Software Engineering Internship at Actian Amsterdam

We are an R&D team based in Amsterdam working on Actian Vector and VectorH products – an analytical database system, that delivers cutting edge performance through constant innovation. Every day we’re confronted with all kinds of interesting and challenging problems that range from algorithm design and database engine design to parallel programming, distributed systems, code optimization, and integration with various Hadoop services.

We are now looking for interns, who would like to work on their MSc projects with us.

MSc project topics

1. **Automatic calibration**
   Automate the process of choosing the right DBMS configuration for a given hardware. A set of metrics (e.g., I/O and network speed) would be computed using a suite of micro-benchmarks. Then, a special algorithm would recommend the right settings, that would help the us choose the best query execution strategies.

2. **Load-balanced query execution on a cluster**
   Optimally utilize all the cluster nodes during query execution by dynamically shifting the work between machines. Exploit data partitioning strategies and HDFS replication mechanism to balance the load while maintaining the data locality.

3. **Compact hash-tables**
   The memory footprint of hash-tables used for executing joins and aggregations could be significantly reduced, which would result in enhanced performance. Create algorithm to determine the best layout using the metadata and modify the hash-table implementation to work with compact data format.

4. **HDFS-over-S3 hybrid storage layer**
   Enhance our Amazon Web Services support by implementing a hybrid solution that would combine limited in size HDFS storage with a larger and more cost-effective Amazon S3. The goal is to develop a mechanism that is highly reliable and has minimal impact on query performance.

Let us know if you’d like to learn more about any of these topics, or to discuss other alternatives.

Many of us completed our MScs projects at Actian in previous years. Majority of that research was published, presented at conferences, and put into production. We’ll be happy to tell you about our experiences.

Your tasks

- Get familiar with the system architecture
- Research available related publications and explore different possible solutions
- Design and implement your solution
- Experiment and benchmark
- Document the research in the MSc thesis

Requirements

- Good C/C++ coding and debugging skills
- Familiarity with Linux operating system
- Ability to use scripting languages (Bash, Python, etc.) to create benchmarks and tests
- Interest in database engines, distributed systems, parallel programming paradigms, high-performance computing and/or code optimization
- Sharp mind and an ability to work and learn independently
- Previous experience is a big plus

Perks

- Having prof. dr. Peter Boncz, one of the world’s top researchers in the field, as your MSc supervisor
- Working in friendly and supportive atmosphere, with young and creative team
- Free lunches twice a week, drinks and snacks at will
- Lots of Lego blocks and beanbags for your enjoyment
- Opportunity to join company outings and team-building activities (e.g., carting, cooking, escape rooms)
- Start date and schedule agreed on individually
- Compensation: 1800 € / month full-time

Contact us

- Iris Hesp – iris.hesp@actian.com
- Peter Boncz – p.boncz@cwi.nl

Office location: Science Park 404, Amsterdam