Massively distributed BIG DATA management for Enterprises

Franz Faerber[1], Jonathan Dees[1], Westenberger Eric[1], and Marc Hartz[1]

SAP SE, Germany
{franz.faerber,jonathan.dees,eric.westenberger,marc.hartz}@sap.com

More and more companies recognize the value of digitalized information and data for their business. We see a clear trend that business models are adapted or completely changed by putting data into the center of their operations. Managing these data from infrastructure, data ingestion, data consistency, data manipulation to data consumption becomes therefore even more critical especially when considering the huge amount of data, which are created in enterprises (BIG DATA). There are many systems and infrastructures available which deal with BIG DATA focusing on different aspects of the problem space like massive distribution, machine learning, and other topics. Making BIG DATA technologies available for enterprises in a way, that they can build their business on top is still an open challenge. In this presentation we demonstrate on the example of SAP HANA BIG DATA what enterprise readiness means for BIG DATA solutions and how the architecture of such a system looks like. Beside the enterprise features, special focus will be set to the extensibility of the system and the ingestion process.