



Research Internship at ESILV

Measuring the Environmental Impact of NoSQL Model Transformations on the Cloud

Jihane Mali, Nicolas Travers

Keywords: NoSQL, Distributed DB, Cost Model, Data Modeling

The development of NoSQL solutions on the Cloud is a complex task, and mastering all aspects of optimization is challenging. At the DVRC (the ESILV laboratory), we have developed an approach to help developers measure the impact of their solutions under different dimensions: time, financial, and environmental.

In the context of this research project, we aim to study the impact of data model transformations on environmental costs in the Cloud. A cost model was previously developed in the team, and we intend to use it to better understand the impact of these transformations and create a recommendation system to minimize the carbon footprint of a NoSQL solution.

The first objective of this project will be to familiarize oneself with the cost model and apply it to various benchmarks. The results will be analyzed to study correlations between data models, queries, and data evolution. In the second phase, these results will be used to design a recommendation system to propose data model generation with a low carbon footprint while respecting other dimensions (financial and time costs).

A potential evolution of the subject could involve studying schema evolution as the database state changes, especially in the context of polystores.

This internship includes:

- Study of a cost model specialized in carbon footprint measurement
- Behavioral analysis of NoSQL databases
- Development of an environmental recommendation model

Profile and expected skills

M2 level students (Master or Engineering Schools).

Distributed DB, Cost Model, Data Modeling, Carbon footprint measuring, MongoDB

Location

DVRC at ESILV at (École Supérieure d'Ingénieurs Léonard de Vinci ; Paris, la Défense).

Duration

6 months (from march or before - 900€/month).

Application

Send you CV, last grades (M1/M2), motivation letter and recommendation letters to:

- nicolas.travers@devinci.fr, jihane.mali@devinci.fr



Targeted publications:

Conference A*, Demonstration, Journal Q1

[Jihane Mali, Shohreh Ahvar, Faten Atigui, Ahmed Azough, Nicolas Travers: FACT-DM: A Framework for Automated Cost-Based Data Model Transformation. EDBT 2024: 822-825](#)

[Jihane Mali, Faten Atigui, Ahmed Azough, Nicolas Travers, Shohreh Ahvar: How to Optimize the Environmental Impact of Transformed NoSQL Schemas through a Multidimensional Cost Model? CoRR abs/2311.15406 \(2023\)](#)

[Jihane Mali, Shohreh Ahvar, Faten Atigui, Ahmed Azough, Nicolas Travers: A Global Model-Driven Denormalization Approach for Schema Migration. RCIS 2022: 529-545](#)

[Valerie Restat, Meike Klettke, Uta Störl: Towards an End-to-End Data Quality Optimizer. ICDEW 2024: 262-266](#)

[EL AHDAB, Léa, MEGDICHE, Imen, PENINOUE, André, TESTE, Olivier. "Unified Models and Framework for Querying Distributed Data Across Polystores". In 18th Research Challenges in Information Science \(RCIS\), 2024.](#)

Alotaibi, Rana, Damian Bursztyn, Alin Deutsch, Ioana Manolescu, and Stamatis Zampetakis (2019). ["Towards Scalable Hybrid Stores: Constraint-Based Rewriting to the Rescue"](#). In: *Proceedings of the 2019 International Conference on Management of Data, SIGMOD Conference 2019, Amsterdam, The Netherlands, June 30 - July 5, 2019*. ACM, pp. 1660–1677.