

A First Look at Problems in the Cloud

*Theophilus Benson**, *Sambit Sahu†*, *Aditya Akella**, and *Anees Shaikh†*
*University of Wisconsin - Madison, †IBM Research

Abstract

Cloud computing provides a revolutionary model for the deployment of enterprise applications and Web services alike. In this new model, cloud users save on the cost of purchasing and managing base infrastructure, while the cloud providers save on the cost of maintaining underutilized CPU, memory, and network resources. In migrating to this new model, users face a variety of issues. Commercial clouds provide several support models to aide users in resolving the reported issues

This paper arises from our quest to understand how to design IaaS support models for more efficient user troubleshooting. Using a data driven approach, we start our exploration into this issue with an investigation into the problems encountered by users and the methods utilized by the cloud support's staff to resolve these problems. We examine message threads appearing in the forum of a large IaaS provider over a 3 year period. We argue that the lessons derived from this study point to a set of principles that future IaaS offerings can implement to provide users with a more efficient support model. This data driven approach enables us to propose a set of principles that are pertinent to the experiences of users and that we believe could vastly improve the SLA observed by the users.

1 Introduction

The emergence and growing popularity of cloud computing signals an evolution in the way IT infrastructure and services are delivered and consumed. Clouds have a number of essential characteristics, such as self-service, on-demand consumption, location independence, and rapid elasticity. There are also evolving cloud delivery models, including Infrastructure-as-a-Service (IaaS), Platform-as-a-Service, and Software-as-a-Service, which provide IT infrastructure, application development services, and software applications from the cloud, respectively [6].

The IaaS delivery model provides users with the ability to easily acquire and release infrastructure resources, such as servers, storage, and network bandwidth on-demand using simple Web-based tools with an accompanying pricing model in which users pay for only what they use. The IaaS model of cloud computing is best exemplified in services such as Amazon's Elastic Compute Cloud (EC2) [1].

The high degree of virtualization and resource abstraction afforded by IaaS clouds also comes with a new set of challenges in terms of users' ability to efficiently diagnose and resolve problems with their applications and virtual resources in the cloud. Compared to traditional IT models, cloud users have limited visibility into the underlying infrastructure. When users experience problems, they have little alternative except trial-and-error troubleshooting or searching support forums for similar symptoms. These approaches may eventually be successful, but can take several hours to days to resolve problems.

In this paper, we aim to develop an understanding of the nature of problems experienced by customers of an IaaS cloud, along with their experience in resolving these problems using the forum-based support that is common in compute cloud services. Our study is based on actual user problems and experiences as captured from the open support forum of a large IaaS cloud provider. We examine message threads appearing in the forum over a 3-year period and develop a taxonomy of the problem classes. Using a text analytics-based approach we are able to automatically classify over 9,575 support forum message threads into problem clusters which we further assign to several high-level problem classes.

We found that, with the exception of problems related to application-level issues, the observed problems are roughly evenly divided among the remaining problem classes (e.g., connectivity, virtual image management, performance, etc.). In studying the evolution of problems over time, we find that some classes of problems are

closely related to the introduction of new cloud features as users incorporate them into their deployments. Other problems, such as those related to connectivity, are persistent and are relatively less affected by new functionality. We see evidence that some problem classes (e.g., related to image management) diminish significantly as cloud providers introduce new interfaces and tooling to simplify certain operations. The level of involvement from the cloud operator in solving problems also is dependent on the problem class, and changes over time. Some classes, such as those problems related to virtual infrastructure components, require operator involvement 50% of the time. While operator involvement in solving problems does generally decrease over time, some problems consistently require help from the cloud operator.

In addition to the classification of problems and involvement from administrators, we also use the data to build a sketch of the support model used in the cloud, in terms of staffing and resolution times. The cloud provider grows the number of administrators as the number of customers posting in the forums grows, and adjusts the number of operators participating in the forum based on peak periods. We discovered that while problems can require 20–120 hours to resolve, the operators respond to the initial problem post within 10-20 hours.

Our study is an important first step in developing a view of how to improve the problem management experience for cloud computing customers. Some of the opportunities we identify from our observations include:

Proactive support enhancements: When new features are introduced into the cloud, we observe an increase in the number of reported problems – cloud providers should anticipate this and provide training to their operators and release new debugging tools for the specific feature.

Automation of operator tasks: Providers cannot simply hire new support staff in response to user growth – many problems will likely be solved by a few seasoned administrators. Instead the providers should focus on providing mature tools for automating operator problem determination tasks.

Improved information exchange: In the current model, information is exchanged back and forth between the user and provider through the open forum – this constricted information channel can significantly inflate problem resolution times. An automated way to collect relevant information from virtual instances could short-circuit much of the administrator-customer exchange in the forum.

These are examples of mechanisms that can make problem-solving more efficient for both customers and cloud providers. We plan to pursue these and other opportunities in our future work to demonstrate how the cloud self-service paradigm can be extended beyond in-

frastructure acquisition to problem management.

2 Background and Motivation

In a typical commercial cloud, the cloud provider is mainly responsible for problems associated with its own infrastructure. The provider monitors its physical resources such as servers, storage and network systems to provide reasonably stable resources up to the hypervisor level. While the cloud provider tries to ensure a highly available infrastructure, it typically does not provide guarantees on individual instance availability. Users should expect that the provided virtual resources may become unavailable at times, thus requiring users to restart their application instances on a new server.

In addition to the unexpected outages, users of the cloud can experience problems for several different reasons, including invalid assumptions about the operating environment, poor isolation between users, hardware degradation, and misconfiguration of software. To help users deal with these problems, commercial clouds expose a few different support options, the most common ones being a free best effort options offering no SLA guarantees and a premium commercial grade options that provide users with SLA guarantees.

Best Effort Support Model: The most common version of this support model is the “user forum”, a forum through which users can receive help on diagnosing problems both from cloud operators as well as other users. With the user forum, the provider makes no guarantee on the response time of the operators and no guarantee on the resolution time of the problem.

Premium Support Model: In the premium support model, the cloud provider guarantees the user that problems will be resolved within a certain period of time. The SLA provided to a user is inversely proportional to the price paid by the users; a higher price demands a guaranteed for a shorter resolution time. Certain providers also ensure that users have access to dedicated operators who are familiar with the user’s environment and needs.

Our goal is to understand the properties of the problems encountered by users, and to further understand how well the best effort support model tackles and resolves the reported problems. Our ultimate goal is to derive lessons that can help build better troubleshooting support.

3 Data & Methodology

In this section, we provide an overview of the dataset used to drive our study on the problems that cloud users experience. To understand the problems encountered, we analyze message threads, or “problem tickets”,

posted in the user forum for a prominent IaaS provider. The provider uses this forum to facilitate troubleshooting. The forum spans three years (Aug '06 to Dec '09) and contains over 9575 message threads. Several of the threads contain descriptive information about the problem users faced; we analyze such threads in this work. Each thread starts with symptoms the user observed. This is followed up by suggestions (by operators or other users) for debugging actions to perform and a description of the results of these actions. Finally, the threads end with the resolution of the problem and an explanation of the root cause either by an operator or the user.

To analyze the message threads, we borrowed techniques from Information Retrieval to cluster the tickets based on the data they contain. We then examined trends common to each of the discovered clusters, thereby unearthing various “classes of common problems”. We describe our approach next.

3.1 Extraction of Problem Clusters

Our first goal is to automatically cluster message threads into a small number of “problem classes” based on the nature of the underlying problems that users faced. There are several challenges in deriving such clusters. First, the number of tickets is large – hence the grouping must be performed automatically. Second, the number and nature of clusters is not known to us beforehand. Third, the tickets are specified using (unstructured) English text which makes automatic clustering hard.

To address these challenges, we leverage Information Retrieval algorithms for imposing structure on, and extracting clusters from, unstructured text. We evaluated several such algorithms (e.g., [4]) on our dataset and we found that the Lemur IR package [5] provided us with the best accuracy on several small random test sets of problem tickets. Briefly, the Lemur tool supports indexing of each document as a set of descriptive words; Lemur supports clustering operations on this index using the k -means algorithm. To avoid using generic words as descriptions of the problem ticket and to obtain meaningful clusters, Lemur, like most other IR approaches, uses a variety of tricks such as eliminating words that match a “stop word” list (including frequently used nouns, pronouns and verbs), and word standardization (eliminating tenses and plurals) etc.

From the 9575 message tickets, the Lemur tool discovered 194 clusters. We limit our attention to 24 of these clusters, each of which had at least 50 problems; Furthermore, we eliminated 4 clusters pertaining to non-technical questions such as billing queries, questions about future releases, and feature requests. Roughly 91% of all problems reported mapped to the 20 clusters we studied. To understand these 20 clusters, we generated a

“summary” of each cluster, which consists of the top 20 words (in terms of the frequency counts) in the cluster. These summaries help us analyze the problems reported.

A key limitation of our study is that the range of problems we examine is limited to problems reported by users through the forum. In particular, it excludes problems that are reported directly to IaaS provider by customers with the premium service. While it is difficult to quantify the impact of this limitation on our study, we do believe that our preliminary evaluation sheds light on the most common problems faced by a typical user of IaaS clouds.

4 Problems Faced by Users

In this section, we analyze the problems users faced based on the clusters of problem tickets derived by the aforementioned approach. We start by grouping the 20 clusters into dominant higher-level *problem classes*. We then dig deeper within each class to answer questions pertaining to the prevalence and evolution of problems observed over time and across categories, and the level of assistance needed and offered. Our goal is to develop an understanding of the nature of problems that can guide the design of appropriate support mechanisms that decrease problem resolution times.

Our key observations are: (i) Users encounter many problems in trying to setup their instance and to keep the instance running. (ii) Of the many type of problems faced by the users, we observe that those related to managing virtual resources and instance performance grow over time. In addition, we find that these two types of problems require the most involvement from cloud administrators because users are ill-equipped with the appropriate tools to debug these two classes. (iii) The addition of new features results in a temporary increase in the number of problems reported – the number of problems subside as user become familiar and the provider perfects the method of delivery for these features. (iv) Our investigation of the cloud support model shows that the support staff grows in proportion to the increasing number of customers posting in the forum. Administrators usually respond to posts within 10-12 hours, but problem resolution can take days.

4.1 Problem Categories

We manually inspected the summaries for the 20 mined clusters mentioned above to further group these into 5 logical problem classes based on their similarity – either in being related to the same set of functional components, or having similar problem semantics.

We assign each cluster to one of the following classes: *Application-related*, *Virtual Infrastructure-related*, *Im-*

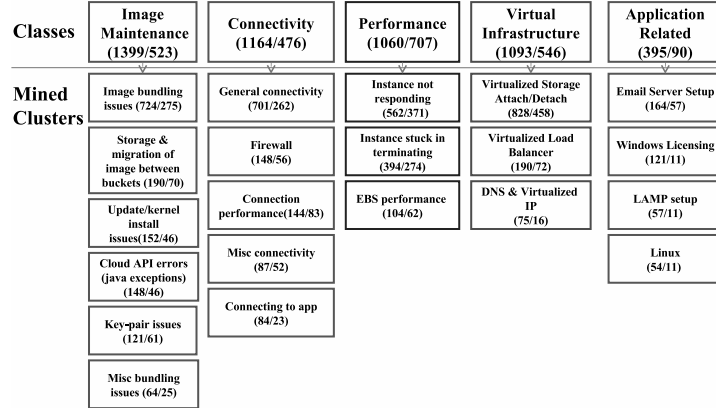


Figure 1: A tree depicting the taxonomy of the problem classes

age Management-related, Performance-related, and Connectivity-related.

Figure 1 illustrates the logical grouping of the mined clusters into these 5 classes of problems. With each cluster under a class, we include two numbers: The first number inside each cluster denotes the total number of corresponding problem threads, and the second number denotes problems that required assistance from the cloud operator to find a resolution. For example, there are 828 message threads that are related to storage volume attach and detach, and out of these, 458 problem threads required operator involvement.

From this, we observe that users may be afflicted by a variety of problems spanning these five categories. In a significant fraction of cases, operator involvement was needed to resolve the problem.

We now examine the relative prevalence of problems across the categories. The top boxes in Figure 1 illustrates the breakdown of the 5 problem classes. We find that 93% or 4716 out of the 5111 problem threads are roughly equally split between the four classes of Performance, Image Maintenance, Connectivity, and Virtualization-related problems. Only a small number of problem threads are Application-related. One possible explanation for the relatively low frequency of application problems is that they are related to higher-layer configuration settings, licensing issues, and installation problems – these issues are typically out of scope for an IaaS provider, and are better dealt with by the application’s support staff.

4.2 Evolution of Problems

We now turn our attention to the evolution of user problems over time. Specifically, we study how the relative frequency of different categories of problems changes over time, and whether any specific events contribute to

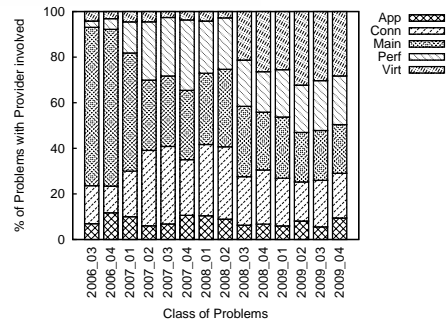


Figure 2: Percentage of problems discovered in each class over time.

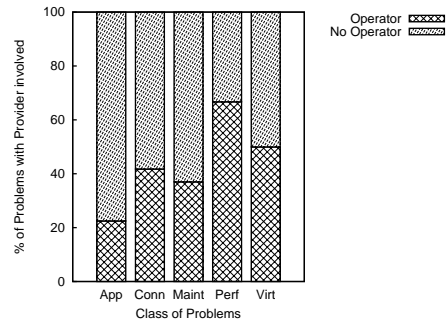


Figure 3: Percentage of problems in each class that required operator intervention.

any of the observed changes. This analysis also provides some insight into problem classes that have persisted over time, and hence require additional effort to help resolve them.

Figure 2 illustrates the evolution of different classes of problems between the 3rd quarter of 2006 and 4th quarter of 2009, where the number of tickets in each class is normalized by the total number of problems in the corresponding quarter. We observe several interesting trends:

- Problems in the *Virtual Infrastructure* category in-

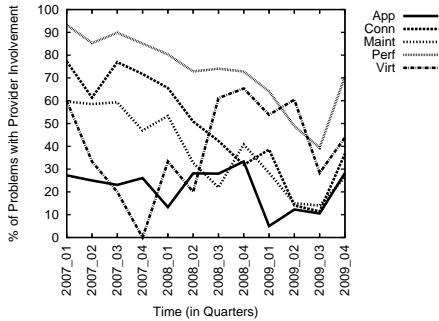


Figure 4: Fraction of problems in each class which required cloud operator intervention over time.

crease over time

- Problems in the *Maintenance* category decrease significantly with time
- *Connectivity* problems are relatively *stable* and persistent

We examined the feature release history for the cloud provider and found that sharp increases in problems in the Virtualized Infrastructure class are correlated with the release of new features in the cloud. For example, we observed that the increase in Virtual Infrastructure problems in the 3rd quarter of 2008 coincided with the introduction of a new virtual storage service. Following this release, another infrastructure-related feature was released in the 2nd quarter of 2009 and followed by another significant increase in the set of Virtual Infrastructure problems. To further verify these observations, we went back to the problem threads in the Virtual Infrastructure class and saw that they were largely related to the new features. Similarly, with the Connectivity and Performance categories, we were able to also correlate certain increases with the release of new features or modification of existing features.

The significant decrease in the Image Maintenance problem class is explained by the fact that better APIs and tooling were released over time to better manage the images. For example, automatic image capture and reboot tooling was made available in 1st quarter of 2007.

4.3 Problems with Operator Involvement

We now examine which problems needed operator involvement for problem resolution, and to what extent. Figure 3 illustrates the percentage of problems in each problem class that required operator involvement for resolution. Note that operator involvement ranges between 20% to 60% of the problems within a class.

Among the 5 classes, we find that Virtualized Infrastructure and Performance require operator involvement

at least 50% of the time. We examined the clusters for these two classes of problems and found that, under Virtualized Infrastructure, attaching and detaching storage volumes requires the greatest degree of operator involvement. Similarly, we find that “unresponsive instance” was the major cause of operator involvement in the Performance problem class. This observation is significant in that users do not have enough information or control to identify and resolve these problem scenarios. For example, without any internal access or information about cloud resources, it is not possible to determine why an instance is unresponsive. Likewise without the ability to inspect or change the internal state of the virtualized storage, users are constrained to API calls which are ineffective in changing their storage volume state.

We also analyze operator involvement in various classes as a function of time. Our intent is to see if there is evidence that users are able to gradually understand and diagnose certain problems, and if there are some problems that are always difficult to self-diagnose due to lack of visibility into the cloud (arising from virtualization). From Figure 4, we find that in general operator interventions decrease over time with the exception of the Virtual Infrastructure class. Although intervention decreases for rest of the categories, it never disappears altogether and gradually settles to a stable level. This suggests that while users become more familiar with the system and accumulate a knowledge-base of solutions, there are a significant fraction of problems that persistently require provider involvement to resolve.

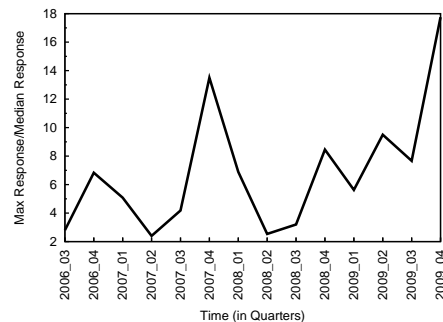


Figure 5: Max thread closed divided by median threads closed by an operator.

4.4 Cloud Support Model

For traditional IT service providers, the support model consists of a staffing plan, which includes the number of support staff and their skills, hours of availability, and their locations. In addition, the model includes targets such as the time to respond to reported problems of different severity levels, and the time to resolve problems. In this section, we analyze the support forum messages

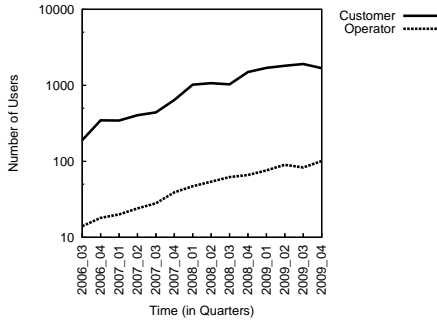


Figure 6: Number of cloud operators and cloud customers on the forums.

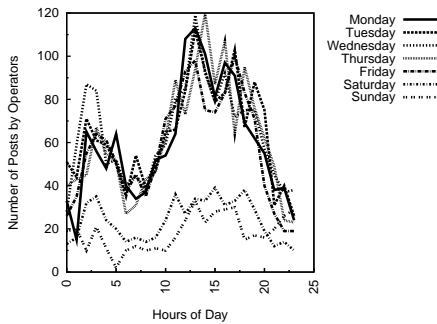


Figure 7: Number of forum post by cloud operators.

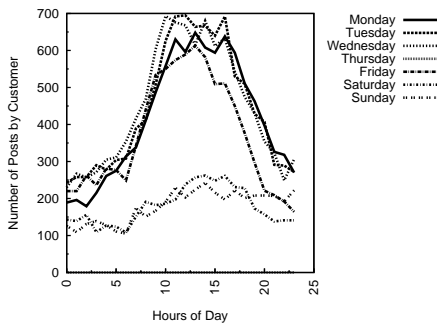


Figure 8: Number of forum post by cloud customers.

to see what can be observed about the support model for IaaS cloud providers.

We first examine the number of administrators responding to forum threads, based on their unique usernames across all messages. We observed that over the full 3-year period, 166 administrators participated in the message threads. A small number of administrators (around 10) were involved in over 150 message threads, while 100 answered fewer than 20 threads. The degree to which there are a few dominant administrators may be an indication of the skills distribution in the staffing plan. To investigate further, we examined the ratio of the maximum number of threads answered to the median over the observation period. Figure 5 shows this on a quarterly basis, and we see that in some periods there are administrators who participate in many more threads than the

median, but there is significant variability in the amount of skew in the forum participation.

In Figure 6, we can see the evolution of the number of customers and administrators active in the forum over time. The size of the support team grows proportionally with the number of customers posting in the forums. The number of administrators remains roughly an order of magnitude smaller than customers over the lifetime of the cloud we observed.

In Figure 7, we examine the number of messages posted by the support staff over the course of 24 hours for each day of the week. The cloud provider’s support team is active 24x7 with the main peak of activity roughly matching the peak times of customer postings (Figure 8). The support team activity on weekends (Saturday and Sunday) is relatively low, likely due to a smaller staffing. Interestingly, there is also a smaller peak of administrator activity in the early morning hours (around 2am PST), with no corresponding peak in customer posting activity. This may indicate that a global support team is being used, with most administrators located in a timezone appropriate for peak North America business hours, and a smaller number deployed to staff the forum during off-peak weekday hours. The number of unique administrator usernames observed during each hourly period also indicates that there are more operators online during this secondary peak period.

We also examined the amount of time taken for resolution of support threads, estimated as the time difference between the first and last timestamp of the messages in a thread. If the last message appears well after the problem was actually resolved, this would overestimate the resolution time. If the problem is never resolved, for example, this would be an underestimate of the resolution time. Nevertheless, the statistics still provide useful estimates of the resolution time. In Figure 9, we show the CDF of the resolution time across all collected threads, and find that about 60% of the problems are resolved in 20 hours while the next 20% of the threads can take as much as an additional 100 hours.

We also considered the initial response time for administrators to post an answer in Figure 10. We find that administrators respond to 60% of the problems in less than 9 hours while for the next 20%, administrators may take as long as an additional 20 hours to respond. We observe that problems seem to be resolved within 11-110 hours of the administrator’s first response. It is likely that much of the time after the first response from an administrator is spent in an iterative trial and error process as customers explore possible root causes.

In trying to understand the support model for this large cloud provider, we find that although the number of support staff increased over time and forums are manned 24x7, there is still considerable variability in problem

resolution times. Users still must often engage in lengthy exchanges to solve their problems, sometimes lasting several days.

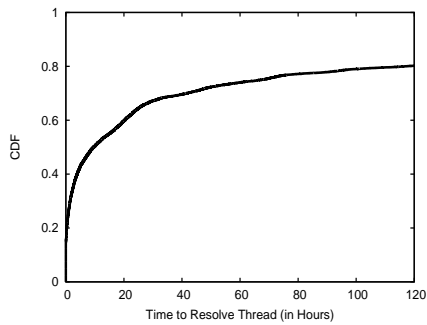


Figure 9: Problem Resolution Time.

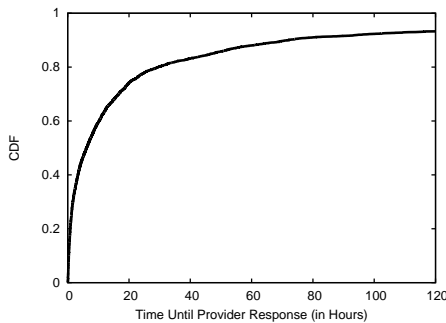


Figure 10: Time for the operator to respond to the first user query.

5 Related Work

The idea of cloud computing has been around for several decades in the form of utility computing; however, the lack of mature virtualization tools and powerful processors has prevented its growth.

Recent advancements in both the virtualization and the processor fields have created an environment for the deployment of cloud computing. Although relatively new, a fair amount of work [3, 8] has been done to examine current and future challenges for both users and provider of cloud computing. However, little has been done to understand the range of operational challenges faced by users as they attempt to run applications within the cloud.

In particular, work on cloud management [2, 7] has focused on the provisioning and scaling of services within infrastructure clouds. Unlike prior work, we believe that users face a significant challenge in merely trying to keep the instance up and running. To this end we studied the frequency of problems encountered by users, the amount

of time required to debug these problems, and the properties of the support model required to resolve these problems.

6 Conclusion

IaaS clouds provide a variety of support models to aide users in debugging the problems that they encounter. In this paper, we conducted a study of the problems encountered by users and examined the effectiveness and the efficiency of the most popular support model, namely best effort, in resolving these problems. Our goal was to understand the missing mechanisms that can be added to allow cloud providers to offer support in a more effective fashion.

We found that of the problems faced by users, performance and virtualized problems are the most persistent and prevalent problems owing in part due to the fact that users have no visibility into the cloud and are thus forced to consult the cloud operators for help. In examining how the best effort support model handles these problems, we discovered that 10 operators are responsible for resolving most problems and that a significant delay of 20-110 hours exists between the initial operator involvement and the problem resolution. Our measurements indicate that to offer more effective support, clouds should (1) take a proactive approach by developing tool targetted at debugging new features, (2) develop tools to automate operator task, and (3) provide a vehicle to gather and transfer information between operator and user.

7 Acknowledgements

This work is supported in part by an NSF FIND grant (CNS-0626889), an NSF CAREER Award (CNS-0746531), an NSF NetSE grant (CNS-0905134), and by grants from the UW-Madison Graduate School. Theophilus Benson is supported by an IBM PhD Fellowship.

References

- [1] Amazon ec2. <http://aws.amazon.com/ec2/>.
- [2] 3Tera. <http://3Tera.com/>.
- [3] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia. Above the clouds: A berkeley view of cloud computing. Technical Report UCB/ECS-2009-28, Feb 2009.
- [4] S. Cunningham and G. Holmes. Developing innovative applications of machine learning. In *Proc Southeast Asia Regional Computer Confederation Conference*, Singapore, 1999.
- [5] K. Fitz, L. Haken, and B. Holloway. Lemur – A Tool for Timbre Manipulation. In *Proceedings of the International Computer Music Conf.*, 1995.
- [6] P. Mell and T. Grance. Draft NIST working definition of cloud computing, June 2009.
- [7] rightscale. <http://rightscale.com/>.
- [8] L. M. Vaquero, L. R. Merino, J. Caceres, and M. Lindner. A break in the clouds: towards a cloud definition. *SIGCOMM CCR*, (1):50–55, 2009.