Report on the First International Workshop on Cloud Data Management (CloudDB 2009)

Xiaofeng Meng[1] Jiaheng Lu[1] Jie Qiu[2] Ying Chen[2] Haixun Wang[3]
{xfmeng,jiahenglu}@ruc.edu.cn, {qiujie, yingch}@cn.ibm.com, haixunw@microsoft.com
[1]School of Information and DEKE, MOE, Renmin University of China, Beijing China
[2]IBM Research - China
[3]Microsoft Research Asia, Beijing

1. INTRODUCTION

The first ACM international workshop on cloud data management was held in Hong Kong, China on November 6, 2009 and co-located with the ACM 18th Conference on Information and Knowledge Management (CIKM). The main objective of the workshop was to address the challenge of large data management based on cloud computing infrastructure. The workshop brings together researchers and practitioners in cloud computing and data-intensive system design, programming, parallel algorithms, data management, scientific applications and information-based applications interested in maximizing performance, reducing cost and enlarging the scale of their endeavors.

The workshop attracted 11 submissions from Asia, Canada, Europe and the United States, out of which the program committee finally accepted 5 full papers and 3 short papers. The accepted papers focused on cloud-based indexing and query processing, cloud platform availability, cloud replication and system development.

2. KEYNOTE PRESENTATION

The keynote speech, titled "EMC Decho: A Real World Use Case of Cloud Computing" was delivered by Bill Sun, Engineer Manager in EMC Center of Excellence China. He presented EMC's perspective about cloud computing, and shared some of their experience in building a reliable infrastructure to provide personal cloud services for millions of users. Bill Sun highlighted some of challenges they are facing in building robust and reliable infrastructures and described their potential remedies.

The first challenge is the increasing importance that personal information has to the eyes of the users. As a simple example, the first picture of one's newborn son is a digital artifact that a user is likely to want to preserve through one or several generations. The value of digital information to users has become such important that preserving one's digital data cannot be tied to a particular device or application. The information saved in these devices are more valuable than devices and should live much longer.

The second challenge is to search effectively in various devices. Current-day technologies for searching personal archives of digital data still have strong limitations. For example, how can you easily find your pictures in Las Vegas taken in July last year? How can you find a presentation you got from a colleague six months back about a particular project? Most of time, we have to organize the information by five "C": i.e. Context, Content, Calendar, Coordinates and Contacts. Effective management of such personal information with five "C" is a big challenge.

The major conclusion is that personal cloud is what they believe the most effective approach to address those challenges in personal information management, where all the information are saved in a secure well managed cloud storage system. Decho will work as the center or the hub to synchronize all users' information across multiple devices through the personal cloud, including PC, cell phone, or even NetBook.

3. RESEARCH PAPERS

The technical paper session consisted of eight presentations, whose main points are summarized next. Together, they give a glimpse to the exciting new developments spurred by data management in the cloud. These papers cover a variety of topics. We believe that these papers will provide researchers and developers with a brief glimpse into this exciting new technology, specifically from the perspective of cloud data management.

The paper entitled *Personalization as a Service: Architecture and Case Study* focuses on how to provide personalized services for individual users in the cloud environment. H. Guo, J. Chen, W. Wu and W. Wang first analyzed the main issues and challenges of using the traditional server-side user profiles for personalized services in the cloud. Then they presented the architecture of Personalization as a Service (PaaS) in which the client-side user modeling method is employed to support personalized cloud services. The

main idea is to decouple user modeling components from cloud services by observing the user's interactions on all of their cloud client devices collectively. As a result, user model can be shared across cloud services and used in a pay-as-you-go way. The client-side user modeling avoids the server overhead and provides unique user experiences with minimal user intervention. They finally give a case study of a personalized cloud search service solution according to the PaaS architecture.

X. Zhang, J. Ai, Z. Wang, J. Lu and X. Meng proposed an efficient approach to build a multi-dimensional index for cloud computing systems in the paper *"An Efficient Multi-Dimensional Index for Cloud Data Management"*. Their approaches can process typical multi-dimensional queries including point queries and range queries efficiently. Besides, frequent change of data on big amount of machines makes the index maintenance a challenging problem. To cope with this problem they proposed a cost estimation-based index update strategy that can effectively update the index structure. They describe experiments showing that their indexing techniques improve query efficiency by an order of magnitude compared with alternative approaches, and scale well with the size of the data. Their approach is quite general and independent from the underlying infrastructure and can be easily carried over for implementation on various cloud computing platforms.

The topic considered in *Packing the Most Onto Your Cloud* by A. Aboulnaga, Z. Wang and Z. Zhang is one particular optimization problem, namely scheduling sets of Map-Reduce jobs on a cluster of machines (a computing cloud). They present a scheduler that takes job characteristics into account and finds a schedule that minimizes the total completion time of the set of jobs. Their scheduler decides on the number of cluster nodes to assign to each job, and it tries to pack as many jobs on the machines as the machine resources can support. To enable flexible scheduling and packing of jobs onto machines, they run the Map-Reduce jobs in virtual machines, although their scheduling approach can be applied in any Map-Reduce scheduler. Their scheduling problem is formulated as a constrained optimization problem, and they experimentally demonstrate using the Hadoop open source Map-Reduce implementation that the solution to this problem results in benefits up to 30%.

Query Processing of Massive Trajectory Data based on MapReduce is addressed by Q. Ma, B. Yang, W. Qian and A. Zhou. Traditional trajectory data partitioning, indexing, and query processing technologies are extended so that they may fully utilize the highly parallel processing power of large-scale clusters. They also showed that the append-only scheme of MapReduce storage model can be a nice base for handling updates of moving objects. Preliminary experiments show that this framework scales well in terms of the size of trajectory data set. The limitation of traditional trajectory data processing techniques and their future research direction are also discussed.

The approach considered in *Leveraging a Scalable Row Store to Build a Distributed Text Index* by N. Li, J. Rao, E. Shekita and S. Tata is a distributed text index called HIndex, by judiciously exploiting the control layer of HBase, which is an open source implementation of Google's Bigtable. Such leverage enables them to inherit the good properties of availability, elasticity and load balancing in HBase. They also present the design, implementation, and a performance evaluation of HIndex.

F. Wang, J. Qiu, J. Yang, B. Dong, X. Li, and Y. Li proposed a metadata replication based solution to enable Hadoop high availability by removing single points of failures in Hadoop in the paper titled *Hadoop High Availability through Metadata Replication*. Single points of failures mean that the whole system becomes out of work due to the failure of critical nodes where only a single copy of data exists. The solution involves three major phases. In the initialization phase, each standby/slave node is registered to active/primary node and its initial metadata (such as version file and file system image) are caught up with those of active/primary node. In the replication phase, the runtime metadata (such as outstanding operations and lease states) for fail-over in future are replicated. Finally, in the fail-over phase, standby/new elected primary node takes over all communications. The solution presents several unique features for Hadoop, such as runtime configurable synchronization mode. The experiments demonstrate the feasibility and efficiency of their solution.

In the Paper entitled *How Replicated Data Management in the Cloud can benefit from a Data Grid Protocol - the Re:GRIDiT Approach*, L. Voicu and H. Schuldt developed, implemented and evaluated the Re:GRIDiT protocol for managing data in the grid. Re:GRIDiT provides support for concurrent access to replicated at different sites without any global component and supports the dynamic deployment of replicas. Since it has been designed independent from any underlying grid middle-ware, it can be seamlessly transferred to other environments like the cloud. They present the Re:GRIDiT protocol, show its applicability for cloud data management, and provide performance results of the evaluation of the protocol in realistic cloud settings.

The topic in *The Design of Distributed Real-time Video Analytic System* by T. Yu, B. Zhou, Q. Li, R. Liu, W. Wang and C. Chang is to propose a distributed scalable infrastructure VAP (Video Analytic Platform) for supporting real-time video stream analysis. In VAP, the application requirements are represented as a Directed Acyclic Graph (DAG), where nodes stand for video analysis computation modules and links show data flow and dependencies between nodes. VAP leverages UIMA (Unstructured Information Management Architecture) framework as the data flow control engine and multiple commodity databases as the storage and computation resources. The actual executions of video analysis computation modules have been pushed down into database engine to minimize the data movement cost.

4. CONCLUSION

CloudDB 2009 was the first CIKM-associated workshop addressing the challenges of large database services based on the cloud computing infrastructure. Whilst these emerging services have reduced the cost of data storage and delivery by several orders of magnitude, there is significant complexity involved in ensuring large data service can scale when one needs to ensure consistent and reliable operation under peak loads. Cloud-based environment has the technical requirement to manage data center virtualization, lower cost and boost reliability by consolidating systems on the cloud.

A first conclusion that can be drawn from this workshop is that the cloud systems should be geographically dispersed to reduce their vulnerability due to earthquakes and other catastrophes, which increase technical challenge on a great level of distributed data interpretability and mobility. Data interoperability is even more essential in the future as one component of a multi-faceted approach to many applications.

A final conclusion is that existing research works in the area of cloud-based data management are still somehow immature and significant room for progress exists. The works presented in the workshop mainly focused on adapting existing Grid and Map/Reduce techniques to the cloud environment. The participants agreed that many open challenges still remain such as cloud data security and the efficiency of query processing in the cloud. The participants also expressed interest in the organization of a conference dedicated to the issues raised by data management in the cloud.

5. ACKNOWLEDGEMENT

We would like to thank the program committee members, keynote speakers, authors and attendees, for making CloudDB 2009 a successful workshop. Jiaheng Lu was supported by 863 National High-Tech Research Plan of China (No: 2009AA01Z133). Finally, we also express our great appreciation for the support from Renmin University of China and IBM research-China.