

面向智能电网的电力大数据存储与分析应用

崔立真¹, 史玉良¹, 刘磊¹, 赵卓峰², 毕艳冰³

1. 山东大学计算机科学与技术学院, 山东 济南 250101;
2. 北方工业大学云计算研究中心, 北京 100041;
3. 国网信息通信产业集团有限公司, 北京 102211

摘要

阐述了智能电网面临的挑战以及大数据关键技术对电力行业的可持续发展和坚强智能电网建立的重要意义。分别从智能电网主数据管理、用电信息统一存储管理、电能质量分析、配网运营能力分析等几个典型大数据系统分析了大数据关键技术 in 智能电网中的应用。

关键词

智能电网; 大数据; 存储; 智能分析

中图分类号: TP391

文献标识码: A

doi: 10.11959/j.issn.2096-0271.2017060

Applications of key technologies of storage and analysis in electric power big data for smart grid

CUI Lizhen¹, SHI Yuliang¹, LIU Lei¹, ZHAO Zhuofeng², BI Yanbing³

1. School of Computer Science and Technology, Shandong University, Ji'nan 250101, China
2. Research Center for Cloud Computing, North China University of Technology, Beijing 100041, China
3. State Grid Information & Telecommunication Group Co., Ltd., Beijing 102211, China

Abstract

Firstly, challenges and improvements of big data in smart grid were elaborated. Then the background of big data and smart grid industry was presented. Finally, big data applications in smart grid were illustrated, and analysis on the systems of big data in smart grid and their key technologies were given.

Key words

smart grid, big data, storage, intelligent analysis

1 引言

电力工业正朝着以物理电网为基础的智能电网发展,我国物理电网是以特高压电网为骨干网架,各电压等级电网协调发展的坚强电网为基础,将现代先进的传感测量技术、通信技术、信息技术、计算机技术与物理电网高度集成而形成的新一代现代化电网。智能电网以充分满足用户对电力的需求和优化资源配置,确保电力供应的安全性、可靠性和经济性,满足环保约束,保证电能质量,适应电力市场化发展等为目的,面向用户提供可靠、经济、清洁、互动的电力供应和增值服务,提高电网的可靠、安全、经济、高效和环境友好等特性。

智能电网一般具有自愈、互动、优化、兼容以及集成等特点^[1],其“智能”主要体现在实时调度和管理、双向信息流、新能源发电的智能接入。其中,实时调度和管理指的是对电网进行实时的管理,在此基础上进行主动的节能与增效,同时能够对安全隐患进行及时发现并且诊断和修复。双向信息流指的是实现发电和用电的实时交互,从而对两者进行综合调度,实现高设备利用率的目标。以上特点以及目标都是基于对电网的观察和控制,为了实现对电网的观察和控制,必须获得电网全景实时数据。这些数据能够反映系统的运行状态,系统能够快速处理和分析这些数据,然后将其转换成可以指导电网运行的决策信息,从而实现电网的智能管理和实时调度。

为了实现对电网的智能管理和实时调度的目标,就必须在智能电网的发电、输电、变电、配电和用电五大环节安装大量的信息采集设备和信息管理系统。例如,在用电网中采用智能电表代替传统的老式机械电表,采集数据的频率从15 min/次

变为1 s/次;在输电网中需要采集各种开关信号量信息以及遥测信息,其刷新频率也能达到1 s/次。在智能电网安装这些信息采集设备和信息管理系统,对电网各个环节进行实时而精确的监控,必将在智能电网中产生大量的数据。这些在电网运行和设备检查、检测过程中产生的数据量呈指数级增长(从TB级逐渐增长为PB级),使得电力行业也进入大数据时代^[2-3]。智能电网中的大数据具有典型的“4V”特征,即数据体量巨大(volume)、数据种类繁多(variety)、价值密度低(value)和处理速度快(velocity)^[4]。利用智能电网中的大数据为电网的发展和运行控制提供科学的决策,不仅是智能电网发展的迫切需求,也是实现智能电网坚强、自愈、兼容、经济、集成、优化的必由之路^[5]。

电网业务数据大致可分为3类:电网运行和设备检测、实时状态数据;电力企业营销数据;电力企业管理数据。其中,电力企业营销数据又包括交易电价、售电量、用电客户等方面的数据^[6]。随着我国智能电网的建设越来越深入,大数据技术成为支撑智能电网安全运行最重要的方法。为此,本文将从面向智能电网应用的电力大数据存储管理、智能分析的4个实例出发,阐述大数据技术在智能电网中的具体应用。

2 基于主数据管理的智能电网全业务统一数据中心

在智能电网信息化和自动化建设过程中,不同部门之间分散式地开发、运行和管理信息系统,系统之间的信息无法互联,造成“信息孤岛”现象,带来硬件冗余、数据多源、格式不一致等问题,使不同电力企业单位及部门之间数据不能及时共享、访问、管理与分析挖掘的矛盾变得突出,难

以制定企业级决策,增加了电力部门的运营成本,甚至造成与用户之间的交流障碍^[7]。数据融合与管理是智能电网大数据的应用基础。在电力行业,最早提出的电力系统公共信息模型(common information model, CIM)^[8]系统性地描述了电力企业尤其是与电力运行有关的所有主要对象,介绍了面向电力生产与电力交易全环节实体及关系的建模方法,并被国际电工委员会(International Electrotechnical Commission, IEC)采纳,成为IEC 61970、IEC 61968、IEC 61925 系列标准的一个重要组成部分^[9]。目前已形成了国家电网公司公共数据模型(SG CIM)^[10]等多个地区性、公司性私有模型,但是对于多领域的完整应用架构与系统调优来说,这些模型与拓展方法并不适用。

国家电网各业务条线信息系统建设和应用的深入发展暴露出跨专业业务协同与信息共享不足,数据多头输入,数据反复抽取、冗余存储、质量不高等一系列问题,对数据的准确性、实时性等要求逐渐提高。为加快构建全球能源互联网,全面建

成电网坚强、资产优良、服务优质、业绩优秀的现代公司,企业需提高全业务协同性和全流程贯通性,深入挖掘数据价值,实现用数据管理企业、用信息驱动业务。数据是信息化的核心,建设全业务统一数据中心^[11]是源端全业务融合、后端大数据分析的必然选择,对建设信息化企业具有重要意义。同时,大数据、云计算等新技术日趋成熟,为全业务统一数据中心的建设提供了技术保障。通过建设全业务统一数据中心,实现对公司全业务数据资源的统一规划、管理和使用,提高企业信息化水平,为公司开展跨专业数据综合利用,实现用数据管理企业、用信息驱动业务的目标,奠定了坚实基础。

2.1 全业务统一数据管理架构

从对基于主数据管理的现有数据中心的进一步发展和完善的角度出发,本文提出一种全业务统一数据中心,主要包括数据处理分中心、数据分析分中心和数据管理分中心3部分,其总体架构如图1所示。

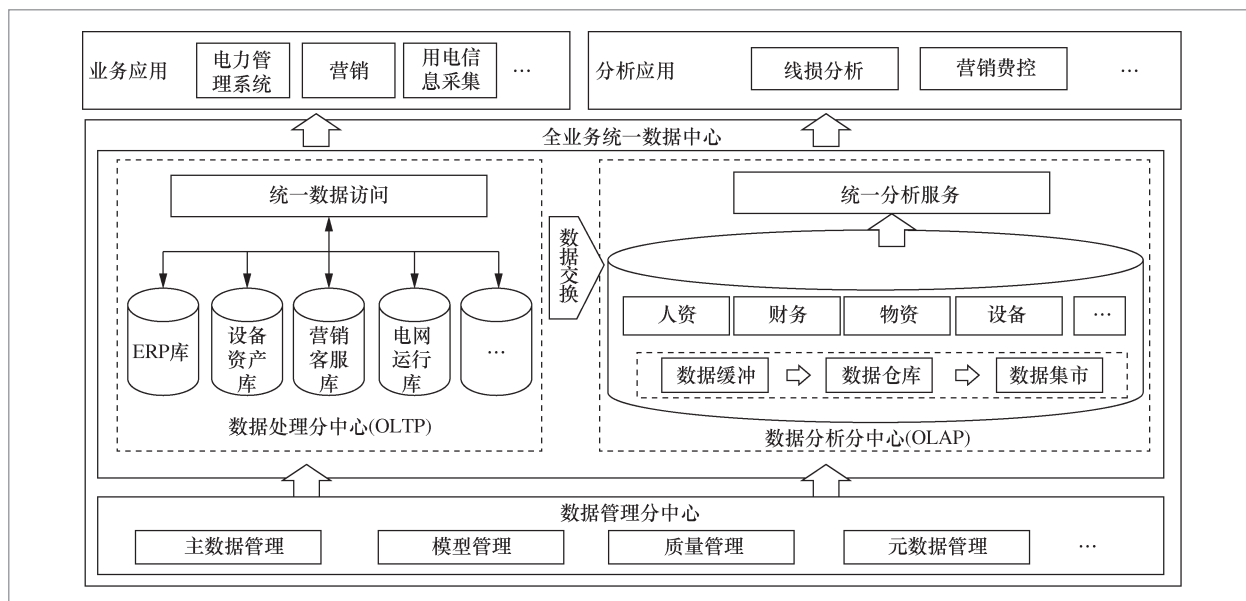


图1 全业务统一数据中心总体架构

2.2 数据处理分中心

数据处理分中心可对公司生产经营管理过程中各类业务数据进行存储、处理和融合,是对原业务系统各个分散数据库的归并、发展与提升,为公司各业务应用提供逻辑统一的数据支撑,使过去数据复制的业务集成方式向共享使用方式转变,实现企业级端到端流程的真正贯通,同时解决系统之间数据集成及数据复制过程中存在的数据安全、效率低下和资源浪费等问题,逐步实现源端数据的干净透明。数据处理分中心包括业务处理数据库和统一数据访问服务两部分,物理上实现两级部署。遵循公司统一数据模型和数据架构要求设计业务处理数据库,按照业务主线对业务处理数据库进行合理划分、部署。为了隔离应用与数据库的直接连接,构建统一数据访问服务,为不同类型数据库构建统一接口提供灵活的访问权限管理、数据路由与调度能力,实现统一的数据管控。数据处理分中心目标架构如图2所示。

2.3 数据分析分中心

数据分析分中心汇集了全业务、全类型、全时间维度的数据,可为公司各类分析决策应用提供完备的数据资源、高效的分析计算能力及统一的运行环境,改变过去需要反复抽取分析型应用数据和数据冗余存储的局面,由“搬数据”转变为“搬计算”,促进企业级数据分析应用的全面开展。数据分析分中心依托企业级大数据平台构建,由统一存储服务、企业数据仓库和统一分析服务3部分组成,物理上两级部署。统一存储服务实现对结构化数据、非结构化数据、采集监测类数据和外部数据的统一存储和管理。企业数据仓库支撑结构化数据的抽取、清洗、存储和多维分析模型的构建,适用于多维分析应用。统一分析服务为数据分析应用提供计算能力和应用构建的支撑,具备高效、便捷访问数据分析分中心数据的能力。数据分析分中心目标架构如图3所示。

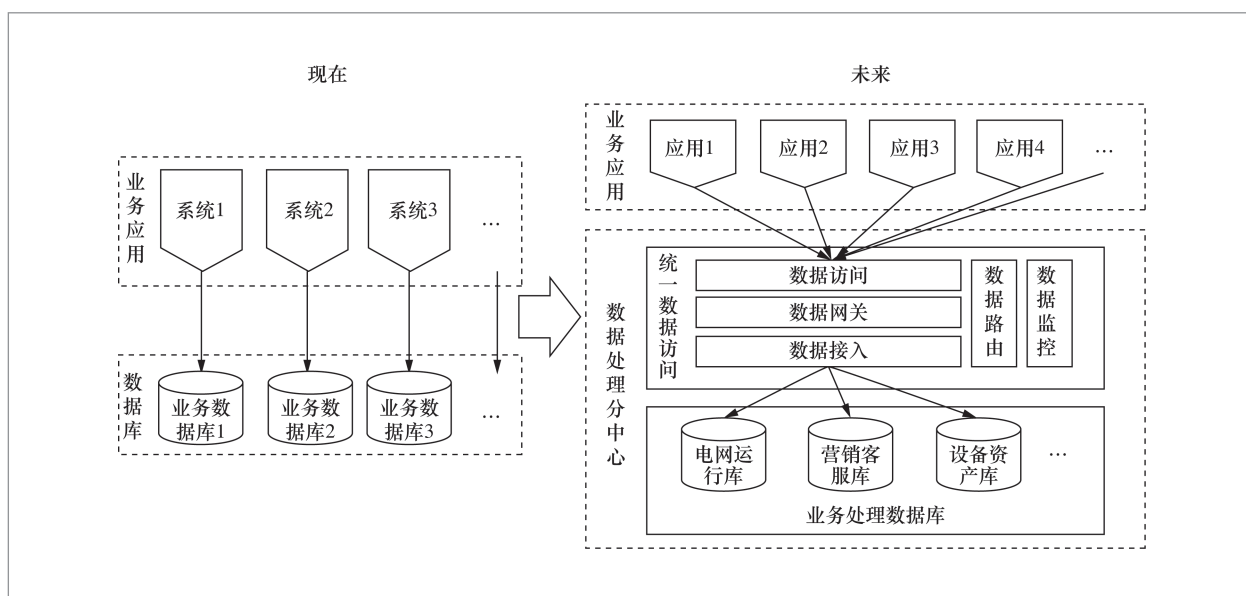


图2 数据处理分中心目标架构

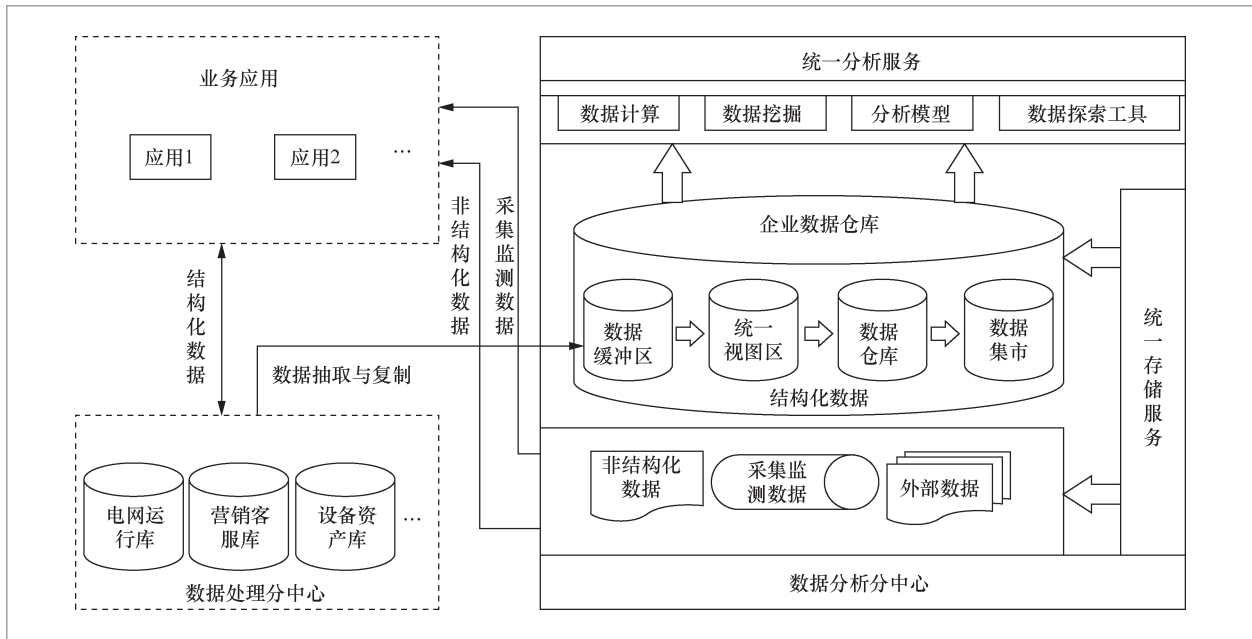


图3 数据分析分中心目标架构

关外部数据的本体^[12, 13]。

2.4 数据管理分中心

数据管理分中心从企业业务全局出发，对企业数据的定义、存储、访问等进行统一规划和管控，为全企业范围内数据的一致性、准确性和可靠性提供保障，为企业内跨专业、跨系统的数据集成与应用提供有力的支持。数据管理分中心物理上一级部署、两级应用，以统一数据模型及主数据管理为建设核心。基于本体论的思想，可以将现实世界中的事物抽象为实体及实体之间的关系，主数据管理可以在此基础上建立信息模型，实现数据含义的表达、共享与重现，利用信息技术对数据进行加工处理，使数据之间建立交互关系，并转化为能回答特定问题的信息，对信息进行概率统计、分类与数据挖掘，可进一步形成有助于决策规划和行动指导的知识。通过对电力系统的公共信息模型(common information model, CIM)建模，得到系统中实体的抽象表示，它表述的对象及其关系构成电力数据及相

3 基于MongoDB的用电信息大数据存储技术

智能电网以电力数据的采集和存储为基础^[4]，电力用户数量和终端数量的快速增长使用电数据成为典型的行业大数据^[14]。用电数据具备大数据的规模大、种类多、要求处理速度快和价值密度低等特性^[15]，为满足大数据管理需求，以非关系型数据库NoSQL为代表的大数据存储技术应运而生，NoSQL技术的优点包括非关系型、分布式数据存储和可横向扩展等，一般分为：基于键值对存储技术，如Redis、Voldemort等；基于数据列分组存储技术，如Cassandra、HBase等；基于文档存储技术，如CouchDB、MongoDB等；基于图存储，如Neo4J、InfoGrid等^[16]。MongoDB^[17-19]作为典型的面向文档的数据库，支持的数据结构非常松散，因此可

以存储复杂的数据类型,保留了SQL一些友好的特性(如索引),另外还支持自动分片、自动故障转移等功能。MongoDB的上述特性满足了用电大数据对存储容量、存储速率等方面的要求,其自动分片机制增强了集群水平扩展能力^[20],可解决用电大数据基本的存储问题;MongoDB的高并发读写性能可实时缓存高速采集到的数据流,解决数据流到达速度与生产库写入速度不匹配的问题;其自动故障转移机制为平台的高可用性提供了有效保障;对于存储模式灵活、时效性低且利用率相对较低的通信源帧、采集状况等数据,可采用模式自由的键值对作为文档存储结构,而对于存储时效性高、利用率高的数据,可采用MongoDB内置的分布式文件存储结构。

3.1 用电信息系统大数据存储架构

用电信息采集系统对用电信息的自动采集、计量异常和电能质量监测、用电分析和管理的提供了技术支持。为适应大规模用电信息数据的存储要求,设计了一种具有高并发、高可靠性和高效存储等特点的

存储架构,加快数据访问速度,本节提出如图4所示的面向用电信息大数据的存储架构,为实现用电信息的自动采集、计量异常等功能,并适应用电数据种类繁多的特点,将数据平台划分为前置通信平台数据库、生产数据库和分析数据库。

3.2 采集数据的存储过程

MongoDB能够存储并支持大数据集的部署和高并发吞吐量的操作,对数据的高效存储便于实时、低时延地访问数据。本节介绍了基于面向用电信息大数据的存储架构对数据的存储过程:前置通信平台负责数据格式转换,当启动数据采集服务和数据处理单元时,从前置通信平台关系数据库加载档案数据,用于数据帧解析和原始数据转存。首先将原始的二进制数据帧转换为JSON(JavaScript object notation)^[21]格式的数据,再进一步转换为业务数据,并存储至应用系统数据库,如图5所示。分布式存储技术具有良好的可扩展性,并对数据充分共享,有助于电力大数据的管理和存储,综合利用分布在各处的资源,能避免由于单个节点失效而使整个系统崩溃的情况出现。

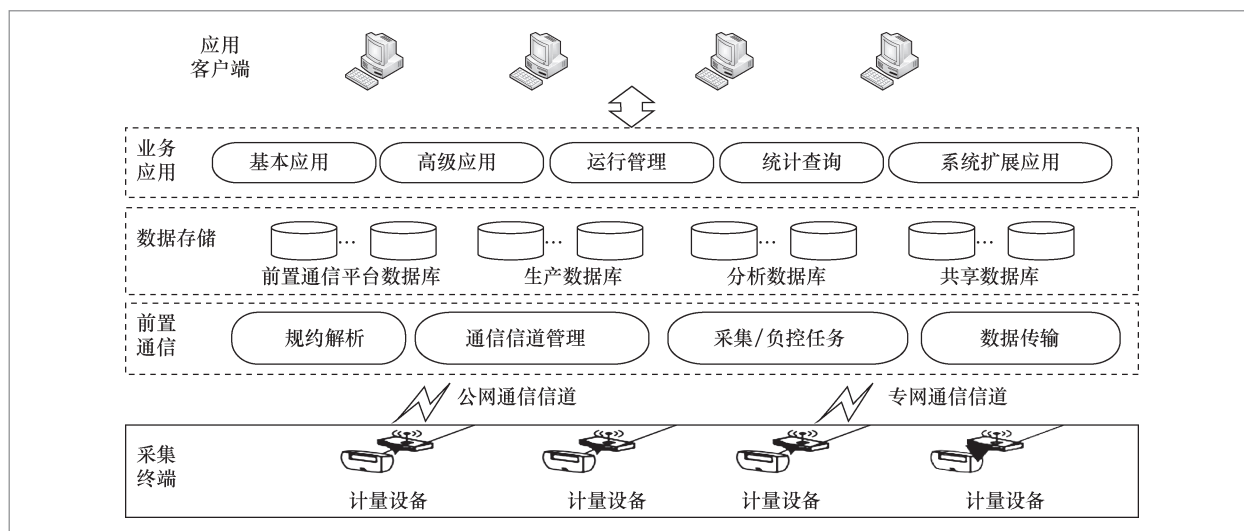


图4 电力用户用电信息采集系统逻辑架构

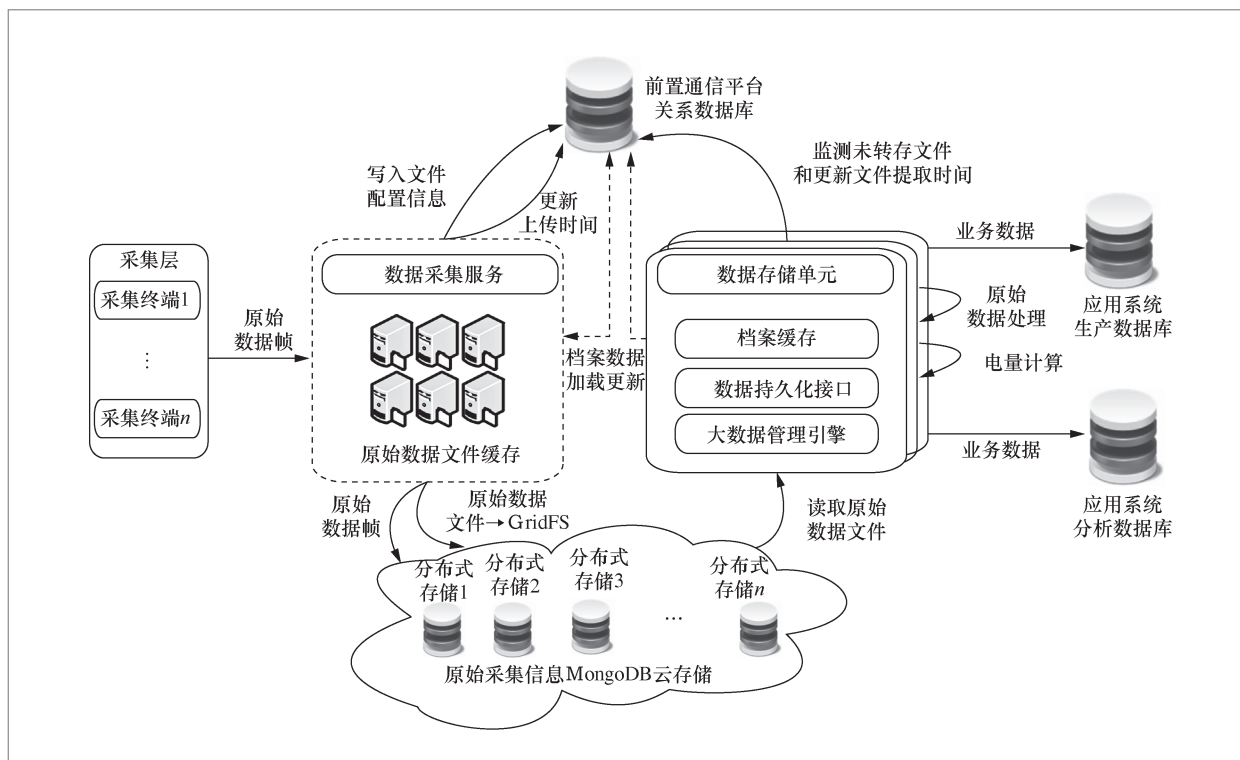


图5 前置通信平台采集数据存储

3.3 基于MongoDB私有云的电力大数据存储关键技术

3.3.1 存储模式

作为一种分布式文档存储数据库，MongoDB可以存储比较复杂的数据类型，本节提出通过对不同类型的采集数据采取不同的存储方式来优化大规模数据的存储和查询效率，采用多集群存储方式提高数据读写速率。MongoDB的自动分片机制支持集群扩展，将9台服务器组成一个MongoDB集群，3台服务器提供路由服务，3台服务器提供配置服务，另外包括由副本集构成的3组数据分片。经大量测试发现，路由服务和配置服务对内存的依赖程度较低，因此可将路由服务和配置服务置于数据分片所在的服务器上，从而提高云资源的利用率。

3.3.2 分片负载均衡

MongoDB的自动分片机制促进了分布式存储的水平扩展，均衡器在一定程度上确保了数据块在每个分片上的均匀分布。MongoDB将数据按用户指定的分片键划分为多个chunk（均衡器进行数据迁移的基本单位），使用均衡器检查各分片内的chunk数，若拥有chunk最多的分片和拥有chunk最少的分片的chunk数之差超过某个阈值（例如8），均衡器则对这些不均匀的分片进行迁移，将前者的chunk移至后者。仅考虑各分片内chunk的数目无法从根本上解决云存储数据访问的动态均衡问题，本文考虑分片所在数据节点负载差异，提出从数据量和负载两方面对负载均衡进行优化，通过路由mongos获取分片所在节点负载，考虑负载因素在迁移限制条件判断、迁移源分片与目标分片选取等环节带来的影响。

3.3.3 读写分离

云存储和副本集技术促进了从副本对读扩展的适用。在从副本上执行查询请求时，会按实际负载情况均衡“读写请求”，增大数据吞吐量。采用响应速度均衡策略执行请求的从副本，向从副本发出探测请求，并将请求分发给最短时间内给出响应的从副本，以较准确地反映节点的运行状态。

4 基于关联分析的电网电能质量监测分析

目前，电能质量干扰源的发展呈现多样化、大容量、高电压等趋势。同时，由于大电网之间高度互联，电能质量扰动的传播和影响范围增大。例如，2011年青藏直流发生了多起直流闭锁事件，均是由700 km外的750 kV主变充电引起的^[22]。电网电能质量往往受多个动态随机干扰源的共同影响。传统的仿真建模方法在电网范围确定、参数获取以及干扰源特征模拟等方面均存在较大困难，不利于分析电能质量扰动事件发生的具体原因。

面向电网扰动事件的提取、定位和原因分析，本节提出了一种基于异常指标关联分析的电能质量扰动事件挖掘方法，通过挖掘频繁共现的异常指标组，形成能表征特定类型扰动事件的特征集合，进而发现电网中可能存在的电能质量扰动事件。

4.1 电能质量监测与分析方法流程

本文提出的电能质量扰动事件特征挖掘方法面向电网监测点采集的电能质量监控指标，在不依赖于特定业务背景和业务知识的情况下，通过识别和关联电能质量

异常指标来定位电能质量扰动事件。由于若干监测点往往受到某个特定干扰源（如电铁、光伏电站、风电站）的影响而产生电能质量扰动，本文提出的挖掘方法关注不同监测点周期出现的具有相同特征的电能扰动事件之间的关联。

图6给出了电能质量监测与分析方法的流程示意。首先，基于监测点实时获取的三相电压电流提取分析电能质量的监测指标；然后，根据电能质量监测指标提取指标异常数据，形成异常指标时序数据；接着，基于异常指标时序数据分析指标间的关联关系，形成一组扰动事件特征的关联指标；最后，通过分析多个检测点的空间分布情况和监测点之间的关联，过滤不相干的扰动特征，提高扰动特征的可用性。

4.2 异常指标提取

本文提出通过监测指标数值的异常来定位电能质量扰动事件。公用电网对谐波的变化范围进行了规范，一般来讲，可将超出规范范围的指标作为异常电能质量指标。然而，在实时获取的电能质量监测数据中，多数情况下并不存在超标数据，导

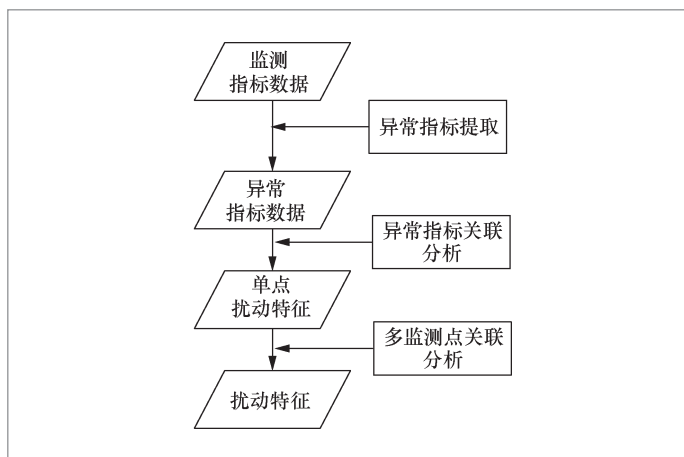


图6 电能质量监测与分析方法流程

致可用于电能质量分析的数据很少。事实上,监测数据中存在大量数据孤立点,这些孤立点虽然没有超标,但是可以用于区别正常数据,因而可将这些数据孤立点作为分析电能质量数据的异常电能质量指标。

电网某监测点C功率因数的时间曲线如图7所示。根据电能质量公用电网谐波GB/T 14549-93规范,当功率因数小于阈值0.9时,指标异常。如图7所示,实际数据中功率因数曲线始终高于0.9,而功率因数小于0.93的指标数据极少,这些数据从一定程度上也可以用于异常指标分析。因此从整体来看,选取0.93而非0.9作为阈值更为合适。

本文通过对指标数据的数值分析来提取异常指标,基本思想为:找出数据中分布稀疏的数值区间,并将这些数值区间的边界作为判定指标异常的阈值。一般来讲,这些阈值为指标数值的上下界。本文提取异常指标的过程以指标时序数据为输

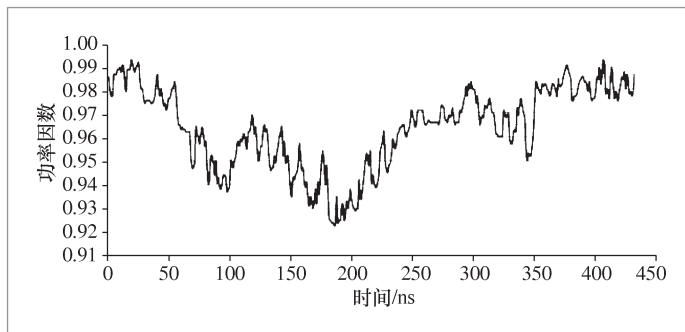


图7 功率因数时间曲线

表1 异常指标限值举例

指标	上限	下限
电压偏差B相	-73.567 5	-73.849 8
电压偏差C相	-73.625 8	-73.958 7
电压总谐波畸变率A相	0.225 334	0.146 073
电压总谐波畸变率B相	0.232 557	0.126 436
电压总谐波畸变率C相	0.222 89	0.145 829

入,选取异常指标时序数据作为输出。首先,计算每个监测指标数值的累计分布概率,以得到累计分布曲线;然后,计算该曲线拐点,以训练出该监测指标所属区间的上下界;最后,将不在数值区间的时序数据定义为异常指标数据,并输出异常指标时序数据。表1为应用此提取异常指标方法得到的一组异常指标。

4.3 异常指标关联分析

某一个监测指标的异常可能由多个干扰源造成,而特定的干扰源也往往会导致多个检测指标出现异常,故通过单一监测指标异常无法直接定位扰动事件及干扰源,本文考察异常指标之间的关联关系,将有关联的若干异常指标视为被同一个干扰源扰动的结果,并将这些异常指标视为电能质量扰动特征。

本文将一组同时出现异常的电能质量监测指标视为一组电能质量频繁项,因此可以将电能质量异常指标的关联分析定义为异常电能质量监测指标的频繁项挖掘问题,并采用FP-Growth算法^[23]挖掘频繁项,图8为异常指标关联分析示意,横坐标为异常指标时序,纵坐标为用于分析的异常指标,指标“电流有效值C相”和指标“电流有效值A相”在 t_2 、 t_3 和 t_4 时刻同时出现异常,因此可将这两个指标视为一组异常电能质量监测指标的频繁项。FP-Growth算法为数据挖掘领域的经典算法,由于篇幅限制,本文不进行详细描述。

5 基于大数据的配网运营能力业务模型

配电网直接面向终端用户,是保障电力能源“落得下、配得出、用得上”的关键

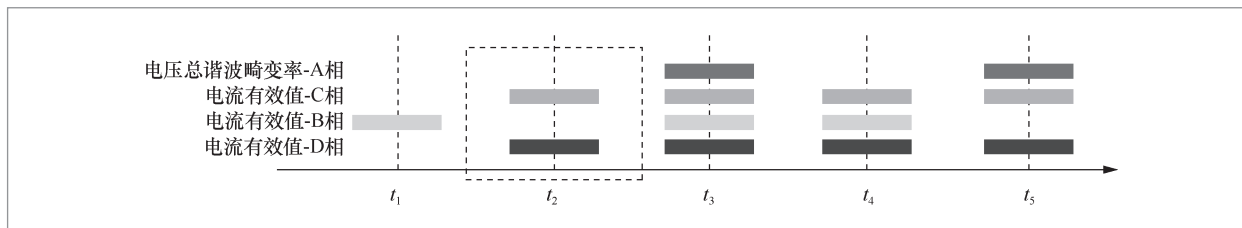


图 8 异常指标关联分析示意

环节。但由于运营调配信息共享不足，管理协同不够，无法及时判断和定位配电故障，停电原因、停电范围分析困难，故障抢修效率低，客户投诉处理不及时等问题，无法满足国家电网公司提出的面向用户提供优质服务的要求。如何利用大数据提高配网运营服务质量和配网运营效能是实现配电网精益化管理的关键，研究如何利用大数据技术提升公司分析决策水平，对推动电力行业有效、可持续发展具有重要的理论与现实意义。

本文通过对配网业务现状进行需求调研，提出了基于大数据的配网运营能力的业务架构，对配网运营服务分析、运营效能分析两大业务域进行设计，完成终端运行情况分析、配变负载情况分析、业扩报装流程分析、配网投入产出分析等10个业务主题的研究与设计，并在此基础上实现基于大数据的配网运行服务情况、客户服务情况、配网项目全过程执行效率、配网

故障抢修效率、配网投入产出分析等方面的大数据应用，挖掘国网信息通信产业集团公司数据的潜在价值，实现其在战略决策、业务应用、管理模式等方面的创新^[10]。配网运营在服务和效能两个业务域的问题及需求见表2。在配网运营服务大数据分析方面，主要可以提供终端运行情况分析、配变负载情况分析、配网电压情况分析、业扩报装流程分析、故障抢修流程分析、投诉举报流程分析、用户停电情况分析。在配网运营效能大数据分析方面，主要提供配网项目全过程执行效率分析、配网故障抢修效率分析、配网投入产出分析。

6 结束语

本文给出了基于主数据的统一数据中心建设、用电信息的大数据存储、电网智

表 2 配网运营能力业务应用需求梳理

指标	应用需求内容
运营质量	终端运行质量、设备间负载均衡度分析、用户停电情况分析、业扩报装流程分析、配变负载情况分析、配网电压质量分析、投诉举报流程分析、数据互联质量分析、系统深化应用成熟度分析、故障抢修流程分析、设备供电能力与供电裕度、信息系统基础数据的完整性、信息系统可用率
运营效率	配网项目全过程执行效率分析、配电系统运行效率协调度分析、配网故障抢修效率分析、配电系统运行效率分析、配网应急物资配送效率分析、配网带电作业效率分析、配网停复电执行效率分析、配电网增资及时性分析
运营效益	配电网资产全寿命周期费用分析、配网资产报废处置分析、配网投入产出分析、配网运营支出分析、配网资金流动合理性分析、配网资产全寿命周期费用分析、配网建设资本支出分析

能监测与分析、配网运营能力大数据分析等方面的几个典型实例,讨论如何利用大数据技术应对智能电网挑战,详细介绍了典型智能电网应用的大数据系统与关键技术以及基于大数据的配网运营能力业务模型。目前智能电网应用发展趋势猛烈,但是电力大数据在电力系统中的应用才刚刚起步,因此结合大数据的技术优势和电力系统的应用需求,发挥电力大数据的价值,增强“信息孤岛”之间的互操作,将为智能电网的建设带来新的发展契机。电力企业应该牢牢抓住这个契机,积极与第三方大数据分析平台厂商合作,从数据政策、人才培养、关键技术研发等层面最大化挖掘电力大数据的市场潜力,充分地挖掘电力大数据具有的资产价值,促进企业未来的可持续发展。

参考文献:

- [1] 李兴源,魏巍,王渝红,等. 坚强智能电网发展技术的研究[J]. 电力系统保护与控制, 2009, 37(17): 1-7.
LI X Y, WEI W, WANG Y H, et al. Study on the development and technology of strong smart grid[J]. Power System Protection and Control, 2009, 37(17): 1-7.
- [2] 王涓,崔健. 智能电网大数据处理技术现状与挑战[J]. 电子技术与软件工程, 2016(1): 188.
WANG W, CUI J. Present status and challenges of big data processing in smart grid[J]. Electronic Technology & Software Engineering, 2016(1): 188.
- [3] WANG P, RAO L, LIU X, et al. Dynamic data center operations with demand-responsive electricity prices in smart grid[J]. IEEE Transactions on Smart Grid, 2012, 3(4): 1743-1754.
- [4] 宋亚奇,周国亮,朱永利. 智能电网大数据处理技术现状与挑战[J]. 电网技术, 2013, 37(4): 927-935.
SONG Y Q, ZHOU G L, ZHU Y L. Present status and challenges of big data processing in smart grid[J]. Power System Technology, 2013, 37(4): 927-935.
- [5] 刘广一,朱文东,陈金祥,等. 智能电网大数据的特点、应用场景与分析平台[J]. 南方电网技术, 2016, 10(5): 102-110.
LIU G Y, ZHU W D, CHEN J X, et al. Characteristics application scenarios and analysis platform of smart grid big data[J]. Southern Power System Technology, 2016, 10(5): 102-110.
- [6] 邓炜瑛. 智能电网大数据处理技术现状与挑战[J]. 中外企业家, 2015(6): 126.
DENG W Y. Present status and challenges of big data processing in smart grid[J]. Chinese and Foreign Entrepreneurs, 2015(6): 126.
- [7] 韩笑,狄方春,刘广一,等. 应用智能电网统一数据模型的大数据应用架构及其实践[J]. 电网技术, 2016, 40(10): 3206-3212.
HAN X, DI F C, LIU G Y, et al. A big data application structure based on smart grid data model and its practice[J]. Power System Technology, 2016, 40(10): 3206-3212.
- [8] USLAR M, SPECHT M, ROHJANS S, et al. The common information model CIM: IEC 61968/61970 and 62325—a practical introduction to the CIM [M]. Berlin: Springer Berlin Heidelberg, 2012.
- [9] SPECHT M, ROHJANS S. ICT and energy supply: IEC 61970/61968 common information model[M]//Standardization in smart grids. Berlin: Springer Berlin Heidelberg, 2013: 99-114.
- [10] 国家电网公司. Q\GDW703-2012 国家电网公司公共信息模型(SG-CIM)[M]. 北京: 中国电力出版社, 2012.
State Grid Corporation of China. Q\GDW703-2012 state grid common information model (SG-CIM)[M]. Beijing: China Electric Power Press, 2012.
- [11] 朱碧钦,吴飞,罗富财. 基于大数据的全业务

- 统一数据中心数据分析域建设研究[J]. 电力信息与通信技术, 2017(2): 91-96.
- ZHU B Q, WU F, LUO F C. Research on the construction of data analysis domain of unified data center based on big data[J]. Power Information and Communication Technology, 2017(2): 91-96.
- [12] 曹晋彰. 面向智能电网的公共信息模型及其若干关键应用研究[D]. 杭州: 浙江大学, 2012.
- CAO J Z. Reaearch on common information model and its key applications for smart grid[D]. Hangzhou: Zhejiang University, 2012.
- [13] 乔金凤. 智能电网语义数据集关键技术研究与应用[D]. 北京: 华北电力大学, 2014.
- QIAO J F. Smart grid semantic data integration research and application[D]. Beijing: North China Electric Power University, 2014.
- [14] 赵腾, 张焰, 张东霞. 智能配电网大数据应用技术与前景分析[J]. 电网技术, 2014, 38(12): 3305-3312.
- ZHAO T, ZHANG Y, ZHANG D X. Application technology of big data in smart distribution grid and its prospect analysis[J]. Power System Technology, 2014, 38(12): 3305-3312.
- [15] 马建光, 姜巍. 大数据的概念、特征及其应用[J]. 国防科技, 2013, 34(2): 10-17.
- MA J G, JIANG W. The concept, characteristics and application of big data[J]. National Defense Science & Technology, 2013, 34(2): 10-17.
- [16] 覃雄派, 王会举, 李芙蓉, 等. 数据管理技术的新格局[J]. 软件学报, 2013(2): 175-197.
- QIN X P, WANG H J, LI F R, et al. New landscape of data management technologies[J]. Journal of Software, 2013(2): 175-197.
- [17] STEVIC M P, MILOSAVLJEVIC B, PERISIC B R. Enhancing the management of unstructured data in e-learning systems using MongoDB[J]. Program Electronic Library & Information Systems, 2015, 49(1).
- [18] BOICEA A, RADULESCU F, AGAPIN L I. MongoDB vs Oracle -- database comparison[C]//Third International Conference on Emerging Intelligent Data and Web Technologies, September 19-21, 2012, Bucharest, Romania. New Jersey: IEEE Press, 2012: 330-335.
- [19] LIU Y, WANG Y, JIN Y. Research on the improvement of MongoDB auto-sharding in cloud environment[C]//International Conference on Computer Science & Education, November 12, 2012, Yogyakarta, Indonesia. New Jersey: IEEE Press, 2012: 851-854.
- [20] 金培权, 郝行军, 岳丽华. 面向新型存储的大数据存储架构与核心算法综述[J]. 计算机工程与科学, 2013, 35(10): 12-24.
- JIN P Q, HAO X J, YUE L H. A survey on storage architectures and core algorithms for big data management on new storages[J]. Computer Engineering and Science, 2013, 35(10): 12-24.
- [21] 仇小花, 秦栓栓, 邱果. 基于WEB开发中的XML与JSON数据传输格式研究[J]. 信息技术与信息化, 2017(4): 123-125.
- QIU X Y, QIN S S, QIU G. Research on XML and JSON data transmission format based on WEB development[J]. Information Technology & Informatization, 2017(4): 123-125.
- [22] 种芝艺, 粟小华, 刘宝宏. 西北电网主变充电引起青藏直流闭锁的原因分析及对策[J]. 电力建设, 2013(3): 88-91.
- ZHONG Z Y, LI X H, LIU B H. Analysis and countermeasures of Qinghai-Tibet DC blocking caused by main transformer charging in northwest power grid[J]. Electric Power Construction, 2013(3): 88-91.
- [23] HAN J W, PEI J, YIN Y W. Mining frequent patterns without candidate generation[J]. Data Mining and Knowledge Discovery, 2004, 8(1): 53-87.

作者简介



崔立真 (1976-), 男, 博士, 山东大学计算机科学与技术学院教授、博士生导师, 主要研究方向为大数据科学与工程、智能数据分析与大图深度学习、服务计算与协同计算等。



史玉良 (1978-), 男, 山东大学计算机科学与技术学院教授、博士生导师, 主要研究方向为云计算、大规模数据管理、隐私保护等。



刘磊 (1981-), 男, 山东大学计算机科学与技术学院副教授、硕士生导师, 主要研究方向为网络性能工程、软件定义网络、网络舆情监测系统。



赵卓峰 (1977-), 男, 北方工业大学云计算研究中心副研究员、副主任, 主要研究方向为云计算、海量感知数据处理、服务计算、智慧城市建设。



毕艳冰 (1980-), 女, 就职于国网信息通信产业集团有限公司, 主要研究方向为软件与数据工程、云计算数据管理、电力系统监控和配电自动化。

收稿日期: 2017-09-15

基金项目: 国家自然科学基金资助项目 (No.71271043); 国家重点研发计划基金资助项目 (No.2016YFB1000602); 创新方法工作专项基金资助项目 (No.2015IM010200)

Foundation Items: The National Natural Science Foundation of China (No.71271043), National Key R&D Program (No.2016YFB1000602), The Innovation Method Fund of China (No.2015IM010200)