TEXAS TECH UNIVERSITY
DEPARTMENT OF COMPUTER SCIENCE
WHITACRE COLLEGE OF ENGINEERING

U-REASON SEMINAR SERIES FALL - 2013

Locality-driven High-level I/O Aggregation for Processing Scientific Datasets

By Jialin Liu
Texas Tech University

Date: October 15th, 2013 (Tuesday)
Time: 3:40pm-4:40pm
Venue:ECE 108 (Pulsed Power Conference Room)

Note: Please note that the venue is changed due to unavailability of bullen room for only this seminar and we will continue to have the next seminar series in bullen room.

Faculty Coordinator: Dr. Yong Chen (yong.chen@ttu.edu)
Student Coordinators: Navaneeth Thiagarajan, Dan Ferguson, Lakhan Jhawar

Abstract:

Scientific I/O libraries, like PnetCDF, ADIOS, and HDF5, have been commonly used to facilitate the array-based scientific dataset processing. The underlying physical data layout in formation, however, is usually hidden from the upper layer¹slogical access. Such mismatching can lead to poor I/O. In this research, we have observed performance degradation in the case of concurrent sub-array accesses, where overlaps among calls that access sub-arrays led to high contention on storage servers due to the logical-physical mismatching. We propose a locality-driven high-level I/O aggregation approach to addressing these issues in this work. By designing a logical-physical mapping scheme, we try to utilize the scientific dataset¹s structured formats and the file systems¹ data distribution to resolve the mismatching issue. Therefore the I/O can be carried out in a locality-driven fashion. The proposed approach is effective and complements existing I/O strategies, such as the independent I/O or collective I/O strategy. We have also carried out experimental tests and the results

Speaker Bio:

Jialin Liu is a 3rd year PhD student in the Data-Intensive Scalable Computing Laboratory. He received his bachelor degree in 2011. His research focuses on the scientific data management, parallel I/O, computer architecture and system software support for big data analysis.