

大量非结构化数据带来存储和管理挑战

广东省农村信用社联社成立于 2005 年 8 月 5 日，拥有 99 家联社（农村商业银行），5622 个营业机构，是广东省内营业网点最多、服务面最广的金融机构。2009 年，广东农信正式上线了数据大集中项目。“虽然广东农信实现数据大集中的时间并不早，但由于广东农信的数据规模庞大、覆盖全省营业机构，该项目成为全国迄今为止最大的农村信用社数据大集中项目，克服了巨大挑战。”广东农信信息技术部总经理韦旦表示。

数据大集中之后，数据特别是非结构化数据的存储和管理就成为了广东农信必须解决的问题。目前，广东农信在大系统运行的账务数据达到一亿一千万，每天交易量约一千万左右。

为配合大数据量影像文件的存储、在线查询和管理，同时兼顾日后可能出现的电子文档、影像、多媒体、报表、邮件类等非结构化数据的存储和管理，建设集中的非结构化数据存储管理系统的需求日益突出。2011 年，广东农信开始建设后督系统，在当时是国内业界最大数据量的非结构化数据存储管理系统应用案例。

通力合作打破架构和技术瓶颈

广东农信的的后督系统可供全省各农合机构事后监督中心和网点使用，能够监督各项业务处理的正确性、合规性、真实性和完整性，及时发现各种核算差错事故，暴露业务核算中发生的各种违章、违纪、违法行为，完善业务操作风险控制体系，实现差错处理的电子化流转，实现凭证的电子化管理等。该系统主要包括影像处理、重点监督、再监督、风险预警、差错处理流程管理，实物档案流程管理等功能。

经过对后督系统数据存储的需求进行分析，广东农信发现，如此大规模的文件存储将打破操作系统文件系统的限制。按照目前的数据量进行存储，三五年后就会达到数据库的上限，而根据行业规范很多票据影像要保存十几年甚至几十年。

韦旦先生表示：“在广东农信应对大数据挑战的过程中，[IBM ECM](#) 在广东农信后督系统的成功运行起到了非常关键的作用。经过广东农信信息技术部的广泛考察和调研，发现只有 [IBM 的 ECM 解决方案](#) 能够适应广东农信如此大规模数据的‘苛刻’需求，同时完成省级集中部署、分批上线各级农商行业务。在后督系统项目的整个规划、架构、调优过程中，IBM 和广东农信技术专家的通力合作，最终实现了技术数据库和系统架构的优化。”

后发制人获得大数据竞争力

目前，基于 [IBM 的 ECM](#) 实施的广东农信后督系统能够满足每日新增 400 万张 100KB 影像文件的存储和生命周期管理的需求。在 2013 年 7 月内容管理平台正式上线后，系统内的数据量为千万级（约为 6000 万），每天新增约 80 万张凭证图像，装载至高速存储花费 40 分钟，迁移至近线盘阵花费 30 分钟，性能表现达到预期。

韦旦先生表示：“尽管广东农信的数据大集中以及非结构化数据存储管理做得并不早，但是后发优势也非常明显。目前该系统覆盖了全省三分之一的农商行，而在 2014 年计划将该系统实现全省农商行的覆盖。那时，将实现全国农信系统最大规模的全省集中的非结构化数据存储查询以及利用票据影像进行风险控制的企业内容管理系统。”

作为一个应用广泛的企业内容管理平台，[IBM ECM](#) 的核心优势在于其领先的数据模型设计，能够保证系统内的大量非结构化数据可以很好地被存储、管理和使用。IBM 中国开发中心企业内容管理、业务分析及能源电力解决方案开发总经理吉燕勇先生表示：“IBM ECM 不仅在国际上有很多金融机构具有深刻实践，还在中国具有强大的研发和技术支持实力。IBM 中国开发中心承担 IBM ECM 产

品的核心开发, 覆盖 **ECM** 几乎所有产品线, 包括文档影像化和捕获、案例管理、信息生命周期管理等。在应用系统研发和实施过程中, **IBM** 的支持团队始终提供及时的支持以及必要的现场服务来建议最佳的产品实施经验, 帮助广东农信顺利完成后督系统的全面覆盖应用, 在大数据时代提升竞争力。”