

# Cloud Data Management



Liang Zhao • Sherif Sakr • Anna Liu  
Athman Bouguettaya

# Cloud Data Management

Foreword by Albert Y. Zomaya

 Springer

Liang Zhao  
NICTA Kensington  
NSW, Australia

Anna Liu  
NICTA, Eveleigh  
NSW, Australia

Sherif Sakr  
Software Systems Research Group  
NICTA, Eveleigh, NSW  
Australia

Faculty of Computers and Information  
Cairo University, Egypt

Athman Bouguettaya  
School of Computer Science  
and Information Technology  
RMIT University, Melbourne  
VIC, Australia

ISBN 978-3-319-04764-5 ISBN 978-3-319-04765-2 (eBook)

DOI 10.1007/978-3-319-04765-2

Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014932980

© Springer International Publishing Switzerland 2014

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media ([www.springer.com](http://www.springer.com))

*To my parents, Dianchi Zhao and Lijuan  
Chu, and my wife, Huan Wang, for their love,  
understanding, and encouragement.*

*Liang Zhao*

*To my wife, Radwa, and my lovely daughter,  
Jana, for their love, encouragement, and  
support.*

*Sherif Sakr*

*Lifelong gratitude to my family.*

*Anna Liu*

*To my wife, Malika, and lovely sons, Zakaria,  
Ayoub, and Mohamed-Islam.*

*Athman Bouguettaya*



# Foreword

The rapidly expanding generation of Internet-based services such as e-mail, blogging, social networking, search, and e-commerce has substantially redefined the behavior and trends of web users when it comes to creating, communicating, accessing content, sharing information, and purchasing products. Information technology professionals are witnessing a proliferation in the scale of the data generated and consumed because of the growth in the number of these systems; this ever increasing need for scalability and new application requirements has created new challenges for traditional relational database management systems (RDBMS). Currently, the apparent goal of the system and tool manufacturers is to facilitate the job of implementing every application as a distributed, scalable, and widely accessible service on the web (e.g., services from Facebook, Flickr, YouTube, Zoho, and LinkedIn).

Cloud computing technology is a relatively new model for hosting software applications. The cloud model simplifies the time-consuming processes of hardware provisioning, hardware purchasing, and software deployment; therefore it revolutionizes the way computational resources and services are commercialized and delivered to customers. In particular, it shifts the location of this infrastructure to the network in order to reduce the costs associated with the management of hardware and software resources. This means that the cloud represents the long-held dream of envisioning computing as a utility, a dream in which the economy of scale principles help to effectively drive down the cost of the computing infrastructure. In practice, cloud computing promises a number of advantages for the deployment of software applications such as pay-per-use cost model, short time to market, and the perception of (virtually) unlimited resources and infinite scalability.

The rise of the cloud technology has been somewhat disruptive. The advantages of the cloud computing model open up new avenues for deploying novel applications that were not economically feasible in a traditional enterprise infrastructure setting. Therefore, the cloud has become an increasingly popular platform for hosting software applications in a variety of domains such as e-retail, finance, news, and social networking. The proliferation in the number of applications also delivers a tremendous increase in the scale of the data generated and consumed by

these applications. This is why a cloud-hosted database system powering these applications forms a critical component in the software stack of these applications.

To meet the challenges posed by hosting databases on cloud computing environments there are a plethora of systems and approaches. This book is the first that approaches the challenges associated with hosting databases on cloud computing environments from different but integrated perspectives; it connects the dots. The authors deal with the problems that may be encountered in every cloud-based data hosting solution: NoSQL storage services, database-as-a-service (DaaS), virtualized database servers in addition to batch-based processing systems for big data. The book is useful for many database researchers or practitioners because the inherent change in hosting database in cloud environment is fundamental on many perspectives as it originates from new foundations and models of thinking.

I found the book to contain a lot of timely and useful information. The book has many gems that inspire the readers as they go through the different chapters which are covering an area that is currently changing the data management field in a fundamental way. It covers an impressive array of topics with great clarity that will excite any reader wishing to understand this emerging technology. It also provides extensive references which will help the interested reader find out more information about the discussed topics. All in all, this is an impressive piece of work and an invaluable source of knowledge for advanced students and researchers working in or wishing to explore this exciting field.

Darlington, NSW, Australia

Albert Y. Zomaya



# Preface

Cloud computing technology represents a new paradigm for the provisioning of computing resources. Cloud computing is with us and for the foreseeable future. This paradigm shift allows for the outsourcing of computing resources to reduce the ownership costs associated with the management of hardware and software. Cloud computing simplifies the time-consuming processes of hardware provisioning, hardware purchasing, and software deployment.

Cloud computing is not a passing trend but a stubborn reality that is rooted on an emerging trend leading computing into a technological quantum leap. It builds on decades of research in virtualization, autonomic computing, grid computing, and utility computing, and ubiquity of the web as the network and delivery medium.

Central to the success of cloud computing is the ability to provision data using different quality of service requirements, including latency, performance, and reliability. Unfortunately, most cloud providers do not guarantee, and let alone, provide information about actual quality of service for data access. This is a complex exercise that depends on many factors, including the location of the data store, type of data, network congestion and data store platforms.

This book fills a gap in that it provides an in-depth analysis of major data cloud platforms using an exhaustive series of tests and experiments to unlock the unanswered questions surrounding the performance of each cloud data platform that is considered. The work presented in this book focuses on evaluating cloud databases in the presence of very little information from cloud providers. This can also be interpreted as reverse-engineering the performance of cloud databases with its own risks in interpretation.

The data cloud platforms considered in this book include the leaders in the field, including, Amazon, Microsoft, and Google. Amazon offers a collection of services, called Amazon Web Services, which includes Amazon Elastic Compute Cloud (EC2) as cloud hosting server, offering infrastructure as a service and Amazon SimpleDB and Simple Storage Service (S3) as cloud databases.

Microsoft Azure is recognized as a combination of infrastructure as a service and platform as a service. It features web role and worker role for web hosting tasks and computing tasks, respectively. It also offers a variety of database options including

Windows Azure Table Storage and Windows Azure Blob Storage as the NoSQL database options and Azure SQL Database as the relational database option.

Google App Engine supports a platform as a service model, supporting programming languages including Python and Java and Google App Engine Datastore as a Bigtable-based, non-relational, and highly sharable cloud database.

We propose a performance evaluation framework of cloud platforms as a uniform testing environment for all the cloud data environments. We describe novel frameworks and architectures to address the following issues: (1) the performance characteristics of different cloud platforms, including cloud hosting servers and cloud databases, (2) availability and reliability characteristics that cloud platforms typically exhibit, (3) type of faults and errors that may be encountered when services are running on different cloud platforms under high request volume or high stress situations, (4) reasons behind the faults and errors, (5) the architecture internal insights that may be deduced from these observations, and (6) the software engineering challenges that developers and architects could face when using cloud platforms as their production environment for service delivery.

Kensington, NSW, Australia  
Eveleigh, NSW, Australia  
Eveleigh, NSW, Australia  
Melbourne, VIC, Australia

Liang Zhao  
Sherif Sakr  
Anna Liu  
Athman Bouguettaya

# Acknowledgements

I would like to thank my parents, Lijuan Chu and Dianchi Zhao, for their constant love, support, and encouragement during my Ph.D. study. I would also like to express my appreciation to my wife, Huan Wang, for her patience, understanding, and encouragement. The work would not have been possible without their support and care.

*Liang Zhao*

I would like to thank my parents, Ali Sakr and Amira Awad, for their encouragement and support. I want to thank my daughter, Jana, for the happiness and enjoyable moments she is always bringing to my life. My most special appreciation goes to my wife, Radwa Elshawi, for her everlasting support and deep love.

*Sherif Sakr*

I would like to sincerely thank the excellent hard work of my coauthors, without their dedication, commitment, and persistence, this book would not be here. I would also like to thank the wider NICTA Dependable Cloud Computing research team, who have provided wonderful companionship through our research journey. Lastly, many thanks go to our industry partners, who have provided much of the use inspiration for our work.

*Anna Liu*

I would like to thank my family for their love and understanding during my work on this book.

*Athman Bouguettaya*

The authors of this book would like to extend their sincere gratitude and appreciation to their collaborators for the contribution to this book; in particular, we would like to mention Alan Fekete, Jacky Keung, Kevin Lee, Hiroshi Wada, Xiwei Xu, Zhen Ye, Xiaofang Zhou, and Liming Zhu. Thank you all!



# Contents

- 1 Introduction** ..... 1
- 2 Cloud Computing**..... 9
  - 2.1 Definitions ..... 9
  - 2.2 Related Technologies for Cloud Computing ..... 11
  - 2.3 Cloud Service Models..... 14
  - 2.4 Cloud Deployment Models ..... 15
  - 2.5 Public Cloud Platforms: State-of-the-Art ..... 16
  - 2.6 Business Benefits of Cloud Computing ..... 19
- 3 Cloud-Hosted Data Storage Systems** ..... 21
  - 3.1 Introduction..... 21
  - 3.2 NoSQL Key Systems..... 23
  - 3.3 NoSQL Open Source Projects ..... 29
  - 3.4 Database-as-a-Service..... 32
  - 3.5 Virtualized Database Servers ..... 39
  - 3.6 Web Scale Data Management: Trade-Offs ..... 39
  - 3.7 Discussion and Conclusions ..... 42
- 4 Performance Evaluation Framework of Cloud Platforms** ..... 47
  - 4.1 The CARE Framework..... 48
  - 4.2 Application of CARE to Cloud Platform Evaluation ..... 54
  - 4.3 Experiment Results and Exception Analysis..... 55
  - 4.4 Discussion ..... 65
- 5 Database Replication of NoSQL Database-as-a-Service** ..... 67
  - 5.1 Architecture of Benchmark Application ..... 67
  - 5.2 Staleness of Data on Different Cloud Platforms ..... 69
  - 5.3 Trade-Off Analysis of Amazon SimpleDB ..... 74
  - 5.4 Discussion ..... 78

<b>6</b>	<b>Replicating Virtualized Database Servers</b> .....	81
6.1	Design of Benchmark Application .....	82
6.2	Implementation of Benchmark Application.....	86
6.3	Trade-Off Analysis of Virtualized Database Servers .....	87
6.4	Discussion .....	95
<b>7</b>	<b>SLA-Driven Database Replication on Virtualized Database Servers</b> .....	97
7.1	SLA Management for Virtualized Database Servers .....	98
7.2	Architecture of SLA Management Framework .....	101
7.3	Implementation of SLA Management Framework .....	105
7.4	Evaluation of SLA Management Framework .....	107
7.5	Provisioning the Database Tier Based on SLA of Transaction Response Times.....	114
7.6	Related work.....	117
7.7	Discussion .....	118
<b>8</b>	<b>QoS-Aware Service Compositions in Cloud Computing</b> .....	119
8.1	Preliminaries.....	120
8.2	QoS-Aware Service Composition in Cloud Computing .....	124
8.3	Experiment and Evaluation .....	128
8.4	Related Work .....	131
8.5	Conclusion.....	133
<b>9</b>	<b>Big Data Processing Systems</b> .....	135
9.1	Introduction.....	135
9.2	MapReduce Framework: Basic Architecture .....	137
9.3	Extensions and Enhancements of the MapReduce Framework ....	140
9.4	Systems of Declarative Interfaces for the MapReduce Framework.....	163
9.5	Conclusions.....	175
<b>10</b>	<b>Conclusions</b> .....	177
10.1	True Elasticity .....	177
10.2	Data Replication and Consistency Management.....	179
10.3	SLA Management .....	182
10.4	Transaction Support .....	184
10.5	Summary.....	186
	<b>References</b> .....	191

# List of Figures

Fig. 1.1	Characteristics of Big Data .....	2
Fig. 2.1	The evolution towards cloud computing in hosting software applications.....	11
Fig. 2.2	Exploitation of virtualization technology in the architecture of cloud computing .....	13
Fig. 2.3	The service models of cloud computing .....	14
Fig. 3.1	Database scalability options .....	22
Fig. 3.2	Sample BigTable structure .....	25
Fig. 3.3	PNUTS system architecture .....	27
Fig. 3.4	Partitioning and replication of keys in dynamo ring.....	29
Fig. 3.5	Basic GQL syntax .....	34
Fig. 3.6	Coexistence of multiple data management solution in one application .....	45
Fig. 4.1	Time measurement terminologies .....	49
Fig. 4.2	The flow chart of evaluation strategies .....	51
Fig. 4.3	Contract-First Web Service based client application .....	55
Fig. 4.4	The cumulative distribution function of high stress round-trip between the end-user and the Amazon EC2 cloud hosting servers .....	57
Fig. 4.5	The cumulative distribution function of high stress round-trip between the end-user and the Microsoft Windows Azure cloud hosting servers .....	57
Fig. 4.6	The cumulative distribution function of high stress round-trip between the end-user and the Google App Engine cloud hosting servers .....	58
Fig. 4.7	The average read time in cloud databases with low stress database read test set .....	58
Fig. 4.8	The average write time in cloud databases with low stress database write test set .....	59

Fig. 4.9	The cumulative distribution function of read throughput in cloud databases with low stress database read test set .....	60
Fig. 4.10	The cumulative distribution function of write throughput in cloud databases with low stress database write test set .....	60
Fig. 4.11	The cumulative distribution function of read and write throughput in cloud databases with high stress database read and write test sets .....	60
Fig. 4.12	The database processing time of read, write, and delete in cloud databases with low stress large file read, write, and delete test sets.....	61
Fig. 5.1	The architecture of NoSQL database as a service benchmark applications .....	68
Fig. 5.2	Probability of reading freshest value .....	70
Fig. 5.3	Time to see freshness with eventual consistent read.....	72
Fig. 5.4	The average, 95 percentile, and 99.9 percentile response time of reads at various levels of load .....	75
Fig. 5.5	Response time of reads at various read/write ratios on Amazon SimpleDB .....	76
Fig. 5.6	Response time of writes at various read/write ratios on Amazon SimpleDB .....	76
Fig. 5.7	Processed requests of Amazon SimpleDB .....	77
Fig. 5.8	Throughput percentage of Amazon SimpleDB .....	77
Fig. 5.9	Request failure rate of Amazon SimpleDB .....	78
Fig. 6.1	The architecture of relational database as a service benchmark application .....	83
Fig. 6.2	End-to-end throughput with 50/50 read/write ratio and 300 initial data size in the same zone .....	87
Fig. 6.3	End-to-end throughput with 50/50 read/write ratio and 300 initial data size in different zones .....	88
Fig. 6.4	End-to-end throughput with 50/50 read/write ratio and 300 initial data size in different regions .....	88
Fig. 6.5	End-to-end throughput with 80/20 read/write ratio and 600 initial data size in the same zone .....	89
Fig. 6.6	End-to-end throughput with 80/20 read/write ratio and 600 initial data size in different zones .....	89
Fig. 6.7	End-to-end throughput with 80/20 read/write ratio and 600 initial data size in different regions .....	90
Fig. 6.8	Average relative replication delay with 50/50 read/write ratio and 300 initial data size in the same zone .....	92
Fig. 6.9	Average relative replication delay with 50/50 read/write ratio and 300 initial data size in different zones .....	92



Fig. 6.10	Average relative replication delay with 50/50 read/write ratio and 300 initial data size in different regions .....	93
Fig. 6.11	Average relative replication delay with 80/20 read/write ratio and 600 initial data size in the same zone .....	93
Fig. 6.12	Average relative replication delay with 80/20 read/write ratio and 600 initial data size in different zones .....	94
Fig. 6.13	Average relative replication delay with 80/20 read/write ratio and 600 initial data size in different regions .....	94
Fig. 7.1	SLA parties in cloud environments.....	99
Fig. 7.2	The SLA management framework architecture.....	102
Fig. 7.3	The implementation of the SLA management framework in the setup of experiments.....	106
Fig. 7.4	The performance of the replication delay for fixed 3 replica servers with the framework disabled.....	110
Fig. 7.5	The performance of the replication delay for fixed 6 replica servers with the framework disabled.....	110
Fig. 7.6	The performance of the replication delay for up to 6 replica servers with the framework enabled, $delay_{tolerance} = 1,000$ ms, and $intvl_{mon} = 60$ s.....	111
Fig. 7.7	The performance of the replication delay for up to 6 replica servers with the framework enabled, $delay_{tolerance} = 1,000$ ms, and $intvl_{mon} = 120$ s .....	111
Fig. 7.8	The performance of the replication delay for up to 6 replica servers with the framework enabled, $delay_{tolerance} = 1,000$ ms, and $intvl_{mon} = 240$ s .....	112
Fig. 7.9	The performance of the replication delay for up to 6 replica servers with the framework enabled, $delay_{tolerance} = 1,000$ ms, and $intvl_{mon} = 480$ s .....	112
Fig. 7.10	The performance of the replication delay for up to 6 replica servers with the framework enabled, $delay_{tolerance} = 500$ ms, and $intvl_{mon} = 120$ s .....	113
Fig. 7.11	The performance of the replication delay for up to 6 replica servers with the framework enabled, $delay_{tolerance} = 2,000$ ms, and $intvl_{mon} = 120$ s .....	113
Fig. 7.12	The performance of the replication delay for up to 6 replica servers with the framework enabled, $delay_{tolerance} = 4,000$ ms, and $intvl_{mon} = 120$ s .....	114
Fig. 7.13	Comparison of SLA-based vs resource-based database provisioning rules. (a) Workload: 80/20 (r/w). (b) Workload: 50/50 (r/w) .....	116
Fig. 8.1	Cloud system .....	120
Fig. 8.2	Control flows .....	121
Fig. 8.3	Data flow graphs .....	122
Fig. 8.4	Aggregation functions for each QoS attribute .....	123

Fig. 8.5	Composition solution .....	124
Fig. 8.6	Crossover and mutation operators .....	126
Fig. 8.7	Example of scheduling string. (a) Example 1. (b) Example 2 for data forwarding .....	127
Fig. 8.8	Fitness vs Dataset .....	130
Fig. 8.9	GA vs Random selection .....	130
Fig. 8.10	Time vs concrete services.....	131
Fig. 8.11	GA vs Integer programming approach. (a) GA vs Integer programming on time. (b) GA vs Integer programming on fitness .....	132
Fig. 9.1	An example MapReduce program.....	137
Fig. 9.2	An overview of the flow of execution a MapReduce operation ....	139
Fig. 9.3	Decision tree for choosing between various join strategies on the MapReduce framework .....	142
Fig. 9.4	An overview of the Map-Reduce-Merge framework .....	142
Fig. 9.5	A sample execution of the Map-Reduce-Merge framework .....	144
Fig. 9.6	An overview of HaLoop architecture .....	146
Fig. 9.7	An example structure of <i>CFile</i> .....	151
Fig. 9.8	An example structure of <i>RCFile</i> .....	153
Fig. 9.9	Example file colocation in CoHadoop .....	155
Fig. 9.10	An example Sawzall program .....	164
Fig. 9.11	An example SQL query and its equivalent Pig Latin program ....	165
Fig. 9.12	Pig compilation and execution steps .....	166
Fig. 9.13	An example HiveQL query .....	167
Fig. 9.14	Basic syntax of SQL/MR query function .....	170
Fig. 9.15	The architecture of HadoopDB .....	173
Fig. 9.16	Sample Jaql script .....	174
Fig. 9.17	Jaql system architecture.....	175

# List of Tables

Table 2.1	Feature similarities and differences between related technologies and cloud computing .....	12
Table 2.2	Summary of cloud deployment models.....	17
Table 3.1	Design decisions of various web scale data management systems .....	41
Table 4.1	Building a test set.....	53
Table 4.2	Total error detail analysis .....	63
Table 4.3	Average error (rates) of high stress database read over different time periods .....	64
Table 4.4	Average error (rates) of high stress database write over different time periods .....	64
Table 5.1	Probability of reading freshest value .....	71
Table 5.2	Successive eventual consistent reads .....	72
Table 7.1	The effect of the adaptive SLA management framework on the end-to-end system throughput .....	108
Table 7.2	Number of provisioned database replicas .....	117
Table 10.1	Open research challenges of cloud-hosted database systems.....	187