ORACLE

Intern with Oracle Labs PGX!

The PGX team at Oracle Labs focuses on high-performance shared-memory and distributed graph processing and has open internship positions available.

Oracle

Oracle, a global provider of enterprise cloud computing, is empowering businesses of all sizes on their journey of digital transformation. Oracle Cloud provides leading-edge capabilities in software as a service, platform as a service, infrastructure as a service, and data as a service.

Oracle's application suites, platforms, and infrastructure leverage both the latest technologies and emerging ones – including artificial intelligence, machine learning, blockchain, and Internet of Things – in ways that create business differentiation and advantage for customers. Continued technological advances are always on the horizon.

Oracle Labs

Oracle Labs is the advanced research and development arm of Oracle. We focus on the development of technologies that keep Oracle at the forefront of the computer industry. Oracle Labs researchers look for novel approaches and methodologies, often taking on projects with high risk or uncertainty, or that are difficult to tackle within a product-development organization. Oracle Labs research is focused on real-world outcomes: our researchers aim to develop technologies that will someday play a significant role in the evolution of technology and society. For example, chip multithreading and the Java programming language grew out of work done in Oracle Labs.

Parallel Graph AnalytiX (PGX)

PGX [1] is a toolkit for graph analytics that supports *graph algorithms,* such as PageRank, *graph queries* with PGQL (an SQL-like graph query language), and *graph ML*. PGX includes both a single-machine in-memory engine and a distributed engine for very large graphs and is already available as an option in Oracle products and an active research project at Oracle Labs.

ORACLE

Internship Details

The goal of this project is to extend PGX, both the single-machine runtime (PGX.SM [2]) and the distributed runtime (PGX.D [3]) with new capabilities. We offer various topics depending on the skills and the interests of the candidate (topics are not limited to the ones below; see also the "Related Topics" sub-section below):

Extended distributed computations

PGX.D implements distributed PGQL [4] queries using an asynchronous depth-first runtime [5-6]. In this project, we will generalize and explore how to leverage, and possibly extend, this asynchronous depth-first runtime to support a broader scope of computations.

Distributed fault tolerance & graph snapshots

Fault tolerance in data-analytics systems often relies on techniques such as snapshots (storing the data in persistent storage) and replication. In this project, we will explore various options for enhancing distributed fault tolerance for PGX.D, including snapshots and replication.

Distributed shortest/cheapest paths

PGQL enables users to query graphs for (top K) shortest or cheapest paths (in combination with all other patterns offered in PGQL). In this project, we will design and implement such functionalities in a distributed setting.

Distributed scalable engine for graph-based ML

Recent research shows that machine learning workloads can benefit from information encoded in the graph to achieve higher accuracy and faster convergence when learning models. In this project, we will explore, given the distributed nature of the graph, how it is possible to retrieve embeddings for ML algorithms from such distributed graphs efficiently for processing in external ML frameworks.

Extension of an SQL-like graph query processing engine (PGQL)

In this project, we will extend the semantics and implementation of the PGQL graph query language. Example topics include: (i) improving the composability of PGQL queries (i.e., starting a PGQL query from the results of a previous one, or from graph algorithms results) and optimizing the execution of such composed queries, and (ii) designing and implementing pipelined versions of the PGQL operators to reduce the peak memory consumption during query.
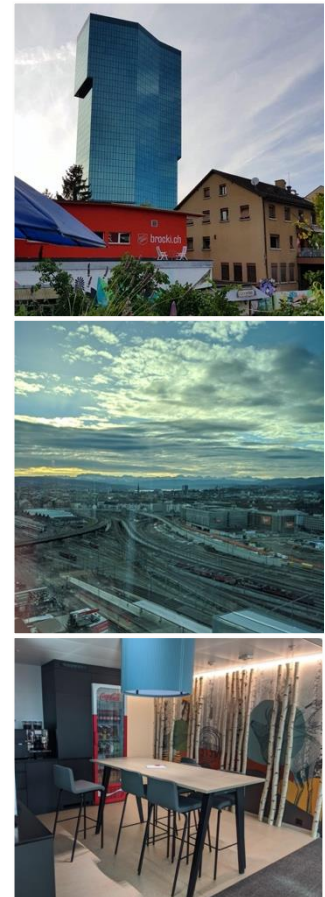

Image Caption 1. Our beautiful offices at Prime Tower, Zurich.

"Working at Oracle Labs is a great experience that I recommend to any student in computer science. It has been one of the best parts of my studies, as I was able to research, design and implement new features in a state-of-the-art distributed graph processing engine. Furthermore, the diversity of the people working at Oracle Labs, as well

ORACLE

Dynamic data loading for very large graphs

Main memory is a limited resource. Consequently, in a data-analytics engine, such as PGX.SM, only the most recent or most important data should can be kept in memory, and other data can be offloaded to external storage/systems. During this internship, we will extend PGX.SM support of dynamically loading of data that is present in offloaded systems, in a graceful, efficient and transparent manner.

References

[1] Parallel Graph AnalytiX (PGX), http://www.oracle.com/technetwork/oracle-labs/parallel-graph-analytix/overview/index.html
[2] Using domain-specific languages for analytic graph databases, VLDB'16
[3] PGX.D: A fast distributed graph processing engine. SC'15
[4] PGQL: Property Graph Query Language, http://pgql-lang.org
[5] PGX.D/Async: A Scalable Distributed Graph Pattern Matching Engine. GRADES'17
[6] aDFS: An Almost Depth-First-Search Distributed Graph-Querying System. USENIX ATC'21

Required Skills

The successful candidate is expected to complete the internship using a wide and diverse set of skills.

- Basic understanding of parallel, concurrent, and distributed programming (having completed relevant courses, such as Distributed / Concurrent algorithms, is a plus);
- For PGX.D: C/C++ programming skills. Experience with Java is a plus;
- For PGX.SM: Java programming skills;
- Ability to design and implement reliable and documented high-performance software, including tests;
- Good problem-solving skills;
- Experience with Linux (e.g., bash scripts);
- Familiarity with graph algorithms is a plus;
- A high average grade in master studies is a plus.

Internship Facts

The duration of the internship can vary based on the candidate's constraints. The usual duration is 6 months. We pay a competitive salary. The research topics listed are informative, we are open to suggestions depending on your skills and qualifications. By sending in your application, you opt-in for processing your personal information.

If this sounds like you, we hope to meet you!

as its great working atmosphere, made my internship even more enjoyable."

<Arnaud Delamare>
EPFL student, now employed by Oracle Labs, 6-months intern

ORACLE

How to Apply

After reviewing the Oracle Labs internship program, please provide us with the following information to apply:

- Your CV or link to your home page containing your curriculum
- Description of your motivation and area(s) of interest
- Availability and preferred internship duration
- Preferred location

If available, please apply to this topic via the internship portal of your university. Alternatively, send an email with the aforementioned information to graphs-labs-hiring_ww@oracle.com

About Us

Innovation starts with inclusion at Oracle. We are committed to creating a workplace where all kinds of people can be themselves and do their best work. It's when everyone's voice is heard and valued, that we are inspired to go beyond what's been done before. That's why we need people with diverse backgrounds, beliefs, and abilities to help us create the future, and are proud to be an affirmative-action equal opportunity employer.

Oracle is an Equal Employment Opportunity Employer. All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, national origin, sexual orientation, gender identity, disability and protected veterans status, age, or any other characteristic protected by law. Oracle will consider for employment qualified applicants with arrest and conviction records pursuant to applicable law.

Life at Oracle

An Oracle career can span industries, roles, countries and cultures, giving you the opportunity to tackle new roles and challenges, while blending work and life. Oracle has thrived through 40+ years of change by innovating and operating with integrity while delivering for the top companies in almost every industry. To nurture the talent that makes this happen, we work hard to build a vibrant and inspiring workplace that celebrates diverse, hardworking teams where everyone can contribute. We take care of each other, and value giving back to the community. We have flexible work arrangements and offer benefits including generous paid parental leave and comprehensive premium medical insurance.

We will ensure that individuals with disabilities are provided reasonable accommodation to participate in the job application or interview process, to perform crucial job functions, and to receive other benefits and privileges of employment. Please contact us to request accommodation.

At Oracle, we don't just respect differences—we celebrate them. We believe that innovation starts with inclusion and to create the future we need people with diverse backgrounds, perspectives, and abilities. That's why we're committed to creating a workplace where all kinds of people can do their best work. It's when everyone's voice is heard and valued that we're inspired to go beyond what's been done before.

More information on Oracle's stance on diversity and inclusion

ORACLE