DIAS: Data-Intensive Applications and Systems Laboratory
School of Computer and Communication Sciences
Ecole Polytechnique Fédérale de Lausanne
Building BC, Station 14
CH-1015 Lausanne
URL: http://dias.epfl.ch/

Smarter Data Compression for Cloud Databases

Keywords	Disaggregated Architectures, Adaptive Data Compression, FlexPushdownDB, Apache IPC
Problem:	With the rise of cloud-based, disaggregated architectures, where storage and compute resources are separated, efficiently managing data movement across networked components has become increasingly critical. In such architectures, the network is often the primary bottleneck, which leads to inefficiencies in data-intensive applications. Strategies like caching, computation pushdown, and data compression have been employed to minimize data transfer, yet achieving an optimal balance remains a challenge. FlexPushdownDB [1-3] is a promising OLAP cloud DBMS prototype that addresses these challenges by dynamically balancing caching and computation pushdown, leveraging a unique hybrid query execution model. However, data compression has so far been statically enabled or disabled without consideration of runtime workload characteristics. This project's goal is to address this gap by implementing adaptive compression strategies for data transmission, dynamically adjusting compression configurations to reduce network traffic and improve overall performance.
Project:	The project will involve setting up and running FlexPushdownDB on AWS EC2 instances, following the reproducibility instructions provided by the authors. This implementation will focus specifically on the data transmission layer between compute nodes and storage nodes, which leverages the Apache Arrow ecosystem, including Apache Flight and Apache IPC, for efficient data exchange. Key tasks will include configuring and tuning Apache IPC (https://arrow.apache.org/docs/cpp/api/ipc.html), experimenting with different compression algorithms (e.g., LZ4, ZSTD), and evaluating their impact across various compression levels.The goal is to design and implement a simple, adaptive solution that dynamically selects the most effective compression strategy based on workload characteristics. This solution will be tested to assess its effect on both data transfer efficiency and query performance.
Plan	**Literature Review** ○ Review the FlexPushdownDB architecture and existing work on hybrid pushdown and caching strategies [1-3].

DIAS: Data-Intensive Applications and Systems Laboratory
School of Computer and Communication Sciences
Ecole Polytechnique Fédérale de Lausanne
Building BC, Station 14
CH-1015 Lausanne
URL: http://dias.epfl.ch/

- Study recent advancements in data compression within disaggregated architectures, particularly adaptive compression techniques used in high-performance storage systems [4].

System Setup and Initial Benchmarking

- Configure FlexPushdownDB on AWS EC2 instances.
- Run baseline tests on data transmission performance using static compression settings in Apache IPC.

Compression Strategy Development

- Implement and evaluate different compression algorithms (LZ4, ZSTD) with varying configurations in Apache IPC.
- Design and test an adaptive compression solution, based on workload patterns and transmission requirements.

Experimental Evaluation

- Measure the impact of adaptive compression on query performance and data transfer efficiency.
- Compare with baseline (static compression) to demonstrate network utilization and response time improvements.

Analysis and Reporting

- Document findings and analyze the effectiveness of adaptive compression strategies within FlexPushdownDB.
- Discuss implications for future research and optimizations in disaggregated architectures.

References

[1] Yifei Yang, Matt Youill, Matthew Woicik, Yizhou Liu, Xiangyao Yu, Marco Serafini, Ashraf Aboulnaga, Michael Stonebraker, *FlexPushdownDB: Hybrid Pushdown and Caching in a Cloud DBMS*, VLDB 2021.

[2] Yifei Yang, Hangdong Zhao, Xiangyao Yu, Paraschos Koutris, *Predicate Transfer: Efficient Pre-Filtering on Multi-Join Queries*, CIDR 2024.

[3] Yifei Yang, Y., Yu, X., Serafini, M. et al. *FlexPushdownDB: Rethinking Computation Pushdown for Cloud OLAP DBMSs.* The VLDB Journal 33, 1643–1670 (2024).

DIAS: Data-Intensive Applications and Systems Laboratory
School of Computer and Communication Sciences
Ecole Polytechnique Fédérale de Lausanne
Building BC, Station 14
CH-1015 Lausanne
URL: http://dias.epfl.ch/

Supervisor:	Prof. Anastasia Ailamaki, anastasia.ailamaki@epfl.ch
Responsible collaborator(s):	Antonio Boffa, antonio.boffa@epfl.ch
Duration:	6 months