

基于云计算技术的 政府统计平台设计的研究*

辛金国¹、龚恺²

(1杭州电子科技大学信息化与经济社会发展研究中心、2杭州电子科技大学经济学院/1教授,浙江 杭州 310018)

摘要:在政府统计工作中,如何充分利用云计算的各种技术,以提高统计数据处理效率,具有重要的实践指导意义。本文基于云计算的理论和政府统计数据的特点,分析云计算在政府统计工作中的具体应用,探索政府统计部门云计算中心的系统架构和关键技术,提出基于云计算的政府统计数据平台的设计层次和应用模式,并建议政府统计云计算平台在建设上统一规划、统一建设等。

关键词:云计算;政府统计;平台设计

一、引言

随着统计信息化的发展,政府统计部门的数据报送和处理工作日渐增加,海量数据的处理超出了现有小型处理机的承受范围,政府统计工作的准确性、及时性、安全性等面临前所未有的挑战。云计算技术作为一种综合应用各种统计分析方法、数据库、智能语言来分析庞大数据资料的技术就应运而生,其在IT行业的研究和应用已取得成功,但在政府统计部门的应用研究还处于探索阶段,研究内容主要集中在系统构想、实现思路和前景展望方面。因此有必要建设网络化、统一化、开放化的统计信息化平台,以适应统计部门多变的业务需求变化。本文基于云计算的理论和政府统计数据的特点,分析云计算在政府统计数据处理中的具体应用,探索统计部门云计算中心的系统架构和关键技术,提出基于云计算的政府统计平台的整体建设层次和管理部署设想模型。

二、相关文献回顾

对于云计算的理解,不同的学者、研究机构和企业提出不同的观点,有的强调虚拟化技术,有的强调客户端,有的强调服务器,因此云计算定义至今也没有统一认识。

美国国家标准与技术研究院(NIST)(2010)认为云计算平台是一种可以提供可用、便捷、按需网络访问的计算模式,进入可配置的计算资源共享池(资源包括网络,服务器,存储,应用软件,服务),这些资源能够被快速提供,只需投

入很少的管理工作,或服务供应商进行很少的交互。

中国网络计算、云计算专家刘鹏(2010)认为:“云计算将计算任务分布在大量计算机构成的资源池上,使各种应用系统能够根据需要获取计算力、存储空间和各种软件服务”。

关于云计算技术在政府统计中的应用,郑慧勇(2011)提出在总结历年使用的统计软件、电子政务及办公自动化系统经验的基础上,利用先进的云计算技术模式,以规范的统计业务流程、统计信息标准为前提,开发集网上直报、数据管理、统计查询、业务自定义为一体的专业通用统计信息管理软件平台,从而构成一个面向统计调查对象、统计工作者、政府相关部门和社会公众的“统计专业数据库云计算平台”。

国家统计局数据管理中心主任许剑毅(2011)提出要加快部署“统计云”。他认为国家统计局和地方政府统计部门需要建立一个国家统一的核心业务系统,即统计云。

综上,云计算由于技术宽广性、实现综合性导致企业从事云计算研发的技术、资金门槛很高,所以目前从事云计算的政府部门、企业还不是很多,而且现阶段云计算的服务都是将自己原有技术和资源中类似云计算的部分(如主机租用、网络存储等)称作云计算。国家统计局的“统计云”正在部署和建设,而地方政府统计平台如何应用云计算技术的探索尚未开展。

三、政府统计平台的基本架构及管理部署

(一)实地访谈情况

为了解云计算在政府统计应用情况,2012年8月以来,笔者实地访谈了浙江省统计局、杭州市统计局、台州市统计局、

注:基金项目:国家统计科学研究(计划)项目“网络直报条件下地区统计数据质量影响因素实证研究”(编号:2009LZ038)和浙江省统计局2012年学术类重点课题(获得一等奖)。

华立仪表等单位,初步了解到大多数政府统计信息化部门的建设模式是一个服务器运行单个应用,大多数时候,这些服务器的负载在15%以下。随着各级政府对统计调查频率的提高,同时统计调查与分析覆盖面的迅速扩大,带来以下问题:一是报送流程长;二是统计人员录入等问题上花费太多时间,难以在统计分析方面做出贡献;三是数据易丢失,且难以共享,难以适应统计业务发展的需要,需要不断购买新的应用系统及硬件设备。

因此,有必要探索使用目前日臻成熟的云计算技术、虚拟化技术,对现有的政府统计设备资源和数据资源进行整合,通过集群的方式组成系统资源池,来满足海量数据、高并发处理的需求,有效解决IT基础设施资源无法共享和协同调度、资源利用率不高、管理和维护成本高等问题,建设网络化、开放化并具有高并发业务处理能力的政府综合统计信息化平台,实现更强的策略执行力度和更高的数据安全。

(二) 政府统计平台的基本架构

Windows Azure是微软所开发的一套PaaS云计算模式操作系统,它具有对数据中心架构的自动管理能力,可自动监控划分在数据中心不同的区块的所有服务器和资源,自行更新补丁,运行虚拟机部署与镜像备份(Snapshot Backup)等能力,系统管理人员只要通过Windows Azure Fabric Controller就能够掌握所有服务器的运行状态。因此本文结合发展较为成熟的微软云计算平台Windows Azure进行设计。

Windows Azure平台提供了多种数据存储方式,主要有Windows Azure存储和SQL Azure。本文设计的平台系统兼顾使用了两种存储方式,示意图如图1所示。

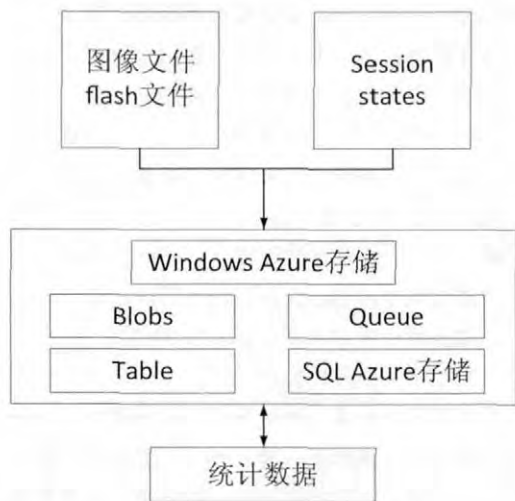


图1 系统云存储方式Blobs

图1中的Blobs数据是用来存储文件、图片、视频、可执行文件、压缩档等二进制格式的文件,另外Blobs可以储存大容量的文件或数据。Table是结构化数据的应用程序存储与管理的一种存储服务,它在Windows Azure存储区中以Key-Value键值对方式存储,并且由Partition来切割在Windows

Azure存储区的存储位置。Queue是一种先到先服务(First-Come, First-Serve),或称为FIFO(先入先出)的存储服务,它可以允许应用程序将信息存储到队列中排队,然后由负责处理的应用程序(通常是Worker)由队列提取信息并处理以后,将信息由队列中移除。

根据统计数据的特点和政府统计系统的实际要求,设想的云计算的政府统计平台系统架构如图2所示。

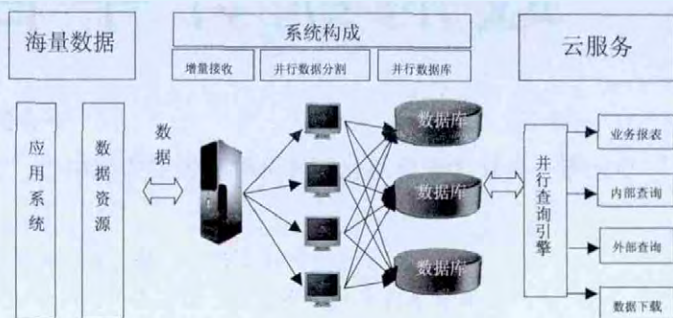


图2 基于云计算的政府统计平台系统

(三) 管理部署

基于云计算的政府统计平台系统以省级政府统计部门作为数据存储中心,可以减少购买和管理数据库服务器的成本和资源投入,另外将统计数据和相关文件存储在云平台后,在云端将由专业人员对存储系统计算机集群进行数据多副本容错、计算节点同构可互换等来保障服务,无需用户管理员进行维护,其专业性和高可靠性使系统更加稳定和健壮。

云计算中心可以为各级政府统计部门提供多种服务,包括为终端用户提供计算业务服务和通过程序接口为需要数据的企业或者高校提供数据查询等服务。通过云计算中心的人机界面,省级、市级和县级政府统计部门及下属各单位,在云计算中心建立各自的账号,分配相应的资源,就可以建立自己的数据空间,然后通过浏览器,录入数据,进行计算和分析。用户在应用程序进程中不需要明白具体的细节问题,以及如何配置计算资源。基于云计算的政府统计平台系统的整体建设层次和管理部署设想如图3所示。



图3 基于云计算的政府统计平台系统的整体建设层次和管理部署设想模型图

图3所示,从系统整体建设和服务的角度,需要具备五个层次要素和三个支撑保障体系,横向层次要素的上层对其下层具有依赖关系,纵向支撑体系对于五个横向层次要素具有约束和保障的作用。分别为:

数据来源层:提供整个系统的原始数据,这其中包括已有的人口、地理、宏观统计数据、企业实时报送数据、下属部门直报数据、统计工作人员调查数据等各类统计数据。

数据传输层:包括互联网、电信网、广播电视网以及三网之间的融合(如:移动互联网),为整个系统提供大容量、高带宽、高可靠的光网络和无线宽带网络所组成的网络通信基础设施。

计算与存储层:包括分布式计算与存储、集中式计算与存储和移动式计算与存储。为整个系统提供数据的存储和计算,保障上层对于数据的相关需求。

数据及服务支撑层:通过数据融合和利用,并通过服务对其进行统一的封装处理和管理,为构建上层各类统计数据应用提供支撑,本层处于整个系统的中上层,具有重要的承上启下的作用。

数据应用服务层:在下面各层基础层之上建立的各种统计数据应用和应用整合,如数据查询、数据存档、数据纠错、数据分析、数据挖掘和预警服务等,为社会公众、企业用户、城市管理决策用户等提供全方位统计数据应用和服务。

标准规范体系:云计算政府统计数据网络直报系统建设中整体所需的标准体系,涉及各横向层次,指导和规范其整体建设,确保建设的开放性、柔性和可扩展性。

安全保障体系:为云计算政府统计数据网络直报系统建设构建统一的安全平台,实现统一入口、统一认证,涉及各横向层次。

运营管理体系:为云计算政府统计数据网络直报系统建设提供整体的运营管理机制,涉及各横向层次,确保整体的建设和管理和长效运行。

四、结论及建议

(一) 基于云计算的政府统计平台服务模式

应该按照对内的内部服务和对外的公共服务两大领域,分别建立政府统计部门公共云和私有云,为不同的对象提供各自需要的云计算应用和服务。

1. 依托互联网建立政府统计部门公共云。从建设模式角度看,政府统计部门的公共云可以是租赁或自建;从运营管理角度看,可以是公益性的,也可以是有偿服务;从应用角度划分,政府统计部门通过外包云服务来提供给公众的各类服务,或者由政府统计部门自行搭建云计算中心,然后向各大企业、政府部门及科研机构提供数据服务。笔者认为比较合理的建设方式是应由政府主导,由政府的IT机构或云提供商依托政府外网来搭建本土化的云计算平台。

2. 依托政府统计部门内部网络建立政府私有云。政府统计

部门的私有云应针对各个统计部门及内部工作人员,以政府统计部门内部网为依托,由政府统计部门的IT机构或云服务提供商来承担建设。政府统计部门拥有一定基础设施和数据基础,可根据政府统计部门管理和需求,在统一的基础设施和数据基础上开发高层应用平台,对数据安全保障、运营管理和服务质量实现最有效的控制。

(二) 政府统计云计算平台建设应着重关注的问题

1. 统一规划、统一建设。平台的基础就是首先统一建设基础设施,然后根据需要各取所需,最后实现推广复制。因此,政府统计部门云计算平台须统一规划、统一建设。需要分析各部门的应用需求,针对不同的需求来开发应用模块,统一按照政府统计部门云服务架构,搭建统一的云计算平台,将政府统计各个部门的各类业务系统整合建设和利用,避免各自为政、分头建设、重复建设而造成资源浪费。

2. 完善管理体制和制度建设。完善政府统计部门云计算建设的配套管理制度。政府统计云计算平台的建设和服务提供将彻底改变现有的政府统计部门信息化管理体制和模式,按照这种模式,政府统计部门应成立统一的信息机构,负责各个部门的信息化建设,各部门不再需要单独配置信息化建设人员,而只需要提出应用需求;信息化项目的管理和统筹需要进一步加强,财政资金的使用将有的放矢、集中管理;部门间的信息共享和业务协同可以方便地实现;政府统计数据服务能力会显著提高。这些都需要在管理体制和制度建设上提前谋划,营造环境。

[参考文献]

- [1] 辛金国,王琳燕,韩秀春.网络直报条件下地区统计数据质量影响因素研究[J].调研世界,2010(08):14-17.
- [2] 刘鹏.云计算[M].电子工业出版社,2010:15-17.
- [3] 王秀花.企业统计报表数据网上直报的几点思考[J].财会与决策,2008(09):55-56.
- [4] 郑慧勇.关于统计数据中心云计算平台的思考[J].调研与观察,2011(03):25-27.
- [5] 侯建,帅仁俊,侯文.基于云计算的海量数据存储模型.中国科技论文在线.http://www.paper.edu.cn.
- [6] 孙健,贾晓菁.Google云计算平台的技术架构及其成本的影响研究[J].电信科学,2010(1):87-88.
- [7] 张和生,张毅,胡东成.海量数据管理框架与方法研究[J].计算机工程与应用,2004,40(11):26-29.
- [8] 王海涛,石晶,周立柱.海量存储系统的可扩展性[J].计算机工程与应用,2002,38(15):72-73.
- [9] 辛金国,严兴良.网络直报条件下企业统计数据质量影响因素分析[J].统计与决策,2013(13):31-33.
- [10] 刘欣洋.复杂网络环境下基于模型驱动的软件进化方法研究[D].内蒙古:内蒙古大学,2011.

(责任编辑:张巧燕)