

基于云计算的数字图书馆 信息共享管理

李淮斌 著



雅外借

辽海出版社

基于云计算的数字图书馆信息 共享管理

李淮斌 著

辽海出版社

图书在版编目(CIP)数据

基于云计算的数字图书馆信息共享管理 / 李淮斌著. --沈阳:
辽海出版社, 2017.12

ISBN 978-7-5451-4625-7

I. ①基… II. ①李… III. ①数字图书馆—情报服务
—研究 IV. ①G250.76

中国版本图书馆 CIP 数据核字(2018)第 000200 号

责任编辑: 丁 凡 高东妮

封面设计: 瑞天书刊

责任印制: 程 祥

责任校对: 齐巧元

北方联合出版传媒(集团)股份有限公司

辽海出版社出版发行

(辽宁省沈阳市和平区 11 纬路 25 号沈阳市辽海出版社 110003)

廊坊市国彩印刷有限公司

全国新华书店经销

开本: 710mm×1000mm

1/16

印张: 7.5

字数: 110 千字

2019 年 1 月第 1 版

2019 年 1 月第 1 次印刷

定价: 38.00 元



前言

数字图书馆是依赖于现代信息技术而产生与发展的一种图书馆的现代化发展形式,随着互联网技术和计算机技术的逐渐普及和飞速发展,数字图书馆也迅猛地向前发展。云计算技术是一系列计算机技术发展 to 一定阶段的融合性产物,将为包括数字图书馆在内的各个领域带来革命性的深远影响。云计算技术还处于高速发展阶段,本书就云计算的起源与概念进行了详细的归纳总结,重点分析了云计算技术的技术特点与基本结构,以及其在各个领域的应用及影响。在云计算环境下,数字图书馆的建设与服务也有着极大的变化,从基于数字图书馆的云存储,到基于云服务的数字图书馆基础设施整合与建设,再到基于云计算的图书馆自动化等,均是云计算时代数字图书馆领域需要深入研究的。随之而来的,基于云计算的数字图书馆服务模式的变化也是重点内容之一。

Google 是云计算的倡导者,也最早提出互联网应用模式,分布式处理、并行处理和网格计算等计算机概念在网络上实现。云计算是一个计算模型,其计算任务大量分布在计算池上,客户可以按需获取信息服务。“云”也是多种混合演进的产物,因此,采用云计算的模式作为数字图书馆的底层构架,应用平台、数据共享等诸多方面可以极大的改变图书馆的服务。可见,它也是最可靠、扩展性较好的解决办法。我国数字图书馆建设已经取得了较大成绩,但其信息资源整合与共享目前还处在初级阶段,用于指导数字图书馆信息资源整合与共享的管理和保障更处于起步与试点时期。同时和国外相比,我国数字图书馆信息资源整合与共享体制与运作存在着较大的差异性,缺乏可以直接利用的国外资料和经验。因而有必要从我国数字图书馆的实际出发,研究其信息资源共建共享保障体系,以整合有限的人力、物力和财力,加速对数字图书馆信息资源的挖掘和研究,实现其社会效益最大化。

本书在编写过程中参考借鉴了一些专家学者研究成果和资料,在此特向他们表示感谢。由于编写时间仓促,编写水平有限,不足之处在所难免,恳请专家和广大读者提出宝贵意见,予以批评改正,以便改进。

目 录

第一章 云图书馆和云服务.....	1
第一节 云计算及其相关概念.....	1
第二节 云计算与图书馆.....	21
第三节 云计算给图书馆带来的变化.....	26
第四节 云计算环境下数字图书馆的发展构想.....	28
第二章 云计算在数字图书馆中的应用.....	37
第一节 云计算在数字图书馆存储中的应用.....	37
第二节 云计算在联合编目中的应用.....	43
第三节 云计算在数字图书馆信息资源服务中的应用.....	45
第三章 数字图书馆及其服务模式.....	47
第一节 数字图书馆概述.....	47
第二节 数字图书馆的主要服务模式.....	51
第三节 新技术环境下的数字图书馆.....	54
第四章 数字图书馆中云服务的应用策略分析.....	58
第一节 数字图书馆云服务发展现状.....	58
第二节 云服务在数字图书馆中的应用.....	61
第三节 基于云服务的数字图书馆信息安全问题与策略.....	66
第五章 云计算平台下数字图书馆服务模式.....	70
第一节 基于云计算的数字图书馆服务新趋势.....	70
第二节 基于用户需求的个性化服务模式.....	73
第三节 用户定题的主动推送式服务模式.....	76
第六章 数字图书馆信息资源整合与共享之保障.....	78
第一节 数字图书馆及其信息资源建设现状.....	78
第二节 数字图书馆信息资源整合与共享.....	83
第三节 数字图书馆信息资源整合与共享保障的意义.....	86
第七章 我国数字图书馆信息资源整合与共享保障层次.....	88

第一节 数字图书馆信息资源整合与共享宏观保障.....	88
第二节 数字图书馆信息资源整合与共享中观保障.....	90
第三节 数字图书馆信息资源整合与共享微观保障.....	93
第八章 我国数字图书馆信息资源整合与共享保障机制.....	95
第一节 数字图书馆信息资源整合与共享保障原则.....	95
第二节 数字图书馆信息资源整合与共享之法制保障.....	98
第三节 数字图书馆信息资源整合与共享之技术保障.....	102
第四节 数字图书馆信息资源整合与共享之伦理保障.....	106
第五节 数字图书馆信息资源整合与共享之评价体系.....	110

第一章 云图书馆和云服务

第一节 云计算及其相关概念

一、云计算的概念与特点

云计算是分布式处理（Distributed Computing）、并行处理（Parallel Computing）和网格计算（Grid Computing）的发展，是透过网络将庞大的计算处理程序自动分拆成无数个较小的子程序，再交由多部服务器所组成的庞大系统经计算分析之后将处理结果回传给用户。通过云计算技术，网络服务提供者可以在数秒之内，处理数以千万计甚至亿计的信息，达到和“超级计算机”同样强大的网络服务。云计算系统的建设目标是将运行在 PC 上、或单个服务器上的独立的、个人化的运算迁移到一个数量庞大服务器“云”中，由这个云系统来负责处理用户的请求，并输出结果，它是一个以数据运算和处理为核心的系统。

云计算（Cloud Computing）从最初的概念逐步进入现实的应用，在短短的时间里引起了众人的关注。云计算是当前信息行业比较流行的话题，有关它的定义虽不是众说纷纭，但并非完全一致。总体上讲，云计算是一种计算行为或技术风格，特点是在互联网上提供一种动态可扩展的虚拟资源服务。为了满足这种动态可扩展性的要求，云计算服务商必须建立和依靠大型数据中心，它们通常分布在一个国家的各个地区，甚至世界其他国家和地区。可以想象，这样的计算网络有多大、多复杂和多昂贵。云计算中的“云”字是相对互联网而言，用以比喻互联网的复杂结构。

“云”是指一些可以自我维护 and 管理的虚拟计算资源，通常为一些大型

服务器集群,包括计算服务器、存储服务器、宽带资源等。云计算将所有的计算资源集中起来,并由软件实现自动管理,无需人为参与。

云计算概念最早是由 Google 提出的,Google 的创始者因为买不起昂贵的商用服务器来提供搜索服务,只得采用了众多廉价 PC 来替代。他们成功地把这种 PC 集群做到比商用服务器更强大,而成本却远远低于商用的硬件和软件,形成了所谓的云计算技术。2003—2006 年,Google 发表了 4 篇关于分布式文件系统(GFS)、并行计算(MapReduce)、数据管理(BigTable)和分布式资源管理(Chubby)的文章,奠定了云计算发展的基础。

根据维基百科(Wikipedia)的定义,云计算是一种动态的、易扩展的、且通常是通过互联网提供虚拟化的资源计算方式;是分布式处理(Distributed Computing)、并行处理(Parallel Computing)和网格计算(Grid Computing)的发展,是透过网络将庞大的计算处理程序自动分拆成无数个较小的子程序,再交由多台服务器所组成的庞大系统经计算分析之后将处理结果回传给用户。通过云计算技术,网络服务提供者可以在数秒之内,处理数以千万计的信息,达到和“超级计算机”同样强大的网络服务。云计算系统的建设目标是将运行在 PC 上、或单个服务器上的独立的、个人化的运算迁移到一个数量庞大服务器“云”中,由这个云系统来负责处理用户的请求,并输出结果,它是一个以数据运算和处理为核心的系统。

二、云计算的工作原理与核心技术

1. 云计算的工作原理

用户通过终端接入网络,向“云”提出需求;“云”接受请求后组织资源,通过网络为“端”提供服务。

在云计算的模式下,用户终端的功能大大简化,诸多复杂的功能都将转移到终端背后的网络上去完成。用户所需的应用程序并不需要运行在用户的个人电脑、手机等终端设备上,而是运行在互联网的大规模服务器集群中;用户所处理的数据也无需存储在本地,而是保存在互联网上的数据中心里。提供云计算服务的企业负责这些数据中心和服务器正常运转所需的管理和维

护,并保证为用户提供足够强的计算能力和足够大的存储空间。在任何时间和任何地点,用户只要能够连接至互联网,就可以访问云,实现按需随用。

从技术层面上讲,云计算基本功能的实现取决于两个关键的因素,一个是数据的存储能力,另一个是分布式的计算能力。因此,云计算中的“云”可以再细分为“存储云”和“计算云”,也即“云计算=存储云+计算云”(存储云:大规模的分布式存储系统;计算云:资源虚拟化+并行计算)。

云计算并行计算的作用是首先将大型的计算任务拆分,然后再派发到云中节点进行分布式并行计算,最终将结果收集后统一整理,如排序、合并等。云计算虚拟化最主要的意义是用更少的资源做更多的事。在计算云中引入虚拟化技术,就是力求能够在较少的服务器上运行更多的并行计算,对云计算中所应用到的资源进行快速而优化的配置等。

2. 云计算的核心技术

云计算系统运用了许多技术,其中以编程模型、数据管理技术、数据存储技术、虚拟化技术、云计算平台管理技术最为关键。

(1) 编程模型

MapReduce 是 Google 开发的 java、Python、C++编程模型,它是一种简化的分布式编程模型和高效的调度模型,用于大规模数据集(大于 1TB)的并行运算。严格的编程模型使云计算环境下的编程十分简单。MapReduce 模式的思想是即将执行的问题分解成 Map(映射)和 Reduce(化简)的方式,先通过 Map 程序将数据切割成不相关的区块,分配(调度)给大量计算机处理,达到分布式运算的效果,再通过 Reduce 程序将结果汇整输出。

(2) 海量数据分布存储技术

云计算系统由大量服务器组成,同时为大量用户服务,因此云计算系统采用分布式存储的方式存储数据,用冗余存储的方式保证数据的可靠性。云计算系统中广泛使用的数据存储系统是 Google 的 GFS 和 Hadoop 团队开发的 GFS 的开源实现 HDFS。

GFS 即 Google 文件系统(Google File System),是一个可扩展的分布式文件系统,用于大型的、分布式的、对大量数据进行访问的应用。GFS 的设

计思想不同于传统的文件系统,是针对大规模数据处理和 Google 应用特性而设计的。它运行于廉价的普通硬件上,但可以提供容错功能。它可以给大量的用户提供总体性能较高的服务。

一个 GFS 集群由一个主服务器(master)和大量的块服务器(chunk server)构成,并被许多客户(Client)访问。主服务器存储文件系统所有的元数据,包括名字空间、访问控制信息、从文件到块的映射以及块的当前位置。它也控制系统范围的活动,如块租约(lease)管理,孤儿块的垃圾收集,块服务器间的块迁移。主服务器定期通过 Heart Beat 消息与每一个块服务器通信,给块服务器传递指令并收集它的状态。GFS 中的文件被切分为 64MB 的块并以冗余存储,每份数据在系统中保存 3 个以上备份。

客户与主服务器的交换只限于对元数据的操作,所有数据方面的通信都直接和块服务器联系,这大大提高了系统的效率,防止主服务器负载过重。

(3) 海量数据管理技术

云计算需要对分布的、海量的数据进行处理、分析,因此,数据管理技术必需能够高效的管理大量的数据。云计算系统中的数据管理技术主要是 Google 的 BT (BigTable) 数据管理技术和 Hadoop 团队开发的开源数据管理模块 HBase。

BT 是建立在 GES, Scheduler, Lock Service 和 MapReduce 之上的一个大型的分布式数据库,与传统的关系数据库不同,它把所有数据都作为对象来处理,形成一个巨大的表格,用来分布存储大规模结构化数据。

Google 的很多项目使用 BT 来存储数据,包括网页查询,Google earth 和 Google 金融。这些应用程序对 BT 的要求各不相同:数据大小(从 URL 到网页到卫星图像)不同,反应速度不同(从后端的大批处理到实时数据服务)。对于不同的要求,BT 都成功地提供了灵活高效的服务。

(4) 虚拟化技术

通过虚拟化技术可实现软件应用与底层硬件相隔离,它包括将单个资源划分成多个虚拟资源的裂分模式,也包括将多个资源整合成一个虚拟资源的聚合模式。虚拟化技术根据对象可分成存储虚拟化、计算虚拟化、网络虚拟

化等，计算虚拟化又分为系统级虚拟化、应用级虚拟化和桌面虚拟化。

(5) 云计算平台管理技术

云计算资源规模庞大，服务器数量众多并分布在不同的地点，运行着数百种应用，如何有效地管理这些服务器，保证整个系统提供不间断的服务是巨大的挑战。

云计算系统的平台管理技术能够使大量的服务器协同工作，方便地进行业务部署和开通，快速发现和恢复系统故障，通过自动化、智能化的手段实现大规模系统的可靠运营。

三、云计算与网格计算云

云计算和一般数据中心的服务器托管听起来很相似，但实际上存在着差别。

首先，工作环境建立有所不同。目前的数据中心提供的托管环境有享用的，也有专用的，有硬件服务器，也有虚拟服务器，但计算资源对于每个托管的软件都是有限的。如果需要更多的资源，就得增加服务器。而云计算的环境可以随时提供所需资源。比如微软的云计算，开发者不需要和服务器直接打交道，而是与服务模块打交道。为了服务更多的客户，开发者只需指定有多少个软件同时运行。至于数据中心的服务器的启动和管理，由体系管理器来负责。

其次，两者的收费方式也有所不同。服务器托管服务环境通常是按月向用户收取固定费用；云计算服务商则根据计算的时间、信息存储量、计算量等向用户收费。存储量增大，用量增大，信息流量增大，收费也随之增加。

从严格的定义上讲，网格计算（Grid Computing）和云计算不同，但是它们之间的界限越来越模糊。

网格计算是一群计算机或超级计算机来完成一项大型的运算任务，如处理全日食数据。它是将大的运算工作量分成若干块，交给网格中的每台机器去分别完成。微软有 Windows2008 HPC 产品技术就属于网格计算这一类。而云计算一开始并不是针对这种复杂的大型运算，主要是为用户提供具有弹性

的计算资源和托管环境。不过，目前存在着这样的趋势，那就是将网格计算“迁移”到云计算中。比如，有些特殊用户需要做网格计算，他们就会希望云计算服务商能够提供网格计算服务，来帮助他们完成需要的运算工作。至于这一天什么时候到来，很大程度上取决于市场的需要。

云计算的未来怎样？云计算将会对信息工业带来巨大的影响，将给信息技术整体结构带来改变。今后，许多的软件将围绕云计算技术来开发，更多的软件会逐步转移到云计算环境中，更多的用户将受益于云计算服务。

云计算的兴起，对许多公司来说，既是机会，也是挑战。无论是老公司还是新公司，最重要的是能否抓住这样的机会，根据市场的需要，提出具有创造力和技术含量的服务，在竞争中占据胜者之地。

四、云计算服务

除了微软以外，目前提供云计算服务的美国公司还有亚马逊（Amazon）、谷歌（Google）、国际商用机器（IBM）、Salesforce.com 和 EMC 等几大公司为用户提供不同的服务。例如，亚马逊提供虚拟的服务器和托管环境，以及储存空间，但用户必须自己提供或购买服务器和数据库软件；谷歌提供的是编辑语言 Python 和 Java 的托管环境，用户可以将自己的软件送到谷歌云计算环境中运行；IBM 则是为客户提供开发和测试环境；salesforce.com 为用户提供网上管理客户的软件服务；EMC 提供储存技术服务。将来越来越多的公司会提供更多不同的云计算服务。

相比之下，微软的云计算服务应该说是比较全面的。作为一个软件开发和托管平台，用户可以在自己的电脑桌面上利用服务模块编写软件，之后将软件送入微软的数据中心进行运行。在微软云计算环境中，用户不再需要自己拥有服务器和操作系统，也不需要提供或购买服务器和数据库软件，只需根据软件的服务和使用要求来付费就可以了。此外，微软还会将现有的一些网上服务，比如顾客服务（Dynamics CRM），办公室服务（Office Live），邮件服务（Exchange）和文档管理服务（Share Point），加入到视窗天蓝平台上。微软的合作伙伴也将推出建立在视窗天蓝平台的各种服务。

根据云计算平台所提供服务的类型,可以将云计算服务细分为如下七类:

1.IaaS (基础设施即服务)

基础设施即服务 (Infrastructure as a Service), 这种类型的云计算通过创造虚拟的数据中心使得其能够把内存、VO 设备、存储和计算能力集中起来成为一个虚拟的资源池来为整个网络提供服务。用户无需购买服务器、网络设备、存储设备, 只需租用硬件进行应用系统的搭建即可。也有学者将这种计算称之为公用计算或效用计算 (Utility Computing)。

2.SaaS (软件即服务)

软件即服务 (Soft as a Service) 这种类型的云计算通过浏览器或程序客户端为成千上万的用户提供服务, 用户不必购买软件, 只需按需租用软件。在用户眼中看来, 这样会省去在服务器和软件授权上的开支; 从供应商角度来看, 这样只需要维持一个程序就够了, 这样能够减少成本。Salesforce.com 是迄今为止这类服务最为出名的公司。SaaS 在人力资源管理程序和 ERP 中比较常用。中信所与万方数据联合推出的科技创新辅助决策支持系统也是类似的服务。

3.云计算领域的 Web 服务

同 SaaS 关系密切, 网络应用服务商通过提供一些预先定义的函数或 API (Application Programming Interface, 应用程序编程接口) 让开发者能够开发更多基于互联网的应用, 而不是提供单机程序。淘宝提供的 API 服务和开心网提供的 API 服务等就属于此类的服务。

4.PaaS (平台即服务)

平台即服务 (Platform as a Service) 是软件即服务 (SaaS) 的变种, 这种形式的云计算将开发环境作为服务来提供给用户。也就是说, 用户可以在供应商的基础架构上创建自己的应用软件来运行, 然后通过网络直接从供应商的服务器的上传递给其他用户。

5.MSP (管理服务提供商)

管理服务 (Managed Service) 是云计算最古老的形式之一, 这种应用更多的是面向 IT 行业而不是最终用户, 常用于电子邮件的病毒扫描服务、应用

软件监控服务等。

6. 商业服务平台

这种云计算服务融合了 SaaS 和 MSP，它实际上为用户提供了一种交互性服务平台。这在日常的商业贸易领域是非常普遍的，比如用户个人开支管理系统，能够根据用户的设置来管理其开支并协调其订购的各种服务。

7. 互联网整合服务

将互联网上提供类似服务的公司整合起来，以便用户能够更方便的比较和选择自己的服务供应商。

五、云计算的应用平台

由于云计算技术范围很广，目前各大 IT 企业推出了很多种不同种类的云服务，微软、EMC、亚马逊、谷歌等就是代表，下面将简要介绍这几个企业的云服务平台产品。

1. 微软应用平台

(1) Windows Azure

Windows Azure 是微软基于云计算的操作系统，和 Azure Services Platform 一样，是微软“软件和服务”技术的名称，Windows Azure 的主要目标是为开发者提供一个平台，帮助开发可运行在云服务器、数据中心、Web 和 PC 上的应用程序。云计算的开发者能使用微软全球数据中心的储存、计算能力和网络基础服务。Azure 服务平台包括了以下主要组件：Windows Azure；Microsoft SQL 数据库服务，Microsoft.Net 服务；用于分享、储存和同步文件的 Live 服务；针对商业的 Microsoft Share Point 和 Microsoft Dynamics CRM 服务。

2008 年 10 月 27 日，在洛杉矶举行的专业开发者大会 PDC2008 上，微软首席软件架构师 Ray Ozzie 宣布了微软的云计算战略以及云计算平台——Windows Azure。The Azure Service Platform (Azure) 是一个互联网级的运行于微软数据中心系统上的云计算服务平台，它提供操作系统和可以单独或者一起使用的开发者服务。Azure 是一种灵活和支持互操作的平台，它可以被

用来创建云中运行的应用或者通过基于云的特性来加强现有应用。它开放式的架构给开发者提供了 Web 应用、互联设备的应用、个人电脑、服务器、或者提供最优在线复杂解决方案的选择。

Windows Azure 以云技术为核心, 提供了软件+服务的计算方法。它是 Azure 服务平台的基础。Azure 用于帮助开发者开发可以跨越云端和专业数据中心的下一代应用程序, 在 PC、Web 和手机等各种终端间创造完美的用户体验。

Azure 能够将处于云端的开发者个人能力, 同微软全球数据中心网络托管的服务, 比如存储、计算和网络基础设施服务, 紧密结合起来。这样, 开发者就可以在“云端”和“客户端”同时部署应用, 使得企业与用户都能共享资源。Ray Ozzie 说道。“今天, 无论对于开发社区还是对于微软来说, 都是一个转折点。我们所提出的技术将改变原来的游戏规则, 同时为 Web 开发者和企业开发者带来新的机遇。

微软会保证 Azure 服务平台自始至终的开放性和互操作性。企业的经营模式和用户从 Web 获取信息的体验将会因此改变。最重要的是, 这些技术将使用户有能力决定, 是将应用程序部署在以云计算为基础的互联网服务上, 还是将其部署在客户端, 或者根据实际需要将二者结合起来。

(2) Windows Azure 服务平台

Windows Azure 服务平台包括 5 个主要部分: Windows Azure, Live Services, Microsoft SQL Services, Microsoft.NET Services, Microsoft Share Point Services & Dynamics CRM Services。

Windows Azure, 用于服务托管, 以及底层可扩展的存储, 计算和网络的管理。

Microsoft SQL Services, 可以扩展 Microsoft SQL Services 应用到云中的能力。

Microsoft NET Services, 使得可以便捷地创建基于云的松耦合的应用程序。另外还包含访问控制机制可以保卫你的程序安全。

Live Services, 提供了一种一致性的方法, 处理用户数据和程序资源,

使得用户可以在 PC、手机、PC 应用程序和 Web 网站上存储、共享、同步文档、照片、文件以及任何信息。

Microsoft Share Point Services and Microsoft Dynamics CRM Services，用于在云端提供针对业务内容、协作和快速开发的服务，建立更强的客户关系。

Azure 就是微软云计算所有服务的基础平台，从 Live 服务，到数据服务，到提供 Share Point 和 Microsoft Dynamics CRM 的空间服务。应用程序即可以运行在云中，也可以运行在本地系统。

(3) Windows Azure 优点

Azure 服务平台的设计目标是用来帮开发者更容易地创建 web 和互联设备的应用程序。它提供了最大限度的灵活性、选择和使用现有技术连接用户和客户的控制。

利于开发者过渡到云计算：世界上数以百万计的开发者使用 Net Framework 和 Visual Studio 开发环境。利用 Visual Studio 相同的环境创建可以编写、测试和部署的云计算应用。

快速获得结果：应用程序可以通过点击一个按钮就部署到 Azure 服务平台，变更相当简单，不需要停工修正，是个试验新想法的理想平台。

想象并创建新的用户体验：Azure 服务平台可以让你创建 Web、手机、使用云计算的复杂应用。与 Live Service 连接可以访问 4 亿 Live 用户，新的使用新方式与用户交流的机会。

基于标准的兼容性：为了可以和第三方服务交互，服务平台支持工业标准协议，包括 HTTP、REST、SOAP、RSS，和 Atom Pub.，你可以方便地集成基于多种技术或者多平台的应用。

(4) Windows Azure SDK 开发工具

Azure 服务平台使开发者能够在原有的技能、工具和技术（如微软的 .NET Framework，Visual Studio 等）的基础上，进一步提高应用程序开发的灵活性和有效性。开发者也可以选择其他的商业或者开源的开发工具或技术，然后通过 Azure 服务平台提供的通用 Internet 标准，比如 HTTP、REST、WS-* 和 Atom Pub 来进行操作。微软为我们提供了云计算服务 Azure 相关的开发工具

包，包括以下三个组成部分：

Complete offline development environment, including computation and storage services; Complete command-line SDK tools and samples; Visual Studio add-in that enables local debugging.

Windows Azure SDK 应运而生，对微软云计算感兴趣的程序员们可以适当关注。微软的 SDK（软件开发工具包）工具包运用 Java 和 Ruby 语言。这个工具包可使 Java 和 Ruby 开发者依托 Azure 平台应用 .NET 服务。不过，到目前为止，该工具包只允许 Java 和 Ruby 开发者在微软数据中心编写服务程序。官方下载地址：Windows Azure Software Development KIT (January 2009 CTP)。

2. EMC ATMOS

EMC ATMOS 是第一套容量高达数千兆字节（petabyte，简称 PB）的信息管理解决方案。Atmos 能通过全球云存储（cloud storage）环境，协助客户将大量非结构化数据进行自动管理。凭借其全球集中化管理与自动化信息配置功能，可以使 Web 2.0 用户、互联网服务提供商、媒体与娱乐公司等安全地构建和实现云端信息管理服务。

Web 2.0 用户正在创造越来越多的丰富应用，文件、影像、照片、音乐等信息可在全球范围共享。Web 2.0 用户对信息管理服务提出了新需求，这正是“云优化存储”（cloud optimized storage，简称 COS）面世的主要原因，COS 也将成为今后全球信息基础架构的代名词。EMC Atmos 的领先优势在于信息配送与处理的能力，采用基于策略的管理系统来创建不同层级的云存储。例如，将常用的重要数据定义为“重要”，该类数据可进行多份复制，并存储于多个不同地点；而不常用的数据，复制份数与存储地方相对较少；不再使用的数据在压缩后，复制备份保存在更少的地方。同时，Atmos 可以为非付费用户和付费用户创建不同的服务级别，付费用户创建副本更多，保存在全球范围内的多个站点，并确保更高的可靠性和更快的读取速度。

EMC Atmos 内置数据压缩、重复数据删除功能，以及多客户共享与网络服务应用程序设计接口（API）功能。服务供应商通过 EMC Atmos