Emerging Issues & Challenges in Cloud Computing—A Hybrid Approach

Yaser Ghanam, Jennifer Ferreira, Frank Maurer

Department of Computer Science, University of Calgary, Calgary, Canada.
Email: yghanam@gmail.com

Received October 16th, 2012; revised November 17th, 2012; accepted November 26th, 2012

ABSTRACT

As cloud computing continues to gain more momentum in the IT industry, more issues and challenges are being reported by academics and practitioners. In this paper, we aim to attain an understanding of the types of issues and challenges that have been emerging over the past five years and identify gaps between the focus of the literature and what practitioners deem important. A systematic literature review as well as interviews with experts have been conducted to answer our research questions. Our findings suggest that researchers have been mainly focusing on issues related to security and privacy, infrastructure, and data management. Interoperability across different service providers has also been an active area of research. Despite the significant overlap between the topics being discussed in the literature and the issues raised by the practitioners, our findings show that some issues and challenges that practitioners consider important are understudied such as software related issues, and challenges pertaining to learning fast-evolving technologies.

Keywords: Cloud Computing; Issues; Challenges; Systematic Literature Review

1. Introduction

With unprecedented adoption in industry over the past few years, cloud computing continues to be one of the most vital and fast-growing models in IT. Cloud computing is “a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction [1].” Cloud computing is based upon a service-based architecture wherein services are provided mainly at the infrastructure level (e.g., virtual machines, storage) platform level (e.g., database, web server), or software level (e.g., email, ERP solution).

Despite the widespread adoption of cloud computing, researchers and practitioners have been actively reporting issues and challenges with this new technology. Some of the challenges seem to be fundamental such as issues with privacy and security. Other challenges such as sub-optimal performance and limited bandwidth are a natural result of pushing the boundaries of this new model to achieve more. The goal of our research is to gain an understanding of the type of issues and challenges that have been emerging over the past five years. In this paper, we aim to answer three research questions:

- RQ1. What issues and challenges with cloud computing have researches been focusing on over the past five years?
- RQ2. How has the interest in different types of issues and challenges evolved over the past five years?
- RQ3. Are there any gaps between the issues and challenges researchers have been focusing on and the issues and challenges practitioners deem to be important?

In order to answer these questions, we conducted our research in two parts. The first part is a systematic literature review based on the guidelines proposed by Kitchenham [2]. The review focuses on published peer-reviewed papers that explicitly discuss issues and challenges with cloud computing. In the second part, we conducted in-depth interviews with experts in the field of cloud computing to solicit their opinions pertaining to issues and challenges.

In the following sections, we elaborate on each study separately (including the detailed research methodology, data collection strategy, results, and analysis). Then, we compare and integrate the findings of the two studies. After that, we discuss the limitations and threats to validity concerning each of our studies. And finally, we draw our conclusions and discuss future work.

2. Systematic Literature Review

2.1. Search Process

The search process covered journal articles and confer-
ence papers available in four major electronic databases namely: ACM Digital Library, IEEE Explorer, SpringerLink, and ScienceDirect. These databases were selected because they are known for including the proceedings of key conferences and journals in the area of computer science and engineering. We believe that these databases include a representative sample of the literature produced in the subject matter as pertinent to this research. Since we were only interested in recent articles, we limited our search to articles published in the year 2007 or later.

We constructed the following search query to look for articles that specifically focus on discussing issues, challenges or problems in cloud computing:

(Cloud AND (issue OR issues OR challenge OR challenges OR problem OR problems)).

We searched the four abovementioned databases for articles whose titles matched the search query. We used Google Scholar to conduct our queries because we found the results returned by Google Scholar to be more reliable and accurate than most other search engines in the selected electronic databases. The search was conducted on August 28, 2012; therefore, results that were indexed after this date have not been included in this study.

To reproduce the raw set of results, the following steps are to be followed:

1) The following query should be entered into Google Scholar search engine: allintitle: issue OR issues OR challenge OR challenges OR problem OR problems “cloud”.
2) The date range has to be set to: 2007-2012;
3) The “published in” field in the advanced search box has to be set to: IEEE OR Springer OR ACM OR ScienceDirect.

2.2. Inclusion and Exclusion Criteria

We included peer-reviewed journal articles and conference papers that clearly focus on discussing issues, challenges or problems in the cloud computing domain. Research papers and articles were eligible for inclusion in this review if they discussed or addressed issues and challenges facing cloud computing in general. Articles that discussed challenges in a specific application domain were not included unless the article presented the challenges in a fashion that could be generalized to other domains. Detailed results on excluded articles will be provided in the results section. Because of the fast evolution of cloud computing over the past decade, this review does not include research published before the year 2007. Non-English contributions were also excluded from this review.

2.3. Data Collection

For each of the papers included in our review, we have extracted the following pieces of information:

- The title;
- The type of contribution;
- Year of publication;
- The authors’ countries;
- The issues and challenges discussed in the contribution;
- The general topic under which the issues and challenges could be categorized.

The researchers manually extracted the data from Google Scholar into a shared Excel sheet. After listing the meta-data of the papers as described above, the researchers went through the papers one by one to extract the pieces of information that are not readily available from the title such as the issues and challenges being discussed, and the topic under which they could be categorized.

2.4. Data Filtering & Analysis

The raw set of results was filtered using a multi-step filtering process to account for the inclusion and exclusion criteria. The different steps and their outcomes will be explained in the results section. All the data was collected and formatted in a tabulated fashion to allow for analysis and charting. The following analyses were conducted:

- The topics that were addressed in the literature and the percentage of publications that address each of the identified topics;
- The number of publications in each year over the time period of interest;
- The change in focus of publications over the time period of interest;
- The number of publications coming from different countries;
- The focus of publications coming from different countries.

2.5. Results

The initial set of results included 185 items. The list went through a filtering process as shown in Table 1. The outcome of this process was a set of 110 items to be considered in the analysis. The final set is available in the references section from [3] to [112].

Taking into consideration the subjectivity of our selection of data sources, construction of the query as well as the definition and application of the inclusion and exclusion criteria, we do not claim that the papers listed are the only relevant papers that were published in 2007 onwards. However, we consider this to be a representative sample of the papers published during that period of time.
Table 1. Data filtering.

<table>
<thead>
<tr>
<th>Filtering step</th>
<th>Data analyzed</th>
<th>Excluded results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exclude results that do not represent a journal article or a conference paper (e.g. book, citation, talk, abstract, panel, workshop schedule).</td>
<td>Title</td>
<td>37</td>
</tr>
<tr>
<td>Exclude non-relevant results that came up in the result set due to:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>˗ Ambiguity in the terms used in the search string (e.g. “issue” as in “special issue”, “cloud” as in “the Land of the Long White Cloud”).</td>
<td>Title</td>
<td>16</td>
</tr>
<tr>
<td>˗ The word “challenges” coexisting with the term “cloud” in the same title but without being associated with it (e.g. “Using the Cloud to Facilitate Global Software Development Challenges”).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exclude non-English results</td>
<td>Title, abstract, and full content</td>
<td>1</td>
</tr>
<tr>
<td>Exclude application-specific papers (e.g. “Performance issues in cloud computing for cyber-physical applications.”)</td>
<td>Title and abstract (plus 11 borderline)</td>
<td>14</td>
</tr>
<tr>
<td>Exclude borderline papers if they do not discuss issues and challenges in a fashion that could be generalized to other domains</td>
<td>Full content</td>
<td>7</td>
</tr>
<tr>
<td>Total results excluded</td>
<td></td>
<td>75</td>
</tr>
<tr>
<td>Remaining results</td>
<td></td>
<td>110</td>
</tr>
</tbody>
</table>

2.6. Discussion

In this section, we present our analysis of the results obtained through the systematic literature review as described earlier in the paper. We conduct each part of the analysis in light of our research questions.

RQ1. What issues and challenges with cloud computing have researches been focusing on?

Looking at the collected data, we have identified 10 categories of issues that have been the focus of research in the past five years, in addition to one category of miscellaneous topics:

Security & privacy. This category includes organizational and technical issues related to keeping cloud services at an acceptable level of information security and data privacy. This includes ensuring security and privacy of sensitive data held by banks, medical and research facilities [33]. Security and privacy issues become even more serious when governmental institutions use the cloud [32]. Despite the known need for Service Level Agreements [62] between Cloud service providers and users, standards for safety have not yet been established [64] and more research in this area would be beneficial [67]. Security and privacy of data spans issues such as authentication [62], encryption [58], and detection of malware, side channel attacks [103] and other kinds of attacks [12]—both internal and external to an enterprise [65]. There exists current research on detection and handling of security breaches [36] to guard against tampering, loss and theft of data [98]. Further, fault tolerant mechanisms for backing up data [75],[104] are required when there are failures in the infrastructure, such as network outages.

Other issues, especially in public clouds, include secure virtualization through effective firewalls, VM isolation [59], and detection of reconnaissance scans [29]. There are security issues with long-term storage correctness [103] and migrating from one vendor to another based on changing needs, i.e., the problem of vendor lock-in [87].

Current research has demonstrated the shortcomings of frameworks such as XML Signature, which highlight the important issue of browser security [11],[105]. Integrity, confidentiality, and non-repudiation of data can be threatened due to the cloud being a multi-tenancy environment [30],[106]. Solutions that segregate user data, manage identity, and governance and regulatory compliance have been investigated to address this [101]. Information leakage resulting from poor information-flow isolation requires robust solutions [31].

As the cloud is used in more contexts, there are specific issues that come with the new domains, such as using cloud services in high-speed rail [66]. Further, research on privacy and security in the cloud has highlighted issues when the cloud is integrated with pervasive systems [60], sensor networks [74] and grid computing [102]. As use of the cloud is scaled up to larger and larger systems, such as cloud networks [61], it will become increasingly important to find effective solutions for the security and privacy challenges already highlighted at a larger scale.

Infrastructure. This category entails issues pertaining to the hardware layer used as a backbone for cloud services as well as the thin layer of software used to operate this hardware. The main issue that dominates this category is performance including topics like SaaS placement problems [23],[94], server allocation optimization [49],
load balancing [48] and many other [46]. Other issues are related to networking such as traffic management [25] 77,81][111], ubiquitous connectivity [20], network speed and cost [79], and network reliability [78]. Another group of challenges pertain to resource management including dynamic resource provisioning [18], scaling [42], and allocation [55]; as well as resource stranding and fragmentation [1].

Furthermore, sustainability stands out as another important issue given the amount of energy needed to operate large-scale hardware infrastructure [42]. Quality attributes of the hardware infrastructure have also been an area of interest including issues like availability [27, 91,92], reliability [109], and scalability [7,96]. Other issues under this category include infrastructure design issues [13][18] and virtualization[15,88,89].

**Data management.** As cloud computing is enabling more data-intensive applications at the extreme scale, the demand is increasing for effective data management systems [16]. One main topic in this category is data storage [19][43] and all the issues that come with it such as data federation (i.e. storage across different providers) [42], data segmentation and recovery [28], data resiliency [109], data fragmentation and duplication [18], and data backup [75]. Other issues include data retrieval [77] and processing [5], data provenance [17], data anonymization [50], and data placement (across different data centers) [1].

**Interoperability.** Most research on issues and challenges with cloud computing recognize interoperability as a major adoption barrier because of the risk of a vendor lock-in [110][87]. Amongst the many problems being discussed are: the lack of standard interfaces [19][42] and open APIs [38], and the lack of open standards for VM formats [13] and service deployment interfaces [91]. These issues result in integration difficulties between services obtained from different cloud providers as well as between cloud resources and internal legacy systems [38].

**Legal issues.** The notion of using cloud resources as a utility has brought about a number of legal issues. The most discussed issue in the literature we surveyed is related to data placement [78]. Laws and regulations vary widely across different regions and jurisdictions as to where and how data should be stored, processed, and used [26],[39],[76]. For example, the European Union requires that all personal data be physically stored within the jurisdictions of the European Union [52]. Also, compliance requirements might vary in regards to the disclosure of data in general and sensitive data in specific (e.g. financial data, health insurance records) [13], in addition to variations in the regulations around transaction logging and taxation [52]. Traceability of data and alterations made on it has also been reported as a legal concern [26][28]. Another important issue is the lack of comprehensive legislation on liability in the cloud [110] as well as identity definition (i.e. users versus systems), and issues related to authentication and authorization [112].

**Economic challenges.** This category includes issues related to the cost-benefit aspect of the cloud from a financial point of view. Some research is focused on producing cost models that reflect more accurately the actual cost of building and operating a cloud [38]. For cloud providers, the cost of the hardware infrastructure and the administrative costs associated with it are key to understanding the economic viability and sustainability of the business [42]. Cloud providers also need to work on effective monetization strategies that would provide a reasonable return on their investments [78]. This includes producing profitable pricing models, resource bundling options and licensing strategies [44]. Moreover, the way billing and payments are currently handled by different cloud providers lacks clarity as to what the customer is paying for in terms of type of service, quality and availability which makes financial benchmarking and comparison across different providers rather difficult [77,91].

From a cloud user’s perspective, there is the cost of migration from legacy systems to cloud-based systems—especially when huge investments have gone into building the legacy systems [110]. Other issues include predicting the potential cost of administering the business aspects that are hosted remotely, upgrading the network bandwidth to achieve practical performance (to use cloud resources), and reliability and backup measures [13]. For applications that are data-intensive such as media applications, cost issues are exacerbated [111].

**Service management.** The Cloud as a service-based IT model created a number of challenges pertaining to service management. Service provisioning seems to be at the core of such challenges [80]. The literature suggests that there is an urgent need for automating service provisioning [15][81] and making it more dynamic [91]. Automatic combination of services has also been suggested [90]. Another challenge is related to the ability to provide customizable [79] and more context-aware [91] services. The authors in [27] have recognized that managing longer-standing service workflows is a major challenge considering the impact of service failure on numerous complex applications into which the service is integrated. Furthermore, managing service lifecycle and service registry and subscription has proven to be challenging for various reasons [50].

**Quality.** The literature in this category makes it abundantly clear that the main challenge regarding quality of service in the cloud is the definition and use of service level agreements (SLAs). The literature identifies as a challenge the lack of SLAs between parties in the cloud.
which lowers consumers’ confidence in the reliability and availability of services and makes adoption more difficult [38]. The definition of quality in a cloud SLA is an important yet challenging aspect [97]. According to [38], preparing SLAs requires careful considerations such as quality at different layers (i.e., infrastructure, platform, software), the tradeoff between complicatedness and expressiveness in the agreement, and the evaluation and feedback mechanisms that keep the SLA relevant and up-to-date. Moreover, the lack of a standard set of service level objectives and quality of service metrics makes negotiation and benchmarking rather difficult [91]. Efforts are being made by researchers to address how an SLA should take into account service monitoring and elasticity [44] and resource provisioning [54].

Another challenge discussed in the literature is the quality of user experience. Especially in multimedia-intensive cloud services such as video streaming services [111] as well as online games [97], video quality and network delays have a huge impact on the overall experience of using the cloud as perceived by end-users.

Software. Literature on software challenges in the Cloud is not abundant but it covers a wide range of topics. Mohagheghi et al. [84] discussed the challenges associated with migrating legacy systems to the cloud including: modernizing the architecture to be more service-oriented, providing a cleaner data access layer, dealing with non-functional requirements such as quality, and using agile methods in the migration process. Software evolution is another major challenge according to [85] given that it could result in service unavailability and version inconsistency in multi-tiered systems. Riongu et al. [41] discussed challenges concerning using the cloud for software testing such as measuring the reliability of testing frameworks on the cloud, standardizing testing processes, fulfilling the 24/7 availability promise, and accessing test data.

At a more technical level, challenges were identified by [14] including standardizing application interfaces, the use of data structures and concurrency models, and capturing execution stages. Moreover, software aging (i.e., growing degradation of the internal state of software due to accumulated faults such as memory leaks) has been identified as a challenge [82]. Patidar et al. [83] discussed challenges with software development for the cloud emphasizing the complexity of the communication and coordination between software engineering and cloud providers during every stage in the software development process (i.e., requirement, design, implementation, testing).

Trust. Trust is recognized as a key obstacle in the way of adopting the Cloud and becoming dependent on its resources [28]. Challenges arise around long-term viability (i.e., trusting that stored data will not vanish as a result of a vendor going bankrupt or getting acquired) [39], having full control over mission critical activities [38], and trusting the supply chain of a given service provider at all different levels [40]. Harsh et al. [87] discussed trust as an issue in cloud federations since every participant needs to trust that all others do filter out malicious users, follow the code of conduct, and report accurate resource usage and billing information.

Khorshed et al. [86] identified trust as a barrier to providing effective remedies against cyber-attacks in the cloud. They mentioned lack of transparency between service providers, malicious insiders, and vulnerable shared technologies as some of the trust issues that need to be addressed.

Other. Miscellaneous topics that were briefly discussed in the literature as emerging challenges include modeling and simulation of cloud computing environments [8], outsourcing business to cloud computing services [9], IT management [53], change management [110], and user freedom in the cloud [39].

Table 2 lists all the categories and the number of publications (of the ones included in our analysis) in each category. The numbers in the categories do not necessarily have to add up to the actual number of the papers included in our analysis because a single paper would typically discuss more than one issue.

RQ2. How has the interest in different types of issues and challenges evolved over the past five years?

Our systematic review covered the period from the year 2007 until August 2012. The number of publications in each year is shown in Figure 1.

The key observation from this analysis is that on the year 2007, we could not find papers that explicitly discussed issues and challenges with cloud computing. In fact, having made this observation, we went back to look for papers that were dated before 2007 using the same search query. A total of 13 results were returned—none of them was relevant to cloud computing2. Looking at the Year 2008, we notice only one result. Noticeably, the number of results grows steadily in the years to follow until it reaches a maximum of 48 in 2011. This number declines in the year 2012.

The sudden increase of papers after 2008 is interesting given that cloud computing was not a new concept then. However, looking at the development of cloud computing as a field, we notice a few events that might have contributed to the wide-spread adoption of the concept which was also accompanied with more attention in academic circles. For example, in 2006, Amazon launched Amazon Web Service (AWS) which allowed easy access to hardware resources as a utility [113]. In late 2007,
Table 2. Categories of issues identified in the study.

<table>
<thead>
<tr>
<th>Category</th>
<th># of papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security &amp; privacy</td>
<td>66</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>46</td>
</tr>
<tr>
<td>Data management</td>
<td>15</td>
</tr>
<tr>
<td>Interoperability</td>
<td>12</td>
</tr>
<tr>
<td>Legal issues</td>
<td>9</td>
</tr>
<tr>
<td>Economic issues</td>
<td>9</td>
</tr>
<tr>
<td>Service management</td>
<td>8</td>
</tr>
<tr>
<td>Quality</td>
<td>8</td>
</tr>
<tr>
<td>Software</td>
<td>6</td>
</tr>
<tr>
<td>Trust</td>
<td>6</td>
</tr>
<tr>
<td>Other</td>
<td>5</td>
</tr>
</tbody>
</table>

Figure 1. Number of papers published from 2007-2012.

IBM and Google supported a joint university initiative called the Academic Cloud Computing Initiative (ACCI) as a project aiming to enhance students’ knowledge and skills to build cloud infrastructure and address the challenges of cloud computing [114]. In early 2008, Eucalyptus became the first open-source platform for deploying private clouds, and it was compatible with Amazon’s cloud service API [115]. In the same year, OpenNebula became the first open-source software for deploying private and hybrid clouds [116]. By mid-2008, a press release by Gartner affirmed that “organizations are switching from company-owned hardware and software assets to per-use service-based models... The projected shift to cloud computing, for example, will result in dramatic growth in IT products in some areas and in significant reductions in other areas. In general, assets will be utilized with greater efficiency.”

On the other hand, the sudden decline in 2012 can be attributed to two possible reasons:

1) The number of papers published in 2012 is not accurate given that some proceedings may not have been indexed yet in the electronic databases we have used. Therefore, an increase in the number of papers may occur over the four months not covered by this survey (September-December), and the positive trend may continue;

2) The number of papers has already hit a maximum and we might start to observe a decline in the trend over the next few years—which might indicate saturation in the issues and challenges being identified. This likelihood is supported by the fact that as of the date of our search, the number of papers (23) is significantly lower than where it should be at the end of 2012. Based on the linear trend obtained from the data (\( R^2 = 0.976 \)), the number of papers by the end of 2012 is projected to be about 61 if the positive trend was to continue.

Looking at the sources of publications over the past few years as illustrated in Figure 2, we notice that most publications were from the USA, followed by China and then India. Australia and the UK come next. The statistics here refer to the location of the first author’s institution.

For example, in our survey, 25 papers originated from US institutions. Although the number of publications varies between the top 3 countries, we notice a level of consistency when we look at the topics distribution as illustrated in Figure 3. That is, in all three countries, the top three topics of interest were security and privacy, infrastructure, and data management. It is also noteworthy that we could not find papers originating from the USA that address legal and trust issues. Meanwhile, China is lacking on the topic of quality and India is lacking on a number of categories.

Under RQ2, we also investigate when the different topics started to appear in the literature and when they gained the most attention. Figure 4 shows the evolution of interest across different topics from 2007 to 2012. In 2007, there were no papers in our dataset as discussed previously. The results for the year 2012 may not be reliable given that the date of our search query was prior to the end of the year 2012.

We notice that the first topics to be explicitly discussed as challenges were related to infrastructure and data management. Infrastructure continued to be an important topic throughout the past five years including 2012. Attention to data management peaked in 2010 and then declined in 2011 and 2012. In 2009, security and privacy started to be a major focus and continued a steady growth throughout the past five years including 2012. This might be due to the fact that late 2008 witnessed a huge adoption of cloud-based services like

---

3This is not a very accurate number given the small number of data points.
Facebook (monthly growth of 178%)\textsuperscript{5}. Other topics also started to appear in 2009 including interoperability, legal issues, economic challenges, and software challenges. These topics continued to be discussed in the years to follow. In 2010, the issue of trust started to be visible in the literature alongside quality and service management challenges. All of the abovementioned issues continued to appear in the Year 2011. Up until the date of our search query, there were no publications in the year 2012 dedicated to discuss challenges related to software challenges or trust issues.

3. Interviews

In this part of the study, we used interviews as our data collection method in order to understand issues and challenges with cloud computing from a practitioners’ perspective. At the end of this section, we identify gaps between what researchers have been focusing on and what practitioners deem important.

3.1. Data Collection

The authors conducted 8 in-depth semi-structured interviews with individuals of different teams and roles in two different organizations. Both organizations are Canadian organizations that specialize in cloud computing as an emerging technology. The first few interviewees were selected collaboratively by the researcher and liaisons in the organizations. Then, to select the next group of interviewees, we used snowball sampling, i.e., we used suggestions from the interviewees to guide the selection process of other interviewees.

Generally, interviewees were first asked questions to describe their role and team responsibilities. Then, they were asked questions about the definition of cloud computing from a technical perspective, their motivation behind adopting the Cloud, the advantages of being part of the cloud, and the issues and challenges they or their customers have been facing with the Cloud. The interviews lasted between 15 and 35 minutes each. The interviews were audio-taped and transcribed.

The data collection phase stopped when we started to get no new insights from new interviews.

3.2. Data Analysis

For the purpose of this paper, we only included the analysis of the part on issues and challenges. The collected data was analyzed by iterating over the interview data to assign codes, and we refined these codes as more data was coded. By the end, we had used 30 codes in total to tag the collected data. These codes effectively reflect issues and challenges found in the data (for example: performance, portability, security). However, all of the codes were identified prior to identifying the categories of challenges in the literature review.

In the next step, we classified the 30 codes under the categories developed through the literature review (as shown in Table 3). If there was no suitable category under which we could naturally classify a code, we created a new category. This theme of categorization and classification was intended so that we could compare the findings in each part of the study and identify similarities and differences.

3.3. Results & Discussion

In this section, we present the results of this part of the study and we analyze them in light of the fourth research question:

RQ3. Are there any gaps between the issues and challenges researchers have been focusing on and the issues and challenges practitioners deem to be important?

Table 3 shows a list of categories including the ones that were identified in the systematic literature review part of this study. Next to each category is an array of codes that fall under the given category and the number of times an issue related to the given category has been brought up by the interviewees. The two categories highlighted in green (namely: change management and learning) were added in the second part of the study to encapsulate issues that could not be classified under the
Table 3. Classification of issues identified in interviews with practitioners.

<table>
<thead>
<tr>
<th>Category</th>
<th>Codes</th>
<th>Number of occurrences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security &amp; privacy</td>
<td>Security, privacy</td>
<td>6</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>Performance, transfer rates, IP address space, limited capability,</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>fine-tuning hardware</td>
<td></td>
</tr>
<tr>
<td>Data management</td>
<td>Data design and management</td>
<td>2</td>
</tr>
<tr>
<td>Interoperability</td>
<td>Portability across platforms, lack of standards</td>
<td>10</td>
</tr>
<tr>
<td>Legal</td>
<td>Legal issues</td>
<td>4</td>
</tr>
<tr>
<td>Economic</td>
<td>Rent vs. buy, access to funding, cost of bandwidth</td>
<td>3</td>
</tr>
<tr>
<td>Service management</td>
<td>Service usability, service discovery</td>
<td>2</td>
</tr>
<tr>
<td>Quality</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Software</td>
<td>Software engineering process, instability, design for the cloud,</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>legacy applications</td>
<td></td>
</tr>
<tr>
<td>Trust</td>
<td>Trust</td>
<td>1</td>
</tr>
<tr>
<td>Change management</td>
<td>Change management, previous investments, perceived loss of jobs</td>
<td>3</td>
</tr>
<tr>
<td>Learning &amp; education</td>
<td>Education about the cloud, learning DevOps, lack of experience in the</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>cloud, learning curve, learning APIs, continuous learning, complexity of operations</td>
<td></td>
</tr>
</tbody>
</table>

existing categories. One category from the literature review (namely: quality), highlighted in blue, was not assigned any codes.

The practitioners we interviewed touched on issues related to all the categories identified in the literature review except for one category on Quality.

On software related issues, participants discussed porting legacy applications as a major challenge. As one participant stated that “you can’t just put some software that’s intended to be run on a desktop and just put it in the cloud and just automatically it’s easy to distribute.” Another participant asserted that “there is a lot of legacy applications that will not scale well in the cloud.” Instability of important platforms in the Cloud such as OpenStack has also been raised as a critical issue. One participant stated that “Sometimes [OpenStack] is just completely unstable and there’s bugs in it that would just stop everything.” Another participant described the platform as “a moving target.” He continued: “Everything is moving. It’s not stable.” The software engineering process itself is being challenged by new roles such as DevOps (aka. development and operations): “We definitely identify as DevOps and with the stuff that we’re doing in particular, seems to be a bit more ops heavy and that doesn’t lend itself as well to Agile.” Other issues that were raised include: integration problems, development versus deployment environments, and dealing with external and virtual resources.

On interoperability, participants brought up the issue of the lack of standards and open APIs. As one participant stated, “there’s a total lack of open APIs that a company could conform to.” Another participant asserts that even for a single provider, this seems to be a major challenge: “There does not seem to be an API architecture that Amazon says OK all our APIs should follow the same way-they will all work like this. They all work slightly differently and don’t always return the same results in a consistent way.” Other issues under this category include moving virtual machines among different platforms, and vendor lock-in.

On security and privacy, participants assert that security and privacy are amongst the key inhibitors to adopting the cloud. From a management perspective, one participant states that “management is very worried about the security of the cloud.” Nonetheless, one of the participants argued that although “privacy and security are big issues..., the cloud is safer. You have teams of people dedicated to security at these big cloud providers and they have to conform to all sorts of regulations.” Participants also expressed their concerns about the privacy of any data stored abroad mainly due to legislation issues.

On infrastructure, participants raised a range of issues. For example, as one participant noted, performance is major issue in the cloud: “there are notable issues with shared I/O that are still kind of being worked out where a physical server is still faster.” Another participant stated: “In a lot of cases it still makes far more sense to grab a stack of disks, tons of terabyte disks, and throw them into your trunk and drive to your destination, than it would to actually send it over the wire.”
The interviewees also touched on other challenges such as change management and learning. While the interviewees mentioned three different issues under change management (i.e., change management, previous investments, perceived loss of jobs), in our literature review, change management was only discussed in one paper (and therefore it was categorized under the “Miscellaneous” category). The learning and education aspect comes across in the result as one of the most important challenges; however, the literature does not discuss it as a challenge. This might be due to the nature of the business some of the practitioners we interviewed were involved in which entailed setting up and maintaining private clouds. One participant asserts that “there is definitely a learning curve to using the cloud and doing DevOps stuff.” Another participant states that “the major challenge for the developers is learning all of that knowledge that comes from the systems areas.” The fast evolution of cloud computing technologies has also been a challenge: “You have to keep learning. And as soon as you stop, at that point I think you have to find another line of work.” Other challenges under this category include educating the public about the capabilities and limitations of the cloud, and the lack of sufficient expertise in the area of cloud computing.

Despite the significant overlap between the identified issues in the two studies, we notice a difference in emphasis between what we found in the literature and what our participants focused on. That is, the most three frequently discussed issues during the interviews were software related issues, learning and experience, and interoperability. We then ranked the categories based on their frequency of occurrence in the literature review as well as the practitioner interviews. Figure 5 shows the difference in ranks between the systematic literature review and the interviews. Green bars denote a higher rank of the category in the practitioner interviews, whereas red bars denote a higher rank in the systematic review. For example, security and privacy ranked first in our literature review and dropped to fourth in the interviews. On the other hand, software ranked ninth in the literature review, but jumped to first in the interviews. For categories that did not show in one of the categories, we assumed it ranked last.

4. Limitations & Threats to Validity

Although the literature review was conducted in a systematic way, there is a number of factors that might have increased the subjectivity of the findings. First, our choice of the search string might have excluded papers that discussed issues and challenges but without necessarily expressing that explicitly in the title. Using the title (as opposed to the full text) in our search helped reduce noise significantly; because otherwise any paper that had the word “cloud” and the word “issue”, for example, would have been returned resulting in more than 24,000 mostly non-relevant papers. Also, our selection of specific databases might have excluded good contributions in the area that exist elsewhere. The categorization process of the surveyed papers is subjective and will vary from researcher to researcher. Similarly in the interviews, coding and classification are both subjective processes. We tried to mitigate this issue by involving more than one researcher in such activities. The literature review is also limited in that it does not take into account the last few months of the year 2012—which might have resulted in an incomplete picture of trends.

Like other qualitative studies, interviews are subject to a number of biases during the data collection stage as well as the analysis stage. The interviews are mainly limited due to the small number of participants and their affiliations.

5. Conclusion & Future Work

As the adoption of cloud computing is becoming increasingly common, issues and challenges are still emerging at the various levels of the cloud architecture. Our findings show that the literature on issues and challenges has increased steadily over the past five years (not including the year 2012). Including the partial results of the year 2012 shows a decline in the number of publications which might indicate that fewer issues are being identified and the focus of research is shifting towards finding solutions. Researchers have been mainly focusing on issues related to security and privacy, infrastructure, and data management. Interoperability across different service providers also seems to be an active area of research. From the practitioners’ perspective, issues related to learning and keeping up with new cloud technologies seemed to be more prominent alongside software related issues and interoperability issues. Despite the significant overlap between the topics being discussed in the literature and the issues raised by the practitioners, there clearly is a gap between the two camps. For example,
researchers have not been giving enough attention to the challenge of learning fast-evolving technologies in the cloud (e.g., virtualization tools). Also, software related issues are understudied. To fill this gap, our future work includes conducting intensive case-studies wherein we aim to understand challenges related to software design and testing in the cloud.

**REFERENCES**


[22] Z. Yusof and M. Tang, “A Penalty-Based Genetic Algorithm for the Composite SaaS Placement Problem in the
Emerging Issues & Challenges in Cloud Computing—A Hybrid Approach


Issues and User-Centric Solution,” Changsha, pp. 927-932.


Emerging Issues & Challenges in Cloud Computing—A Hybrid Approach

225-234.


