

云数据管理研究进展与展望

赵婧 胡亨梅

1 引言

随着互联网的日益普及和 IT 技术的迅猛发展, 互联网数据急剧膨胀, 如何存储和管理海量数据已成为一个亟待解决的挑战性问题。云计算的概念应运而生, 它改变了数据存储的基础架构。伴随着云计算的产生, 云计算的研究备受关注, 云计算的概念对于每一个 IT 界的人都不陌生。然而, 关于云计算的发展前景, 所有的专家学者意见并不统一。

一方面, 伴随着的云计算应用的出现, 改变了数据的存储方式, 催生了一批新企业的诞生, 这极大地促进了整个科学的进步与发展。云计算的研究话题也引起了极大地关注, 在刚刚已召开的国际顶级学术会议 SIGMOD2010、VLDB2010 以及 CIKM2010 CloudDB workshop 上都有大量的关于云计算相关的研究成果。

另一方面, 倘若我们仔细分析研读, 会发现这些云数据管理的相关研究可以归结为以下两类: (1)将原有的研究问题使用 Map-Reduce 框架进行实现。(2)在云数据管理这个新的应用背景下, 使用传统的索引、查询等技术来解决云数据管理的问题。

综上所述, 云计算确实是一个划时代的变革, 对科技地发展起着极大地推动作用, 然而, 云数据管理的研究并没有什么新的实质内容产生, 无异于新瓶装旧酒。这是否意味着云计算的研究毫无意义呢? 当然不是! 我们知道美国国家自然基金投入大量资金进行云计算的研究, 而且随着 Google、Yahoo!、Facebook 等企业的推动, 出现了不少基于云计算平台的数据管理系统, 而且大部分系统已经投入生产环境使用。因此, 云计算是真实的, 但目前的云数据管理研究现状却表明云数据管理的研究并不真实。导致这一现象的原因归结起来主要就是云计算的应用并没有明确显现, 这也就导致了云计算上的需求不够明确, 进而导致创新性研究的缺乏。研究的创新性依赖于需求的变革。以互联网为例, 由于 TCP/IP 体系结构的发展, 互联网在七十年代迅速发展起来, 最初的应用仅局限于 e-mail(电子邮件)、FTP(文件下载)和 telnet(远程登录)等在大学和特殊领域推广的互联网应用, 互联网技术研究也在起步阶段, 随着应用的不断明晰, 各种技术研究不断成熟, 进而又产生各种应用需求, 各大互联网公司逐渐发展起来, 循环地推动了研究与技术的进步。因此, 只有新的应用模式逐渐明晰, 才有可能带动相关技术的发展, 云计算的内涵才能够真正明晰, 创新的研究模式才能够带动起来, 进而达到云计算真正的成功。

本文接下来将主要从三个方面进行详细介绍云数据管理的相关研究。主要包括: (1)云数据管理的研究现状; (2)云数据管理的主要研究问题; (3)总结和展望。

2 研究现状

云计算的核心思想，是将海量的通过网络互连起来的计算资源通过统一的管理和调度，形成一个抽象的计算资源向用户按需提供服务，对于用户来讲，云计算的使用就好比用电一般，需要多少用多少，随时取用。这是一个美好的愿景，但是从系统管理的角度讲仍然存在许多亟待解决的问题：系统结构、数据模型、可扩展性、一致性、容错性等等。其中最为突出的两个问题在于海量和容错。海量包含多个角度：集群规模、数据量以及用户量。无论从那个角度来讲，都使管理系统面临着严重的挑战。另一方面，云计算平台往往采用廉价、不可靠的 PC 机来搭建 shared-nothing 集群，因此出错几率高于传统的分布式数据库中的高性能服务器。而这个问题随着集群规模的增大显得尤为突出[7]：查询设计的云数据规模越大，需要工作的节点就越多，而在查询中节点出问题的概率也会越大。文献[8]指出，Google 公司平均每个分析任务都会遇到 1.2 个节点错误。如果每次出错查询操作都需要重新启动的话，那么一个分析任务很有可能永远无法完成。

伴随着云计算概念的兴起，尤其是在传统的关系数据库面临各种压力与挑战的情况下，NoSQL 数据库应运而生。顾名思义，NoSQL 数据库即打破了传统的关系数据库的范式约束。由于传统的关系数据库在应付高并发读写、海量数据的高效存储和访问以及对数据库的高可扩展性和高可用性的需求显得力不从心，暴露出许多难以克服的问题，进而引发了 NoSQL 运动，其拥护者认为，关系数据库的许多主要特性面对当前的挑战非但无用武之地，反倒掣肘系统的功能及性能。比如对于数据库事务一致性需求、写实时性和读实时性的需求以及复杂的 SQL 查询，特别是多表关联查询等等。因此，各种 NoSQL 数据库放弃了关系数据库强大的 SQL 查询语言和事务一致性及范式的约束，或者采用 Key-Value 数据格式的存储以满足极高的并发读写性能，或者采用面向文档的方式以保证系统满足海量数据存储的同时具备良好的查询性能，又或者针对可扩展性展开的可伸缩数据库以增强其弹性的扩展能力。近年来，随着 NoSQL 运动的蓬勃发展，人们从初期的打破传统的关系数据库约束逐渐演变成对当今数据存储及管理可行且高效灵活的方案的探求，这与云数据管理的目的是极为相似的。在云数据管理中，我们同样要解决的是传统的关系数据库在数据及查询压力下所暴露出的实时插入性能、海量存储能力、迅速的查询检索速度以及无缝扩展等问题。NoSQL 数据库与云数据管理两者殊途同归，从满足应用需求的角度来说，最终都渴求找到一种集一致性、可用性和高容错性于一身的数据存储及管理方案以应对日益高涨的数据管理需求。

但是，根据 Eric Brewer 教授所提出著名的 CAP 理论，一致性（C: Consistency），可用性

(A: Availability), 分区容错性 (P: Tolerance of network Partition) 三者不可兼得, 必须有所取舍。关注一致性, 就需要处理由于系统不可用而导致的写操作失败的情况, 而如果关注的是系统可用性, 就需要做好读出的数据并不一定是最新值的准备。传统数据库保证了强一致性 (ACID) 和高可用性 (侧重 CA), 所以分布式数据库的集群实现非常困难, 其扩展性受到了一定程度的限制。而近年来不断发展壮大的 NoSQL 数据库, 尤其是 Key-Value 数据库就是通过牺牲强一致性, 采用 BASE 模型, 用最终一致性的思想来设计系统, 使得系统达到高可扩展性和高可用性 (侧重 AP)。但是, 对于 CAP 理论也有一些不同的声音, 数据库大师 Michael Stonebraker 在[9]中提出, 为了 P 而牺牲 C 是不可取的。事实上, 数据库系统最大的优势就对一致性的保证, 如果我们放弃了一致性, 也许 NoSQL 比数据库更有优势。

与云计算在工业界引起的热潮相似, 近年来在学术界也涌现了大量云数据管理方向的炙手可热的话题。各种研讨会相继召开, 汇集了大批对云计算充满兴趣的研究者、开发者、用户以及前沿实践人员。包括 The ACM Symposium on Cloud Computing 2010 (ACM SOCC 2010), VLDB'2010 的座谈会 Cloud Databases: What's New?, CIKM'2010 的第二届 CloudDB Workshop 等。我们在此简要介绍一下第二届 CloudDB Workshop 的内容。第二届 ACM 云数据管理国际研讨会 (CloudDB'10) 于今年 10 月 30 日在加拿大的多伦多作为 CIKM'10 的相关会议召开。此次研讨会的主要目标是讨论处理基于云计算框架下的大规模数据管理的挑战问题。研讨会共收录 8 篇学术论文, 研究方向包括查询效率的提高、数据挖掘计算、MapReduce 和 DBMS 的自适应算法、云安全、云框架、数据备份应用、系统测评以及二分图数据管理等。可以看出, 在云数据管理领域内现有的研究工作在某种意义上来说还不成熟, 还有很大的上升空间。本次研讨会中的工作重点大多集中在将现有的网格和 MapReduce 技术采用到云环境下进行开发。与会者一致认为云计算中依然存在着许多开发性的挑战, 比如云数据的安全性和查询处理的高效性等。总体而言, 本次 workshop 上次工作主要还是集中在如何将现有的技术如何应用在云计算环境中, 研究工作仍然不够成熟, 留有很大的上升空间。同时与会者一致认为, 在云数据管理领域中还有许多开放性问题的研究, 如云安全、高效查询等等。

3 主要研究问题

经过我们实验室接近两年的研究与探索, 我们总结出云数据管理的研究问题主要包括以下几个方面: (1) 云数据存储问题的相关技术研究: xml 数据存储、key-value 存储以及 Object 存储问题。(2) 查询处理问题的相关技术研究。(3) 索引管理的相关技术研究。(4) 事务处理的相关技术研究。(5) 云数据管理的隐私保护问题的相关技术研究。下面将就这些问题进行详细的阐述。

3.1 云数据存储

随着云计算的不断发展,互联网数据的不断膨胀,逐渐呈现出数据的多样化,云数据存储面临的问题将会越来越明显,经过我们的详细调研与研究,我们总结出未来云数据存储主要面临的问题分为以下三大类: XML 存储、Key-value 存储、Object 存储。

XML 作为 Internet 数据表示和交换的工具,已经成为了众多应用领域中的标准数据格式。XML 文档的存储方式极大地影响了查询处理的效率,成为一个非常重要的研究方向。云数据管理系统的出现为 XML 文档的存储提供了一个极好的存储平台,那么究竟如何在云数据管理系统中进行 XML 文档的存储必将是一个亟待解决的研究课题。

针对绝大部分的检索都是基于主键的查询应用,使用 Key-value 存储将会是一个很好的选择。它被广泛应用于缓存、搜索引擎等领域。同时,一个好的 key-value 存储系统需要满足一些条件:可用性、可扩展性、故障恢复以及高性能。简单来说,就是数据不能丢失,服务不能中断,能对故障进行感知并能自动恢复,读写性能极高。目前已有一些开源的云数据管理系统默认就是以 key-value 格式存储数据,譬如 Cassandra,然而这些系统针对实际的应用仍然有一些问题。针对实际应用,结合开源系统进行存储的研究开发具有极大的意义。

对象存储提供了具有高性能、高可靠性、跨平台以及安全的数据共享的存储体系结构。其组成包括智能存储接口和设备,以及分布的元数据管理。在对象存储系统中,客户端可以直接访问存储设备,减少了数据存储路径中的控制路径。在对象存储中,使用对象存储设备(Object-based Storage Device, OSD)进行物理的数据存储。OSD 是连接到网络上的存储设备。它可以是磁盘、磁带或者其他的存储介质,并具有自我管理功能。伴随着云数据管理的兴起与发展,在云数据管理系统上进行对象存储也是一个极具意义的研究课题。

3.2 查询处理

从对查询接口的支持来看,目前的云数据管理系统大多只支持简单的关键字查询,不能支持数据分析工作中经常使用的聚集、连接等查询操作。目前的云数据管理系统大多基于采用键-值模型存储数据的分布式文件系统构建。当查询被提交时,系统利用键值对数据进行搜索,这样限制了查询接口的种类,也不利于查询的处理和优化。推动云数据管理系统的动力是对于大规模海量数据的高级管理,而且其面向的用户也具有多元化的特征。未来的云数据管理系统在处理海量数据时,必须能够提供高性能查询处理。如何针对云计算平台上数据的海量性、分布性以及冗余的特性设计查询优化策略及算法显得尤为重要。

查询优化是分布式数据库系统,尤其是规模巨大的云数据库系统中的核心问题。查询优化的最终目标就是要找到最小的通信代价和最低的算法复杂度。分布式查询处理就是将一个分布式数据库上的高级查询转换成局部数据库上的一个有效的低级执行计划。转换主要包括两个方面:第一,产生输入查询的正确表示,以便执行计划能产生预期的结果;第二,对执行计划进行优化,即必须对代价函数最小化。对以上两个方面孤立成两个顺序的步骤:数据的局部化和全局化。数据局部化是将一个分布式数据库上的代数查询转换成一个等价的段查询,并通过代数转换来作进一步的简化。全局查询优化通过决策操作的顺序,结点间的数据

移动，以及数据库操作的分布和局部算法的选择来为输入的分段查询计划产生一个优化，其关键在于如何选择操作的执行顺序。选择查询操作的重点又在于连接操作的顺序，因为此类操作往往涉及多个节点上，操作尤为复杂。

3.3 索引管理

在云数据管理平台上，用户查询面临的是海量的数据，查询效率必然是一个亟待解决的问题。于是，受传统关系数据库的启发，我们可对云存储管理系统的数据建立索引，以提高用户查询的效率。那么，究竟选择何种索引就是一个关键问题。目前，针对云数据管理的索引已有一些研究工作

针对单维度的数据，我们可以建立 B+树索引，hash 索引等，文献[1][2][3]都是有关一维数据索引的研究，它们均在实际存储数据的物理节点上建立本地索引，然后在服务器端建立全局的索引，当用户查询数据时，通过服务器端的全局索引定位到本地索引，然后进行数据的查询，这样可以大大减少查询的时间。文献[1][2][3]也分别通过实验验证了论文中提出的算法的有效性。针对多维数据，我们可以建立 R 树，k-d 树等多维索引。文献[4][5]都是有关多维索引的研究。文献[4]采用 k-d 树建立本地索引，采用 R 树建立全局索引，它主要的应用于主-从体系结构中。文献[5]采用 R 树建立本地索引，采用 CAN 建立全局索引，它主要应用于点对点体系结构中。最后，二者都通过实验验证了文章中提出的索引算法的有效性。基于查询效率在海量数据处理中的关键地位，索引的研究仍将是一个值得研究的关键性问题。

3.4 multi-key 事务处理

随着 Web 应用数据规模的发展，可扩展的云数据管理系统在云计算研究中占有越来越重要的位置，而 key-value 存储方式是这种存储系统的一种主要架构，包括 BigTable, PNUTS, Dynamo 等等。在这些系统中，数据被存储成 key-value 对，但是对于大多数系统来讲原子性操作的粒度只能达到单个键值。这种特性在当前的很多 Web 应用中工作良好，但是在下一代 Web 应用中却力不从心：在线游戏，社会网络，合作编辑等等。因为这些应用需要在一组键值上进行读写操作，所以一致性的处理及事务原子性的保证要以一组键值为单位而不仅仅是单个键值。目前有一些系统可以支持 multi-key access，例如 Google 的 AppEngine Store，可以支持一组键值的事务处理，但是这种键值组是固定的，无法动态修改，无法满足当前的 Web 应用。一种可行的解决方案是在 key-value 存储层上增加一层键值组管理层，负责键值组的建立与撤销，同时管理键值组中所有数据的读写一致性[6]，在这种方案中，如何保证键值组管理的容错性和一致性将是研究的重点所在。

3.5 云数据管理的隐私保护

云计算已成为未来海量数据管理的必然趋势，从成本和性能两方面考虑，越来越多的企业更愿意把自己的数据中心从昂贵的高性能计算机转移到公有云或私有云中。然而，中小企业以及个人用户在使用云计算带来的便利的同时，也经受着隐私泄露的风险。业界越来越多的人认为，云计算中的隐私问题已经危及到云计算未来的发展。

云计算模型中一般有三个角色：数据提供者、查询用户以及云计算平台---云数据库系统。

数据提供方将数据提供给云数据库系统存储, 查询用户通过云数据库系统的平台对数据提供方的数据进行查询。对于云计算的各个参与者而言, 面向隐私保护的查询处理都是迫切需要解决的问题: 针对查询用户来说, 如果在查询处理中隐私保护机制不完善, 用户担心自己的隐私泄露, 将会尽量减少使用云计算服务; 针对数据提供者来说, 如果用户查询将暴露其它用户的隐私, 那么用户会对其服务可靠性产生质疑; 对于云数据库系统而言, 如果用户隐私遭到泄露, 用户可能放弃该公司的云计算服务。因此, 迫切的需要一种在云计算中的隐私保护查询处理技术, 既能保护用户的查询隐私、数据隐私以及三方交互隐私。

4 总结和展望

综上所述, 各种现象都表明, 云计算是真实存在的, 只是现阶段在云计算上的研究并不真实, 大有新瓶装旧酒的趋势。放眼当前形势, 中国应当架构符合我国实际情况的云计算平台, 探索新的应用模式, 从而展开云计算相关研究。我们认为, 这种新的应用模式应该是大数据和云计算相结合的产物。云计算扮演的角色是应用的基础, 它作为应用的载体存在, 而大数据才是真正产生价值的所在。如今的应用需求实际是对服务的需求, 我们如何为用户提供某种服务? 能够为用户提供什么样的服务? 这都需要数据来进行支撑, 没有数据服务就无从谈起。因此, 积累数据财富应该成为我们的习惯, 在不远的将来, 数据的积累也许会成为国力的标志。从这个角度讲, 对“数据思维”的研究可能会带来新的意义。在传统的探索中, 人们把精力集中于复杂算法的研究, 试图设计出复杂的算法来解决问题, 但是, 当有了数据, 也许配合了大规模的数据后, 原有的问题只需要非常简单的算法即可得到很好的解决。如今风风火火的 NoSQL 运动给无论是学术界还是产业界都带来了新的机遇, 人们对数据库的需求发生了改变, 人们不再需要一个大而全但自己只能用到一小部分的数据存储及管理工具, 这就带来了打破大一统数据管理现状的机会, 我们需要做的是抓住机遇, 研究出既满足实际需求, 又恰如其分的高效灵活的云数据管理系统。

参考文献

- [1] M . K . Aguilera, W .Golab , and M.A.Shah : A Practical Scalable Distributed B-Tree . (VLDB08)
- [2] S.Wu and K.-L. Wu : An Indexing Framework for Efficient Retrieval on the Cloud.
- [3] S. Wu, D. Jiang, B. C. Ooi , K. L. Wu: CG-Index: A Scalable B+-tree Based Indexing Scheme for Cloud Data Management Systems(PVLDB10)
- [4] X. Zhang, Z. Wang, J. Ai, J. Lu, X. Meng: An Efficient Multi-Dimensional Index for Cloud Data Management. In Proceedings of the CIKM Workshop on Cloud Data Management (CloudDB2009), November 2, 2009, Hong Kong, China.
- [5] J. Wang,S.Wu,H. Gao,J. Li, B. C. Ooi :Indexing Multi-dimensional Data in a Cloud System(sigmod10)
- [6] Sudipto Das Divyakant Agrawal Amr El Abbadi. G-Store: A Scalable Data Store for Transactional Multi key Access in the Cloud. SoCC'10
- [7] J. Hamilton. Cooperative expendable micro-slice servers (cems): Low cost, low power servers for internet-scale services. In Proc. of CIDR, 2009.

- [8] J. Dean and S. Ghemawat. MapReduce: Simplified Data Processing on Large Clusters. In OSDI, 2004.
- [9] Michael Stonebraker, Errors in Database Systems, Eventual Consistency, and the CAP Theorem,
<http://cacm.acm.org/blogs/blog-cacm/83396-errors-in-database-systems-eventual-consistency-and-the-cap-theorem/fulltext>