Intelligent Re-deployment Feedback Loop for Hybrid Applications

Kalman Meth
IBM Research - Haifa
meth@il.ibm.com

Indika Kumara
Jheronimus Academy of Data Science
i.p.k.weerasingha.dewage@tue.nl

Giovanni Quattrocchi
Politecnico di Milano
giovanni.quattrocchi@polimi.it

Published in [SYSTOR'21] ACM International Systems and Storage Conference, June 2021, Haifa, Israel.

ACM ISBN 978-1-4503-8398-1/21/06

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825480.

Intelligent Re-deployment Feedback Loop for Hybrid Applications

Kalman Meth
IBM Research - Haifa
meth@il.ibm.com

Indika Kumara
Jheronimus Academy of Data Science
i.p.k.weerasingha.dewage@tue.nl

Giovanni Quattrocchi
Politecnico di Milano
giovanni.quattrocchi@polimi.it

ABSTRACT

We propose enabling continuous performance optimisation of distributed hybrid applications in heterogeneous cloud, Edge, and HPC environments by employing an intelligent re-deployment feedback loop.

CCS CONCEPTS

• **Software and its engineering** → *Development frameworks and environments*; *Monitors*; • **Computer systems organization** → **Heterogeneous (hybrid) systems**; *Cloud computing*.

KEYWORDS

Distributed computing, monitoring, application optimisation

1 DISTRIBUTED APPLICATION REDEPLOYMENT

In hybrid / multiple cloud, Edge and HPC environments, distributed applications consist of numerous components that work together. Based on the requirements of each component and the available resources in the computing environments, a deployment plan is developed to run the application.

Once the application is deployed, monitoring information is continuously collected at the infrastructure level. If available, performance information is also collected from the application. If the performance of the application is not satisfactory with its current deployment, it may be possible to improve its performance by deploying it differently on the existing or on other hardware (perhaps on a different cloud). The monitoring information is fed to the Refactoring-Redeployment component, which recommends a new deployment of the distributed application.

Figure 1 illustrates the intelligent re-deployment feedback loop. A lightweight environment runs on each hardware component to capture metrics and other items being monitored. The monitoring information is consumed by an intelligent Refactoring component to determine a better deployment plan that is input to the Orchestrator component. The Orchestrator orchestrates a re-deployment of the application.

The refactoring employs a machine learning based planner and a control theory based planner, each operating at different timescales. The machine learner can estimate the impacts of a given deployment option selection (a new deployment model) on the performance metrics under different workloads, and thus generate a new alternative deployment plan and assign the performance goals to the individual nodes in the deployment plan. The hierarchical control-theoretical planner, given the goals selected by the machine learner, optimizes GPU and CPU allocation for the different application components running on each node in the deployment plan.

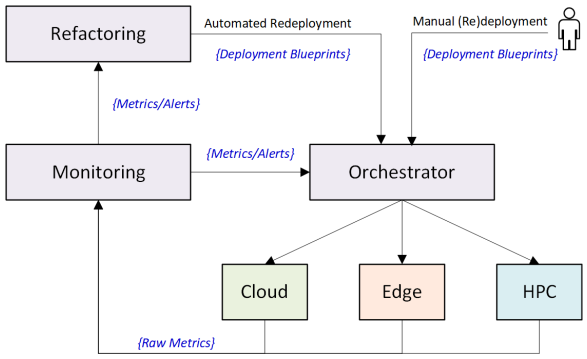

Figure 1: Deployment and Control Architecture.

ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825480.

SYSTOR '21, June 14–16, 2021, Haifa, Israel

ACM ISBN 978-1-4503-8398-1/21/06…$15.00
https://doi.org/10.1145/3456727.3463830