

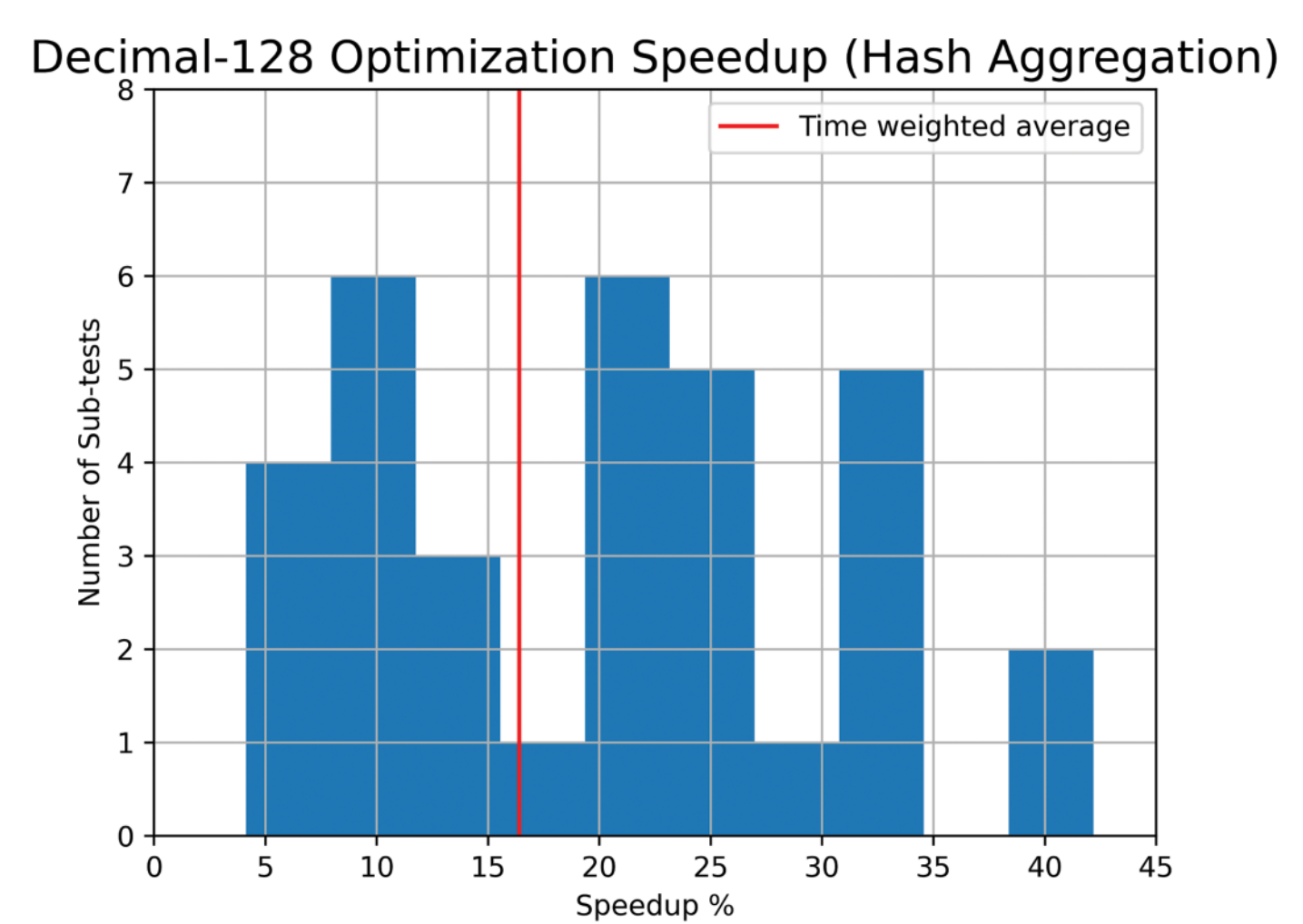
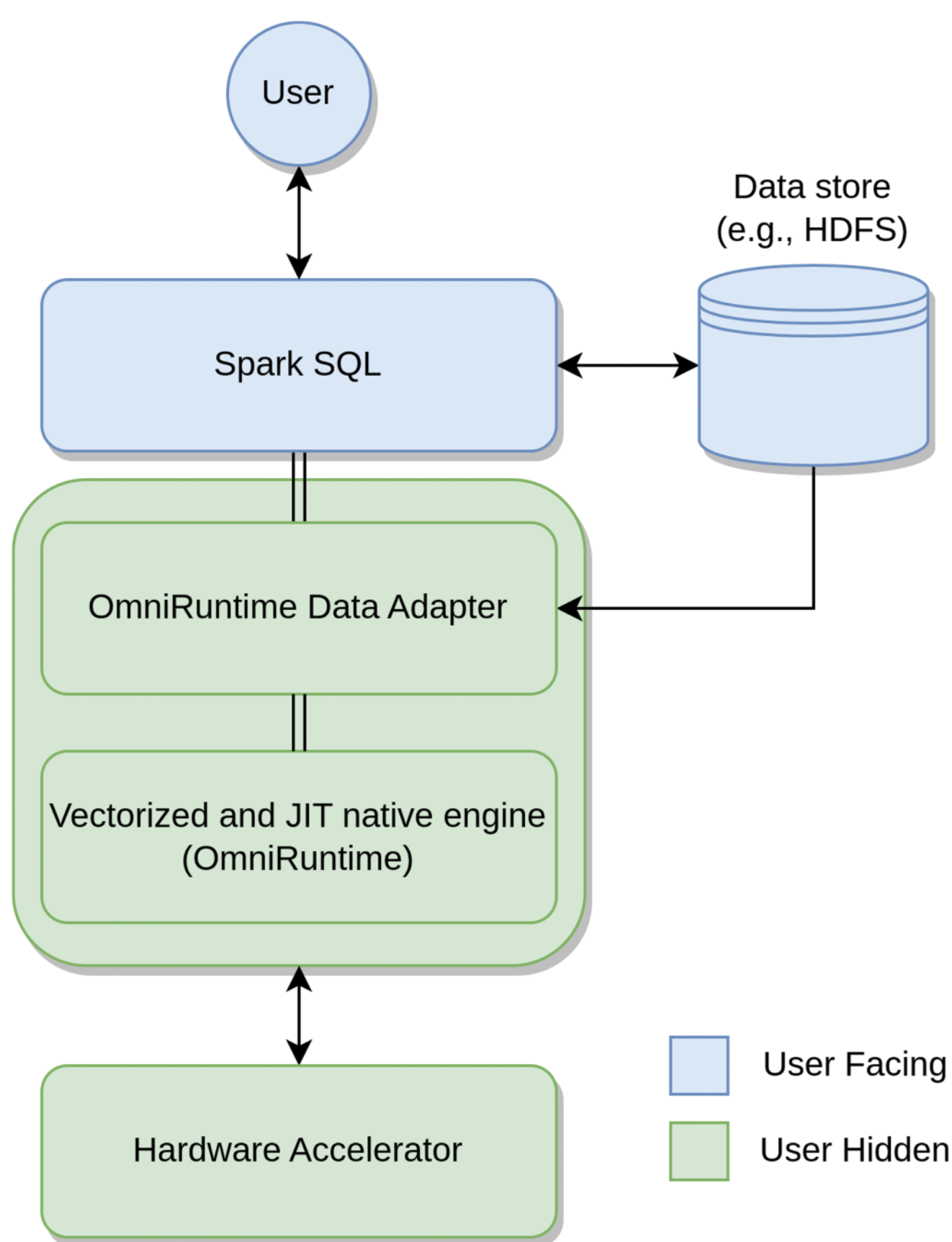
## OmniRuntime: Squeezing Every Last Drop of Performance for In-Memory Data Analytics

Hand optimizing a SQL analytics engine through low level analysis of operators and integrating hardware accelerators.

**Benjamin Correia**

**Dr. Nandita Vijaykumar**  
ACADEMIC SUPERVISOR

**Sundar Annamalai**  
INDUSTRY SUPERVISOR



### PROJECT SUMMARY

OmniRuntime (ORT) is an existing data analytics engine used for the acceleration of Spark SQL on Huawei’s proprietary hardware. In this project we explore how to increase runtime performance to make ORT the industry leader. We are interested in any optimization possible; from low-level changes targeting micro-architectural features, to novel indexing of data to reduce IO overhead. We also lay the groundwork for future optimizations and hardware acceleration via hardware accelerator offloading. We are targeting hardware acceleration due to the support for fine-grain parallelism common to data analytics algorithms, and availability of vendor libraries targeted towards data manipulation.

Our methodology includes low-level binary analysis, software profiling, and algorithmic analysis for hardware suitability. Our methods for benchmarking involved small scale micro-benchmarks, data and query excerpts from TPC-DS, and full scale end-to-end customer-like environments. After benchmarking we find our optimizations increase performance up to 16% depending on the exact test and environment.

