

A Cloud Computer Architecture Based on Enterprises Cloud*

Shi Zheng

School of Economic and Management, Dalian Nationalities University, Dalian
Email: zhengshi1993@yahoo.com.cn

Received: Jul. 25th, 2012; revised: Aug. 8th, 2012; accepted: Aug. 17th, 2012

Abstract: This paper presents a cloud computer architecture based on enterprises cloud. The architecture is suitable for the requirements of the enterprises. Its application framework of is elaborated by description of enterprises cloud notation, analysis of the difficulties and critical technologies, construction of flow chart, and system functional analysis.

Keywords: Cloud Computer Architecture; Enterprises Cloud; Pattern Recognition; Public Clouds; Business Clouds; High Performance Computing

一种基于企业级云的云模式构架*

郑 时

大连民族学院经济管理学院, 大连
Email: zhengshi1993@yahoo.com.cn

收稿日期: 2012年7月25日; 修回日期: 2012年8月8日; 录用日期: 2012年8月17日

摘 要: 本文提出了一种基于企业级云的云模式构架, 目的是为了提出一个适合企业需求的云平台方案。通过对企业级云模式概念的描述、对其难点与关键技术的分析及流程图构建与系统功能分析, 给出了该模式的应用框架。

关键词: 云模式; 企业级云; 模式识别; 公共云; 商业云; HPC(高性能计算)中心

1. 引言

随着并行计算、分布式计算与网格计算研究的不断深入, 以排队式任务模式的工作流, 目前已被新的计算模式所取代, 这就是云计算模式。在 2006 年亚马逊推出弹性计算服务(即让中小企业按照自己需要(on_demand)购买亚马逊数据中心的计算能力)后, Google、微软、雅虎和 IBM 等国外 IT 巨头纷纷推出了自己的云计算服务^[1-3]。国内多家研究院也纷纷推出了自己的云计算方案与云计算服务或试验中心。

云计算模式此后就以其按需自动服务、宽带接入、虚拟池化资源、快速弹性架构、可测量服务吸引

着大量的业内人士的眼球^[4-8]。然而在云计算诱人模式的背后, 它的安全性、性能、可用性、集成性和可订制能力等都成为了人们担心的问题^[9-11], 而在笔者看来, 云计算由于“公共云”应用中存在着一个理论上的“悖论”而必须大力发展“商业云”, 其方向也必然沿着企业级云的模式发展。

关于“公共云”应用中的“悖论”, 可以沿着如下逻辑去推理。按照云模式的逻辑, 其技术体系结构为物理资源层、资源池层、管理中间件层和 SOA (Service Oriented Architecture)构建层^[12]。其中物理资源层包括计算机、存储器、网络设施、数据库和软件等。在企业传统的度 C-S 模式下物理资源层主要指向了依赖于局域网建设的服务器, 而在云计算模式下,

*基金项目: 大连市科技局科学基金项目(项目号: 2011J21DW020); 中央高校科研业务经费资助项目(项目号: DC10030207)。

物理资源层则指向了所有在线的计算机、移动设备和相关软硬件设备。这就意味着线上必须拥有大量的物理资源。然而云计算模式同时提出支持用户在任何位置,使用各种终端获得服务。从“经济人”角度,就意味着人们会选择类似瘦客户端的模式。极端的讲,只连接显示器而免费占有他人资源。试想如果在线的用户都选择这一模式,则网络上就会出现物理资源不足的情况,这将是云模式的一个致命威胁。因此,云计算模式的发展需要依靠一些新的构想来避免上述极端现象的出现。

本文将提出一种基于企业级云的云模式构架,希望籍此为云计算模式增添一些新的内涵和操作空间。

2. 企业级云模式

云模式的概念一经提出,就衍生了诸多相关概念,如:“公共云”、“私有云”、“商业云”等等^[13]。

国家标准与技术研究院(NIST)提出了一套关于云的定义。该定义提出了4种不同的云配置:

公共云:这种类型的云服务通常遍布整个因特网,能够服务于几乎不限数量的、拥有相同基本架构的客户。亚马逊、Rackspace、Salesforce.com、微软以及谷歌等公司推出的一些流行的公共云产品竞相争妍,提供了丰富多彩的IT服务和商业应用。

私有云:这种类型的云针对单个机构特别定制,例如一些金融机构或政府机构。一般来讲,这类机构都会采用一些虚拟化操作系统和网络技术,因此能够降低使用服务器和网络设备的数量——或者至少能够使这些设备的管理更为明晰。

社区云:这类云专为一系列互不相连的、严格界定的机构而设立。提到这种云,我们会想到供应链或是多个政府机构的联合体这类比较好的使用实例。

混合云:这种云表现为以上多种云配置的组合,数个云以某种方式整合在一起,为一些商业计划提供支持。有时用户可能需要用一套单独的证书访问多个云,有时数据可能需要在多个云之间流动,或者某个私有云的应用可能需要临时使用公共云的资源(这种情况称为“云爆”)。

然而这种分类往往是站在服务商角度的。从企业级客户角度作出分类的概念很少见。

对企业而言,云模式的出现将带来企业管理模式

的变革。作为一个企业客户往往会做成本收益分析(benefit cost analysis),调整企业的管理模式,暴露企业的关键数据来获得高速的数据处理能力,这在普遍对数据处理时间要求不高的企业来讲是得不偿失的。企业管理模式转变和由此带来的企业业务流程重组其根本目的在于提高企业效率。以最常见的营销信息系统为例,其基本模块由内部报告系统、营销调研系统、统计分析系统、情报系统四部分组成。其中统计分析系统需要大量数据存储功能和较高的数据处理能力。如果企业认为有必要,可将统计分析功能交于云计算中心处理。这样在运作中首先需要对云模式加以重新架构。

2.1. 企业级云模式的基本方案

企业级云模式主要由企业级数据处理中心和公共云两部分组成。公共云模式由云计算服务商提供。企业级数据处理中心由云服务商提供服务方案并由该服务商联合小型软件开发商提供。企业的机密数据处于传统的C-S模式下,企业级数据处理中心的服务器或服务器集群构建成自己的局域网。通过分布式结构访问外网,并通过权限设置与认证、许可方式被企业外部用户访问。企业的非机密数据可以通过虚拟传送交于公共云的HPC(高性能计算)中心。具体机制主要是当提出计算任务时,其机密数据将通过一套完整的数据转换技术加以处理。例如:采用设定规则集的方式,转换为对应的另一种形式的“稀疏矩阵”。

简而言之,企业级云模式就是由云计算服务商提供外部计算能力的共享方案,企业自行组织内部数据存储、处理、转换方案的企业商业模型。

2.2. 企业级云模式的难点与关键技术

2.2.1. 机密与非机密数据聚类分析

随着人工智能技术的发展,模式识别已成为一项广泛应用的技术,采用多种聚类分析算法的模式识别技术将数据分为机密与非机密数据,并用明确标识加以区分。

2.2.2. 机密与非机密数据参与的数据计算

企业大部分的业务流程的信息化,都需要机密与非机密数据的共同参与。例如:企业销售利润的获得需要企业的机密数据——成本和非机密数据销售数

据。一旦借助于“公共云”，机密数据就容易暴露给企业外部的人员。因此，需要为参与运算的企业机密数据设计一整套数据转换技术，使其在保证信息安全的前提下，参与“公共运算”。

2.2.3. 数据参与“公共云”或“企业级云”的判别

在数据参与在“公共云”HPC中心的数据计算之前，要进行判别该项作业是否需要利用HPC中心的运算能力，以此确定其运算级别，加以明确标识。一旦某项作业被判别为不需要利用HPC中心的运算能力，则在企业级数据处理中心中运行。

2.2.4. HPC中心的数据回传中的数据管理与清洗

云模式一般都会提供软件系统来保证在运算执行后数据会很容易被找到等一类管理数据服务。在企业接入“公共云”的接口层，需要一道类似“安全检查”的程序，该项程序运行后，冗余及奇异性数据将被去除。

2.2.5. 数据传输中的自动识别与人工审核

为了提高业务流程的运行速度，应尽量减少人工审核，而主要以人工智能技术中的自动识别技术为基础。但必要的流程环节，如：数据是否利用“公共云”HPC中心环节，需要自动识别与人工审核相结合的方式，以防止机密数据外泄。

2.3. 企业级云模式的流程图

图1的流程图企业显示企业数据经过人工与自动识别形成了机密与非机密数据的分类，机密数据需加密后利用HPC中心的运算能力，非机密数据则直接利用HPC中心的运算能力。这类似于混合云的应用，但它的突出区别在于它立足于企业视角，由云模式提供商的“证书导向”模式转换为了企业角度的“数据机密类型导向”模式，从而提高了云模式的实用性与操作性。

3. 企业级云模式的功能分析

表面上看，企业级云模式与商业云模式的概念区别不大，但其功能拓展方面，不仅使企业数据的安全性得到了保证，从深层次的理论意义上看，企业级云模式对于打破云计算模式的“悖论”有着深远意义，其实用价值在这一层面上得到了真正的彰显。具体而言，可以从四方面将其与“商业云”加以对比。

3.1. 理论基础对比

云计算的理论基础是分布式计算、并行计算和网格计算。企业级云计算的基础是模式识别、数据转换技术、分布式计算、并行计算和网格计算。它侧重于人工智能技术的应用方面，相比之下，它要比现行的

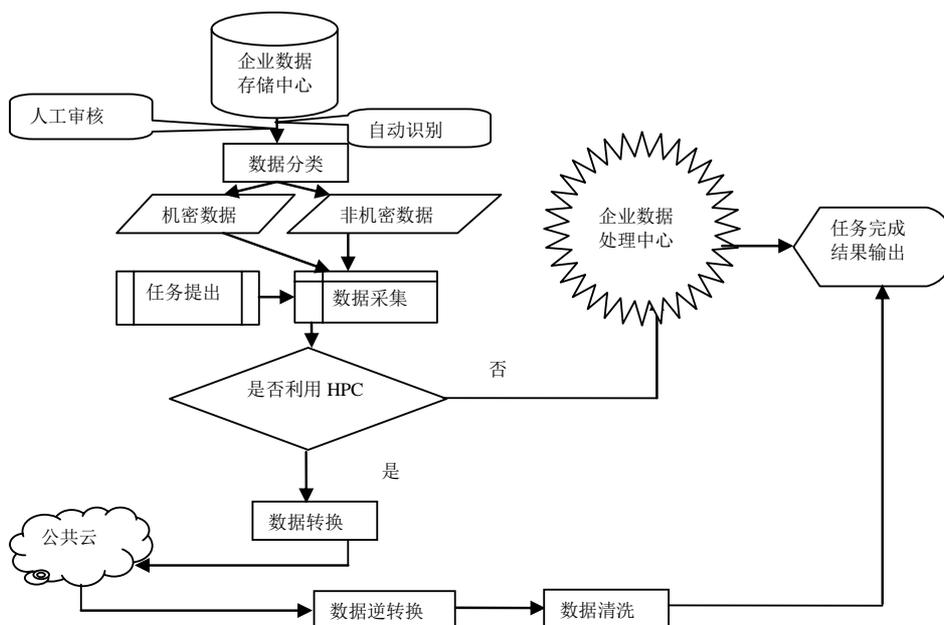


Figure 1. Flow chart of enterprises cloud mode
图 1. 企业级云模式的流程图

基于将现存技术与工具缝合在一起的工程或商业导向的商业云方案更具有理论价值。该模式对于更大范围的拓展云模式内涵具有重要的意义。

3.2. 技术架构对比

现行的云计算系统架构包括分布式文件系统层、分布式计算层、数据分析中间件层、平台应用层。本文的系统架构则增加了基于中间件的企业内部的分布式系统和模式识别、数据转换等应用接口方案。基于中间件的企业内部的分布式系统有效的避免了商务逻辑非必要时在虚拟机上的驻留，也防止了商业代码暴露的问题。应用接口方案的应用则增强企业内部对机密数据的掌控力。

3.3. 应用价值对比

目前对云计算的质疑主要集中在它能否落地的问题，也就是是否可以真正的为广大企业的实际需求服务的问题。截至目前为止，云计算在国外已经被应用到了生物 DNA 信息分析、教育研究平台、实时金融系统等领域^[8]。国内则尚处于起步摸索阶段。据不完全统计，辽宁省目前有 3 家公司以 Paas(平台即服务)的模式研发云计算模式。综上可知，目前云计算技术目前仍掌握在少数 IT 巨头手中。它们利用其长久以来的客户信任度来进一步赢得客户。百度在 2011 年百度世界大会上宣布开放百度云计算平台，将逐步对外提供云计算、存储资源，以及开发环境、应用 API 接口等全方位的云平台服务。这意味着，继亚马逊、Google 之后，百度也将其云计算能力和平台开放为“公共云”。数据显示，迄今为止，百度数据开放平台已拥有资源合作方 3500 多家，应用开放平台则收到众多开发者和运营商提交的应用 3 万多个，审核上线超过 2 万个 APP 应用。

本文提出的方案则有助于更多的软、硬件企业参与到云计算模式中来，这既有利于云计算的“落地”，又有利于从企业的需求角度重新思考技术方案。因此，更具有实用价值。

3.4. 成本收益对比

当前云模式厂商提供给客户的方案包括构建私有云或与公共云相集成两种方案。私有云的构建成本

可能比较高昂，私有云通常是不具备大规模公共云提供商所提供的规模经济效益的。遗留硬件需要手动配置才能迁移到私有云中的服务器，自动化和协同编排管理对于这些老旧机器来说可能无法适用。构建了私有云之后，如果还需要公共云服务的话，也可以转向一种混合模式。这会牵涉到如何保障系统的安全，以及如何在公共云和私有云中核实能够运行的工作负载。这种成本收益上的冲突很难使客户作出迅速的购买决策。

本文提出的方案则立足于利用原有的硬件基础通过软件的智能化来实现系统的少投入和规模经济效益。由于多家软件企业参与到该模式中，可以抑制技术性垄断，满足中小企业定制要求。

4. 一个相关应用实例

某成长型小型制造服务企业，需要对在全国各地的三十余家分公司与工厂实现实时数据的销售管理、库存管理等标准的 ERP 管理，采用国外 IT 巨头的方案，将承担云计算方案设计与架构大约数千美金甚至数万美金的成本以及业务及数据被其完全控制的风险。采用服务器聚集模式大约需要数台应用服务器和数据库服务器及 30 余台客户端机器，相关费用可达数十万元，这其中不包括方案设计费用。采用本文提出的方案，服务器数量可以采用两台服务器分布多个应用的模式，其中 1 台服务器同时作为数据库备份服务器。非机密数据运算采用利用公共云的方式，租用虚拟计算机成本大约为 150 元人民币，如果利用最新的 LOTUSLIVE 方案则只需花费 3.75 美元即可获得 1 GB 的存储空间，这也足以完成基本的数据运算工作。这样，除上述费用，该企业将只需花费方案设计、实施及两台服务器的购置或改造费用几万元即可实现业务流程的升级。值得一提的是本文提出的方案由于数据安全性较高，较之云计算模式最初提出的数据托管方案更易为广大中小型企业用户所认可。

5. 结束语

网络计算之父 Ian Foster 认为，云计算是一种大规模分布式计算的模式，其推动力来自规模化所带来的经济性。从这个意义上分析，企业应用云模式必须在其具有大规模分布计算的需求的前提下。本文的目

的旨在提出一个适合企业需求的云平台方案。通过对企业级云模式概念的描述、对其难点与关键技术分析及流程图构建与系统功能分析,给出了该模式的应用框架。本项工作的优点在于将云计算直接指向企业应用尤其是中小企业应用。不足之处在于对这一问题的研究尚未进入实际操作与应用层面。目前,前期的研究已经具备了较扎实的基础,随后将进行更深入的研究。

参考文献 (References)

- [1] FutureGrid. <http://FutureGrid.org>
- [2] Google App Engine. <http://code.google.com/appengine>
- [3] Amazon. Elastic compute cloud.<http://aws.amazon.com/ec2>
- [4] S. Hazelhurst. Scientific computing using virtual high-performance computing: A case study using the Amazon elastic computing cloud. ACM International Conference Proceeding Series, 2008: 94-103.
- [5] Y. H. Gu, R. L. Grossman. Sector and sphere: The design and implementation of a high-performance data cloud. Science all Hands Meeting, 2008.
- [6] R. L. Grossman, Y. H. Gu. Data mining using high performance data clouds, experimental studies using sector and sphere. In: G. Fenu, S. Surcis, Eds., A Cloud Computing Based Real Time Financial System. 8th International Conference on Networks, 2009, 12: 48.
- [7] J. Gideon, D. Ewa. Scientific work flows and clouds. Crossroads, 2010, 16(3): 14-18.
- [8] M. O. Schatz. Cloudburst: Highly sensitive read mapping with MapReduce. Bioinformatics, 2009, 6(1): 1363-1369.
- [9] V. Ymir, C. Gregory. Clouds at the crossroads: Research perspectives. Crossroads, 2010, 16(3): 10-13.
- [10] 蔡建新, 徐迪威. 基于云计算平台的海量数据挖掘技术在塑料电子商务平台中的应用[J]. 广东科技, 2011, 8: 58-59.
- [11] 谢四江, 冯雁. 浅析云计算与信息安全[J]. 北京电子科技大学学报, 2008, 12: 1-3.
- [12] 王鹏, 董静宜. 一种云计算架构的实现方法研究[J]. 计算机工程与科学, 2009, 31(A1): 11-13.
- [13] 程苗. 基于云计算的经济预测系统研究[J]. 激光杂志, 2011, 32(2): 37-39.