

---

## Resource Allocation for Scale-Out Database Systems and Cloud Computing

---

### Description

Growing data volumes and the desire for analyzing this data requires multi-node data management systems, e.g., scale-out database systems. Such systems are increasingly deployed in cloud environments. With that, questions arise about efficiently distributing data and processing load within a multi-node system and assigning cloud resources to this system. Therefore, in this master's project, we develop approaches for resource-efficient allocations in the context of scale-out database systems and cloud computing.

Allocation problems are optimization problems and omnipresent in database and enterprise systems. Many of these problems are NP-hard, i.e., if the input sizes increase, they quickly cannot be solved in a reasonable amount of time anymore – particularly when using a brute-force approach, examining all possible solution candidates. In the field of mathematical programming, we can use off-the-shelf solvers, which efficiently search for optimal solutions, to mitigate the increase of calculation time. Further, we can use the power of these solvers to build efficient heuristics.

We have previously developed a decomposition-based allocation approach using mixed-integer linear programming (a subclass of mathematical optimization) for partially replicated database clusters.

In this master's project, we want to investigate mathematical programming approaches for adapted problems, which are characterized by other optimization goals and constraints:

- For scale-out database systems, we want to improve the memory-efficiency (i.e., data reuse) when queries are distributed across multiple database nodes.
- In the context of cloud computing, we want to optimize the resource utilization when placing virtual machines with allocation constraints (e.g., co-location and fault tolerance) in a cluster.

### Skills

There are no particular prerequisites for this project. Prior knowledge in database or distributed systems may help to understand the problem domain. We will give you an overview of the problem domain and teach you to solve optimization problems using solvers, such as Gurobi and CLPEX, via the modeling language AMPL. For the rest of the programming work, we will use Python.

## Learning Goals

Through active participation in this project, you will:

- Learn to solve optimization problems with integer linear programming
- Apply techniques to control the problem complexity and, thus, calculation time
- Understand the possibilities to scale database systems
- Understand challenges for virtual machine placement
- Improve your research methodology and academic writing

After this project, there will be research opportunities to dive deeper into identified issues in the form of master's theses and PhD positions.

## Resources

- <https://ampl.com>
- Halfpap and Schlosser: Workload-Driven Fragment Allocation for Partially Replicated Databases Using Linear Programming. ICDE 2019
- Halfpap and Schlosser: Memory-Efficient Database Fragment Allocation for Robust Load Balancing when Nodes Fail. ICDE 2021
- Rabl and Jacobsen: Query Centric Partitioning and Allocation for Partially Replicated Database Systems. SIGMOD 2017
- Schlosser and Halfpap: Robust and Memory-Efficient Database Fragment Allocation for Large and Uncertain Database Workloads. EDBT 2021

## Organization

- Master's programs: ITSE, DE
- Extent: 8 SWS / 12 ECTS
- Content:
  - Group work (starting first week of lecture period; until end of semester)
  - Programming project
  - Final presentation (July/August)
  - Research report (due: September 30)

## Contact

You are welcome to contact us via email.

- Stefan Halfpap ([stefan.halfpap@hpi.de](mailto:stefan.halfpap@hpi.de))
- Dr. Rainer Schlosser ([rainer.schlosser@hpi.de](mailto:rainer.schlosser@hpi.de))
- Dr. Michael Perscheid ([michael.perscheid@hpi.de](mailto:michael.perscheid@hpi.de))