Find the Way in the Jungle of Quality of Service in Industrial Cloud: A Systematic Mapping Study

Malvina Latifaj[1], Federico Ciccozzi[1][a] and Séverine Sentilles[1][b]
[1]*School of Innovation, Design and Engineering – Mälardalen University, Sweden*
{*malvina.latifaj, federico.ciccozzi, severine.sentilles*}*@mdh.se*

Keywords: Industrial cloud computing, Quality of service, Quality attributes, Systematic mapping study

Abstract: The rapid development of Industry 4.0 and Industrial Cyber-Physical Systems is leading to the exponential growth of unprocessed volumes of data. Industrial cloud computing has great potential for providing the resources for processing this data. To be widely adopted, the cloud must ensure satisfactory levels of Quality of Service (QoS). However, the lack of a standardized model of quality attributes hinders the assessment of QoS levels. This paper provides a comprehensive systematically defined map of current research trends, results, and gaps in quality attributes and QoS in industrial cloud computing. An extract of the main insights is as follows: (i) the adoption of cloud technologies is closely related to performance indicators, however other quality attributes, such as security, are not considered as much as they should; (ii) solutions are most often not tailored to specific industrial application domains; (iii) research largely focuses on providing solutions without solid validation, unsuitable for effective and fruitful technology transfer.

1 Introduction

Industrial cloud computing, which aims to provide industrial digital information integration and collaboration between enterprises based on a shared understanding of concepts (Wlodarczyk et al., 2009), is considered to be a favorable solution to cope with limitations of hardware and software for multiple enterprises. Dependability and software quality are crucial factors in the success of industrial applications (Guo and Deze, 2019). Thus, appropriate quality of service (QoS) level must be attained for cloud platforms to supply the necessary resources for processing industrial data.

However, the quality attributes needed to assess QoS have no formal definition and there does not exist a complete list of them (Chung et al., 2012; Mairiza et al., 2010). There have been some attempts to the standardization of these attributes, but the classification schemes are inconsistent with each other terminologically and also categorically, and therefore their definitions still lack consensus. The absence of a generic model of quality attributes that can be used as a reference for specific domains leads to the lack of a fundamental set of quality attributes to be used to evaluate QoS in industrial cloud computing, which is necessary due to the peculiarities of the latter. The existing literature does not provide the necessary evidence that could help in identifying the appropriate quality attributes to evaluate QoS in industrial cloud computing. To fill this gap, we conducted a systematic mapping study in the context of QoS in industrial cloud computing. The overall goal was to identify and classify the quality attributes that are used to evaluate QoS in industrial cloud computing and the ways to assess them. The **main contributions** of this study are as follows:

- an overview of the publication trends on the topic;
- a classification of the quality attributes used to evaluate QoS in industrial cloud computing and the quality metrics used to assess them;
- a classification of the factors that have an impact on QoS in industrial cloud computing;
- a classification of the domains in which QoS in industrial cloud computing is investigated;
- a classification of the primary studies' objectives and a description of the means by which these objectives are achieved.

The target audience of this paper includes researchers and practitioners in the field of cloud computing, seeking a better insight on the QoS specifically dealing with quality attributes, or willing to adopt cloud technologies but needing a reference for

[a] https://orcid.org/0000-0002-0401-1036
[b] https://orcid.org/0000-0003-0165-3743

the evaluation of related QoS.

The remainder of this paper is organized as follows. Section 2 provides a background on QoS and industrial cloud computing and Section 3 presents the study design. Section 4 reports on the vertical analysis results, while Section 5 reports on the horizontal analysis results. Section 6 provides a discussion of the main results and Section 7 discusses threats to validity. Section 8 introduces the related work and Section 9 concludes the paper.

2 Background

Cloud computing is a cutting-edge technology that makes possible the sharing of resources between multiple users. The popularity that cloud computing has obtained over the last decades, attributable to the advantages it provides compared to traditional approaches, has stimulated consumers to adopt it. Moreover, during the last few years, the digital information produced by industry has increased rapidly and the need for collaboration mechanisms is higher. Enterprises try to handle this flow of information using their existing structures, but they have boundaries on their capabilities. Even though cloud computing can tackle the complexity of these collaborative approaches, none of its deployment models can overcome all the occurring obstacles. For that reason, the concept of Industrial Cloud Computing was introduced as a solution in the form of a platform to exchange, process, and analyze digital information (Wlodarczyk et al., 2009). But, to stay competitive as technology evolves, a certain QoS should be developed to meet the customers' expectations. QoS in cloud computing indicates "*the levels of performance, reliability, and availability offered by an application and by the platform or infrastructure that hosts it*" (Ardagna et al., 2014, p. 1), and its main objective is to make the services acceptable for users. Dromey (1995) states that to increase the software quality, there should be a defined quality model that can make clear and direct links amongst the high-level quality attributes and specific product characteristics. Quality attributes are indeed fundamental elements of quality models and have a crucial impact on software development, but they have been left in the background compared to functional requirements, and the results are scattered (Chung et al., 2012). This leads to quality attributes not being addressed correctly and software of poor quality. The reasons as to why quality attributes, which are so pivotal to the quality of software systems, can be so hard to address correctly are as follows: (i) their nature is subjective as different stakeholders interpret and evaluate them differently, leading to ambiguity, (ii) their achievement is relative, as there might always be different ways to reach a satisfactory level, (iii) they can be interactive and dependent on one another, so there can not be a localized solution. This set of issues makes it difficult to deal with and even more challenging to measure and verify quality attributes. To objectively measure these quality attributes, we use quality metrics (Fernando et al., 2014). Through this quantitative measurement, the actual QoS can be compared to the expected QoS and future steps for improvement can be determined.

3 Study design

In this section, we present the design of our study. To form a good understanding of the state-of-the-art on QoS in industrial cloud computing with a focus on quality attributes, we carry out a systematic mapping study according to the guidelines of Petersen et al. (2015).

3.1 Research goal and questions

The fundamental goal of this study is to: *achieve a classification of the quality attributes addressed the most in industrial cloud computing and to provide researchers and practitioners with a mapping of key QoS attributes and methods to assess them.* This will enable them to make informed decisions in the context of QoS for industrial cloud computing and will highlight research gaps, laying the groundwork for future research. The research questions (RQx) to be answered by this study are the following:

RQ1. What are the publication trends regarding the quality of service in industrial cloud computing research?
Rationale and outcome: Identify the existing state of research on QoS in industrial cloud computing and assess the density of scientific publications.
RQ2. What are the most addressed quality attributes and related aspects in industrial cloud computing research?
Rationale and outcome: The concept of QoS is rather broad and it includes a multitude of quality attributes, therefore it is important to identify (i) a set of software quality attributes addressed in industrial cloud computing, (ii) a set of factors that affect quality attributes and the overall QoS, and (iii) the industrial domains in which cloud computing is used.
RQ3. What are the most common quality metrics used to assess quality attributes in industrial cloud

computing research?
Rationale and outcome: Identify the metrics used to assess the quality attributes in industrial cloud computing and map them to the latter.
RQ4. What are the research objectives of the primary studies in relation to QoS in industrial cloud computing and how are these objectives achieved?
Rationale and outcome: Obtain a classification of the objectives of the primary studies in relation to QoS, and ways on how these objectives can be achieved.

3.2 Search and selection

The search and selection process is a decisive multi-step process that provides wide coverage of the topic under investigation. Figure 1 illustrates the complete process and provides information on the size of our corpus during the entailed four steps.

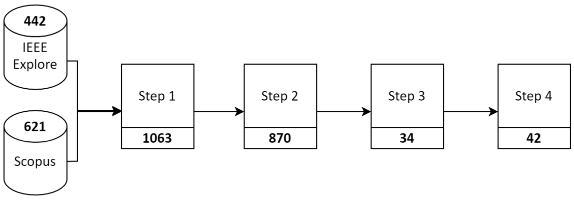

Figure 1: Search and selection process

Step 1: Initial search. To obtain the initial set of primary studies, automatic searches were carried out on two of the largest and most complete electronic databases in software engineering that are well-established and include a wide spectrum of peer-reviewed publications: IEEE Xplore Digital Library and SCOPUS. The search was performed considering the title, abstract, and keywords of studies and it was conducted on 25 February 2020.

The search string is constructed using the PICOC Petticrew and Roberts (2008) criteria and it is the following:

“cloud” **AND** “industr*” **AND** (“quality of service” **OR** “QoS” **OR** “quality model” **OR** “software qualit*” **OR** “quality propert*” **OR** “quality attribut*” **OR** “non functional” **OR** “extra functional” **OR** “NFR” **OR** “EFR” **OR** “NFP” **OR** “EFP”)

Step 2: Merging and impurity removal. A multi-database search provides a broader spectrum of studies and minimizes publication bias, but simultaneously leads to duplicates and repetitive studies. For IEEE Xplore and Scopus, we identified 193 duplicates based on an exact match of title, authors, and publication year; we removed the duplicates. No studies with different versions were identified.

Step 3: Application of selection criteria. The selection criteria are crucial elements to the selection process and therefore, an explicit definition of these criteria should be provided to allow the replication of the study. An initial set of selection criteria was defined and the papers were screened based on title, abstract, and keywords. However, there was a set of 150 primary studies that were uncertain with respect to the inclusion criteria. The inclusion criteria were refined and the final criteria are introduced in the following paragraph.

Inclusion criteria
I1. The study reports on an approach that aims to define, measure, analyze, evaluate or improve QoS in the scope of cloud computing.
I2. The study investigates quality attributes, quality metrics, or quality models and their impact on QoS in cloud computing.
I3. The study reports on the use or need of cloud technologies in industry (e.g., banking) or it is published at an industrial venue.
I4. The study is written in English and is available in full text.
I5. The study is peer-reviewed.

Exclusion criteria
E1. The study does not focus on any specific quality attribute.
E2. Secondary or tertiary study.
E3. The study is published as tutorial paper, short paper, poster paper, editorial, book, keynote, tutorial summary, tool demonstration, panel discussion, technical report, or other non-peer-reviewed publications.
E4. The study is not in the computer science context.

Step 4: Backward snowballing. The initial corpus of primary studies obtained after the application of the selection criteria is complemented with a fully-recursive backward snowballing activity, according to the guidelines by Wohlin (2014). Out of 790 results, only eight publications were included in our corpus for a total of **42 primary studies** that can be found in the appendix in the replication package[1]. The main reasons for the exclusion of publications were as follows: i) the publications retrieved from the reference list of a specific study, do not report on QoS, or quality attributes and metrics, therefore not satisfying I1 and I2, ii) the publications do not report

[1] https://bit.ly/2LvkORL

on the use or need of cloud technologies in industry, therefore not satisfying I3.

3.3 Data extraction and synthesis

As a means of extracting and classifying the relevant information from the primary studies, we constructed the classification framework illustrated in Table 1. The parameters of the classification framework were defined based on (i) the research questions, and (ii) the keywording method. After the definition of the classification framework, the full text of the primary studies was read to collect information according to the categories of the defined framework, and additional information that did not fit in any category but was considered relevant. The categories with the most additional information were the quality attributes, quality metrics, and means of reaching the objectives. These categories had to be refined, as the keywording process's initially defined parameters were not representative enough of the primary studies. At the end of this process, no studies were excluded.

The data synthesis activity is conducted according to the recommendations by Cruzes and Dybå (2011), and it is divided into two main phases: vertical analysis in which we analyze the extracted data from the primary studies in order to collect information regarding each defined parameter of our classification framework by applying the line of argument synthesis recommended by Wohlin et al. (2012), and horizontal analysis in which we cross-tabulate the data among the different categories defined for each research question, to find similarities, noteworthy differences, and recurring patterns.

4 Vertical analysis results

The purpose of the vertical analysis is to provide quantitative results regarding each category in isolation. These results are provided in the following subsections in a textual and graphical form.

RQ1 – Publication trends

Regarding the distribution of primary studies by type of **venue**, the most common publication type is *conference papers* (21), followed by *journal papers* (15) and *workshop papers* (2). Four primary studies pertain to industrial venues (P1, P2, P4, and P5), and only one venue hosts more than one primary study (P1 and P5 published at IEEE Transactions on Industrial Informatics).

Moreover, according to the extracted data, the main **contribution** types are *model* (18) and *method* (18), with the rest distributed as follows: *framework* (4), *tool* (2), *process* (1) and *metric* (1).

With respect to the **research** type, among our primary studies, there were only occurrences of *solution proposal* (40) and *evaluation research* (3) papers, where the majority were solution proposals that proposed novel ways on how to define, measure, analyze, evaluate, or improve QoS in industrial cloud computing. P28 and P37 are evaluation research papers, while P38 is both a solution proposal and an evaluation research paper.

In terms of **validation**, the majority of primary studies validate their approaches by means of *evaluation* (24), but these approaches are not validated formally nor empirically. A set of 11 primary studies conducts a more extensive validation of the proposed approaches through *analysis*, while a smaller set of primary studies focus on validation by *example* (4). Only two primary studies fall in the *persuasion* validation category and only one primary study falls in the *blatant assertion* category. The absence of experience papers may suggest a lack of maturity in this research area, as existing solutions have not been validated by the community yet.

RQ2 – Quality attributes and related aspects

Quality attributes. Figure 2 illustrates the details of our quality attributes categorization by presenting the quality attributes mentioned in more than one primary study. Our investigation identified 18 considered quality attributes in total. Nine of them are mentioned only once through all the primary studies and seven of these attributes are mentioned in P35, which is also the study that identifies most quality attributes (14/18). This paper contributes with a framework that aims to evaluate the quality of cloud services by systematically measuring QoS attributes and using them as a basis for ranking cloud providers. The majority of the primary studies addresses *performance* as a relevant quality attribute in industrial cloud computing. The overall results indicate a significant gap between *performance* and all other quality attributes. This suggests that the adoption of cloud technologies in industry is very related to the performance indicators offered by these technologies.

Factors: Figure 3 provides an illustration of the factors that affect quality attributes. Being that cloud technologies operate on the grounds of *virtualization*, it is not surprising that quality attributes are frequently affected by virtualization solutions. Virtualization is followed by *data storage* archi-

RQ	Category	Possible values
RQ2	Quality attributes	performance (P), security (SE), efficiency (E), reliability (RL), availability (A), scalability (SC), usability (U), elasticity (EL), stability (ST), privacy (PR), trust (T), responsiveness (RS), transparency (TP), interoperability (I), adaptability (AD), sustainability (SUS), suitability (SUI), accuracy (AC)
	Factors	virtualization (V), data storage (DS), network (N), energy (E), other (O)
	Domain	gaming (G), mobile (M), power trading (PT), big data (BD), sensor cloud (SC)
RQ3	Quality metrics	response time (RT), resource utilization (RU), make span (MS), latency (L), request completion time (RCT), waiting time (WT), CPU utilization (CU), failure rate (FR), throughput (TH), execution time (ET), rejection probability (RP), request arrival rate (RAR), delay (D), jitter (J), performance deviation (PD), mean time to failure (MTTF), packet delivery ratio (PDR), guarantee ratio (GR), drop rate (DR), time to adapt to upgrades (TAU), access rate (AR), request loss probability (RLP), scheduling time (ST), system overhead rate (SOR), mean time between failures (MTBF) , bandwidth (B), deadline miss ratio (DMR), job loss ratio (JLR)
RQ4	Objectives	improve (I), evaluate (E), select provider (SP), measure (M), define (D)
	Means	scheduling (S), security mechanisms (SM), load balancing (LB), error-prone conditions (EPC), allocation (A), edge servers (ES), queuing model (QuM), request traffic control (RTC), workload variability (WV), SPs coalition (SPC), ranking cloud services (RCS), quality model (QlM), dynamic optimal routing (DOR), VM live migration (LM)

Table 1: Classification framework overview

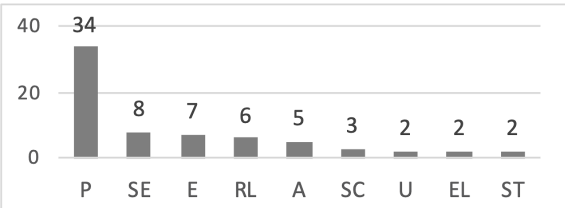

Figure 2: Quality attributes

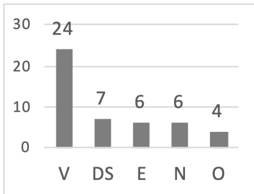

Figure 3: Factors

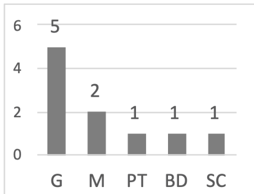

Figure 4: Domains

tecture which is not surprising either, considering the fact that enterprises use cloud to store sensitive data, and that raises numerous security concerns. Furthermore, six studies mention that cloud providers often have to trade-off between offered levels of QoS and energy consumption. *Network* architecture is also mentioned in six primary studies and is mostly investigated in cloud gaming, because the majority of computational operations are performed in the cloud, and, in high-action games, network latency is the greatest concern. A total of four primary studies (P1, P4, P11, P35) fall into the *other* category as they focus on providing quality models and mechanisms to rank service providers and are not affected by specific factors.

Domains. From the set of 42 primary studies, 32 of them investigate QoS in cloud computing by providing insights on the use or need of cloud technologies in industry, but without a clear indication of specific industrial domains. The other ten primary studies identify five industrial domains, as illustrated in Figure 4. Overall, the results indicate that even though cloud technologies are very beneficial to various industries and enterprises, research on their use in specific industrial domains still lacks maturity.

RQ3 – Quality metrics

Our investigation identified 28 relevant quality metrics. Figure 5 illustrates the quality metrics that are mentioned in more than one primary study. Top three quality metrics are *response time*, followed by *resource utilization* and *make span*. *Response time* is repeatedly emphasized for the impact it has on the performance of cloud technologies. Thus, being that performance is the quality attribute mentioned the most in our primary studies this result is not surprising. It is worth mentioning that, while all other studies only suggest the use of specific metrics, P35 proposes new approaches for assessing quality attributes through an extensive set of formulas, but with no specific terminology regarding the quality metric, thus they are not included as part of the quality metrics list.

RQ4 – Objectives and how to reach them

Objectives. Regarding the objectives of primary studies the results indicate that the most common objective is to *improve* the QoS offered by the service providers as it is crucial for the adoption of cloud technologies. Even though a considerable difference is observed between the number of studies that try to *improve* and *evaluate*, research efforts on the latter do not go unnoticed. A total of nine primary studies fall

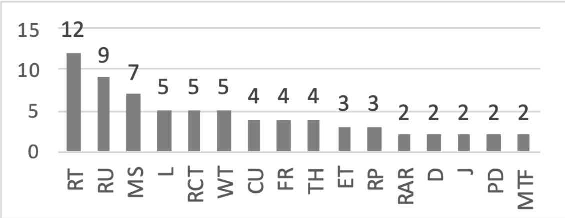

Figure 5: Quality metrics

into this category. Three primary studies investigate approaches on how to *select* the most suitable cloud provider according to specific requirements and only two primary studies fall into the *define* category, where the authors aim to define quality attributes and a quality model. The low number of primary studies in this category highlights the fact that the community is still lacking agreement when defining QoS and quality attributes.

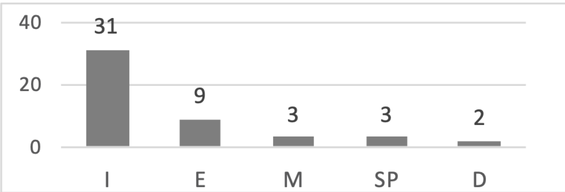

Figure 6: Objectives

Means. Our investigation identified 14 approaches that can help researchers achieve the objectives specified above. As shown in Figure 7, the five top approaches are: *scheduling*, *security mechanisms*, *load balancing*, *error-prone conditions*, and *allocation*. These results imply that the most effort has been put into *scheduling*, which is not surprising since, with the increasing amount of data that is being processed in the cloud, effective scheduling mechanisms are crucial to providing satisfactory levels of QoS. Furthermore, being that security was the second most mentioned quality attribute in industrial cloud computing (see Figure 2), it was expected that the security mechanisms would be ranked among the top approaches.

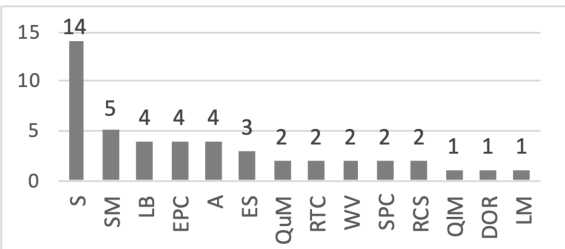

Figure 7: Means of reaching the objectives

5 Horizontal analysis results

The purpose of the horizontal analysis is to identify possible connections between the categories of our classification framework. To do that, we cross-tabulated the extracted data across different categories and analyzed the relevant insights.

Quality attributes and metrics. Figure 8 provides evidence on the quality metrics that are used to assess specific quality attributes. We have only included the metrics that have more than one occurrence among the primary studies and the quality attributes assessed by them. Only five quality attributes are assessed by quality metrics that are mentioned in more than one primary study. The majority of quality metrics are used to assess *performance* and only a few metrics are used to assess other quality attributes. This was in part expected, considering the significant gap in occurrence between performance and other quality attributes identified during the vertical analysis.

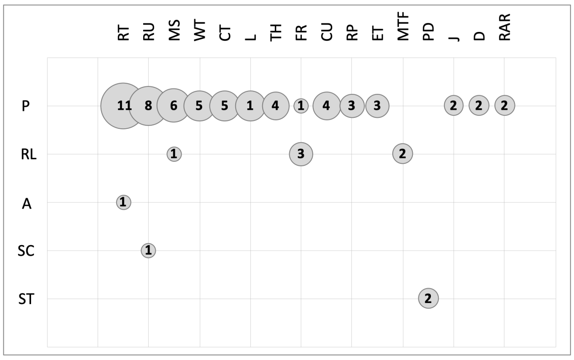

Figure 8: Quality attributes and metrics

Factors and quality attributes. Figure 9, provides evidence that *performance* is mostly affected by the *virtualization* solutions of the cloud provider. Different virtualization solutions provide varying levels of performance due to the technologies and strategies used. Another factor that has a significant impact on *performance* is the *network* infrastructure, as a poor infrastructure is bound to negatively affect the application's performance. Furthermore, *security* is mostly affected by the *data storage* architecture. Considering the higher exposure that data encounters on the cloud than on-premises, it is surprising to see that only one primary study mentions the impact that the virtualization solutions have on security. This result suggests the lack of research on how to mitigate security threats when providing virtualization solutions. *Efficiency* is affected by the energy management solutions offered by the cloud provider, as the high consumption of energy for cloud data centers is raising economic and environmental concerns.

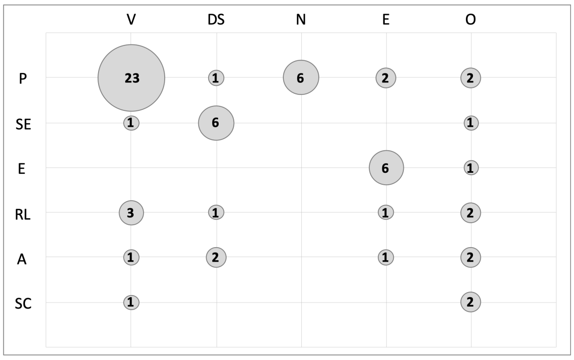

Figure 9: Factors and quality attributes

Objective of study and means of reaching it. Figure 10 illustrates the relation between the objectives of primary studies and means of reaching them. From the results it can be seen that the most mentioned approach to *improve* QoS is *scheduling*, followed by *security mechanisms* and *load balancing*. Through scheduling, the resources can be utilized efficiently, and the overall performance of the system can experience a significant improvement. In order to *evaluate* QoS, the majority of primary studies use simulation of error-prone conditions, such as VM and server failure, followed by *edge servers*, *queuing model* and *workload variability*. The most common approach to help *select a provider* is by ranking the cloud services based on the users' requirements and providers' capabilities. This approach is also used to *measure QoS* of different cloud providers.

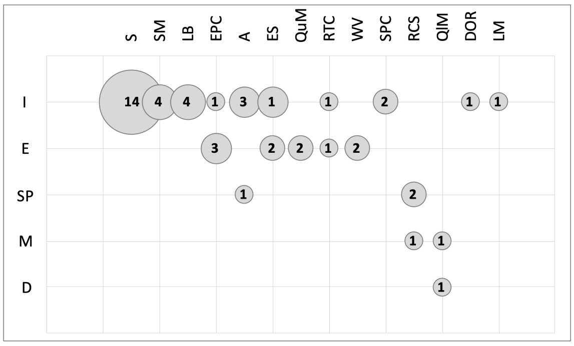

Figure 10: Objectives and means

Contribution type and validation type. Figure 11 illustrates a high concentration of primary studies that use *evaluation* to validate the proposed *models* and *methods*. As a consequence, the scientific evidence that they provide cannot be considered particularly strong. However, *models* surpass *methods* when it comes to more extensive and systematic validation, such as *analysis*. This suggests that *models* have been provided with higher quality evidence in terms of their applicability. *Metrics*, *tools*, and *processes* have only been validated through *analysis*, while *frameworks* have an even distribution with one occurrence in each category, except for *blatant assertion*.

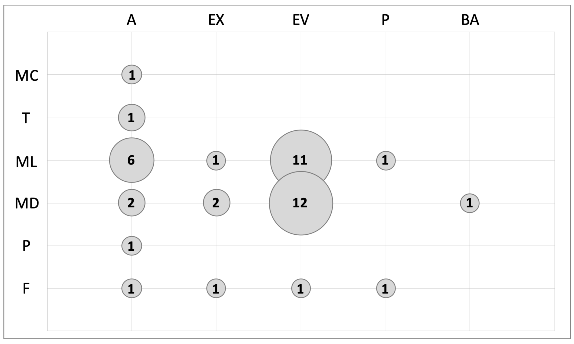

Figure 11: Contribution and validation type

6 Discussion

In order to describe and interpret the significance of our findings, this section discusses the obtained results of the vertical and horizontal analysis introduced in the previous section.

Regarding **publication trends** (*RQ1*), the majority of primary studies are published at conferences, followed by journals and workshops, which is common in the software engineering literature. Furthermore, the research on this topic can be categorized as solution-seeking, with $\sim 95\%$ of the primary studies being *solution proposals*. Among them, the majority contribute with a *model* or *method*. We believe that more effort should be put on the *evaluation* and *validation* of solutions in industry. More specifically, more attention should be put on formal or empirical validation to properly assess the limitations and benefits of the proposed solutions.

Regarding **quality attributes** (*RQ2*), the results provide evidence that one particular quality attribute is addressed more frequently than others: *performance*, addressed in $\sim 80\%$ of the primary studies. More effort should be directed to other quality attributes since desired levels of QoS are usually achieved through a combination of quality attributes (Software & Systems Engineering Standards Committee, 1998). Furthermore, different primary studies report on different factors that affect quality attributes and, consequently, the overall QoS. *Virtualization* is covered in more than $\sim 50\%$ of the primary studies, and it is mostly mentioned as a factor affecting *performance*. This result was expected as *virtu-*

alization is a ground technology that enables the use of cloud computing and improves the utility of physical machines. Furthermore, a notable number of primary studies investigate *data storage* architecture as a factor affecting QoS and more specifically, *security*. This was expected as numerous enterprises are moving their data to the cloud because of its simplicity, but fear the security aspect. We believe additional research must be conducted in how *virtualization* solutions affect *security* because each virtualization layer can be a target to malicious attacks and compromise the entire cloud infrastructure. Moreover, our study revealed that only $\sim 24\%$ of the primary studies reported on QoS with respect to a particular domain, therefore, it is our belief that more research efforts should be spent on investigating QoS in specific industrial domains.

Regarding **quality metrics** (*RQ3*), the majority of primary studies focuses on three quality metrics: *response time*, *resource utilization* and *make span*. These results do not imply that the remaining quality metrics have not received enough attention from the research community, but only that they have a more restricted application in the assessment of quality attributes mentioned in the selected primary studies. The aforementioned quality metrics are, to a large extent, used to assess *performance*. However, $\sim 18\%$ of quality metrics have two occurrences in the primary studies, while $\sim 43\%$ have only one. Furthermore, only one primary study contributed by proposing metrics for the assessment of specific quality attributes. This makes us believe that more research should be conducted on the definition of more specific and enhanced quality metrics. Moreover, the mapping between quality attributes and quality metrics showed that $\sim 67\%$ of the quality attributes mentioned in the primary studies were not assessed by any quality metric. We foresee a need for further research targeting these quality attributes with the goal of proposing and identifying quality metrics to assess them.

Regarding **objectives** (*RQ4*) of the primary studies, we found that the majority were interested in *improving* QoS. We could discuss how the research community could have focused more on other objectives such as *measuring* or *defining* QoS, rather than on its *improvement*, but being that QoS is crucial for the adoption of cloud technologies, we understand the rationale behind this. Moreover, the approaches that strive to improve QoS involve the wide use of scheduling techniques and mechanisms. However, our suggestion would be to focus on providing scheduling solutions that take into consideration multiple QoS attributes rather than a single one. Furthermore, we observed a noteworthy focus on using security mechanisms to improve QoS.

7 Threats to validity

External validity. One major threat that could limit the generalizability of our study is having a set of primary studies that is not representative enough of the research on this topic. To mitigate this threat, we conducted an automatic search on two of the largest and most complete electronic databases in software engineering, which was complemented with a fully-recursive backward snowballing activity to eventually enrich the initial corpus of primary studies. Another threat that can jeopardize the external validity of our study is the exclusion of primary studies in languages other than English. Regardless, being that English is the de-facto standard language used for scientific papers, this threat is negligible. Moreover, we only targeted peer-reviewed publications as they are expected to provide scientific work of certain quality assessed by peers, thus excluding gray literature.

Internal validity. To mitigate biases regarding the degree of influence of external variables on the design of the study, we rigorously defined and validated a detailed study protocol that follows the guidelines proposed by Petersen et al. (2015), Kitchenham and Charters (2007), and Wohlin et al. (2012). Furthermore, we defined a classification framework for the extraction of data from the primary studies, that was iteratively revised. The keywording process was also used to transform qualitative data into quantitative data, while quantitative data was analyzed using descriptive statistics, making the data analysis validity threats minimal. Finally, a complete replication package[1] is available for the independent replication of our study.

Construct validity. To mitigate the threat of having a set of primary studies that is not representative enough of the population defined by the research questions, we conducted an automatic search on two electronic databases and complemented the search with a fully-recursive backward snowballing activity. Furthermore, we conducted preliminary searches and refined the search string based on the analysis of a set of sample studies. Finally, all relevant studies were screened according to deterministic and clear selection criteria.

Conclusion validity. Our study has been conducted using widely accepted systematic methods documented in the study design. The study design and the replication package can be used by other researchers to verify and replicate our work. We are aware that they may design a classification framework with different categories and data items than ours, thus leading to different results. However, we mitigated this potential threat by: (i) documenting how we constructed the classification framework, (ii) refining the categories and data items throughout the entire data extraction process, and (iii) making the replication package publicly available for transparency and replicability.

8 Related work

This section discusses other existing systematic studies related to QoS in cloud computing. However, as to our best knowledge, there is no existing systematic mapping study on QoS in industrial cloud computing with a focus on quality attributes. Therefore, in the following, we provide only an overview of the few existing studies related to our topic.

Lehrig et al. (2015) conducted a systematic literature review in the cloud computing context with respect to three quality attributes (scalability, elasticity, and efficiency), to recommend relevant definitions and metrics. Their findings can be used to evaluate the QoS of cloud technologies and as a starting point to derive new quality metrics. Abdelmaboud et al. (2015) conducted a systematic mapping study to offer insights on the state-of-art of QoS approaches in cloud computing. Most of the studies focus on Infrastructure-as-a-Service and Software-as-a-Service. The challenges and gaps identified in the study confirm that more research must be conducted on QoS approaches. Prakash et al. (2019) conducted a literature review and a comparative study of QoS management techniques in cloud computing. They present a list of eight quality attributes and their respective metrics that can help with the improvement of performance in cloud systems and compare various QoS techniques. Nevertheless, they state that there are still lots of possibilities on how QoS in cloud computing can be further enhanced.

The related works presented in this section highlight the fact that the existing research conducted at the intersection of cloud computing and QoS is generic and does not focus on the industrial side of cloud computing, which we instead focus on.

9 Conclusions

This paper reports on a systematic mapping study on QoS in industrial cloud computing with the goal of identifying and classifying the most commonly addressed quality attributes and related metrics, and highlighting the research gaps. The initial set of potentially relevant studies consisted of 1063 publications. The selection of the primary studies through a rigour and well-documented process led to a final set of 42 primary studies. For the extraction of relevant information from the set of studies, we defined and incrementally refined a classification framework that can also be used for future research. For data synthesis, we carried out a vertical and horizontal analysis.

The main results of this study are the following:

- Research on QoS in industrial cloud computing largely focuses on providing *solution proposals* that contribute with novel *models* and *methods*.
- There is a need for more solid *validation* of proposed solutions that can contribute to the maturity of the research area and the adoption of these solutions in practice.
- The adoption of cloud technologies in industry is closely related to the *performance* indicators offered by these technologies. However, research on other quality attributes is quite limited.
- Despite the undeniable importance of security in cloud technologies, the results suggest a lack of research on *security* as a quality attribute and on how to mitigate the security threats when providing *virtualization* solutions.
- The most mentioned quality metrics in industrial cloud computing are *response time*, *resource utilization* and *make span*, that to a large extent are used to assess *performance*.
- The approaches are in most cases not targeting explicitly a specific industrial domain, potentially hampering the identification of relevant quality attributes and metrics for specific applications.
- There is strong scientific research on the impact the *virtualization* solutions have on QoS and specifically, *performance*, that can help cloud providers make more informed QoS-aware architectural decisions.
- The focus of research in this area seems to be oriented towards the *improvement* of QoS by using better *scheduling* techniques to tailor quality attributes according to users' expectations.

To summarize, these conclusions provide evidence that there have been research efforts on this topic, but the focus of the current literature still leaves open research challenges. Overall, we believe that our study will be helpful to the research community for the following reasons: (i) the catalog of quality attributes and metrics can be useful for evaluating QoS in industrial cloud computing, and (ii) the catalog of factors affecting QoS can help make more informed decisions regarding cloud architecture with respect to how it affects specific quality attributes.

Acknowledgments

This work was supported by Vinnova through the ITEA3 BUMBLE project (rn. 18006).

REFERENCES

Abdelmaboud, A., Jawawi, D. N., Ghani, I., Elsafi, A., and Kitchenham, B. (2015). Quality of service approaches in cloud computing: A systematic mapping study. *Journal of Systems and Software*, 101:159–179.

Ardagna, D., Casale, G., Ciavotta, M., Pérez, J. F., and Wang, W. (2014). Quality-of-service in cloud computing: modeling techniques and their applications. *Journal of Internet Services and Applications*, 5(1):11.

Chung, L., Nixon, B. A., Yu, E., and Mylopoulos, J. (2012). *Non-functional requirements in software engineering*, volume 5. Springer Science & Business Media.

Cruzes, D. S. and Dybå, T. (2011). Recommended steps for thematic synthesis in software engineering. In *2011 international symposium on empirical software engineering and measurement*, pages 275–284. IEEE.

Dromey, R. G. (1995). A model for software product quality. *IEEE Transactions on software engineering*, 21(2):146–162.

Fernando, W., Wijayarathne, D., Fernando, J., Mendis, M., and Guruge, I. (2014). The importance of software metrics: Perspective of a software development projects in sri lanka. In *SAITM Research Symposium on Engineering Advancements*.

Guo, S. and Deze, Z. (2019). *Cyber-Physical Systems: Architecture, Security and Application*. Springer International Publishing, 1st edition.

Kitchenham, B. A. and Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering technical report. *Software Engineering Group, EBSE Technical Report, Keele University and Department of Computer Science University of Durham*, 2.

Lehrig, S., Eikerling, H., and Becker, S. (2015). Scalability, elasticity, and efficiency in cloud computing: A systematic literature review of definitions and metrics. In *Proceedings of the 11th International ACM SIGSOFT Conference on Quality of Software Architectures*, pages 83–92.

Mairiza, D., Zowghi, D., and Nurmuliani, N. (2010). An investigation into the notion of non-functional requirements. In *Proceedings of the 2010 ACM Symposium on Applied Computing*, pages 311–317.

Petersen, K., Vakkalanka, S., and Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, 64:1–18.

Petticrew, M. and Roberts, H. (2008). *Systematic reviews in the social sciences: A practical guide*. John Wiley & Sons.

Prakash, S. et al. (2019). Review of quality of service based techniques in cloud computing. In *Data Science and Big Data Analytics*, pages 255–265. Springer.

Software & Systems Engineering Standards Committee (1998). IEEE Std 1061-1998—IEEE standard for a software quality metrics methodology. *IEEE Computer Society, Tech. Rep*.

Wlodarczyk, T. W., Rong, C., and Thorsen, K. A. H. (2009). Industrial cloud: Toward inter-enterprise integration. In *IEEE International Conference on Cloud Computing*, pages 460–471. Springer.

Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proceedings of the 18th international conference on evaluation and assessment in software engineering*, pages 1–10.

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., and Wesslén, A. (2012). *Experimentation in software engineering*. Springer Science & Business Media.