

Compute optimiser for containerised and function-based cloud applications

In den letzten zehn Jahren war Cloud Computing eines der wichtigsten Schlagworte in den IT-Abteilungen von Unternehmen. Es ist ein Ziel für jedes Unternehmen, das seine Infrastrukturkosten senken und einen Teil seiner Ausgaben von CAPEX auf OPEX umstellen will.

Die Hauptidee besteht darin, die Infrastrukturnutzung in Dienste aufzuteilen, die je nach Bedarf verkauft oder gekauft werden können.

Die beiden Hauptschwerpunkte des Cloud Computing sind die Optimierung der Kosten und die Optimierung der Leistung.

Mit der Verbreitung von Machine Learning Tools wurde die Möglichkeit geschaffen, diese zu nutzen, um die bestmögliche Kombination zwischen Preis und Leistung von Cloud Services zu finden.

In unserer Bachelorarbeit werden wir ein alternatives Cloud-Optimierungstool erforschen und entwickeln, das nicht an eine proprietäre Cloud-Plattform gebunden ist. Dabei versuchen wir uns den Leistungen der bereits bestehenden Lösungen anzunähern.

Um unser Ziel zu erreichen, haben wir einen Mixed-Methods-Ansatz verwendet, bei dem wir zunächst die relevante bestehende Literatur recherchiert und eine Umfrage bei Schweizer Cloud-Anbietern durchgeführt haben. Auf dieser Basis haben wir unser Lösungskonzept entwickelt. Schliesslich verifizieren wir unsere Forschung und unser Konzept durch die Implementierung eines Prototyps, um unsere Vermutungen zu testen und die Machbarkeit des Konzepts zu zeigen.

Während unserer Forschung konzentrierten wir uns aufgrund der für Schweizer KMU hauptsächlich genutzten Software-Landschaft auf die Optimierung von Container-Technologien, genauer gesagt auf die vertikale Skalierung.

Wir wollten daher die Anpassung der Vorhersagen mit Hilfe eines Machine Learning Ansatzes verbessern.

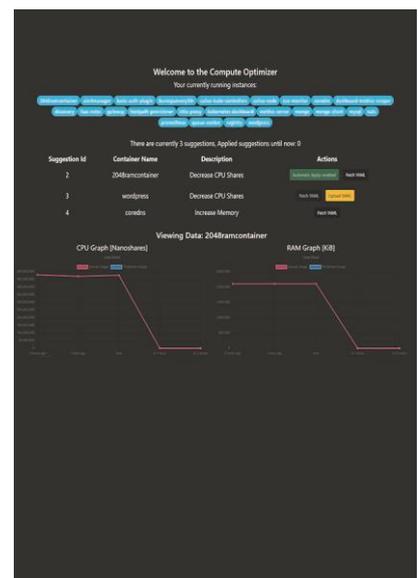
Mit unserer Arbeit haben wir eine funktionierende Lösung erreicht. Wir fragen im Minutentakt Messdaten ab, aus denen wir Vorhersagen generieren. Diese werden stündlich von unseren APIs auf mögliche Vorschläge überprüft, automatisch mit einer Meldung auf dem Dashboard abgeglichen und wenn möglich automatisch angewendet.

Wir sehen Potenzial in der Weiterentwicklung dieser Arbeit, mit dem hinzufügen anderer Vorschlagsarten, der erweiterten Unterstützung für andere Kubernetes-basierte Orchestrierungslösungen oder sogar der Integration unseres Machine Learning Algorithmus für die Ressourcenanpassung in die Kubernetes Metrics Vertical Pod Autoscaler Komponente.



Diplomierende
Sergio Conti Rossini
Nick Marti

Dozierende
Josef Spillner
Andreas Weiler



WebUI Dashboard - Compute
Optimizer