KairosDB

InfluxDB

Bachelor Thesis

Benchmarking of Time Series Databases by extending the Yahoo Cloud Serving Benchmark

Context

In the context of Cloud Computing or the Internet of Things an upcoming task is the storage and analysing of monitoring data produced by applications or sensors. Typically the monitored data is kept for a certain amount of time to provide the possibility of long time analyses or machine learning. The monitored data is commonly referred as time series data which consists for example of the metric name, metric value and a timestamp.

A common way to store such monitoring data are Time-Series-Databases (TSDB). A TSDB typically consists of a datastore back-end and a front-end providing the aggregation capabilities and a graphical representation. Common TSDBs like KairosDB or InfluxDB use NoSQL databases as back-end datastores to handle even large amount of data by distributing the data.

The Yahoo Cloud Serving Benchmark (YCSB) is a popular framework for benchmarking NoSQL databases. The framework is written in Java and offers an extensible architecture. It already offers adapters for various databases like Cassandra, Couchbase or MongoDB.

Scope of the Thesis

This thesis focuses on the following challenges, defining TSDB related benchmark queries and extending the YCSB framework with at least one TSDB adapter. Possible TSDB could be InfluxDB, KairosDB or any other open source TSDB.

Requirements and Comments

If this thesis achieves good progress and outcome, its results are to be integrated in the PaaSage research project (http://paasage.eu) which is released under an OpenSource license. For that reason, we appreciate if you are ready to OpenSource your results.

If you are interested in this or similar theses, please contact Daniel Seybold either by mail or in directly in his office.

mail: daniel.seybold@uni-ulm.de
office: Uni West, 43.2.209

Faculty of Engineering and Computer Science

Institute of Information Resource Management