PhD Proposal

Towards a Generic Storage and an Adaptive Query Optimization for Astronomical Data Management at Scale

Proposed by: Professor Karine Zeitouni, co-advised by Laurent Yeh (Associate Professor)

Hosting Laboratory: DAVID - University of Versailles Saint-Quentin, Paris-Saclay University

Context

The amount of data daily collected or produced has dramatically exploded over the last decade. Exploiting such big data will benefit to many businesses and scientific areas, and is expected to have a great societal impact. In spite of the extraordinary advance of its supporting technology, Big Data remains a hot topic of database research. At the European level, the Commission outlined a new strategy on Big Data, supporting and accelerating the transition towards a data-driven economy in Europe. Exploratory analysis of these data is indeed crucial to enable scientists and practitioners to better understand their data and optimize various processes. Many research programs address the topic of Big Data in H2020 program. Applications to universe science are among the most demanding of this technology. At the European level, ESA, the European Space Agency, has launched several new programs for sky and earth surveying. The H2020 Cost action BIG-SKY-EARTH [1] investigated the challenges of Big Data in astronomy and earth observation. At the international level, LSST project is about to be launched in 2020, where a giant telescope will produce peta bytes of data. In France, the CNRS launched MASTODONS, an interdisciplinary program on Big data, which founded a Petasky collaborative project involving astronomer and data scientists. A working group MAESTRO in the GdR MADICS is an outgrowth of Petasky project.

Big Data has also captured a lot of interest in the research and academic programs, in particular at the University of Paris Saclay (UPSay). The Master “Datascale” of UPSay located at the University of Versailles Saint-Quentin (UVSQ) is part of this program.

In the framework of a co-funded PhD thesis by UVSQ and CNES, UVSQ/ADAM group has proposed and implemented a distributed framework ASTROIDE for efficient querying of astronomical surveys. Drawing on past experience, this thesis proposal aims at extending ASTROIDE framework towards more genericity, more functionality, and better adaptivity of the system. It has also been selected for a co-funding from the CNES.

Problem statement and objectives

The Big Data phenomenon is radically changing the terms of data management because it introduces new challenges due to the volume, the transfer speed (velocity), and the type (variety) of data. Today, parallel shared-nothing architectures using commodity machines have been established as the de facto standard in Big Data management, and map-reduce as a new programming paradigm. This paradigm has been adopted by the major solutions for big data in such a distributed environment, with some variants, among which we can cite Spark [6,8], Flink and Dask (dask.org).

Whatever is the technology used, efficient query processing of astronomical data leads to optimize the data storage and organization. Today, the most used formats in astronomy are FITS, HDF5, or simple csv/gzip, mainly for data exchange purpose. There disadvantage w.r.t. big surveys lies either in their complexity, in the lack of compacity and scalability, or their suitability to the Cloud environment. More importantly, they all require a significant over-cost to be loaded and used in the target framework / library, such as Spark or pandas. Nowadays, Parquet format [2], recommended by the Apache consortium, is becoming a de facto standard adopted by a large variety of Big Data tools, and NoSQL systems. Indeed, it is a compact columnar storage format, which is auto-descriptive, allows compression, data partitioning, and "rows group" indexation. However, there exists a gap between the astronomical standards and Parquet, as a matter of fact. One of the objectives of this PhD thesis is to fill this gap by proposing an optimized generic storage and exchange model for astronomical data in a distributed processing environment. The use of Parquet, or equivalent formats (e.g., kudu), will favor its adoption, since various systems use them as a native storage.

To this end, several issues must be addressed: How to partition these data across processing nodes? How to index these data? How to deal with updates? How to measure the performance in term of ingestion, access & filtering, updates, etc.? What is the impact of the parameters, and how to tune them?

An adapted data storage structure should optimize a system overload. This is guided by the response time, and throughput for a query load as well as the resources variation. A possible solution is to monitor the activity of the system, learn the performance behavior from previous execution traces, which allows optimizing the current or the next executions. This can be done by keeping track of the system dynamics, in term of execution performance, and the resource consumption. A significant decrease of performance may automatically trigger either data re-organization, such as partitioning, local indexing, caching, or the adaptation of resources allocation (number of executors, memory, and number of CPU per executor).

This raises several questions among which: What are the parameters to collect, and what are the performance metrics? How to establish a cost model and a distributed caching technique? Which access path to choose when different indexing and/or storage methods are possible? How to adapt to the data and the workload profiles? How to optimize complex queries, and mixed query and update workloads?

Hence, the objective of this PhD thesis is twofold. The first is to propose an optimized generic storage and exchange model for astronomical data, and to implement it under modern distributed processing environment. The second objective is to improve the overall system performance throughout its execution. At this end, we envision applying machine learning techniques based on the previous execution traces in order to optimize the current query execution.

Overview of the state of the art

One of the problems when dealing with scientific data is handling of native formats, while avoiding the need and the costs of transformation and loading [9,10,13]. For instance, in [9], the authors propose a view on HDF5 file format from SciDB without conversion. Within Paris-Saclay University, spark-fits [13] library allows the support of FITS format with Spark. These works focus on multidimensional array or image data, based on the traditional standards. Our objective is rather to propose a new generic format, tailored for the needs of astronomical queries, which adapts to modern big data processing systems.

Tuning storage has been studied in the past for traditional database systems [3]. But this problem raises a new dimension for Spatial Big Data systems for which we can distinguish approaches according to whether it's disk-oriented systems or main memory-oriented systems. For the first, the focus is targeted on logical data organization on disk [4] where they often use global and local index data structure. Global index allows to prune irrelevant partitions during the map task. In main-memory oriented systems, like GeoSpark [5] and ASTROIDE [12], the data organization is built on the fly during the data load. Both approaches aim at reducing the processing time, but often at the cost of additional storage space or partitioning time. This shows the importance of considering the overall system performance and balance the query optimization with the resources consumption. However, the task of parameters tuning becomes complex, as their number increases, which is the case of these systems.

Beyond tuning tools, self-tuning has emerged recently to cope with the problem of the complexity of the tuning task by the administrators. For instance, *TuneIn* is an auto tuning framework developed by *LinkedIn* on top of Dr. Elephant, a performance monitoring and tuning tool for *Hadoop* and Spark [14], mainly based on a cost function, and not fully automatic. But the most promising direction of research is leveraging advanced machine learning algorithms for complex query optimization [15,16,17]. In this thesis, we will investigate these approaches and adapt them to the context of astronomical queries.

Application scenario and dataset

The Gaia space mission, launched by the European Space Agency in the end of 2013, will ultimately provide an unprecedented census of our galaxy in size, scope, and accuracy, encompassing astrometry for over one billion objects in the sky. The main mission goal for the Gaia archive is to provide a comprehensive repository of the rich data products to be generated by Gaia, and a range of access mechanisms and associated helper applications to enable effective access to the Gaia data by the end user science community. The Gaia mission will introduce several challenges over current technologies of data management, mainly due to the unprecedented amount of data that will be produced (around 1 PB). In this context, basic operations like data retrieval, analytic or visualization are becoming increasingly difficult and, in many cases, almost impossible. Simple database queries can now return results so big that they are incomprehensible — slow to handle, extremely hard to analyze and impossible to visualize with available tools.

This thesis will exploit the catalogs that are used and/or will be produced within GAIA mission, as well as other catalogs such as EUCLID, LISA, etc.

Work Plan

- 1st Semester: State of the art and familiarization with the corpus and tools
- 2nd Semester: Survey Report and first proposal
- 2nd Year: Refinement and validation, 1st paper or demo submission
- 3rd year: Publication and writing of the manuscript

Collaboration

This joint work is led by DAVID Lab and the CNES DNO/SC and DNO/ISA structures. The CNES co-funds this PhD scholarship. It will involve external collaborations with:

- the Observatory of Paris / GEPI group - http://gepi.obspm.fr
- Strasbourg astronomical Data Center (CDS) - http://cds.u-strasbg.fr
- The Laboratory of the Linear Accelerator (LAL) in Paris-Saclay University- https://www.lal.in2p3.fr

The PhD candidate can also benefit from the exchange program in the framework of the H2020 project MASTER (http://www.master-project-h2020.eu), and visit other groups among the partners.

References:

1. Big Data Era in Sky and Earth Observation (BIG-SKY-EARTH): https://bigskyearth.eu
2. Parquet : https://en.wikipedia.org/wiki/Apache_Parquet
3. Database tuning: Principles, Experiments, and Troubleshooting Techniques. Dennis Shasha, Philippe Bonnet. Morgan Kaufmann. 2002.
4. Aji, Ablimit et al. (2013a). Hadoop GIS: A High Performance Spatial Data Warehousing System over Mapreduce. In:Proc. VLDB Endow.6.11, pp. 1009–1020.
5. Yu, Jia, Jinxuan Wu, and Mohamed Sarwat (2015). GeoSpark: A Cluster Computing Framework for Processing Large-scale Spatial Data. In: Proceedings of SIGSPATIAL ’15. ACM, 70:1–70:4
6. NoSQL Databases, http://nosql-database.org/, accessed Mai 29, 2018

7. Brahem, Mariem and Zeitouni, Karine and Yeh, Laurent, HX-MATCH: In-Memory Cross-Matching Algorithm for Astronomical Big Data, International Symposium on Spatial and Temporal Databases (SSTD), pp 411–415, 2017, Springer

8. Spark SQL, https://spark.apache.org/sql/, accessed October 6, 2018

9. Xing, H., Floratos, S., Blanas, S., Byna, S., Wu, K., & Brown, P. ArrayBridge: Interweaving declarative array processing in SciDB with imperative HDF5-based programs, IEEE ICDE 2018.10. D. Han, Y.-M. Nam, and M.-S. Kim. A Distributed In-Situ Analysis Method for Large-Scale Scientific Data. In BigComp 2017.

11. Brahem, Mariem and Zeitouni, Karine and Yeh, Laurent, ASTROIDE: A Unified Astronomical Big Data Processing Engine over Spark, IEEE Transactions on Big Data, October 2018.

12. ASTROIDE Project: https://github.com/MBrahem/ASRTOIDE

13. Spark-fits: https://astrolabsoftware.github.io/spark-fits/24.

14. Manoj Kumar, and Arpan Agrawal, TuneIn: How to Get Your Hadoop/Spark Jobs Tuned While You're Sleeping, SPARK+AI SUMMIT 2018.

15. Krishnan, Sanjay, et al. Learning to optimize join queries with deep reinforcement learning. arXiv preprint arXiv:1808.03196(2018).

16. Kraska, Tim, et al. The case for learned index structures. Proceedings of the 2018 International Conference on Management of Data. ACM, 2018.

17. Ortiz, Jennifer, et al. Learning State Representations for Query Optimization with Deep Reinforcement Learning. Proceedings of the Second Workshop on Data Management for End-To-End Machine Learning. ACM, 2018.

Application

Please submit your application including:
- cover letter
- CV
- copies of the relevant certificates
- the academic transcripts of the 2 last years
- list of references
- any complementary document: recommendation letters, relevant publications if exist.

On the website of the doctoral School. See:
https://www.universite-paris-saclay.fr/en/education/doctorate/sciences-et-technologies-de-linformation-et-de-la-communication-stic-0#the-doctorate

and in parallel by email to:

Prof. Karine Zeitouni
DAVID Lab - University of Versailles Saint-Quentin
Paris Saclay University
www.david.uvsq.fr/zeitouni
www.universite-paris-saclay.fr
E-mail: Karine.Zeitouni@uvsq.fr