

EINLADUNG

Zeit: Freitag, 12.05.2017, 13:00 Uhr

Ort: Raum 5054, Ahornstraße 55

Referent: Edgar Gabriel, Associate Professor
Computer Science, University of Houston

Titel: File I/O in the Era of Large Scale Data Analytics

Abstract:

File I/O increasingly dominates the overall execution time of many data intensive applications and has consequently been identified as one of the major obstacles towards Petascale systems. This talk discusses the I/O infrastructure and data access methods for very large data files on High Performance Computing (HPC) systems and in Big Data environments.

The first part of the talk focuses on the architecture and the main features of the OMPIO parallel I/O library, the default MPI I/O implementation used by Open MPI starting from version 2.0 for most file systems. We discuss the feedback received with the library since its public release and outline the currently ongoing work.

In the second part, we explore the I/O infrastructure and data access methods used in Big Data environments and applications. While there are numerous similarities between traditional HPC and Big Data environments, the talk will highlight a few important differences between these two eco-systems and the resulting consequences.

Es laden ein: Die Dozenten der Informatik