

Operating Heterogeneous Server Systems - Improving Power Management in the Future SOC Lab

Arbeitslasten wie Maschinelles Lernen, In-Memory Computing sowie ähnlich komplexe Anwendungen stellen vermehrt eine Herausforderung für den Betrieb von Rechenzentren dar. Diese intensive Nutzung von Ressourcen rückt das Energiemanagement im Betrieb in den Fokus. Einerseits übersteigen die laufenden Kosten wie z.B. Energiekosten eines Rechenzentrums bei weitem die Anschaffungskosten der Hardware. Andererseits ist der Energieverbrauch von IT-Systemen auf vielfältige Weise ein limitierender Faktor geworden. Dieser bestimmt wie viele Server betrieben und welche Ausstattung diese haben können. Die erforderliche Kühlleistung ist ebenfalls eng an dem Energieverbrauch der zu kühlenden Systeme gekoppelt. Elektrische Energie wird zu einem großen Teil als Wärme abgegeben und muss abtransportiert werden. Die Rechenzentrumslandschaft wird immer heterogener und der Trend zeigt, dass Server-Systeme immer mehr Energie benötigen. Das wird bereits im Future SOC Lab deutlich.

Das HPI Future SOC (Service-Oriented Computing) Lab ist eine Kooperation des Hasso-Plattner-Instituts mit den Industriepartnern Hewlett Packard Enterprise (HPE), Dell EMC, Fujitsu, SAP. Seine Mission ist die Ermöglichung und Förderung des Austauschs zwischen Forschungsgemeinschaft und Industrie. Am Lab wird interessierten Wissenschaftlern eine Infrastruktur von neuester Hard- und Software kostenfrei für Forschungszwecke zur Verfügung gestellt. Dazu zählen teilweise noch nicht am Markt verfügbare Technologien, die im normalen Hochschulbereich in der Regel nicht zu finanzieren wären. Einige der Schwerpunkte sind Cloud Computing, Maschinelles Lernen, Parallelisierung und In-Memory Technologien.

Studierende sollen in diesem Masterprojekt Methoden, Techniken und Werkzeuge evaluieren, mit denen es möglich ist den Energieverbrauch zu erfassen und auszuwerten. Hierfür soll zunächst eine Übersicht zu dem Stand der Kunst erstellt werden. Ein oder mehrere dieser Verfahren sollen ausgewählt und in Betrieb genommen werden. Mit den gesammelten und analysierten Daten soll die praktische Anwendung durch Platzierung von virtuellen Maschinen, Containern oder Jobs evaluiert und mögliche Überlast-Situation vermieden werden. Ziel ist es die Ergebnisse in einem wissenschaftlichen Papier oder einem technischen Bericht zu veröffentlichen.

Kontakt

Bei weiteren Fragen stehen wir gerne zur Verfügung. Projektverantwortlich sind Bernhard Rabe, Max Plauth, Felix Eberhardt sowie Andreas Polze.