Anschubfinanzierung Globaler Süden 2023 – Projektbericht

Energieeffizientes Maschinelles Lernen

- **Fachgebiet an der TU Berlin:** Verteilte Systeme und Betriebssysteme (DOS)
- **Partnerländer:** Nordmazedonien
- **Partnerinstitutionen:** University of Kiril and Metodij, Skopje, Faculty of Computer Science and Engineering
- **Vorrangig adressierte Nachhaltigkeitsziele (SDGs):**

© UNITED NATIONS

SDG 12: Nachhaltige/r Konsum und Produktion

Cloud Computing ist zu einem integralen Bestandteil sowohl der industriellen IT-Infrastruktur als auch der täglichen menschlichen Interaktionen geworden und hat den Zugang zu Rechenleistung revolutioniert. AIOps (Artificial Intelligence for IT Operations) hat sich als Schlüssellösung für die Unterstützung des Cloud-Betriebs durch die Automatisierung von IT-Aufgaben erwiesen. Trotz ihrer Bedeutung erhöhen sie jedoch den Energiebedarf von Cloud-Systemen. Mit der zunehmenden Verbreitung der Cloud verbrauchen die Rechenzentren weiterhin viel Energie und tragen damit indirekt zu den Treibhausgasemissionen bei.

Im Rahmen dieser Global-South-Initiative haben wir auf das Ziel der nachhaltigen Entwicklung "Nachhaltiger Konsum und nachhaltige Produktion" reagiert und eine Zusammenarbeit mit Professor Gjorgji Madjarov von der Universität St. Cyril und Methodius in Skopje, Nordmazedonien, begonnen. Ziel der Zusammenarbeit ist die Untersuchung und Verbesserung der indirekten Umweltauswirkungen der derzeitigen Praktiken zur Ausführung von AIOps-Workloads. Unser primäres Ziel ist es, die Beziehung zwischen den bestehenden Planungsstrategien bei der Ausführung von AIOps-Workloads und deren Nachhaltigkeit zu verstehen. Wir untersuchen die potenziellen Kosten, die sich aus dem zeitkritischen Aspekt der AIOps-Workloads unter Berücksichtigung ihrer Umweltauswirkungen ergeben. Wir konzentrieren uns zunächst auf die Untersuchung der aktuellen Praktiken, wann und wo die AIOps-Workloads geplant werden. Ziel ist es, die Produktionspraktiken zu verstehen, wenn AIOps-Methoden ausgeführt werden. In der gängigen Praxis werden sie als Teil der regulären Rechnerauslastung betrachtet und meist in regelmäßigen Abständen (z. B. einmal pro Tag oder Woche) manuell nachtrainiert. Zusätzliche automatische Strategien, wie z. B. die Erkennung von Konzeptabweichungen, werden zwar eingesetzt, aber in der Praxis nicht häufig verwendet, da sie zusätzliche Fehler verursachen. Wird die erste Strategie angewandt, können die Cloud-Anbieter die Planungsphasen auf Zeiten mit verfügbarer grüner Energie verlegen

und so eine kostenneutrale Vermeidung ohne Leistungseinbußen erreichen. Bei der letztgenannten Strategie sind die AIOPs mit höheren Kosten verbunden, da die Methode zur Erkennung von Konzeptdrift ständig ausgeführt werden muss. Unsere Forschung unterstreicht die Notwendigkeit, die Emissionsprofile von AIOps-Workloads unter Berücksichtigung von Faktoren wie geographische Region und Berechnungszeit zu quantifizieren.

Im Verlauf der Studie tauschten wir Ideen und Ansätze mit den weltweit führenden AIOPs-Experten auf zwei AIOPs-Workshops aus, die wir in Zusammenarbeit mit Prof. Gjorgji Madjarov organisierten. Der erste AIOPs-Workshop fand im Mai 2023 in Berlin in Räumen des Einsteinzentrums Digitale Zukunft statt und versammelte 25 Experten aus der ganzen Welt. Der zweite AIOPs-Workshop fand im Dezember 2023 im Rahmen der renommierten internationalen Konferenz über Data Mining (ICDM) in Shanghai, China, statt. Sieben Vorträge über AIOPs unterstrichen die Bedeutung von AIOPs und die Notwendigkeit einer nachhaltigen Entwicklung. Um diese Herausforderungen weiter zu adressieren, haben wir uns für ein EU-Projekt beworben und planen derzeit weitere gemeinsame Forschungsvorhaben. Den Grundstein für diese Aktivitäten legten wir mit der Stärkung der Zusammenarbeit durch gegenseitige Austauschbesuche, die mit dem Seed Funding der TU Berlin finanziert wurden: Prof. Madjarov verbrachte 2023 fast zwei Monate in Berlin, während Prof. Kao und Dr. Bogatinovski rund einen Monat in Skopje, Nordmazedonien, weilten. Die Zusammenarbeit wird mit dem nächsten geplanten Besuch in Skopje im März 2024 fortgesetzt.