J. Parallel Distrib. Comput. 74 (2014) 3087–3097

Contents lists available at ScienceDirect

J. Parallel Distrib. Comput.

journal homepage: www.elsevier.com/locate/jpdc

Towards efficient and fair resource trading in community-based cloud computing☆

Han Zhao, Xinxin Liu, Xiaolin Li *

Scalable Software Systems Laboratory, University of Florida, Gainesville, FL 32608, USA

HIGHLIGHTS

- We create a distributed market-oriented framework for cloud resource trading.
- We propose a multiagent technique for budget-free optimal resource allocation.
- We propose a new directed hypergraph model for budget-limited resource trading.
- We develop effective heuristic-based protocols for budget-limited resource trading.

ARTICLE INFO

Article history:
Received 16 June 2013
Received in revised form
9 July 2014
Accepted 29 July 2014
Available online 7 August 2014

Keywords:
Cloud computing
Resource trading
Social welfare
Envy-free
Hypergraph

ABSTRACT

In this paper, we investigate the resource trading problem in a community-based cloud computing setting where multiple tenants communicate in a peer-to-peer (P2P) fashion. Enabling resource trading in a community cloud unleashes the untapped cloud resources, thus presents a flexible solution for managing resource allocation. However, finding an efficient and fair resource allocation is challenging mainly due to the heterogeneity of tenants. Our work first develops a market-oriented model to support resource negotiation and trading. Based on this model, we adopt a multiagent-based technique that allows a group of autonomous tenants to reach an efficient and fair resource allocation. Further, when budget constraint presents, we propose a directed hypergraph model to facilitate resource trading amongst heterogeneous tenants. We analyze the application of the directed hypergraph model to trading decision making, and design a series of heuristic-based resource trading protocols for both budget-unaware and budget-aware scenarios. The performances of the proposed protocols are validated through simulations. The results are in tune with the theoretical analysis and provide insights into practical application issues.

Published by Elsevier Inc.

1. Introduction

The ever-increasing demand for computing power motivates recent advances in multi-core processing and high-speed networking, and drives the emergence of distributed computing platforms that span a variety of heterogeneous devices across the Internet. With the aid of hardware virtualization, these platforms are capable of offering "rent-on-demand" resource leasing services that let end users instantly access a vast pool of resources, known as the "cloud computing" model. Current cloud computing model is mostly vendor driven, with users having no control over the data or the technology supported by the cloud. Such a vendor-driven model, although convenient to use, also brings many issues to light, e.g., failure of monocultures, tradeoff between convenience and control, and concerns about environmental impact [6]. To address these issues, researchers have proposed an alternative model that provides a collaborative resource sharing platform called community cloud [18,19,27]. Different from the centralized vendor model, community-based cloud computing leverages under-utilized networked private resources for infrastructure support. Tenants within the same community cloud typically share common security and compliance concerns, and may delegate management to some trusted third-party organization.

Similar to the centralized vendor-driven model, the community-based model provides computation and storage resources as metered services. Therefore, the design goal of the community cloud should not only focus on the quality of computing service, but

☆ The work presented in this paper is supported in part by National Science Foundation (grants ACI 1245880, ACI 1229576, CCF-1128805, OCI-0904938, and CNS-0709329).

* Corresponding author.
E-mail addresses: han@cise.ufl.edu (H. Zhao), xinxin@cise.ufl.edu (X. Liu), andyli@ece.ufl.edu, andyli.ece@gmail.com (X. Li).
URL: http://www.andyli.ece.ufl.edu/ (X. Li).

http://dx.doi.org/10.1016/j.jpdc.2014.07.005
0743-7315/Published by Elsevier Inc.

should equally address the economic aspect such that tenants receive cost-effective cloud service provisioning. While managing resource allocation is relatively straightforward in the centralized vendor-driven model (e.g., Amazon's on-demand and spot instance pricing), it is particularly challenging due to the heterogeneity in the multitenancy environment. In a community cloud, we are facing a free market where tenants are only incentivized to accept profitable resource exchange. As a result, a well designed multitenancy resource trading protocol is highly desirable to effectively regulate the management of resource allocation.

In this paper, we study the distributed resource trading problem in a community-based cloud computing environment, and propose a set of multitenancy resource trading protocols to jointly optimize resource allocation efficiency and fairness. Specifically, better efficiency refers to the increased aggregate valuations of all the tenants, and better fairness is interpreted as reduced envy between every pairwise combination of tenants. Our solution follows a market-oriented design principle, and uses a directed hypergraph model to integrate these two seemingly conflicting design objectives into one unified resource trading framework. It directly extends Chevaleyre et al.'s work [11], and further addresses the challenge of budget limited resource trading. With systematic analysis of the resource trading market, a set of heuristic-based distributed resource trading protocols are developed and evaluated.

The comprehensive study presented in this paper has broad utility in the growing world of "everything-as-a-service", because it characterizes the extent to which independent and self-interested tenants interact with each other. Our analysis shows that incentive preserving resource exchanges tend to benefit the system, both from a global view of the overall service efficiency and from a local view of the improved service quality valuation. Moreover, the proposed resource trading approaches are complementary to the vendor-driven cloud computing. For example, consider user Alice rents a virtual machine from Amazon with reserved instance pricing. After Alice finishes her job and before the lease expires, Alice might "sublease" this virtual machine to user Bob in order to partially compensate for her resource rental cost.

The contribution of this paper is primarily four-folded, and is summarized as follows.

- Create a distributed market-oriented framework for modeling the multitenancy resource trading problem in community cloud.
- Leverage a multiagent-based technique to solve the optimal resource allocation for distributed budget-free resource trading.
- Propose a novel directed hypergraph model to facilitate the analysis of budget constrained resource trading.
- Develop effective heuristic-based protocols for resource trading given the presence of budget limitation.

The reset of the paper is organized as follows. Section 2 presents an overview of the related work. In Section 3, we describe the problem setting and formulate the resource allocation objectives. In Section 4, we introduce a multiagent-based technique to achieve optimal resource trading efficiency and fairness. Section 5 further investigates allocation strategies with limited budget. We propose a novel directed hypergraph model and develop a series of distributed resource trading protocols based on heuristic approaches. Section 6 shows simulation results and analyzes their implications. Finally, in Section 7 we summarize our solution and conclude the paper.

2. Related work

This paper presents distributed protocol design to jointly optimize resource trading efficiency and fairness. As the organization of distributed resource evolves towards a more hierarchical architecture [23], distributed algorithms designed for solving combinatorial multi-criteria optimization problems become more attractive. Common optimization techniques include machine learning [30], evolutionary algorithms [14], swarm intelligence [29], and socialeconomy approaches [31,24]. All these approaches share a common flavor that involves interacting entities evolving towards the optimal solution (by following certain learning or negotiation rules). Our proposed approach falls into the category of socialeconomy approaches. They are built based on the observation that resource management in distributed systems shares common features with commodity allocations driven by market power in the economic study. It is widely adopted to create a computational economy for grid computing [1,7] and the emerging cloud computing [8,34]. In an early study, Wolski et al. [37] presented two different market strategies for controlling resource allocation, namely commodities markets and auction. The commodities markets strategy treats disparate resources as interchangeable commodities, while auction requires orchestration from a centralized auctioneer for collecting bids and determining winners. Our proposed resource trading framework is designed for a community cloud environment, and belongs to the commodities market category. In particular, we propose a P2P resource trading market for managing cloud resource allocation. Example research related to this notion includes [13,36]. In [13], a P2P data replication system was proposed to improve fault-tolerance of digital collections in library. In [36], the authors proposed a multiple currency economy that any peer can issue its own currency. Different from their design, peers directly exchange resources in our distributed resource trading design.

In this paper, two economic metrics are used to quantify the quality of an allocation: efficiency in terms of overall social welfare, and fairness in terms of envy-freeness. The metric of efficiency is important to characterize the achievable system performance, and was studied in a number of publications [20,39,4]. On the other hand, the metric of fairness highlights individual's utility such that each individual achieves the maximum contentment of its allocated share [16]. Compared to efficiency, the envy-free fairness has generally received far less attentions. A related work targeting grid computing was found in [32]. Using game theory, the authors tackled with a multicriteria optimization problem with the aid of axiomatic theory of equity. The authors concluded that for fair and feasible scheduling on global scale computational grid, a strong community control is required. The research conducted in this paper approaches the multicriteria optimization problem from a different angle, and further investigates how to balance the two metrics amongst budget-aware distributed tenants.

Our proposed protocols build on a directed hypergraph model. A hypergraph is an extension of the graph concept that one edge (called a hyperedge) can connect an arbitrary set of vertices rather than two. A hypergraph model is flexible and informative to use in algorithm design as it generalizes the graph. For that reason, it becomes attractive to improve algorithm performance in various research domains, e.g., page reputation computation for search engines [3], cellular mobile communication [33] and memory management [22]. For large-scale scientific computing, Çatalyürek and Aykanat [10] proposed a multilevel partitioning approach for mapping repeated sparse matrix–vector computations to multicomputers using hypergraph. Their approach significantly reduces communication overheads while achieving drastically improved mapping results. In their hypergraph model, hyperedges represent affinity among subsets of the data, and the weights reflect the strength of this affinity. We model the resource trading problem in a similar manner that aims to optimize the aggregate weights of the directed hypergraph model.

3. A distributed resource trading framework for community cloud

This section presents the design overview of a distributed resource trading framework for the community cloud. In Section 3.1,