

Expressing and Running Big Data Analytics

Dr. Berthold Reinwald

IBM Almaden Research Center, San Jose (CA)

Tuesday, 06 August 2013

1:00 pm - 2:00 pm

Room: INF 3105

(building of the Faculty of Computer Science, Nöthnitzer Straße 46)

Abstract: Analytics on Big Data are ubiquitous in virtually every industry ranging from test analysis in manufacturing, information correlation in health sciences, churn analysis in telecommunications, fraud analysis in finances, and traffic analysis in transportation. Expressing and running these analytics for varying data characteristics and at scale is challenging. In this talk, we will discuss some of the challenges in building large-scale analytical systems. We share our approach of implementing declarative machine learning on MapReduce, describe our design principles and goals, and discuss some of our lessons learned.

Bio: Dr. Berthold Reinwald is a Research Staff Member in the database group at IBM's Almaden Research Center. His current research areas include scalable analytics platforms and database technology. He currently leads a research team on large scale machine learning focusing on building platforms to enable the implementation of algorithms to run on large amounts of data. Dr. Reinwald holds a Ph.D. in Computer Science from the University of Erlangen-Nuernberg, Germany.

