

# CS554 Project Ideas

## FusionFS:Sched - Data-Aware Scheduling for Distributed File Systems

### Overview

FusionFS [1] is a new distributed file system designed for exascale systems. In the current implementation, data movement is not optimized to increase data locality. For example, a job's data is written into the local storage and will be transferred to the requested node for reading, on the fly. While this strategy is optimal for some specific workload (e.g. checkpointing), a more general and intelligent mechanism is desired for achieving higher end-to-end data throughput and shorter wall time runtime of applications. In this project we will design and implement a scheduling algorithm, implement it in the Matrix [2] framework, and create a communication interface between Matrix and FusionFS in order to allow tasks to be scheduled close to the data. Work from 2009 [3] has insight into the centralized data-aware task scheduling system Falkon. Paper [4] outlines the support in MATRIX for data-aware scheduling.

### Relevant Systems and Reading Material

Please refer to the following papers for the current status of this project:

[1] Dongfang Zhao, Zhao Zhang, Xiaobing Zhou, Tonglin Li, Ke Wang, Dries Kimpe, Philip Carns, Robert Ross, and Ioan Raicu. "FusionFS: Towards Supporting Data-Intensive Scientific Applications on Extreme-Scale High-Performance Computing Systems", IEEE International Conference on Big Data, 2014.

Available online: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7004214>

[2] Anupam Rajendran. MATRIX: Many-task Computing Execution Fabric, *Master thesis, Illinois Institute of Technology*, 2013. [http://datasys.cs.iit.edu/publications/2013\\_MS-IIT\\_MATRIX.pdf](http://datasys.cs.iit.edu/publications/2013_MS-IIT_MATRIX.pdf)

[3] Ioan Raicu, Ian Foster, Yong Zhao, Philip Little, Christopher Moretti, Amitabh Chaudhary, Douglas Thain. "The Quest for Scalable Support of Data Intensive Workloads in Distributed Systems", ACM HPDC 2009.

[http://www.cs.iit.edu/~iraicu/research/publications/2009\\_HPDC09\\_data-diffusion.pdf](http://www.cs.iit.edu/~iraicu/research/publications/2009_HPDC09_data-diffusion.pdf)

[4] Ke Wang, Xiaobing Zhou, Tonglin Li, Dongfang Zhao, Michael Lang, Ioan Raicu. "Optimizing Load Balancing and Data-Locality with Data-aware Scheduling", IEEE International Conference on Big Data 2014

[http://datasys.cs.iit.edu/publications/2014\\_BigData14\\_data-aware-scheduling.pdf](http://datasys.cs.iit.edu/publications/2014_BigData14_data-aware-scheduling.pdf)

### Preferred/Required Skills

- Principles: operating system, distributed systems, computer network
- Programming: Shell Script, Perl/Python, C, C++, PThread, sockets, FUSE
- Operating systems: Linux

### Evaluation and Metrics

I/O for a variety of workloads, data locality achieved; experiments are expected to be conducted on the Amazon EC2 cloud on up to 128 VM instances.

### Project Mentors

- Dongfang Zhao [dzhao8@iit.edu](mailto:dzhao8@iit.edu)
- Ke Wang [kwang22@hawk.iit.edu](mailto:kwang22@hawk.iit.edu)