Oracle and Oracle Labs

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 invests heavily in research and development: US$6.2 billion in FY 2017. 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 is the only organization at Oracle that is devoted exclusively to research.

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)

Relationships in the data are becoming a key feature to enable knowledge discovery from large datasets. Graphs are a powerful abstraction to support this analysis, thanks to their explicit representation of relationships as edges. Graph analysis lets you reveal latent information that is encoded, not as fields in your data, but as direct and indirect relationships between elements of your data – information that is not obvious to the naked eye, but can have tremendous value once uncovered.
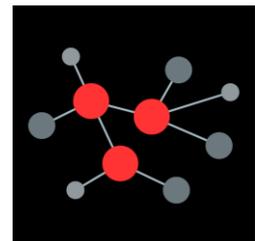
PGX is a toolkit for graph analysis that supports both *running algorithms* such as PageRank on graphs, and performing SQL-like *pattern-matching* on graphs, using the results of algorithmic analysis. Algorithms are parallelized for extreme performance. The PGX toolkit includes both a single-node in-memory engine, and a distributed engine for extremely large graphs. Graphs can be loaded from a variety of sources including flat files, SQL and NoSQL databases and Apache Spark and Hadoop; incremental updates are supported.

PGX is both already available as an option in Oracle products and an active research project at Oracle Labs, with a world-class team of researchers further advancing the capabilities of the toolkit. The PGX team has various openings for internships on themes including:

Software language engineering: Extend and improve our domain-specific languages for graph algorithms (Green-Marl) and graph queries (PGQL), including syntactic and static analysis, optimization, code generation, and high-performance interpretation.

Distributed graph analytics: Distributed graph algorithms and query execution in a cloud environment. Improve resiliency and fault tolerance in a cloud environment.

Heterogeneous graphs: Optimize in-memory graph representation to reduce memory footprint and to improve performance.

For more information, contact Tomas Faltin (tomas.faltin@oracle.com).