

基于监控大数据的设备监控业务评价指标体系的设计与实现

程 道¹, 孙立时¹, 徐春雷², 季 剑³, 吴海伟², 张 明⁴

(1. 国家电网有限公司, 北京 100031; 2. 国网江苏省电力公司, 江苏 南京 210024;

3. 国网淮安供电公司, 江苏 淮安 223001; 4. 国网南京供电公司, 江苏 南京 210000)

摘 要:随着智能变电站的逐步覆盖,以及智能电网监控运行大数据分析系统数据接入范围的扩大,监控大数据系统具备海量数据的基本特征。文章设计的基于监控大数据系统的调控机构设备监控业务评价指标体系,主要包括数据管理、业务指标计算、GIS展示等功能,其中业务指标计算是以全网设备运行数据为基础,业务指标的计算具备数据规模大、计算复杂的特点。为了解决海量数据的统计效率,系统采用去中心化的建设思路,设计采用分级复用、面向服务和内存计算策略的多级指标计算框架,实现各类监控业务指标的高效统计,能够满足系统快速响应的非功能性需求,同时结合面向服务的分布式架构设计,实现不同系统之间的数据贯通,提升设备监控的业务协同能力,系统具备良好的扩展性。

关键词:监控大数据;分级复用;内存计算;面向服务

文章编号:2096-4633(2019)06-0047-05 **中图分类号:**C39 **文献标志码:**B

国网调度中心“十三五”规划中明确提出应用大数据、云计算等技术提升调控技术支撑能力,提高海量、多源、分布式接入、异源异构的数据进行联合分析和挖掘能力,提升大数据质量的技术支撑能力^[1-4]。随着“大运行”的推进,变电站向智能化无人值守方向的不断发展,监控大数据系统需要采集的数据范围逐步扩大,目前基于监控数据的变电站设备运行大数据分析系统从监控系统接入设备模型、监控数据(五遥)、输变电在线监测信息、二次设备在线监测数据(录波数据等)等电网及设备运行数据,从OMS系统接入设备检修、设备故障、设备缺陷、设备台账等基础管理数据,从气象、操作票、雷电监测、视频、GIS系统及其它系统接入相关辅助信息。

面对多源、高维、异构等特征的电力监控大数据,采用中心化数据、分布式计算等技术能够解决海量数据的计算问题^[5,6],但是在系统应用过程中,还存在以下问题:①系统建设成问题:目前江苏省调控中心管辖范围的变电站3 107个,通过数据抽取的中心化数据技术路线、分布式计算改造需要主站(子站)进行改造,中心化数据和分布式计算的运维成本高;②功能建设重复问题:主站端需要的统计指标来源于子站侧,子站系统建设厂家众多,指标计

算、统计数据报送等功能如果各自独立建设,重复建设导致建设成本浪费;③系统建设协调难度大:面对多个业务系统的数据集成工作,需要跨部门、跨公司的协调工作,协调难度大。

针对以上问题,面向监控大数据的设备监控业务评价模块采用去中心化的建设思路,根据系统性能和扩展性的需求,监控大数据系统基于数据服务构建分布式计算框架,采用分级复用、内存计算等策略,优化计算性能,并以面向服务的方式支撑应用服务器的扩展^[7-9],能够满足监控大数据的设备监控业务评价指标的统计要求。

1 系统功能

1.1 数据管理

数据管理主要实现主站与子站之间的数据同步,通过主站端配置定义工具提供的数据源信息、定时任务配置信息和映射表信息实现对子站基础数据、管理数据、运行数据的抽取。依据不同类型数据的特点,采用数据库抽取、文件抽取、服务接口三种不同的抽取方式,数据抽取完成后通过数据封装服务,封装成带数据类型信息的规范化字符串发送到数据传输组件,主站侧的数据抽取组件通过适配器解决子站数据格式不一致的问题。

数据抽取核心组件分为数据感知组件、数据抽取组件、数据转换组件、数据封装组件。数据感知组件以不同的方式获取数据源信息的变化,抽取组件根据感知信息进行数据抽取,利用转换规则形成中间数据格式,并由数据封装组件进行打包后发送给传输组件。中间数据格式可由系统日志进行数据暂存,以便查阅或重发。

1.2 业务评价指标管理

根据系统功能要求,从设备监控业务评价的指标计算需求分析,主要包括基础评价、运行评价和站端监控评价等三类指标,其中基础评价涵盖变电站集中监控覆盖率、监控信息总量、监控运行人员到位率、监控效率等,运行评价涵盖监控信息接入完备率、站日均告警信息量、告警信息频发率、监控缺陷处理率、母线电压合格率等,站端监控评价涵盖监控信息接入完备率、开关操作次数、变电站顺控功能情况、监控缺陷处理及时率等,站端监控评价以运维单位、运维站、监控变电站为对象的监控业务量化分析评价,为年度评价指标。

为了保证数据准确性及一致性,具备对业务指标数据进行锁定的功能,锁定后的数据无法更新。系统具备按照周期范围、统计单位等多维度的统计分析,满足综合统计对比。针对监控缺陷处理率、监控效率等业务评价指标设置阈值,超限情况下具备告警统计等功能。

1.3 GIS 综合展示

系统 GIS 功能以电网资源的空间结构及位置信息和图形可视化为核心,注重电网设备的空间位置与电网的拓扑信息,为各种电力业务应用提供电网空间信息服务^[10]。

设备监控业务评价的 GIS 综合展示功能包括变电站分布展示、子站自动监测、运行告警展示等功能,同时系统在 GIS 上具备变电站与核心业务指标的联动可视化展示,在功能上主要实现变电站的站日均告警信息量、告警信息频发率、监控缺陷处理率和监控缺陷处理及时率等信息的展示。

2 系统架构设计

根据系统的业务需求,从系统架构设计的角度,系统划分为数据层、计算层、应用层和展示层。数据层描述统计相关数据的存储信息;计算层按照指标特征建立指标库,实现指标的计算模型、调度计算处

理指标数据,数据展现模型定义、数据展现接口接合指标库完成用户的展现;应用层是按照数据源管理、业务指标管理等应用功能区分;展示层是对变电站各类评价指标的相关统计数据进行图形化展示的相关功能,系统功能架构如图 1 所示,其中内存计算框架为外部引用模块。

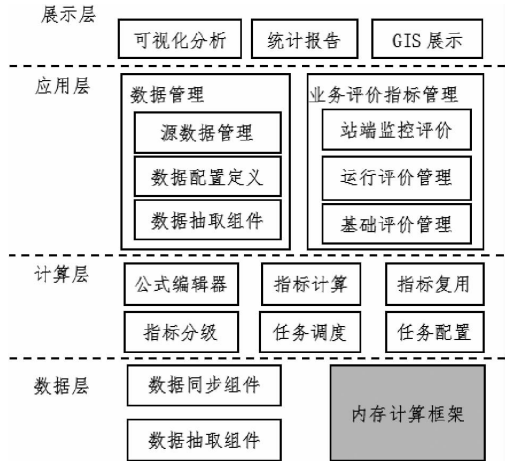


图 1 系统功能架构

Fig. 1 System functional architecture

系统功能架构采用分层式的设计方式,各层之间逻辑独立,采用独立的交换服务接口方式,完成不同的逻辑处理,通过松耦合的结构,从数据层逐步向展现层实现平滑过度,消除不同功能之间的干扰。

3 关键技术

从业务角度来讲,系统数据可以划分为采集数据、汇总数据和指标数据。其中采集数据描述从子站直接获取的源数据;汇总数据描述在采集层的基础上,按照不同周期进行汇总的累加数据,最大汇总周期为月份;指标数据描述不同单位的汇总数据按照设备监控业务评价指标对应的公式描述计算的各类指标数据。

从统计指标的计算需求来看,指标计算的 90% 时间耗费在采集数据的汇总阶段,而大数据量的汇总效率依赖于数据库的 I/O 速度算法开发要求减少数据的重复计算,并且充分利用中间应用服务器资源,通过分级复用的计算策略实现系统性能的优化。

3.1 分级复用策略

下面以母线电压合格率为例,阐述分级复用计算策略的过程,地市供电局的母线电压合格率的计算公式如下所示,其中 n 为电压监测点数量:

$$V_{\text{母线}} = \left(1 - \frac{\sum_{i=1}^n \text{超上限时间} + \sum_{i=1}^n \text{超下限时间}}{\sum_{i=1}^n \text{电压监测总时间}} \right) \times 100\%$$

区域电网公司、省(市、直辖市)电网公司的母线电压合格率:

$$V_{\text{网省(电网)}} = \left(\frac{\sum_{i=1}^n (V_{\text{地市(电网)}} \times n_{\text{地市(电网)}})}{\sum_{i=1}^n n_{\text{地市(电网)}}} \right) \times 100\%$$

从上述计算公式可以看出,由于母线电压合格率的指标具备分级统计的特征,根据业务要求,系统首先计算最小统计单位的电压合格率指标,然后利用下级单位的统计指标递归计算上级单位的电压合格率指标,避免下级单位统计指标的重复计算。同时根据电压合格率统计周期的规律性,在相同统计条件下,如果采集层的数据版本保持稳定,系统自动采用复用策略,进一步减少指标重复计算的可能性,优化系统整体性能。系统计算流程如图2所示,简单描述系统的计算过程。

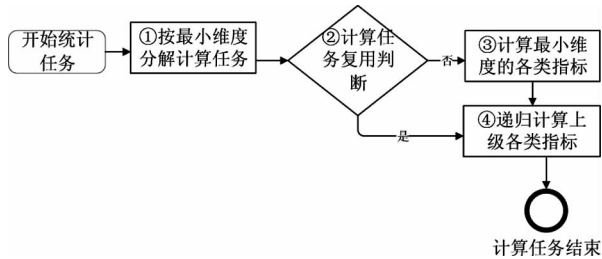


图2 系统计算流程

Fig. 2 The flow of system calculates

上述流程具体描述如下:

(1)按照最小统计单位、最小统计指标等两个维度分解省级电网电压合格率统计的计算任务,计算任务表按照任务单号、计算单位、统计时间、指标名称、数据版本、计算时间进行存储。

(2)根据计算任务表的相关信息,结合采集数据版本判断是否存在复用的可能,在上述条件下,假设地调变电站母线电压监测的运行数据版本与计算任务表一致,系统自动判断电压合格率相关数据不需要计算,直接从指标数据表获取电压合格率统计指标值。如果地市供电局的母线电压监测点的运行数据版本发生变化,系统仅计算对应地市供电局的母线电压合格率;

(3)系统在地市供电局的母线电压合格率基础上,计算上级单位的电压合格率指标。

3.2 面向服务的计算策略

面向服务系统架构(SOA)技术发展迅猛,已发展成为一种高度开放性、层次化、支持并行/分布式信息交互、松耦合的组件模型^[11]。根据指标计算的海量数据特征,系统具备应用服务器的扩展功能,避免数据快速膨胀带来性能下降的风险。本文设计的轻量级计算框架,通过面向服务的方式对外公开指标计算功能,实现按照区域划分递归计算省公司的业务统计指标。

面向服务的计算策略主要是为了解决分级复用策略中的(2)、(3)的潜在性能问题,同时采用流程化的方式实现指标计算,提供了一个灵活的、易集成、易实现的框架,实现思想如图3所示。

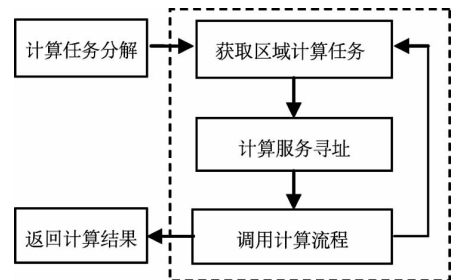


图3 面向服务的计算过程

Fig. 3 Service-oriented computing process

对于简单计算框架来说,计算过程的流程化采用XML配置的方式,做到简单实用,实现计算过程的并行与串行化,优化计算性能。具有面向服务架构的计算策略能够满足供电电压指标计算的性能扩展要求,计算功能按照统一接口定义,通过松耦合的架构设计,提供集中、统一的管理方法,为日后的维护及升级提供了便利的条件。

3.3 内存计算策略

虽然系统通过分表、分区存储,按照区域分表存储,减少周期统计的数据检索量,提高统计效率,但是面对海量数据的计算需求,传统数据库的I/O读取速度仍然是系统性能的瓶颈,本系统设计

经良机计算框架利用内存计算的方式能够有效解决 I/O 瓶颈,系统的简单计算框架与内存计算框架的关系如图 4 所示,本系统采用 HANA 内存计算框架。

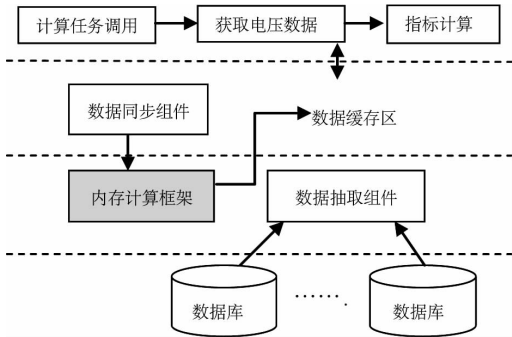


图 4 内存计算策略

Fig. 4 Memory computing strategy

内存计算是指 CPU 直接从内存,而不是硬盘或者数据库上读取数据,并进行计算、分析,是对传统数据处理方式的一种加速,内存计算非常适合处理海量数据,以及需要实时获得结果的数据^[12]。系统采用的轻量级计算框架与内存计算框架采用松耦合的方式,通过适配器模式能够适应不同内存计算框架的配置。当指标计算任务不能复用的情况下,系统需要调用简单计算框架完成指标统计工作,从上述计算步骤可以看出,指标统计是在变电站运行数据的基础上实现。由于变电站运行数据的海量特征,当基础数据首次加载或者更新,需要通过内存计算框架把变电站运行数据同步到内存中。数据同步组件按照周期模式实现关系数据库与内存数据的同步,数据抽取组件可以通过 ETL 复制的方式,把表结构和内容复制到内存数据库里。

4 结束语

随着电力系统对大规模统计的性能要求,日益迫切。面向监控大数据的设备监控业务评价指标管理系统通过分级复用、面向服务的部署、内存计算等多种计算策略,在江苏省电力公司管辖范围下 3 107 座变电站每天的 5 000 万的母线电压运行数据的测试条件下,统计计算耗费时间基本满足用户使用要求。对于母线电压合格率的关联大数据分析,进一步挖掘分析,以及通过应用服务化的方式实现在调控云上的部署和应用,支撑智能电网安全、自愈、绿色、坚强及可靠运行^[13-15],是下一步重点工作。

参考文献:

- [1] 许洪强,姚建国,南贵林,等. 未来电网调度控制系统应用功能的新特征[J]. 电力系统自动化,2018,42(01):1-7.
XU Hongqiang, YAO Jianguo, NAN Guilin, et al. New features of application function for future dispatching and control systems [J]. Automation of Electric Power Systems, 2018, 42(01): 1-7.
- [2] 郭建成,南贵林,许丹,等. 大电网全局监控内涵与关键技术[J]. 电力系统自动化,2018,42(08):1-7.
GUO Jiancheng, NAN Guilin, XU Dan, et al. Connotation and key technology of global monitoring for large power grid [J]. Automation of Electric Power Systems, 2018, 42(08): 1-7.
- [3] 李占英. 智能配电网大数据应用技术与前景分析[J]. 电力大数据,2017,20(11):18-20.
LI ZHANYING. Application technology of big data in smart distribution grid and Its prospect analysis [J]. Power Systems and Big Data, 2017, 20(11): 18-20.
- [4] 中国电机工程学会信息化专委会. 中国电力大数据发展白皮书[M]. 北京:中国电力出版社,2013.
- [5] 周邨飞,徐石明. 智能变电站数据中心初探[J]. 电力系统自动化,2011,35(18):57-61.
ZHOU Yefei, XU Shiming. A preliminary investigation on data center in smart substation [J]. Automation of Electric Power Systems, 2011, 35(18): 57-61.
- [6] 辛耀中,米为民,蒋国栋,等. 基于 CIM/E 的电网调度中心应用模型信息共享方案[J]. 电力系统自动化,2013,37(08):1-5.
XIN Yaozhong, MI Weimin, JIANG Guodong, et al. Scheme of application model information sharing between control centers based on CIM/E [J]. Automation of Electric Power Systems, 2013, 37(08): 1-5.
- [7] 李峰,李虎成,于益军,等. 基于并行计算和数据复用的快速静态安全校核技术[J]. 电力系统自动化,2013,37(14):75-80.
LI Feng, LI Hucheng, YU Yijun, et al. Fast computing technologies for static security checking based on parallel computation and data reuse [J]. Automation of Electric Power Systems, 2013, 37(14): 75-80.
- [8] 吴佳妮,刘露,陈萃等. 高性能计算环境下地理计算服务流程并行处理方法[J]. 计算机科学,2012,39(11):111-115.
WU Jiani, LIU Lu, CHEN Luo, et al. Parallel implementation approach for the geo computation service process in the high performance computing environment [J]. Computer Science 2012, 39(11): 111-115.
- [9] 熊玮,夏文龙,余晓鸿,等. 多核并行计算技术在电力系统短路计算中的应用[J]. 电力系统自动化,2011,35(08):49-52.
XIONG Wei, XIA Wenlong, YU Xiaohong, et al. Multi-core parallel computation technique for power system short-circuit faults analysis [J]. Automation of Electric Power Systems, 2011, 35(08): 49-52.
- [10] 冯晋军,王颖. 基于 Web 服务的 WebGIS 技术研究[J]. 北京:

华北电力大学学报,2006.33(02):101-104.

FENG Jinjun, WANG Ying. Research on WebGIS based on web services[J]. Journal of north china electric power university, 2006.33(02):101-104.

- [11] 王兴志,严正,沈沉,等.基于面向服务架构的调度计划安全校核网格计算[J].电力系统保护与控制.2011,39(24):90-95.

WANG Xingzhi, YAN Zheng, SHEN Chen, et al. Grid computing of operation scheduling with security constraint based on SOA [J]. Power System Protection and Control. 2011, 39(24):90-95.

- [12] 关晓林,黄拓,凌德祥,等.智能配电网大数据应用技术与前景分析[J].电力大数据,2017,21(08):1-7.

GUAN Xiaolin, HUANG Tuo, LING Dexiang, et al. Research and application of big data architecture in electric power Enterprise [J]. Power Systems and Big Data, 2017, 21(08):1-7.

- [13] 宋亚奇,周国亮,朱永利,等.智能电网大数据处理技术现状与挑战[J].电网技术.2013,37(04):927-934.

Song Yaqi, Zhou Guoliang, Zhu Yongli. Present status and challenges of big data processing in smart grid[J]. Power System

Technology, 2013, 37(04):927-934.

- [14] 许洪强.调控云架构及应用展望[J].电网技术,2017,41(10):3104-3111.

XU Hongqiang. Architecture of dispatching and control cloud and Its application prospect [J]. Power System Technology, 2017, 41(10):3104-3111.

- [15] 梁寿愚,胡荣,周华锋,等.基于云计算架构的新一代调度自动化系统[J].南方电网技术,2016,10(06):8-14.

LIANG Shouyu, HU Rong, ZHOU Huaifeng, et al. A new generation of power dispatching automation system based on cloud computing architecture [J]. Southern Power System Technology, 2016, 10(06):8-14.

收稿日期:2019-02-13

作者简介:



程道(1973),男,硕士,教授级高级工程师,从事电网调度与运行管理工作。

(本文责任编辑:龙海丽)

Design and implementation of evaluation index system for equipment monitoring business based on monitoring big data

CHENG Xiao¹, SUN Lishi¹, XU Chunlei², JI Jian³, WU Haiwei², Zhang Ming⁴

(1. State Grid Electric Power Company, Beijing 100031, China; 2. State Grid Jiangsu Electric Power Company, Nanjing 210024 Jiangsu, China; 3. State Grid Huaian Electric Power

Company, Huaian 223001 Jiangsu, China; 4. State Grid Nanjing Electric Power Company, Nanjing 210000 Jiangsu, China)

Abstract: With the gradual coverage of intelligent substation and the expansion of the data access range of the large data analysis system of the smart grid monitoring and operation, the monitoring large data system has the basic characteristics of the massive data. The system mainly includes data management, calculation of business indicators, GIS display. The calculation of business indicators is based on the operation data of equipment in the whole network, and the calculation of business indicators has the characteristics of large-scale data and complex calculation. In order to solve the statistical efficiency of massive data, decentralization is adopted in the construction of the system, the system designs a multi-level index calculation framework of hierarchical reuse, service oriented and memory computing strategy to realize various kinds of monitoring service indicators. High efficiency statistics can meet the non functional requirements of the fast response of the system, and implementing data connection between different systems, enhancing business collaboration of equipment monitoring, and the system has good scalability through the design of service oriented distributed architecture.

Key words: monitoring large data; hierarchical reuse; memory computing; service-oriented architecture