一种数据加密存储服务的设计探究

郭建超

（工业和信息化部网络安全产业发展中心（工业和信息化部信息中心），北京　100846）

摘要：随着云计算、移动互联网、大数据等技术的发展，信息大数据带来巨大便利，催生了新机遇，数据红利热潮正在加速到来。但同时，也带来了新的隐患，严重威胁个人、企业乃至国家机关单位网络信息安全的事件屡见不鲜，数据安全、隐私保护等挑战日趋严峻，如何确保数据安全是当前面临的重要而紧迫的任务。基于此，本文设计了一种数据加密服务，可针对结构化数据和非结构化数据进行有效加密，确保数据安全。

关键词：数据安全；加密服务；结构化数据；非结构化数据

中图分类号：TP309　　**文献标识码：**A

DOI：10.12230/j.issn.2095-6657.2022.29.027

目前，我国云计算、移动互联网、大数据等技术得到了快速发展，从行业云、企业云、政务云等各类云平台到智慧城市，公众、企业和政府机构等对 IT 系统的依赖也日趋加强，社会每天产生的数据更是以几何倍数增长 [1]。据中国社科院发布的《中国数字经济规模测算与“十四五”展望研究报告》显示，2019 年中国数字经济增加值规模为 170293.4 亿元，在同期 GDP 中的占比达 17.2%。

作为数字技术的关键要素，全球数据爆发增长、海量集聚，成为实现创新发展、重塑人们生活的重要力量，事关各国安全与经济社会发展。数据意义也因此从原先虚拟的字节符号成了如今核心的“生产要素”和“数字黄金”,颠覆着我们的社会生活、商业、产业模式，改写着城市乃至地球的未来。

对某项目管理系统（以下简称“某系统”）进行数据加密，某系统实现了信息技术与行政管理的有机结合，但伴随的数据安全问题也更加突出，确保系统的数据安全是当前面临的一大挑战。

1 数据安全风险分析

1.1 结构化数据安全风险分析

结构化数据简单来说就是存储在数据库中的数据。2020 年国家互联网应急中心（CNCERT）发布《2019 年中国互联网网络安全报告》，其中专门对我国境内数据库隐患排查及处置情况进行公布，据安全内参分析：CNCERT 针对境内的 Oracle、MongoDB、ElasticSearch、MySQL 及 Redis 等数据库进行排查，发现大量存在数据泄露隐患，共涉及 20720 个数据表、1330.3TB 数据量、1.78 万亿余条数据 [2]。

本次要做数据加密的系统数据库为 Oracle，数据采用明文存储，存在拖库、泄密、非法入侵窃取数据、内部高权限用户窃取数据、合法用户违规访问数据等数据安全风险，解决目前系统的数据安全问题，需要采用更加安全可靠的信息安全保护手段。

1.2 非结构化数据安全风险分析

非结构化数据指数据结构不规则或不完整、没有预定义的数据模型、不方便用数据库二维逻辑表来表现的数据，包括所有格式的办公文档、文本、图片、各类报表、图像和音频 / 视频信息等 [3]。

某系统非结构化数据安全面临的威胁主要有三点：（1）系统用户身份泄漏，文件拖库；（2）存储介质中的数据是以明文方式保存，使得内部或者外部的入侵者可以轻易定位并非法获取和篡改数据；（3）内部高权限用户窃取数据，合法用户违规访问数据。

2 数据安全需求分析

2.1 结构化数据加密存储服务需求

某系统主要分为项目管理、成果管理、评奖管理等模块，所涉及数据库表主要有科技项目计划信息表、科技项目申报书信息表、项目执行情况年报信息表、科技项目变更信息表、科技项目结题信息表、科技项目验收信息表、科技项目成果登记信息表、科技项目成果应用信息表、科技项目成果转化信息表、项目评奖推荐信息表。

以上数据库表所涉及的项目名称、项目实施方案与进度安排、成果形式与主要内容、本年度计划进度、技术考核指标等字段信息，需要进行数据加密。

2.2 非结构化数据加密存储服务需求

经梳理归纳，某系统中非结构化数据主要包括内容见表 1。

表 1　非结构化数据内容表

序号	内容	格式
1	结题验收附件	文本文档
2	项目成果登记附件	文本文档
3	项目成果转化登记附件	文本文档
4	QC成果登记附件	文本文档
5	项目评奖申请附件	文本文档
6	项目评奖发放附件	文本文档
7	专家组维护附件	文本文档

本项目中，结题验收、项目成果登记、项目成果转化登记、

QC 成果登记等附件需要进行加密存储。

2.3 加密存储服务及加密数据使用情况监测及分析需求

为方便系统管理人员及时了解数据加密情况，掌握系统数据安全动态，需要对数据库表加密情况和字段情况、加密数据访问量和运行效率等做统计分析。

3 数据加密存储服务建设方案

3.1 建设目标

数据加密存储服务作为保障某系统数据安全的核心建设内容，要突破传统数据安全保护技术瓶颈，应用国密算法对结构化和非结构化数据加密存储及访问，实现数据安全、应用透明、密文高效调用的建设目标[4]。

3.2 建设思路

（1）结构化数据加密

某系统中结构化数据主要是数据库表和字段，我们基于透明代理加密技术，通过对处理流程上的创新和密码算力上的突破，对系统数据库字段加密存储，并实现全密态状态下数据的增、删、改、查功能。

（2）非结构化数据加密

某系统中非结构化数据主要是文本文档的形式，我们采用面向块的安全存储网络（SAN 安全存储）。

（3）实施过程

针对结构化数据，我们提供结构化数据加密服务。结构化数据加密存储服务启动时，首先将实体数据库中的数据、表项等各类数据映射到结构化数据加密存储服务内存镜像中，保证结构化数据加密存储服务中的数据与实体数据库中的数据一致。通过结构化数据加密存储服务对数据加密并进行入库处理，应用直接连接结构化数据加密存储服务完成数据操作。

针对非结构化数据，我们提供非结构化数据加密服务。某系统链接到非结构化数据加密存储服务中发送存储数据，非结构化数据加密存储服务接收到数据文件后，调用内部的密码服务中间件，采用国密算法进行数据加密。数据加密为密文，并将密文同步至数据服务存储区。

3.3 建设方案

项目分为结构化数据加密建设和非结构化数据加密建设，系统逻辑架构图如图 1 所示。

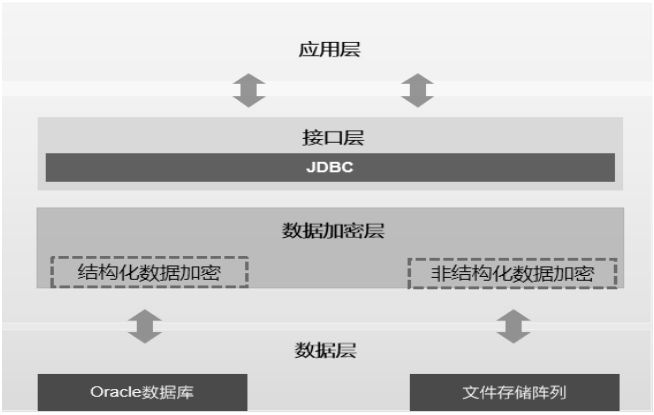

图 1　系统逻辑架构图

（1）结构化数据加密存储服务建设

结构化数据加密存储服务采用透明代理技术，实现数据的系统数据库字段级、全库的数据加密存储，并实现全密态状态下数据的增、删、改、查功能。应用系统通过透明代理技术提供的 JDBC 数据库应用接口连接数据库，调用接口的时候，根据设定的加密策略，对数据库字段级、全库的数据加密存储。

1）服务核心模块功能

a）密码服务模块

结构化数据加密存储服务内置高性能国密密码卡，并自主研发密码服务中间件，用于提供身份认证和加解密服务，保障数据访问的安全性、数据存储的保密性和完整性。其中身份认证服务主要面向管理员用户和各类客户端用户，均支持采用 USB-KEY+ 随机数签名方式进行强身份认证。数据加解密服务通过内置的高性能国密密码卡，采用国密 SM4 算法进行数据加密，密钥安全存储在密码卡安全硬件密码模块中。

b）密钥管理模块

密钥管理为结构化数据加密存储服务的核心模块，密钥管理是保障数据库安全的重要保障。结构化数据加密存储服务内部有 2 个随机源，位于内置密码卡上。结构化数据加密存储服务会在上电自检过程中对所有随机源进行检测，确保所有随机源的正确性。在服务过程中，服务会有单独线程进行随机源检测，保证所有随机源的正常工作，若发现有不正常的随机源后，通知主服务不再使用该随机源进行工作。在每次生成随机数时，都会对产生的随机数进行检查，符合随机性。通过以上几步操作可以有效地保证产生高质量的随机数。

c）内存镜像模块

该模块主要对数据库原库、结构化数据加密服务镜像库、镜像库状态等进行统一的管理，用于结构化数据加密存储服务与数据库建立连接，数据通过结构化数据加密存储服务加密后自动连接数据库系统进行存储。其中数据库原库功能，可通过输入数据库的 IP 地址、端口号、数据库名称内容，与结构化数据加密存储服务进行连接。

d）服务管理模块

支撑结构化数据加密存储服务的管理与运维功能，包括权限管理、日志管理、审计管理、系统管理、网络管理、设备管理和用户管理等内容。其中日志管理模块支持日志等级设置，可根据级别进行管理；支持日志导出功能，系统日志可通过接口下载到本地；支持日志审计功能，根据日志状态可进行删除、恢复、查询和审计功能；支持日志批量导出功能。

权限管理除超级管理员 admin 用户外，其余用户需要在权限管理模块进行添加和权限授权，即先添加用户，再将用户添加到组，最后以组的形式进行授权，授权后再以新用户身份使用结构化数据加密存储服务。所有用户均需通过 key+ 随机数签名方式进行身份认证。

e）结构化数据加密监控模块

通过数据加密监控模块，可以监控数据库加解密统计（如数据库数量、加密的表数量和加密的字段数量）等。

2）结构化数据加密设计

根据某系统的数据库表结构，考虑到各字段的敏感性，本着既能减少对数据库访问效率的影响，又能较好地提高数据库安全性，我们对项目管理、成果管理、评奖管理等模块所涉及的数据库表字段选择敏感字段进行加密。

3）系统对接改造

结构化数据加密存储服务支持 Oracle 数据库的 JDBC 接口集，内置 SQL 引擎可直接支撑数据库的增、删、改、查等操作。针对某系统的改造内容，不涉及系统代码层和架构层的改动，仅需将原访问 Oracle 数据库的地址连接到结构化数据加密存储服务即可。

4）Oracle 数据库对接改造

Oracle 数据库层面的对接，涉及数据库自身与结构化数据加密存储服务对接，及数据库目前已建的表格式的改造。

目前结构化数据加密存储服务已适配 Oracle 数据库，可直接对接数据库管理系统，Oracle 数据库本身无需进行改造。结构化数据加密存储服务可对数据库的表项进行加密，加密后进行密文存储，则密文的数据格式统一为 VARCHAR，字符长度根据默认长度调整为 16 的整数倍。

5）数据库加密模式

结构化数据加密存储服务启动时，首先将实体数据库中的数据、表项等各类数据映射到结构化数据加密存储服务内存镜像中，保证结构化数据加密存储服务的数据与实体数据库中的数据一致。通过结构化数据加密存储服务对数据加密并进行入库处理，应用直接连接结构化数据加密存储服务完成数据操作。

功能架构分为前端的系统管理模块、后台支撑模块以及密码服务模块。

前端系统管理模块包括：设备监控、日志管理、系统设置、密钥管理、权限管理、网络配置以及模式的配置管理。

后台支撑模块为系统运行提供支撑服务，如高可用性模块提供双机热备、资源调度数据备份 / 恢复功能，保障数据的高可用性（防止单点故障）和数据容灾能力；系统监控模块提供系统运行信息、操作信息的监控报警能力，密码加速以及数据同步和密态检索的能力。

密码服务模块主要用于提供身份认证和加解密服务，保障数据访问的安全性、数据存储的保密性和完整性。

（2）非结构化数据加密存储服务建设

针对本系统中非结构化数据，我们提供针对面向块的安全存储网络（SAN）功能，存储粒度基于卷和扇叶分区。

1）服务核心模块功能

a）安全存储模块

安全存储模块可以提供 SAN 安全存储服务，SAN 安全存储支撑 IP–SAN 和 FC–SAN 协议的块级访问，支持应用非结构化数据等数据的加密存储。

非结构化数据加密存储服务以内置的高速密码卡为密码核心，实现高效率的文件数据加解密，并保证文件加密密钥的安全可靠。支持细粒度文件加密存储，实现一文一密加密方式。

b）密钥管理模块

与结构化数据加密存储服务的密钥管理模块类似。

c）服务管理模块

与结构化数据加密存储服务的服务管理模块类似。

d）非结构化数据加密监控模块

与结构化数据加密存储服务的数据加密监控模块类似。

2）工作方式

非结构化数据加密存储服务 SAN 提供块级存储，存储粒度基于卷和扇叶分区，用户通过 Ukey+ 随机数签名认证后，会在本地设备中存在一个共享磁盘映射。用户将需要加密的非结构化数据存放到该共享磁盘中，数据会调用密码服务对数据进行加密，存放到共享磁盘的数据就是已经加密好的密文数据。同理，当用户进入共享磁盘操作时，打开文件系统会自动调用解密服务，把需要的密文解密成明文数据呈现给用户，实现了透明加解密。当其他非法用户直接访问磁盘阵列时，数据未经过解密，那么该用户将无法获取明文数据内容，从而实现数据的安全存储 [5]。

4 结语

综上所述，随着云计算、移动互联网、大数据等技术的快速发展，公众、企业和政府机构等对 IT 系统的依赖日趋加强，但同时也带来了新的安全隐患。本文提出的一种针对数据安全的解决方案，对某系统进行了数据安全风险分析，分为结构化数据和非结构化数据，并提出了有效加密方法，确保了数据的安全。

参考文献：

[1] 蔡跃洲，牛新星，Cai Yuezhou，等 . 中国数字经济规模测算及“十四五”预测 [M]. 北京：社会科学文献出版社，2021.

[2] 孙友添 .2019 年中国互联网网络安全报告出炉 [J]. 计算机与网络，2020，46（20）：4.

[3] 范利辉 . 一种非结构化数据的处理方法及防复制加密传输系统：CN111143342A[P].2020.

[4] 魏巍，周薇，邵千芳，等 . 一种基于国密算法的数据加密保护方法：CN110704839A[P].2020.

[5] 张芸，黄炎，高文胜 . 面向不动产登记业务应用的数据库安全加密方案 [J]. 国土资源信息化，2020，（05）：5.

作者简介：郭建超（1989–），男，山东东平人，硕士研究生，主要从事信息化项目建设管理、数据安全研究。