Effectively Task Containerization in High Performance Clustering using Artificial Intelligence

आर्टिफिशियल इंटेलिजेंस का उपयोग करके

उन्नत सक्षम क्लस्टरिंग में प्रभावी ढंग से कार्य कंटेनरकरण

Thesis submitted in partial fulfillment
for the Award of Degree

Doctor of Philosophy

by

Chitranjan Singh

चितरंजन सिंह

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY
(BANARAS HINDU UNIVERSITY)
VARANASI - 221005

Roll No. 17071512

Year 2023

Chapter 7

Conclusion and Future work

This chapter provides an overview of the thesis's context, summarizing its primary contributions and outlining future research directions in the field of "Effectively Task Containerization in High Performance Clustering using Artificial Intelligence".

Conclusion

The research conducted in the thesis delves into the integration of artificial intelligence (AI) techniques within task containerization for High Performance Clustering, yielding significant findings and valuable contributions. This thesis makes substantial contributions to the domains of secure and efficient task execution, privacy-preserving machine learning, and resource optimization across Containerization, Federated Learning, and High Performance Computing. These contributions propose pioneering solutions to address critical challenges in these fields, with the potential to fundamentally transform and redefine existing practices. Importantly, they prioritize the crucial aspects of security, efficiency, and data privacy, ensuring that as technology evolves, it does so with a strong commitment to safeguarding sensitive information and delivering optimal performance in diverse applications and contexts. These contributions represent a significant step forward in enhancing the reliability and effectiveness of contemporary computing

and machine learning paradigms.

In chapter 3, we introduce a novel task containerization approach uniquely tailored to the industrial context, where both security and stringent deadline constraints are paramount. Unlike prior works, our system is purpose-built for industrial applications, addressing the critical coexistence of security and timely task execution. Our approach delves into the cost considerations of individual compute nodes in high-performance clusters, encompassing task execution and security expenses. To optimize this multi-faceted scenario, we employ a Stackelberg game framework, where Worker Machines (WMs) act as followers, guided by a Master Machine (MM) as the leader. This strategic framework enables efficient task allocation while accounting for cost and security dynamics. The validity and efficacy of our approach are rigorously validated through mathematical analysis and practical experimentation, outperforming existing methods within the high-performance cluster context. It is worth noting that our framework's flexibility allows for variations in node acceptance of task allocations, accommodating real-world scenarios. Additionally, while our approach assumes independent task partitions within a single cluster node, it can readily adapt to handle interdependencies, ensuring efficient task execution in complex settings.

In chapter 4, we have introduced a comprehensive and intelligent patient monitoring system that harnesses Federated Learning (FL) within the Internet of Medical Things (IoMT). Our contributions extend beyond existing work, primarily in the improved participant partitioning, where we consider factors like bandwidth, dataset size, and data freshness. We have also detailed the creation of a large-scale Deep Neural Network (DNN) model tailored for patient activity monitoring, effectively utilized in a federated learning framework. Our approach uniquely addresses participant heterogeneity by clustering them according to their resources and constraints, ensuring efficient alignment with each cluster's capabilities. This innovative clustering optimizes the model training process, making our system more efficient and resource-aware. Furthermore,

our iterative model compression using Knowledge Distillation (KD) significantly boosts performance, especially in clusters with limited resources. Continuously compressing the server model and deploying it to different clusters optimizes Federated Learning, adapting to diverse strengths and requirements while safeguarding privacy and data security. Our patient monitoring system underwent rigorous evaluation on publicly available and collected datasets, affirming its effectiveness. The results validate practical applicability, with improvements in personalized and generalized accuracy, reduced memory consumption, and enhanced model stability, making our system a valuable asset in real-world patient monitoring scenarios.

In chapter 5, we highlighted the critical importance of Optimized Container Selection Using Shared Memory Architecture in the realm of containerized learning. The shared memory architecture presented here plays a pivotal role in enhancing OS loading and memory utilization, demonstrating the tangible benefits of containerization and structured memory sharing. This innovative approach not only enhances performance but also optimizes resource utilization and scalability, positioning it as a promising solution for privacy-conscious and efficient machine-learning applications in decentralized settings. Our findings affirm the innovation and potential inherent in our containerized learning framework. By introducing and validating the shared memory architecture, we have made a significant contribution to the expanding knowledge base in the domain of optimized container selection. This research underlines the continued relevance of shared memory architecture, reaffirming its role in the pursuit of superior containerized learning solutions.

Finally, in chapter 6, we proposed a synergistic union of containerization and Federated Learning, resulting in a comprehensive and highly efficient approach to model development, optimization, and evaluation for chest condition diagnosis. The extensive testing and fine-tuning process undertaken in this study have ensured the readiness of the machine learning model for real-world deployment, marking a significant mile-

stone in the realm of medical assessments, particularly for conditions like COVID-19. Our experiment not only showcased the model's proficiency but also emphasized the paramount importance of collaboration and consistency in the fields of machine learning and healthcare. This collaborative approach, leveraging the power of distributed learning, brings us closer to harnessing the full potential of AI in the medical domain. It not only enhances diagnostic accuracy and efficiency but also serves as a model for future research and application, fostering a more collaborative, precise, and robust healthcare ecosystem. Our work opens the door to further advancements in medical AI, promising more accurate and timely diagnoses and ultimately improving patient care. Moreover, it has been successfully validated on the PARAM Smriti HPC System, further attesting to its practical applicability.

Future Directions

This thesis work has made significant contributions in the domains of secure and efficient task execution, privacy-preserving machine learning, and resource optimization within the intersections of Containerization, Federated Learning, and High Performance Computing. These contributions propose pioneering solutions to address critical challenges, with the potential to transform existing practices. Our contributions prioritize security, efficiency, and data privacy, reflecting a commitment to safeguarding sensitive information and delivering optimal performance in diverse applications.

Firstly, our novel task containerization approach addresses security and timely task execution in industrial contexts, providing tailored solutions to optimize task allocation. This approach considers both task execution and security costs, using a Stackelberg game framework. The approach's validity and efficacy have been rigorously validated through mathematical analysis and practical experimentation, demonstrating its superiority in high performance cluster contexts.

Secondly, our patient monitoring system harnessing Federated Learning within the

Internet of Medical Things (IoMT) introduces improved participant partitioning, large-scale DNN model development, and iterative model compression. The system demonstrates superior accuracy, reduced memory consumption, and enhanced model stability, making it a valuable asset in real-world patient monitoring scenarios.

Thirdly, we have highlighted the critical importance of Optimized Container Selection Using Shared Memory Architecture in the realm of containerized learning, showcasing its benefits for enhanced performance, resource utilization, and scalability.

Lastly, our synergistic union of containerization and Federated Learning for chest condition diagnosis promises more accurate and timely diagnoses in the medical domain. This collaborative approach fosters a more precise and robust healthcare ecosystem, supported by successful validation on the PARAM Smriti HPC System.

The work described here can be extended further in the following directions:

- Diverse Applications: Extend research findings to various domains like finance, energy, and logistics, exploring the adaptability of methodologies.
- Advanced Security: Evolve security protocols, intrusion detection, and privacy measures to proactively address emerging threats.
- Scalability and Efficiency: Enhance solutions to scale seamlessly in high-performance computing, cloud, and edge environments while optimizing resource utilization.
- Integrated Technologies: Investigate the synergy of task containerization, federated learning, and optimized container selection for comprehensive problem-solving.
- Fair AI: Prioritize fairness in machine learning applications, mitigating bias and discrimination in real-world scenarios.
- Human-AI Synergy: Explore collaborative models where AI enhances human capabilities, enabling better decision-making and problem-solving.

Looking ahead, future directions include exploring applications of these methodologies in diverse domains, refining the security aspects further, data privacy, efficiency,

and advancing collaborations across healthcare and machine learning to bring these innovations to real-world settings.