【文件系统】从2014年进入研究生阶段开始接触文件系统，就对文件系统有了浓烈的兴趣。刚开始接触的是单机文件系统，例如Journaling FS、Log-structured FS和Copy-On-Write FS，熟读过EXT2的源代码，并经常给实验室的同学分享。随着新型存储介质的火热，也了解一些针对新型存储器（Flash和NVM）的文件系统，例如F2FS，SFS等。后来开始接触分布式文件系统，也是目前正在研究的主要方向，带领实验室团队负责整个系统的搭建、运维和研发，目前主要研究Ceph上的元数据管理问题，包括元数据的预取和分布、负载均衡、热点优化等问题。作为实验室第一批接触文件系统的同学，深知做层次系统的艰辛，但又有很强烈的兴趣，希望可以为文件系统的发展出一份力。

【基于关联性的分布式文件系统元数据预取方案】本项目主要是基于文件在数据方面的关联性引出的存取关联性，通过利用轻量级的预取机制以优化元数据访问流程，从而缩短元数据请求的存取延迟。项目主要分为以下几个方面的工作：

1.inode 结构扩展 – 将关联文件信息以 KV 对的形式存储到文件 inode 的扩展属性中以透明支撑关联性；

2.关联性提取 – 对写入的数据进行语义分析，并提供语义分析模块接口和用户接口提高操作效率；

3.关联性更新 – 针对文件覆盖写和删除等操作引入延迟更新等轻量级更新方法；

4.关联元数据预取 – 引入客户端动态反馈机制以提高预取的准确率，并预取关联文件的目录信息；

5.原型系统实现 – 基于 Ceph 实现 SpeedyMeta，关联元数据请求延迟缩短 70%，整体的存取延迟缩短 35%。

本人作为该项目的负责人，从最初的发现文件数据关联性，验证关联性的重要性，到系统设计中的关联性的表现方式、在文件系统中的存储模式、对文件系统操作的影响和高效预取方案的设计，都亲自动手完成，尤其在系统实现方面，通过阅读和调试ceph源码了解其运行流程、在ceph里实现该项目的整体架构以及最终的性能测试方面都自己动手完成。从一步步调试源码、查看log文件发现和解决问题，在失望和希望中不断的迭代、直到最终项目完成，通过这几个阶段的工作加强了自身的学习能力和解决问题的专注力。

要说最期望去的部门应该是阿里云-盘古吧，盘古作为国内最先进的分布式存储部门，一直致力于存储系统、文件系统的设计和研发，为公司提供可靠的存储环境。