云存储架构与关键技术的研究

**曾文英**1,2**，赵跃龙**1**，宋 威**1,3

[wyzeng@126.com](mailto:wyzeng@126.com) [ylzhao1@scut.edu.cn](mailto:ylzhao1@scut.edu.cn) [song.wei@scut.edu.cn](mailto:song.wei@scut.edu.cn)

（1.华南理工大学计算机科学与工程学院，中国广州 510640；2.广东科学技术职业学院计算机工程与技术学院，中国珠海 519090；3.广东工业大学计算机学院，中国广州 510640）

**摘 要：**本文提出了云存储系统的通用架构，分析了各个模块的功能，并讨论了其中的关键技术，等。云存储是一种新型的存储服务模式，它由服务提供方通过互联网向客户提供存储空间和数据存储服务，同时客户不需要了解其中的细节、底层结构和机制。本文推荐的云存储架构是层次化的、相互协作的，且讨论的关键技术包括部署、存储虚拟化、数据组织、转移和安全性等。运行机制包括生态链、博弈论、蚁群优化、数据生命周期管理、维护和更新，同时也分析了收敛和演变机制。所以本文从总体上阐明了云存储系统，并加入了新的观点。

**学科分类：**操作系统；存储管理；分配与收集策略；辅助存储；存储层次；虚拟内存

**一般术语：**算法；管理；度量；文档；性能；设计；经济学；可靠性；实验；安全；人为因素；标准化；语言；理论；法律方面；验证

**关键词：**云存储架构；关键技术；运行机制；生态链；博弈论；蚁群优化

**Keywords**

Cloud Storage Architecture, Key Technologies, Operation Mechanism, Ecology Chain, Game Theory, Ant Colony Optimization.

1. **概述**

云计算是一种新兴的计算平台与服务模式，它基于互联网组织并安排服务。云存储就是服务之一，它使用基于云计算的远程存储服务器来提供存储资源和服务。云存储可以提供更为廉价且高可靠性和安全性的存储服务。

云存储系统是一种协作型存储服务系统，它拥有多台设备，多个应用域和多种服务形式。云存储的发展利益于网络带宽、Web 2.0，存储虚拟化，存储网络，应用存储与服务器及存储设备的集成，集群技术，网格计算，分布式文件系统，内容分发网络，对等网络，数据压缩，数据加密等技术的发展。

本文后面的内容按如下安排：第2节介绍相关工作；第3节提出云存储的通用架构；第4节讲述构建策略；第5节讨论云存储服务器的关键技术；第6节分析云存储系统的运行机制；第7节做个总结。

1. **相关工作**

现在有许多云计算和云存储提供商，如国际商业机器（IBM），谷歌（Google），太阳微系统公司（Sun Microsystems），微软（Microsoft），亚马逊（Amazon），EMC，NetApp，惠普（HP），Nirvanix，日立数据系统（HDS），赛门铁克（Symantec）等。

另外也有越来越多的云存储平台，如，HDFS，GFS，Sun Network.com，SkyDrive，Amazon S3，EMC Atoms，Hitachi Content Platform，FileStore，和KFS等。

存储网络行业协会（SNIA）提议云存储首创方（CSI）将云数据管理接口（CDMI）标准采纳为云服务标准。

Yunhong Gu等人拟议的文献[1] 允许用户像使用本地磁盘文件一样地与建立在大规模分布式节点上的大量数据集打交道。用户不必在多个节点中去定位数据，管理数据，也不必向系统中添加新的节点或删除已有节点。

文献[2] 介绍了MetaCDN，一个利用“存储云”资源来创建能为内容创建人员提供廉价且高性能的内容分发网络（CDN）的集成重叠网的系统。MetaCDN可以通过依据服务质量、覆盖范围及预算以智能地匹配和替换用户的内容到一个或多个存储供应商来去除处理多个存储供应商的复杂性。MetaCDN提供一个单独统一的命名空间使得内容创建人员及消费者驾驭多个“存储云”的性能及覆盖范围工作变得琐碎，统一命名空间也使得它更容易集成到原始网站，且对终端用户来说更为透明。

文献[3] 提出了一种在广域网之上及时存储转移机制，在存储的分布式转移方面具有参考价值。

云计算与云存储的飞速发展将产生云资源市场，并带来云服务选择的挑战，文献[4] 提出过一些这个方面相关的算法。Ying Zhan等[5] 提出的云存储管理技术也是个值得关注的紧要问题。

Albert Greenberg 等人[9] 讨论了数据中心的云服务的费用，包括服务器（45%），基础设施（25%），电源（15%）和网络（15%）。所有的资源要根据用途从资源池中动态地分配出来以适应需求和费用。为了减少花费且提高灵活性，所采取的策略有位置无关寻址，统一带宽和延迟和安全与性能隔离等。用于资源消耗塑型的市场机制被采纳来提高效率，同时地理多样化数据中心也被用来提高点对点性能和可靠性。这些想法也可做为云存储的参考。

在2009年的存储峰会[10] 中，许多人认为关于云的主题已经引起了众多信息技术的教授，开发人员，市场人员，出版社和分析专家们的注意，诸如云计算，云服务或者云存储，这里仅仅列出不同变体。

云存储产生了一些新的应用类型，包括面向服务的体系架构，基于网络虚拟化的低成本、能通过单一命名空间提供随时随地访问、大量数据存储，共享和协作的网络服务应用程序接口及统一服务接口，等。

目前云计算和云存储还没有全球标准的规格说明和通用架构。本文将分析云存储的需求，提出架构，并分析其中的关键技术。

主要想法是通过集成并提高当前的架构、分布式模式、应用域等来构建一个低成本、容错性好、可靠、可扩展、高性能、公平的云存储联盟系统。

本文创造性地利用了生态链、服务市场、博弈论、蚁群优化的概念，并讨论了一些关键技术点。

1. **云存储架构**

不同的云存储服务平台上有许多云存储架构制式。它们通常复杂且互不相容。我们提出一种分层的普遍的云存储架构。而云存储是基于云计算的一种服务类型。

* 1. **需求分析**

云存储是为普遍的存储需求和大量的存储愿望而筹备的。它的操作和持久性是由云系统提供支持的。程序、数据、文本、图片、视频等都需要保存到云系统中。移动终端、个人电脑、消费类电子产品如相机、智能手机、mp3、mp4等需要越来越多的存储资源。通常，本地存储空间有限，且容易丢失，多个设备之间的数据一致性也难以得到保证。

因此，存储需求的必要性及普遍性使得云存储必须具有廉价、容易维护、可靠性好、安全、可恢复等特点。

* 1. **云存储架构**

云存储由上千个通过网络集群的存储设备，分布式文件系统和其它为用户提供存储服务的存储中间件构成。传统的云存储结构包括资源池，分布式文件系统，服务等级协议（SLA）和服务接口等。全球范围来说，它们可以在物理和逻辑的功能边界和关系上进一步划分以提升兼容性和交互性。基于这个想法，提出如下架构设想。这是一种分层的模型。架构从底向上依次是网络和存储基础设施、存储管理、元数据管理、存储叠加层、服务接口。下面将会讨论功能细节。

|  |
| --- |
| 服务接口 |
| 存储叠加层 |
| 元数据管理 |
| 存储管理 |
| 网络和存储基础设施 |

**图表 1云存储的分层模型**

在网络和存储基础设施层中，有分布式的有线、无线网络，存储网络。

存储管理层中，地理上分散的存储资源通过域及逻辑入口被组织起来，数据可以以文件或数据块的形式保存在存储介质上。

元数据管理层集群全球范围内域的数据存储元数据信息并协调不同的域以达到负载均衡。

在存储叠加层中，虚拟化和服务检索及重定向将被实现。它可以被看成是一个中间件，连接分散的存储设备到一个虚拟的存储网络并简化的、标准的数据结构暴露给服务接口。

在服务接口层，云存储系统提供客户端统一接口来访问和过滤非法客户端。

服务分发模式是云存储的一个重要方面。存储资源可以看成是商业产品，有很多商业理论和经验可以引导云存储服务。

* 1. **模块功能**

根据云存储的架构，基础设施层包括网络、硬件节点、本地操作和文件系统。存储管理层包括本地存储组织和远程存储重定向。元数据管理层意味着数据可以集中，也可以分散到多个节点上以提升查询服务性能。存储叠加层用于云存储系统的抽象、存储资源的虚拟化及调试优化。服务接口层提供客户端访问接口和应用接口域API调用。

1. **构建策略**

云存储实际上就是把存储作为服务的一种实现。构建云存储的策略包括需求分析、容量预估及计划的性能、部署、验证、分配、维护和更新。主要目标是在经济且实用的条件下构建一个可用的、可靠的、协作的、可扩展的、安全的、可并发的存储系统。

为了构建云存储系统，带有相关管理软件的分布式的存储设备应该通过虚拟化，集群和整合联合起来向用户提供一个统一的容量无限的虚拟存储资源池。

在构建云存储过程中，服务质量（QoS）是存储性能中一项重要指标。云存储的服务质量是指存储率、存储延迟、带宽、职责、可靠性、卷大小、生命周期、可恢复性、安全性等方面。

1. **云存储服务的关键技术**

云存储的关键技术包括来自服务器、网络、客户端及相关的控制度量，如可用性、可靠性、虚拟化、反馈、信用度、安全等的诸多方面。云存储系统应该支持自动管理，分布式合作，数据整合，服务等级协议（SLA）匹配，服务质量（QoS），认证，访问控制，权力分配，审查等。

以下讨论主要着重于云存储服务器和一些相关的控制度量。

* 1. **云存储部署**

云存储的部署阶段包括以下项目：需求分析，存储资源重定向，优化和演变等。

* + 1. *需求与基础*

云存储的规模应当由应用需求和技术基础决定。通用的存储网络由中间件和叠加层集成，而分布式的存储资源可能是NAS，SAN，对等网络（P2P）存储等。

* + 1. *存储资源重定向*

地理位置由应用的数据需求来选取。基本规则是就近原则及冗余的自主复制管理以适应访问过载。

* + 1. *优化与演变*

花费应该基于部署模式进行优化。历史过程和服务器及客户端的反馈可收集起来用于高速分布式策略和访问控制。

* 1. **云存储虚拟化与可用性**

虚拟化的概念被应用到诸多领域，如操作系统、服务器、网络、存储等等。存储虚拟化是指在在访问过程中将逻辑存储映射到物理存储上去。云存储虚拟化有助于对用户或客户端隐藏存储的地理位置、存储方式及其他技术细节。

云存储的可用性包括持久性和可恢复。高可用性就需要保障应用服务质量。

标准通用的文件系统，如NFS，CIFS和GFS等均都采用。就像Linux的虚拟文件系统（VFS）一样，新的云存储文件系统在不久的将来也会推进。

* 1. **云存储数据组织**

云存储的数据组织方式可以是数据库式的，文件级的，或块级的。数据库可以是商业的数据库产品，也可以是开源的数据库。信息以记录的形式组织起来，以提高检索速度。但是数据库只能管理一些特定类型的数据。文件级组织方式可以很灵活并能根据应用处理方式而改变。块级组织方式是一种底层的数据格式，数据库和文件都是基于数据块的。纯粹的块级数据没有语言，它必须与其它存储管理模式结合起来。面向对象的存储是一种新兴的存储模式，如果加入一些自主的操作它将会变得智能。

* 1. **数据转移与负载均衡**

云存储的数据转移是指将数据从一个云系统移动到另一个地方的云系统。它旨在云系统间的协作及负载均衡。当存储容量使用量超过一定的阈值时，数据应当被转移到另外的云存储单元上，并在原来的位置上保留指针，或者同时修改更新元数据。

负载均衡就是在云系统中保持不同存储设备均有可用容量提供后面的应用使用。它可以全面提高存储职责和可用性。数据转移是负载赵均衡的一种有效机制，但是会给网络带宽和I/O进程带来额外的工作量，而且它不能解除并发客户端的访问瓶颈。

数据复制是数据转移的一个特例，它将原始数据保留着。数据复制是分布式云存储系统的单点故障的优秀的解决方案，它在不同设备和地点保留相同内容的多份拷贝。理想的云存储系统应能根据客户端的访问频率和存储服务器的工作量自主地创建必要的拷贝。

数据转移可能会花费相对较长的时间且可能因为存储介质和信道的错误而导致数据丢失。这个问题引起了许多关注。所以对于企业级存储[6] 来说对用灾难恢复的中央存储站点就显得十分必要。

* 1. **重复数据删除**

重复数据删除[7] [8] 是一种的新技术，主要应用于存储备份、恢复和压缩打包以减小内部重复数据所占用的空间。重复数据删除是显著减小数据容量、削减存储需求和使数据保护的花费及风险[7] 最小化的最好的方法。为了适应指数级增长的企业和科学数据，就需要巨大的存储空间，而重复数据删除将节约相对大的存储空间并减小开支。在巨型规模的云存储中，重复数据删除将是节省存储容量和使数据移动更安全可靠的优良解决方法。但是仍存在一个问题就是：在哪里进行重复数据删除？云服务器上，还是客户端？

* 1. **存储安全**

存储安全牵涉到存储介质的物理安全性和数据安全性。像一般的网络存储一样，云存储的安全性包括认证、授权、审查和加密等。通过自发的冗余复制数据，一旦宕机可以很容易得利恢复。

云存储安全性也可以扩大到云服务的整个流程中，包括硬件、软件、数据、信息、网络安全和客户端隐私安全等。

云存储趋向于联合云安全以提供更为强健的安全机制。

1. **运行机制**

云存储系统的运行机制是指动态的、长周期的存储组织，服务器分发和演变并且开支随着时间和存储空间的增长而增大。在不同的时期和地点，应用可能有不同的存储需求。整个云存储系统可为看成是一个存储资源在不断消耗和产生的生态系统。我们认为博弈论和蚁群优化可能可以提高性能和效率。

* 1. **云存储生态链**

云存储可以看作是一个存储在不断产生和消耗的生态链，它包括云存储提供方（硬件提供方，软件提供方），存储整合方，存储破坏方，数据中心，内容提供方和客户端。上层与下层之间可以通过租赁合同和加入的超值服务交互。生态链可以被分为两条子链，即存储资源生产链和存储资源消费链。，如**图表 2**所示。生产链将为消费链提供云服务。

云存储破坏方

云存储整合方

云存储提供方

云存储客户端

内容提供方

数据中心

云资源生产链

云资源消费链

**图表 2云存储生态链**

作为云存储中的一个服务，云资源可以用灵活的方式进行分配，如客户订购，内部用户的企业存储云服务或外部用户的第三方存储云服务，和来自多个云存储服务提供商的联合存储云服务。

* 1. **云存储中的博弈论**

事实上，现在已经有很多云服务提供商，并且他们已经组成市场。云服务市场需要合作与竞争以平衡并推动云存储技术的发展。博弈论在云存储的构建和协作阶段得到应用。我们现在正在研究这个。

假设有多个云存储服务提供商，如果他们在向客户提供存储服务时合作组成联盟可以得到更多利益。在这个存储联盟中，谁向客户提供存储资源可以看成是一个博弈论问题。他们会向地理位置上最近的提供资源以减小传输延迟和交流上的开销。

* 1. **蚁群优化**

蚁群优化可以在云存储的消耗阶段得到应用。在这一节中将牵涉到存储服务在多个云存储提供商中的选择问题。当存储客户端请求云服务时，选择哪个存储提供商或存储服务将基于蚁群优化理论。

初始时客户端可以随机地选择一个云存储服务，并且在本地和服务路径（路由器或存储服务器）上保留体验与满意度。当下一个请求发送出去时，上次的体验与满意度将被当前的存储服务器用作选择的参考并更新相关参数，就蚁群优化算法中的信息素一样。信息素可以根据一些规则和服务器负载情况来调整以适应客户端和云存储用户的选择行为。

* 1. **数据生命周期管理**

云存储可以为任何类型的客户端提供存储资源，并且费用可以依据存储容量和带宽进行周期性调整。云存储的数据生命周期可以基于服务器的配置，或者基于存储服务初始时服务器与客户端之间的合约。无论访问模型是桌面电脑、笔记本、移动终端、应用服务器，还是应用程序接口（API）[12] ，它们可以采用与云存储服务器交涉过的关于存储资源周期的租用时间，当要过期时，存储空间就被释放掉。对于重要的数据，生命周期可以通过续租延长。

* 1. **维护与更新机制**

云存储将业界领先的技术——包括DAS，FTP/NAS，iSCSI/SAN，TAPE，分布式文件系统（如NFS，GFS，HDFS），集群和网络计算，面向服务的体系架构，等——结合为一个终极安全、可靠性[12] 的统一存储环境。维护与更新由云存储管理和监控系统集中化、自动化控制。所以客户端不需要了解复杂存储管理的细节和复杂的运行机制，只需要通过统一接口按需求来申请存储资源。

* 1. **存储收敛与演变**

从不同云存储提供商提供的存储资源往往是分级且不兼容的。当一个客户端从一Storage resources from different cloud storage providers are often hierarchical and incompatible. When the clients get storage service from one cloud storage providers, they often can’t migrate the stored object or data to other ones, else if they ask for another storage service request again. It is not flexible and may be expensive if some application needs the same data based on different cloud computing platforms. So the cloud storage providers should make interactive interfaces to share with each other’s data and contents seamlessly, and the fee of communication and convention mayn’t be paid by the clients. This is the ideal situation of cloud storage convergence and evolution. For this goal, open source cloud storage platforms should be encouraged to deployment although security and risk of the data in cloud storage must be considered and ensured at first. There will be multiple convergence modes, i.e. centralized convergence, distributed convergence; pre-organized or ad hoc; local area or wide area or global area, etc.

According to the Sun Microsystems [13], cloud computing support every facet, including the server, storage, network, and virtualization technology that drives cloud computing environments to run virtual application to assemble in minimum time. Cloud storage as one kind service based on cloud computing, i.e. Storage as a Service (SaaS), can be used to store and hold application, business, and personal data, and integrate photos, maps, and GPS information, science data, etc. Cloud storage should be combined with other cloud services such as cloud database, cloud data, cloud security, etc. More and more valued services on cloud storage may be developed, e.g., mobile business, mobile learning, mobile game, etc.

Another evolution trend may be that there will be more wide range of cloud storage providers not just from information technology, but also from communication operators, enterpriser data centers, common community, personal storage sharing pools, etc.

1. **总结**

本文提出一云存储的架构并讨论了相关的关键技术。尽管云存储的应用已经得到实用且快速的发展，商业集成和运行机制仍然需要统一的规格和标准。本文提出云存储的分层架构，讨论了部署、虚拟化和可用性、数据组织、数据转移和负载均衡、冗余数据删除、存储安全等。在运行机制中，生态链、博弈论、蚁群优化和存储资源收敛与演变被提出，并值得未来进一步的研究。

1. **致谢**

本次研究工作受到中国国家自然科学基金委员会编号为60573145的拨款，湖南省自然基金会05JJ30120号拨款，广州科学技术项目2007J1-C0401号拨款，及中国高等教育博士学科点专项基金200805710019号拨款的支持。

1. **参考文献**
2. Yunhong Gu, Robert L. 2009. Grossman. Sector: A high performance wide area community data storage and sharing system. Future Generation Computer Systems, 20 May 2009.
3. James Broberg, Rajkumar Buyya, Zahir Tari. 2009. MetaCDN: Harnessing ‘Storage Clouds’ for high performance content delivery. Journal of Network and Computer Applications 32 (2009), 1012–1022.
4. Takahiro Hirofuchi, Hidemoto Nakada, Hirotaka Ogawa, Satoshi Itoh, Satoshi Sekiguchi. 2009. A live storage migration mechanism over wan and its performance evaluation. Proceedings of the 3rd international workshop on Virtualization technologies indistributed computing, Barcelona, Spain, 2009, 67-74.
5. Wenying Zeng, Yuelong Zhao, Junwei Zeng. 2009. Cloud service and service selection algorithm research.GEC '09: Proceedings of the first ACM/SIGEVO Summit on Genetic and Evolutionary Computation, Shanghai, China, June 2009, 1045-1048.
6. Ying Zhan, Yong Sun. 2009. Cloud Storage Management Technology. Second International Conference on Information and Computing Science. Manchester, England, UK, May 21-May 22, 2009, icic, vol. 1, 309-311.
7. Henry Newman. 2009. Why people don’t like to use cloud storage? http://www.cnw.com.cn/storage-Technology/htm2009/20091013\_183980\_2.shtml, 2009-10-13.
8. FalconStor Software, Inc. 2009. Demystifying Data Reduplication: Choosing the Best Solution. http://www.ipexpo.co.uk/content/download/20646/353747/file/DemystifyingDataDedupe\_WP.pdf, White Paper, 2009-10-14, 1-4.
9. Mark W. Storer Kevin Greenan Darrell D. E. Long Ethan L. Miller. 2008. Secure Data Deduplication. StorageSS’08, October 31, 2008, Fairfax, Virginia, USA. 2008, 1-10.
10. Albert Greenberg, James Hamilton, David A. Maltz, Parveen Patel. 2009. The Cost of a Cloud: Research Problems in Data Center Networks. ACM SIGCOMM Computer Communication Review, Volume 39, Number 1, January 2009:68-73.
11. SNIA CLOUD Storage Summit. 2009. http://www.snia.org/events/wintersymp2009/cloud/, Held at the WINTER SYMPO SIUM 2009.
12. Steve Lesem. 2009. Cloud Storage and The Innovator's Dilemma. http://cloudstoragestrategy.com/cloud-ecosystem/, July 19, 2009.
13. Soft Layer Technologies. 2009. CloudLayer™ Storage. http://softlayer.com/cloudlayer\_storage.html, 2009-10-15.
14. Sun Microsystems, Inc. 2009. Introduction to Cloud Computing architecture. http://www.sun.com/featured-articles/CloudComputing.pdf, White Paper, 1st Edition, June 2009:1-32.