CS554 Project Ideas

FusionFS:Lib – Improving FusionFS Performance through User-level Library Interfaces

Overview

FusionFS [1] is a new distributed file system designed for exascale systems. FusionFS achieves a scalable data and metadata throughput by co-locating the storage to its computation, as well as a completely distributed metadata management component built upon ZHT [2]. A key issue in the current implementation of FusionFS is that it requires the FUSE kernel module to support the POSIX interface, wherein extra user-kernel context switches are introduced that result in huge overhead. In this project, you will implement a user-level library to bypass the FUSE module, so that the performance could be further improved. We expect to integrate this implementation in the next release of FusionFS.

Relevant Systems and Reading Material

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

[1] Dongfang Zhao, Chen Shou, Zhao Zhang, Iman Sadooghi, Xiaobing Zhou, Tonglin Li and Ioan Raicu. FusionFS: a distributed file system for large scale data-intensive computing, *2nd Greater Chicago Area System Research Workshop*, 2013. Available online: http://datasys.cs.iit.edu/~dongfang/download/poster_fusionfs_revised.pdf

[2] Tonglin Li, Xiaobing Zhou, Kevin Brandstatter, Dongfang Zhao, Ke Wang, Anupam Rajendran, Zhao Zhang, Ioan Raicu. ZHT: A Light-weight Reliable Persistent Dynamic Scalable Zero-hop Distributed Hash Table, *IEEE International Parallel & Distributed Processing Symposium*, 2013. Available online: http://datasys.cs.iit.edu/~dongfang/download/2013_IPDPS13_ZHT.pdf

Preferred/Required Skills

Principles: operating system, distributed systems, computer network, key-value stores

Programming: Shell Script, Perl/Python, C, C++, PThread, sockets

Operating systems: Linux

Project Mentor

Dongfang Zhao

Email: dzhao8@hawk.iit.edu