



Postprint Version	1.0
Journal website	http://www.sciencedirect.com/science/article/pii/S0167739X14001502
Pubmed link	
DOI	10.1016/j.future.2014.08.002

This is a NIVEL certified Post Print, more info at <http://www.nivel.eu>

Chargeback for cloud services

THIJS BAARS ^{A,*}, RAVI KHADKA ^A, HRISTO STEFANOV ^{A,B}, SLINGER JANSEN ^A, RONALD BATENBURG ^A, EUGENE VAN HEUSDEN ^B

a Utrecht University, PO Box 80.089, 3508TB Utrecht, The Netherlands

b IBM, Amsterdam, The Netherlands

ABSTRACT

With pay-per-use pricing models, elastic scaling of resources, and the use of shared virtualized infrastructures, cloud computing offers more efficient use of capital and agility. To leverage the advantages of cloud computing, organizations have to introduce cloud-specific chargeback practices. Organizations have to allocate IT service costs to business users in a way that reflects service consumption. To help organizations become effective users of cloud services, this article provides an overview of the factors that influence chargeback in the cloud services. This is an initial work that determines the factors influencing the chargeback in the cloud services. The findings of this research facilitate organizations to realize the implications of the cloud for their chargeback .

1. INTRODUCTION

In Information Technology (IT) management, chargeback refers to the practice of charging the costs of IT back to the different departments and business units that use IT [1–3]. Chargeback makes service consumers aware of the costs of IT and is generally used to control escalating IT costs, improve decision making, align behavior with organizational goals, and facilitate a more effective use of IT [4]. Chargeback models can be described as having a delivery and a consumption side. The delivery side represents a producer, such as processing power or a VPN connection [2], whereas a consumption side is a department using a service or good [2,5], such as bandwidth or hardware expenditure. However, in comparison to chargeback for physical products, chargeback for intangible products is still poorly understood by many organizations [2,3] and is rarely applied to their advantage. Because of this, there is a lack of successful chargeback models that are well aligned with organizational objectives, and are clear and acceptable to all the involved stakeholders [6] .

To make the situation more complicated, organizations are adopting cloud computing. With pay-per-use pricing models, elastic scaling of resources, and the usage of shared virtualized infrastructure, the cloud fundamentally changes the economics of IT [7] .

It enables more efficient use of capital, great cost reductions, and facilitates flexibility. However, to fully leverage those benefits organizations also need to employ appropriate form of pay-per-use based chargeback, something that is seldom done in the current chargeback practices [8]. Let alone the chargeback regarding the temporary provisioning of additional CPUs and the allocation of several megabytes of RAM .

Failure to allocate cloud-based costs in a per-use manner can lead to an explosion of unnecessary consumption that can offset the cost reductions and to an inability to leverage the business advantage offered by flexible pricing [8]. The current chargeback models are oriented towards more traditional IT environments in which costs do not vary with usage. This allows for relatively simple chargeback models that are easy to understand by those being charged. With cloud computing, chargeback models become inherently more complex and opaque for the consumers. To the best of our knowledge, there are no cloud related chargeback models and the underlying factors reported in literature. Hence, there is a need to determine the factors that influence the chargeback in cloud computing environment. This leads to the formulation of the following research question: Which factors influence the chargeback of cloud services? The research in this article identifies the factors that influence the chargeback for cloud services on the consumption side. We use empirical research methods to determine the factors that influence the chargeback for the cloud services. In particular, we employ three approaches. Initially, a systematic literature review is performed to identify the factors reported in the literature. Subsequently, twenty-five semi-structured interviews with chargeback specialists are held to identify new factors for chargeback and to evaluate the findings of the systematic literature review. Finally, the findings are validated through in-depth interviews and six workshops .

From the systematic literature review, five factors are determined and from the expert interviews four additional factors are determined. The factors are consolidated to a single list of factors that influence chargeback in the cloud. To further validate the findings, we conducted two evaluation interviews and a series of focused workshops, in which chargeback experts review the findings of the research .

The overall findings of the research can be summarized with the following contributions: (i) This article provides a list of factors influencing the chargeback for cloud services as a pioneering work in the chargeback domain .

(ii) This research facilitates organizations to realize the implications of the cloud for their chargeback and how to improve chargeback in cloud services .

This article is structured as follows: Section 2 presents the theoretical background of this research, Section 3 details the research methods applied, Section 4 describes the results and Section 5 presents the discussion of the finding and the threats to validity of the research. Finally, the article concludes in Section 6 with recommendation for future work .

2. THEORETICAL BACKGROUND

In the following sub-sections, we briefly discuss the theoretical background to position this research .

2.1. Cloud computing

In this research, we have adopted the definition of cloud computing from NIST [9] as: “Cloud computing is a model for enabling convenient, on-demand network

access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction .

This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models’’. This definition provides a clear overview of what the cloud entails and provides three service models namely: (i) Software as a Service (SaaS), (ii) Platform as a Service (PaaS), and (iii) Infrastructure as a service (IaaS). Table 1 lists the characteristics of cloud computing .

These characteristics are especially interesting for cloud chargeback as they differentiate the cloud from traditional IT which in turn influences revenue - and thus cost - models .

Two high level business advantages of the cloud are relevant for understanding impact of cloud computing on chargeback. First, the cloud offers substantial cost reductions that are enabled by efficient utilization of resources due to virtualization, efficient use of capital due to pay-per-use, automation, standardization, and self-service [7,10]. Second, the cloud also offers business flexibility, because of the pay-per-use models and the characteristics near instant scalability and flexibility, and near instantaneous provisioning [10,11], which allows for accurate metering and charging of the used resources, without fixed costs. This implies that users that have peak usage, i.e. short moments in time with intensive use followed by longer periods of low use, users pay only for peak usage when they need it. This avoids significant investments in capital to be able to service those peak moments which will remain dormant most of the time as the capacity is not needed in slower periods. As the cloud allows for automatic elasticity, there should be no user involvement nor lag in up and down scaling .

2.2. Chargeback models

Chargeback models are presented under a variety of names such as chargeback systems [4,12,13], chargeback practices [2], and chargeback approaches [14–16]. However, they all tend to comprise the same concepts, without precisely defining them. According to the notion used behind these terms of chargeback [2,3,15,5, 17,18], “chargeback for cloud services” can be defined as the process of allocating the costs of the cloud services that an organization procures to the organizational units using those services. A chargeback model in general is a conceptual representation of how the costs are allocated to the organizational units using those services .

Regardless of the terminology, chargeback is mainly recognized as a means of IT governance [2,5] that enables IT cost reductions .

On the consumption side, it makes users cost-aware and results in a more cost-efficient choice of services [2,19]. On the delivery side, increased understanding of service costs facilitates more effective IT investment and provisioning decisions [4,12]. Table 2 provides various types and definition of chargeback reported in literature .

As is evidenced by the list presented in Table 2, the types of chargeback models do not have the same structure and scope .

Some of the models deal with price formation (e.g., market based pricing, cost plus charging, negotiation between provider and consumer, management set pricing), some are accounting methods (e.g., activity based costing and fixed/user costs

charging), others focus on the choice of allocation units (e.g., metered usage and tiered subscriptions), while there are also some that just provide supplementary chargeback techniques, such as notional charging .

2.3. Impact of cloud computing on chargeback

Recent research on the impact of cloud computing on chargeback shows that organizations that want to fully leverage the benefits of the cloud should apply pay-per-use chargeback practices [8,22]. Per-use charging for cloud services prevents an explosion in consumption and the subsequent increase in costs that can offset the cost reduction benefits of the cloud. Furthermore, for internal users, no pay-per-use chargeback for cloud services diminishes the business value of cloud computing. From a business perspective, pay-per-use is an advantage because costs follow value generation and demand, subsequently allowing efficient use of capital and flexibility in consumption behavior .

To sum up, the usage of chargeback for cloud computing is compelling . However, there are numerous challenges to it [23].

[TABLE 1]

[TABLE 2]

For costs of the shared virtualized infrastructure behind the cloud? What pricing models (e.g., subscriptions or some forms of metered service consumption) are most suitable for the different types of cloud services? How to charge the costs of the over-provisioned resources that are necessary to enable elastic scaling?

2.4. Related work in chargeback for cloud services

As cloud services are starting to reach a stable platform of adoption [24], cost analysis of cloud solutions are getting more and more important. These so-called “Cloud Economics” or “Clouconomics” [25] are being covered in scientific research. Talukder, Zimmerman and Prahalad [25], for instance, provide an overview of cloud economics. Several works have contributed towards the development of methods to perform cloud economics analysis, such as Martens et al. [24], evaluated the possibility of a total cost of ownership approach. Tak et al. [26] evaluated cloud adoption for in-house developed applications and Wang et al. [27] described pricing fairness of cloud computing. However, none of these works discuss chargeback. Activity based costing, which tends to be applied in cloud services has a history of over two decades, but current literature does not mention cloud chargeback services either [28]. This may be because of the difficulty of the actual metering in cloud environments, as Iyer et al. show [29]. Sekar and Maniatis [30] also note that the accountability of metering can be an issue in the cloud .

3. RESEARCH METHOD

The research described in this article applies the Design Science Research methodology, specifically the design science in information system research method by Hevner et al. [31]. The research relevance is formulated due to the fact that there are no chargeback models and/or guidelines to help organizations to design or adapt their chargeback models for a cloud environment in the scientific and management literature .

At the same time, the problem is growing with importance as large organizations, the typical users of chargeback models, are moving towards cloud computing as a way to provisioning IT capabilities .

Consequently, the need to do chargeback for a cloud environment is inevitable .

The overall research process, depicted in Fig. 1, is adopted in this research. In Phase I, a systematic literature review (SLR) is conducted to identify the chargeback factors mentioned in the literature. In Phase II, twenty-five interviews with chargeback experts are conducted to identify factors that influence chargeback and to cross-validate those with the initial list of chargeback factors aggregated from Phase I .

At the time of the investigation, there was no prior research in chargeback models for cloud services. Results from an SLR would thus only provide a theoretical basis that is inspired or influenced by traditional (IT) chargeback models. To extend this theoretical basis, and discover non-documented factors from practices, expert interviews (Phase II) are held. The expert interviews not only added new factors that were not found in the literature, but also facilitated the successful validation of the findings of the SLR. Phase II results in an extensive list of factors that influence chargeback: the factors from both the SLR and the interviews .

Finally, in Phase III, the identified chargeback factors are empirically validated through focused workshops and two in-depth evaluation interviews .

In the following sub-sections, each phase is explained in detail .

3.1. Phase

I In Phase I, an SLR is conducted, as suggested in [32,33] to identify the factors that influence the chargeback explicitly mentioned in the literature. Initially, a search query was developed by combining the key terms and their synonyms. Table 3 depicts the key terms used to develop the search query .

The search string was then constructed using Boolean “AND” to connect the keywords and Boolean “OR” to allow synonyms and word variants of each keyword. The resulting search string is then executed in digital libraries/indexing services such as Google scholar, ISI web of knowledge .

The abstracts of the list of papers retrieved were then studied to find the relevance. We employed exclusion criteria to filter out unnecessary papers.

[FIGURE 1]

[TABLE 3]

For instance, more than 50% of the retrieved papers were related with chargeback in the banking field, which is out of our scope .

The initial set of papers is termed as “initial articles”. Additionally, references of the initial articles were followed from the discovered papers to increase the number of relevant works. Subsequently, the list of look-up terms was expanded to include synonyms of the term chargeback and additional search iteration was conducted.

Queries on chargeback for cloud services were showing a rather limited set of results and all found works were thoroughly reviewed. On the other hand, broader terms such as chargeback model or cloud types were listing thousands of results .

To cope with the large amounts of data, scientific literature found by scholarly search engines were reviewed only until a theoretical saturation point was reached. Constant Comparison Analysis (CCA) [34], which is used to create knowledge from the data

source by avoiding subjective interpretation (i.e., interpretation of the data in accordance with the research objectives) technique, was employed to determine the saturation point. This is the point at which reviewing more works is hardly adding any new relevant information [35]. For cloud computing concept search terms, a saturation point was reached at about 100 articles. For chargeback concept search terms, saturation was reached on average at 250 articles. The higher number was due to the need to investigate the charging topic in greater depth, and also due to the usage of the term chargeback in the banking field to refer to the process or reversing a bank transaction [36].

The analysis of SLR resulted in five factors that influence that chargeback, namely: (i) Accuracy, (ii) Cost of Costing, (iii) Transparency and Understandability, (iv) Controllability, and (v) Fairness.

These factors are described in the Results section (Section 4).

3.2. Phase II

In Phase II of the research process, twenty-five semi-structured interviews are conducted, aimed at identifying an extensive list of factors that influence chargeback success. The results from these interviews are evaluated with the findings of the SLR

[TABLE 4]

The field experts interviewed come from IBM and IBM's client and partner network, and have a variety of backgrounds and positions to ensure the generalizability of the findings. Five types of stakeholders concerning chargeback were identified: executives and board members, chargeback specialists, charges recipients, IT specialists, and consultants. Despite the differences in background, the interviewees also shared some common characteristics. They are from Dutch origin and work for seven different large organizations in The Netherlands. In Table 4, details of each organization in terms of industry and size are provided.

The objective of the interviews is to uncover the factors of chargeback in cloud services. The respondents have the opportunity to freely discuss the topic and the researcher used this freedom to clarify the questions. The need for free bi-directional communication precluded the usage of questionnaires or formally structured interviews [37,38]. Semi-structured interviews are preferred over unstructured interviews to allow comparison between the responses of the different participants and to keep the interviews focused [38].

A two-page interview protocol was developed to impose a common structure on all interviews. On top of an introduction and a wrap-up, it contained sections about the background and experience of the interviewee, unguided questions, guided questions, and cross-evaluation. In the introduction, the researcher and the research topic were introduced, while the background check was used to determine the stakeholder category of the interviewee. In the guided and unguided sections the respondents are asked questions such as: "What characteristics of a chargeback model do you find relevant for a chargeback model for cloud services?" and "How do these characteristics influence the model?" No directions were provided by the interviewer, and subsequent questions were asked mainly to clarify what the interviewee meant by a certain concept. In the guided part of the interview, the concepts identified from literature were introduced and the respondents were asked to comment on them. They were also queried for any additional factors that came to their mind.

Finally, the cross-evaluation phase was used to validate the findings. The wrap-up was used to verify whether the researcher had properly captured all the points made by the interviewee .

Furthermore, the twenty-five interviews are used to empirically cross-evaluate the findings for correctness and completeness .

Since it is impossible to determine how many interviews would be necessary to reach an “extensive” list a priori, a theoretical data saturation principle is applied to empirically determine the required number [35,39]. Interviews are conducted until the moment a data saturation point is reached, i.e. the list of discovered forces started to converge and no additional factors are brought up in subsequent interviews. At this point it was deemed that the probability for additional findings from questioning more respondents was too low to justify the required research efforts. The list was considered complete at this point .

Correctness is evaluated by asking the respondents whether all the discovered factors could influence the chargeback of a cloud service. Completeness is determined by asking the participants to comment on the comprehensiveness of the compiled list. This list started at first interview with factors discovered from the literature review and the factors that are mentioned in that interview. In each subsequent interview, the final list from the previous interview was extended with the factors derived from that interview. Since the latter is always larger than the list identified by each individual interviewee, the participants mostly made positive comments on comprehensiveness . The saturation point is reached at the twenty-fifth interview: after interview twenty-one, each discovered factor had been mentioned at least four times, and no new factor had been identified since the twelfth interview .

Based on the analysis of the interview, four additional factors were identified, namely: (i) Measurability, (ii) Predictability, (iii) Accountability, and (iv) Comparability. These factors are described in the Results section (Section 4) .

3.3. Phase III

Phase III evaluates the results from the earlier phases. Two approaches are utilized: in-depth evaluation interviews and focused workshops, as depicted in Fig. 1 .

The two interviews, lasting 2 hrs in a semi-structured setting, are held with a chargeback expert each. These experts are selected as they are familiar with all concepts of chargeback, and have extensive chargeback experience. Both are responsible for the design and operation of chargeback models in more than one enterprise, including multi-nationals and organizations in the public sector .

Both interviewees are Dutch and at the time of the interview were working in organizations in the Netherlands, one in the public sector, and the other one in the financial industry .

In each interview, the factors gathered from the earlier phases are shown, including the relations between one another, to analyze how different factors affect chargeback success in a cloud setting .

[FIGURE 2]

Finally, participants were asked to make general comments about the model and to provide further improvement recommendations .

Additionally, six workshops are conducted featuring 32 participants in total. The number of participants (with the exception of the researcher) in each workshop varied between three and ten participants, allowing all participants to actively

participate in the discussions. Each workshop lasted between 60 and 90 min, with a tendency for the larger workshops to last longer .

Twelve of the participants in the workshops had previously taken part in Phase II. The other twenty participants were new to this research. The nine organizations in which the participants work have extensive experience with chargeback. Three participants are senior professionals from the French and Belgian IBM offices who took part in our fifth workshop through a telepresence conferencing system, while the rest are Dutch participants. The details of the role of the workshop participants is depicted in Fig. 2 .

The participants of the evaluation phase are asked to provide feedback on correctness and completeness of the identified factors, and to evaluate the relationships between the factors. The respondents agreed on the correctness and completeness of both the factors and the relationship between the factors and effectiveness and acceptability .

4. RESULTS

The findings of the SLR, the interviews of Phase II and the crossevaluation in the two interviews and workshops of Phase III are explained in the following sub-sections. First, factors from the SLR are explained, and based on analysis of those factors inferences are discussed. Second, the factors from the analysis of the interviews are discussed, including the inferences between those factors .

Third, all those factors and their inferences are evaluated and consolidated to provide a complete overview of the findings .

4.1. Factors derived from the structured literature review

A review of the literature reveals five factors that are explicitly mentioned in the literature. Furthermore, a few more factors are inferred through detailed analysis. The five concepts that are explicitly recognized in the scientific literature to influence chargeback are: accuracy, cost of costing, transparency and understandability, controllability, and fairness. These factors are described below. (i) **Accuracy**: Accuracy is a factor influencing chargeback success that is universally recognized in academia [1,12,5,40,41]. This is a property of a chargeback that describes to what extent the charges allocated to an organizational unit for each service accurately approximate the actual costs incurred by the organization for delivering the service to the unit. The better the costs are approximated with the charges, the higher the accuracy of the chargeback .

High accuracy has a profound two-fold effect on the chargeback. First, it is conducive to realizing cost reductions, because it bases both provisioning and consumption decisions on actual costs [12,40]. Second, it makes chargeback more acceptable to the involved stakeholders, because it motivates the correctness of the charges and prevents attempts for overthrowing the model based on low accuracy arguments [5] .

(ii) **Cost of costing**: While high accuracy has positive impact on the chargeback model, it might be expensive to achieve. The term cost of costing is used to quantify the costs of operating chargeback and accounting models, and to offset these costs against potential benefits [21,42,43]. Operating a chargeback model is associated with significant design, implementation, labor, and IT systems costs. Therefore, chargeback designers should take the costing factor into account in order to develop chargeback models of which the benefits outweigh the related expenses .

Cost of costing has a negative impact on chargeback, because those costs offset the cost savings realized through chargeback. However, costs incurred by, for example, investing in improved accuracy can positively influence chargeback as accuracy has a positive influence on chargeback models (as described above) .

Therefore, the design of a chargeback model should be optimized to balance the cost of costing with factors that positively influence chargeback model such as accuracy [42,43] .

In the context of IT chargeback, cost of measurement is synonymously used to refer to the costs related to the precise measurement and allocation of the costs of IT services. The more elaborate the accounting methods used, the higher the accuracy and the cost of costing. Another synonym is the cost of errors which refers to the costs of missed benefits or errors made due to low accuracy of the chargeback model [42] .

For instance, such errors might lead to sub-optimal provisioning or consumption decisions. High cost of errors is a product of low accuracy. Therefore, the lower cost of measurement that accompanies low accuracy is negatively compensated by high cost of errors. In line with this reasoning, the design of a chargeback model should be optimized to balance between costs and accuracy in order to reach an optimal balance that best suits the organizational goals [42,43] .

(iii) **Transparency and understandability**: Researchers who investigate organizational behavior in relation to chargeback note that a chargeback should be understandable and transparent to the involved stakeholders [12,16].

Understandability describes whether the recipients understand what they are charged for i.e., do they understand the units of service and their prices? Charges for cloud services based on a user subscription are more understandable than charges based on utilized CPU cycles. Transparency, on the other hand, describes whether the involved stakeholders understand how the charges are formed. Therefore, Transparency characterizes the capability of stakeholders to comprehend the chargeback model and the opportunity to enjoy non-obscure chargeback processes, while understandability deals with the capability to understand the charges, the end product of applying the chargeback model .

Transparency and understandability are closely related and discussed together in the literature [12,16]. Understandability is necessary in order to have transparency, because lack of understanding of the charges themselves (low understandability) leads to inability to comprehend the process of forming the charges (low transparency). The major effect of transparency and understandability on chargeback is related to obtaining stakeholders' buy-in. Low transparency and understandability lead to resentment of the chargeback [12,16] .

(iv) **Controllability**: Nolan [44] uses the Controllability concept to denote to what extent consumers are in control of their IT costs. Chargeback models that enable users to have impact on IT bills by changing consumption behavior have high controllability .

On the contrary, if users cannot influence IT bills, then controllability is low . Controllability has profound impact on chargeback. On the one hand, it is essential to enable cost reduction opportunities on the consumption side, because it allows managers to reduce their IT bills by changing consumption behavior. On the other hand, controllability influences the users' acceptability for the chargeback model [44]. Low controllability leads to resentment, because chargeback is perceived as an

unnecessary overhead that does not benefit managers, while high control allows them to realize cost reductions and accept the chargeback model .

(v) **Fairness**: Fairness is another concept investigated by chargeback researchers [17,41]. Those authors dub Perceived Fairness (PF) as “the key to chargeback systems effectiveness”. In their works, Allocative Fairness is used as a synonym to accuracy, while PF refers to user’s perception of how fair the method is. It is unclear, however, whether in this case the term fair can be used as a synonym to the words “just” or “unprejudiced” .

Managers’ self-interest and opportunism lead to perception of high fairness only when the chargeback model is consistent with their goals (e.g. allows them to get a higher bonus because their profit increases due to lower IT costs). On the other hand, if a chargeback model is highly accurate, or “allocatively” fair, but leads to higher IT costs for a manager, he/she might be likely to perceive the model as unfair. It may be inferred that Fairness and transparency are synonymous, however Hufnagel and Birnberg [17,41] do not mention transparency in their work .

The term PF could be considered a misnomer, because of its slightly contradictory meaning to the word fair (just, unprejudiced) .

Nevertheless, it is a proper concept to describe stakeholders’ attitude towards the chargeback model and whether they are likely to accept (high PF) or reject it [17,41] Table 5 summarizes the factors identified from the literature study .

Inferred factors: A close inspection of the above factors influencing chargeback reveals that their impact is realized in two separate ways. The examined factors either affect the effectiveness of the chargeback model [1,12,19,44,45], or the acceptability to the stakeholders [12,16,17,41,44]. These two aspects of chargeback are used in the literature (see above) to describe how accuracy, cost of costing, transparency and understandability, and controllability influence chargeback. Following the findings of the SLR, Fig. 3 visualizes how acceptability and effectiveness can be recognized as two high level dimensions that explain how the other forces influence chargeback. Solid arrows visualize a positive relationship, while the dashed ones depict a negative relationship. Fairness is merged into acceptability and is discussed further below .

(i) **Effectiveness** determines whether the chargeback stimulates the desired chargeback results, such as cost awareness and control, behavior steering, minimization of internal conflicts, more competitive costing and pricing .

[TABLE 5]

(ii) **Acceptability** describes whether all the involved stakeholders find the chargeback acceptable. High acceptability means that stakeholders agree and are supportive of the chargeback process, while low acceptability stands for the lack of support and results in resentment of the chargeback .

In Fig. 3, fairness has been merged into acceptability because the latter excellently captures the meaning implied by Hufnagel and Birnberg [17,41]. The substitution is appropriate, because Hufnagel and Birnberg equate low perceived fairness to the resentment of the chargeback, while high perceived fairness is considered equivalent to the acceptance of the chargeback .

4.2. Factors derived from the interviews

The results of the interviews confirmed the relevance of the factors discussed in Section 4.1 and depicted in Fig. 3. Furthermore, four additional factors were found to

be conducive to chargeback, and to be especially relevant in a cloud environment. These are Measurability, Predictability, Accountability, and Comparability . Table 6 displays how the respondents mentioned each factor during the interviews. Subsequently the newly discovered factors are explained. Fig. 4 shows the factors newly discovered from the interviews .

(i) **Measurability**: Is a concept that is mentioned mostly by the interviewed IT specialists (43%) who have in-depth experience with cloud computing. The term is used to refer to the degree of ease with which the chargeback allowed measuring the amount of consumption units used and the ability to allocate those consumed units. For example, if a certain service is charged based on completed transactions, it should be possible to count the amount of completed transactions over a charging period and by whom they are executed. High measurability implies that it is possible to measure usage without highly specialized or custom made metering systems, while low measurability requires such technology .

The interviewees motivated the importance of measurability, by explaining that it was difficult to measure usage in the shared virtualized infrastructure in the cloud and that there are still a number of technical limitations on metering systems, especially when it came to charging units such as CPU cycles or memory usage. Yet, as one respondent remarked “(. . .) *usage based charging [is] possible only through measuring*” .

Regarding the impact of measurability on chargeback, the interviewees suggested that it positively influenced both acceptability and effectiveness. From an effectiveness perspective the capability to gather detailed usage data allowed for effective decision making and more accurate charges. On the other hand, high measurability reassured the stakeholders in the accuracy of the chargeback, consequently improving acceptability .

(ii) **Predictability**: It describes to what extent chargeback recipients are able to predict future bills. Predictability becomes far more significant in a cloud environment, because pay-per-use and elastic scaling of resources could lead to fluctuating bills that tend to bother budget minded managers. The latter do not mind bills that are smaller than expected, but it turns out that they are afraid of bills that greatly exceed expectations. Despite these concerns, the interviewees commented that predictability did not impact the effectiveness of the chargeback . It affected only the acceptability, as unpredictability led to resentment towards the chargeback .

(iii) **Accountability**: Is the extent to which recipients are able to verify the correctness of the bill. The easier it is to verify the charges, the higher the accountability, and consequently, the higher the acceptability. Fixed subscription fees on a per user basis offer the highest accountability, because managers are able to multiply the number of people in their team by the subscription fee and verify the correctness the bill. However, accountability decreases when metering of IT resource consumption is involved, such as transactions made and bandwidth consumed, because that requires higher technical competence from managers and access to specialized IT systems .

The respondents motivated the relevance of accountability by explaining that they often witnessed disputes, sometimes well-grounded, over the correctness of the bill. One interviewee summarized the general opinion as follows: “*The lower the accountability, the more often you see disputes over the bill*” .

(iv) **Comparability**: It is the degree to which consumers are able to compare the prices of internally provisioned services to similar or equivalent services offered on the market. The standardization characteristic of the cloud leads to the availability of highly similar standardized services available on the market .

Therefore, in a cloud environment it becomes much easier to benchmark IT costs on a per-service basis. Yet, this still depends on how the chargeback is designed, and whether internal service pricing follows the pricing patterns of public providers. If the chargeback is designed in a way that allows easy price comparison, then comparability is high. This results in cost reductions on the delivery side, because internal users who are able to compare prices start to exercise pressure for more efficient IT delivery on the IT department .

4.3. Consolidated factors from the SLR and the interviews

As described in Section 4.1, the factors resulting from Phase I are evaluated during the semi-structured interviews in Phase II .

This happened as interviewees would mention those factors individually or because the interviewer explicitly evaluated them .

Nonetheless, the interviewees confirmed the factors found in the literature (depicted in Fig. 3) and the relations between the factors and the two high-level dimensions: Acceptability and Effectiveness .

Furthermore, the results from the interviews extended those of the SLR by providing 4 additional factors not found in the literature (see Fig. 4). As the saturation method assured recurring factors and excluded isolated artifacts, the factors presented in Fig. 3 can thus be merged with the factors presented in Fig. 4 .

The consolidated factors for the chargeback in cloud service are depicted in Fig. 5. This consolidated model is presented to the experts in the workshops and during the two in-depth interviews. Specific questions are asked regarding the validity of the factors and the relationships presented. The participants could not add factors

[FIGURE 4]

[TABLE 6]

even though they were specifically asked to think whether the presented findings reminded them of factors not listed. On the contrary, affirmative comments stated that the list contained factors that were relevant for a variety of stakeholders. One of the interviewees summarized that *“a great thing about the list [is] that it contain[s] the concerns of different types of people. [He] would not have thought of all those”* . The cross-validation shows that the compiled list of characteristics influencing chargeback model could be considered as comprehensive not only from a theoretical, but also from an empirical perspective. Moreover, the evaluation confirmed the expectations that the attitude of different types of stakeholders towards chargeback might differ, depending on their expertise and professional and personal interest. Finally, the interviewees approved the findings and positively commented on their value as a foundational work that can be used to guide chargeback design decisions. Table 7 summarizes these results, describing each factor and its relation to Acceptability and/or Effectiveness .

Below the final factors are described in detail .

Predictability was added to the model as a factor influencing acceptability as it was mentioned to be a determining factor in payper- use services. It thus influences acceptability of these services can influence not only chargeback model but the acceptance of the complete service offering. New methods are being developed for predicting cloud costs on the architecture level [46–49] and on a managerial level [50,51] .

Transparency and understandability was mentioned by 48% of the respondents, with the majority being Chargeback specialists. It influences the acceptability of a chargeback model as it measures the understanding users have of the measured units in the model and how they reflect the costs involved. It has an interesting relation with accuracy: respondents said that, for example, measuring network bandwidth is very accurate, but users tend to find bandwidth less comprehensible and transparent. This of course differs per user base .

Accountability prevents disputes in the chargeback practice, providing the ability for users to verify the bills incurred, and is therefore clearly a factor influencing acceptability. Fixed subscription fees on a per user basis offer the highest accountability, because managers are just able to multiply the number of people in their team by the amount of the subscription fees and verify the correctness of the bill. However, when metering the consumption of IT resources is involved, such as transactions made and bandwidth

[FIGURE 5]

[TABLE 7]

consumed, accountability decreases. Verifying the bill in such cases requires the capability to verify the amount of used consumption units .

Measurability adds transparency. The better the measurability of services, the more effective the chargeback is, thus influencing both effectiveness and acceptability. As certain costs are hard to measure, such as cache usage [52], high level methods of metering can be used to improve measurability. These could be based on SLAs [53–56], or so-called Smart Metering as introduced by Singh and Vara [57], or take a higher level unit as base for measuring such as the amount of users registered .

Accuracy was mentioned by 21 interviewees, 84% of the total. This is in line with the literature that mentions accuracy universally .

Although multiple respondents referred to accuracy as precision, correctness or fairness, unanimous they stated it covers the extent to which the charges allocated to an organizational unit for each service accurately approximate the actual costs incurred by the organization for delivering the service to the unit. As it influences the ability of an organization to promote cost-awareness and cost reduction and to avoid subsidization due to the accurate results of cost allocation, it directly influences effectiveness of a chargeback .

As the correctness is influenced by accuracy, stakeholders will trust the billing and thus accuracy improves acceptability. The accuracy differs per measuring units, since some units are easier to measure than other (cf. measurability). Choosing the correct measuring unit is of paramount importance to reach a high level of accuracy .

Controllability is both being discussed in the literature as by eight respondents in the interviews. It determines whether chargeback recipients are in control of their IT costs. Charges recipients should be able to understand how specific changes in usage

behavior would influence their bill. This is essential for the effectiveness of the model, because this is the driving force most conducive to changes in consumption behavior and subsequent to reductions in costs .

The interviewees explained that if changes in usage were immediately reflected in IT bills, chargeback recipients found the chargeback method more acceptable and as less overhead, because they saw the change as a direct consequence of their actions. This further stimulated them for more changes in consumption behavior and also led to an increase in effectiveness. Current research investigates cost-awareness in cloud computing, such as the work of Han et al. [58], explaining the cost-awareness elasticity to keep costs in control without losing elasticity of the infrastructure. Similarly, the work of León, Trinh and Navarro [59] investigates local resource congestion prevention .

Comparability only affects the effectiveness of a chargeback model as it describes the measure of which different cloud services are comparable to one another. This effectively means that if a service can be easily compared against its competitors, it will make the chargeback model more effective. One of the characteristics of the cloud is standardization and this leads to the availability of (near) similar and standardized services available publicly and on demand. In a traditional IT environment where services are customized and rarely publicly provisioned in a standardized manner, it is difficult to benchmark IT costs on a per service basis. It is usually only possible to compare IT budgets in the same industry .

However, due to the transition to cloud computing it becomes possible to easily compare prices of standardized services. Yet, this still depends on how the chargeback is designed, and whether internal service pricing follows the pricing patterns outlined by public providers. If the chargeback is designed in a way that allows easy price comparison, then comparability is high .

High comparability allows users to compare prices to external providers and exercise pressure on the internal IT department for more cost-effective service provisioning. In turn, this results in an increase in effectiveness due to cost reductions on the delivery side. Breskovic, Altmann and Brandic [60] present an overview of electronic markets and how products can be standardized of such markets .

Finally, **Cost of Costing** influences **Effectiveness** negatively. As a property of a chargeback, the Cost of Costing refers to the expected costs of chargeback design, implementation, and operation .

Such costs offset part of the cost reduction benefits. This decreases effectiveness and lowers the usefulness of chargeback as a cost reduction tool .

5. DISCUSSION

A closer analysis of the results provides additional insights about the chargeback landscape. Section 5.1 describes the difference between chargeback for cloud services and chargeback for traditional IT environments and Section 5.2 describes the different perspectives from the various stakeholders of chargeback models . Finally, Section 5.3 describes the threats to validity of this research .

5.1. The influence of cloud services on chargeback

The interview results show that the current chargeback literature is almost oblivious in regards to cloud computing and that further research in the field is necessary. This is suggested by the fact that all the factors identified through the SLR hold for both traditional IT and cloud environments. However, the interviewees put forward the

additional factors (measurability, predictability, accountability and comparability) for chargeback in the cloud .

- Measurability pertains to cloud computing as new units of measurement such as CPU time and memory usage are introduced .

Some of these are harder to measure than others .

- Predictability is introduced because new units of measurement (CPU time, memory usage, etc.) can spike at unforeseen moments, possibly resulting in high costs. The level and occurrence of these spikes determine the predictability measure. ·

Accountability is important in cloud chargeback because the measurability can be fine grained (e.g. megabytes of RAM) and it can be difficult to determine who is using which resources .

- Comparability pertains to cloud chargeback as in the cloud many service providers offer a comparable service. This allows users to compare these services and select the best fit. As services in the cloud are mostly virtual, comparing and switching are relatively easy .

Cloud computing brings highly abstracted, shared resources at near instantaneous scalability and flexibility to IT environments [61] .

As found from the literature research, current chargeback models do not account for this. The aforementioned factors extend chargeback to account for the characteristics that cloud computing brings to the traditional IT environment .

These findings contribute to the understanding of how cloud impacts chargeback models and assist organizations that need to implement cloud-specific chargeback models as part of the transition to the cloud .

5.2. Stakeholder motivations

An intriguing conclusion that can be drawn from Table 6, is that different types of stakeholders tend to recognize different factors .

This can be used to better explain the chargeback landscape, the rationale behind the behavior of individual stakeholders, and thus to remind chargeback designers to address the unique concerns of all stakeholders .

Some notable patterns that can be identified from the results are the following: ·

Chargeback specialists do not seem to be concerned with the cost of costing as only one out of six specialists mentioned it .

- Executives on the other hand do not seem concerned with accountability or controllability but 50% do mention cost of costing .

- Measurability is a force recognized mainly by IT specialists, all consultants who did mention it are chargeback specialists as well .

- Predictability tends to be mainly a concern of chargeback recipients (44% mentioned it) and IT specialists (79%), but not of chargeback specialists (only one, 17%) .

- Comparability is mentioned by 4 respondents (16%) of whom none are recipients .

The fact that chargeback specialists did not mention the cost of costing at first sight is rather perplexing, because these professionals usually have strong accounting backgrounds and their task is to keep an overview of the involved financials. One plausible explanation that was supported during the interviews is that the negative impact of cost of costing is largely offset by the benefits arising from investing in improving other factors. Yet, this view seems to contradict previous research [42] .

Opportunism might provide an explanation why chargeback specialists tend to avoid recognizing the importance of cost of costing .

First, higher cost of costing could be expected to provide more work for the chargeback specialists and to lead to increased job security .

Second, complex and expensive chargeback model is more challenging and interesting from a practitioner's point of view, and thus could be regarded as a source of professional satisfaction .

This opportunism might also be a reason for the executives to mention cost of costing, accuracy and effectiveness on the top of their list (50% of the executives mentioned these), but not to mention controllability and accountability, the methods to provide proof of the three factors (costing, accuracy and effectiveness). This could also indicate that executives look at the results, which influences their pay, without being concerned how these results are made. The relative lack of interest in predictability and measurability from chargeback specialists further suggests that chargeback for cloud computing is poorly understood, and that the current practices fail to leverage pay-per-use. If it were otherwise, then the concerns of IT specialists about both predictability and measurability would have had proliferated also to chargeback specialists .

While comparability could help chargeback recipients in their work and reduce costs of their teams, none mentioned comparability as a factor. This is striking as they are the user group that should be able to leverage comparability the most. Consultants, IT and chargeback-specialists did mention comparability. This could indicate the aforementioned opportunism, as comparability involves an overhead to implement it, thus more effort by those parties .

On top of information about the current chargeback practices, the discussion from Phase III also confirms that chargeback involves a lot of political play and opportunistic behavior. For example, chargeback specialists might be inclined to adopt heavy chargeback processes both for job security and out of professional interest. Chargeback recipients might show resentment for the charging practices if the latter put them at a disadvantage, and external consultants and solution providers might have incentives to push through expensive technology and services. Therefore, chargeback designers should be well aware of possible opportunistic behavior and should take into account the positions of all stakeholders, as well as all the factors influencing the chargeback .

Also, the rate of change in this field is high, i.e. these experts have to change all their smart rules (and the software systems that depend on those rules), which is both challenging and costly .

5.3. Threats to validity

The findings reported in this research are the results solely based on empirical observations. The result of this research is a preliminary step towards documenting factors for a successful chargeback model in a cloud environment. Hence, the findings are suggestive rather than conclusive. Due to empirical nature of the research approach, the research is subjected to bias. In the following sub-sections, the most relevant threats to validity are discussed with the countermeasures that have been undertaken to minimize the threats .

Reliability (construct) validity [62] refers to the extent to which concepts being studied are operationalized and measured correctly .

In our research, we have followed specific guidelines to conduct an SLR to find out the factors from the literature, which reduces the threat to reliability validity.

Furthermore, the interviews were also conducted with a standard guideline which reduced the chances of misunderstanding. In addition, we have used data saturation principle to support the validity of the findings in both the SLR and the interviews, thereby minimizing the threats to reliability (construct) validity .

Generalizability (External) validity [62] refers to the extent to which the findings can be generalized. In our research, the potential threat to generalizability is due to the fact that the research was conducted within the scope of IBM, IBM's client and partner network. Nevertheless, there are a number of reasons that allow for a reliable generalization of the results .

First, no evidence for significant differences between chargeback practices in European and non-European environments was found in the literature. Second, the results from this work confirm the findings from the predominantly US based literature used for this research. Third, the majority of the newly discovered criteria was mainly related to the technology characteristics of the cloud and is globally valid. Finally, the findings are motivated by location independent arguments why certain forces impact success. Due to all those reasons, it can be considered relatively safe to assume that the findings are valid at a global level. Another potential threat to generalizability is that the results rely mostly on IBM's chargeback expertise, because more than half the respondents were affiliated with IBM. However, the aim of this research was to provide an extensive set of factors influencing chargeback success, rather than to rank the different factors in importance or to describe how chargeback concerns differed between organizations. Therefore, any corporate bias was prevented and expanding the list of interviewees with additional respondents with broad chargeback and cloud experience from IBM helped to create a complete list of influential chargeback factors. Also, the diverse industrial categories of the interview participants are a countermeasure to the threats to generalizability .

6. CONCLUSION

This research provides an overview of the factors that should be taken into account in the process developing a chargeback model for cloud computing. Eight factors are empirically validated which influence chargeback concerning cloud services – accuracy, cost of costing, transparency and understandability, controllability, measurability, predictability, accountability, and comparability – and their impact on chargeback models is explained through two higher level variables, acceptability and effectiveness. The findings suggest that to develop a successful chargeback model, chargeback designers should try to optimally balance between the presented dimensions and should address the concerns of the different stakeholders .

The findings can promptly assist organizations that need to implement cloud-specific chargeback models as part of the transition to the cloud. Rather than summarizing the accumulated chargeback experience once the transition is in a more developed stage, this research provides timely empirical findings that can enable immediate actions and future research that could further facilitate the move to the cloud .

This research contributes to the understanding of the impact of cloud computing on chargeback. The findings can be leveraged by researchers to develop cloud-friendly chargeback models and methods that assist chargeback designers to create such models .

This work empirically confirms the findings of a large body of literature, and extends their work with additional factors specific for chargeback in cloud computing. Finally, it enables further research on the development of more cloud-specific chargeback models that can further improve how organizations benefit from chargeback models in a cloud computing setting .

In times of common views that IT investments do not pay back [63,64], leveraging the findings could help improve the effectiveness of delivering and using IT in contemporary organizations .

To start with, cloud adopters can benefit from this work by implementing chargeback models that can help them leverage the cost advantages and business flexibility of the cloud and further reduce IT costs. Moreover, the results can be useful to cloud providers and providers of chargeback systems and services. Cloud providers can benefit by more knowledge about the chargeback requirements of their clients in order to offer competitive pricing schemes. Vendors of chargeback software can update their products to better support cloud-specific chargeback models. Finally, providers of chargeback services can also use it to convince their clients of the value of external help in chargeback design and implementation .

Several directions have been identified as of future work. To start with, this study could be replicated with different companies and in different geographic areas to explore whether there are additional factors that influence chargeback in cloud services in those environments and thus, contributing towards the generalizability of the findings. However, the findings presented here enable further research with far more significant contributions. An interesting future work is to quantitatively analyze the effects of the individual factors on chargeback factors leading towards the ranking of the factors based on importance and allow for an investigation why certain factors in chargeback are effective and what can be done to effectively improve chargeback. It would be interesting to see if models that exploit non-dedicated resources for cloud/grid computing [48,65–67] could be integrated with chargeback models. This would allow for a synergy between economic and technical aspects which would be compelling, but also integrating just the economic perspectives of both non-dedicated resource exploitation with chargeback modeling would be an interesting case as it potentially could show the economic value of both models. It would also be interesting to relate the identified factors with the service level agreements of the cloud services and to identify how this would impact the chargeback of the cloud services .

Finally, the findings of this research should be validated with a case study so as to provide an experimental evaluation result .

REFERENCES

- [1] F. Bergeron, Factors influencing the use of DP chargeback information, *MIS Q.* 10 (1986) 225–237 .
- [2] J. Gerlach, B. Neumann, E. Moldauer, Determining the cost of IT services, *Commun. ACM* 45 (2002) 61–67 .
- [3] C.A. VanLengen, J.N. Morgan, Chargeback and maturity of IS use, *Inf. Manage.* 25 (1993) 155–163 .
- [4] D.H. Drury, The dialectic of IT chargeback systems, *Int. J. Technol. Manage.* 14 (1997) 496–512 .
- [5] J.W. Ross, M.R. Vitale, C.M. Beath, The untapped potential of IT chargeback, *MIS Q.* 23 (1999) 215–237 .
- [6] T.D. Oleson, Price of precision, in: *CIO Magazine*, 1998, pp. 34–38 .

- [7] R. Harms, M. Yamaritino, The economics of the cloud, in: Microsoft, 2010, pp. 1–21 .
- [8] H. Stefanov, How to charge for the cloud? Towards a structured view of successful IT chargeback models for Cloud services, in: Utrecht University, Utrecht, The Netherlands, 2011 .
- [9] P. Mell, T. Grance, The NIST definition of cloud computing (v. 15), in, National Institute of Standards and Technology, Gaithersburg, Maryland, US, 2009, pp. 1–2 .
- [10] IBM, Smarter planet: Using cloud computing to deliver innovation and efficiency, in: IBM (Ed.), IBM Corporation, New York, 2010 .
- [11] M. Armbrust, I. Stoica, M. Zaharia, A. Fox, R. Griffith, A.D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, A view of cloud computing: Clearing the clouds away from the true potential and obstacles posed by this computing capability, *Commun. ACM* 53 (2010) 50 .
- [12] D.H. Drury, Assessment of chargeback systems in IT management, *Informatica* 38 (2000) 293–313 .
- [13] C. Lin, System for charging computer services, *J. Syst. Manage.* 34 (1983) 6–10 .
- [14] J.V. Bon, Best Practices: Introduction to ITIL, Van Haren Publishing, Zaltbommel, The Netherlands, 2007 .
- [15] M. Iqbal, M. Nieves, *Service Strategy Book*, 2007 .
- [16] B. Stiller, G. Fankhauser, B. Plattner, N. Weiler, Charging and accounting for integrated Internet services—state of the art, problems, and trends, in: *INET'98*, Geneva, Switzerland, 1998, pp. 1–9 .
- [17] E.M. Hufnagel, J.G. Bimberg, Perceived chargeback system fairness: A laboratory experiment, *Account. Manag. Inf. Technol.* 4 (1994) 1–22 .
- [18] C. Finden-Brown, C. Long, Introducing the IBM process reference model for IT: PRM-IT sequencing the DNA of IT management, in: IBM Global Services, Armonk, New York, vol. 3, 2008 .
- [19] D.H. Drury, Conditions affecting chargeback effectiveness, *Inf. Manage.* 5 (1982) 31–36 .
- [20] N. Abdat, M. Spruit, M. Bos, *Software as a service and the pricing strategy for vendors*, Business Science Reference, Hershey, New York, 2010 .
- [21] A.A. Atkinson, R.S. Kaplan, S.M.S. Young, *Management Accounting*, Prentice Hall, New Jersey, US, 2003 .
- [22] C.S. Yeo, S. Venugopal, X. Chu, R. Buyya, Autonomic metered pricing for a utility computing service, *Future Gener. Comput. Syst.* 26 (2010) 1368–1380 .
- [23] H. Stefanov, S. Jansen, R. Batenburg, E. van Heusden, R. Khadka, How to do successful chargeback for cloud services, in: *Economics of Grids, Clouds, Systems, and Services*, Springer, Phafos, Cyprus, 2012, pp. 61–75 .
- [24] B. Martens, M. Walterbusch, F. Teuteberg, Costing of cloud computing services: A total cost of ownership approach, in: *Proceedings of the 45th Hawaii International Conference on System Sciences*, IEEE, 2012, pp. 1563–1572 .
- [25] A.K. Talukder, L. Zimmerman, H.A. Prahald, Cloud economics: Principles, costs, and benefits, in: N. Antonopoulos, L. Gillam (Eds.), *Cloud Computing: Principles, Systems and Applications*, Springer, London, London, UK, 2010, pp. 343–360 .
- [26] B.C. Tak, B. Urgaonkar, To move or not to move: The economics of cloud computing, in: *Proceedings of the 2011 Conference on Hot Topics in Cloud Computing*, Portland, OR, 2011 .
- [27] H. Wang, Q. Jing, R. Chen, B. He, Z. Qian, Distributed systems meet economics: Pricing in the cloud, in: *Proceedings of the 2nd USENIX Conference on Hot Topics in Cloud Computing*, USENIX Association, Boston, MA, 2010, p. 6 .
- [28] G. Ooi, C. Soh, An activity based costing approach to systems development and implementation, in: *International Conference on Information Systems, ICIS'98*, Helsinki, Finland, 1998, pp. 341–345 .
- [29] G. Iyer, A. Pazgal, Internet shopping agents: Virtual co-location and competition, *Mark. Sci.* 22 (2003) 85–106 .
- [30] V. Sekar, P. Maniatis, Verifiable resource accounting for cloud computing services, in: *Proceedings of the 3rd ACM Workshop on Cloud Computing Security Workshop—CCSW'11*, ACM Press, New York, New York, USA, 2011, p. 21 .

- [31] A.R. Hevner, S.T. March, J. Park, S. Ram, Design science in information systems research, *MIS Q.* 28 (2004) 75–105 .
- [32] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, S. Linkman, Systematic literature reviews in software engineering—a systematic literature review, *Inf. Softw. Technol.* 51 (2009) 7–15 .
- [33] J. Webster, R.T. Watson, Analyzing the past to prepare for the future: writing a literature review, *MIS Q.* 26 (2002) .
- [34] B.G. Glaser, The Constant comparative method of qualitative analysis, *Soc. Probl.* 12 (1964) 436–445 .
- [35] B.G. Glaser, A.L. Strauss, *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Transaction Books, 2009 .
- [36] K. Kuttner, J. McAndrews, Personal on-line payments, *Econ. Pol. Rev.* 7 (2001) .
- [37] P. Corbetta, *Social Research: Theory, Methods and Techniques*, Sage Publications, London, 2003 .
- [38] A.B. Kajornboon, Using interviews as research instruments, *E-J. Res. Teach.* 2 (2005) .
- [39] K.M. Eisenhardt, Building theories from case study research, *Acad. Manage. Rev.* 14 (1989) 532–550 .
- [40] Cisco, Managing the real cost of on-demand enterprise cloud services with chargeback models: A guide to cloud computing costs, in: *Server Costs, Pricing Plans, and Chargeback Implementation and Systems*, Cisco Systems, 2010, pp. 1–10 .
- [41] E.M. Hufnagel, J.G. Bimberg, Perceived chargeback system fairness in decentralized organizations: An examination of the issues, *MIS Q.* 13 (1989) 415–430 .
- [42] R. Cooper, Cost management concepts and principles, *J. Cost Manag.* (1987) 45–49. Spring .
- [43] P. Holzer, H. Norreklit, Some thoughts on cost accounting developments in the United States, *Manag. Account. Res.* 2 (1991) 3–13 .
- [44] R.L. Nolan, Effects of chargeout on user/manager attitudes, *Commun. ACM* 20 (1977) 177–185 .
- [45] B. Raghunathan, T.S. Raghunathan, A discriminant analysis of the relationship between IS charging systems and organizational variables, *Omega* 22 (1994) 321–330 .
- [46] R. Van den Bossche, K. Vanmechelen, J. Broeckhove, Cost-optimal scheduling in hybrid IaaS clouds for deadline constrained workloads, in: *2010 IEEE 3rd International Conference on Cloud Computing*, IEEE, 2010, pp. 228–235 .
- [47] J. Wilkes, PRESS: PRedictive Elastic ReSource Scaling for cloud systems, in: *2010 International Conference on Network and Service Management*, IEEE, Niagara Falls, ON, 2010, pp. 9–16 .
- [48] A. Andrzejak, D. Kondo, D.P. Anderson, Exploiting non-dedicated resources for cloud computing, in: *2010 IEEE Network Operations and Management Symposium—NOMS 2010*, IEEE, Osaka, 2010, pp. 341–348 .
- [49] A. Di Stefano, C. Santoro, An economic model for resource management in a Grid-based content distribution network, *Future Gener. Comput. Syst.* 24 (2008) 202–212 .
- [50] M. Mattess, C. Vecchiola, R. Buyya, Managing peak loads by leasing cloud infrastructure services from a spot market, in: *2010 IEEE 12th International Conference on High Performance Computing and Communications (HPCC)*, IEEE, 2010, pp. 180–188 .
- [51] N. Grounds, J. Antonio, J. Muehring, Cost-minimizing scheduling of workflows on a cloud of memory managed multicore machines, in: M. Jaatun, G. Zhao, C. Rong (Eds.), *CoudCom 2009*, in: LNCS, vol. 5931, Springer, 2009, pp. 435–450 .
- [52] R. Iyer, R. Illikkal, L. Zhao, D. Newell, J. Moses, Virtual platform architectures for resource metering in datacenters, *ACM SIGMETRICS Perform. Eval. Rev.* 37 (2009) 89–90 .
- [53] T. Benson, A. Akella, D.A. Maltz, Mining policies from enterprise network configuration, in: *Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement Conference*, ACM, 2009, pp. 136–142 .
- [54] P. Patel, A. Ranabahu, A. Seth, Service level agreement in cloud computing, in: *Cloud Workshops at OOPSLA*, IEEE, 2009 .
- [55] I. Petri, O.F. Rana, G.C. Silaghi, Service level agreement as a complementary currency in peer-to-peer markets, *Future Gener. Comput. Syst.* 28 (2012) 1316–1327 .

- [56] Í. Goiri, F. Julià, J.O. Fitó, M. Macías, J. Guitart, Supporting CPU-based guarantees in cloud SLAs via resource-level QoS metrics, *Future Gener. Comput. Syst.* 28 (2012) 1295–1302 .
- [57] T. Singh, P.K. Vara, Smart metering the clouds, in: 2009 18th IEEE International Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises, IEEE, 2009, pp. 66–71 .
- [58] R. Han, M.M. Ghanem, L. Guo, Y. Guo, M. Osmond, Enabling cost-aware and adaptive elasticity of multi-tier cloud applications, *Future Gener. Comput. Syst.* 32 (2014) 82–98 .
- [59] X. León, T.A. Trinh, L. Navarro, Using economic regulation to prevent resource congestion in large-scale shared infrastructures, *Future Gener. Comput. Syst.* 26 (2010) 599–607 .
- [60] I. Breskovic, J. Altmann, I. Brandic, Creating standardized products for electronic markets, *Future Gener. Comput. Syst.* 29 (2013) 1000–1011 .
- [61] D.H. Catteddu, Giles, benefits, risks and recommendation for information security, in: ENISA, Crete, Greece, 2009 .
- [62] R.K. Yin, *Case Study Research: Design and Methods*, Sage, 2003 .
- [63] N.G. Carr, *Does IT Matter? Information Technology and the Corrosion of Competitive Advantage*, 2004 .
- [64] N.G. Carr, IT doesn't matter, *Harv. Bus. Rev.* 81 (2003) 41–49 .
- [65] D. Kondo, G. Fedak, F. Cappello, A.A. Chien, H. Casanova, Characterizing resource availability in enterprise desktop grids, *Future Gener. Comput. Syst.* 23 (2007) 888–903 .
- [66] W.-K. Cheng, B.-Y. Ooi, H.-Y. Chan, Resource federation in grid using automated intelligent agent negotiation, *Future Gener. Comput. Syst.* 26 (2010) 1116–1126 .
- [67] K. Ramachandran, H. Lutfiyya, M. Perry, Decentralized approach to resource availability prediction using group availability in a P2P desktop grid, *Future Gener. Comput. Syst.* 28 (2012) 854–860 .

TABLES AND FIGURES

Table 1

The five characteristics of the cloud that set it apart from traditional IT (based on the work of Mell and Grance [9]).

Cloud characteristic	Description
Highly abstracted resources	Using virtualization, resources can be 'created' and scaled on the spot over one or more physical resources.
Near instant scalability and flexibility	The ability to add or remove virtual resources with the click on a button.
Near instantaneous provisioning	The ability to supply resources, services and such nearly-instantaneous.
Shared resources	Resources can be shared by multiple tenants.
Service on demand	Get the services needed on demand, and pay only for what you use (pay per hour basis, pay per use etc.).
Programmatic management	APIs provide interfaces to manage the cloud environment. E.g. via web interfaces, on a step for step basis.

Table 2

Chargeback types described in literature.

Model name	Description
No chargeback	The organization applies no chargeback and the IT costs form a general organizational expense [19].
Notional charging	A model according to which organizational units are only notified about the charges, but not charged in practice [14,15].
Tiered subscription	Services are offered to consumers in different tiers, e.g. bronze, silver, gold, and user are charged a corresponding subscription fee [15].
Metered usage	Service usage is metered and consumers are charged based on actual usage [20].
Direct plus	The costs directly related to a service are augmented with part of the general IT overhead and are charged back to users [15].
Fixed costs charging	Costs are divided between organizational units based on a common denominator such as the number of employees or the revenue of the department [15].
Market based pricing	When services are provisioned externally from the market, the market price could be directly allocated to consumers [2].
Negotiation charging	A chargeback approach according to which the prices are set through negotiation between the IT department and the different consumers [2].
Management set prices	Internal charges are determined by executive management [2].
Activity based costing	IT costs are attributed to activities based on the rate with which the latter consume IT resources [2,21].

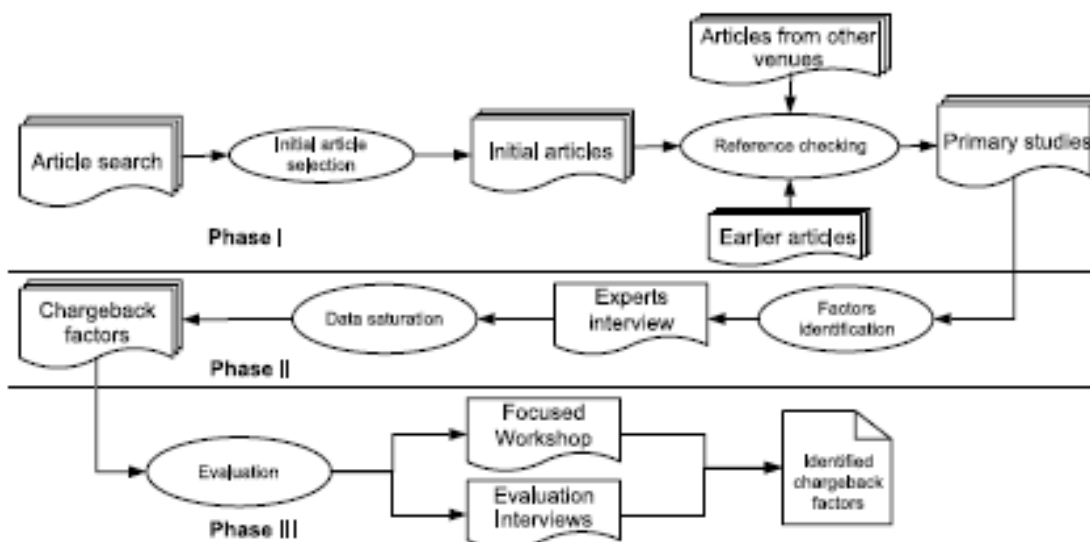


Fig. 1. Overview of the research process.

Table 3

Key terms used in the literature review.

Chargeback concept	Cloud computing concepts
(IT) Chargeback (charge-back, charge back)	Cloud computing
(IT) Charging	(Cloud) services
(IT) Chargeout	Cloud types
Chargeback models/types	Cloud computing economics
Cost allocation/accounting	Cloud computing pricing model

Table 4

Organizations participating in the research by industry, size and number of respondents.

Alias ^a	Industry	Size	No. of respondents
IBM	Consulting technology	> 10,000	16
INS1	Insurance	1000–10,000	1
INS2	Insurance	1000–10,000	2
MUN	Government/Municipality	1000–10,000	1
GOV	Government	> 10,000	3
TRANS	Transport	> 10,000	1
HOUS	Housing/Construction	1000–10,000	1

^a For confidentiality reasons aliases are used for the names of partners and clients of IBM.

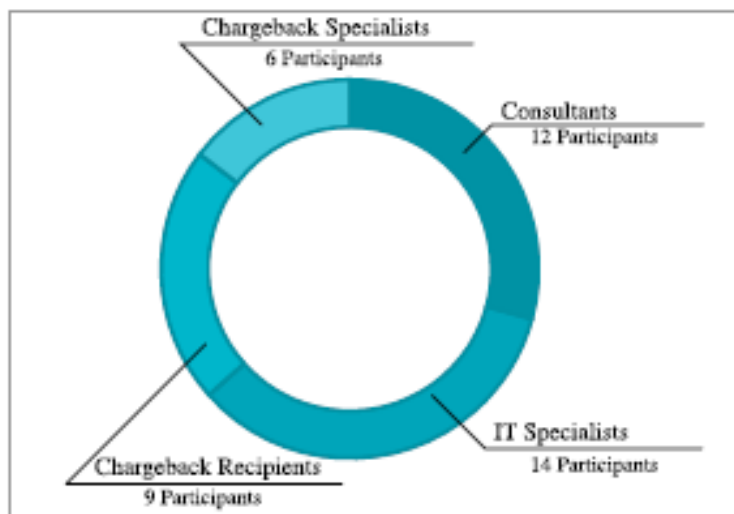


Fig. 2. Details of the role of the participants of the evaluation phase.

Table 5

Summary of the factors identified from the literature.

Factors	References
Accuracy	Bergeron et al. [1], Drury [12], Ross et al. [5], Cisco [40], Hufnagel and Birnberg [41]
Cost of costing	Atkinson et al. [21], Cooper [42], Holzer et al. [43]
Transparency and understandability	Drury [12], Stiller et al. [16]
Controllability	Nolan [44]
Fairness	Hufnagel and Birnberg [17,41]



Fig. 3. Factors that impact chargeback according to the scientific literature.

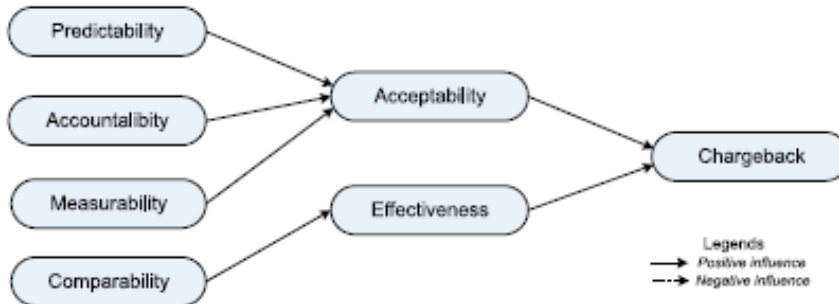


Fig. 4. Newly discovered factors from the semi-structured interviews. These exclude those found in both the SLR and in the interviews.

Table 6

Number of times each factor was mentioned during the interviews by background of the twenty-five respondents and the responding percentages per group total.

	Total ^a	Executives	Chargeback recipients	IT specialists	Chargeback specialists	Consultants
Effectiveness	13 (52%)	2 (50%)	0 (0%)	4 (29%)	6 (100%)	5 (42%)
Acceptability	12 (48%)	1 (25%)	1 (11%)	5 (36%)	4 (67%)	6 (50%)
Measurability	6 (24%)	0 (0%)	0 (0%)	6 (43%)	0 (0%)	3 (25%)
Accuracy	21 (84%)	2 (50%)	4 (44%)	9 (64%)	6 (100%)	9 (75%)
Transparency and understandability	12 (48%)	1 (25%)	3 (33%)	4 (29%)	5 (83%)	4 (33%)
Controllability	8 (32%)	0 (0%)	3 (33%)	4 (29%)	5 (83%)	1 (8%)
Predictability	14 (56%)	0 (0%)	4 (44%)	11 (79%)	1 (17%)	7 (58%)
Accountability	7 (28%)	0 (0%)	1 (11%)	4 (29%)	3 (50%)	1 (8%)
Cost of costing	8 (32%)	2 (50%)	0 (0%)	3 (21%)	1 (17%)	5 (42%)
Comparability	4 (16%)	0 (0%)	0 (0%)	2 (14%)	2 (33%)	2 (17%)

^a Since a few respondents fell into two stakeholder categories due to their diverse job responsibilities, the sum of the five rightmost columns is greater than the total count. Percentage is calculated based on the actual group totals.

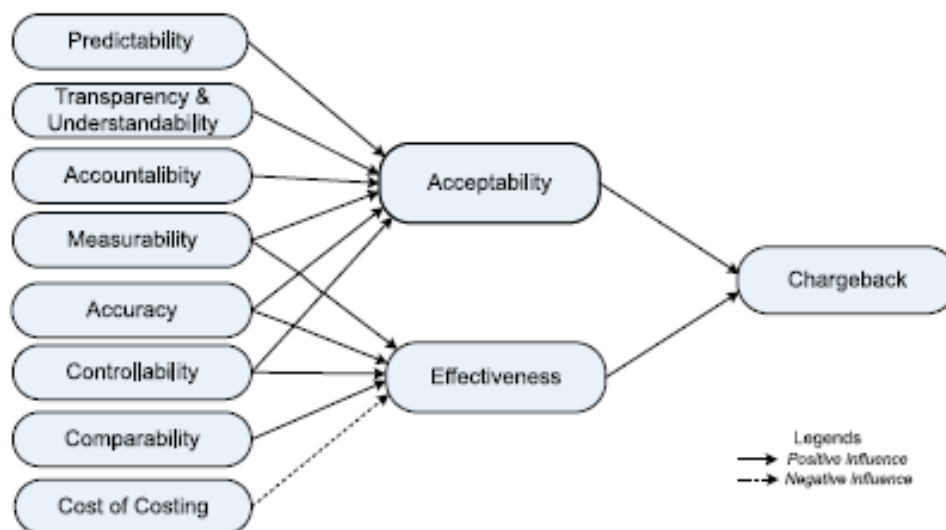


Fig. 5. All factors combined that influence cloud chargeback.

Table 7
The final factors defined and their relation to acceptability and/or effectiveness explained.

Factor	Definition	Influences
Measurability	The capability to precisely measure the amount of used consumption units of a service and to determine who has used those consumption units.	Acceptability: Charging on the basis of something that cannot be clearly measured, but is approximate, leads to mistrust for the model. Effectiveness: The more precisely usage is metered, the more effective the charging.
Accuracy	The extent to which the charges allocated to an organizational unit for each service accurately approximate the actual costs incurred by the organization for delivering the service to the unit.	Acceptability: Fundamental for business consumers to consider the method fair and to trust it. Effectiveness: The more accurate the charges, the more effective the chargeback subsidization is avoided; decision making and behavior are based on actual costs.
Transparency and understandability	The capability to understand what the service units mean (understandability) and how their prices are formed (transparency).	Acceptability: Consumers are more likely to accept highly transparent bills that they can understand well.
Controllability	Determines whether chargeback recipients can control their IT bills and cost.	Acceptability: Inability to control the bill leads to frustration and perception of chargeback as unnecessary overhead. Effectiveness: Behavior can be changed, and cost reductions realized only if users can control the bill.
Predictability	The extent to which chargeback recipients are able to predict future bills.	Acceptability: Unpredictable bills could eschew the budget minded manager, leading to resentment for the chargeback model.
Accountability	To what extent chargeback recipients are able to easily verify the correctness of the bill.	Acceptability: Users find it easier to accept the chargeback model if they can easily verify bill correctness.
Cost of costing	All the chargeback related costs, including chargeback design, integration, and operation.	Effectiveness: Weighing the expected chargeback benefits against costs is key to designing an effective chargeback model.
Comparability	The degree to which consumers are able to compare the prices of internally provisioned services to similar or equivalent services offered on the market or in other organizations.	Effectiveness: Easy benchmarking leads to consumers exercising pressure for more efficient provision of IT services and a subsequent cost reductions on the provisioning side.