UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft
Faculty for Computer Science, Electrical Engineering and Mathematics
Department of Computer Science
Research Group DICE Group

Bachelor's Thesis Proposal

Submitted to the DICE Group Research Group
in Partial Fullfilment of the Requirements for the Degree of

Bachelor of Science

Basilisk – Continuous Benchmarking for Triplestores

by
Fabian Rensing

Thesis Supervisor:
Prof. Dr. Axel-Cyrille Ngonga Ngomo

Paderborn, October 18, 2021

1 Description

- halbe Seite

2 Motivation

The motivation

The topic I have to offer has the working title is "Basilisk – Continuous Benchmarking for Triplestores". At the core, it is developing and deploying a CI/CD tool that hooks into github and/or docker registries.

Described in more detail:

Triplestores – the database backend of knowledge graphs – are typically developed in long iterations and are bench-marked – if at all – only in a very late stage of such an iteration. Benchmarking and evaluation of benchmarking results are typically done manually and binds developer's time. Thus, performance regressions are found very late or never.

With Basilisk we started to develop a continuous benchmarking platform for triplestore which hooks into github and docker image registries.

On events like pull requests or newly published versions of triplestores, a benchmarking suite is run automatically.

The first version of Basilisk (https://github.com/dice-group/Basilisk, https://github.com/dice-group/basilisk-frontend) is already implemented. It is based on the benchmarking tool IGUANA (https://github.com/dice-group/IGUANA) and Docker. (It requires triple stores to be dockerized).

The thesis task is to:

1. describe and review the software architecture 2. deploy Basilisk and its frontend on a publicly available VM 3. benchmark 2 versions of Tentris (https://github.com/dice-group/tentris, via a github hook) and one version from another triple store (via a docker image registry, e.g. https://hub.docker.com/r/ontotext/graphdb/ or https://hub.docker.com/r/openlink/virtuoso-opensource-7). 4. fix critical bugs in 1.-3. and document non-critical

The thesis can be extended to a paper in a scientific journal (e.g. ISWC) on the resource track with you as first author.

3

Formulation of Target Setting

4

Tasks Descriptions

5

Schedule

Bibliography