Zürcher Hochschule
für Angewandte Wissenschaften

School of
Engineering

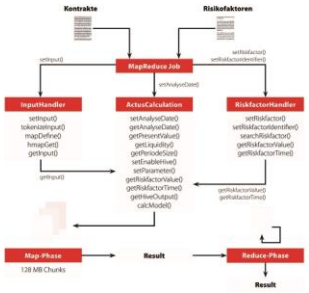
InIT Institut für angewandte
Informationstechnologie

Skalierbare Finanzanalysen unter Verwendung von Big Data und Data Warehouse Technologien

Die Analyse von Liquidität, Ertrag und Risiko ist für Finanzdienstleister sehr wichtig. Da die heutigen Ansätze und Fähigkeiten seit der letzten Finanzkrise in der Kritik stehen, soll das Projekt ACTUS [projectactus.org] ein zukunftsorientiertes, aussagekräftiges Werkzeug für Entscheidungsträger werden. Dabei steht die Modellierung einzelner Finanzinstrumente und deren erwarteten Zahlungsströme in Abhängigkeit der Risikofaktoren im Zentrum. Damit müssen Millionen von Finanzkontrakten in ACTUS abgebildet und die erwarteten Zahlungsströme algorithmisch abgeleitet werden. Dies macht die Finanzanalyse zu einem Big Data Problem.
Im Fokus der vorliegenden Arbeit steht die Anwendung von ACTUS in Kombination mit der Big Data Technologie Hadoop. Anhand eines theoretischen Einstiegs und des bekannten „WordCount"-Beispiels erhält der Leser einen praktischen Einblick in die Programmierung mit MapReduce. Danach werden Informationen zu ACTUS erläutert. Das MapReduce-Programm wird in drei Szenarien ausgearbeitet: Das Erste errechnet den Barwert eines Portfolios von Finanzinstrumenten, das Zweite die Liquidität und das Dritte erzeugt einen generischen Output für das Hadoop-basierte Data Warehouse Hive. Mithilfe von Hive können die Ergebnisse kontrolliert und weitere Analysen durchgeführt werden.
Performance und Skalierbarkeit der MapReduce-Programme werden mit Hilfe von unterschiedlich grossen Portfolios analysiert. Dabei wird die Performance anhand der Metrik Megabyte pro Sekunde gemessen. Zum Vergleich wird die maximale MapReduce Performance auf dem Cluster anhand von zwei Referenzprogrammen ermittelt.
Die Resultate zeigen, dass die Analyse grosser Finanzportfolios mit ACTUS und Hadoop möglich ist. Die gewählte Konfiguration erlaubt es, in einem Zeitraum von 9.5 Stunden 100 Mio. Kontrakte zu verarbeiten. Kleinere Portfolios mit 2 bis 5 Mio. Kontrakten werden innert 12 respektive 30 Minuten berechnet. Ein Grossteil der aufgewendeten Analysezeit ist auf den ACTUS Algorithmus zurückzuführen. Die Optimierung dieses Algorithmus soll ausserhalb dieser Arbeit nochmals behandelt werden. Ein weiterer zu berücksichtigender Aspekt ist die Monte-Carlo-Simulation, die in Kombination mit ACTUS die Entscheidungsfindung unterstützt.
Auf den letzten Seiten beschäftigen wir uns mit dem Hive Output, führen einige Diskussionspunkte auf und steuern im Ausblick unsere Gedanken zum zukünftigen Umgang mit Hadoop/ACTUS bei.

Diplomierende
Andreas Bösch
Manuel Kofler

Dozierende
Kurt Stockinger
Nils Andri Bundi



Konzeptioneller Aufbau des ACTUS MapReduce-Programms. Grundsätzlich geht es darum, dass ein grosses Inputfile mit Kontraktdefinitionen und diverse Parameter innerhalb des RunStatements, dem MapReduce-Job übergeben werden.

Auswertung des Szenario 1 auf dem Hadoop CLuster.