

# 图书馆在云计算时代的机遇与挑战

黎意慧

(广西生态工程职业技术学院, 广西 柳州 545004)

[摘要]对基于云计算的新型信息处理环境进行了深入的分析。结合图书馆发展过程中面临的服务内容不足、信息服务速度太慢等问题,研究了基于云计算模式的图书馆海量信息存储与共享的解决思路。分析了计算资源非常有限的小型图书馆面临的发展机遇,探讨了在云计算环境下图书馆较为可行的发展模式。

[关键词]云计算;云存储;图书馆;挑战;发展模式

[中图分类号]G250.7

[文献标志码]A

[文章编号]1005-6041(2011)04-0010-03

## 1 前言

云计算的概念自2007年由IBM公司提出以来,受到了学术界和工业界的广泛重视<sup>[1-2]</sup>。云计算从提出至今,其定义并不一致,目前比较能够接受的定义是:云计算是将各种动态、易于扩展且被虚拟化的计算资源,通过互联网提供的一种服务<sup>[7-8]</sup>。目前对云计算的研究已不仅仅停留在理论上,一些典型的云计算应用系统已经开始投入使用。比如Google公司开发的基于云计算的搜索引擎,通过遍布在全球各地的数据中心直接对用户输入的搜索请求进行处理,并将各地数据中心搜索的结果,按照排名规则发送给用户。Google公司也正是利用其数据中心的强大计算和存储能力,使得其在信息搜索领域处于绝对的技术领先层面<sup>[3-4]</sup>。不过Google公司构建的云环境只为Google公司提供服务,因此,属于相对私有的“云”。普通用户如果期望也能享受到云计算带来的便利,则需要使用公共“云”。目前IBM公司推出的“蓝云”(Blue Cloud)计算平台和亚马逊公司推出的“弹性计算云”(Elastic Compute Cloud)都是面向普通用户的<sup>[5]</sup>,是一种即买即用的通用云计算服务。我国也于2008年在无锡建立了国内第一个云计算中心,并投入运营<sup>[6]</sup>,在随后,阿里巴巴集团、中科院等单位纷纷建立云计算实验平台或应用环境。总之,云计算的服务离现实应用已经越来越近,因此,有必要研究云计算对图书馆传统信息服务带来的变革与挑战。

## 2 云计算环境中图书馆的发展机遇

对于图书馆而言,在云计算环境中将会获得非常难得的发展机遇,但在发展过程中仍会遇到以下问题。

问题1:急剧增长的信息需求与图书馆馆藏资

源缓慢增加的矛盾。信息时代赋予了人们更多的知识,读者的知识面相比以前要宽广得多,求知欲望也要强烈得多。然而,目前图书馆主要的资源大部分还是集中在纸质文献上,无论是图书数量还是图书质量都远远满足不了读者急剧增长的信息需求,按目前的发展模式,短期内较难解决这个矛盾。

问题2:图书馆之间信息共享欠缺,资源重复建设现象严重。由于图书馆纸质资源与网络中的电子资源难以融合,因此,图书馆在采购图书时往往是孤立地考虑问题,缺乏全局上统筹规划。即便是数字资源的建设,由于图书馆之间交流较少,也同样存在重复建设的问题。这样的后果将直接导致图书馆中的有限资源难以得到充分利用,进一步限制了图书馆的服务范围。

问题3:图书馆建设资金利用率低。目前图书馆建设过程中,占据经费开支的前几项是:图书馆场地建设费、图书采购费、图书维护和管理费、图书馆信息化费等。其中读者能够享受到的相关开支主要是图书采购和信息化的开支,但从图书馆发展的角度看,另外两项开支又是不可避免的。因此,传统图书馆这种建设和建设模式下,图书馆有限的资金利用率普遍不高,建设经费的投入产出比太低。

图书馆面临的这些问题在中小型图书馆中显得更为突出,如果在图书馆发展中引入云计算,则可以在很大程度上缓解这些问题。首先,通过云计算分布式的存储和资源计算,可以为图书馆提供一个宽广的虚拟应用平台。在云计算环境中,云服务提供者通过动态组合各种计算资源,拥有近乎海量的计算资源,各个图书馆根据需要可从“云”中申请资源。目前“云”中的数据中心是计算和存储的有机结合体,在“云”中提供各种海量信息存储功能的同时,也提供了大量通用功能的计算和服务。因此,借

助“云”的拓展平台,图书馆在信息处理平台中上升到了一个全新的高度<sup>[9-10]</sup>。

其次,云计算通过网络可以整合各个独立图书馆的资源,为用户提供更广泛的信息服务。图书馆的纸质图书能够提供的信息量是有限的,因此,绝大多数图书馆都开始引入数字图书资源。除此之外,很多图书馆或者社会上的一些商业机构纷纷将纸质的图书通过扫描或者直接数字化处理,制作成为电子图书,这对传统图书馆的馆藏资源又做了进一步的扩充。然而,单个图书馆或商业机构的力量是有限的,而通过云计算平台,可以将分散在各地的电子图书资源进行有机的整合,剔除重复的图书信息,合并形成一个巨大的、组织有序的电子图书信息库。当某个图书馆接入云计算平台中,向“云”发送某一专题图书资源检索时,可以利用云计算中的数据中心进行并行、分布式的信息检索和结果筛选,最终提供给用户的是一个经过整体筛选后的信息集合。由此,用户在这种应用环境中,图书馆的信息资源得到了极大地扩充,而且在信息扩充时并不是简单的信息堆砌,而是由云计算后台程序将所有信息进行有机组合后提供给用户访问,由此图书资源的数量和质量都能得到极大的提升。

再者,云计算的强大计算、存储、应用平台还能够支持中小型图书馆开发和使用深层次功能的应用服务。云计算环境提供的计算和存储服务大部分都是通用的一些应用服务。基于这些服务条件,图书馆可以极大地提高目前的工作效率,节省下来的精力可以为图书馆提高服务质量创造条件。目前,一些发展比较领先的图书馆纷纷开展了基于数据挖掘的图书管理服务,利于数据挖掘的相关理论和技术,可以发现读者的一般借阅规律和需求变化,及时且有针对性地调整图书馆的馆藏资源。然而在这个数据挖掘的过程中,其前提条件是获得大量真实有效的数据。一些小型图书馆产生的原始数据量比较小,应用传统的技术手段,即使对其进行数据挖掘也难以得到很精确的结果。而在云计算平台中,图书馆能够得到大量其他图书馆的原始数据,大大丰富了可供分析的数据源,一些深层次的数据挖掘和分析工作由此都可以得到开展。

### 3 云计算环境中图书馆发展面临的挑战

云计算时代的到来虽然能够为图书馆的发展提供很多全新的机遇,但与此同时,我们也应该看到这项新兴技术在应用过程中面临的各种挑战。首先是云计算中的数据安全性问题。在云计算完全开放的

结构中,各个用户在方便地申请云计算资源的同时必然存在隐私信息安全的问题,尤其是在云计算环境中,用户可能对同一领域的其他用户愿意公开和共享信息资源,而对于领域之外的其他用户则并不愿意共享资源。并且由于可加入的用户总是处在动态变化中,信息资源存储和处理的位置也在不断发生变化,因此,如何尽可能让“云”中所有用户高效、透明、安全地使用“云”资源,已经不再是传统的权限管理那么简单了。

其次是云计算中的信息管理问题。在云计算系统中由于需要对各种用户、各种信息资源进行存储和计算处理,因此对于这些信息的管理会变得非常复杂,有的管理过程和模式需要根据实际应用系统对数据的访问情况来制定。对于一些更新频率较低、访问频率很高的信息,则需要设计和优化云计算系统中的信息管理模式,使其更能适合于这种应用需求,比如 Google 针对大量信息搜索过程中,用户对信息的访问呈高访问、低更新频率的现象,开发了专门的 BiGTable 的数据管理技术,该技术采用列存储的方式管理数据,能够提供较高的应用效率<sup>[11]</sup>。这种数据管理方式对于图书馆中的信息应用环境同样非常适合。在图书馆提供的信息服务中,更多的是集中在对数据的读取,即使是需要对数据进行数据挖掘处理,同样还是先对原始数据进行读取。当然,在云计算环境中的图书馆发展过程中,面临的数据管理问题不仅仅是数据访问模式的问题,还包括数据分布式存储、数据备份和容灾处理等问题,这些都需要在今后的实际应用过程中加以解决。

最后是云计算的应用程序开发问题。目前对图书馆信息服务系统的开发主要是通过 Visual C++、C#、.NET、Java 等编程工具来实现的。这些编程工具基本上都是属于单机环境中的程序开发工具。而在云计算环境中,由于可利用的条件和资源都有了巨大的变化,因此其编程模式也发生了较大的改变。目前云计算中的编程模式主要采用 Map-Reduce 思想<sup>[12]</sup>,这是一种专门针对并行环境(如多核、多处理机、网络环境等)开发的编程模式,并在这个编程思想指导下,一些 IT 厂商已经推出了一些实用的编程开发工具。但是这种新型的程序开发工具的编程理念和以往相比变化非常大,程序员在这种环境下开发程序,要对云计算的核心思想、组织结构有比较深刻的理解后才能开发出实用的应用程序。这也是云计算环境下图书馆发展过程中必须要解决的一个问题。

#### 4 可行的发展模式

在当前的条件下,图书馆基于云计算的服务可以率先从以下三个方面开展:

(1) 创造接入环境,将图书馆资源融入“云”环境中。云计算作为一个开放的环境,能够为所有想加入到“云”的用户提供接入的环境,与此同时,作为公共的云计算平台,其平台上资源的丰富程度也与各个用户加入的程度有关,因此,图书馆加入“云”环境中得到服务的过程同时也在为其他用户提供资源服务。然而,图书馆当前的接入条件和资源组织方式还并不能完全符合云计算环境的需要,因此,首先需要对图书馆中的资源进行合理的整合与重组,按照云计算环境中的媒体接入要求,将图书馆现有的各种资源形成“云”中资源的一部分,然后对图书馆的网络接入条件,包括通信带宽、接入服务器等设备进行改造,创造能够胜任云计算接入的条件。

(2) 积极应对,学习和掌握在云计算环境中的资源管理技术。当图书馆接入“云”之后,此时图书馆内部自有的资源相对“云”中的海量信息已经显得微不足道了。图书馆需要给读者提供服务的内容绝大部分也来自“云”中的资源,此时图书馆相当于是建设在“云”之上的一个服务机构。为了使图书馆能够为所有的读者提供有价值的信息,图书馆管理人员的工作重点将是对“云”中的信息资源进行整理和调度,将同类型、同主题的信息资源抽取出来为读者提供专一化的服务。这种服务模式相当于将图书馆的规模进行了急剧扩张,图书馆工作人员可管理的信息资源变得极为庞大,如何将这些庞大的信息资源合理组织,使之不是零散的信息堆砌,是图书馆工作人员需要重点解决的问题。因此,在面对这个挑战时,图书馆的工作人员应该以积极的态度应对,加强对自我素质的培养与提高,使得自身的整体能力能够适应云计算环境的需要,更好地为读者提供服务。

(3) 充分利于“云”中丰富的资源,开发基于云计算的图书服务系统。按照“云”的体系结构,在“云”中有一个专门的媒体接入层,该层能够对目前绝大多数的媒体格式兼容,使得各种各样的媒体信息能够方便、快速地接入到“云”资源库中。对于图书馆而言,在发展过程中由于媒体格式不断增多,导致后台处理的任务量日益增加,同时各种媒体之间的信息共享、信息交换等都可能存在兼容问题。为了解决这一问题带来的局限性,图书馆需要在云计算环境中,采用基于“云”的编程模式,开发专业的

图书馆应用软件,为读者提供更为深入和丰富的信息服务。

#### [参考文献]

- [1] IBM 公布“云计算”[J]. 软件世界,2007(23):10.
- [2] IBM 将投资3亿美元备战“云计算”[J]. 信息系统工程,2008(10):76—77.
- [3] NSF 斥资数百万资助14所大学的云计算研究[J]. 现代图书情报技术,2009(6):95—95.
- [4] 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]. 电信交换,2010(2):12—20.
- [5] 彼得·伯罗斯. 亚马逊进军“云计算”[J]. 商业周刊,2008(5):39—40.
- [6] 陆建豪. 云计算:开始实践—TaaS 应用—无锡云计算中心实践[J]. 程序员,2009(12):49—53.
- [7] 唐川. 浅谈云计算的概念问题[J]. 科技情报开发与经济,2010(10):90—92.
- [8] 张建勤,古志民,郑超. 云计算研究进展综述[J]. 计算机应用研究,2010(2):429—433.
- [9] 胡新平. 云图书馆构想[J]. 情报理论与实践,2010(6):29—32.
- [10] 孙卫. 图书馆面对云计算[J]. 中国图书馆学报,2010(3):79—82.
- [11] 吴吉义,傅建庆,张明西,等. 云数据管理研究综述[J]. 电信科学,2010(5):34—41.
- [12] 陈全,邓倩妮. 云计算及其关键技术[J]. 计算机应用,2009(9):2562—2567.

[收稿日期]2011-04-18

[作者简介]黎慧慧(1976—),女,馆员,大学本科,广西生态工程职业技术学院图书馆科技信息室工作,曾公开发表论文20多篇,研究方向:图书馆信息化管理与读者服务。

