

doi: 10.12052/gdutxb.180035

智能工厂工业大数据云平台的设计与实现

孙为军¹, 谢胜利¹, 汪谷银², 刁俊武², 阮航²

(1. 广东工业大学 自动化学院, 广东 广州 510006; 2. 中海油信息科技有限公司 智能制造分公司, 广东 惠州 516000)

摘要: 针对智能工厂工业大数据应用需求, 以炼化企业为例, 通过搭建智能工厂大数据云平台, 实现全网全流程多源异构数据采集, 提供多层次分析方案, 建立数据挖掘方法模型库, 承载智能生产、网络化协同制造等智能服务, 提升企业自身的核心竞争力。

关键词: 工业大数据; 智能工厂; 炼化企业; 网络化协同制造

中图分类号: TP319

文献标志码: A

文章编号: 1007-7162(2018)03-0067-05

Design and Implementation of Industrial Big Data Cloud Platform for Smart Factory

Sun Wei-jun¹, Xie Sheng-li¹, Wang Gu-yin², Diao Jun-wu², Ruan Hang²

(1. School of Automation, Guangdong University of Technology, Guangzhou 510006, China; 2. CNOOC (Intelligent Manufacturing Branch) Information Technology Co. Ltd., Huizhou 516000, China)

Abstract: According to application demand of industrial big data for smart factory, the big data cloud platform was built to achieve multi-source heterogeneous data acquisition in whole networks and whole processes, provide multi-level analysis scheme, establish data mining model library, and support the intelligent bearing production, network collaborative manufacturing and intelligent services. With the platform, petroleum refineries enhanced their core competitiveness.

Key words: industrial big data; smart factory; petroleum refinery; networked collaborative manufacturing

云计算、大数据、物联网等新一代信息技术和工业化的“深度融合”, 正在加速传统制造企业的转型升级。智能工厂作为智能制造的重要实践模式, 核心在于工业大数据的智能化应用^[1]。炼化企业是制造业的重要组成部分, 工业大数据处理技术成为提升炼化企业核心竞争力的关键力量。

1 工业大数据的需求

1.1 工业大数据的概念和特征

《工业大数据白皮书(2017)》将工业大数据定义为^[2]: 在工业领域中, 围绕典型智能制造模式, 从客户需求到销售、订单、计划、研发、设计、工艺、制造、采购、供应、库存、发货和交付、售后服务、运维、报废或回收再制造等整个产品全生命周期各个环节

所产生的各类数据及相关技术和应用的总称。

工业大数据具有一般大数据的数据体量大、数据类型多、价值密度低和处理速度快的特征(4V特征), 还有区别于其他行业大数据的自身特点和挑战^[3-6]: (1) 非结构化数据比例大, 数据来源广泛且分散: 工业大数据有来源于产品生产现场工业控制网络的监控数据, 有来源于企业现有的信息化系统的数据, 也有来源于互联网的供应商、消费者数据; (2) 数据相关性强, 既存在关联关系, 也存在因果关系: 工业大数据的产生和应用都围绕产品全生命周期、优化生产线、改善供应链等, 数据之间存在很强的关联性, 并且分析准确性要求高; (3) 时空序列特性: 工业大数据来源于工业控制网络和各种传感设备, 具有产生频率快、严重依赖采集时间、测点多、信息量

收稿日期: 2018-03-05

基金项目: 国家自然科学基金资助项目(61703112, 61773128); 广东省自然科学基金资助项目(2014A030308009); 广东省科技计划项目(2016B030308001, 2013B091300009, 2014B090907010, 2015B010131014, 2017B010125002); 广州市天河区科技计划项目(201603YH093)

作者简介: 孙为军(1975-), 男, 讲师, 博士, 主要研究方向为大数据、云计算和软件工程。

大和稳定性要求高等特点,需要采用可靠的数据采集、高效的数据存储、快速的海量数据处理工具进行管理;(4)面向具体工业领域,专业性强:工业大数据需要解决面向智能设备和智能产品的全生命周期、故障检测、健康预测等深层次分析和应用的问题。

1.2 炼化企业的工业大数据

炼化企业作为典型流程制造模式,生产过程极其复杂,具有原料物性、生产工艺、生产装备复杂和安全环境要求高等特点^[7]。炼化企业工业大数据应用的难点是打通企业数据采集、集成、管理、分析和应用的产业链条,采用数据驱动业务的数据思维方式进行生产过程管理和创新业务活动。以数据采集为例,炼化企业石油加工生产过程不仅涉及多变量协调,不同工艺路线、工艺参数会产出不同目标产品及收率,而且涉及塔类、罐区、机泵、管线、换热设备等类型众多的复杂生产装备,此外,生产计划多变、技术人员经验不足等都是导致数据采集困难的主要原因。

炼化企业智能工厂建设将以统一工业大数据云平台为中心,提高工厂对不同设备收集的海量信息进行梳理的能力,提高企业信息系统的计算能力和数据消化能力,实现对企业的产品数据、运营数据、销售数据、客户数据的实时而有针对性的分析,对企业的研发、生产、运营、营销和管理方式进行智能化创新,实现信息流、物流、资金流、知识流、服务流的高度集成与融合,使得企业持续创新,并不断开发新产品、新服务。

2 工业大数据云平台

统一的工业大数据云平台能提供云计算和大数据能力,建立数据共享集群、计算服务集群、数据仓库集群,实现PB级数据存储。平台使用OpenStack实现硬件资源虚拟化,使用内存计算和内存数据库技术解决性能问题,以有效地支撑各类资源和数据实现面向炼化企业智能工厂的按需聚合应用,支撑海量数据的分析处理,统一的大数据云平台物理结构如图1所示。

平台基于Lambda架构搭建支持多计算模式的大数据系统,整合离线、实时和流计算,可集成Hadoop、Kafka、Storm、Spark、HBase、MPP数据库等各类大数据组件^[8-10]。平台使用新一代的MPP并行数据库集群作为大数据处理平台的核心,实现对海量数据的存储、集成和分析等操作,支持在线查询、实

时交互分析、异构数据处理、高频历史数据融合探索、可视化数据展现,支撑炼化业务大数据分析应用。使用基于Hadoop的数据处理技术管理非/半结构化大数据,并用事务型数据库支撑智能工厂在线业务系统,为平台提供数据支撑,构建智能工厂大数据分析 and 处理平台。

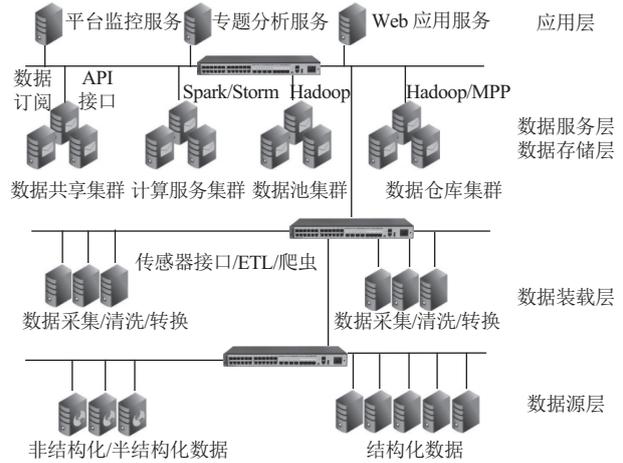


图1 统一的大数据云平台物理结构

Fig.1 Unified physical structure of big data cloud platform

3 炼化企业工业大数据采集

3.1 数据采集范围

炼化企业工业大数据种类繁多,数据采集的范围包括:生产经营相关的业务数据,来自于企业内部高度集成的各类信息化系统^[11],企业信息化系统如图2所示;设备物联数据,来自于炼化企业自动化控制程度相对较高的DCS、PLC等系统的感知数据;外部数据,来自于互联网的行业、市场和竞争对手等数据。

3.2 数据采集技术

3.2.1 现场数据采集

融合各种物联网感知技术,以温度、光敏、视觉等传感器为主要采集工具,结合RFID、条码扫描器、生产监测设备、PDA、智能终端等手段采集多源、异构数据信息,全面感知工厂的生产过程要素(物料物性参数、工艺参数、设备运行等)和安全环境(可燃有毒有害气体、污染物、空气质量等),通过互联网或现场总线等技术进行实时准确传输,存储于实时数据库。

3.2.2 业务数据采集

融合Sqoop、Flume、Fluentd、Logstash、Chukwa、Scribe和并行ETL等多种大数据采集技术和工具,从

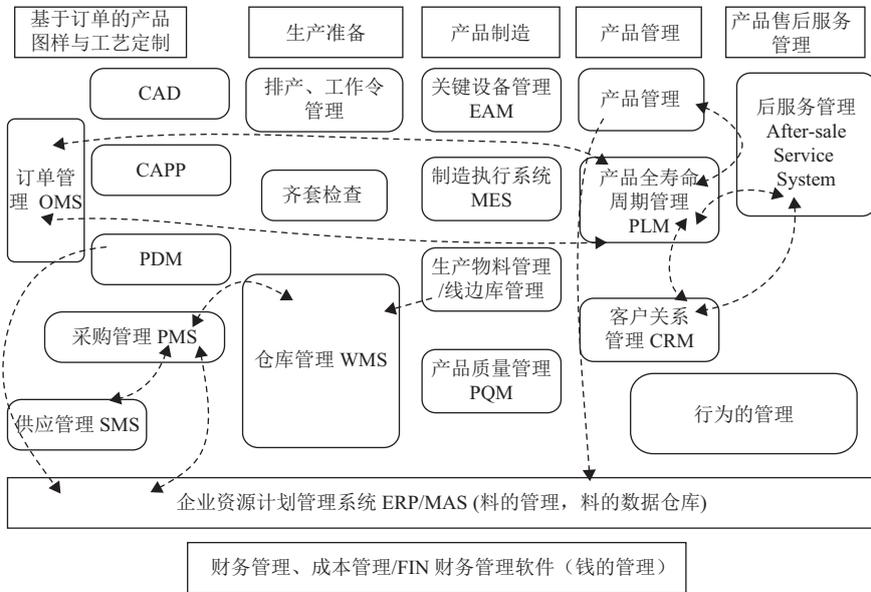


图2 企业信息化系统

Fig.2 Enterprise IT System

MES、ERP、PDM、PLM、SAP和CRM等信息化系统采集生产经营相关业务数据,提取系统间共享数据,实现炼化企业主数据管理。

3.2.3 外部数据采集

融合Nutch、Snoopy等搜索引擎和Web爬虫技术,从电商、WEB、自媒体、微信、微博、博客等采集用户的全网接触数据,全网全流程数据采集示意图如图3所示。

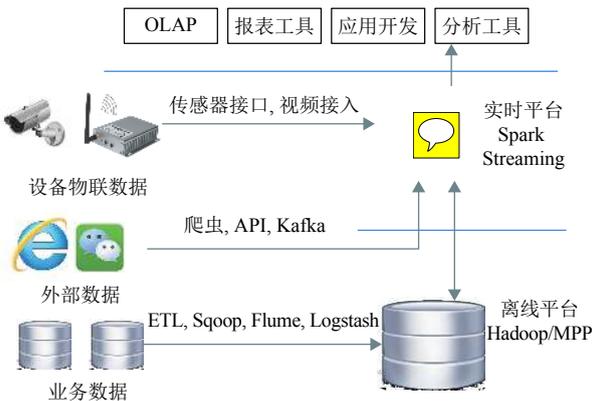


图3 全网全流程数据采集

Fig.3 Whole-networks whole-processes data acquisition

4 大数据分析处理

4.1 多层次的方案

大数据分析提供多层次多场景的数据分析方案:即席查询、统计分析、数据挖掘和知识自动化决策,以满足不同级别的需求。(1) 即席查询(AdHoc):

用户根据特定的需求,灵活地选择或组合查询条件,系统能够快速生成相应的统计报表;(2) 统计分析:运用数据统计分析方法及与分析对象有关的知识,将定量分析与定性分析相结合,适用于上层应用具有明确的业务目标的场景;(3) 数据挖掘:基于人工智能、深度学习和统计学等技术,高度自动化地对数据做出归纳性的推理演绎分析,挖掘出数据的潜在价值,适用于预测预警^[12];(4) 知识自动化决策:集成一系列知识自动化算法模型,包括知识表示、知识获取、知识关联、知识重组和知识推理等,适用于发现和解决生产、管理、调度流程中的关键决策点的决策问题^[13]。借助于可视化工具进行即席查询、多维分析、钻取溯源和推理演绎等数据分析活动。

4.2 基础算法模型库

在关联分析、分类算法、聚类算法、时序分析和决策树等通用分析挖掘算法的基础上,构建炼化流程制造领域专业的的基本算法模型库。利用统计分析、深度学习等理论方法,设计领域专业知识的获取、表示、关联方法,深度挖掘领域相关的物理化学原理、工艺、制造等知识,使得工业大数据的分析结果能够满足高置信度的要求^[14]。算法模型库提供常用的适合R语言和Spark Mlib的分析模型,如Copula (风险分析常用)、ExpSmooth (指数平滑模型,是比较通用的预测模型)、MovingAVG (移动平均模型,产品需求增速预测常用)和Trend (趋势分析)等,此外还有预警预测、滚动预测等服务。利用可视化技术

和工具进行多维分析、钻取溯源和推理演绎,可视化展现分析结果,可根据不同场景可以选择不同的分析框架,支持包括SQL、Restful服务等通用分析接口.研发基本算法模型库,可对数据挖掘分析程序和模型算法进行维护,将模型和算法静态化保存以供随时调用.

5 工业大数据智能化应用

5.1 智能生产

利用炼化企业生产线实时感知的热能、温度、压力、振动和噪声数据,可以实现设备诊断、能耗分析、实时环境分析、质量事故分析(包括违反生产规定、零部件故障)等很多形式的分析.通过对生产线和生产设备配备的传感器进行数据感知和抓取,对生产本身进行实时监控(视频监控、可燃及有毒有害气体报警等).实时生产数据经过大数据平台快速处理后,反馈到生产过程中,实现智能网络管理和自适应调整,使工业控制和管理最优化,最大限度使用有限资源,降低工业和资源的配置成本,使生产过程能够高效地进行^[15-16].

5.1.1 产品故障诊断与预测

设备运行过程中的自然磨损会导致产品品质发生一定的变化.通过实时数据感知,监控系统运行状态和异常情况,诊断故障类型、故障部位及原因,预测产品的不良概率,预测评估故障风险水平,提前更换配件,从而避免发生生产故障,实现产品质量控制和生产风险控制.通过生产过程中故障相关因素的精确控制,真正实现智能化生产.

5.1.2 生产计划与排程

生产计划与排程是生产管理之核心.大数据帮助发现历史预测与实际的偏差概率,在有限产能的基础上,综合考虑人员市场约束、技能约束、物料约束、工序流程约束、员工行为约束等对生产的影响,通过智能优化算法,得到合理有效的计划排产,并监控计划与实际偏差,现场动态调整计划排产.

5.1.3 生产设备参数最优化

基于产品生产中各个环节的大量运行数据的挖掘,如生产中的各个生产设备数据、外部环境数据和产品自身数据(如扫描的图像数据、参数数据等)等,根据各设备的属性数据和生产连续监测数据,实现生产设备参数最优化(见图4).

5.1.4 实时优化能耗

从生产能耗角度看,通过监控生产流程中的耗电、耗水、耗气(天然气或者煤气)、集中供热耗热、

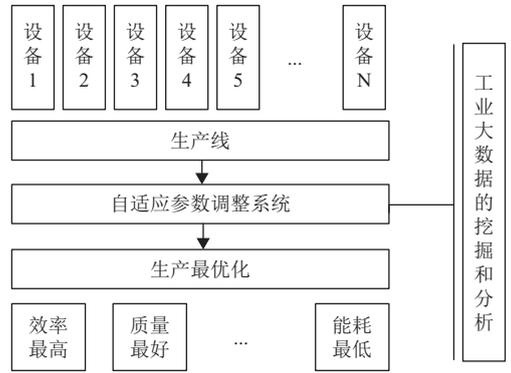


图4 生产设备参数最优化

Fig.4 Optimization of production equipment parameters

集中供冷耗冷的实时数据,能够发现能耗的异常或峰值情况,对所有流程的大数据进行分析,实时优化生产过程中的能源消耗,会整体上大幅降低生产能耗.

5.2 网络化协同制造

利用互联网技术,通过制造过程与信息化系统的深度集成,与产业链各环节紧密协同,构建网络化协同制造服务平台^[17](见图5).

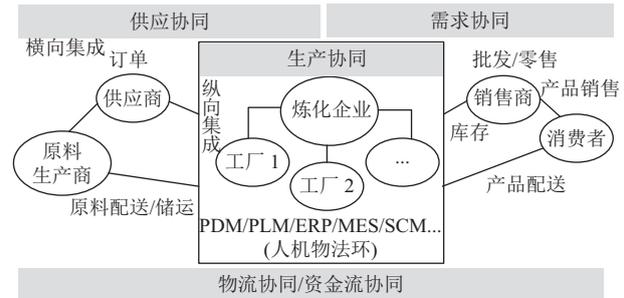


图5 网络化协同制造

Fig.5 Networked collaborative manufacturing

从消费者需求开始,到接受产品订单、寻求合作、采购原材料或零部件、共同进行产品生产、物流、分销和零售,整个环节都通过互联网联接起来并进行实时通信,从而确保炼化企业的最终产品满足大规模客户的差异化、个性化定制需求.

5.2.1 纵向集成: 企业内部的协同制造

纵向集成实现企业内部各种信息化系统之间进行网络协同、实时通信和数据共享. PLM、ERP、MES、SCM的协同集成实现一个产、销、研、用深度结合于一体的协同制造,有效地解决炼化企业产品结构与工艺的设计问题,制定合适的企业资源计划. PLM和ERP的集成主要是通过产品结构的传递完成,而ERP与MES的集成则主要通过生产计划的

传递完成。

5.2.2 横向集成: 企业间的协同制造

炼化企业通常设置有不同的生产基地及多个工厂,横向集成不但包括不同工厂之间的网络协同,也包括产业链上、下游企业之间的网络协作。一方面,消费者与炼化企业通过网络化协同共同进行石油化工产品设计与研发,消费者介入产品的设计生产过程,满足个性化定制需求;另一方面,原材料、资本、设备等生产资源通过网络化协同进行合理配置,组织动态的石油加工过程。通过供应链上各个环节的大数据采集和分析,消费者数据、炼化企业业务数据、供应商数据被融合到供应链体系中,使供应链改进和优化能够持续进行,保证了炼化企业对消费者的敏捷响应,满足差异化市场需求。

6 结束语

随着炼化行业信息化和工业化的深度融合,炼化企业智能工厂将成为炼化行业的发展愿景。炼化企业为应对解决生产和经营所亟需解决的难题,利用统一的大数据云平台支撑智能生产、网络化协同制造等智能化应用,实现产供销存一体化运营,是建设炼化企业智能工厂的有益探索。

参考文献:

- [1] 夏茂森. 流程工业智能工厂建设技术的研究[J]. 信息技术与信息化, 2013(6): 46-52.
XIA M S. Research on intelligent plant construction technology of process industry [J]. Information Technology & Informatization, 2013(6): 46-52.
- [2] 卫凤林, 董建, 张群. 《工业大数据白皮书(2017版)》解读[J]. 信息技术与标准化, 2017(4): 13-17.
WEI F L, DONG J, ZHANG Q. Interpretation of industrial big data white paper (2017) [J]. Information Technology & Standardization, 2017(4): 13-17.
- [3] Industrial Big Data. Know the future-automate processes. Software for data analysis and accurate forecasting[EB/OL]. (2015-10-23)[2018-3-5]. <https://www.differentia.consulting/qlik/docs/Blue-Yonder-White-Paper-Industrial-Big-Data.pdf>.
- [4] 王建民. 工业大数据技术综述[J]. 大数据, 2017, 3(6): 3-14.
WANG J M. Survey on industrial big data [J]. Big Data Research, 2017, 3(6): 3-14.
- [5] 刘强, 秦泗钊. 过程工业大数据建模研究展望[J]. 自动化学报, 2016, 42(2): 161-171.
LIU Q, QIN S Z. Perspectives on Big Data modeling of process industries [J]. Acta Automatica Sinica, 2016, 42(2): 161-171.

- [6] 工业互联网产业联盟工业大数据特设组. 工业大数据技术与应用实践(2017)[M]. 北京: 电子工业出版社, 2017.
- [7] 罗敏明. 流程企业智能制造实践与探讨[J]. 石油化建设, 2016, 38(1): 16-18, 69.
LUO M M. The practice and discussion of intelligent manufacturing in process enterprise [J]. Petroleum and Construction, 2016, 38(1): 16-18, 69.
- [8] MICHAEL H, NATHAN B. Lambda architecture[EB/OL]. (2013-10-23)[2018-3-5]. <http://lambda-architecture.net>.
- [9] 陈尧. 支持多计算模式的大数据系统的研究[D]. 成都: 电子科技大学计算机科学与工程学院, 2015.
- [10] 倪昀炜. 大数据时代动环集中运维管理浅析[J]. 信息技术, 2015(11): 144-146, 151.
NI Y W. Analysis of power and environment centralized monitoring in the big data era [J]. Information Technology, 2015(11): 144-146, 151.
- [11] 覃伟中, 冯玉仲, 陈定江, 等. 面向智能工厂的炼化企业生产运营信息化集成模式研究[J]. 清华大学学报(自然科学版), 2015, 55(4): 373-377, 469.
QIN W Z, FENG Y Z, CHEN D J, *et al.* Study on the information integration of production and operation for smart refinery [J]. Journal of Tsinghua University (Science and Technology), 2015, 55(4): 373-377, 469.
- [12] WU X D, ZHU X Q, WU G Q, *et al.* Data mining with big data [J]. IEEE Transactions on Knowledge and Data Engineering, 2014, 26(1): 97-107.
- [13] 陈晓方, 吴仁超, 桂卫华. 工业生产中的知识自动化决策系统[J]. 中兴通讯技术, 2016, 22(5): 42-46.
CHEN X F, WU R C, GUI W H. Knowledge automatic decision making system in industrial production [J]. ZTE Technology Journal, 2016, 22(5): 42-46.
- [14] 章红波. 工业大数据挖掘分析及应用前景研究[J]. 科技创新与应用, 2016(24): 90.
ZHANG H B. Analysis and application prospect of industrial large data mining [J]. Technology Innovation and Application, 2016(24): 90.
- [15] 王喜文. 大数据驱动制造业迈向智能化[J]. 物联网技术, 2014(12): 7-8.
WANG X W. Big data drives manufacturing towards intelligence [J]. Internet of Things Technologies, 2014(12): 7-8.
- [16] DAYAL U, GUPTA C, VENNELAKANTI R, *et al.* An approach to benchmarking industrial big data applications[C]//WBDB 2014, August 5-6, 2014, Potsdam, Germany. Berlin: Springer International Publishing, 2014: 45-60.
- [17] 宁振波, 吴元良. 从先进制造业的发展看实施工业4.0的前提条件[J]. 航空制造技术, 2014(18): 37-40.
NING Z B, WU Y L. Precondition of industrie 4.0 implementation based on the development of advanced manufacturing industry [J]. Aeronautical Manufacturing Technology, 2014(18): 37-40.