



# 政务大数据交换中心的架构研究

喻 健,刘美伶

(湖北工程学院 计算机与信息科学学院,湖北 孝感 432000)

**摘 要:**传统的政务信息系统一般由各地区、各部门自行建设,没有统一的数据中心,系统集成度低、存储分散、标准不一,再加上各地区、各部门间信息共享和业务协调困难等原因,信息孤岛和信息壁垒问题长期存在且难以解决。政务大数据交换中心的建设就是要联通城市数据资源,构建一个“政务大数据中心”,为城市的各类智慧服务提供统一的平台支撑,让数据多跑腿,群众少跑腿。政务大数据交换中心将成为人民群众和整个城市实时在线的连接纽带,提升政务服务满意度。

**关键字:**电子政务;大数据;数据中心;数据交换

中图分类号: TP311.52

文献标识码: A

文章编号: 1671-931X (2020) 01-0108-05

108

随着大数据和移动计算等技术的应用普及,越来越多的组织机构开始关注大数据技术相关的研究和应用工作,在医疗卫生、教育培训、行政管理等领域产生了大量的成功范例。政府机构管理着大量的数据,且这部分数据准确度高、关联性强,将其与大数据技术结合应用于政务服务及业务管理方面具有重要的价值和意义。

## 一、政务数据的现状

传统的电子政务信息系统在建设过程中,多以纵向单位为网络,建设垂直系统,逐渐形成了各部门信息系统互相隔离的局面。由此就产生了一个个网络不能互联、数据难以共享、重复投资建设的信息系统,进而形成了“信息孤岛”,造成大量软硬件资源的冗余浪费<sup>[1]</sup>。各单位的数据孤立静止,处于存档的状态,其中能保持更新和有效利用的只占小部分<sup>[2]</sup>。为了有效的利用这部分信息资源,应加强建设和打通横向网络体系,为信息互通和数据共享创造有利条

件。建设政务大数据交换中心的一个重要目标就是要借助政府和市场力量盘活尘封的数据,让数据流向它的需求方,发挥数据的社会和经济效益。

## 二、系统功能需求

政务大数据交换中心的建设完善是一个长期过程,建设内容多、涉及面广,各个环节相互关联,需要按照政府各个部门的需求和业务特点有计划、有步骤、有策略地实施。系统主要包括大数据基础支撑平台、政务大数据资源平台、政务数据共享交换平台、政务大数据标准规范体系、政务大数据分析展示应用、与省市其他信息平台接口等模块。

### (一)政务数据标准规范体系建立

建立政务数据标准规范体系,包含基础数据标准规范、数据资源目录编制标准规范等;建立数据交换技术规范,包含数据接入接口规范、数据服务接口规范等;建立运行维护规范,包含数据共享交换工作规范、数据中心管理办法、数据交换中心运维手册

收稿日期:2020-01-15

基金项目:教育部产学研合作协同育人项目“基于 OBE 理念的计算机基础实践课程改革研究”(项目编号:201901292015)。

作者简介:喻健(1989-),男,湖北孝感人,湖北工程学院计算机与信息科学学院实验师,研究方向:计算机软件与应用;刘美伶(1977-),女,湖南郴州人,湖北工程学院计算机与信息科学学院实验师,研究方向:计算机教育。

等。确保在后续操作中有据可循,确保数据的统一性,操作的规范化。

### (二)大数据基础平台建设

搭建可靠、高效、可伸缩的分布式大数据承载平台,提供标准化、开放式的大数据支持接口,为上层应用提供大数据交换及处理能力,为后续各部门业务系统的扩展、数据挖掘分析提供结构化支撑。大数据能力包括了大数据收集、存储、计算、检索和数据挖掘等全方位的基础能力。

### (三)信息库建设

搭建大数据平台,保证硬件的高可用,实现可视化、自动化、集中化的快捷运维平台、开发平台及数据地图等基础模块支撑,构建基础库、部门库、专题库、共享库、交换库等多类数据库,其中基础库又以人口、法人、信用信息、电子证照、空间地理等为主体。大数据资源中心的数据库框架如图1。

### (四)交换平台建设

分散的、不规范的信息难以满足电子政务业务需求,为形成服务于电子政务的统一数据平台,需要

对分散的数据进行整合及标准化改造。建立支持各类信息库的一体化管理功能,在保障信息安全和信息共享的同时,实现信息的集成调度和管理。

## 三、关键技术

### (一)硬件平台高可用建设

数据的安全性和完整性是大数据交换中心最重要的考量之一,硬件平台的高可用首先要尽量消除设备的单点故障,同时要求在单点故障发生时实时切换主备设备,保证大数据业务的连贯和数据的安全完整。为此要为大数据平台所需的硬件预留足够的备份冗余和升级空间,保证存储、计算、网络资源的负载均衡和设备的高可用;设置齐备的多级备份规划和策略,保证整体系统逻辑不受部分设备的问题而导致不可用。大数据硬件平台的物理网络拓扑如图2。

### (二)大数据存储与管理技术

关系型数据库系统是结构化数据存储方案的首选,可以满足大多数应用需求。但是在大数据环境

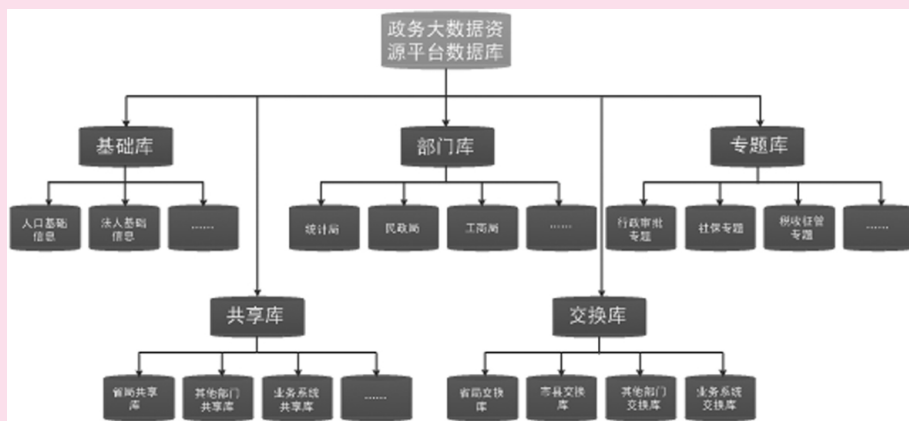


图1 数据库框架图

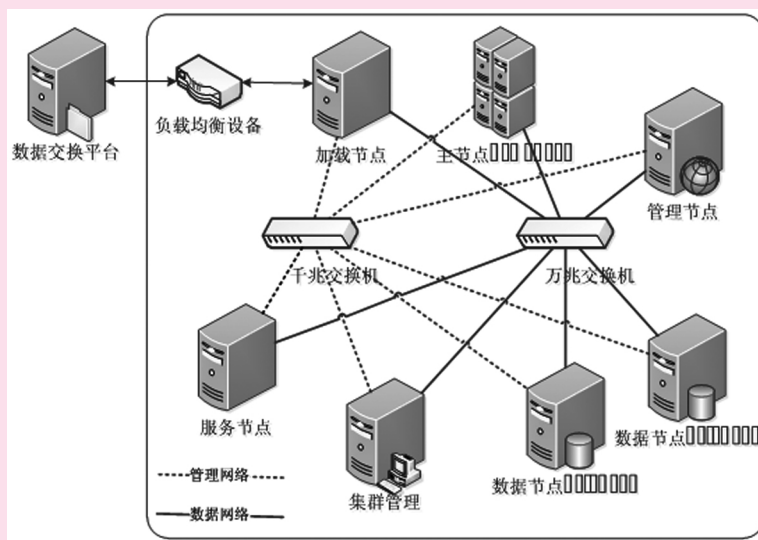


图2 高可用物理网络拓扑

下,半结构化和非结构化数据占了很大的存储比例,而且很多大数据应用是查询多更新少,常常需要针对不同类型的数据内容交叉检索、比对分析和深度挖掘。面对这类应用需求,传统的关系型数据库在技术和功能上都难以为继,由此促成了 OldSQL、NoSQL 与 NewSQL 近年并存发展的局面<sup>[3-4]</sup>。

根据数据的存储结构的不同,大数据管理的技术路线一般可分为三种:一是针对大规模的结构化数据,通常采用新型数据库集群。通过列存储、行列混合存储、粗粒度索引等技术,结合大规模并行计算架构(Massive Parallel Processing, MPP)的分布式计算模式,实现对 PB 量级数据的存储和管理,这类集群性能好、扩展性高,在各类数据分析类领域广泛应用<sup>[5]</sup>。二是针对半结构化或非结构化数据为主的大数据,采用 Hadoop 等分布式集群来进行高速运算和存储。三是针对结构化和非结构化混合的大数据,采用 MPP 并行数据库集群与分布式集群的混合来实现对百 PB 量级、EB 量级数据的存储和管理。一方面用大规模并行计算来完成结构化数据的计算和管理,对外提供强大的 SQL 和 OLTP 型服务;另一方面用分布式集群实现对半结构化或非结构化数据的处理,以支持诸如内容检索、综合分析、数据挖掘等应用。这种混合模式将是未来大数据存储和管理发展的趋势<sup>[6]</sup>。

### (三)数据交换引擎

数据交换引擎的主要功能包括提供共享域内交换节点之间的数据和文件自动交换服务,包括采集、分发、汇总和转发;提供节点与业务系统之间的交换桥接服务,实现数据发布和获取;提供跨域交换服务,实现共享域之间通过对接节点进行数据交换<sup>[7]</sup>。数据交换引擎接收和处理服务请求的流程如图 3 所示。其中主要的模块有服务请求监听端、服务定义文档、日志监控数据库、接口控制模块、管道管理模块、服务解释执行端、传输管道等,各模块的基本功能如下:

服务请求监听端:采用监听机制,能同时接收外界的多个服务请求并作出及时响应。

服务定义文档:文档记录了服务的配置参数、操作流程和其他说明,能指导服务分类和解释执行,是服务解释执行端工作的基础依据。

日志监控数据库:日志模块对整个交换引擎的行为进行实时监控,根据事件的分级把交换引擎在运行过程中的行为和通过管道的数据流向记录到日志中。

接口控制模块:负责控制和协调交换节点的数据接口,驱动数据交换的执行,把操作命令和交换数据高效完整的传送到交换双方的对等接口。

管道管理模块:协同接口控制模块,实现管道的

创建、初始化、动态分配、变更、撤消等操作,控制传输管道完成数据交换的传输任务。

服务解释执行端:负责解释和执行服务命令。服务请求监听端收到服务请求后,把有效的请求传递给解释执行端,解释执行端查询服务定义文档后,根据对应的服务定义执行操作命令,协调其他模块完成数据交换工作。

传输管道:传输管道是数据交换的场所,数据交换的过程可概括为:数据源输出数据、中间件传输数据、需求方输入数据。输出接口把数据提供者的数据读出,再根据服务定义利用数据过滤模块完成数据加密、数据压缩等步骤;处理后的数据经过消息中间件传输到目的节点。目的节点的接口收到数据后利用反向过滤器完成数据解压、数据解密等操作把原始数据还原出来,再通过数据需求方的输入接口读取数据。

## 四、系统架构

政务大数据交换中心的总体系统架构设计如图 4,自底向上包括基础层、数据层、支撑层、应用层和接入层,采用面向服务的体系架构(Service-Oriented Architecture, SOA)进行应用的设计、开发、集成,力求摆脱面向技术解决方案的束缚,集中精力构建政务数据模型,从而提高软件重用、加快应用系统的实施。

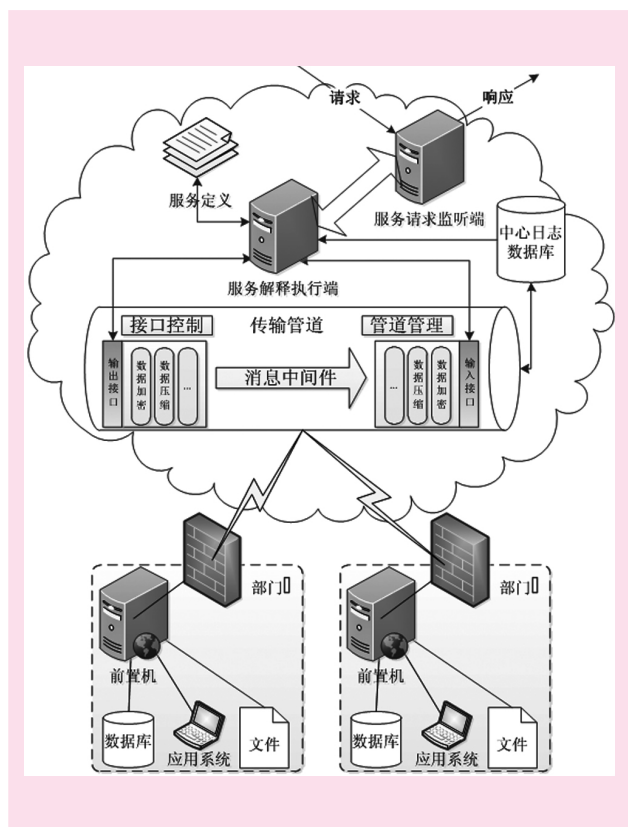


图3 数据交换引擎流程图



### (一)基础层

基础层的工作是将经过虚拟化的计算资源、存储资源、网络资源、安全设施、虚拟设备、负载均衡等以基础设施即服务(Infrastructure as a Service, IaaS)的方式通过网络提供给上层用户使用和管理<sup>[9]</sup>。

### (二)数据层

数据层又可细分为数据源层、数据整理层和数据存储层,数据采集层负责把各种异构数据源中的数据汇集起来,包括传统关系型数据库、数据仓库、分布式数据库、NOSQL 数据库、半结构化数据、无结构化数据、爬虫、日志系统等,是平台的数据产生机构;数据整理层的工作包括数据清洗、转换、加工、关联、标注、预处理、加载、抽取等,作用是将原始数据(Raw data)加工成产品数据(Product data);数据存储层负责存储经过清洗处理后形成的标准化数据,比如元数据、业务数据库、模型数据库等,数据直接面向上层服务,具有高可靠、高并发、高精度的特性。

### (三)支撑层

支撑层采用面向对象和组件式设计技术,将分散、异构的应用和信息资源进行聚合,提供跨领域、与业务无关、通用的基础服务,能适应电子政务系统后期的发展变化进行扩展伸缩。为上层应用系统跨数据库、跨系统平台的无缝接入和集成,提供一个支

持信息访问、传递、以及协作的集成化环境<sup>[9]</sup>。

### (四)应用层

本层基于标准化数据服务,结合各委办局的实际需求建立对应的智慧应用,包括业务系统的统一身份认证系统、政务资源交换系统、“天网”视频共享系统、政府领导驾驶舱、“互联网+”放管服、数据开放、城市运行分析、数据增值应用等。

### (五)接入层

接入层制定相关系统接入平台和提供数据的方式、数据格式和接口标准等,以指导相关系统通过统一的接口将数据归集到数据中心平台。具体包括前置机接入标准规范,数据采集接口使用规范,消息交换接口规范等。

## 五、小结

政务大数据交换中心的建设有利于改变目前存在的“信息孤岛”“信息荒漠”现象,有助于政府各部门实现“职责清”“数据准”的目标。通过对数据信息资源的梳理、建设和共享整合,为今后信息资源的开发利用奠定了基础,有助于实现跨部门的高效信息共享和业务协作,减少各部门在相关领域的重复建设,节约大量的人力物力,促进部门业务流程的优化、整合,提高政府的运行效率。

喻健,刘美伶:政务大数据交换中心的架构研究

111

武汉职业技术学院学报 2020 年第十九卷第二期(总第一百零六期)

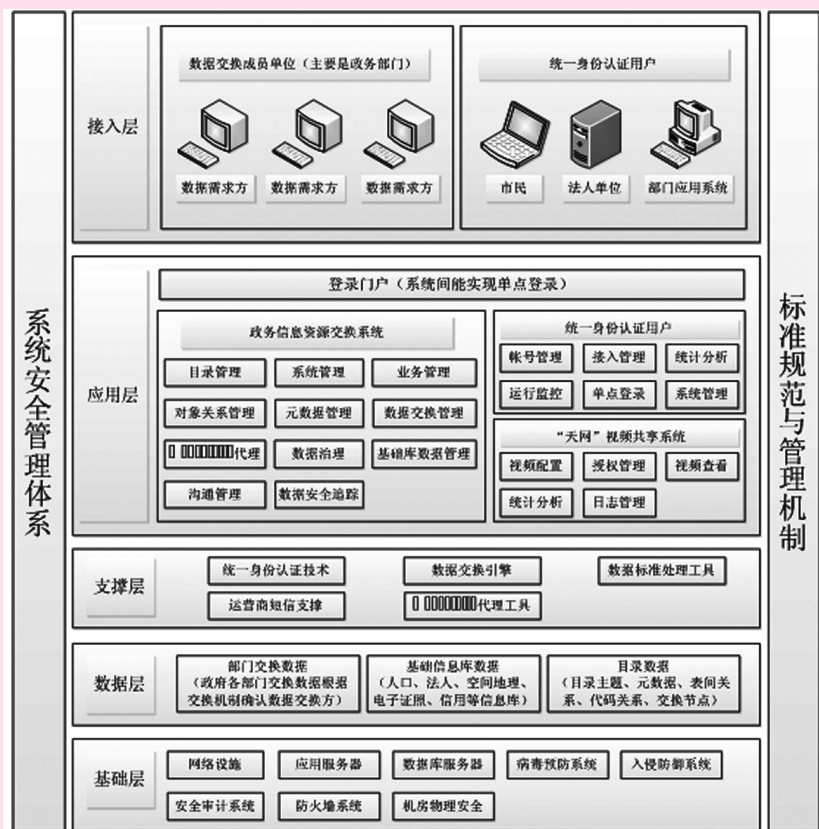


图4 政务大数据交换中心总体系统架构图

#### 参考文献:

- [1] 娄枫.中国电子政务建设的现状和发展趋势[J].中国科技信息杂志,2008,(23).
- [2] 朱锐勋.政府信息资源开发模式比较研究[M].北京:国家行政学院出版社,2016,(3):101-102.
- [3] 闫素霞,班秀萍,叶云龙,等.高校在大数据环境下如何开展数据分析和应用[J].河北北方学院学报(自然科学版),2018,(1):12.
- [4] 陆可,洪绍明,方晗琦.基于大数据的建模方法与理论在社科普及中的应用研究[J].安徽冶金科技职业学院学报,2018,(1).
- [5] 程学旗,靳小龙,杨婧,等.大数据技术进展与发展趋势[J].科技导报,2016,(14).
- [6] 闫素霞,班秀萍,叶云龙,等.高校在大数据环境下如何开展数据分析和应用[J].河北北方学院学报(自然科学版),2018,(1):28.
- [7] 匡晓红,汤丹,刘志峰.一种引擎驱动的电子政务数据交换平台[J].计算机时代,2015,(1).
- [8] 宋梦馨,冯得福.基础设施即服务层企业私有云平台构建[J].信息系统工程,2015,(9).
- [9] 侯莉莎.云计算与物联网技术[M].成都:电子科技大学出版社,2017,(05):10-12.

[责任编辑:胡大威]

## Research on the Architecture of Big Data Exchange Center of Government Affairs

YU Jian, LIU Mei-ling

(College of computer and information science, Hubei Engineering University, Xiaogan 432000, China)

**Abstract:** The traditional government information system is generally built by different regions and departments, without a unified data center, with low system integration, scattered storage and different standards. In addition, due to the difficulties in information sharing and business coordination among regions and departments, the problems of information islands and information barriers exist for a long time and are difficult to solve. The construction of the government big data exchange center is to connect the city's data resources and build a "government big data center" to provide a unified platform support for various smart services in the city, so that the data can run more and the masses run less. The big data exchange center for government affairs will become a real-time online connection link between the people and the entire city, improving the satisfaction of government services.

**Key Words:** e-government; big data; data center; data exchange