

Aufbau und Einsatz eines Lehrclusters für Big Data Anwendungen

Prof. Dr. Bertil Schmidt, Dr. Christian Hundt (FB08, Institut für Informatik)

Das Erheben großer Informationsmengen verändert zunehmend unseren Alltag, z.B. im Verkehr, im Klimaschutz, in der Medizin. Das Themenfeld Big Data ist somit nicht nur von hoher praktischer und akademischer Relevanz in vielen Fachbereichen der JGU Mainz, sondern auch bei Studierenden sehr beliebt.

In diesem Projekt soll ein moderner Compute-Cluster auf Spark-Basis in der Lehre nachhaltig eingesetzt werden. Der Cluster soll im Rahmen mehrerer Lehrveranstaltungen zum Einsatz kommen, um die effiziente Programmierung in den Sprachen Spark und Hadoop zu erlernen. Hierzu werden entsprechende Programmieraufgaben basierend auf aktivem Lernen in Kleingruppen und Paarprogrammierung erstellt, die innerhalb von Vorlesungen oder Übungen basierend auf dem Web-Interface basierten SAUCE System zum Einsatz kommen.

Die Etablierung eines Big Data-Clusters auf Basis von Hadoop und Spark für Lehrzwecke ist ein erheblicher Beitrag zu einer modernen, innovativen und nachhaltigen Ausbildung. Die Studierenden können auf einem solchen System Hands-on Erfahrung in einem der derzeit wichtigsten informatischen Themenfelder erlangen.