

DOI:10.3969/j.issn.1001-4551.2018.01.015

基于 OpenStack 的输电线路数据云平台架构设计

李洁珊, 赵志勤*

(中国南方电网有限责任公司 超高压输电公司, 广东 广州 510000)

摘要:针对电力企业对于输电线路大数据处理能力要求不断提高的问题,对输电线路大数据特征及需求进行了深入研究,对 OpenStack 云平台技术进行了介绍分析,提出了一种基于 OpenStack 的输电线路数据云平台架构,共分数据采集层、数据中心层、应用层以及接入层 4 个层次。该数据云平台架构充分利用了 OpenStack 技术灵活性、松耦合、开源性的优势以应对输电线路大数据多来源、大存量、高增速的特点。研究表明:该架构能够提供实施简单、可大规模扩展的云计算管理平台,实现了云主机、云硬盘、对象存储及关系型数据库等服务,有效提高了输电线路数据平台的运行效率。

关键词:输电线路数据;OpenStack;数据云平台;大数据

中图分类号:TM71;TP24

文献标志码:A

文章编号:1001-4551(2018)01-0079-04

Transmission line data cloud platform architecture based on OpenStack

LI Jie-shan, ZHAO Zhi-qin

(EHV Transmission Company, China Southern Power Grid Co., Guangzhou 510000, China)

Abstract: Aiming at the increasing demand for big data processing capability of transmission lines in power enterprises, the characteristics and requirements of transmission line big data were deeply studied and the technology of OpenStack cloud platform was introduced, a data cloud platform structure for transmission line based on OpenStack was proposed, consisting of data collection layer, data center layer, application layer and access layer. This structure was designed to cope with the characteristics of transmission line big data such as multiple sources, large stocks and high growth speed relying on flexibility, loose coupling and open source advantages of OpenStack. The results indicate that a simple and scalable cloud management platform can be established based on the proposed structure, offering services including cloud hosts, cloud hard disks, object storage and relational databases to solve problem of insufficient processing capability of power enterprise data centers and effectively improve the operation efficiency of transmission line data platform.

Key words: transmission line data; OpenStack; data cloud platform; big data

0 引 言

随着现代电力行业的不断发展,输电线路数据呈现数量大、维度高、来源多、联系紧等特点。与之相比,电力大数据具有容量大、类型多、速度快的特点^[1-3]。随着各种智能终端接入电网,输电线路数据数量、数据存储类型、信息传输方法不断增多^[4],现有数据处理方式不利于智能电网的发展。

目前,国内外已有大量相关的研究。文献[5]在数据量爆炸性增长的背景下,从大能源与大数据融合

的角度描述大数据对电力企业数据处理方式的影响;文献[6]提出了适合电力大数据应用的数据技术框架和关键技术体系;文献[7]针对广东电网发展的阶段性特点,探讨广东电网大数据的发展思路。许多研究基于单一或多个大数据计算技术提出相应的电力系统或其他行业的大数据平台解决方案。文献[8]利用分布式处理架构与电网实际分布式架构的相似性,提出了一种综合数据平台建设模式和基于多代理技术的系统框架;文献[9]基于 Map-Reduce 模型设计了可以并行计算正域中元素个数的属性约简算法;文献[10]根

收稿日期:2017-03-21

作者简介:李洁珊(1985-),女,广东潮州人,工程师,主要从事电力企业管理工作。E-mail: 554207565@qq.com

通信联系人:赵志勤,男,工程师。E-mail: zy_alex_sh@163.com

据 OpenFlow 的技术特点,提出未来国家电网公司大数据平台应用的方法,具有较好的可扩展性和加速比,可以处理电力大数据集;文献[11]针对海量设备状态数据和动态信息的在线收集问题,基于 Hadoop 技术设计了一套数据储存处理系统,具有高效查询的优势。以上研究大多集中于总体架构和电力系统总的信息平台设计,覆盖面较大,缺少针对发输配用特定平台的研究。

本文提出以大数据技术支撑的统一云平台架构设计方案。

1 输电线路大数据特征及需求分析

1.1 输电线路数据分析

输电线路数据根据文件格式可分为文档、纯文本、线路实时图像、XML、HTML、报表、视频文件;按照数据存储方式可以分为结构化数据、非结构化数据和半结构化数据;按照数据单位大小分为 Byte、Block、Chunk 等级别;按照数据内容可分为台账、日志、记录、测量等数据。

具体线路数据来源及数据分类如表 1 所示。

表 1 输电线路数据源表单

系统平台应用	主要数据
生产管理系统	台账、日志、工况、操作记录数据
输电精益作业管理系统	台账、记录、表单数据
变电现场作业数据采集系统	台账、记录数据
运维自动化平台	量测、监测、记录数据
技术监督数据管理平台	记录数据

1.2 输电线路云平台需求分析

输电线路数据云平台是表 1 中各个系统平台应用的基础,在上述平台系统的数据添加、查询、更新和管理中起到了关键性的作用。

输电线路数据云平台中,输电线路半结构化、非结构化数据主要由运维自动化平台产生,包括现场图像、照片数据,微气象数据,调度自动化系统(Power dispatching Automation System, PAS)测量数据和波记录数据。不同测量数据的规模总量和增量大体可以分为 3 类:(1)日增量大,但产生频率较低,雷电定位系统数据;(2)日增量较小,但产生频率较高,如保护装置动作记录及录波记录;(3)日增量较大且产生频率较高,如 SCADA 量测数据和 PAS 量测数据。因此需要对现有的数据存储进行优化,以提高数据的利用效率。

各类数据存量及增量速度表如表 2 所示。

表 2 各类数据存量及增量速度表

数据名称	总数据量	数据日增量	产生频率
SCADA 量测数据	70 G	500 M	秒级
PAS 量测数据	70 G	500 M	秒级
在线监测装置量测数据	30 G	100 M	10 分钟
保护装置动作记录及录波记录	2 G	20 M	秒级
雷电定位系统数据	100 G	1 G	分钟级

从表 1 和表 2 可以看出数据源分散,无法将现有的数据进行多部门、跨专业、以设备为中心的有效融合,而应用大数据技术的云数据平台可以有效地解决数据整合优化的问题。

2 OpenStack 云平台技术分析

2.1 OpenStack 技术介绍

OpenStack 是云计算服务商 Rackspace 和美国航空航天局一起开发的开源的基础设施及服务云计算平台,它能够帮助服务运营商或者大型企业来管理物理资源设备,搭建和运行云计算平台^[12-14]。该技术具有灵活、松耦合以及开源特点 3 大优势。

2.2 OpenStack 技术在数据云平台服务分析

OpenStack 主要由计算服务,存储服务、镜像服务、认证服务和 UI 服务组成,其具体架构如图 1 所示。

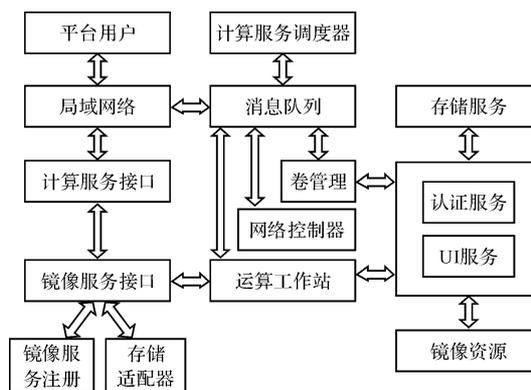


图 1 OpenStack 概念架构

(1)计算服务。主要提供虚拟主机服务,用于为单个用户或使用群组管理虚拟机实例的整个生命周期,根据用户需求来提供虚拟服务;

(2)存储服务。提供一种分布式、持续虚拟对象存储,通过在软件层引入一致性散列技术和数据冗余性,提供了很好的非结构化数据存储方案;

(3)镜像服务。一套虚拟机镜像查找及检索系统,支持多种虚拟机镜像格式,有创建上传镜像、删除镜像、编辑镜像基本信息的功能;

(4)认证服务。对于用户管理可以实现用户与用户行为跟踪,同时提供一个服务目录,包括服务项与相

关接口的位置;

(5)UI 服务。使用 python 语言开发的 Web 端管理门户,封装了大部分主要的 API,用户可通过图形化的界面对 OpenStack 中的资源进行操控,简化了 OpenStack 的交互方式。

3 基于 OpenStack 的输电线路数据云平台架构设计

3.1 输变电线路云平台设计思路

输电线路数据云平台核心思想如图 2 所示。

3.2 数据云平台整体架构

输电线路数据云平台以数据为中心,将实际数据作为应用基础,使用 OpenStack 搭建云平台,统一管理和调度硬件资源,为上层的大数据处理提供资源透明

的弹性服务。平台整体框架由数据采集层、数据中心层、应用层以及接入层共 4 个层次组成,如图 3 所示。

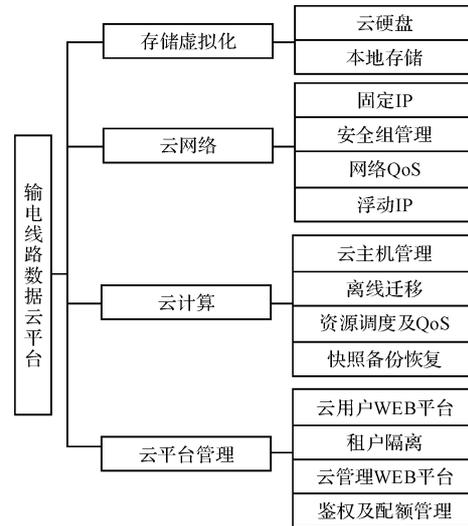


图 2 输电线路数据云平台核心思想

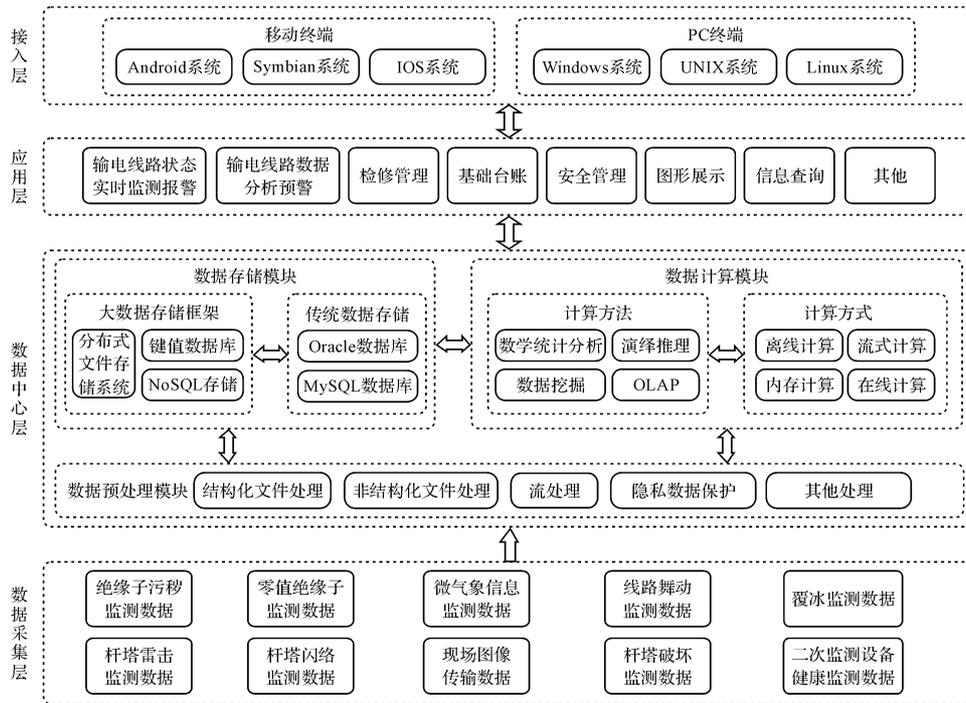


图 3 总体架构图

数据采集层是输电线路数据私有云平台的最底层,作为整个平台的智能感知层,实现对结构化、半结构化、非结构化的海量数据的智能化识别、定位、跟踪、接入、信号转换、监控、初步处理和管理,再通过网络将各种方式采集的数据或人为加工的数据统一传输和存储到上层。

数据中心层是私有数据云平台的核心,数据预处理模块对底层数据进行数据清洗,并将多个数据源中的相似数据结合起来统一存储。数据存储模块分为大数据存储和传统存储两个框架,分别存储非结构化和结构化数据。数据计算模块以 Hadoop/MapReduce 并

行处理引擎为基础,支持调度和分配计算资源,结合数据挖掘技术,利用多种计算方式,计算推演不同来源数据之间的关联关系,提供给上层前端应用。

应用层包括输电线路状态实时监测报警、基础台账、图形展示、安全管理等高级应用,采用基于面向服务的体系结构(service oriented architecture, SOA)的轻量级服务总线,实现不同粒度服务的灵活配置。

接入层即用户终端,平台数据及分析结果等可以通过移动终端或者 PC 终端被用户实时访问,并且支持现有主流操作系统,提高平台的适应性与兼容性。

3.3 输电线路数据云平台基础服务体系分析

输电线路数据云平台包括云主机服务、云硬盘服务、对象存储服务和管理服务。本研究以安全生产的海量数据进行梳理分析和图 3 所示架构为基础,提供可以灵活部署的分层基础服务体系,其具体组件服务架构如图 4 所示。

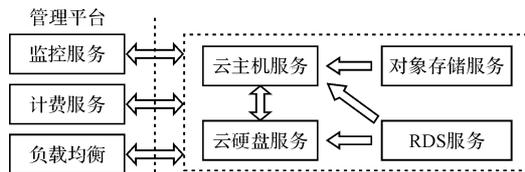


图 4 服务架构图

云平台基础服务主要由 4 种服务组成:

(1) 云主机服务。计算服务的主要提供者,主要部署在可以提供 CPU 和内存服务的计算节点上,由于计算节点具有伸缩性,可以按照实际需求控制物理节点数;

(2) 云硬盘服务。提供虚拟磁盘服务,在默认的机架式服务器集群中,与云主机服务都部署在计算节点上,因此也可以按照应用需求实现弹性伸缩,并且节点之间互不干扰;

(3) 对象存储服务。提供安全可靠的存储服务,每个对象存储设备具有一定的智能,基于 OpenStack 的对象存储服务即 Swift 子项目,具有很强的数据持久性和无单点故障的特点,适用于静态数据的长期存储;

(4) 关系型数据库服务 (relational database service, RDS)。现有数据中大量数据还是以二维形式存储,并且 RDS 具有保持数据一致性、数据新的开销小等优点,因此与非关系数据库互为补充。

除了主要的服务组件,还有相应的管理服务,主要由监控服务、计费服务和负载均衡这 3 个服务组成。管理节点部署管理员平台和产品用户平台,同时也部署了云监控以及各个云服务的 API 组件。这是云平台的唯一入口。在用户请求量较大的情况,可考虑万兆网卡。

4 结束语

本研究根据输电网实际数据特点和 OpenStack 技术的优点,提出基于 OpenStack 的输电线路数据云平台架构和服务架构。

(1) 现有输电线路数据同时存在结构化数据和半

结构化数据,且数据量增长较快,需要根据数据来源不同和数据特性对其进行区分;

(2) OpenStack 服务架构本身具有灵活、松耦合、开源等特点,可以有效提高输电线路数据平台服务框架搭建效率;

(3) 基于 OpenStack 的输电线路大数据平台可以有效地提高数据关联度、处理结构复杂的输电数据,由于服务松耦合性,可以缩短服务响应时间,提高平台运行效率。

参考文献 (References):

- [1] 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.
- [2] 彭小圣,邓迪元,程时杰,等. 面向智能电网应用的电力大数据关键技术[J]. 中国电机工程学报,2015,35(3):503-511.
- [3] 赵云山,刘焕焕. 大数据技术在电力行业的应用研究[J]. 电信科学,2014,50(1):57-62.
- [4] 朱朝阳,王继业,邓春宇. 电力大数据平台研究与设计[J]. 电力信息与通信技术,2015,13(6):1-7.
- [5] 薛禹胜,赖业宁. 大能源思维与大数据思维的融合(一) 大数据与电力大数据[J]. 电力系统自动化,2016,40(1):1-8.
- [6] 李栋华,耿世奇,郑建. 能源互联网形势下的电力大数据发展趋势[J]. 现代电力,2015,32(5):10-14.
- [7] 蔡徽. 广东电网电力大数据现状及主要发展思路[J]. 广东电力,2014,27(12):11-14.
- [8] 畅广辉,镐俊杰,刘涤尘,等. 基于多代理技术的电力控制中心综合数据平台设计[J]. 电力系统自动化,2008,32(1):85-89.
- [9] 曲朝阳,陈帅,杨帆,等. 基于云计算技术的电力大数据预处理属性约简方法[J]. 电力系统自动化,2014,38(8):67-71.
- [10] 尹晓华,胡楠,刘为,等. OpenFlow 在电力大数据平台中的应用[J]. 电力建设,2015,36(3):43-48.
- [11] 李伟,张爽,康建东,等. 基于 hadoop 的电网大数据处理探究[J]. 电子测试,2014(1):74-77.
- [12] 李知杰,赵健飞. OpenStack 开源云计算平台[J]. 软件导刊,2012,11(12):10-12.
- [13] 李小宁,李磊,金连文,等. 基于 OpenStack 构建私有云计算平台[J]. 电信科学,2012,28(9):1-8.
- [14] CHEN S C, HWANG R H. A scalable integrated SDN and OpenStack management system[C]. International Conference on Computer and Information Technology, Nadi: IEEE,2016.

[编辑:周昱晨]

本文引用格式:

李洁珊,赵志勤. 基于 OpenStack 的输电线路数据云平台架构设计[J]. 机电工程,2017,35(1):79-82.

LI Jie-shan, ZHAO Zhi-qin. Transmission line data cloud platform architecture based on OpenStack[J]. Journal of Mechanical & Electrical Engineering, 2017, 35(1):79-82.

《机电工程》杂志;http://www.meem.com.cn