

探索多元化发电数据驱动的发电企业数据中心建设方法

曹春兰, 黄彪斌, 郑志文, 林巧锋, 陈静霖, 翁疆

中国华电集团有限公司福建公司, 福建 福州 350001

DOI: 10.61369/TACS.2025050005

摘 要 : 本文旨在阐述发电企业大数据中心研究及应用的设计规划方案。其内容涵盖了各发电领域数据管理现状分析、数据中心架构设计、建设体实践方案等多个方面, 旨在为发电企业数据中心建设实施提供规划意见和参考。

关 键 词 : 发电企业; 数据标准; 数据治理; 数据中心

Exploring the Construction Methods of Diversified Power Generation Data-Driven Data Centers for Power Generation Enterprises

Cao Chunlan, Huang Biao-bin, Zheng Zhi-wen, Lin Qiao-feng, Chen Jing-lin, Weng Jiang

China Huadian Corporation Limited Fujian Company, Fuzhou, Fujian 350001

Abstract : This article aims to elaborate on the design and planning scheme for the research and application of big data centers in power generation enterprises. Its content covers multiple aspects such as the analysis of the current situation of data management in various power generation fields, data center architecture design, and construction practice plans, aiming to provide planning opinions and references for the implementation of data center construction in power generation enterprises.

Keywords : power generation enterprises; data standards; data governance; data center

引言

当前, 国际油价波动、科技信息化水平提高和环境问题, 给能源开发利用带来新挑战。各国开始新能源开发, 推动能源产业转型。云计算、人工智能等关键技术加速转型。新能源大规模开发利用、低能耗低污染低排放、先进信息技术驱动的新能源产业革命正在兴起, 能源朝着多元化、智能化时代迈进。

发电企业需要具备更强的感知决断行动能力, 以适应发展方式的重大变革。全球气候变化和生态环境保护对能源生产消费的绿色低碳化提出了更高要求。我国正在大力推进能源革命, 电力市场化改革也在推进售电侧放开。同时, 电力需求多样化使得市场环境变化加快, 电力交易的不确定性因素增加。国有企业改革要求企业转变发展方式和管理方式, 提高发展质量与效益。这要求企业能够更加敏锐地感知外部市场变化和内部运营状况, 更加优化地配置资源, 更加有效地响应市场变化。

发电企业的未来发展需要从信息化向智慧化转变。信息技术为企业发展和经营管理提供重要支撑。为了适应发展战略的要求, 发电企业需要从传统信息化向未来智慧化提升, 以更全面、及时准确地掌握内部运营状况和外部环境变化态势, 更好地分析预测市场变化与企业运营发展趋势。通过积累数据信息、转化为知识资产, 增强企业核心竞争力^[1]。

一、各发电领域数据管理现状分析

电力企业的业务与信息化现状与未来发电企业发展的要求仍有一定的差距, 具体体现在如下几个方面:

火电领域: 在火电领域数字化智能化方面, 相关标准尚不统一, 对火电厂数据的挖掘深度和广度仍有待提升。由于受到机组负荷、环境温度、燃料成分以及操作方式等多种因素的影响,

缺乏大数据算法的应用使得准确反映机组运行特性及效率变得困难。

风电、光伏领域: 目前集控中心仅实现部分监视功能, 且缺乏自动化、智能化监视手段, 数据汇集后的优势尚未完全挖掘, 对于机组可靠性、稳定性、安全性及经济性的提升尚未实现实际意义。同时, 在集控中心的建设与运行过程中, 数据接入仍缺乏统一标准。

水电领域：要求通过远程发电集控、智能水情预报、智能大坝安全监测，以实现全流域水情信息自动预警，并实现跨流域大坝群及库区地质灾害的全天候安全监测，从而避免地质灾害带来的安全问题。因此，生产数据治理及应用是发电企业向着发电业务集约化、精益化、标准化发展的重要基石^[2]。

建立发电企业大数据中心是顺应目前电力信息化技术水平发展、服务发电企业及下属各电厂的架构平台。它致力于打造一个以 PaaS 为核心技术的一体化大数据中心，通过融合分布式大数据技术、内存计算技术、数据仓库技术以及微服务开发等技术，提供完整 PaaS 云资源部署及管理的能力，能将现有各种业务能力进行整合，提供全面的技术支撑能力，满足企业数字化转型所需的各类应用开发场景，为业务开发、数据开发等各类应用开发者提供图形化、可视化和自动化地开放式开发、测试和运行环境，持续提升发电企业的业务应用开发速度、标准化实施和持续集成能力，全面支撑企业数字化、智能化转型。

同时，构建一套标准化的数据共享机制，以推动企业内部的数据流通与业务协同，从而为决策制定提供实时、精确且可信的信息支撑。此举不仅有助于提升企业工作的前瞻性与精准性，而且还能进一步强化宏观调控力度，有力推动企业的持续健康与稳健成长。

二、发电企业数据中心架构设计

(一) 总体架构

发电企业大数据中心的构建，是依托于工业大数据体系架构的技术基础。自下而上，细分为中心的各个关键层次：数据源层数据层是各类多源异构的设备或系统，为本项目大数据中心提供风、光、水、火、楼宇等领域各类生产实时数据、经营管理数据、综合业务管理数据以及相关外部数据资源；IaaS 层则负责平台运行所依赖的软硬件环境，包括服务器资源、存储资源、操作系统、虚拟化，以及网络资源等；平台层作为发电企业大数据中心的核心，其由数据采集、大数据基础平台、数据存储、应用开发、空间数据等子系统和组件构成；应用层是基于数据中心平台，以微服务的方式构建领导驾驶舱，电力交易相关数字化应用场景。同时，应用层也支撑未来更多的数字化应用；而门户层则为用户提供统一、便捷的访问入口。整个大数据中心建立在统一的标准规范体系、安全管理体系以及运维保障体系之下，确保了其稳定、高效地运行^[3]。如下图所示：

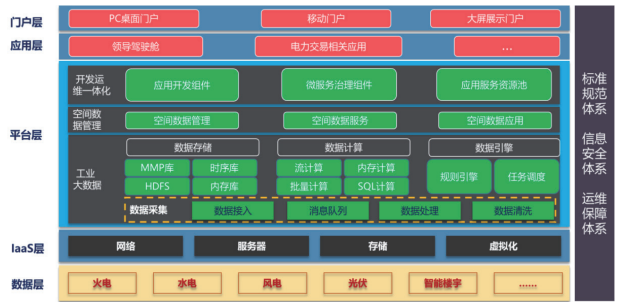


图1 数据中心总体架构

Fig. 1 Data Center Architecture Overview or Overall Data Center Architecture

(二) 数据架构

根据对发电企业业务需求的分析，大数据中心数据架构包括数据采集区、数据仓库区、原始数据区、实时分析区、数据分析区、数据访问区等大数据基础数据平台各个系统功能区，能实现并完全满足企业大数据中心建设的需求。



图2 数据架构

Fig. 2 Data Architecture

三、建设实践方案

(一) 数据采集

发电企业多元化发电数据的采集范围通常包括发电厂发电设备监控系统、计算机系统、设备状态监测系统、气象系统、电力交易平台以及营销、财务、综合业务管理等信息管理系统数据。

由于发电企业信息化系统建设时间跨度较大，技术架构以及数据架构存在形式各异，因此需要采用不同的数据接入方式，数据接入主要采用消息队列、ETL 数据接入、文件交换、网络采集程序、ESB 总线、消息接口、人工录入等方式。

在数据采集周期方面，数据采集系统需按照业务要求，实现实时、非实时业务对应的数据采集需要。实时业务即满足查询当前业务状态或将实时业务数据提供给在线监察、应用的业务类别；非实时业务是满足查询和分析处理实时性要求较低的业务类别。

(二) 数据处理

数据的处理和抽取需要利用自动化软件或自动化辅助工具实现，结合实际情况，在软件设计上应该支持数据的在线填报、电子表格导入、自动抽取等方式，满足不同的应用场景和客观条件。

(三) 数据存储

发电企业的生产经营管理活动会产生各种类型的数据，其中主要分为生产实时类数据和业务系统中的关系型数据。

在工业物联网系统的构建过程中，实时数据是一种极为常见且至关重要的数据类型。为了高效地管理和存储这类数据，通常采用实时数据库技术。实时数据库是基于实时数据模型构建而成的，它专为处理持续更新、快速变化的数据流以及具有时间约束的事务处理而设计。

关系型数据是发电企业在综合业务管理和经营财务管理中常用的数据结构类型。这种数据结构具有结构化程度高、关联性强、数据时间跨度长、数据量大以及对实时性要求相对较低等特

点。基于当前发电企业的信息化发展规模和业务体量，在建设数据中心时，我们不妨选用分布式存储型关系数据库。这类数据库具备动态扩展、高可用性和易维护性等优点，能够满足发电企业日益增长的业务需求^[4]。

（四）数据建设

1）数据分区设计

根据数据存储方式以及数据采集、处理、存储、挖掘分析和访问流程，确定数据分区架构设计，主要包括数据采集区、实时分析区、数据仓库区、原始数据区、数据分析区、数据访问区。数据分区设计图示如下：

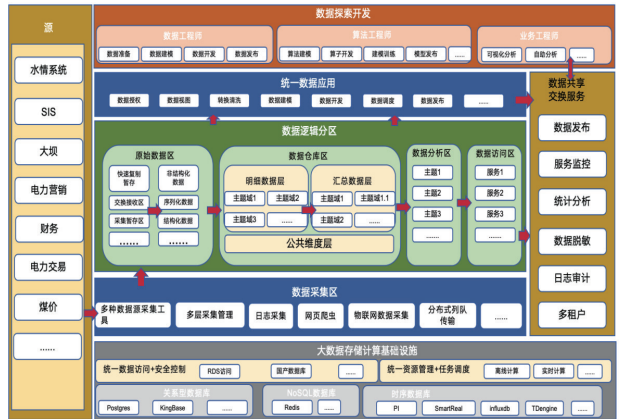


图3 数据分区架构

Fig. 3 Data Partition Architecture

数据采集区：主要承担着数据抽取、协议归一化、数据传送以及采集周期设定等核心任务。该区域针对多源异构数据的采集需求，提供了消息队列、实时数据同步、ETL、爬虫、API 接口等多种技术手段，确保了数据的高效、准确采集。采集的数据类型涵盖了结构化、半结构化以及非结构化等多种类型。

数据仓库区：作为数据平台的基石，负责将业务系统的各类数据进行有效集成，以满足海量数据的管理需求。同时，通过数据接口，将数据安全地提供给数据分析、挖掘以及各类应用系统使用。

原始数据区：主要功能是提供全量、历史、多元异构数据的存储，以及业务逻辑相对简单的数据加工处理服务。该区域基于 Hadoop 生态构建，数据类型丰富多样，包括图片、音频、视频、日志等所有数据。

实时分析区：旨在满足实时 / 准实时的业务需求，提供结构化或非结构化数据的实时加载、清洗、转换、汇总、统计分析等功能。该区域支持对设备的在线实时监测、电力实时负荷监测等应用，为业务决策提供实时支持。

在诸多行业领域中，建设、调试、运营等阶段往往会产生大量带有时间戳的传感器数据，这些数据典型地构成了时间序列数据。时间序列数据主要由各类实时监测、检查与分析设备所采集或产生，涉及火电、水电、风电、光伏等多元业态。

数据分析区：主要功能是提供一整套面向数据分析的统计和分析算法，以及支持大规模数据处理的快速通用计算引擎。此外，数据分析区还应具备支持基于算法的各类分析模型开发、测试、存储的能力，同时实现计算结果的可视化、可定制化和交互

性功能，从而提供完整的分析解决方案^[5]。

数据访问区：旨在满足各级用户、各业务部门不同权限用户的访问需求。该区域提供数据接口，实现多业务系统数据共享，并具备主管政府部门和外部协作单位获取数据的功能。

2）数据分层设计

按照数据流入流出的过程，涵盖集团主要业务线路，数据仓库架构分为五层：原始数据层 (ODS)、数据明细层 (DWD)、数据集市层 (DM)、维表层 (DIM)、数据应用层 (ADS)。

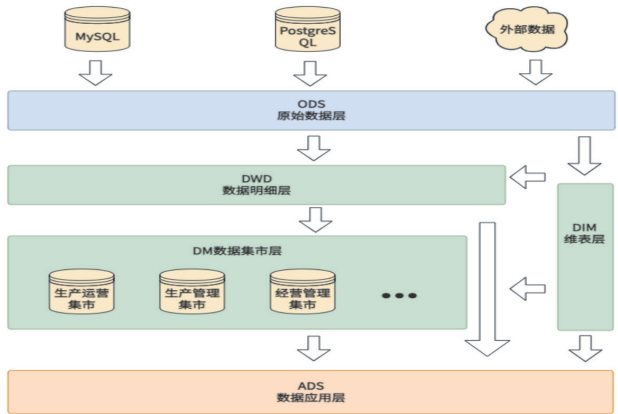


图4 数据分层架构图

Fig. 4 Data Layering Architecture Diagram

3）数据模型设计

数据模型以企业架构为核心，将数据资产、业务流程、数据标准通过科学的数据模型建模方法论和长期的数据治理实践经验以及企业实际业务需求结合，从而实现数据从“盘点数据”到“规范数据”再到“治理数据”的落地过程。

（五）数据标准化体系建设

数据标准化的行业标准可以根据不同行业和应用领域的需求而有所不同。电力行业是国家能源发展的关键领域，涉及能源安全和公共利益，网络安全问题备受重视。为提高电力监控系统信息安全的防范能力，国家能源局、电监会等监管单位基于国家标准规范，结合电力行业的业务特点，对电力行业的标准规范进行了不断的完善；同时各大发电集团在行业标准的基础上，也制定了符合各集团业务特点的标准文件，指导发电场站和集控中心的信息安全建设工作^[6]。

在发电企业中，数据分类对于实现高效的生产运营管理和优化具有举足轻重的作用。一个健全的数据分类体系能够全面覆盖发电企业从基础信息到生产运营管理的各个环节，确保数据的有效利用和管理的精细化。以下是针对发电企业核心业务的数据分类描述：

基础信息：

场站与机组信息：包括场站和机组的地理位置、规模、设备型号、容量、投运时间等关键参数。

特殊属性：涉及机组或场站的特殊技术要求、环保标准、能源效率等特定属性信息。

生产运行：

实时运行数据：涵盖发电量、设备运行状态、实时负荷、燃

料消耗等实时动态数据。

历史运行数据：包括历史发电量、历史负荷曲线、设备历史故障记录等用于分析和对比的历史数据。

生产管理：

设备维护数据：记录设备的维护周期、维护内容、维护成本以及维护效果等信息。

生产流程数据：涵盖从原料采购到能源生产的整个流程数据，如原材料质量、生产过程监控、能源转换效率等。

经营管理：

销售数据：包括电力销售量、客户信息、销售合同执行情况等。

财务数据：涉及收入、成本、利润等财务数据，以及预算执行情况、资金流动情况等。

安全管理：

安全监测数据：包括设备安全监测数据、环境安全监测数据等，用于评估和预防安全风险。

事故记录与分析：记录安全事故的详细信息，分析事故原因，提出改进措施。

计划管理：

生产计划：涉及发电量预测、设备检修计划、原料采购计划等。

资源计划：包括人力资源计划、财务预算计划等，用于指导企业资源的合理配置。

这样的数据分类体系不仅有助于发电企业更全面地掌握生产运营状况，还能为企业提供有力的决策支持。通过对各类数据的深入分析和挖掘，企业可以更加精准地把握市场动态、优化资源配置、提高生产效率，进而实现可持续发展。

同时，为了确保数据分类体系的有效实施，发电企业还需要建立相应的数据管理制度和技术支持体系。通过制定数据分类标准、建立数据共享机制、加强数据安全保护等措施，可以确保数据的准确性、完整性和安全性，为企业的数据化管理提供有力保障。

（六）数据可视化

数据可视化，是一种将复杂的数据信息以直观、形象的图表和图形形式进行呈现的技术手段，旨在让非专业人士也能迅速洞察数据背后的趋势、关联及模式。通过这一方式，人们能够更加直观地理解数据，快速提炼出关键洞察，从而为决策提供有力的数据支持。数据可视化的实现通常借助一系列工具和技术，包括但不限于各类图表库、可视化软件以及编程语言等。

在实际应用中，常见的数据可视化方案繁多，如折线图、柱状图、饼图、散点图、地图可视化及热力图等，每种图表都针对不同类型的数特有的数据类型。

（七）数据共享

为了实现发电企业大数据中心的数据共享，需要采取一系列的技术手段和措施。例如，建立统一的数据标准和共享平台，确保数据的安全性和隐私性；加强与其他企业、机构或政府部门的合作与沟通，推动数据共享的落地实施；加强数据治理和监管，确保数据的准确性和可靠性等。

数据共享组件需具备实时数据库与关系型数据库的双重兼容性，通过灵活配置数据源，实现共享数据的轻松访问与发布。相关流程审批管控严谨，确保共享数据的安全性与可信度。我们提供消息加密/解密功能，基于安全基础服务，采用加密、数据脱敏等手段，确保信息保密性，并实现操作不可抵赖性。同时，基于日志功能，实现数据交换任务的追踪与事后审计，为数据安全提供有力保障。访问权限控制服务更是对业务流程中的操作、数据和资源访问进行精细化的权限控制，为整体业务流程的安全保驾护航。

（八）数据安全

发电企业数据中心的数据安全，主要包括网络安全、系统及平台安全、数据安全以及数据脱敏和加密功能几个方面：

网络安全：需要借助发电企业信息化网络基础设施进行防护，采用防火墙、安全隔离网闸等措施对进入内部网络的数据包进行扫描过滤，根据网络访问情况判断入侵行为并进行阻止。

系统及平台安全：是指系统或平台本身具有的安全管理功能或防信息泄露机制，包括访问控制安全、身份认证、通讯安全及账号密码的加密和管理机制，对数据存储、缓存的防泄漏机制。

数据安全：需基于数据管理制度进行开展，建立数据安全评估体系、多层次的立体防护体系、完善数据库运维管理制度等方面，具备完善的防范、告警、拦截和审计手段。

数据脱敏和加密：需要对敏感数据进行定义，制定脱敏和加密规则，采用先进的数据脱敏和加密技术对数据进行转换、修改、隐藏，使数据在安全可控的环境下进行交换和共享^[7]。

四、结论

本文探讨了发电企业建设数据中心的规划设计与实施方法。通过数据中心的建设，发电企业能够实现数据的集中存储、统一管理和控制，解决数据分散、管理混乱的问题，提升数据的可靠性和一致性。数据中心还能整合不同系统和平台的数据，形成标准化的数据格式和模型，支持企业内部各部门之间的数据共享和协作，并可与外部合作伙伴进行数据交换和共享。此外，数据中心建设还能提升企业的数据管理能力、优化运营流程、提高经济效益和竞争力。通过数据安全技术，确保数据的安全性和完整性，防止数据泄露和滥用。数据挖掘和分析能力则有助于发现数据中的潜在价值，优化生产流程、提高运营效率、降低运营成本。因此，发电企业应将数据中心建设视为提升数据管理水平的重要手段。

参考文献

- [1]《信息安全等级保护基本要求》GB/T22239-2008.
- [2]《信息技术安全信息系统等级保护安全技术要求》GB/T25070-2010.
- [3]《涉及国家秘密的信息系统分级保护技术要求》BMB17-2006.
- [4]《IT 主流设备安全基线技术规范》（Q/CSG11804-2010）.
- [5]《信息系统应用开发安全技术规范》（Q/CSG11805-2011）.
- [6]《大数据安全白皮书（2018年）》.
- [7]《电力监控系统安全防护规定》发改委2014年14号令.