



**Studienarbeit  
Forschungsarbeit  
Masterarbeit**

zu vergeben

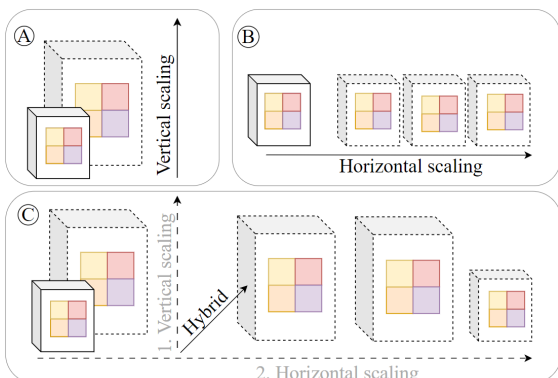
# Erweiterung Kubernetes Autoscaler um multi-level Monitoring

## Hintergrund

Um flexible Nutzung von Rechen und Kommunikationsressourcen, wie im IT-bereich, auch in der Automatisierungstechnik zu ermöglichen, werden Konzepte aus der IT entsprechend angepasst, um die Vorteile auch in der Automatisierung zu nutzen. Kubernetes ist ein open-source Container Orchestrierungs Framework welches z.B. bei den Cloud Providern Google, Amazon und Microsoft Azure eingesetzt wird, um Ressourcen zu verwalten und Anwendungen entsprechend ihrer Anforderungen zu skalieren.

## Problemstellung

Kubernetes horizontaler Pod Autoscaler (HPA) greift für Skalierentscheidungen standardmäßig auf die Metriken die durch den Metricsserver bereitgestellt werden zurück. Überwacht werden hierbei lediglich relative CPU und Speichernutzung. Cloud gehostete Anwendungen, idR. server-basierte Web-anwendungen, unterscheiden sich allerdings in ihren Anforderungen deutlich von Steuerungsapplikationen, weshalb diese Metriken zur Überwachung nicht ausreichen. Um Skalierentscheidungen für Echtzeitanwendungen treffen zu können, müssen unter anderem die absolute Ressourcennutzung des Hosts, sowie Task-level Metriken der Anwendung überwacht werden.



## Aufgabe

- Einarbeitung in Cloud Skalierung, Echtzeit Workloads und Kubernetes
- Anforderungsanalyse für Monitoring von Echtzeit-tasks
- Vergleich mit Kubernetes Autoscaler und Herausarbeiten der Lücken
- Konzeptionierung des multi-Level Monitoring
- Umsetzung des Konzepts und Einbindung in Kubernetes HPA.
- Validierung anhand Testapplikation

## Anforderung

- Interesse an verteilten Systemen und Cloud Computing
- Interesse an Kubernetes und Multi-level Monitoring

## Kenntnisgewinn

- Cloud Computing und Skalierung
- Performance und Ressourcenmonitoring
- Kubernetes

