

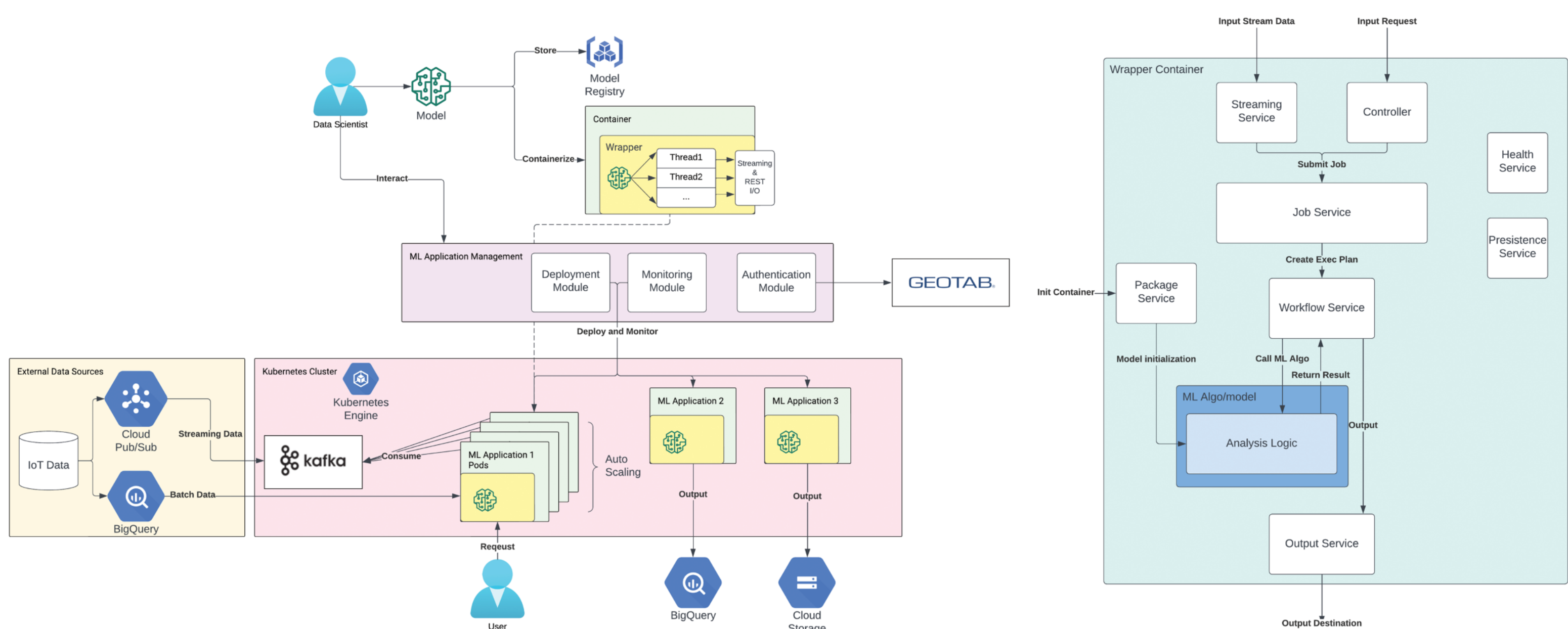
From ML Model to Production: Streaming Data Analytics Infrastructure for IoT Big Data

A modern approach for deploying machine learning applications on cloud-based infrastructure integrating distributed computing and streaming capabilities for IoT Big Data use cases

Zhifan Wang

Hans-Arno Jacobsen
ACADEMIC SUPERVISOR

Mengliao Wang
INDUSTRY SUPERVISOR



PROJECT SUMMARY

With the continuous expansion of the Internet of Things (IoT) market, industries are recognizing that by applying real-time analytics alongside machine learning models on this data, they can extract valuable insights from the big data, and make rapid responses to real-world incidents. Considering the nature of IoT data, the analytics workflow needs to handle data that arrives with high volume and high velocity. To bridge the gap between data scientists' experimental algorithm and a robust, high-performance big data service in production, we propose a cloud-based infrastructure solution that is scalable, distributed, and seamlessly integrated with streaming capabilities.

Our solution leverages the power of containerization technology and provides a development toolkit to enable parallelism, configure the I/O and ensure the adaptivity of the ML model on different system environments, with minimum effort from model developers. Then, we provide a cloud infrastructure based on Kubernetes to deploy the models in a distributed manner, with fault-tolerance and scaling capabilities. Furthermore, a streaming module built on Apache Kafka is introduced for delivering IoT data directly into the ML models.

We evaluate our approach on several industrial-level big data analytics scenarios and demonstrate that the solution can perform more stably and efficiently compared to traditional ML application deployment approaches.