Preface

The concept of task containerization has ushered in a revolutionary transformation in software deployment and scalability. This innovative technology, exemplified by Docker and other containerization technologies, simplifies the packaging, distribution, and management of software. It abstracts and encapsulates applications and their dependencies within isolated containers, offering numerous advantages for software development and deployment. At its core, containerization abstracts the underlying infrastructure and operating system, allowing developers to bundle applications and all required dependencies into self-contained units known as containers.

This thesis explores the design, implementation, evaluation, benefits, challenges, and future prospects of task containerization. It showcases the integration of task containerization with artificial intelligence techniques and underscores its significance in modern software engineering practices. In the realm of High Performance Clustering, the efficient allocation and execution of tasks within containerized environments have become a critical challenge. This thesis presents an innovative approach to address this challenge by leveraging the power of Artificial Intelligence (AI) techniques. Task containerization, which plays a pivotal role in orchestrating workloads across clusters, is optimized to enhance resource utilization, meet varying application demands, and ensure secure task execution. The thesis explores the intersection of containerization technology and federated learning to address the challenge of privacy-preserving task offloading. It proposes methodologies that leverage AI-driven algorithms, including

machine and deep learning models, to allocate tasks to containers based on real-time resource availability and workload characteristics.

The thesis introduces a novel approach, Secure Task Containerization with Deadline Constraint (S-TCDC), which leverages containerization to address resource allocation challenges and workload distribution in cluster computing environments. S-TCDC aims to facilitate secure task offloading from resource-constrained devices to more powerful compute servers within an allowable response time. Extensive experiments and simulations validate the effectiveness of S-TCDC, enhancing overall system performance. The study also proposes a semi-distributed algorithm for distributed task resolution within a containerized framework. It employs the Stackelberg game theory approach, federated learning, and knowledge distillation to fine-tune containerization decisions, estimating the fraction of the task to be containerized while minimizing costs and ensuring the desired security level. In the context of task offloading with sensitive user data, the thesis explores container-based federated learning as a privacy-preserving approach for task allocation. Furthermore, the thesis introduces a novel federated learning mechanism through Docker, creating connections among multiple Docker containers to facilitate memory sharing within a structured graph framework. These interconnected containers collaborate to create a global model and transmit updates to a designated shared memory repository. Subsequently, the containers retrieve these updated models from shared memory and employ them to train their respective datasets. This approach accommodates three distinct types of devices and empowers the efficient and secure execution of machine learning within High Performance Computing (HPC) environments. It streamlines resource management, dependency handling, and library usage, enabling agile and secure machine learning execution within HPC workflows. The successful implementation of Federated Learning for precise COVID-19 predictions demonstrates its effectiveness in preserving data integrity and privacy within data-intensive challenges.