12-2009

Where are we at with Cloud Computing?: A Descriptive Literature Review

Haibo Yang
School of Information Management Victoria, University of Wellington, sodatea@gmail.com

Mary Tate
School of Information Management Victoria, University of Wellington, Mary.tate@vuw.ac.nz

Where are we at with Cloud Computing?: A Descriptive Literature Review

Haibo Yang
School of Information Management
Victoria University of Wellington
Email: Sodatea@gmail.com

Mary Tate
School of Information Management
Victoria University of Wellington
Email: Mary.Tate@vuw.ac.nz

Abstract

Cloud computing is an exciting area for research, because of its relative novelty and exploding growth. In this paper, we present a descriptive literature review and classification scheme for cloud computing research. The former consists of 58 articles published since the recent inception of cloud computing. Clearly, there is an explosively increasing amount of cloud computing research has been conducted this year. The articles are classified and results are presented, based on a scheme that consists of four main categories: technological issues, business issues, applications, and general. The results show that although current cloud computing research is still skewed towards technological issues, such as performance, network, and data management, new research theme regarding the social and organisational implications of cloud computing is emerging. We hope this review will provide a snapshot and reference source of the current state of cloud computing research and stimulate further research interest.

Keywords: Cloud computing, descriptive literature review

INTRODUCTION

In an age of information and globalisation, massive computing power is desired to generate business insights and competitive advantage (Liu and Orban, 2008). A traditional way for enterprises to process their data is to use the computing power provided by their own in-house data centres. However operating a private data centre to keep up with the rapid growing data processing requests can be complicated and costly.

Cloud computing offers an alternative. “Cloud computing”, as a term for this Internet based service, was launched by industry giants (e.g. Google, Amazon.com, etc.) in late 2006. It promises to provide on-demand computing power with quick implementation, little maintenance, less IT staff, and consequently lower cost (Aymerich et al., 2008). As projected by market-research firm IDC, IT cloud-service spending will grow from about USD16 billion in 2008 to about USD42 billion by 2012 (Leavitt, 2009).

The relative novelty and exploding growth of cloud computing makes it an exciting area for research. The present paper aims to assess the state of cloud computing research. We portray a current landscape of this research stream, where it is today, and, most importantly, where effort should be focused in the future in order to produce more “consumable research” (Robey and Markus, 1998).The reminder of this paper is organised as follows. First a brief overview of cloud computing is given. Next the research methodology and our classification schema are presented. This is followed by the results of our literature review and classification. Finally, some discussions and conclusion are drawn.

LITERATURE REVIEW

Since this is a literature review paper, the literature we evaluate is mainly discussed in the body of the paper. This section offers a short introduction to what cloud computing is, and how it can be distinguished from related concepts such as grid computing.

Cloud computing can be regarded to a certain degree, as the evolution of grid computing. Such a close relationship has caused confusion. The grid framework is originally driven by scientific purposes(e.g. SETI@home project), and aimed at coordinating resources that are not subject to centralised control under standard, open, general-purpose protocols and interfaces (Foster et al,2008). Cloud computing is born for commercial purposes and naturally service oriented. It is based on centralised data centres. The protocols and interfaces used may not be the same across clouds providers.

Cloud computing has completely distinct business model. It offers clear SLAs (Service Level Agreements) and is based on a "pay per use" pricing model (Weinhardt et al., 2009). Therefore it is promised that with nothing but a credit card, one can get on-demand access to 100,000+ processors from the clouds (Foster et al., 2008). Grid computing on the other hand is based on a sharing system, that is, one needs to join the grid and contribute computing power to gain access to the computing power of other members. In this co-operative model, SLAs are not required or enforceable (Weinhardt et al., 2009).

In this paper, we see cloud computing is based on "a large pool of easily usable and accessible virtualised resources (such as hardware, development platforms and/or services)" (Vaquero et al., 2009, p.51). These resources are typically built in centralised data centres and are dynamically re-configured to achieve an optimum utilization. Clouds are provided by a pay-as-you-go model in which guarantees are offered by the providers by means of customized SLAs (Service Level Agreements). This turns computing power into a public utility which will bring a profound "paradigm shift" to the IT industry and even to society as a whole.

Cloud computing has promised many technological and sociological benefits. The computing power is generated from highly centralised and standardized data centres which contain up to millions of servers, with a considerable economy of scale. From an enterprise standpoint, cloud computing can deliver on-demand computing power at a very low if not any cost of the upfront infrastructure and ongoing maintenance. Cloud computing also promises to provide better performance, reliability and scalability (Erdogmus, 2009). There is some evidence that these are being delivered, (Fox et al., 2009). From an environmental standpoint, owing to the advanced electrical and cooling systems used by its centralised data centres, cloud computing has promised to bring low environmental cost and high energy efficiency, compared to the traditional scattered enterprise data centres (Katz, 2009). All in all, these seductive promises have drawn drastically increasing attention of a world wide scale.

RESEARCH METHODOLOGY

A Descriptive Literature Review

Literature review is an essential approach to "conceptualise research areas and survey and synthesize prior research" (Webster and Watson, 2002). It directly contributes to a cumulative culture which is desired in any research field, especially in IS. It is suggested that the lack of review articles has been hindering the progress of IS field (Webster and Watson, 2002).

A literature review can be conducted in different ways. Figure 1, shows four methods of literature review: Narrative Review, Descriptive Review, Vote Counting, and Meta Analysis. These four review methods are placed in a qualitative-quantitative continuum to illustrate their different focuses (King and He, 2005).

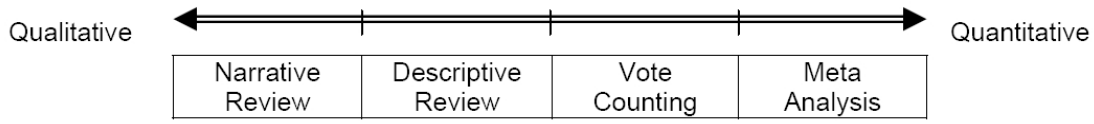

Figure 1 Review Methods on a Qualitative-Quantitative Continuum

The narrative review is the traditional way of reviewing the literature and is skewed towards a qualitative interpretation of the literature. It is conducted by verbally describing the past studies, focusing on theories and frameworks, elementary factors and their research outcomes, with regard to a hypothesized relationship (King and He, 2005). However, there is no consensus towards a standardized procedure of narrative review. How to conduct a narrative review largely depends on the reviewer's personal preference, thus this approach is easily vulnerable to subjectivity. It is not rare for "two reviews to arrive at rather different conclusion from the same general body of literature" (Guzzo et al., 1987, p. 408).

A descriptive review focuses on revealing an interpretable pattern from the existing literature (Guzzo et al., 1987).It produces some quantification, often in the form of frequency analysis, such as publication time, research methodology, research outcomes etc. Such a review method often has a systematic procedure including searching, filtering and classifying processes. First a reviewer needs to conduct a comprehensive literature search to collect as many relevant papers as possible in an investigated area. Then the reviewer treats an individual study as one data record and identifies trends and patterns among the papers surveyed (King and He, 2005). The outcome of such a review is often claimed to be representative of the state of art of a research domain.

Vote counting is generally used to draw inferences upon focal relationships by combining individual research findings (King and He 2005). Here a tally is made of the frequency with which existing research findings support a particular proposition. Most likely it is applied to generate insights from a series of experiments. The premise underlying this approach is that repeated results in the same direction across multiple studies, even some of them are non-significant, may be more powerful evidence than a single significant result (King and He, 2005).

Meta-analysis aims at statistically providing supports to a research topic by synthesizing and analysing the quantitative results of many empirical studies (King and He, 2005). In most cases, it may specifically examine the relationships between certain Independent Variables (IVs) and Dependent Variables (DVs) derived from existing research findings. Qualitative studies have to be excluded by a meta-analysis due to its extremely quantitative nature. Only similar quantitative studies are collected for a meta-analysis. The incentive of this approach is to generate a much less judgmental and subjective literature review upon a specific research context.

Our objective is to portray a landscape of cloud computing as an emerging research area and provide a snapshot or reference point to guide future development. Given the nascence of this research area, we do not and could not aim at examining any variables, correlations, or theories. We find a descriptive review approach was most appropriate. The procedure of this descriptive review is described in the next section.

Scope of the Literature Search

The first step of a literature analysis study is to locate relevant literature through computer and manual searches. Traditionally this is done by targeting some prominent journals and conferences. This approach is relevant to other research topics like Electronic Commerce where some major publication outlets have been formed by the long development of the research area (Bharati and Tarasewich, 2002). However focusing on limited outlets cannot be justified for a literature review on cloud computing as this is a recent phenomenon which just emerged two years ago, therefore the publication channels are still largely scattered. In the meantime, using online databases search as a primary literature collecting approach has become an emerging culture among IS researchers who are interested in contemporary phenomena (Petter and Mclean, 2009; Sabherwal et.al, 2006; Hwang and Thorn, 1999). Therefore for a literature review on cloud computing, it is safe and practical to focus on online databases rather than library collections.

Four prominent online databases, ACM Digital Library, IEEE Xplore, ProQuest (ABI/INFORM), and ScienceDirect (Elsevier) were targeted. These four databases cover almost all of the ISWorld's top 50 IS journals and most of top 10 IS conferences (Levy and Ellist, 2006). We therefore believe these databases are comprehensive enough to produce a literature set which can represent the current status of IS research literature.

We conducted keyword and abstract search across all the four databases and all years (till 3rd July 2009) with the phrase "cloud computing". The search aimed at peer-reviewed, scholarly articles, therefore restrictions were used if available (e.g. the "scholarly journals, including peer reviewed" option was selected in ProQuest; the "journal, proceeding" options were selected in ACM digital library). The initial search resulted in 161 hits.

Filtering Process

The 161 articles were then analysed in 2 steps. Firstly by scanning the title and abstract of the articles, duplicates, editorials, keynotes and panel talks were excluded; this left 87 articles. Secondly, following a procedure outlined by Grover et al. (1993), the full text of each article was reviewed; those which were not focused on cloud computing, instead merely mentioning cloud computing as a concept or a relevant technology, were discarded. Finally 58 articles, whose phenomenon of interest is clearly the technologies, business applications and impacts of cloud computing remained in the Endnote database for further classification.

Classification Scheme

The classification framework, shown in Table 1, was based on the literature review, the interest of cloud computing research, and the existing classification schemes of IS (Barki et al., 1993). The 58 articles were full text reviewed and grouped into 4 broad domains: Technological issues, Business issues, Applications, and General. This grouping is based on assigning the single most applicable topic-domain to an article. Then each domain was divided into sub domains according to the specific research interest of the articles. It is inevitable that any piece of research may contribute to several of these sub domains. However, by assigning one article to only one primary sub domain, we are able to offer a simplified classification of the major domains within current cloud computing research and conceptualise the relationships between these domains.

A: Technological issues: This domain focuses on the cloud computing technology. Articles in this domain are produced by researchers who see cloud computing as a white-box and are interested in its components and mechanism. Five categories are related to technological issues.

1. Cloud Performance: This sub domain covers articles focusing on the assessment, measurement, and optimization of the performance of the clouds. This includes studies that attempt to quantify the performance of cloud computing (Napper and Bientinesi, 2009; Fox et al., 2009), refine workflow scheduling and load balancing (Dornemann et al., 2009; Singh et al., 2009), and improve CPU performance (You et al, 2009, Woo and Lee, 2009).

2. Network: This includes design and management of the physical and logical network supporting special requirements of cloud computing. This sub domain consists of studies addressing problems about cloud computing network topologies and constructions (Streitberger and Eymann, 2009; Hou et al, 2009), large scale inter-cloud networks protocols (Bernstein et al., 2009), and cloud level traffic control (Matos et al., 2009).

3. Data Management: This sub domain aims at specific issues associated with the unprecedented scale of data processed in the clouds. Such as data parallelism (Gu and Grossman, 2008), data consistency (Vogels, 2009), and optimised algorithms and methods for cloud level data mining (Johnson, 2009; Grossman and Gu, 2008).

4. User control: Giving up user side control is widely criticized to be a downside of using cloud computing service. Articles in this sub domain attempt to technologically mitigate such concerns by providing solutions for retaining user control over data (Descher et al, 2009) and quick migration across clouds (Hirofuchi et al ,2009), .

Table 1 Classification of topics in Cloud Computing

Topics	Sub-Topics
Technological Issues	Cloud Performance, Network, Data Management, User Control, Software Development
Business Issues	Cost/Benefit, Pricing/Billing, Legal issues, Privacy
Applications	E-Science, Knowledge Management, Search Engines
General	Foundation/Introductions, Implementations

5. Software Development: This sub domain represents a stream of research focusing on developing distributed and parallel software in cloud computing environment (Rellermeyer et al, 2009).

B: Business issues: This domain concerns the business models and implications of cloud computing technology. Articles in this domain treat cloud computing as a black-box technology which can generate business value to both providers and users.

1. Cost/Benefit: This sub domain focuses largely on users' side. Researchers aim at quantifying the cost and benefit for migrating computing tasks onto the clouds (Assuncao et al., 2009). Such effort can further help users with selecting cloud services (Zeng et al., 2009).

2. Pricing/Billing: Articles in this sub domain mainly focus on providers' side. They propose pricing and billing models for cloud providers in order to retain customers and guarantee profit (Yeo et al., 2009).

3. Privacy: It is inevitable that in a cloud computing paradigm, privacy is always a concern as the cloud users have to upload, and in most case, store their data in publicly accessible data enters. Research in this sub domain argues that privacy protection better should be considered early on, at the cloud system design phase rather than post policy making stage (Pearson, 2009).

4. Legal issues: Cloud computing also involves many legal issues regarding data protection, confidentiality, copyright, and audit right etc. (Joint et al., 2009)

C: Applications: This domain consists of studies which tend to speculate or implement fresh applications on cloud computing platforms.

1. E-Science: A major consumer group of computing power is the E-Science community. E-Science refers to the scientific disciplines (i.e. earth science, bio-informatics, particle physics, etc.) where explosively increasing data gathered from sensors and instruments (i.e. the CREN Large Hadron Collider) need to be processed in a timely manner. Cloud computing, with its tremendous computing power and inexpensive cost has drawn considerable attention from E-Science community. Articles in this sub domain aim at understanding the impact of cloud computing on current computing infrastructure of E-Science (Hazelhurst, 2009).

2. Knowledge Management: KM is critical to knowledge intensive industries. These industries generally require large scale data processing and computing power as well. Research in this sub domain speculates how to

architecturally blend knowledge management practices with cloud computing and web 2.0 (Delic and Riley, 2009).

3. Search Engines: Web search engines naturally need to deal with arbitrary amounts of metadata cross the whole Internet. Research in this sub domain aims at the utilization of a cloud computing framework in the semantic search engines (Mika and Tummarelo, 2009).

D: General: This domain contains articles that discuss general aspects of cloud computing research, such as foundational concepts or introductions to cloud computing, or general implementation, etc.

1. Foundational/Introductions: This sub domain includes articles that introduce the definitions and components of cloud computing, and compare them with other computing paradigms (Lenk et al., 2009).

2. Implementation: This contains literature which illustrates the implementation of cloud computing framework and shares experience about its pros and cons (Aymerich et al., 2008).

This review takes a descriptive approach. We provide an overview of the current developments in cloud computing research, by conducting a systematic literature classification using the classification scheme presented above. The results of the classification are presented next.

RESULTS AND ANALYSIS

A total of 58 articles were classified according to our scheme. We also analysed the articled by year of publication, research methods, primary contribution, and the publication outlets.

Distributions of the articles by year

There are no studies related to "cloud computing" before 2007 because it did not exist under this name. As previously mentioned, the term "Cloud computing" was coined by industry practitioners in 2006. Academic researchers started to embark on this "bandwagon" later in 2007. The number of articles published in the first half of 2009 (39) has already doubled that of the whole of 2008 (18). This explosive growing of publications reflects the increasing enthusiasm on the cloud computing paradigm.

Distributions of articles by topics [3]

"Technological issues" clearly stands out as the most heavily published research topic (28 articles, 48%), followed by "General issues" (15 articles, 26%), "Business issues" (9 articles, 16%) while the least published topic is "Applications" (6 articles, 10%). This probably is because that there are still many technological obstacles for the growth of cloud computing which need to be removed, such as Data Lock-In, Data Confidentiality and Auditability, Data Transfer Bottlenecks, and Performance Unpredictability, etc. (Armbrust et al, 2009).

Table 2 lists the number of articles for **each** technological issue. We can see that the majority, 13 articles (46%) are based on cloud performance. This is unsurprising as the performance of cloud computing is always a critical factor to be considered by any user groups. Therefore the measurement, assessment and improvement of the performance of a certain cloud computing service is of great interest to the researchers. Network issues are second (7 articles, 25%) as currently the bandwidth in many areas is still a bottleneck which hinders the full actualization of cloud computing power. Researchers attempt to release the maximum power of the clouds by improving the efficiency and capacity of current network architecture (Jinno and Tsukishima, 2009; Bernstein et al., 2009).

Table 2 Number of Technological Issues Articles

Technological Issues	Number of articles
Cloud Performance	13(46%)
Network	7(25%)
Data Management	4(14%)
User Control	3(11%)
Software Development	1(4%)
Total	28(100%)

Table 3 shows the number of articles in topics of business issues. Cost/benefit stands at the first with 56% of the articles, followed by pricing/billing in related to 22% articles. This echoes the trend that most organisations have refocused on cost efficiency with regard to IT investment under the current economic downturn. Evaluating and

[3] A summary table of all articles classified is also provided in the Appendix.

quantifying costs and benefits of cloud computing services is very pertinent for those organisations which are planning to adopt cloud computing mainly for a cost-saving purpose. For example, Assuncao et al., (2009) conducted a series of experiments to assess the cost of improving the performance in a scenario in which a large organisation that has its own computing infrastructure but want to allocate additional resources from a cloud provider. Although legal issues and privacy are still a concern, seemingly a lesser concern than getting the cost and price structure right.

Table 3 Number of Business Issues Articles

Business Issues	Number of articles
Cost/Benefit	5(56%)
Pricing/Billing	2(22%)
Legal Issues	1(11%)
Privacy	1(11%)
Total	9(100%)

Table 4 shows the number of articles in related to applications. 67% of the articles are on E-Science. Although knowledge management and search engines are also believed to be prominent applications (Delic and Riley, 2009; Mika and Tummarello, 2009), the already existing hunger for arbitrary computing power in E-Science communities clearly provokes much more passion.

Table 4 Number of Application Articles

Applications	Number of articles
E-Science	4(67%)
Knowledge Management	1(17%)
Search engines	1(17%)
Total	6(100%)

Table 5 shows the number of articles in general topics. 67% of the articles are general introductions which provide foundational concepts and knowledge of cloud computing. For example, Foster et al., (2008) explain the differences between grid computing and cloud computing by making a comprehensive comparison. Mei et al., (2008) provide a comparison framework to clarify the differences between cloud computing, service computing, and ubiquitous computing. Effort in this direction is understandable as cloud computing, after all, is still a fresh paradigm which needs more time to be well conceptualised.

Table 5 Number of General Articles

General	Number of articles
Foundational/Introductions	10(67%)
Implementations	5(33%)
Total	15(100%)

Distributions of articles by research methods

The research methods of the articles were also identified and classified. The bulk of the articles we reviewed, 36 articles (62%) used experimental, simulation, and design research methods.This can be seen as a direct result of the most focused research topic "Technological Issues" which need to be solved by technological approaches. Research has focused on testing and enhancing the performance of the clouds (Napper and Bientinesi, 2009), adjusting the network design (Matos et al., 2009), improving the data processing (Grossman and Gu, 2008).

A considerable portion of articles (26%) are based on speculation and commentary. The authors of these publications are often senior researchers who have comprehensive experiences in academia or industry, or both. Given their extensive knowledge and sharp insights, their conceptual speculation on any aspects of cloud computing could be seen as valuable. Most articles in foundational/introductions category (see Table 5) were produced through this approach.

"Literature review" is the least used research method in our sample and only 7 articles (12%) are based on this approach. This is unsurprising due to the fact that very limited extant literature is directly related to cloud computing which was only proposed as a fresh computing paradigm approximately two years ago.

Distributions of articles by primary contributions

The primary contributions of these articles were also categorised. Thirty-five percent (20 articles) of the articles provide some insights and visions about the field, derived from speculations (Delic and Riley, 2009), experiments (Hazelhurst, 2008) and literature reviews (Vaquero et al., 2009). Thirty-one percent (18 articles) of the articles contribute some exemplar systems or tools for measuring cloud performance (Yigibasi et al.,2009) and refining cloud computing architecture (Lagar-Cavilla et al., 2009).

Among the exemplar systems, an interesting trend can be seen is that many academic researchers tend to build their own systems under the cloud computing paradigm and demonstrate superior performance compared to the counterparts used by commercial cloud providers. For example Grossman et al. (2009) designed a middleware called Sphere and compared it with Hadoop (Yahoo! Inc.,2009). The results of their experiments showed the former is almost 2 to 3 times faster than the latter (Grossman et al., 2009).

Twenty-two percent (13 articles) of the articles provide some algorithms or methods to solve certain problems, i.e. the algorithm for cloud workflow scheduling (Yang, et al., 2008), the algorithm for cloud service selection (Zeng et al.,2009). Most of these were derived from experiments or simulations. Twelve percent (7 articles) of the articles provide some frameworks or models which were generated by literature review (Youseff et al., 2009) or speculations (Weinhardt et al., 2009).

Publication outlets

The publication outlets of the articles were also analysed. The majority of the literature (42 articles, 72%) is from conference proceedings, while only 28% (16 articles) is from journals. This to some degree reflects the general lifecycle of academic publication, namely, first get through a conference, after that, to a journal. As cloud computing is new research phenomenon, such an unbalance is unsurprising.

Table 6 Distributions of Articles by Journals

Journal	No of publications
Future Generation Computer Systems	3
IT professional	3
Computing in Science & Engineering	2
Communications of the ACM	2
Computer Networks	1
Computer Law & Security Review	1
IEEE Computer	1
IEEE Intelligent Systems	1
SIGCOMM Computer Communication Review	1
SIGOPS Operating System Review	1
Total	16

The 16 journal articles were published across ten journals as shown in Table 6. Clearly most of the major IS journals (Levy and Ellis, 2006) have not embarked on the cloud computing "bandwagon" yet. It is understandable as there are still quite a few technological obstacles in front of a world wide adoption of cloud computing, it may be a bit early for most IS researcher to investigate the social and organisational issues of this new paradigm.

Table 7 Distributions of Articles by Conferences (partial)

Conference	No of publications
IEEE/ACM International Symposium on the Cluster Computing and the Grid	5
IEEE International Conference on E-Science	4
ICSE Workshop on the Software Engineering Challenges of Cloud Computing	4
Grid Computing Environments Workshop	2
ACM/IEEE conference on Supercomputing.	1
ACM European conference on Computer systems.	1
ACM international symposium on High performance distributed computing.	1
IEEE Asia-Pacific Services Computing Conference	1
IEEE International Conference on Data Engineering	1
IEEE International Conference on High Performance Computing and Communications	1
Total	21

Half of the conference articles were presented at ten conferences as shown in Table 7. The other half are published via other 21 conferences (not listed), one at each. We believe such abundant publication outlets echo the broadly relevant nature of cloud computing and also will encourage more researchers to conduct cloud computing related studies.

DISCUSSION AND CONCLUSIONS

This paper has limitations. First, our sample was mainly based on academic publications. As cloud computing is industry-driven in nature, many good professional articles may also embrace this phenomenon. This may hinder the ability of the present paper to delineate a complete picture for the current development in this domain. Secondly, our search criteria may be incomplete, as some papers that do not have the term cloud computing in the title or key words may not have been included.

The intention of this paper is to illustrate a landscape of current academic research from an IS standpoint. We have presented a descriptive review, classifying the literature of extant cloud computing research in a range of categories. The results presented in this paper have several implications.

First, we expect an exponential growth in the number of cloud computing research in a near future. According to our review, the number of research articles in the first half year of 2009 has already doubled that of the whole of 2008. The economic downturn is fuelling the burning interest towards cloud computing, there is no doubt that cloud computing research will increase significantly in future. A wide variety of publication outlets are accepting research on cloud computing, with a focus, at this stage, on more technical outlets.

Second, a unified definition and a set of technical standards of cloud computing are being forged. Many researchers (Buyya et al.,2009; Foster et al., 2008; Mei et al.,2008; Weinhardt et al., 2009) are working on clarifying the salient features of cloud computing and differentiating it from other frameworks (i.e. grid computing, service computing, cluster computing, and ubiquitous computing). Effort in this direction clearly defines the technological components of cloud computing and mitigates the confusions around its boundary.

Third, while currently the majority of the articles focus on cloud providers or brokers and aim at addressing technological obstacles, little has been done from cloud consumer's standpoint. It may be that this is a natural evolution, and that the technology needs to perform reliably before mission-critical business applications can be implemented.

Finally, our review indicated that theory building is still not at the centre of cloud computing research. Instead, most studies focus on praxis. This is consistent with the trend in other nascent research areas such as mobile business (Scornavacca et al.,2006). The missing of solid theoretical foundations is always a concern in IS academics. This is because of a traditional view that the academic legitimacy of a research field hinges on the presence or absence of core theories. However, recently, Lyytinen and King (2004) argue that, to increase legitimacy of an "applied research" field like IS, relevance to praxis, can and should be placed at the centre. Salience and strong results should be major determinants of the academic legitimacy of IS research field. Cloud computing clearly has salience. Producing strong research results related to praxis may be a natural way to strengthen the legitimacy of this research area.

Nevertheless, the global recession is forcing the IT functions of organisations to focus on cost and resource efficiency which is promised as a major benefit of cloud computing. We suggest that information systems researchers could consider the following questions. Should an enterprise adopt cloud computing and when? This could be investigated from the point of view of IS strategy, and organisational diffusion of innovation. If yes, what aspects should be considered when choosing a cloud provider? What criteria can be used to make a comparison across the different cloud services? This could be informed by insights from outsourcing literature. Will cloud computing help to mitigate the information systems management problems typically experienced by small-medium enterprises? All these questions are currently unanswered but could be answerable. Cloud computing has displayed huge potential for IS researchers to produce "consumable research" (Robey and Markus, 1998). By investigating these questions, IS researcher may be able to help the decision making of enterprises towards cloud computing adoption.

Although this review cannot claim to be exhaustive, it provides some insights into the state-of-the-art. Our classification and descriptive review can provide a snapshot and reference base for academics and practitioners with an interest in cloud computing.

REFERENCES

Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R., Konwinski, A., Lee, G., Rabkin, A., Stoica, I., and Zaharia, M.2009. "Above the Clouds: A Berkeley View of Cloud Computing." Retrieved 10th June 2009 from www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.pdf

Assuncao, M.D.d., Costanzo, A.d., and Buyya, R. 2009. "Evaluating the Cost-Benefit of Using Cloud Computing to Extend the Capacity of Clusters," in: *Proceedings of the 18th ACM international symposium on High performance distributed computing*. Garching, Germany: ACM.

Aymerich, F.M., Fenu, G., and Surcis, S. 2008."An Approach to a Cloud Computing Network," *Applications of Digital Information and Web Technologies, 2008. ICADIWT 2008. First International*

Conference on the, pp. 113-118.

Barki, H., Rivard, S., and Talbot, J. 1993. "A Keyword Classification Scheme for Is Research Literature: An Update," *MIS Quarterly* (June), pp 209-225.

Bernstein, D., and Ludvigson, E. 2009. "Networking Challenges and Resultant Approaches for Large Scale Cloud Construction," *Grid and Pervasive Computing Conference, 2009. GPC '09. Workshops at the*, pp. 136-142.

Bernstein, D., Ludvigson, E., Sankar, K., Diamond, S., and Morrow, M. 2009. "Blueprint for the Intercloud - Protocols and Formats for Cloud Computing Interoperability," *Internet and Web Applications and Services, 2009. ICIW '09. Fourth International Conference on*, pp. 328-336.

Bharati, P. and Tarasewich,P. 2002." Global Perceptions of Journals Publishing Ecommerce Research", *Communication of the ACM,* (45:5),pp 21-26.

Buyya, R., Yeo, C.S., Venugopal, S., Broberg, J., and Brandic, I. 2009. "Cloud Computing and Emerging It Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility," *Future Generation Computer Systems* (25:6), pp 599-616.

Cappos, J., Beschastnikh, I., Krishnamurthy, A., and Anderson, T. 2009. "Seattle: A Platform for Educational Cloud Computing," in: *Proceedings of the 40th ACM technical symposium on Computer science education*. Chattanooga, TN, USA: ACM.

Dash, D., Kantere, V., and Ailamaki, A. 2009. "An Economic Model for Self-Tuned Cloud Caching," *Data Engineering, 2009. ICDE '09. IEEE 25th International Conference on*, pp. 1687-1693.

Delic, K.A., and Riley, J.A. 2009. "Enterprise Knowledge Clouds: Next Generation Km Systems?," *Information, Process, and Knowledge Management, 2009. eKNOW '09. International Conference on*, pp. 49-53.

Descher, M., Masser, P., Feilhauer, T., Tjoa, A.M., and Huemer, D. 2009. "Retaining Data Control to the Client in Infrastructure Clouds," *Availability, Reliability and Security, 2009. ARES '09. International Conference on*, pp. 9-16.

Dornemann, T., Juhnke, E., and Freisleben, B. 2009. "On-Demand Resource Provisioning for Bpel Workflows Using Amazon's Elastic Compute Cloud," *Cluster Computing and the Grid, 2009. CCGRID '09. 9th IEEE/ACM International Symposium on*, pp. 140-147.

Erdogmus, H. 2009. "Cloud Computing: Does Nirvana Hide Behind the Nebula?," *IEEE Software* (26:2), pp 4-6.

Foster, I., Yong, Z., Raicu, I., and Lu, S. 2008. "Cloud Computing and Grid Computing 360-Degree Compared," *Grid Computing Environments Workshop, 2008. GCE '08*, pp. 1-10.

Fox, G.C., Ho, A., Chan, E., and Wang, W. 2009. "Measured Characteristics of Distributed Cloud Computing Infrastructure for Message-Based Collaboration Applications," *Collaborative Technologies and Systems, 2009. CTS '09. International Symposium on*, pp. 465-467.

Grossman, R.L. 2009. "The Case for Cloud Computing," *IT Professional* (11:2), pp 23-27.

Grossman, R., and Gu, Y. 2008. "Data Mining Using High Performance Data Clouds: Experimental Studies Using Sector and Sphere," in: *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. Las Vegas, Nevada, USA: ACM.

Grossman, R.L., Gu, Y., Sabala, M., and Zhang, W. 2009. "Compute and Storage Clouds Using Wide Area High Performance Networks," *Future Generation Computer Systems* (25:2), pp 179-183.

Grover, V., Lee, C.C., and Durand, D. 1993. "Analyzing Methodological Rigor of Mis Survey Research from 1980-1989," *Information & Management* (24), pp 305-317.

Gu, Y., and Grossman, R. 2008. "Exploring Data Parallelism and Locality in Wide Area Networks," *Many-Task Computing on Grids and Supercomputers, 2008. MTAGS 2008. Workshop on*, pp. 1-10.

Guzzo, R.A., Jackson, S.E., and Katzell, R.A. 1987. "Meta-Analysis Analysis," *Research in Organizational Behavior* (9), pp 407-442.

Hazelhurst, S. 2008. "Scientific Computing Using Virtual High-Performance Computing: A Case Study Using the Amazon Elastic Computing Cloud," in: *Proceedings of the 2008 annual research conference of the South African Institute of Computer Scientists and Information Technologists on IT research in developing countries: riding the wave of technology*. Wilderness, South Africa: ACM.

Hirofuchi, T., Ogawa, H., Nakada, H., Itoh, S., and Sekiguchi, S. 2009. "A Live Storage Migration Mechanism over Wan for Relocatable Virtual Machine Services on Clouds," *Cluster Computing and the Grid, 2009. CCGRID '09. 9th IEEE/ACM International Symposium on*, pp. 460-465.

Hoffa, C., Mehta, G., Freeman, T., Deelman, E., Keahey, K., Berriman, B., and Good, J. 2008. "On the Use of Cloud Computing for Scientific Workflows," *eScience, 2008. eScience '08. IEEE Fourth International Conference on*, pp. 640-645.

Hou, Y., Zafer, M., Kang-won, L., Verma, D., and Leung, K.K. 2009. "On the Mapping between Logical and Physical Topologies," *Communication Systems and Networks and Workshops, 2009. COMSNETS 2009. First International*, pp. 1-10.

Hutchinson, C., Ward, J., and Castilon, K. 2009. "Navigating the Next-Generation Application Architecture," *IT Professional* (11:2), pp 18-22.

Hwang, M.I., and Thorn, R.G. 1999. "The Effect of User Engagement on System Success: A Meta-Analytical Integration of Research Findings," *Information & Management* (35), pp 229-236.

Jinno, M., and Tsukishima, Y. 2009. "Virtualized Optical Network (Von) for Agile Cloud Computing Environment," *Optical Fiber Communication - includes post deadline papers, 2009. OFC 2009. Conference on*, pp. 1-3.

Johnson, J.L. 2009. "SQL in the Clouds," *Computing in Science & Engineering* (11:4), pp 12-28.

Joint, A., Baker, E., and Eccles, E. 2009. "Hey, You, Get Off of That Cloud?," *Computer Law & Security Review* (25:3), pp 270-274.

Katz, R.H. 2009. "Tech Titans Building Boom," *IEEE Spectrum* (46:2), p 40.

Kesavan, M., Ranadive, A., Gavrilovska, A., and Schwan, K. 2008. "Active Coordination (Act) - toward Effectively Managing Virtualized Multicore Clouds," *Cluster Computing, 2008 IEEE International Conference on*, pp. 23-32.

King, W.R., and He, J. 2005. "Understanding the Role and Methods of Meta-Analysis in Is Research," *Communications of the Association for Information Systems* (16), p 1.

Kroeker, K.L. 2009. "The Evolution of Virtualization," *Association for Computing Machinery. Communications of the ACM* (52:3), p 18.

Lagar-Cavilla, H.A., Whitney, J.A., Scannell, A.M., Patchin, P., Rumble, S.M., Lara, E.d., Brudno, M., and Satyanarayanan, M. 2009. "Snowflock: Rapid Virtual Machine Cloning for Cloud Computing," in: *Proceedings of the fourth ACM european conference on Computer systems*. Nuremberg, Germany: ACM.

Leavitt, N. 2009. "Is Cloud Computing Really Ready for Prime Time?," *Computer* (42:1), pp 15-20.

Lenk, A., Klems, M., Nimis, J., Tai, S., and Sandholm, T. 2009. "What's inside the Cloud? An Architectural Map of the Cloud Landscape," *Software Engineering Challenges of Cloud Computing, 2009. CLOUD '09. ICSE Workshop on*, pp. 23-31.

Levy, Y., and Ellis, T.J. 2006. "A Systems Approach to Conduct an Effective Literature Review in Support of Information Systems Research," *Informing Science Journal* (9), pp 181-211.

Liu, H., and Orban, D. 2008. "Gridbatch: Cloud Computing for Large-Scale Data-Intensive Batch Applications," *Cluster Computing and the Grid, 2008. CCGRID '08. 8th IEEE International Symposium on*, pp. 295-305.

Llora, X., Acs, B., Auvil, L.S., Capitanu, B., Welge, M.E., and Goldberg, D.E. 2008. "Meandre: Semantic-Driven Data-Intensive Flows in the Clouds," *eScience, 2008. eScience '08. IEEE Fourth International Conference on*, pp. 238-245.

Lyytinen, K., and King, J.L. 2004. "Nothing at the Centre? Academic Legitimacy in the Information Systems Field," *Journal of the Association for Information Systems,* (5:6), pp 220-264.

Matos, M., Sousa, A., Pereira, J., and Oliveira, R. 2009. "Clon: Overlay Network for Clouds " in: *Proceedings of the Third Workshop on Dependable Distributed Data Management*. Nuremberg, Germany: ACM.

Mei, L., Chan, W.K., and Tse, T.H. 2008. "A Tale of Clouds: Paradigm Comparisons and Some Thoughts on Research Issues," *Asia-Pacific Services Computing Conference, 2008. APSCC '08. IEEE*, pp. 464-469.

Mika, P., and Tummarello, G. 2008. "Web Semantics in the Clouds," *Intelligent Systems, IEEE* (23:5), pp 82-87.

Napper, J., and Bientinesi, P. 2009. "Can Cloud Computing Reach the Top500?," in: *Proceedings of the combined workshops on UnConventional high performance computing workshop plus memory access workshop*. Ischia, Italy: ACM.

Nurmi, D., Wolski, R., Grzegorczyk, C., Obertelli, G., Soman, S., Youseff, L., and Zagorodnov, D. 2009. "The Eucalyptus Open-Source Cloud-Computing System," *Cluster Computing and the Grid, 2009. CCGRID '09. 9th IEEE/ACM International Symposium on*, pp. 124-131.

Pearson, S. 2009. "Taking Account of Privacy When Designing Cloud Computing Services," *Software Engineering Challenges of Cloud Computing, 2009. CLOUD '09. ICSE Workshop on*, pp. 44-52.

Pallickara, S., Ekanayake, J., and Fox, G. 2008. "An Overview of the Granules Runtime for Cloud Computing," *eScience, 2008. eScience '08. IEEE Fourth International Conference on*, pp. 412-413.

Patchin, P., Lagar-Cavilla, H.A., Lara, E.d., and Brudno, M. 2009. "Adding the Easy Button to the Cloud with Snowflock and Mpi," in: *Proceedings of the 3rd ACM Workshop on System-level Virtualization for High Performance Computing*. Nuremburg, Germany: ACM.

Petter, S., and McLean, E.R. 2009. "A Meta-Analytic Assessment of the Delone and Mclean Is Success Model: An Examination of Is Success at the Individual Level," *Information & Management* (46:3), pp 159-166.

Raghavan, B., Vishwanath, K., Ramabhadran, S., Yocum, K., and Snoeren, A.C. 2007. "Cloud Control with Distributed Rate Limiting," in: *Proceedings of the 2007 conference on Applications, technologies, architectures, and protocols for computer communications*. Kyoto, Japan: ACM.

Ragusa, C., Longo, F., and Puliafito, A. 2009. "Experiencing with the Cloud over Glite," *Software Engineering Challenges of Cloud Computing, 2009. CLOUD '09. ICSE Workshop on*, pp. 53-60.

Rellermeyer, J.S., Duller, M., and Alonso, G. 2009. "Engineering the Cloud from Software Modules," *Software Engineering Challenges of Cloud Computing, 2009. CLOUD '09. ICSE Workshop on*, pp. 32-37.

Robey, F., and Markus, M.L. 1998. "Beyond Rigor and Relevance:Producing Consumable Reserach About Information Systems," *Information Resources Management Journal* (11:1), pp 7-15.

Sabherwal, R., Jeyaraj, A., and Chowa, C. 2006. "Information System Success: Individual and Organizational Determinants," *Management Science* (52:12), p 1849.

Scornavacca, E., Barnes, S. J., & Huff, S. L. (2006). Mobile business research published in 2000-2004: Emergence, Current status, and Future opportunities. *Communications of the Association for Information Systems*, 17, 635-646.

Silva, J.N., Veiga, L., and Ferreira, P. 2008. "Heuristic for Resources Allocation on Utility Computing Infrastructures," in: *Proceedings of the 6th international workshop on Middleware for grid computing*. Leuven, Belgium: ACM.

Sterling, T., and Stark, D. 2009. "A High-Performance Computing Forecast: Partly Cloudy," *Computing in Science & Engineering* (11:4), pp 42-49.

Streitberger, W., and Eymann, T. 2009. "A Simulation of an Economic, Self-Organising Resource Allocation Approach for Application Layer Networks," *Computer Networks* (In Press, Corrected Proof).

Singh, A., Korupolu, M., and Mohapatra, D. 2008. "Server-Storage Virtualization: Integration and Load Balancing in Data Centers," in: *Proceedings of the 2008 ACM/IEEE conference on Supercomputing*. Austin, Texas: IEEE Press.

Vogels, W. 2009. "Eventually Consistent," *Association for Computing Machinery. Communications of the ACM* (52:1), p 40.

Vaquero, L.M., Rodero-Merino, L., Caceres, J., and Lindner, M. 2009. "A Break in the Clouds: Towards a Cloud Definition," *SIGCOMM Computer Communication Review* (39:1), pp 50-55.

Vouk, M.A. 2008. "Cloud Computing; Issues, Research and Implementations," *Information Technology Interfaces, 2008. ITI 2008. 30th International Conference on*, pp. 31-40.

Walker, E. 2009. "The Real Cost of a Cpu Hour," *IEEE Computer* (42:4), pp 35-41.

Wang, L., Tao, J., Kunze, M., Castellanos, A.C., Kramer, D., and Karl, W. 2008. "Scientific Cloud Computing: Early Definition and Experience," *High Performance Computing and Communications, 2008. HPCC '08. 10th IEEE International Conference on*, pp. 825-830.

Webster, J., and Watson, R.T. 2002. "Analyzing the Past to Prepare for the Future: Writing a Literature Review. ," *MIS Quarterly* (26 2), pp iii-xiii.

Weinhardt, C., Anandasivam, A., Blau, B., and Stosser, J. 2009. "Business Models in the Service World," *IT Professional* (11:2), pp 28-33.

Woo, D.H., and Lee, H.-H.S. 2009. "Prophet: Goal-Oriented Provisioning for Highly Tunable Multicore Processors in Cloud Computing," *SIGOPS Operating System Review* (43:2), pp 102-103.

Yahoo!Inc. 2009. "Hadoop and Distributed Computing at Yahoo! ." Retrieved 20th June, 2009 from http://developer.yahoo.com/hadoop/

Yang, Y., Liu, K., Chen, J., Liu, X., Yuan, D., and Jin, H. 2008. "An Algorithm in Swindew-C for Scheduling Transaction-Intensive Cost-Constrained Cloud Workflows," *eScience, 2008. eScience '08. IEEE Fourth International Conference on*, pp. 374-375.

Yeo, C.S., Venugopal, S., Chu, X., and Buyya, R. 2009. "Autonomic Metered Pricing for a Utility Computing Service," *Future Generation Computer Systems* (In Press, Accepted Manuscript).

Yigitbasi, N., Iosup, A., Epema, D., and Ostermann, S. 2009. "C-Meter: A Framework for Performance Analysis of Computing Clouds," *Cluster Computing and the Grid, 2009. CCGRID '09. 9th IEEE/ACM International Symposium on*, pp. 472-477.

You, X., Xu, X., Wan, J., and Jiang, C. 2009. "Analysis and Evaluation of the Scheduling Algorithms in Virtual Environment," *Embedded Software and Systems, 2009. ICESS '09. International Conference on*, pp. 291-296.

Youseff, L., Butrico, M., and Da Silva, D. 2008. "Toward a Unified Ontology of Cloud Computing," *Grid Computing Environments Workshop, 2008. GCE '08*, pp. 1-10.

Zeng, W., Zhao, Y., and Zeng, J. 2009. "Cloud Service and Service Selection Algorithm Research," in: *Proceedings of the first ACM/SIGEVO Summit on Genetic and Evolutionary Computation*. Shanghai, China: ACM.

APPENDIX

Table 8 Classification of reviewed articles

Topics	Reference
Technological Issues	
Cloud Performance	You et al., 2009; Yigitbasi et al.,2009; Woo and Lee, 2009; Napper and Bientinesi, 2009; Lagar-Cavilla et al., 2009; Grossman et al., 2009; Fox et al., 2009; Dornemann et al., 2009; Liu and Orban, 2008; Yang et al.,2008;Singh et al,.2008; Kesavan et al.,2008; Hoffa et al., 2008.
Network	Streitberger and Eymann, 2009; Matos et al., 2009; Jinno and Tsukishima, 2009; Hou et al., 2009; Bernstein et al., 2009; Bernstein and Ludvigson, 2009; Raghavan et al, 2007.
Data Management	Vogels, 2009; Johnson, 2009; Gu and Grossman, 2008; Grossman and Gu, 2008.
User Control	Hirofuchi et al., 2009; Descher et al., 2009; Patchin et al., 2009.
Software Development	Rellermeyer et al., 2009.

Business Issues	
Cost/Benefit	Zeng et al., 2009; Walker, 2009; Dash et al., 2009; Assuncao et al., 2009; Silva et al., 2008.
Pricing/Billing	Weinhardt et al., 2009; Yeo et al., 2009.
Legal Issues	Joint et al., 2009.
Privacy	Pearson, 2009.
Applications	
E-Science	Nurmi et al., 2009; Pallickara et al., 2008; Llora et al., 2008; Hazelhurst, 2008.
Knowledge Management	Delic and Riley, 2009.
Search Engines	Mika and Tummarello, 2009.
General	
Foundational/Introductions	Lenk et al., 2009; Buyya et al.,2009; Grossman, 2009; Kroeker, 2009; Hutchinson et al., 2009; Vaquero et al.,2009; Sterling and Stark, 2009; Foster et al., 2008; Mei et al., 2008; Youseff et al., 2008.
Implementations	Ragusa et al.,2009; Cappos et al., 2009; Wang et al., 2008; Vouk,2008; Aymerich et al., 2008.

COPYRIGHT