



基于云平台的数据采集与统计汇总系统应用研究

石桂芳, 刘陈荣

(北京华电天仁电力控制技术有限公司 北京 100039)

摘 要: 基于云平台的数据采集与统计汇总系统应用研究以某科技环保集团的实际需求为背景, 主要解决数据中心“数据”方面的问题, 如数据收集、数据加工、数据专题分析、数据展现等, 数据覆盖了综合效益、营销管理、生产管理、资产管理、科技管理、项目管理、社会效益等主要业务范围。项目最终实现了数据处理流程、业务处理流程和经营监管指标体系的功能。

关键词: 云平台; 数据中心; 数据采集汇总

doi: 10.3969/j.issn.1000-0801.2013.12.017

Application Research for Data Acquisition and Summary Statistics System Based on Cloud Platform

Shi Guifang, Liu Chenrong

(Beijing Huadian Tianren Electric Power Control Technology Co., Ltd., Beijing 100039, China)

Abstract: On the background of a practical needs of science and technology of environmental protection group, the system of data collection and statistical summaries based on cloud application research were proposed to solve the problem that the main settlement of data center, “data”, such as special topics in data collection, data processing, data analysis, data show, and so on. The major lines of data business covered comprehensive benefit, marketing management, production management, asset management, technology management, project management, social benefits and so on. Finally, project implement process for data processing, business process, and operating function of index system of supervision were realized.

Key words: cloud platform, data center, data collection and summary

1 概述

某科技环保集团(以下简称集团公司)的业务涉及很多不同的业务范畴,二级单位之间的业务没有重复,垂直业务模式不可横向复制。从数据的角度看,各板块业务的数据之间没有任何关联,给集团公司统一数据口径、统一数据报送带来一定难度。因此,建立一套数据在线收集系统,逐渐形成集团公司的数据中心,利用数据中心建设的多维度、多主题智能分析反哺二级单位,实现数据管理的一致性就显得尤为重要。

集团公司自成立以来,十分重视信息化建设,目前已经初步建立了面向整个集团的信息化网络环境和集成环境,实现了财务与审计、企业综合办公管理(OA)的信息化。之前集团公司对下属企业生产经营管理信息的获取与统计汇总,是由下属企业填写 Excel 报表,后经 E-mail 方式报送给集团公司本部对口部门。集团公司本部业务部门通过 Excel 等软件进行整理、核对、汇总,并报送上级部门,相关数据也以 Excel 等文件形式保存在个人电脑上。

这种方式存在如下问题。



- 数据一致性难以保证。各个业务部门数据交叉的现象比较普遍,不同部门填写的相同指标项的结果是不同的,这在后期的数据统计分析中会产生很大的影响。
- 历史数据的查询速度和准确性难以保证。工作中经常出现需要迅速查询某个历史数据,或者快速汇总某个不常用指标的情况。以前是通过电话询问二级单位或其他业务部门,并在历史数据文件中查找,手工进行数据的汇总整理。这种方式的速度和准确性都难以满足实际的工作要求。
- 工作量相对较大。集团公司业务部门之前需要在汇总下属企业的报表后,核对报表数据,汇总相关指标项的工作量较大;下属企业在填报过程中,存在相同的指标项在多张表中重复出现、重复填写的问题,从而增加工作量。
- 基于集团公司之前经营监管信息的采集和管理现状,集团公司迫切需要构建一套集团本部与分/子公司公用的信息采集、整理与统计汇总一体化的平台。平台一方面有效地支持集团本部业务部门对信息的采集、整合和统计汇总,并实现结果的便利导出;另一方面又可实现下属企业的信息报送并与上级企业交互。

鉴于以上现状,开发了基于云平台的数据采集与统计汇总系统。本系统基于云平台上的数据中心技术研究及应用,主要解决数据中心“数据”方面的问题,如数据收集、数据加工、数据专题分析、数据展现等^[1]。

本系统建立的数据中心基于云平台架构,选择高效能服务器,能避免服务器蔓延增长、设备利用率低、运维成本高(管理、维护)、应用可靠性低(安全性、可靠性)、性能功耗比低、管理麻烦的缺点;实现高可用按需分配,可以满足集团公司高效、灵活、快速的业务增长需求。

- 通过对数据中心的整体设计,提高系统的可靠性及可用性,保证数据的统一性,提高工作效率,节省人力,节约资金。
- 通过对数据中心的整体设计,发现问题并将问题消灭在萌芽状态,实现整体数据中心的节能减排。
- 通过对数据中心的整体设计,可实现系统的灵活变更、按需扩展。
- 实现统一工程维护,把机房系统作为一个整体产品交付,全网动力监控更易实现,提供主动性服务,并

实现统一和全面的服务。

- 统一工程界面清晰,降低工程成本,并大大缩短工程周期,使网络提早发挥效益。
- 保证维护的持续性及连贯性,保障服务提供的连续性,减少维护风险。
- 通过管理提升 IT 效率,进一步降低日常维护的 IT 财务支出,实现业务的创新和变革。

2 主要功能结构及关键技术

2.1 主要功能结构

本系统的主要功能结构如图 1 所示。

2.2 应用系统数据处理

图 2 给出了集团公司运营监管信息采集与统计汇总平台数据的处理流程,具体介绍如下。

(1)数据采集与审核、管理

对于采集数据,系统提供对各类数据的获取功能,分/子公司的计划与发展、科技管理、运营管理和国际合作等数据通过在线填报、离线填报方式,报送系统集团公司本部;对于现有业务系统中的数据,如财务数据,提供与现有系统接口的方式实现数据的获取。

(2)数据的存储管理

针对来自分/子公司、现有业务系统的数据,系统提供统一的信息存储和管理功能,确保数据的安全。

(3)数据的统计汇总与展现

根据集团公司对业务监管的需求,提供统计汇总功能,并提供灵活、方便查询、实现多样化图表的展现及结果的下载、导出。

2.3 应用系统业务处理

集团公司数据采集与统计汇总系统的用户主要分为 3 种类型,分别为系统管理员、数据管理员和业务用户。其中,系统管理员维护系统基础信息;数据管理员包括集团公司、二级公司的数据管理人员;业务用户包括集团公司、二级公司、三级公司的数据填报人员和使用人员。应用系统业务处理流程如图 3 所示。

系统主要管控指标的分配与数据填报业务流程、数据的查询与导出业务流程。指标的分配与上报流程分 3 级管控模式,由集团公司首先维护报表模板,并下发给二级单位。二级单位有权限对接收的报表模板进行二次分配,形成新的分发规则发给三级单位填报。三级单位填报完成后,将报表提交给二级单位统一审核汇总并上报给集团公司,集

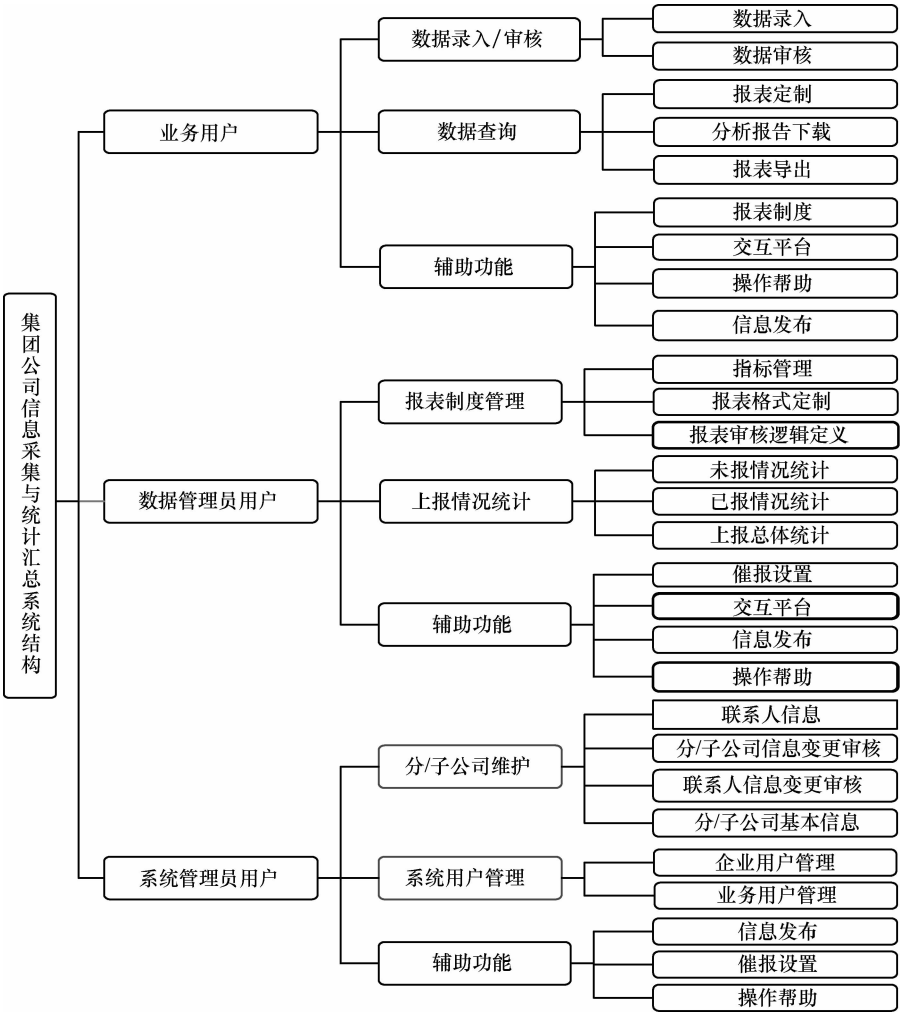


图 1 软件系统功能结构

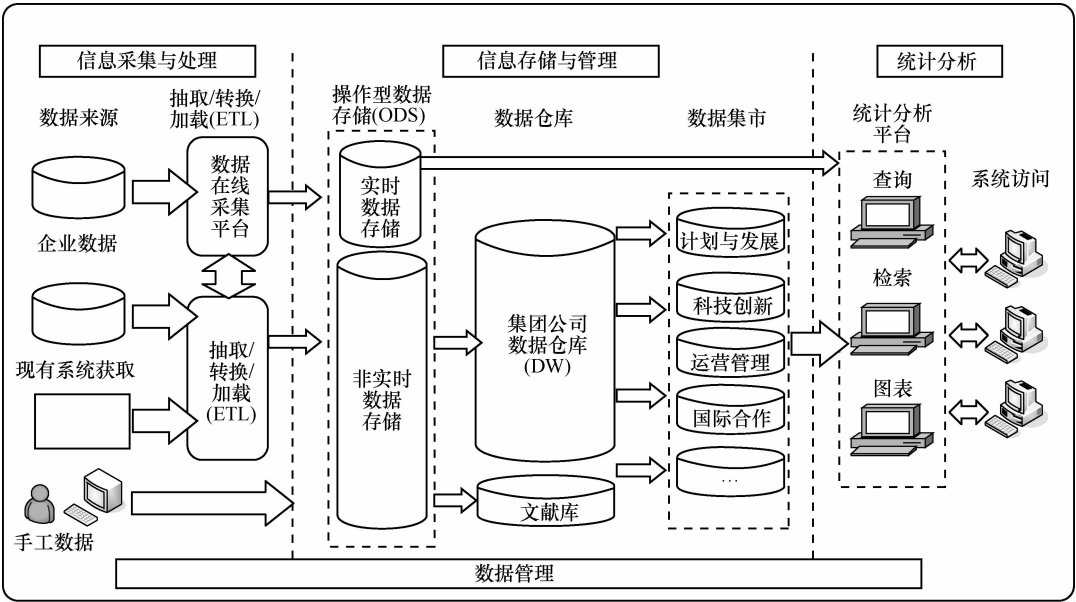


图 2 集团公司运营监管数据中心信息处理流程

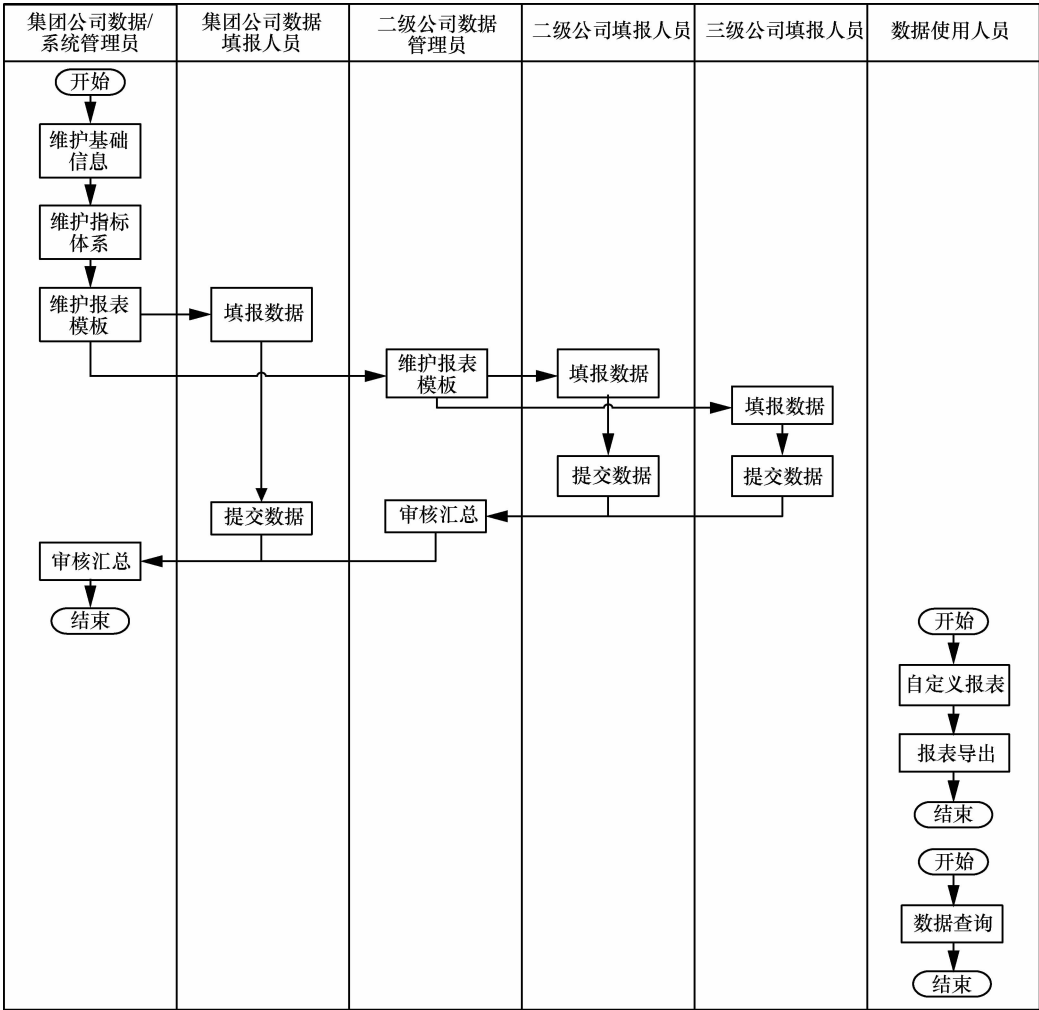


图3 应用系统业务处理流程

团公司进行审核汇总后完成数据的下发与上报流程。定期(按周、月、年等)执行数据分配与填报流程后,系统形成数据池,可供数据使用人员随意组成报表,由系统自动抓取数据以用户希望的表现形式显示数据,完成数据查询流程。

2.4 针对集团经营监管指标体系的功能

针对集团公司的运营监管指标体系,系统提供如下3方面功能。

- 指标属性管理,包括指标属性设置、指标增/删/改/查、指标阈值设置等^[2]。
- 指标体系建模,指标建模的过程,对指标进行编号、命名,并且建立指标之间的逻辑关系、数据来源、数据存储格式、数据展现方式选择等指标属性。
- 指标展现设计,指标数据随时可被已经授权的用户读取,并通过预先定义的各种展现方式,按用户需

要的形式展现给用户,并且支持数据审核、数据导出等数据再利用。

3 系统架构及实现

3.1 系统建设目标

在充分利用现有网络和基础设施环境的基础上,基于集团公司的经营监管指标体系,实现对下属10余家子公司的计划与发展、运营、科技创新和国际合作等业务关键信息的定期(周/月/季/年)跟踪、监测、监控、管理和分析,及时了解各子公司的生产、营销和运行情况的特点、现状和发展趋势,为集团的发展提供决策支持、信息参考,从而实现集团对子公司的高效指导和管控,有效推动集团公司及其下属分/子公司的稳定、可持续发展,进而提升集团的信息化和监管水平。

通过本系统的建设,可以实现如下具体目标:

- 设计公司共享的经营监管指标体系,为集团运营监管提供信息依据和参考;
- 指标信息采集、存储、管理与统计汇总;
- 提供多样化的数据展现方式和导出功能;
- 提供辅助管理功能;
- 提供集团本部与分/子公司的信息交互功能,便于信息的及时沟通、报告下载等。

3.2 系统架构

系统架构如图4所示,包括云平台基础架构及上层平台架构。

推荐配置如下。

- 服务器:长城四子星,3个节点,每个节点4个物理机,单个物理机配置为2个CPU 4核,96 GB内存,2个300 GB硬盘,4个万兆网口。
- 交换机2台:48口,万兆以太网口。
- 盘阵一台:EMC,配12块600 GB盘,可扩展12块(不加扩展柜还可扩展12块2 TB SAS盘),支持iSCSI和SAS混配,8 GB FC 光纤扩展端口或4个10 Gbit/s接口,8个吉比特以太网接口,配40 TB盘,可用20 TB。
- 虚拟磁带库一台,型号为EMC DD140。

3.3 系统实现

系统实现如图4所示的云计算平台,并在此云平台上

实现企业门户系统。

由于集团公司下属有多个分子公司,故本次集团公司的数据中心建设涉及面较广,而且数据中心系统要与企业门户系统集成,与门户系统互相倚靠、相互协调。为确保数据中心项目的成功实施,集团公司企业数据中心系统将根据“统一规划、分步实施”的原则进行建设,具体介绍如下。

(1)数据采集及统计系统开发并与门户系统集成

完成集团公司的数据上报及统计系统的开发实施,并在集团公司企业内部门户中初步展现,并以门户为接入口为集团及各职能部门领导决策提供依据。

(2)优化并完善系统功能

实现各企业在第一阶段没有提出的新功能需求以及需要改进的第一期系统的数据采集及上报、审批、统计的功能(内容和功能优化完善阶段)。

(3)整体优化阶段

实现完整全面的数据采集及统计系统,并与企业门户完美集成,作为企业门户的报表门户和数据采集及统计系统的入口,打造一个真正有意义的综合信息平台,实现企业信息的统一查看、统一入口访问、企业内外流程的统一管理,加强移动应用,深化跨平台协作和企业搜索功能,建立真正的集团公司企业信息业务数据中心,实现信息的交互与共享,消除信息“孤岛”现象,增强企业的市场竞争力;

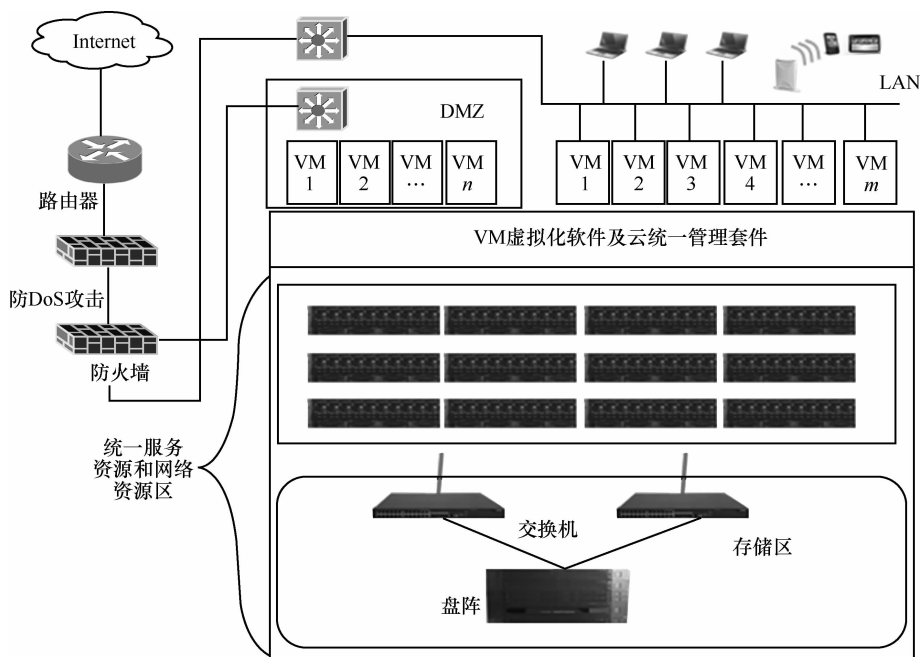


图4 云平台基础架构及上层平台架构



实现与其他企业内部信息系统的全面集成(如财务系统等);实现不同的数据能够为不同公司、不同领导和部门提供参考和决策。

4 创新点及亮点

本系统最大的亮点及创新点是采用了以经营监管指标为体系的指标系统管理模式,随时灵活地增加、修改、减少指标,任意选取不同指标生成不同报表,突破一般以报表为体系的报表管理模式,在数据存储和抽取方面既灵活,又大大提高了数据抽取速度,对于有大量数据需要管理的企业非常适合,对于不断改进和变换经营监管指标的企业也是不可多得的体系管理模式和数据管理平台。

第二大亮点是技术开发平台完全智能化,称为华电天仁智能 Web 平台。智能 Web 平台是一款企业级应用的快速开发平台,可以作为成熟的产品推向市场,也可以在该平台上进行二次开发。开发的报表基于华电天仁智能 Web 报表,是业界第一款基于 J2EE 技术的以纯 HTML 及 CSS 实现的 Web 报表。华电天仁智能 Web 平台采用业界领先的 J2EE 技术构建, MVC 设计模式, B/S 的多层结构,技术上相当领先,具有极高的稳定性、安全性及可扩展性、可伸缩性、可维护性^[3]。

- 利用华电天仁智能 Web 平台,企业及开发商均可以非常方便、快速、高质量地开发复杂的业务系统,包括 CRM、MIS、ERP、HIS、OA 等系统;80%以上的模块不需要编码,甚至零编码,不懂 Java 也可以开发,只需懂数据库即可,开发速度是传统开发模式的 5 倍以上。
- 华电天仁智能 Web 平台内集成了在线自定义 Web 报表引擎,用户除了可以用来开发业务系统外,还可以集成企业内已有的分散的业务数据,制作各类复杂的综合查询、Web 报表、统计分析。

第三大亮点是本数据中心构建在云平台上,采用了最先进的存储技术,加快了数据存取速度,实现了数据存储的安全稳定,节约了硬件资源和网络资源。

5 结束语

本系统云平台由公司统一建设,在云平台上构建虚拟机,在分配的虚拟机上建立数据采集及统计系统应用,并具体实现与以下系统的集成。

与财务系统:数据上报及统计系统需要从财务系统获取相应的财务数据,作为重要指标项的来源。

与集团公司企业门户系统:数据采集及统计系统需要企业门户中的组织机构(LDAP)作为基础登录用户数据,进行单点登录及统一用户管理及身份验证,需要将报表数据以不同形式展现在企业门户系统中,根据用户权限的不同分别展现不同的数据。

与邮件系统:邮件系统作为邮件服务的提供者,需要向企业门户、数据上报及统计系统提供基础的邮件发送及接收服务。此外,也需要同企业门户实现单点登录、统一用户管理及身份验证、邮件来件提醒等功能。

与短信系统:短信系统作为短信服务的提供者,需要向企业门户、数据上报及统计系统提供基础的短信发送及接收服务。此外,也需要同企业门户实现单点登录、统一用户管理及身份验证、短信来信提醒等功能。

总的来说,系统已正式投入使用,目前系统运行良好,用户反馈满意,达到了预期的效果,获得了预期的经济效益。

参考文献

- 1 唐箭.云存储系统的分析与应用研究.电脑知识与技术,2009(20)
- 2 成静静,喻朝新.基于云计算的大数据统一分析平台研究与设计.广东通信技术,2013(1)
- 3 田甜.基于云计算的 GIS 数据采集教学服务平台设计与研究初探.技术与市场,2011(7)

[作者简介]



石桂芳,女,北京华电天仁电力控制技术有限公司信息事业部工程师,主要从事计算机软件系统的开发研究工作。



刘陈荣,男,北京华电天仁电力控制技术有限公司信息事业部副总经理,主要从事计算机软件系统的开发研究工作。

(收稿日期:2013-08-01)