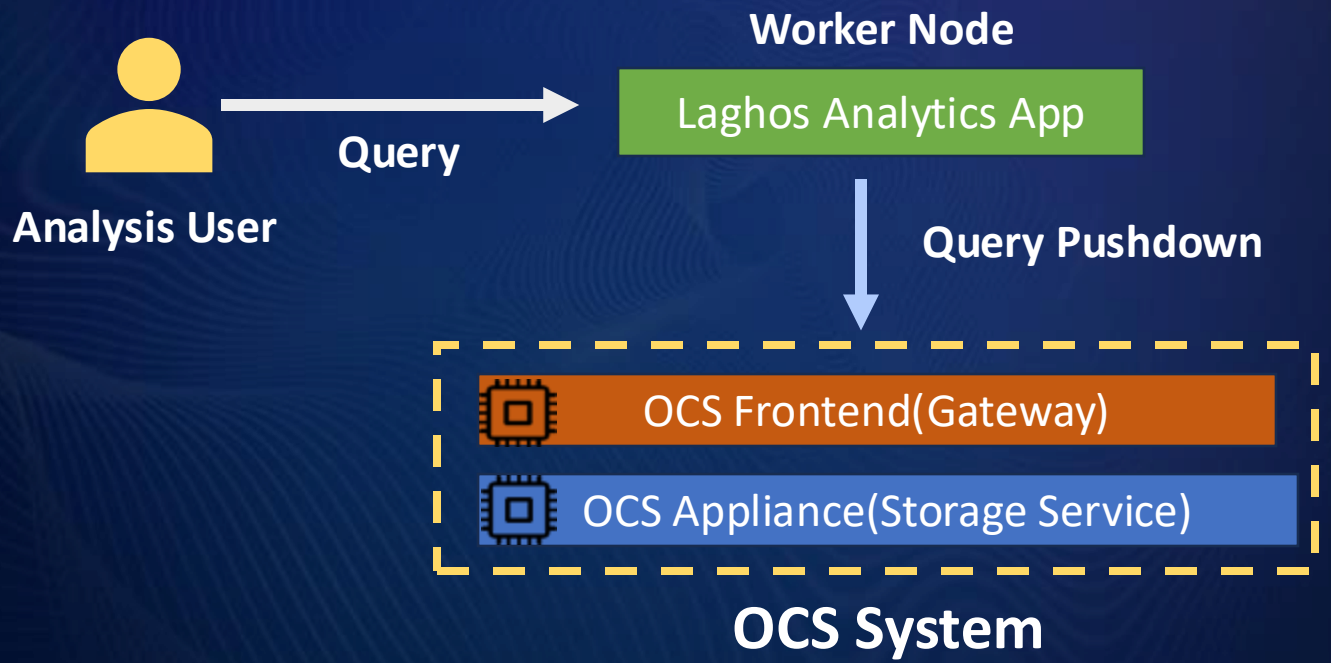


- Pushdown Decision Maker at Presto(Analytics platform) level
  1. Connector-level pushdown decision maker
  2. Connector-level Substrait translator

Preliminary Results

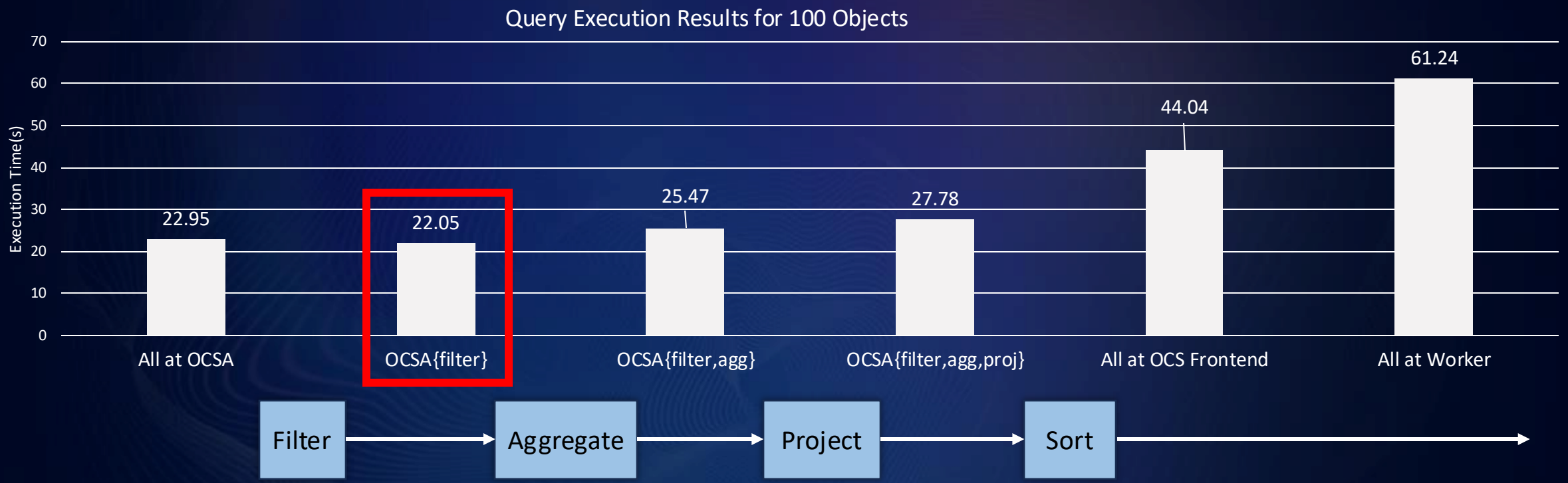
Laghos sample query and dataset[3] are used for evaluation**.**

Laghos[4] is gas dynamics miniapp with unstructured mesh, 2D and 3D elements.



[3]Lanl-Ocs, “Sample Laghos dataset for system prototyping and benchmarking,” GitHub. https://github.com/lanl-ocs/laghos-sample-dataset
[4] Ceed, “GitHub - CEED/Laghos: High-order Lagrangian Hydrodynamics Miniapp,” GitHub. https://github.com/CEED/Laghos

Preliminary Results



- Results with varying query decomposing point shows that vertically pushdown inside the OCS system can be effective.
- **Ongoing work**
  - **Evaluation**) Scientific queries with more computation cost
  - **Implementation**) Presto integration with the OCS system
    Reinforce cost-based algorithm

Learn more about SK hynix

Booth Location

Visit booth #703 to experience SK hynix products and demos