

Analytics@Big Data mit dem SAS High Performance Analytics Server

Schütz, Martin

SAS Deutschland

SAS® High-Performance Analytics Server – kurz HPA Server- basiert auf neuer In-Memory Technologie von SAS und dient dazu, den Prozess der Erstellung komplexer Modelle oder Anwendung analytischer Methoden für Mustererkennung extrem zu beschleunigen. Die dafür neu entwickelten SAS Prozeduren übernehmen dabei die intelligente Steuerung der Hauptspeicherressourcen einer zugrundeliegenden, auf massive Parallelisierung ausgerichteten Infrastruktur für Hardware und Data Storage. Eine solche Infrastruktur lässt sich heute etwa mit Teradata oder EMC Greenplum, aber auch über das Open-Source Framework Hadoop in Verbindung mit Standard-Hardware unter Linux realisieren.

Damit werden Fragestellungen für analytische Modellierung zugänglich, bei denen bisher aufgrund prohibitiv großer Datenvolumen auf Workaround-Strategien wie Stichprobenziehung, Verdichten granularer Datensätze oder Einschränkungen beim Algorithmen-Tuning (z.B. Topologie und Iterationen-Begrenzung beim Trainieren neuronaler Netze) zurückgegriffen werden musste.

In diesem Vortrag erhalten Sie einen Überblick darüber, wie der HPA Server funktioniert und welche neuen Anwendungsmöglichkeiten in Bezug auf Performance-Optimierung bei der Modellentwicklung für analytische Anwender und Data Step Programmierer sich daraus ergeben.