



## HIWI JOB

# The Spork-Join Project

## Validating the Queuing Theory of Map-Reduce Systems

Cloud-based big data processing is tremendously important these days. The map-reduce model of computing is becoming universal. The fundamental properties of these parallelized computing systems have been studied since the 1990's using queuing models such as Fork/Join and Split/Merge. However these popular models misrepresent the architecture of modern map-reduce systems!

We are running controlled experiments on a Spark-cluster. Our goal is to validate or invalidate the use of the popular Fork/Join and Split/Merge models for modeling modern map-reduce systems. Working on this project you can expect to become very familiar with the administration of a Spark system, and instrumentation, analysis, and programming.

Some specific possible tasks include:

- Build and experiment with new task scheduler for Spark.
- Write Spark driver scripts that controllably exercise Spark's map/shuffle/reduce phases.
- Implement synthetic Spark job-generators based on empirical studies of cloud traffic.
- Better instrument Spark task logging, and process and analyze logs.
- Set up containerized Spark workers with a fixed division of network bandwidth.



Qualifications:

- Programming in some relevant language; Scala, Java, C++
- Knowledge of Linux
- Awareness of the existence of queueing theory

---

### Contacts:

Dr. Brenton Walker  
Leibniz Universität Hannover  
Institut für Kommunikationstechnik

Tel.: 511 / 762 – 2827  
E-Mail: [brenton.walker@ikt.uni-hannover.de](mailto:brenton.walker@ikt.uni-hannover.de)