Contribution ID: **16**   Type: **not specified**

Big Data Analytics with Stratosphere: A Sneak Preview

Tuesday 28 May 2013 17:50 (5 minutes)

In this talk I will give a sneak preview of Stratosphere, an open-source software stack for parallel analysis of "Big Data". Stratosphere combines features from relational DBMSs and MapReduce: it enables "in situ" data analysis using user-defined functions, declarative program specification, and automatic program optimization, covering a wide range of use cases, from data warehousing to information extraction and integration. Further, Stratosphere covers use cases such as graph analytics and Machine Learning by integrating support for iterative programs in the system's optimizer and runtime engine. In particular, I will highlight the need for declarative languages and automatic parallelization and optimization of complex data analysis programs that involve iterative computation, and show how to achieve this using a combination of database query optimization and compiler techniques.

Author: TZOUMAS, Kostas (T)

Presenter: TZOUMAS, Kostas (T)

Session Classification: Lightning Talks