

浅谈电力大数据对信息运行的影响

许海清¹, 黄敏²

(1.江苏省电力公司电力科学研究院,江苏南京 211103;2.无锡供电公司信通分公司,江苏无锡 214061)

摘要:在电力生产管理过程中产生了海量异构、多态的数据,随着智能电网的兴起,这一发展态势更为明显。信息运行是电力企业信息化的技术支撑,文中分析了电力大数据将会对信息运行发生怎样的影响,提出了信息运行如何改变来适应这些影响进而更好地支撑大数据平台的应用;此外,电力信息运行本身蕴含着丰富的运行数据和日志信息,提出了电力信息运行如何借助大数据挖掘出数据中隐含的价值。

关键词:大数据;信息运行;Hadoop

中图分类号:TP391

文献标志码:B

文章编号:1009-0665(2015)02-0062-03

近年来,大数据被人们谈论的越来越多,大数据甚至受到了包括美国政府在内的诸多机构组织的重大关注,大数据在互联网、电信运营、金融投资等相关领域的率先应用也已经累积了相当的示范效应^[1,2],不论是传统的 IBM, Oracle, SAP, Intel 等 IT 厂家,还是以 Google, Amazon 为代表的新兴互联网产品供应商,都推出了各种各样的解决方案和软硬件产品。正是在大数据逐步推行到各行各业的背景下,电力行业也开始谋求将大数据应用于电力行业的生产管理,预期切实借助大数据提高电力企业的运营水平。可以预见,在不久的将来,大数据将会广泛应用于电力行业,全方位的覆盖输电、配电、调度、营销等电网运行的诸多环节^[3],作为电力信息化企业支撑的信息运行,无疑将在大数据这一广泛性应用中发挥基础性的作用。

1 大数据及其特点

1.1 大数据的定义

麦肯锡公司认为大数据是指无法在一定时间内用传统数据库软件工具对其内容进行抓取、管理和处理的数据集合^[4]。此外,Gartner 公司也认为:“大数据是需要新处理模式才能具有更强的决策力、洞察发现力和流程优化能力的海量、高增长率和多样化的信息资产”^[5]。简言之,大数据指的就是所涉及的数据量规模巨大到无法通过目前主流软件工具,在合理时间内达到提取、管理、处理、并整理成为帮助企业经营决策更积极目的资讯。

1.2 大数据的主要特点

IBM 公司认为大数据具有 3 个 V (Volume, Variety, Velocity) 的特征,随着发展,业界又增加了一个 Value 特征,简称 4V 特征。其中,Volume 表示数据量巨大,当前一些大企业的数据量已经接近 EB 数量级;Variety 是数据类型多,不仅包括结构化数据还包

括非结构化的数据;Velocity 表示处理速度快,最后一个 V 即表示数据价值密度低,对于海量数据而言,相应有价值的东西并不会多,因此对数据处理技术提出了更高的要求。

2 电力大数据

电力行业的信息时代正处于关键转折点,随着智能变电站系统、现场移动检修系统、测控一体化系统、地理信息系统、智能表计等智能电网系统的建设,以往数据类型较为单一、增长较为缓慢的情况将发生转变,逐渐步入到由复杂及异构数据源广泛存在和驱动的时代,电力大数据将快速的向着异构、多源、PB 级规模发展^[6]。电力大数据究竟该如何定义,目前行业内普遍认为电力大数据是指通过传感器、智能设备、视频监控设备、音频通信设备、移动终端等各种数据采集渠道收集到的,结构化、半结构化、非结构化的海量业务数据的集合^[7]。

挖掘电力大数据,形成真正适用的电力大数据应用系统,如电网安全大数据系统、全社会用电信息大数据系统等,这些系统将产生大量的应用价值,有利于电力企业高效的生产运行和精细化运营管理,是下一代智能化电力系统在大数据时代下价值形态的跃升。

3 大数据关键技术

基于大数据产业链的定义,电力大数据的关键技术既包括数据分析技术等核心技术,也包括数据管理、数据处理、数据可视化等重要技术。

(1)数据分析技术:包括数据挖掘、机器学习等人工智能技术,具体是指电网安全在线分析、间歇性电源发电预测、设施线路运行状态分析等技术^[8]。由于电力系统安全稳定运行的重要性以及电力发输变配用的瞬时性,相比其他行业,电力大数据对分析结果的精度要求更高。

(2) 数据管理技术:包括关系型和非关系型数据库技术、数据融合和集成技术、数据抽取技术、数据清洗和过滤技术,具体是指电力数据 ETL(Extract, Transfer 和 Load)、电力数据统一公共模型等技术。电力数据质量本身不高,准确性、及时性均有所欠缺,也对数据管理技术提出了更高的要求。

(3) 数据处理技术:包括分布式计算技术、内存计算技术、流处理技术。具体是指电力云、电力数据中心软硬件资源虚拟化等技术。近几年电力数据的海量增长使得电力企业需要通过新型数据处理技术来更有效的利用软硬件资源,在降低 IT 投入、维护成本和物理能耗的同时,为电力大数据的发展提供更为稳定、强大的数据处理能力。

(4) 数据展现技术:包括可视化技术、历史流展示技术、空间信息流展示技术等。具体是指电网状态实时监视、互动屏幕与互动地图、变电站三维展示与虚拟现实等技术。电力数据种类繁多,电力相关指标复杂,加以未来的电力用户双向互动需求,需要大力发展数据展现技术,提高电力数据的直观性和可视性,从而提升电力数据的可利用价值。

4 电力大数据对信息运行的影响

随着大数据在电力领域的纵深推广应用,电力信息运行必须要适应这一变化而提前加强自身建设。

4.1 数据中心的影响

数据中心是企业数据物理承载的核心,数据量的数量级增长无疑对数据中心的建设和运维形成巨大影响,这些影响从目前来看主要是 3 个方面。

(1) 首先在企业建设数据中心时应充分考虑扩展性和灵活性,要预留较大的机房空间,数据大幅地增加肯定需要大量的存储,要增加大量的硬件设备如服务器、存储设备和网络交换机,大量硬件设备的增减对机房基础设施肯定带来新的挑战,如机房散热通风、网络带宽等等因素;(2) 大数据下的信息系统会更多地采用分布式的数据库和文件系统,因此集中化的数据中心、云计算中心和超级计算机中心将成为数据中心的新目标。(3) 就是运行维护人员生产技能培训方面,大数据的存储和计算更多建立在 Hadoop 多层次的技术生态圈中,因此运行维护人员学习和掌握 Hadoop 体系的技能知识,将成为未来数据中心的运行维护基础。

4.2 信息网络的影响

大数据的网络特点决定了对信息网络有着更高的要求。(1) 高带宽,典型的网络重载业务,持续大带宽通信;(2) 集群部署,降低成本,与其他业务混合部署,减少突发对其他业务影响;(3) TCP incast,传输性

能优化。因此信息网络基础设施也要逐步升级,满足大数据下的网络要求。首先,网络交换设备要选择有合理缓存的设备;其二,主网络带宽要从 1 G 向 10 G 发展,数据中心的网络应该是全光网络,建设新一代高性能光网络。

4.3 信息安全的新要求

在大数据给企业带来的机遇同时,大数据也给企业信息安全带来了新的挑战。由于企业的更多信息以海量数据的形式存在,数据更加全样本,而且这些数据管理相比以往更为集中,访问的形式也更方便和多样,便利的同时意味着风险更大,一旦出现信息安全事件,后果更为严重,因此必须要加强信息安全管理 and 应为更为先进的安全技术,保护海量数据的安全,提升企业的整体信息安全治理水平^[9]。

5 挖掘电力信息运行大数据

在电力信息运行领域的各个环节也蕴含着海量数据,既然拥有海量数据,就可以借助大数据的平台和技术对这些海量数据进行分析和处理,从而挖掘其中的价值。

5.1 各类信息运行日志信息

构成电力信息运行主体主要有主机设备、网络设备、PC 桌面计算机、各类信息系统,在长期信息运行过程中,各个环节已经累积了巨量的日志信息数据,而且每天还在不断地产生新的日志信息数据,这些数据按照来源可以分为 2 类数据:

(1) 信息设备和信息系统自身的运行日志信息,可以细分为正常运行、异常运行、预警事件日志信息,通过这些运行日志信息,就可以描画出信息设备和信息系统的运行状况,甚至对未来的运行状况进行预测,由这些运行状况大的信息就可以来组织一些运维工作,比如设备采购、运行检修等等;

(2) 信息设备和信息系统自身的访问日志的信息,通过这些访问日志的信息,就可以描画用户的行为,哪些信息系统用户访问的最多,哪些用户对信息系统的使用更多,通过这些数据的挖掘,就可以对整个电力公司的业务状况等等信息进行梳理和统计,从而对公司信息化辅助决策。这些日志信息按照数据类型来分,应该分为结构化数据和非结构化数据,结构化数据主要是专业的监控系统保存在数据库中的数据,非结构化则是一些文本的数据信息,而且这类数据还占大多数。

显然,这些日志信息无法凭借以前的平台和技术进行处理,首先是数据量巨大,其次还存在大量的非结构化数据,因此必须借助于大数据平台和技术来进行的数据的存储、处理、挖掘和可视化展示。

5.2 信息运行语音客服信息

以国网江苏公司为例,江苏公司目前在全省范围内开通了信息客服热线 5186,任何有关信息方面的事宜都可以通过该热线进行反映,据统计,每天语音电话受理量都在 100 起左右,而且还在不断地增长中。这些语音数据仅是存储就是一个问题,如果不借助于大数据的廉价存储,长期保存这些数据,企业将不堪重负,如果没有高性能的计算分析处理能力,这些语音数据也仅限于作为档案资料进行保存,而失去应有的价值,使用大数据平台强大的计算能力和一些语音分析模型就可以对这些语音数据进行分析处理,比如分析录音数据中的客户投诉抱怨、咨询建议、个性特征、服务偏好等等信息。

5.3 信息机房各类监控数据

信息机房更是蕴藏着海量数据。首先机房视频监控系统中的视频数据,也应该像客服语音数据进行相同处理;其次各类设备如门禁、UPS、空调、电源、温湿度等机房动力设备和环境监控设备也将产生大量的数据,这些数据也完全可以应用大数据平台和技术进行处理。

6 结束语

面对大数据时代的发展趋势,应对电力行业海量数据带来的挑战,尤其是智能电网建设的快速发展,作为技术支撑的企业信息化,应更为主动的融入到这一

趋势中,不断适应新形势下的变化,探索出更多路径,努力创建出大数据下的电力信息运行新模式。

参考文献:

- [1] 维克托·迈尔—舍恩伯格,肯尼思·库克耶. 大数据时代[M]. 杭州:浙江人民出版社,2013:25-30.
- [2] 涂子沛. 数据之巅[M]. 南宁:广西师范大学出版社,2013:46-50.
- [3] 赵刚. 大数据:技术与应用实践指南[M]. 北京:电子工业出版社,2013:120-124.
- [4] 道格拉斯·W·哈伯德. 数据化决策[M]. 北京:中国出版集团,2013:52-56.
- [5] 赵国栋,易欢欢,糜万军. 大数据时代的历史机遇[M]. 北京:清华大学出版社,2013:28-25.
- [6] 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.
- [7] 王春毅. 电力行业的大数据发展解析[J]. 电力信息化,2013,11(2):8-9.
- [8] 衡星辰,周力. 分布式技术在电力大数据高性能处理中的应用[J]. 电力信息化,2013,11(9):40-43.
- [9] 李国杰,程学旗. 大数据研究:未来科技及经济社会发展的重大战略领域[J]. 中国科学院院刊,2012,08(9):647-657.

作者简介:

许海清(1972),男,江苏无锡人,高级工程师,从事电力信息管理工作;

黄敏(1978),男,湖北仙桃人,高级工程师,从事大数据项目管理工作。

A discussion on the influence of power big data on information operation

XU Haiqing¹, HUANG Min²

(1.Jiangsu Electric Power Company Electric Power Research Institute, Nanjing 211103, China ;

2. Wuxi Power supply company Information and Communication Corporation, Wuxi 214061, China)

Abstract: Massive and polymorphic data are produced during the process of electricity production. With the development of smart grid, the situation is becoming more obvious. This paper analyses the influence of massive data on information operation and thus proposes some countermeasures to change information operation to support big data platform's operation. In addition, a method using data mining to dig valuable information from operation and log data is proposed.

Key words: big data; information operation; hadoop

从源头治霾 首选核电

核电污染物排放水平极低。核电的碳排放水平仅相当于火电的 1%左右且不会产生硫化物、氮氧化物及粉尘等高污染物,可以从源头上杜绝雾霾的产生。核电是清洁能源中最高效和稳定的发电方式。与水电相比,核电受地理条件的限制相对较少,而核燃料的供应又相对充足,因此不受扩张规模的限制。与风电和太阳能发电相比,其发电的稳定性和效率最高,且不会因大规模并网造成电网的安全稳定性问题。核电安全性高,发生事故只是小概率事件。截至目前,我国已经投运的核电反应堆近 250 个,发生事故的只是很小一部分,事故比例极低,而随着技术的提升,其安全系数也有逐步提高的趋势。从成本收益的角度来看,发展核电具备极高的性价比我国核电推进的步伐在加快。经过福岛核事故之后的安全检查,近两年我国在核电推进方面的步伐在加快,预计未来几年将是我国核电大发展的事情。

伴随一路一带走出国门,获取额外增长点。目前,我国自主知识产权的“华龙一号”核电技术日臻成熟,而国家一路一带政策又为核电技术走出国门提供了充分的条件。2020 年前,核电设备总投资 3000 亿元以上。

摘自《江苏电力信息网》