SpringerBriefs in Computer Science

SpringerBriefs present concise summaries of cutting-edge research and practical applications across a wide spectrum of fields. Featuring compact volumes of 50 to 125 pages, the series covers a range of content from professional to academic.

Typical topics might include:

- A timely report of state-of-the art analytical techniques
- A bridge between new research results, as published in journal articles, and a contextual literature review
- A snapshot of a hot or emerging topic
- An in-depth case study or clinical example
- A presentation of core concepts that students must understand in order to make independent contributions

Briefs allow authors to present their ideas and readers to absorb them with minimal time investment. Briefs will be published as part of Springer's eBook collection, with millions of users worldwide. In addition, Briefs will be available for individual print and electronic purchase. Briefs are characterized by fast, global electronic dissemination, standard publishing contracts, easy-to-use manuscript preparation and formatting guidelines, and expedited production schedules. We aim for publication 8–12 weeks after acceptance. Both solicited and unsolicited manuscripts are considered for publication in this series.

More information about this series at http://www.springer.com/series/10028

Gaurav Baranwal · Dinesh Kumar
Zahid Raza · Deo Prakash Vidyarthi

Auction Based Resource Provisioning in Cloud Computing

Springer

Gaurav Baranwal
Department of Computer Science
Institute of Science, Banaras Hindu University
Varanasi, Uttar Pradesh, India

Dinesh Kumar
School of Computer and Systems Sciences
Jawaharlal Nehru University
New Delhi, India

Zahid Raza
School of Computer and Systems Sciences
Jawaharlal Nehru University
New Delhi, India

Deo Prakash Vidyarthi
School of Computer and Systems Sciences
Jawaharlal Nehru University
New Delhi, India

ISSN 2191-5768 ISSN 2191-5776 (electronic)
SpringerBriefs in Computer Science
ISBN 978-981-10-8736-3 ISBN 978-981-10-8737-0 (eBook)
https://doi.org/10.1007/978-981-10-8737-0

Library of Congress Control Number: 2018944359

Printed on acid-free paper

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Dedicated to
All who are contributing in the just allocation of the resources

Foreword

It gives me a pleasure to write a foreword for this book and gives me an honour to congratulate the authors of this book for a job well done. This is a timely presentation, and the scope of the topics covered in the book provides the reader with up-to-the-moment resources in the area of Cloud resource provisioning.

Cloud computing, a business model of computing, has revolutionized the world by its enormous benefits. Cloud computing resources such as CPU, memory, storage, network are available on demand, and anyone can avail these services on a subscription basis. Cloud computing is a business model of computing unlike grid and cluster, so pricing of the Cloud resources can act as a lever to control the wastage of the computing resources. At present, most of the Cloud services are being offered using static pricing. The drawback with static pricing is its non-compliance of demand and supply principle of economics resulting in an inefficient control of the utilization of computing resources. Dynamic pricing is obviously a better solution for this problem.

Amazon proved that auction results in the efficient allocation of resources compared to static pricing. It is beneficial to both Cloud providers and Cloud customers. But, to my best knowledge, besides Amazon, no provider is offering resources using auction mechanism. This area needs proper attention from professionals and academicians to identify the suitable reasons for this and offer remedies. Recently, Google took a step forward and started offering virtual machines using dynamic pricing on a trial basis. I have proposed resource provisioning models that apply auction. Professor Vidyarthi is also an active researcher and with his group members has proposed a good number of models that apply auction. What is lacking is an organized literature that motivates researchers in the field of Cloud computing to contribute in this area further.

This book introduces auction mechanism for resource provisioning in Cloud computing. The aim of this book is to provide a structured literature of auction based on Cloud market to attract academicians and professionals to contribute in this area. Since various models have also been proposed, it will help readers to formulate the Cloud resource allocation problem using auction. This book also discusses challenges in detail, and the reader will be able to understand them.

In fact, various ingredients of auction-based Cloud market have been combined at one place.

It is to emphasize that auction-based Cloud market needs expertise from both computer scientists and economists. It provides a good opportunity for interdisciplinary research. The developed services must reach the society; otherwise, it would be a failure. This book will act as a bridge between computer scientists with specialization in Cloud computing and economists with specialization in resource allocation using auction.

Melbourne, Australia

Dr. Rajkumar Buyya
Redmond Barry Distinguished Professor
Director, Cloud Computing and Distributed
Systems (CLOUDS) Lab, School of Computing
and Information Systems
The University of Melbourne

Preface

Cloud computing, a business model of computing, has revolutionized the world by its enormous benefits. Computing resources such as CPU, memory, storage are available in the form of services online, and one can avail these services on payment. In 2006, Amazon launched various types of Cloud services and named it Elastic Compute Cloud (EC2). This step accelerated the well acceptance of the Cloud services and established a whole new market of the computing resources. A good number of service providers joined the Cloud bandwagon and started offering various types of Cloud services in order to attract the customers. Of all the early service providers, a common attribute was the pricing of these services as it was based on the principle of static pricing. Various services were available in the Cloud, but at fixed rates. These service providers increased their infrastructure to cater more number of Cloud customers. A common problem, encountered by most of these providers, was less number of requests at off-peak period which kept most of the infrastructure idle. Though virtualization is the main ingredient for creating Cloud services in order to best utilize the resources, even to run virtual machines of low power a physical machine with substantive power is needed. Various professionals argued that scheduling policies can be designed to handle the challenge of wastage of computing resources, but it was observed that scheduling alone would not suffice. Unlike grid and cluster, an important characteristic of Cloud computing is its business model of computing in which pricing of the resources acts as a lever to control the wastage of the computing resources. Thus, the main drawback with static pricing was its non-adherence to the demand and supply principle of economics, making it difficult to control the better utilization of the computing resources. Dynamic pricing is naturally a better option for the pricing of such resources.

Although researchers across have suggested various approaches to implement dynamic pricing, perhaps industries are not ready to accept so. It is because of the ease of static pricing; it is very simple in nature and more understandable to the customers. In 2009, Amazon took a bold step and launched a type of instance (i.e. virtual machine) which would be available using spot mechanism, a type of auction. Cloud customers were smart enough to understand this, as reduction in price is a

motivating factor. Since then, Amazon has created a glorious history of spot instances (read on Amazon EC2 on huge reduction in price paid by Cloud customers using spot instances). The crucial issue with spot instance, which makes it more reasonable and cheaper for Cloud customer, is non-guarantee of resources as Amazon may take back the allocated spot instances whenever there is a shortfall of the resources assigned using static pricing. This issue warrants smartness from the customers in terms of fault tolerance schemes.

Amazon proved that auction has the ability of effective allocation of resources as compared to static pricing. It is beneficial to both Cloud providers and Cloud customers. Unfortunately, except Amazon to our best knowledge, no other provider offers the Cloud resources using auction mechanism. Therefore, proper attention is warranted from the Cloud professionals and academicians to identify the appropriate reasons for this and suggest best remedies for the promotion of applying auction for the Cloud resources. Recently, Google started offering pre-emptive virtual machines similar to Amazon spot instances, though on a trial basis.

While making a survey on resource allocation, using auction in Cloud computing, it is observed that very few researchers have contributed significantly in this area of research. What majorly lacks is well-organized structured literature that can motivate the researchers, working in the area of Cloud computing for more significant contribution.

This book is intended to offer an organized literature that deliberates on auction mechanism for resource provisioning and pricing in Cloud computing. Chapter 1 introduces the fundamental knowledge of Cloud computing that will help naïve in the field of Cloud computing to cultivate sufficient background for understanding Cloud computing, Cloud market, and how pricing acts as a lever, i.e. how economical approaches can be applied for the allocation of computing resources and most importantly why auction is significant in Cloud computing. Chapter 2 provides a comprehensive study of auction, available in the literature, and describes its various types to be applied in the different market scenarios. As literature reports only few types of auction in Cloud computing, this chapter generates sufficient motivation to identify different Cloud market scenarios in which various types of auction can be used. First two chapters are the prerequisites for those readers who are novice in Cloud computing and therefore may be skipped by those equipped with sufficient background in Cloud and auction. Auction has generally been categorized into three classes based on the role of customer and provider: forward auction, reverse auction and double auction. A detailed description of forward auction in Cloud computing has been provided in Chap. 3, of reverse auction in Chap. 4 and of double auction in Chap. 5. These three Chaps. 3, 4 and 5, discuss the advantages and disadvantages of the corresponding auction types along with the detailed framework and a model that provides sufficient background to the

researchers in order to understand and formulate resource allocation problem using auction under various Cloud market scenarios. Chapters 3, 4 and 5 also enlist open research issues to be considered as a potential research problem for the researchers. Chapter 6 summarizes and concludes this brief.

Varanasi, India — Gaurav Baranwal
New Delhi, India — Dinesh Kumar
New Delhi, India — Zahid Raza
New Delhi, India — Deo Prakash Vidyarthi

Contents

About the Authors

Gaurav Baranwal is an assistant professor at the Department of Computer Science, Institute of Science, Banaras Hindu University, Varanasi, UP, India. Before, he served as an assistant professor in Madan Mohan Malviya University of Technology, Gorakhpur, India. He completed his M.Tech. and Ph.D. in computer science from the School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi, India. His research interests include resource provisioning and service coordination in Cloud computing.

Dinesh Kumar is a Ph.D. student at the School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi, India. He completed his M.Tech. in Computer Science at the School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi, India. His research interests include resource provisioning and pricing in Cloud computing.

Zahid Raza is an associate professor at the School of Computer and Systems Sciences, Jawaharlal Nehru University (JNU), New Delhi, India. He received his Ph.D. in computer science from Jawaharlal Nehru University, New Delhi, India. Prior to joining JNU, he served at Banasthali Vidyapith University, Rajasthan, India. His research interests include parallel and distributed systems, evolutionary algorithms, Cloud computing and IoT.

Deo Prakash Vidyarthi is a professor at the School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi. He has published over 90 research papers in various peer-reviewed international journals and transactions (including IEEE, Elsevier, Springer, Wiley, World Scientific) and over 50 research papers in the proceedings of various conferences in India and abroad. He has also contributed many chapters in related disciplines. He has two research monographs to his credit and has also contributed chapters in many edited books. He serves on the editorial boards of many international journals and is a senior member of IEEE;

International Association of Computer Science and Information Technology (IACSIT), Singapore; International Society of Research in Science and Technology (ISRST), USA; and International Association of Engineers (IAENG). His research interests include parallel and distributed systems, grid and Cloud computing, mobile computing and evolutionary computing.

Acronyms

BB	Budget-balance
BO	Bidder's optimality
C-BIC	Cloud-bayesian incentive compatible
C-DSIC	Cloud-dominant strategy incentive compatible
C-OPT	Cloud-optimal
CA	Combinatorial auction
CCIF	Cloud computing interoperability forum
CDA	Continuous double auction
CDARA	Combinatorial double auction resource allocation
CFP	Call for proposal
CMO	Cell membrane optimization
CRC	Cloud resource consumer
CRP	Cloud resource provider
DMAA	Double multi-attribute auction
EE	Economic efficiency
FCC	Federal communications commission
FMCDAM	Fair multi-attribute combinatorial double auction model
GVA	Generalized vickrey auction
IaaS	Infrastructure as a service
IC	Incentive compatible
ICT	Information and communication technologies
IDC	International data corporation
ILP	Integer linear programming
IR	Individual rationality
JADE	Java agent development framework
LP	Linear programming
MA	Multi-attribute
MIP	Mixed-integer programming
MVA	Modified vickrey auction
MVO	Mean variance optimization

ND	Non-dominance
NIST	National institute of standards and technology
OVF	Open virtualization format
PaaS	Platform as a service
PFA	Paddy field algorithm
PIP	Packing integer programming
PMDA	Preston-mcAfee double auction
PTAS	Polynomial time approximation scheme
QoS	Quality of service
RBMA	Reverse batch making auction
SaaS	Software as a service
SAA	Simultaneously ascending auction
SLA	Service-level agreement
SME	Small and medium enterprises
SVM	Support vector machine
SW	Social welfare
TCRA	Truthful combinatorial reverse auction
TDACC	Truthful multi-unit double auction for cloud computing
TPDA	Threshold price double auction
VCG	Vickrey–clarke–groves
VM	Virtual machine
VMPAC	VM provisioning and allocation problem for cloud computing
WDP	Winner determination problem