油田大数据安全与策略

刘志衡（大庆油田有限责任公司采油工程研究院，黑龙江　大庆　163458）

摘　要：在当前互联网时代，大数据在人们的生活生产中早已全面渗透，为各行业领域的发展贡献了一定力量。石油行业信息化建设经过多年发展，已积累了大量生产与经营方面的数据，有效消除了业内信息孤岛问题，全面提高了油田企业的管理水平。然而，我们在享受油田大数据带来的便利与高效率的同时，也要认识到数据安全风险问题的存在，为了保证油田企业网络发生故障或瘫痪时数据的完整性与一致性不会受到影响，则要针对油田大数据安全防范策略展开研究，打造更安全的油田数据信息体系。鉴于此，本文将结合笔者工作经验，对油田大数据安全与策略展开全面分析，希望能为从业同行提供一定参考。

关键词：油田企业；大数据；安全；策略

0 引言

伴随着油田开发的持续深入以及油田企业信息化建设的不断完善，越来越多数据随着开采、经营、管理工作的开展而被收集与存储，呈现出种类繁多、体量巨大的特点，形成了大量不同类型的数据库。而形成这些大数据，能够为油田企业的后续发展提供决策参考，可见数据安全关乎油田企业的发展前景，一旦数据方面出现问题，都有可能造成油田企业信息化的停滞。所以，油田大数据安全问题俨然已成为如今油田企业信息化建设的关键核心。

1 油田大数据概述

大数据（big data）是指难以在特定时间内用常规工具进行收集、处理、管理的数据集，而需要一种新的处理模式才能发挥其决策力、发现力、优化能力的海量信息资产[1]。油田大数据属于属于大数据的概念范畴，但是本质上有一定差别，其并不是统计学的处理结果，而是油田行业经过几十年的勘察、开发、生产、科研等工作开展而积累的真实数据[2]。油田大数据是一项数以千TB计的庞大数据资产，而且每时每刻都在呈几何倍数地增长。油田大数据作为油田企业勘探开发中的重要载体，能够为资源勘察、油气田开采提供保障，具有真实、精准、可靠与价值的特点。

2 油田大数据安全问题与防范思路

2.1 安全问题

油田大数据的安全问题基本上可分为两个方面，一方面是数据存储的安全问题，另一方面是数据使用的安全问题。前者安全问题表现在存储介质、存储环境等方面，需要重点防范“天灾”；而后者作为使用阶段的安全，包含数据传输、安全加密、恶意攻击等，需要重点防范“人祸”。

2.2 防范思路

面对影响油田大数据安全的“天灾”与“人祸”，为了提高安全性则要先明确安全防范思路。

其一，备份是保证油田大数据安全的首要手段，笔者认为可采用异地备份方案以及完善可行的实时备份程序作为保障。异地备份实际上是指在异地构建完全一样的数据管理系统，这套系统在日常工作中不需要提供给用户使用，只负责对数据的同步备份更新。而一旦用户使用的那套系统发生磁盘故障、雷电击打等问题，则要立马启动异地备份系统，实现对损坏系统数据的远程恢复，保证正常运行不受影响。同时，完善可行的实时备份程序可实现对数据的定期更新与备份，能确保数据的实时性和完整性。

其二，随着油田企业信息化建设不断推进，高速网络在企业内部的应用日渐广泛，大量数据资源得到高效利用。而随着数据量增大及使用频率升高，在线服务的需求也逐渐超过离线服务，所以在大数据安全防范中要重点关注网络攻击、病毒入侵等方面。

其三，油田行业的发展需要大量不同层面、多学科的数据，过去点对点、面对面的数据复制模式已无法适应当前行业发展需求。随着油田企业数字化、信息化建设的推进，越来越多的数据可通过网络传输，用户对数据的存取与访问也更加便捷与频繁。所以，在油田大数据安全防范中除了要确保用户的操作需求外，还要防范人为发起的恶意攻击，所以要针对数据库构建防火墙机制。

3 油田大数据安全的具体防范策略

3.1 中心机房安全

油田企业大数据的安全，首先要保证数据运行物理环境方面的安全，这其中中心机房的安全是重中之重。中心机房需安装门禁系统，做好防火、防水工作，设置环境实时监控系统以及利用UPS电源保障电力不间断供应[3]。当然，上述工作仅是基础部分，要重点关注的是中心机房的视频监控系统，通过该系统的搭建，对机房内部状况进行动态化监控，包含所有设备指示灯显示状态、人员进出记录、人员操作行为等，均会以视频信息方式记录在硬盘中，存储设备也会放置在绝对安全的地方，一旦中心机房发生故障，监控视频信息存储设备便能发挥“黑匣子”的作用，通过回放录像找到事故原因。

此外，还要制定完善的机房容灾备份制度，且保证落实到位。因为油田行业存在一定特殊性，投入了大量

人力、财力及物力，才得以形成地质、勘探、开发等方面的宝贵数据信息，不仅是对我国油田发展历程的记录，也能为油田未来发展提供指引，可见油田大数据会影响到油田的发展前景。所以，要提前设想机房发生灾害时要如何快速恢复所存储的数据，其中异地备份属于较为保险的方案，具体来讲可在距离中心机房 30km 以上的地方建立数据备份机房，通过 VPN 专线以及采用远程磁盘景象技术，实现数据的实时同步与备份 [4]。如此一来，即便是发生火灾或地震，也能快速恢复油田企业数据资产。

3.2 网络安全

油田大数据汇聚着油田勘探、开发、生产、经营等数据资产，而数据中心作为对这些数据进行管理、处理、服务的专业化数据管理机构，发挥着信息系统建设、油田大数据管理与服务的多种功能。现阶段油田企业的数据中心所选用的设备多为阵列存储设备、服务器群和若干工作站组成，所有设备基本集成在交换机上，因此不管是从网络环境，还是从应用平台层面来看，均存在一定安全风险。

要想改善油田大数据网络方面的安全风险，最直接有效的方法便是搭建防火墙和设置防毒系统。借助划分 Vlan，将数据库服务器群与其他设备相互隔离，可防止网络病毒的蔓延影响；对数据中心的访问进程中设置防火墙，确保数据中心始终处在保护状态中，而防火墙的设置方面主要加强对进出数据中心的访问请求进行重点审计与控制，有效防范所有恶意攻击和影响数据安全的因素。

当然，防火墙并非万能，无法绝对保护油田大数据的安全，所以还要建立网络入侵防御系统，该系统能够对有恶意攻击意图和行为的用户进行记录，并且找出其攻击目标。此外，还可启用桌面安全管理系统，该系统能对计算机终端的安全性进行全面检查，通过在后台实时更新病毒库与为系统打补丁，能够消除绝大部分计算机终端因为病毒引发的网络安全隐患。通过数据库的补丁更新，能拦截各种影响数据库安全的新病毒，保障了数据库的安全性。此外，也要制定严格的网络准入制度，在制度中明确提出不允许非法计算机接入油田企业网络，所有入网计算机必须满足提前设定的数据库访问条件，避免非法用户窃取、篡改油田大数据。

3.3 数据安全

保障油田企业数据安全，基础在于保障数据库的安全，确保数据库内部数据的独立性、安全性、完整性，还要关注并发控制、故障快速恢复等多个方面。

在对油田数据库进行设计时，一定要重点关注数据在物理与逻辑方面的独立性。其中物理独立性指的是用户在应用程序、数据库中的数据是相互独立的；而逻辑独立性是指用户在应用程序与数据库中的数据在逻辑结构方面是相互独立。先进的数据库设计理念是保护数据安全的首要步骤，同时为了保证数据的完整性与一致性，还要做好对数据库并发事件的有效控制。数据库作为多用户共享资源，用户逐个串行操作，意味着同时间段只能有单个用户进行对数据库的存取操作，其他用户需等待该用户执行完毕才能进入。这一机制的弊端在于如果某个用户需要进行大量数据的存取操作，则会造成数据库系统长时间闲置。所以，为了发挥出数据库资源共享优势，应当允许多用户并行操作，但此举会导致如果对同时间段多用户的并发操作不加以控制，便会出现存取数据无法保持一致的情况，因此还需完善数据管理系统的并发控制机制。油田企业数据库可通过实施封锁机制解决并发操作问题，其能够确保任意时间段都能必行多个用户程序，且所有用户能在相互隔离的环境中操作。

对数据的访问审计与控制也是保证数据泄漏的重要手段。通过系统自动对用户的数据访问行为进行分析，基于用户单天的访问次数与访问量去判定其对数据使用的合法与否。如果系统发现某个用户单天对数据的存取量超过预设值，便会自动剔除用户且即时通知管理人员，以人工方式去审查该用户的行为合法性，如果判定不合法则立即拿回所有下载的数据，防止数据泄露。

数据的完整性及保密性同样是保证数据安全的关键点，其中健全的用户管理机制是保障数据安全的重要手段，通过实施用户授权的方式去控制访问行为，比如权限分配、口令设置等，如此一来权限不同的用户只能对油田大数据实施相应的操作，一定程度上能保证核心数据的安全性。同时，用户对数据库数据的存取行为也会留下痕迹，翻查用户“足迹”能了解其访问数据中心的时间、IP、设备及具体操作等信息，一切判定不合法的操作都会立即终止。

4 结束语

综上所述，在油田企业的信息化建设中，数据中心的建设必不可少，而且其中的数据量会随着油田开采的发展而不断增大，加之互联网技术、云计算技术的推广，油田大数据有了更多应用放线，但同时也对油田大数据的安全提出更高要求。保障油田大数据的安全是一项复杂、系统且对技术要求严格的工程，既要关心中心机房、网络的安全，也要关注数据本身的安全。所以，身为油田数据人，一定要提高自身安全意识，端正工作态度，不断更新专业知识储备，丰富数据安全防御手段，确保油田朝着数字化、信息化方向稳步前进。

参考文献：

[1] 李园园，胡璐．大数据分析技术在油田生产中的应用研究 [J]. 中国管理信息化 ,2019,22(08):56-57.

[2] 黄海燕．大数据环境下油田信息安全体系构建研究 [J]. 中国管理信息化 ,2020,23(04):62-63.

[3] 李海荣．大数据分析在安全生产中的应用 [J]. 石油石化节能 ,2021,11(04):38-42+11.

[4] 张云志．油田大数据分析模式研究 [J]. 信息与电脑（理论版）,2020,32(19):4-6.