

基于 HDFS 的校园云存储系统的设计与实现

作者：于哲

随着云计算技术的快速发展，云存储服务受到越来越广泛的关注。在校园环境中，老师、同学、行政人员仍然使用传统的方式进行数据的存储和共享，缺少对复杂社交关系的灵活支持，存储容量难以弹性扩展，数据存储的可靠性低，存储媒介如硬盘的损坏容易导致数据丢失。

研究针对校园云存储的时代背景，通过对比现有产品如百度云盘、QQ 网盘的功能特性，分析其在校园环境使用的不足。结合校园应用场景，针对校园应用对数据存储基本操作、数据传播与共享、存储空间弹性扩展、多终端访问、以及数据冗余与可靠性的需求，设计系统的整体架构与核心功能。通过元数据对存储内容进行管理，设置存储空间管理、文件去重、延时删除等机制，给出解决方案。最后，根据设计要求进行系统编码实现，完成部署、测试工作。经过测试，系统达到了需求的目标，实现了完整的软件开发过程。

论文设计并实现的 IStudy 校园云存储系统(以下简称 IStudy Campus)，基于 Hadoop 生态系统底层的分布式文件系统（Hadoop Distributed File System，HDFS）实现核心的云存储功能，并与数据库、移动应用技术、以及软件服务技术相结合，提供安全可靠的、易于扩展的、多终端可访问的数据存储和业务访问服务。在业务功能方面，除文件上传下载等核心功能外，增加对校园社交关系的支持，并以此为基础实现文件数据的分享和去重；在终端支持方面，包含 Web 浏览器和 Android 客户端两种终端访问方式，提供便捷的数据存储与访问体验；在后台软件服务方面，采用 SSH 框架实现校园业务逻辑，为校园用户服务提供后台支持。