



Design, monitoring, and testing of microservices systems: The practitioners' perspective[☆]

Muhammad Waseem^a, Peng Liang^{a,*}, Mojtaba Shahin^b, Amleto Di Salle^c, Gastón Márquez^d

^a School of Computer Science, Wuhan University, 430072 Wuhan, China

^b Department of Software Systems and Cybersecurity, Faculty of Information Technology, Monash University, 3800 Melbourne, Australia

^c Department of Information Engineering, Computer Science and Mathematics, University of L'Aquila, I-67100 L'Aquila, Italy

^d Department of Electronics and Informatics, Federico Santa María Technical University, 4030000 Concepción, Chile

ARTICLE INFO

Article history:

Received 10 November 2020

Received in revised form 30 June 2021

Accepted 6 August 2021

Available online 19 August 2021

Keywords:

Microservices architecture

Design

Monitoring

Testing

Industrial survey

ABSTRACT

Context: Microservices Architecture (MSA) has received significant attention in the software industry. However, little empirical evidence exists on design, monitoring, and testing of microservices systems.

Objective: This research aims to gain a deep understanding of how microservices systems are designed, monitored, and tested in the industry.

Methods: A mixed-methods study was conducted with 106 survey responses and 6 interviews from microservices practitioners.

Results: The main findings are: (1) a combination of domain-driven design and business capability is the most used strategy to decompose an application into microservices, (2) over half of the participants used architecture evaluation and architecture implementation when designing microservices systems, (3) API gateway and Backend for frontend patterns are the most used MSA patterns, (4) resource usage and load balancing as monitoring metrics, log management and exception tracking as monitoring practices are widely used, (5) unit and end-to-end testing are the most used testing strategies, and (6) the complexity of microservices systems poses challenges for their design, monitoring, and testing, for which there are no dedicated solutions.

Conclusions: Our findings reveal that more research is needed to (1) deal with microservices complexity at the design level, (2) handle security in microservices systems, and (3) address the monitoring and testing challenges through dedicated solutions.

© 2021 Elsevier Inc. All rights reserved.

1. Introduction

Microservices Architecture (MSA)¹ inspired by Service-Oriented Architecture (SOA) has gained immense popularity in the past few years (Dragoni et al., 2017b). The International Data Corporation (IDC) predicted that by the end of 2021, 80% of cloud software will be developed using the MSA style (Balalaie et al., 2016). With the MSA style, an application is designed as a set of business-driven microservices that can be developed, deployed, and tested independently (Taibi et al., 2021). MSA follows a “share-as-little-as-possible” architecture approach and communicates with each other through an API layer (Tulie, 2020; Richards, 2015). In

contrast, SOA adopts a “share-as-much-as-possible” architecture approach and typically uses the Enterprise Service Bus for communication purpose (Tulie, 2020; Richards, 2015). Another key difference between MSA and SOA is that MSA advocates one data storage per microservice, while SOA uses one data storage for the whole system (Tulie, 2020). Moreover, the systems that adopt the MSA style (i.e., microservices systems) have a better fault-tolerance than SOA-based systems (Rajput, 2018). The advantages of the MSA style are enormous, and the key drivers of adopting MSA are faster delivery, improved scalability, and greater autonomy compared to other architectural styles (e.g., SOA) (Jamshidi et al., 2018).

There is a growing body of literature investigating different aspects of MSA (Waseem et al., 2020b,a; Di Francesco et al., 2019; Alshuqayran et al., 2016; Soldani et al., 2018; Pahl and Jamshidi, 2016; Li et al., 2020). One notable research area is how to migrate a monolithic system to MSA (e.g., Balalaie et al., 2016; Fan and Ma, 2017; Taibi et al., 2017; Di Francesco et al., 2018). Others propose solutions for securing microservices systems (e.g., Dragoni et al., 2017b; Pereira-Vale et al., 2019). Researchers have

[☆] Editor: Fabio Palomba.

* Correspondence to: School of Computer Science, Wuhan University, China.

E-mail addresses: m.waseem@whu.edu.cn (M. Waseem),

liangp@whu.edu.cn (P. Liang), mojtaba.shahin@monash.edu (M. Shahin),

amleto.disalle@univaq.it (A.D. Salle), gaston.marquez@usm.cl (G. Márquez).

¹ All the abbreviations used in this paper are provided in Table 24.

recently shown an increasing interest in understanding how microservices systems are developed and migrated in the industry (e.g., Pahl et al., 2018; Kang et al., 2016; Taibi et al., 2018b). Knoche and Hasselbringa conducted a survey study to investigate the drivers, barriers, and goals of adopting MSA (Knoche and Hasselbring, 2019). Ghofrani and Lübke reported the services boundaries identification techniques (e.g., Domain-Driven Design (DDD)) and notations used to design and develop microservices systems (Ghofrani and Lübke, 2018). The survey study conducted by Viggato et al. reports the popular programming languages, technologies, advantages, and challenges in microservices systems (Viggato et al., 2018). Haselböck et al. surveyed MSA experts to identify essential design areas (e.g., infrastructure design) and design challenges of microservices systems (Haselböck et al., 2018).

In another study, Baškarada et al. empirically explored the practitioners' perspective on the advantages of MSA over monolithic architectures and the challenges faced by MSA practitioners when developing microservices systems (Baškarada et al., 2018). Zhang et al. examined the characteristics, benefits, and costs of MSA in practice (Zhang et al., 2019). Zhou et al. investigated the fault analysis and debugging of microservices through an industrial survey (Zhou et al., 2021). In the same line, Wang et al. conducted a mixed-method study (i.e., interview and survey) with practitioners to collect the best practices, challenges, and solutions for the successful development of microservices systems (Wang et al., 2021). Taibi et al. conducted a mixed-method study (i.e., interview and survey) to investigate the motivations (e.g., maintainability, scalability), issues (e.g., monolith decoupling, database migration, and data splitting), and benefits (e.g., maintainability improvement, scalability improvement) of microservices architecture (Taibi et al., 2017). They also developed a migration process framework which is composed of three parts (i.e., system structure analysis, definition of the system architecture, and prioritization of feature development) (Taibi et al., 2017). The mixed-method study by Di Francesco et al. focused on the activities (e.g., domain decomposition, services identification) and challenges (e.g., high coupling among parts of the legacy systems, identification of the boundary of each service) that practitioners face when migrating towards MSA (Di Francesco et al., 2018).

Despite these efforts, it is little empirically known how the software industry designs, monitors, and tests microservices systems. These three aspects play a significant role in succeeding with microservices systems (Rajput, 2018; Newman, 2015). The design of a microservices system (i.e., MSA design) provides a basis for communicating architectural decisions among stakeholders, exposing system structure, realizing use cases, and addressing both functional and quality requirements (Richardson and Smith, 2016). However, to maintain the benefits of microservices systems (e.g., flexibility and separation of concerns), it is essential to reduce complexity at the design level. Monitoring and testing help to achieve a better understanding of the system and improve quality throughout the entire development process. However, microservices systems are increasingly complex and need more intensive monitoring and testing than other types of systems (e.g., SOA, monolith) (Rajput, 2018). This stems from the unique characteristics of microservices systems: greater complexity, polyglot programming languages, heterogeneous and independent microservices, communication behavior at runtime, deployment infrastructure model (e.g., application server vs. container), frequent changes, and several others (Arijs, 2020; Swersky, 2020). Hence, software development organizations need to develop new monitoring and testing strategies or evolve traditional ones to match the unique characteristics of microservices systems.

This study aims to bridge the gap mentioned above by providing a comprehensive understanding of how microservices systems are designed, monitored, and tested in practice. To this end, we conducted a mixed-methods study, including a relatively large-scale survey with 106 responses and 6 interviews from microservices practitioners. Overall, our results suggest that a combination of DDD and business capability is the most effective strategy to decompose an application into microservices and addressing security and availability is a significant concern for MSA practitioners. We found that “API gateway” and “Backend for frontend” patterns are the most used MSA patterns for providing the entry point to clients' requests of microservices systems. We observed that MSA practitioners consider “clearly defining the boundaries of microservices” as the most critical challenge when designing microservices systems. Concerning monitoring microservices systems, most of the respondents collect metrics that support monitoring quality aspects, such as performance (e.g., resource usage), availability (e.g., ratio of the system available time to the total working time), scalability (e.g., load balancing), and reliability (e.g., service level agreement). About testing microservices systems, we found that End-to-End (E2E) testing and contract related (e.g., consumer-driven) testing are gaining space in the industry. Our findings also indicate that the complexity of microservices systems poses challenges for their design, monitoring, and testing.

This study makes the following three major contributions:

1. It brings quantitative and qualitative evidence of how practitioners design, monitor, and test microservices systems.
2. It identifies the challenges that practitioners face and the solutions employed by them when designing, monitoring, and testing microservices systems.
3. It provides an in-depth discussion on the findings from the microservices practitioners' perspective and implications for researchers and practitioners.

The rest of this paper is structured as follows: Section 2 presents the background of the survey. Section 3 describes the research method in detail. Section 4 report the survey results. Section 5 and Section 6 discuss the findings and threats to the validity of the survey respectively. Section 7 summarizes related work. Section 8 concludes this survey with future research directions.

2. Background

Microservices is an architecture style where an application is built as a set of independent services. The main characteristics of the MSA style are (i) *small*, i.e., a service accomplishes a single business capability; and (ii) *autonomous*, i.e., a single deployable unit contains all needed data to be independently executed (Newman, 2015). In contrast, a monolithic system implements features as a single unit to be deployed as a whole (Newman, 2020). In recent years, microservices have become popular, and the software industry has moved from a monolithic architecture to MSA. A substantial body of literature (e.g., Jamshidi et al., 2018; Ponce et al., 2019; Dragoni et al., 2017a; Esposito et al., 2016; Yarygina and Bagge, 2018; Cerny et al., 2018) has concluded that MSA is significantly different from monolithic architecture and SOA in terms of design, implementation, test, and deployment. MSA contains a suite of fine-grained services working together as a single application that can be executed independently, while in a monolithic architecture, all functionalities are encapsulated into one single application that cannot be executed independently (Ponce et al., 2019). MSA is more fault-tolerant, highly scalable, and loosely coupled than monolithic architecture (Ponce et al., 2019). MSA and SOA are also different in several aspects. For instance,

MSA-based systems can use shared or single database dedicated to a specific microservice, whereas SOA-based systems can only use a shared database (Cerny et al., 2018). In MSA, each service can run its own processes and communicate via lightweight mechanisms (e.g., HTTP, REST), whereas in SOA, each service shares a common communication mechanism (e.g., Enterprise Service Bus) (Richards, 2015).

This section presents different aspects related to design, monitoring, and testing of microservices systems that are investigated in this study.

2.1. Design of microservices systems

According to Bass et al. (2012) “the software architecture of a system is the set of structures needed to reason about the system, which comprises software elements, relations among them, and properties of both”. The design elements of microservices systems are (micro)services, views (i.e., solutions), processes, tools, culture, and organization structure (Nadareishvili et al., 2016). This survey investigates several aspects of the MSA design, such as application decomposition and implementation strategies, architecting activities, description methods, design patterns, Quality Attributes (QAs), skills required to design and implement MSA, and MSA design challenges and their solutions. We briefly describe some of these aspects below.

Architecting activities: The architecture of a software-intensive system is the result of various architecting activities that are performed by related stakeholders. Hofmeister et al. defined a general model of architecture design, which is composed of a set of architecting activities, including architectural analysis, synthesis, and evaluation (Hofmeister et al., 2007). Architecting activities are widely explored in the context of monolithic architecture, but the activities are rarely investigated in the context of MSA. We identified only one Systematic Mapping Study (SMS), i.e., (Di Francesco et al., 2019), that reports architecting activities in the context of MSA. In this study, we investigated the level of agreement and disagreement of the practitioners about the general architecting activities employed in the context of MSA. According to the literature (e.g., Hofmeister et al., 2007; Tang et al., 2010; Li et al., 2013), the general architecting activities we considered for MSA include Architectural Analysis (AA), Architectural Synthesis (AS), Architectural Evaluation (AE), Architectural Implementation (AI), and Architectural Maintenance and Evolution (AME).

MSA description methods: Architecture description methods consist of a set of practices that are used to express the architecture of a system-of-interest (ISO/IEC/IEEE, 2011). To describe the architecture of software systems, architects may use various notations (Malavolta et al., 2012), such as boxes and lines, Unified Modeling Language (UML), Architectural Description Languages (ADLs), and Domain-Specific Languages (DSLs). In this study, we investigated the description methods and diagrams used by microservices practitioners to describe and represent the architecture of microservices systems.

MSA design patterns: Patterns are reusable solutions to occurring problems in specific contexts (Gamma et al., 1994). Since the emergence of MSA, dozens of patterns have been documented for implementing different aspects (e.g., data management, deployment) of microservices systems (Alshuqayran et al., 2016; Richardson, 2018; Márquez and Astudillo, 2018; Marquez et al., 2018; Taibi et al., 2018a). We investigated how often microservices practitioners use various patterns to address the problems related to design of MSA-based systems, such as application decomposition, data management, deployment, services communication, server-side service discovery, and security in the context of microservices systems.

Quality attributes: A QA is a “measurable or testable property of a system that is used to indicate how well the system satisfies the needs of its stakeholders” (Bass et al., 2012). Architecture design of software systems is the first phase of software development, in which QAs can be satisfied. Some empirical studies (e.g., Di Francesco et al., 2019; Alshuqayran et al., 2016; Pahl and Jamshidi, 2016; Li et al., 2020) on MSA have taken QAs into consideration, but the results reported are not from the practitioners’ perspective. The ISO/IEC 25010 standard regulates a comprehensive set of QAs of software systems (ISO/IEC, 2016), but these QAs are considered with different priorities in the design of microservices systems (Waseem et al., 2020b). Thus, we decided to investigate the importance of those QAs that are reported in the existing studies (e.g., Márquez and Astudillo, 2018; O’Brien et al., 2007; Li et al., 2020) in the context of MSA.

Besides the aspects described above, we also investigated several other aspects, such as application decomposition strategies (e.g., DDD), development processes (e.g., DevOps, Scrum, XP), architectural components for microservices systems, and skills required to design MSA. Also, we identified the MSA design challenges, solutions, as well as the impact of the MSA challenges on monitoring and testing of microservices systems.

2.2. Monitoring of microservices systems

A microservices system is primarily based on isolated microservices that can be developed, deployed, and operated independently (Lewis and Fowler, 2019). Microservices interact with each other at runtime to exchange information. The runtime architecture of microservices systems is continuously evolved. Thus, information on the system architecture and service interaction and behavior needs to be generated and updated automatically (Mayer and Weinreich, 2017). The dynamic nature of microservices systems needs monitoring infrastructure to diagnose and report errors, fault, failure, and performance issues. In this study, we investigated the monitoring metrics, monitoring practices, tools, challenges, and solutions for monitoring microservices systems.

Monitoring metrics are numerical representations of data measured over intervals of time that generally come from resource usage and system behavior (Sridharan, 2018; Swersky, 2020). Monitoring metrics can be used to predict the system’s future behavior. We identified 12 monitoring metrics from grey literature, e.g., (Swersky, 2020; Fowler, 2017; Dave and Degioanni, 2020) (see Fig. 15) that are used for infrastructure-level monitoring and microservices monitoring. Specifically, metrics such as resource usage (e.g., CPU, memory), threads, the status of dependencies, and database connections are used for infrastructure-level monitoring. In contrast, language-specific metrics (e.g., garbage collector behavior, database latency), availability, Service-Level Agreement (SLA), endpoint success, and errors and exceptions are used to monitor microservices (Fowler, 2017).

Monitoring practices concern with how to monitor microservices systems. We identified six monitoring practices from grey literature (e.g., Rajput, 2018; Swersky, 2020; Carstensen, 2020), i.e., log management, exception tracking, health check API, log deployment, audit logging, and distributed tracking. Each of these practices consists of several activities. For instance, log management is related to generating, collecting, centralizing, parsing, transmitting, archiving, and disposing log data (e.g., errors, failure) produced by applications servers, network devices, and other components (Carstensen, 2020). Similarly, exception tracking is a monitoring practice with which exceptions are identified, understood, and resolved with monitoring tools. Health check API is a service implemented within microservices systems to get the operational status and dependencies of microservices. It also

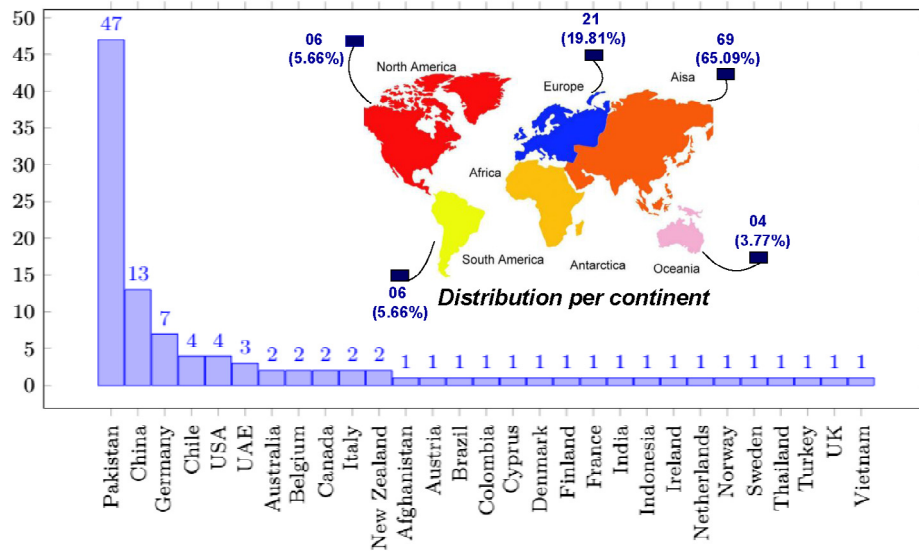


Fig. 1. The number of respondents from each country.

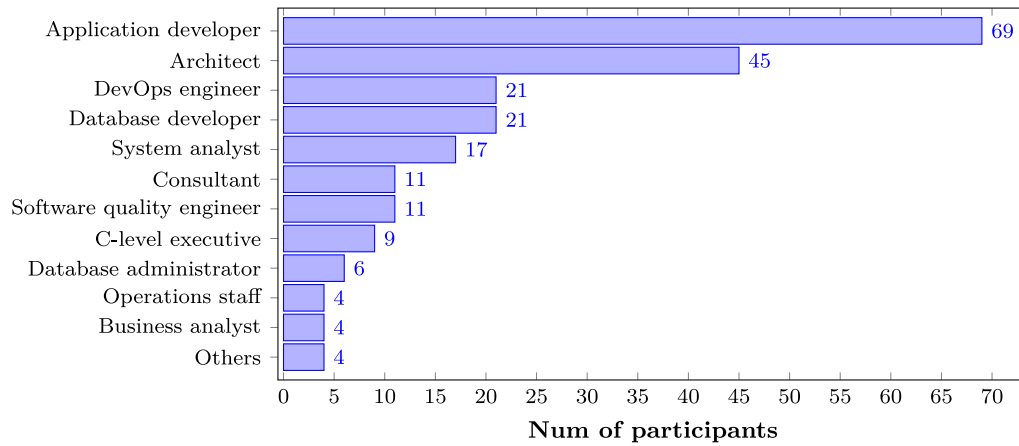


Fig. 2. Survey participants' responsibilities in their companies.

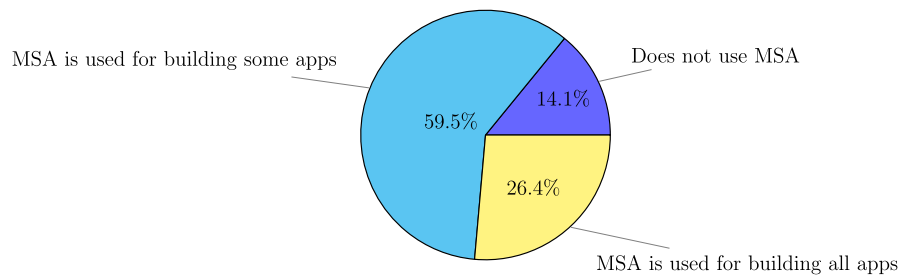


Fig. 3. Status of MSA adoption across organization.

returns the performance information (e.g., connection time) of the microservices (Richardson, 2018).

Monitoring tools can be classified into two categories: libraries and platforms. Monitoring libraries are used during the development of microservices and permit collecting the application data. In contrast, monitoring platforms allow gathering and analyzing data from different sources, such as the hosts, infrastructure, and microservices (Swersky, 2020). We identified 13 candidate monitoring tools (see Table 27), which are listed in our survey question, from the literature (e.g., Rajput, 2018; Waseem et al., 2020b; Swersky, 2020; Carstensen, 2020).

2.3. Testing of microservices systems

Software testing is intended to identify problems and to fix them for improving the quality of software systems (Saha, 2008). However, the testing of microservices systems requires extra effort due to, for example, the polyglot code base in multiple repositories, feature branches, and database per service (Lumetta, 2019). Some other aspects that bring difficulties in microservices testing are inter-process communication, inter-service dependencies, cloud infrastructure, and third-party services components (Waseem et al., 2020a; Sotomayor et al., 2019). We identified several studies (e.g., Rajput, 2018; Lumetta, 2019; Rahman

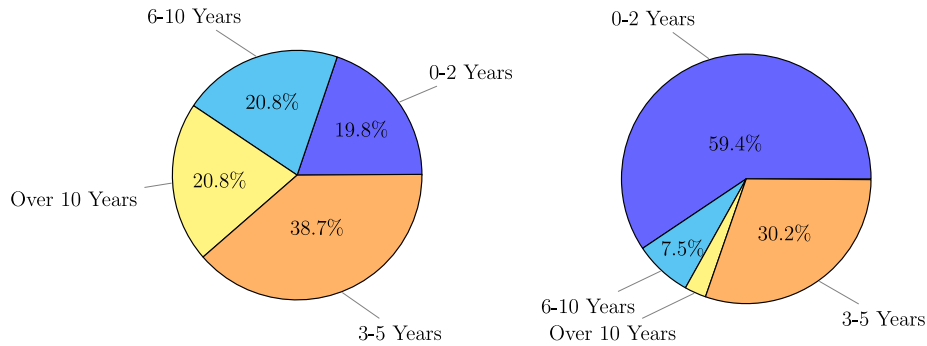


Fig. 4. Experience of the respondents in the IT industry (Left) and the development of microservices systems (Right).

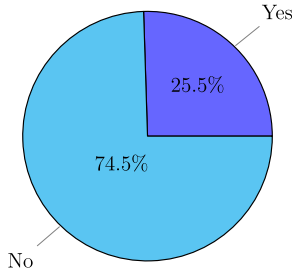


Fig. 5. Participants professional training related to MSA.

and Gao, 2015; Heorhiadi et al., 2016; Meinke and Nycander, 2015; Clemson, 2020; Dawid, 2020) in which microservices testing was discussed. However, we did not find any study that presents the empirical evidence on testing of microservices systems from the practitioners' perspective. Our previous work also indicates that there is a lack of knowing the industrial reality concerning testing for MSA-based applications (Waseem et al., 2020a). Therefore, this industrial survey explores different aspects of testing microservices systems, such as testing strategies and skills required to test microservices systems.

Testing strategies: A testing process for microservices systems includes a set of testing strategies (Clemson, 2020). These strategies contain functional (e.g., unit testing) and non-functional testing strategies (e.g., performance testing) (Dawid, 2020). The aim of using multiple testing strategies in the testing process is to make sure a developed application successfully operates across various environments and platforms. We identified 12 testing strategies (see Table 28), which are listed in our survey question, from the literature (e.g., Waseem et al., 2020a; Rajput, 2018; Lumetta, 2019; Rahman and Gao, 2015; Heorhiadi et al., 2016; Meinke and Nycander, 2015; Clemson, 2020; Dawid, 2020).

Testing tools: Microservices practitioners may use various testing tools to detect bugs and evaluate applications' behavior, depending on the chosen testing strategy, for instance, tools to perform unit, performance, load, and contract testing. We identified 20 candidate testing tools (see Table 28), which are listed in our survey question, from the literature (e.g., Waseem et al., 2020a; Rajput, 2018; Lumetta, 2019; Rahman and Gao, 2015; Heorhiadi et al., 2016; Meinke and Nycander, 2015; Clemson, 2020; Dawid, 2020).

3. Research method

Our research aims to provide a comprehensive insight into how microservices systems are designed, monitored, and tested in practice. Given this objective and inspired by the guidelines for selecting empirical methods for software engineering research (Easterbrook et al., 2008), we decided to use a mixed-methods study, which collected data from a survey and interviews. We first conducted a survey to reach out to a wide range

Table 1
Research questions and their rationale.

Research questions	Rationale
RQ1: How microservices systems are designed in the industry?	The answer to this RQ helps to understand various aspects of designing microservices systems from the practitioners' perspective. This RQ aims to investigate the team structure, application decomposition strategies, MSA implementation strategies, architecting activities, description methods, design components, QAs, MSA design patterns, skills required to design and implement MSA, design challenges and their solutions, as well as the reasons for the key findings related to the design of microservices systems.
RQ2: How microservices systems are monitored in the industry?	Monitoring microservices systems is challenging because a large number of independent services are communicating with other independent services in many ways. This RQ aims to investigate microservices monitoring metrics, practices, tools, monitoring challenges and their solutions, as well as the reasons for the key findings related to the monitoring of microservices systems.
RQ3: How microservices systems are tested in the industry?	Microservices systems are challenging to test because of several factors, such as polyglot code base, feature branches, database per service. This RQ aims to investigate microservices testing strategies, tools, skills, testing challenges and their solutions, as well as the reasons for the key findings related to the testing of microservices systems.

of practitioners to gain a broad overview of how microservices systems are designed, tested, and monitored. We then carried out in-depth interviews to understand the reasons behind the key findings of the survey. We also followed the guideline proposed by Kitchenham and Pfleeger (Kitchenham and Pfleeger, 2008) and used the survey template recommended by evidence-based software engineering (EBSE, 2019).

3.1. Research questions

We formulated the Research Questions (RQs) in Table 1 based on the study aim presented in Section 1. The findings from this set of RQs are expected to contribute to the growing body of evidential knowledge about the state of the practice of design, monitoring, and testing of microservices systems.

3.2. Survey

This section briefly discusses the design of the survey.

3.2.1. Survey type

We conducted this survey as a descriptive survey with a cross-sectional design. Cross-sectional design surveys are appropriate

Table 2

Microservices practitioners social and professional platforms used to post our survey.

#	Group name	URL	Platform
1	Microservice architecture	https://tinyurl.com/42pazdhj	LinkedIn
2	Microservices meet DevOps in a Docker Container and Kubernetes world	https://tinyurl.com/akwt44jj	
3	Cloud, Kubernetes, Docker, microservices, Gitops discussions	https://tinyurl.com/pfatv6s4	
4	.Net people	https://tinyurl.com/3an82ktz	
5	Software architects and enterprise architects Group	https://tinyurl.com/2r7y4cnw	
6	Microservice architecture	https://tinyurl.com/fh2evwwc	Facebook
7	Microservice Java Spring	https://tinyurl.com/35c4xsf2	
8	Developing microservices	https://tinyurl.com/8nbs9p2f	
9	Spring boot and microservices developers	https://tinyurl.com/yuc658tv	
10	Microservices developer	https://tinyurl.com/3tzbmz85	
11	Microservices community	microservices-community@googlegroups.com	Google groups

for gathering information at a set point of time (Kitchenham and Pfleeger, 2008). Surveys can be administered in many ways, including online questionnaires, phone surveys, or face-to-face interviews (Lethbridge et al., 2005). We decided to conduct the survey study through an online questionnaire because (1) this type of survey is time and cost-effective, (2) our potential participants were from different countries, and (3) online questionnaires are convenient for collecting the responses from a large number of participants (Lethbridge et al., 2005).

3.2.2. Sample and population

We did not limit our target population to a specific region or country (e.g., China). We invited the potential practitioners working in the software industry, and whose organizations had adopted or planned to adopt MSA. We asked several questions about the background of the participants (see Table 25), such as their main responsibilities, number of years working in the IT industry, number of years working with microservices systems, and professional training related to MSA. We used the following approaches to reach the target population.

- 1. Personal contacts and authors of industrial track research papers and web blogs:** We sent an invitation email to our personal contacts who work on microservices in the IT industry and the authors of the industrial track research papers (e.g., conference proceedings) and web blogs related to microservices. We also requested the respondents to disseminate our survey to their colleagues who are eligible to answer the questions or share the contacts of relevant practitioners with us.
- 2. Microservices practitioners on GitHub:** We executed a search string (i.e., “microservice” OR “micro service” OR “micro-service”) in GitHub. We sorted the search results as “best match” to our search string and collected 9,000 publicly available email addresses of MSA-based project contributors. Then we sent an invitation email with a survey link to all practitioners whose email addresses were collected.
- 3. Social and professional platforms:** The authors are active members of several microservices groups on LinkedIn, Facebook, and Google, where microservices practitioners from different countries share their issues, experiences, and knowledge. We posted our survey to all the groups in Table 2.

3.2.3. Obtaining a valid sample

To identify valid answers from the collected responses, we applied the inclusion and exclusion criteria presented in Table 3. We used the criterion I1 to include the participants’ responses and evaluated this criterion by the answers of survey questions DQ2 (we asked the participants their major responsibilities in the companies), DQ3 (we asked the participants their usage of MSA

Table 3

Inclusion and exclusion criteria for selecting valid responses.

Inclusion	I1: The respondent must have responsibilities associated with MSA-based system development and operation.
Exclusion	E1: An inconsistent, randomly filled, or meaningless response is excluded. E2: If more than 5 survey questions (SQs) have inconsistent, randomly filled, or meaningless answers, the whole response is excluded.

for developing applications), and DQ5 (we asked the participants their experience (in years) working with microservices systems). On the other hand, we used the criterion E1 to exclude those responses that were inconsistent, randomly filled, or meaningless. We also used the criterion E2 to exclude the entire response from a respondent if more than 5 SQs have inconsistent, randomly filled, or meaningless answers. This is because practitioners had started our survey but he/she might not have sufficient knowledge or intention to answer all the survey questions. Inconsistent responses mean that the answers to two or more survey questions are conflicting. For example, a participant chose “we use MSA style for building all applications in our organization/team” to survey question “DQ3: Regarding the usage of MSA for developing applications.”, and he or she also answered “I have no idea” to survey question “SQ12: How do you decompose an application into microservices?”. We also found several answers that were randomly selected. For example, a respondent answered the usage of all the MSA design patterns as “Very often”, answered all the quality attributes as “Very important”, and rated the severity of all the challenges as “Catastrophic”. Moreover, we considered logically senseless answers to some or all the open-ended survey questions as meaningless. For example, a participant wrote the answer “It is easy to get” to survey question “SQ26: What solutions (e.g., practices, tools, team structures) can be employed to address or mitigate the challenges related to architecting for MSA-based systems?”.

3.2.4. Sampling method

The non-probabilistic sampling method was used to obtain a valid sample for this survey. We used the non-probabilistic sampling because the target population was specific to microservices practitioners, which were hard to identify through a random selection. As a type of non-probabilistic sampling, we also used convenience and snowball sampling. Convenience sampling “refers to acquiring responses from available people who are willing to participate” (Kitchenham and Pfleeger, 2008). The reason for using the convenience sampling method for this survey is that it provides easy access to the target population through personal contacts and social and professional platforms (e.g., GitHub, LinkedIn, Facebook). To maximize the chances of getting more responses, we also used the snowball sampling technique, which asks the

participants in a survey to invite other people who have relevant expertise and agree to participate in the survey (Kitchenham and Pfleeger, 2008).

3.2.5. Survey instrument preparation

In this section, we describe the survey instrument and questionnaire format. Through reviewing the grey literature (e.g., Lewis and Fowler, 2019; Swersky, 2020; Clemson, 2020; Lumetta, 2019), peer-reviewed literature (e.g., Di Francesco et al., 2019; Alshuqayran et al., 2016; Taibi et al., 2018b; Soldani et al., 2018; Pahl and Jamshidi, 2016), and conducting SMSs (i.e., Waseem et al., 2020b,a) on microservices systems, we developed a questionnaire consisting of 39 survey questions (see Appendix B). The survey questions are organized into the following four groups.

- **Demographic information:** We used ten demographics questions (DQ1–DQ10, see Table 25) to collect comprehensive demographic information about practitioners (e.g., country, responsibilities, experience) and their respective companies (e.g., size, work domain, development methods). We asked for demographic information to determine whether the respondents represent microservices practitioners for results generalization purpose.
- **Design of microservices systems:** We formulated 17 survey questions with the purpose of exploring the practitioners' perspective on the design of microservices systems (SQ11–SQ27, see Table 26). The questions are about individuals or teams who are responsible for designing the architecture of microservices systems (SQ11), application decomposition strategies (SQ12), MSA design implementation strategies (SQ13), MSA architecting activities (SQ14–SQ18), MSA description methods (SQ19, SQ20), MSA design components (SQ21), quality attributes (SQ22), and MSA design patterns (SQ23). Moreover, we asked three questions to explore the MSA design challenges (SQ24), the impact of MSA design challenges on monitoring and testing of microservices systems (SQ25), and solutions for the MSA design challenges (SQ26). Finally, the skills required to design and implement MSA are investigated through SQ27.
- **Monitoring of microservices systems:** The third part of our survey consists of six questions that aim to explore practitioner's perspectives on monitoring microservices systems (SQ28–SQ33, see Table 27). We explored the monitoring metrics (SQ28), monitoring practices (SQ29), monitoring tools (SQ30), monitoring challenges (SQ31), the severity of monitoring challenges (SQ32), and the solutions for monitoring challenges (SQ33).
- **Testing of microservices systems:** The survey questions (SQ34–SQ39, see Table 28) are used to solicit the participants' perspectives on the testing strategies (SQ34), tools (SQ35), and challenges (SQ36), the severity of testing challenges (SQ37), skills required to test microservices systems (SQ38), and the solutions for testing challenges (SQ39).

3.2.6. Questionnaire format

The survey starts from the Welcome page to explain the purpose and design of the survey and provides our contact information (see Appendix B). Our survey questionnaire has both closed-ended (e.g., SQ11, SQ12, SQ13) and open-ended questions (e.g., SQ26, SQ33, SQ39). The closed-ended questions of our survey are composed of multiple-choice, rating scale, Likert scale, matrix, and drop-down questions. The open-ended questions in our survey allow the participants to provide their solutions for design, monitoring, and testing challenges. We used Google Forms for the implementation of the survey questionnaire.

3.2.7. Pilot survey

We performed a pilot survey to ensure that (i) the length of the survey is appropriate, (ii) the terms used in the survey are clear and understandable, and (iii) the answers to survey questions are meaningful. We recruited 10 practitioners for the pilot survey through our personal contacts and sending emails to publicly available email addresses of microservices practitioners on GitHub. We got only a few suggestions to refine some adopted terms, and we incorporated them in the final version of the survey instrument.

3.3. Interviews

To get deeper insights into the survey results, we conducted follow-up interviews with microservices practitioners. We present the detailed procedure of the interview process below.

3.3.1. Participant selection

We recruited microservices practitioners from IT companies in five countries: Pakistan, Chile, Netherlands, Sweden, and Colombia. Interviewees were recruited by sending an email to some of the survey's original respondents and personal contacts in each country. Since the interviews were expected to last about 40 to 55 min in order to ensure a sufficient time allocation to each question, we decided to focus on gaining a deep insight into the key findings obtained from the survey. We informed the participants that this interview is entirely voluntary with no compensation. We finally recruited 6 interviewees and asked them to provide their demographics information (see Table 29), and the information about the interviewees is presented in Table 4.

The roles of the interviewees are mainly related to the design and development of microservices systems. However, they also actively participated in other microservices system development stages, such as testing, monitoring, and deployment. Their average experience in the IT industry is 11.7 years (minimum 4 years and maximum 20 years). The interviewees' average experience on microservices is 3.5 years (minimum 1 year and maximum 6 years). We refer to the interviewees as P1 to P6.

3.3.2. Interviewing process

The first, fourth, and fifth authors of the study conducted the interviews. The average time for all interviews is 53 min. The interviews were conducted in Urdu, Italian, and Spanish, and later were translated into English. The interviews are semi-structured and divided into four sections.

- **Demographic information:** We asked five questions regarding interviewees' demographics, such as experience in the IT industry and microservices, responsibilities, and organization domains.
- **Design:** We asked 13 open-ended questions to get a deeper understanding of seven key findings concerning microservices system design. The questions include: (i) confirmation of results about the combination of DDD and business capability, (ii) reasons for using a combination of DDD and business capability strategy, (iii) decomposition strategies other than DDD and business capability, (iv) reasons for creating the only high-level design for microservices systems, (v) possible risks if practitioners do not create a design or detailed design for microservices systems, (vi) reasons for the frequent use of Architecture Evaluation (AE) activity, (vii) reasons for paying more attention to architecting activities close to the solution space instead of the problem space, (viii) reasons for not using ADLs and DSLs, (ix) reasons for using informal and semi-formal languages, (x) reasons for considering security, availability, performance, and scalability as most important, (xi) reasons for using API

Table 4
Interviewees and their information.

#	Responsibilities	Domain	Experience in		Country
			Microservices	IT Industry	
P1	Design and development	Telecom	6 years	17 years	Pakistan
P2	Design and development	E-commerce	5 years	15 years	Sweden
P3	Design, development, and maintenance	Banking	4 years	4 years	Netherlands
P4	Design and development	Telecom	3 years	9 years	Pakistan
P5	Development	Telecom	2 years	5 Years	Colombia
P6	Design and development	Telecom	1 year	20 Years	Chile

gateway and Backend for frontend, Database per service, and Access token patterns as most often, (xii) reasons for the most prominent challenges of designing microservices systems, and (xiii) solutions for the most prominent design challenges.

- **Monitoring:** We asked 5 open-ended questions to get a deeper understanding of four key findings from the survey regarding microservices system monitoring. The questions include: (i) reasons for frequent use of the monitoring metrics, like resource usage, load balancing, and availability, (ii) reasons for frequent use of the monitoring practices, like log management, exception tracking, and health check APIs, (iii) opinions about using Jira for monitoring purpose, (iv) reasons for the most prominent challenges of monitoring microservices systems, and (v) solutions for the most prominent monitoring challenges.
- **Testing:** The interviewees were asked four questions to elaborate on the key findings obtained from the survey regarding microservices system testing. The questions include: (i) reasons for frequent use of the testing strategies, like unit testing, integration testing, and E2E testing, (ii) reasons for frequent use of the testing tools, like Junit, JMeter, and Mocha to test microservices systems, (iii) reasons for the most prominent challenges of testing microservices systems, and (v) solutions for the most prominent testing challenges.

3.4. Data analysis

3.4.1. Survey data analysis

We used two techniques for analyzing the obtained data from survey participants. First, a descriptive statistic technique was used to analyze the closed-ended questions' responses (e.g., SQ11, SQ12, SQ13). Second, we used open coding and constant comparison techniques from Grounded Theory (GT) (Glaser and Strauss, 2017; Hoda and Noble, 2017) to analyze the qualitative response to the open-ended questions (e.g., SQ26, SQ33, SQ39). The relationship between survey questions, Interview Questions (IQs), data analysis methods, and RQs are shown in Table 5.

3.4.2. Interview data analysis

The main goal of this step is to identify the reasons obtained from the interviews for the key findings of the survey results. The interviews were translated and transcribed by the first, fourth, and fifth authors from audio recordings to text transcripts. Before applying the open coding and constant comparison techniques, the second and third authors read the transcribed interviews and highlighted the irrelevant information. We dropped several sentences that are not related to "design, monitoring, and testing of microservices systems". After removing the extraneous information from the transcribed interviews, the first author read and coded the interview transcripts' contents to get the answers to the interview questions. The second and third authors of the study verified the coding results and provided suggestions to improve the quality of coding results. We used the MAXQDA tool² for analyzing and coding the qualitative data.

Table 5
Relationship between questions, data analysis methods, and research questions.

RQs/Section	Survey and interview questions	Data analysis method
Demographics	DQ1-DQ10	Descriptive statistics
RQ1	SQ11-SQ25	Descriptive statistics
RQ2	SQ28-SQ32	Descriptive statistics
RQ3	SQ34-SQ38	Descriptive Statistics
RQ1	SQ27	Open coding and constant comparison
RQ2	SQ33	Open coding and constant comparison
RQ3	SQ39	Open coding and constant comparison
Demographics	IQ1.1-IQ1.5	Descriptive statistics
RQ1	IQ2.1-IQ2.7	Open coding and constant comparison
RQ2	IQ3.1-IQ3.4	Open coding and constant comparison
RQ3	IQ4.1-IQ4.3	Open coding and constant comparison

3.4.3. Result comparison




To better understand practitioners' perspective on design, monitoring, and testing of microservices systems, we divided the responses about Likert scale survey questions (i.e., SQ14-SQ18, SQ19, SQ22, SQ23, SQ29, SQ32, SQ34, SQ37) into three major demographic groups (i.e., Experienced-based comparison, Organization size-based comparison, MSA style-based comparison).

- Participants who have ≤ 2 years of experience with microservices systems (63 responses)
- Participants who have > 2 years of experience with microservices systems (43 responses)
- Participants whose organization or team does not use the MSA style at all (15 responses)
- Participants whose organization or team uses the MSA style for building some or all applications (91 responses)
- Participants who are working in a relatively small and medium organization (≤ 100 employees) (51 responses)
- Participants who are working in a relatively large organization (> 100 employees) (55 responses)

We computed the statistically significant difference (i.e., P -value) for the three groups. We summarized the survey responses to Likert scale survey question statements, SQ#, mean values, P -values, and effect size in Tables 10, 11, 12, 16, 17, 21, and 22. The detail about the gaps between the different demographic groups is reported in Sections 4.2.6, 4.3.5, and 4.4.5. In this survey, we used five points Likert scale for obtaining levels of Agreement (e.g., strongly agree, agree), Frequency (e.g., very often, often), Importance (e.g., very important, important), and Severity (e.g., catastrophic, major) from the practitioners about several aspects of microservices systems (see Likert scale survey questions in Tables 26, 27, and 28). We analyzed each Likert scale level to corresponding survey question option to identify the statistically significant differences between the two chosen groups (e.g., experience ≤ 2 years vs. experience > 2 years) at 95% confidence level. We first use the independent sample T-test by assuming equal variances to test the statistical significance difference between two group means (e.g., experience ≤ 2 years

² <https://www.maxqda.com/>.

vs. experience > 2 years), and then we applied Bonferonni correction for the multiple comparisons on the means of survey question statements (e.g., see Table 10). We also calculated the effect size (Nakagawa and Cuthill, 2007) to quantify the difference between two groups. We used different colors to indicate participants from which groups have more responses about options to survey questions. The responses to survey questions are sorted according to the number of responses. The dark gray color shows that the first group is more likely to agree with survey question options, and the light gray color shows that the second group is more likely to agree with survey question options. We use the following symbols to show the statistically significant difference of the chosen groups.

-  symbol indicates a significant difference between participant groups who have **experience** ≤ 2 years vs. **experience** > 2 years.
-  symbol indicates a significant difference between participant groups who use the **MSA style** vs. **No MSA style**.
-  symbol indicates a significant difference between participant groups who are from an organization that has **Employees size** ≤ 100 vs. **Employees size** > 100.

4. Results

In this section, we provide the our mixed-methods study results according to the collected responses and answer the RQs defined in Section 3.1.

4.1. Demographics

Countries: We received 106 responses to the survey (indicated by R1 to R106), interviewed 6 microservices practitioners (indicated by P1 to P6) from 6 organizations in 6 countries (see Table 4). Respondents came from 29 countries of 5 continents (see Fig. 1) working in diverse teams and roles to develop microservices systems. The majority of them are from Pakistan (47 out of 106, 63.2%), China (13 out of 106, 12.2%), and Germany (7 out of 106, 6.6%), and a close look of the results shows that the respondents from other countries tend to answer similarly to the respondents from the top three countries (i.e., Pakistan, China, Germany).

Responsibilities: Fig. 2 shows that the majority of participants were application developer (69 out of 106, 65.1%), architect (45 out of 106, 42.4%), and DevOps engineer (21 out of 106, 19.8%). Note that one participant may have multiple major responsibilities in the company (DQ2), and consequently the sum of the percentages exceeds 100%.

MSA adoption across organization: As shown in Fig. 3, 59.5% (63 out of 106) of the respondents said that they only used MSA to develop specific applications in their respective organizations. 26.4% (28 out of 106) of the respondents answered that they used MSA to develop all the applications in their organizations. Overall, almost 86% of the respondents are from organizations that have adopted MSA to develop applications. Note that although 14.1% (15 out of 106) of the respondents answered that their organizations did not use MSA to develop applications, these respondents still have experience and expertise with MSA-based systems, which is one of the criteria to answer this survey (see the Welcome page of the survey in Appendix B).

Experience: We asked the participants about their experiences in the IT industry (DQ4) and in the development of microservices systems (DQ5). Fig. 4 (Left) shows that the majority of respondents (41 out of 106, 38.7%) have worked in the IT industry for 3 to 5 years. On the other hand, Fig. 4 (Right) shows that over half of the respondents (63 out of 106, 59.4%)

had little experience (from 0 to 2 years) in the development of microservices systems. We also received three responses in which the participants mentioned that they had more than 10 years of experience working on microservices. However, considering that the term “microservices” was first coined in 2011 (Comunytet, 2020), which is around 9 years ago. We accept these results because (1) responses to all other survey questions from these three participants were valid, (2) the overall working experience of these participants in the IT industry was more than ten years and they were working as “C-level executives”, and (3) we also found the blogs related to microservices (e.g., Stenberg, 2020) in which practitioners claim “microservices have a very long history, not as short as many believe”. For example, the similar term “micro web services” was presented in 2005 (Mauersberger, 2020).

Professional training: Fig. 5 shows that 74.5% (79 out of 106) of the respondents did not receive any professional training to develop microservices systems. Only 25.5% (27 out of 106) of the respondents answered that they had such training. Note that we explicitly mentioned in DQ6 that professional training does not mean higher education (e.g., MSA course at universities).

Organization size and the number of people working on microservices: The size of organizations varies from small (1–20 employees) to large (more than 1000 employees) (see Fig. 6 (Left)). We found that 29.2% (31 out of 106) of the respondents were from medium scale organizations (21–100 employees) and 21.7% (23 out of 106) of the respondents were from large organizations. Similarly, Fig. 6 (Right) shows the number of persons working on one microservice. 43.4% (46 out of 106) of the respondents stated that they were working in a group of 2 to 3 people, and 27.4% (29 out of 106) answered that they were working in a group of 4 to 5 people.

Application domains and development methods: Fig. 7 (Left) shows the domains of the participants' organizations where the survey participants worked. “E-commerce” (35 out of 106, 33.1%), “Professional services” (24 out of 106, 22.6%), and “Financial software” (23 out of 106, 21.6%) are the most popular domains. We also found that Scrum (63 out of 106, 59.4%) and DevOps (53 out of 106, 50.0%) are the two widely used development methods in the participants' organizations to develop microservices systems (see Fig. 7 (Right)). Other popular development methods are iterative development and Rapid Application Development (RAD), which were mentioned by 25 and 14 participants, respectively. We also received 12 responses in which the participants indicated that they did not employ any development method. Note that one organization may adopt one or more development methods in their MSA projects.

Demographics

Finding 1. The majority of participants did not receive any training to develop microservices systems.

Finding 2. The participants indicated that a team of 2–5 people often works on (develop, test, and deploy) each microservice.

Finding 3. Scrum and DevOps are the two most widely used development methods in MSA projects.

4.2. Design of microservices systems (RQ1)

To realize how microservices systems are designed in the industry (RQ1), we asked the respondents to answer 17 survey questions (SQ11 to SQ27) (see Table 26 in Appendix B).

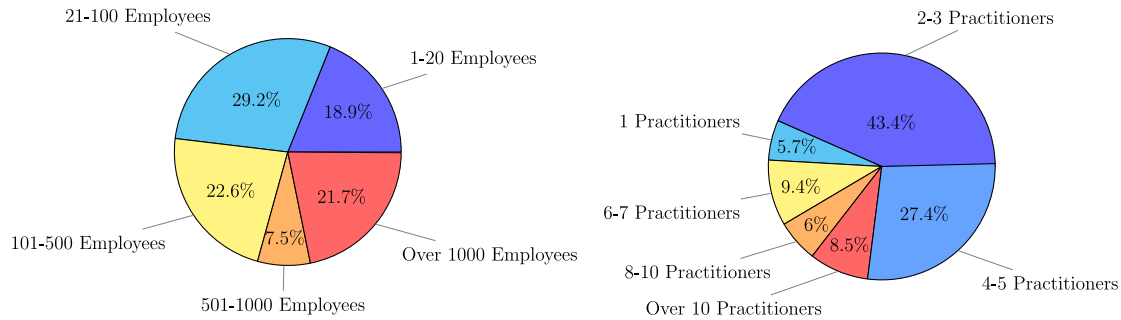


Fig. 6. An overview on organization size (Left) and number of people working on one microservice (Right).

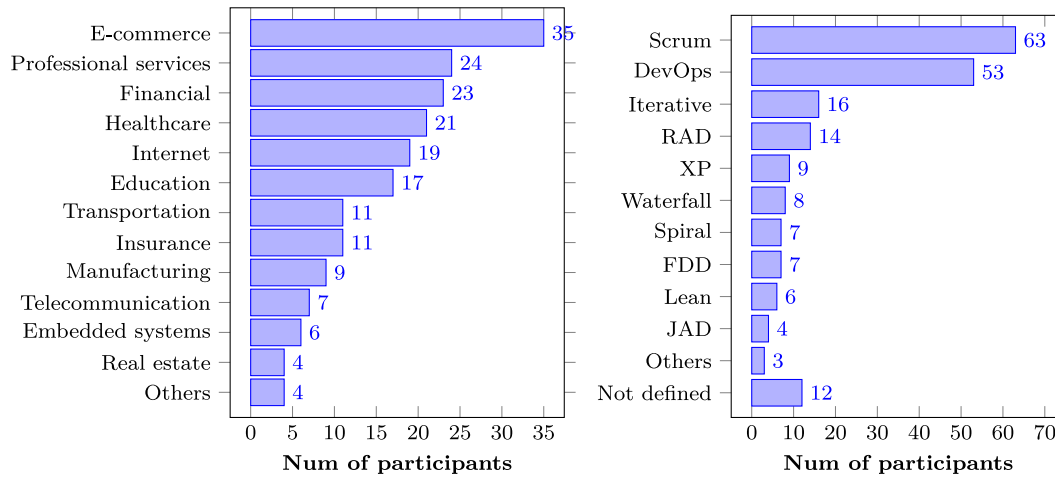


Fig. 7. An overview on the domains of the organizations (Left) and the software development methods employed (Right).

4.2.1. Team structure, application decomposition, and implementation strategies

We asked the survey respondents about whether a dedicated individual or team is responsible for creating and managing the architecture of microservices systems (SQ11) in their organizations. 36.7% (39 out of 106) of the respondents said that their organizations do not have a dedicated person or team. We further investigated our results and found that most of these respondents are from small companies or small teams that possibly cannot afford a software architect. In contrast, 31.1% (33 out of 106) answered that their organizations have such a team or person. Moreover, 32.1% (34 out of 106) of the respondents shared that they have a dedicated team or person designing MSA, who has other responsibilities as well.

We were also interested in understanding how an application is decomposed into microservices in the industry (SQ12) (see Fig. 8). The number of valid responses to this survey question is 104 because two responses are meaningless. 42.3% (44 out of 104) of the respondents answered that they used a combination of business capability and DDD to decompose an application into microservices. 29.8% (31 out of 104) of the respondents decomposed an application into microservices by applying business capability. 27.9% (29 out of 104) of the respondents used the DDD strategy for this purpose. For example, one respondent answered that “we decompose an application into microservices by using the Banking Industry Architecture Network (BIAN)³ domain model, i.e., the DDD strategy” **Architect (R14)**.

On the other hand, all of the interviewees (IQ2.1.1 in Table 30) confirmed that the results of the survey reflect the actual use of application decomposition strategies (e.g., a combination of

DDD and business capability) for microservices systems in the industry. Five out of six interviewees also confirmed that they are not familiar with any other approach that can be used for application decomposition. Four interviewees (i.e., P2, P3, P4, and P6) mentioned several reasons for using a combination of DDD and business capability, such as (i) a single strategy cannot meet the application decomposition goal in a good way, (ii) it supports the creative and iterative collaboration process between system stakeholders (e.g., technical expert, domain expert, and customers) to achieve the decomposition goals, (iii) it is suitable when implementing microservices in an agile or DevOps process, (iv) it provides the controlled flexibility for future changes as most of the changes remain around the particular domain, (v) it reduces the misunderstanding because requirements are gathered from a specific domain perspective, and (vi) it provides better team communication because all of the stakeholders use a shared set of the terminologies. Following is one representative quotation about the application decomposition strategies for microservices systems.

“Combination of DDD and business capability, helps to build microservices with clear responsibilities and thus increase the maintainability and evolvability of microservices systems” **Architect and Application developer (P1)**.

Fig. 9 shows the results of SQ13, in which the practitioners were asked to select the strategies to design and develop microservices systems. One response to SQ13 is invalid as the participant answered that “I have no idea”, therefore, the number of valid responses is 105. From Fig. 9, it shows that 46.7% (49 out of 105) of the participants only modeled the high-level design (the major components) and 30.5% (32 out of 105) of the participants modeled the detail design before writing the code. In the context of this finding, our interview demonstrates

³ <https://bian.org/>.

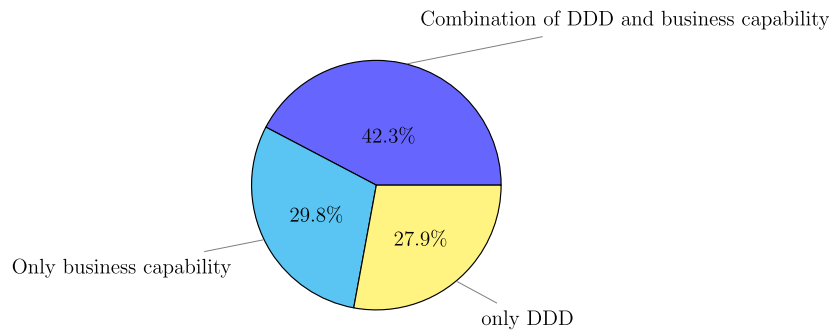


Fig. 8. Strategies of application decomposition into microservices.

that practitioners have stated several reasons for using high-level design for microservices systems. For instance, four interviewees (i.e., P2, P4, P5, and P6) mentioned that high-level design is enough to (i) visualize the identified microservices and their major functionalities, (ii) show communication and orchestration style between microservices, (iii) indicate scalability points for each microservice, (vi) identify vulnerable and insecure points in microservices systems, and (vii) present the boundaries of microservices. Another interviewee (i.e., P3) mentioned that the reasons for only using high-level design are (i) time pressure, which does not allow teams to spend time on low-level design, and (ii) having an experienced development team who does not need low-level design. Following is one representative quotation about the use of only high-level design for microservices systems.

“We always make the high-level design for the whole system (end to end) because it is enough to visualize the identified microservices and their fundamental functionalities, communication or orchestration style between microservices, and the boundaries of microservices. **Architect and Application developer (P1).**”

Moreover, a small percentage (6 out of 105, 5.7%) of the practitioners applied the Model-Driven Development (MDD) approach for generating the code of microservices systems. Surprisingly, 17.1% (18 out of 105) of the participants reported that they wrote the code for microservices systems without having either a high-level or detailed design.

We summarize some of the prominent risks from interviewees' answers that many organizations may face the situation that there is no design or detailed design for microservices systems. The potential risks are (i) teams cannot correctly estimate the amount of work (e.g., person-hour) and time required to complete projects, (ii) implementing poor security standards and solutions, (iii) having no clear or shared vision for teams, (iv) flawed integration plan for microservices, (v) understandability issues for novice developers, (vi) system migration and evolution issues, and (vii) poor communication among the system stakeholders. Two interviewees (i.e., P1, P2) explicitly mentioned that

“Without a design of microservices systems, a system could become a nightmare for developers. Several kinds of risks can affect the microservices system as a whole, including poor work and time management, no clear and common vision about security standards, integration of microservices, and understandability of novice developers” **Architect and Application developer (P3).**

Key Findings of RQ1

Finding 4. Most of the practitioners used a combination of Domain-Driven Design (DDD) and business capabilities to break a monolith application into microservices.

Finding 5. Most of the practitioners only created a high-level design for microservices systems before implementing them.

4.2.2. MSA architecting activities and design description

Five questions (SQ14–SQ18) were asked to understand to what extent the participants agree or disagree to use architecting activities (i.e., Architectural Analysis (AA), Architectural Synthesis (AS), Architectural Evaluation (AE), Architectural Implementation (AI), Architectural Maintenance and Evolution (AME)) in MSA projects. Fig. 10 shows that 53.7% (54+3 out of 106) of the participants either “agree” or “strongly agree” with the use of AE (i.e., evaluating candidate architecture solutions when designing microservices systems to ensure design decisions are right). 52.8% (49+7 out of 106) of the participants “agree” or “strongly agree” with the use of AI (i.e., refining coarse-grained architecture design of microservices systems into fine-grained detailed design with which developers can write code for microservices systems).

Concerning to frequent use of AE activity, four interviewees (i.e., P1, P3, P5) mentioned that they use AE activity to (i) evaluate the functional and quality concerns (e.g., scalability, availability, performance, deployability, and security), (ii) assess the architectural decisions along with their strengths and weaknesses, (iii) get feedback about missing or additional architecturally significant requirements, and (iv) evaluate the business goals in the context of the designed architecture and current development processes. Three interviewees (i.e., P2, P4, P6) stated that they are unfamiliar with these architecting activities and their organizations have dedicated architects or teams to design microservices systems who are responsible for these activities. In the following, we provide one interviewee's response.

“We use AE to determine whether the microservices system's architecture is necessary, beneficial, and useful? Whether the current architecture or design requires continuous deployment? Whether the system requires greater availability and fault tolerance? Whether the system requires scalability? Whether the architecture of the system has or adds complexity to the existing system?” **Application developer (P5).**

Our results also indicate that the survey respondents pay more attention to architecting activities close to the solution space (i.e., AE, AI, AME) instead of the problem space (i.e., AA, AS). About this finding, four out of 6 interviewees (i.e., P1, P2, P3, P5) stated that they mainly focus on the solution space because

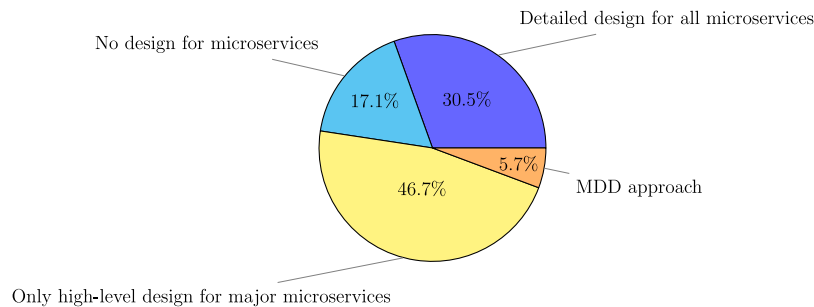


Fig. 9. Strategies from design to implementation of microservices systems.

their primary goal is to meet the customer requirements by providing them with working software. For instance, one participant explained his opinion in the following comment with an example.

“It is because, in the case of the solution space, scenarios are more comfortable to evaluate. For example, if we have a Zookeeper cluster located on two data centers. Losing two nodes within the data centers is an easy scenario to simulate when we have different architectures in the solution space. If we only focus on problem space and design an architecture resilient to the loss of one data center, it is more complicated. In the solution space, we have a set of proven reference architectures, and we use those and enrich them”

Architect and Application developer (P3).

It is interesting to note that a significant number of respondents took a “neutral” position about several MSA architecting activities, for example, AS (50 out of 106, 47.2%), Architectural Analysis (43 out of 106, 40.6%), and AE (34 out of 106, 32.1%). A considerable number of participants “disagree” or “strongly disagree” on the usage of MSA architecting activities. For example, 21.7% (19+4 out of 106) of the participants answered that they do not use AME, and 20.8% (20+2 out of 106) of them responded that they do not use AI.

Key Findings of RQ1

Finding 6. Concerning MSA architecting activities, over half of the participants agreed to use Architecture Evaluation and Implementation when designing microservices systems.

We also investigated about the MSA description methods in this survey. The results show that practitioners prefer to use informal Boxes and Lines for representing MSA. Fig. 11 shows that almost 51% (29+25 out of 106) of the participants answered that they “very often” or “often” used Boxes and Lines for describing MSA. UML is reported as the second most used MSA description method, as 41.5% (31+13 out of 106) of the participants answered that they used UML frequently (often or very often) for describing MSA. Regarding the frequent use of informal and semi-formal description methods, we summarize the responses of all the interviewees: (i) interactively visualize or describe several views of the system for stakeholders, (ii) provide visual representation and flexibility to changes, (iii) are low cost and easy to learn, (iv) are quick to implement and maintain the architecture documents, (v) are not limited to formal notations, (vi) effectively represent the communication between different parts of the system, (vii) are easier to interpret, and (viii) have interactive tools support. One interview participant stated that

“We are not feeling the need for other approaches besides boxes and lines (informal approaches). Informal approaches are simpler to implement and effective to describe microservices architecture” **Architect and Application developer (P3).**

A vast majority of the participants (89 out of 106, 84.7%) responded that they never (41 out of 106, 38.7%), sometimes (25 out of 106, 23.6%), or rarely (23 out of 106, 21.6%) used ADLs for describing MSA. DSLs are also not much popular in the industry, as only 29.2% (8+23 out of 106) of the participants reported that they used (often or very often) DSLs for this purpose. According to interview participants, there are several reasons behind the limited use of ADLs and DSLs, including (i) more effort is required to learn ADLs and DSLs within a tight project delivery schedule, (ii) ADLs and DSLs cannot achieve all the architecting goals (e.g., visualization of architectural views), (iii) ADLs and DSLs lack of generalizability (most ADLs or DSLs are limited for a specific purpose), (iv) ADLs and DSLs have insufficient tooling support, (v) ADLs and DSLs cannot describe the complexity of microservices systems effectively, and (vi) ADLs and DSLs are not delivering benefits to working software. In the following, we provide one example comment from an interviewee about the use of ADLs and DSLs.

“I described architecture for hundreds of systems during my more than fifteen years of work experience from small to very large-scale companies (both monolithic and microservices). However, we did not use ADLs and DSLs either for monolithic or microservices systems because of their low business values and limited visualization to architectural views for the key system stakeholders” **Architect and Application developer (P1).**

Overall, we identified that the use of formal approaches (e.g., ADLs and DSL) for describing MSA is not widespread. In contrast, practitioners prefer to use informal (e.g., Boxes and Lines) or semi-formal (e.g., UML) approaches for describing MSA.

About the diagrams used to represent MSA, Fig. 12 shows the frequency of various types of diagrams used by the respondents to represent the design and architecture of microservices systems (SQ20). Flowchart (56 out of 106, 52.8%) followed by Use case diagram (49 out of 106, 46.2%), Data flow diagram (46 out of 106, 43.3%), and Activity diagram (46 out of 106, 43.3%) are the widely used diagrams to represent MSA in the industry. Other popular diagrams are Sequence diagram (44 out of 106, 41.5%), Class diagram (30 out of 106, 28.3%), and Deployment diagram (27 out of 106, 25.4%). Interestingly, except Microservices circuits,⁴ all the diagrams above have been used for representing the scenario, process, logical, implementation, and physical view of monolithic architecture.

⁴ <https://jolielang.gitbook.io/docs/architectural-composition/aggregation>.

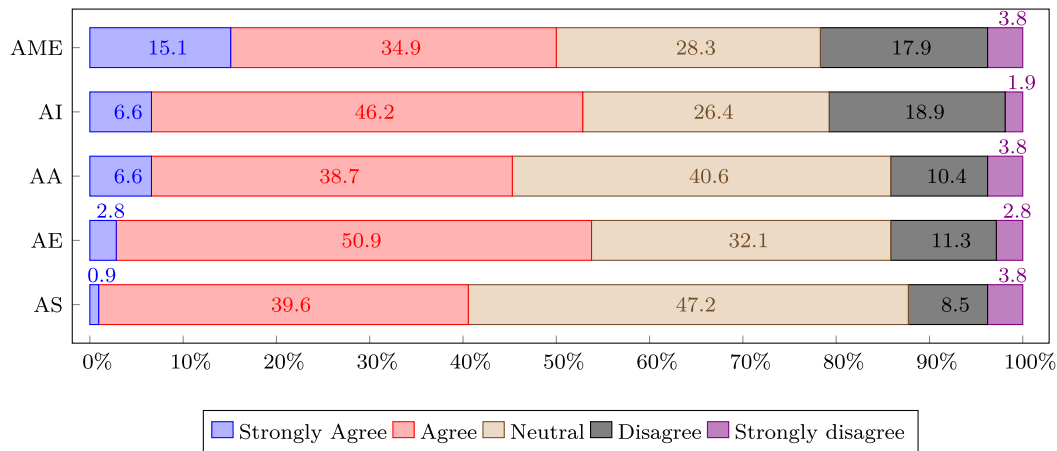


Fig. 10. MSA architecting activities used in industry (AA-Architectural Analysis, AS-Architectural Synthesis, AE-Architectural Evaluation, AI-Architectural Implementation, AME-Architectural Maintenance and Evolution).

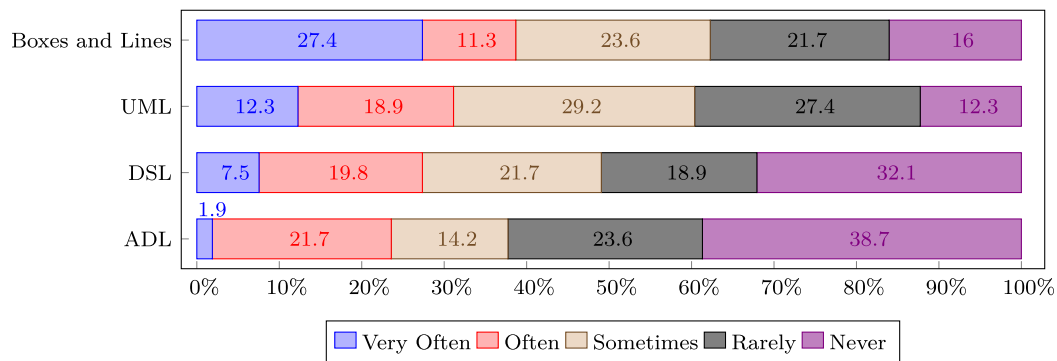


Fig. 11. Methods used for describing MSA in industry (UML-Unified Modeling Language, DSL-Domain Specific Languages, ADL-Architectural Description Language)

Key Findings of RQ1

Finding 7. The use of formal approaches (e.g., ADLs, DSLs) is not widespread in the industry to describe MSA. In contrast, practitioners prefer to use informal (e.g., Boxes and Lines) and semi-formal (e.g., UML) approaches for describing MSA.

Finding 8. Responses of the participants about the diagrams (e.g., flowchart, data flow, activity, sequence) for representing the design and architecture of microservices systems indicate that practitioners prefer to model the process view of MSA.

Another important aspect that we explored through this survey is MSA design components that can be used to describe MSA. We provided a list of potential architectural elements (e.g., domain, service, service interface, process, and storage) to the participants (Hoogendoorn, 2020), and asked them to select which of these elements could be considered as MSA design components. Fig. 13 shows that 45.2% (48 out of 106) of the respondents selected services (e.g., service gateway, service client) as MSA design components whereas 35.2% (37 out of 106) of the participants said that they considered all the architectural elements in the list as MSA design components.

4.2.3. Quality attributes and design patterns

We selected 15 QAs that are popular in the MSA and SOA context (Márquez and Astudillo, 2018; O'Brien et al., 2007; Li et al., 2020). For each QA, the participants were asked to rank their importance for microservices systems using a Likert-scale question rated as “very important”, “important”, “somewhat important”, “not important”, and “not sure”. Table 6 shows that 55.7% (59 out of 106) of the respondents ranked Security as “very important”, followed by Availability (58 out of 106, 54.7%), Performance (56 out of 106, 52.8%), and Scalability (51 out of 106, 48.1%). About these results, we asked the interview participants with IQ2.5.1 “what could be the reasons that practitioners consider security, availability, performance, and scalability the most important quality attributes for microservices systems”. All the interviewees mentioned that they consider Security an important QA because (i) microservices systems are distributed in nature, (ii) microservices systems are complex and have security vulnerabilities that an intruder can exploit. Two of the interviewees (P1, P4) also stated that “Security is equally crucial for microservices or any other type of distributed applications”. Following is one representative quotation about the importance of Security in the context of microservices systems.

“We have to keep up with the various security patches. However, in my understanding, providing security to microservices systems is relatively easy than monolithic applications because normally, we need to secure API gateway from intruders. Also, the isolated nature of microservices helps to limit the access of intruders to only vulnerable services” **Architect and Application developer (P1).**

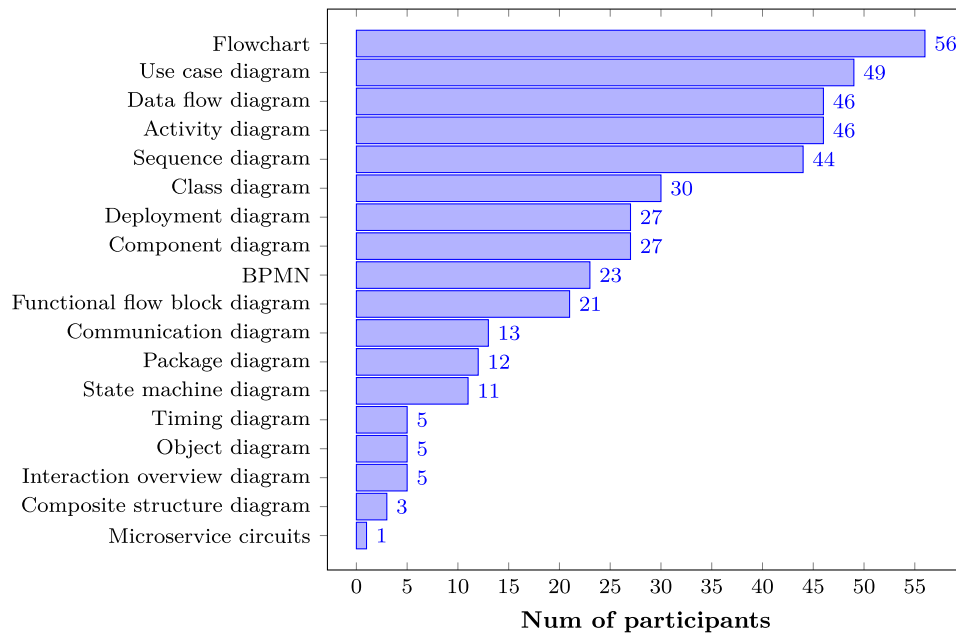


Fig. 12. Diagrams for representing the design and architecture of microservices systems.

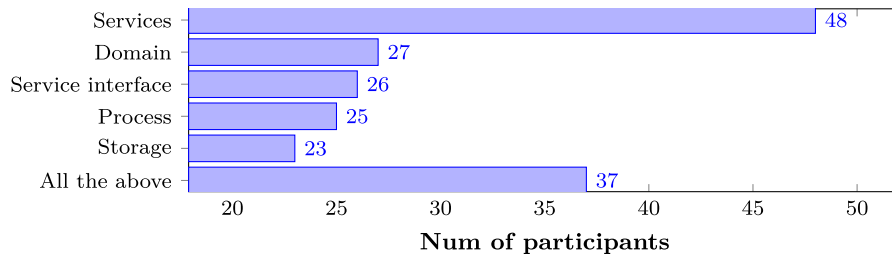


Fig. 13. Design components of microservices systems.

About availability, three interviewees (i.e., P1, P2, P3) stated that end-users usually do not compromise on the non-availability of the systems. One interviewee (i.e., P1) mentioned that end-users could compromise on the system performance to some extent in the peak time of system usage. However, it is not acceptable for the end-users that the system goes down for some time. Moreover, DDD plays a critical role in the availability of microservices systems because applications are decomposed into microservices according to subdomains of the system. If any problem occurs, it remains limited to specific microservice of subdomain and does not propagate in the whole system. We also investigated the reasons for considering Performance as an important QA during the design of microservices systems, and four interviewees (i.e., P1, P2, P3, P5) mentioned that (i) performance should not be degraded due to frequent communication calls and (ii) effective load balancing techniques can be introduced (e.g., creating multiple instances of one microservice and using a load balancer to improve the performance). Regarding Scalability, three interviewees (i.e., P2, P4, P6) stated that it is because scalability is one of the key features of microservices systems that provides a simplified way to scale a particular microservice instance by having a minimal effect on the availability of the system to end-users. Following is one representative quotation about the importance of Availability in the context of microservices systems.

“To achieve reliability for microservices systems, practitioners want to make their systems available 24/7. It is also essential when

the system has several competitors or millions of users” **Architect and Application developer (P5).**

On the other hand, only 26.4% (28 out of 106) and 23.5% (25 out of 106) of the respondents ranked Resilience and Portability as “very important” QAs, respectively. We also found several QAs that were ranked as “somewhat important” by the participants. Such QAs are Reusability (25 out of 106, 23.6%), Compatibility (24 out of 106, 22.6%), and Monitorability (23 out of 106, 21.7%). Overall, the results show that QAs are an essential architectural concern to design microservices systems.

Key Findings of RQ1

Finding 9. Over 80% of the participants indicated that security, availability, performance, and scalability are very important QAs for the design of microservices systems.

The participants were also asked to indicate how often they use patterns when designing microservices systems. The majority of MSA design patterns are decently (i.e., very often, often, sometimes) used in practice, and only some of them are not used or used less often (e.g., rarely, never) to design microservices systems. The results show that the following four MSA design patterns are used most often: API gateway (26.4% very often, 42.5% often, 11.3% sometimes), Backend for frontend (17.9% very often, 35.8% often, 15.1% sometimes), Access token (17.9% very often, 34.9% often, 16.0% sometimes), and Database per service

Table 6

Importance of QAs (in %) for designing microservices systems (VI-Very Important, I-Important, SI-Somewhat Important, NI-Not Important, NS-Not Sure).

ID	Quality attributes	VI	I	SI	NI	NS
QA1	Security	55.7	27.4	9.4	3.8	3.8
QA2	Availability	54.7	29.2	14.2	1.9	0.0
QA3	Performance	52.8	33.0	9.4	2.8	1.9
QA4	Scalability	48.1	38.7	6.6	3.8	2.8
QA5	Reliability	40.6	38.7	12.3	2.8	5.7
QA6	Maintainability	35.8	42.5	20.8	0.0	0.9
QA7	Usability	35.8	40.6	14.2	2.8	6.6
QA8	Compatibility	34.0	34.0	22.6	1.9	7.5
QA9	Testability	33.0	41.5	19.8	0.9	4.7
QA10	Monitorability	32.1	38.7	21.7	1.9	5.7
QA11	Functional suitability	30.2	42.5	15.1	4.7	7.5
QA12	Reusability	30.2	35.8	23.6	5.7	4.7
QA13	Resilience	26.4	30.2	22.6	5.7	15.1
QA14	Portability	23.6	35.8	26.4	7.5	6.6
QA15	Interoperability	21.7	42.5	17.0	5.7	13.2

(12.3% very often, 36.8% often, 21.7% sometimes). These four patterns are used to manage external APIs, application security, and databases for microservices systems. Similarly, we identified some other MSA design patterns that are frequently used for data management, deployment, and communication of microservices. The examples of such patterns are API composition (14.2% very often, 33.0% often, 35.8% sometimes), Service instance per container (16.6% very often, 30.3% often, 24.5% sometimes), and Messaging (14.2% very often, 32.1% often, 24.5% sometimes). Overall, the combined usage (e.g., very often, often, Sometimes) reported by the participants for each MSA design pattern is more than 50.0%.

We also asked IQ2.6.1 to the interviewees about the leading MSA design patterns used in industry. Four out of six interviewees (i.e., P1, P2, P4, P5) provided the following reasons for the frequent use of API gateway and Backend for frontend patterns: (i) API gateway and Backend for frontend are the key patterns for request routing, API composition, and protocol translation for microservices systems, (ii) these patterns have naturally bound with microservices architecture, (iii) API gateway offers a single entry-point for a specific group of microservices, (iv) Backend for frontend is a variant of the API gateway pattern that can be used to separate the frontend from the backend of the applications, and (v) these two patterns enhance the loose coupling property of MSA. About the frequent use of security pattern Access token, four interviewees (P1, P2, P4, P6) mentioned several reasons, such as (i) token-based authentication allowing an application to access an API, (ii) one-time login, (iii) no session is persevered on the server side (stateless), (iv) servers do not need to track and keep the information of users, (v) the security token is short-lived and can be revoked easily, (vi) the security token helps to save the server memory consumption, and (vii) if the service goes down for some reason, the security token is still valid because it is also used on the client side. Regarding the frequent use of Database per service pattern, six interviewees (P1, P2, P3, P4, P5, P6) provided four reasons that the Database per service pattern (i) helps to design and develop the loosely coupled applications, (ii) provides support for selecting databases according to the requirement of each microservice, (iii) helps to control the changes or failure by limiting them to a specific microservice or database, and (iv) provides freedom of using specific databases according to the team's skills. In the following, we provide one representative quotation about the use of API gateway pattern for microservices systems.

“In my understanding, the reasons behind using these patterns are rooted in microservices systems. For instance, instead of the API gateway pattern, we did not use any other communication pattern,

and it is almost considered a ditto standard for communication in the industry” **Application developer (P4)**.

On the other hand, some MSA design patterns, such as Sagas (14.2% rarely, 34.0% never), Event sourcing (21.7% rarely, 23.6% never), and Self-registration (20.8% rarely, 22.6% never), are rarely or never used. More details about the use of MSA design patterns can be found in Table 7.

The participants also mentioned several other patterns via the “Other” field in SQ23. Such patterns are Model-view-controller (5 responses), Two-phase commit⁵ (3 responses), Facade (1 response), Singleton (1 response), Repository (1 response), API key⁶ (1 response), Rate limit⁷ (1 response), Service decorator⁸ (1 response), Session memory (1 response), and Tolerant reader (1 response).

Key Findings of RQ1

Finding 10. API gateway and Backend for frontend as communication patterns, Database per service as data management pattern, and Access token as security pattern are reported as the most often used MSA design patterns in practice.

4.2.4. Skills

We asked the participants about the skills (including hard and soft skills) required to design and implement MSA (SQ27). We identified these skills from the literature, e.g., Xia et al. (2019), Jedrzejewski (2020) and Li et al. (2015). 72.6% (77 out of 106) of the participants believed that practitioners should have the skills to break down a complex coding task into smaller tasks, and 56.6% (60 out of 106) of the participants mentioned that practitioners should have skills to implement the functionality of a microservice by using suitable MSA design patterns. On the other hand, only 30.1% (32 out of 106) of the participants answered that practitioners should know container technologies (see Fig. 14 for more details).

4.2.5. Challenges and solutions

Challenges: The participants were asked about architectural challenges that they may encounter during the design of microservices systems (SQ25). We provided a list of challenges identified in a recent literature review (Waseem et al., 2020b). The results show that DC1: *clearly defining the boundaries of microservices* is the most common challenge in MSA design (74 out of 106, 69.8%). The other two top-ranked challenges in MSA design are DC2: *addressing security concerns* (49 out of 106, 46.2%) and DC3: *managing microservices complexity at the design level* (37 out of 106, 34.9%).

Concerning these design challenges, we asked IQ2.7.1 to the interviewees about their reasons. All the interviewees agreed that these are valid challenges. Regarding the reasons for the challenge DC1, four out of six interviewees (i.e., P1, P2, P5, P6) mentioned that DC1 occurs due to (i) failure in deciding about the responsibilities or functionalities of each microservice, (ii) failure to locate the overlapping and shared functionalities between microservices, (iii) lack of business knowledge when creating a new application, and (iv) no iterative approaches (e.g., sprint meetings) being employed to refine the design of microservices

⁵ <https://tinyurl.com/y9r3sl7q>.

⁶ <https://tinyurl.com/y2sm2jqz>.

⁷ <https://tinyurl.com/yxroch9s>.

⁸ <https://tinyurl.com/y5rphfsx>.

Table 7
MSA design patterns (in %) used in the industry (VO-Very Often, O-Often, S-Sometimes, R-Rarely, N-Never).

ID	MSA patterns	VO	O	S	R	N
MSP1	API gateway	26.4	42.5	11.3	10.4	9.4
MSP2	Backend for frontend	17.9	35.8	15.1	17.9	13.2
MSP3	Access token	17.9	34.9	16.0	17.0	14.2
MSP4	Service instance per container	16.0	30.2	24.5	15.1	14.2
MSP5	Service registry	16.0	29.2	23.6	12.3	18.9
MSP6	API composition	14.2	33.0	35.8	10.4	6.6
MSP7	Messaging	14.2	32.1	24.5	17.9	11.3
MSP8	Remote procedure invocation	13.2	32.1	17.0	17.9	19.8
MSP9	Database per service	12.3	36.8	21.7	17.9	11.3
MSP10	Circuit breaker	11.3	30.2	18.9	19.8	19.8
MSP11	Client-side discovery	10.4	34.9	19.8	19.8	15.1
MSP12	Server-side discovery	10.4	24.5	28.3	17.9	18.9
MSP13	Serverless deployment	9.4	25.5	18.9	17.9	28.3
MSP14	Service instance per host	8.5	29.2	24.5	18.9	18.9
MSP15	Self-registration	8.5	25.5	22.6	20.8	22.6
MSP16	Polling publisher	8.5	17.9	29.2	16.0	28.3
MSP17	Service instance per VM	7.5	30.2	18.9	17.0	26.4
MSP18	Application events	7.5	24.5	33.0	18.9	16.0
MSP19	Sagas	7.5	24.5	19.8	14.2	34.0
MSP20	Service deployment platform	6.6	32.1	23.6	17.0	20.8
MSP21	Multiple service instance per host	6.6	31.1	22.6	24.5	15.1
MSP22	Shared database	6.6	31.1	27.4	19.8	15.1
MSP23	Event sourcing	5.7	20.8	28.3	21.7	23.6
MSP24	CQRS	5.7	17.0	32.1	20.8	24.5
MSP25	Transaction log tailing	4.7	30.2	19.8	24.5	20.8
MSP26	Domain-specific protocol	3.8	33.0	18.9	23.6	20.8

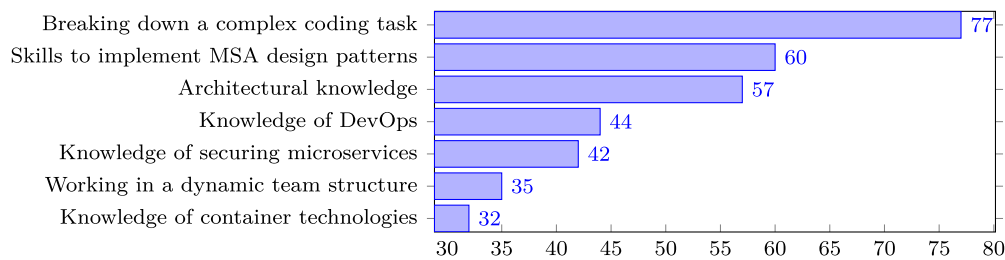


Fig. 14. Skills required to design and implement MSA.

systems. About the reasons for the challenge DC2, three out of six interviewees (i.e., P1, P3, P4) provided the following reasons: (i) insecure communication between microservices, (ii) vulnerable security standards and tools, (iii) insecure cloud infrastructure for deploying and scaling microservices, and (iv) poor design and introducing insecure code for microservices systems. About the reasons for the challenge DC3, four out of six interviewees (i.e., P1, P2, P3, P6) provided the following reasons: (i) hard to identify end-to-end microservices and decide about their communication patterns, (ii) defining microservices for every system feature or activity (too fine-grained), (iii) poorly defining the boundaries of microservices (i.e., DC1), and (iv) too many interactions and dependencies between microservices. Following is one representative quotation about the challenge DC3 during the design of microservices systems.

“We do not have any silver bullet to reduce the complexity. We need to balance the complexity of the system while defining the boundaries of microservices. Complexity cannot be mitigated entirely, but we need to bring it to the level where it does not severely affect the system’s quality (performance, availability, security)” **Architect and Application developer (P3).**

On the other hand, for each design challenge, the participants were further asked to indicate if these challenges impact monitoring and testing of microservices systems. The fourth and fifth columns of Table 8 show how many respondents believed that certain challenge has an impact on monitoring and testing of microservices systems. More than half of the respondents felt that the challenge of DC1: *clearly defining the boundaries of microservices* impacts both monitoring and testing of microservices

systems. Similarly, 61.9% (63 out of 106) and 59.4% (58 out of 106) of the respondents agreed that DC2: *Addressing security concerns* can affect testing and monitoring of microservices systems, respectively. More details can be seen in Table 8.

Solutions for general challenges – from survey results: Through an open-ended survey question (SQ26), we asked the participants what solutions they used to address or mitigate the challenges in MSA design. We received 36 answers, and after evaluating each answer, only 23 responses were found as valid. We excluded those solutions that were incomplete, inconsistent, or do not adequately address the MSA design challenges. We classified the valid answers into two categories (see Table 9) as presented below:

- **Strategy, patterns, and DSLs:** Most of the participants recommended the DDD strategy (10 responses) to address the challenge of DC1: *clearly defining the boundaries of microservices*. This strategy can help in determining the boundaries of microservices through concepts like bounded context (Haselböck et al., 2018). To give an example, one participant explicitly mentioned that “we built a micro services-oriented platform (i.e., Dolittle⁹) that helps in the identification of MSA boundaries by using the DDD strategy” **C-level executive (R5)**. However, one respondent questioned the usefulness of DDD in determining the boundaries of microservices and said that “DDD is not always the right option

⁹ https://dolittle.io/contributing/guidelines/domain_driven_design/.

Table 8

Challenges faced during the design of microservices systems and their impact on monitoring and testing of microservices systems (in %)

ID	MSA Design challenge	Response	Impact of challenges on	
			Monitoring	Testing
DC1	Clearly defining the boundaries of microservices	69.8	59.4	54.7
DC2	Addressing security concerns	46.2	59.4	61.3
DC3	Managing microservices complexity at the design level	34.9	57.5	52.8
DC4	Recovering MSA from existing code	34.9	57.5	50.9
DC5	Reaching scalability	26.4	60.4	51.9
DC6	Finding appropriate modeling abstractions for microservices	23.6	49.1	48.1
DC7	Separating functional and operational concerns	23.6	67.0	49.1
DC8	Addressing microservices communication at the design level	22.6	62.3	50.0
DC9	Addressing data management at the design level	21.7	60.4	50.9
DC10	Aligning team structure with the architecture of MSA-based system	19.8	50.9	60.4

for identification of MSA boundaries as it does not always address problems of scalability and architectural issues for microservices systems. For instance, some boundaries that defined through DDD may not fit with the business model or may not even help to achieve the scalability” **Application developer (R45)**. Vertical decomposition pattern was deemed to achieve appropriate granularity, scalability, and reliability in microservices systems (Hasselbring, 2016; Hasselbring and Steinacker, 2017). One participant stated that “we use vertical decomposition pattern for identification of microservices boundaries” **System analyst (R33)**. One other participant mentioned that “Circuit breaker pattern was considered useful to control the flow of traffic and APIs calls between the services, as well as to protect the system from both transient failures and buggy code” **Architect (R19)**. We also received a few answers in which the participants recommended patterns and DSL as solutions to mitigate the challenges related to the design of MSA, such as microservice API patterns,¹⁰ structural design pattern¹¹ (e.g., re-deployable Jolie Circuit breaker), and DSL Jolie.¹²

- **Adoption of software development methods:** We received one response in which a participant suggested that “developers who have good knowledge of microservices should be involved during the architecting process of the MSA in order to effectively deal with challenges that may arise on design level” **Consultant (R102)**. Furthermore, some participants also proposed to use DevOps (3 responses), Scrum (2 responses), Feature-driven development (1 response), and Team sprints (1 response) to address the challenges (DC1 and DC10 in Table 8) related to the design of microservices systems. However, none of the participants provided detailed descriptions about how software development methods (e.g., DevOps) can address the challenges of MSA design (e.g., security concerns).

Solutions for top three challenges – from interview results: We summarized the solutions from the interview results for the leading three challenges, i.e., DC1: clearly defining the boundaries of microservices, DC2: addressing security concerns, and DC3: managing microservices complexity at the design level.

- **Solutions for clearly defining the boundaries of microservices:** The interviewees proposed several measures to address this challenge: (i) gain a deeper understanding of business requirements, application domains, technologies, business processes, and development processes, (ii) introduce and implement the concept of miniservices (Gupta, 2020) for defining the boundaries of microservices, (iii) identify

Table 9

Proposed solutions to address the MSA design challenges.

Type	Proposed solution	Challenge ID addressed	Count
Strategy, patterns, and DSL	DDD	DC1	15
	Vertical decomposition pattern	DC1, DC5	
	Microservice API patterns (e.g., public API, community API)	DC2, DC5	
	Structural design pattern	DC3	
	Circuit breaker, Jolie (DSL)	DC8	
Software development methods	DevOps, Scrum, FDD, and Team sprints	DC1, D10	8

possible contexts or subdomains of the application by using certain strategies (e.g., DDD), (iv) discover the common and isolated functionality of each subdomain, (v) find the suitable granularity level for microservices, and (vi) ensure the visibility of defined boundaries of microservices.

- **Solutions for addressing security concerns:** Microservices systems can be secured on the network, load balancing, firewall, and application level by simply buying the updated security solutions from cloud providers. For example, AWS provides their own way of securing microservices (e.g., API facade pattern), Kubernetes offers Role-Based Access Control component to limit the access of specific microservice to authorized users. It is suggested that end-users constantly update their firewalls and other security standards to address several security challenges. One of the practitioners (i.e., P2) also recommended employing DevSecOps¹³ and microservices together as a solution to address security challenges, because both technologies have a natural alliance to protect the applications.
- **Solutions for managing microservices complexity at the design level:** There is no silver bullet for addressing the complexity of microservices. However, it can be reduced by taking these measures: (i) do not design or develop microservices for every system’s feature, (ii) define a manageable set of responsibilities for each microservice, (iii) understand the operating environment of microservices (e.g., deployment and monitoring infrastructure), and (iv) clearly define the boundaries of microservices.

¹⁰ <https://microservice-api-patterns.org>.

¹¹ <https://csrc.nist.gov/publications/detail/sp/800-204/draft>.

¹² <https://jolie-lang.org>.

¹³ <https://www.devsecops.org/>.

Key Findings of RQ1

Finding 11. Clearly defining the boundaries of microservices and addressing security concerns are the most reported challenges during the design of microservices systems. Moreover, it has been identified that MSA design challenges have a significant impact on monitoring and testing of microservices systems.

Finding 12. Practitioners indicated that they could better understand and define the boundaries of microservices by using Domain-Driven Design (DDD) strategy.

4.2.6. Differences between the subpopulations of the study for MSA design

We analyzed the answers to Likert scale survey questions in Tables 10, 11, and 12. To better understand practitioners' perspectives on the design of microservices systems, we divided the survey respondents into three groups (i.e., *Experience* ≤ 2 years vs. *Experience* > 2 years, *MSA style* vs. *No MSA style*, *Employees size* ≤ 100 vs. *Employees size* > 100). We calculated the *P*-value to indicate statistically significant difference for each survey question statement and the Effect size to indicate the magnitude of difference between two groups. The dark gray color in table cells shows that the first group (e.g., *Experience* ≤ 2 years) is more likely to agree with the survey question statements. In contrast, the light gray color shows that the second group (e.g., *Experience* > 2 years) is more likely to agree with the survey question statements. In the following, we report the statistical comparison for MSA architecting activities and description approaches, quality attributes, and MSA design patterns.

MSA architecting activities and description approaches (see Table 10): *Experience* ≤ 2 years vs. *Experience* > 2 years: There are two survey question statements (i.e., **Architectural maintenance and evolution** and **Domain specific languages**) with statistically significant differences between *Experience* ≤ 2 years and *Experience* > 2 years groups. The *Experience* > 2 years group is more likely to agree on using Architectural maintenance and evolution activity in MSA projects than the *Experience* ≤ 2 years group. On the other hand, the *Experience* ≤ 2 years group is more likely to use Domain Specific Languages for describing MSA than the *Experience* > 2 years group. *MSA style* vs. *No MSA style*: There are three survey question statements (i.e., **Architectural maintenance and evolution**, **Unified modeling language**, **Domain specific languages**) with statistically significant differences between MSA style and No MSA style groups. The MSA style group is more likely to agree on using Architectural maintenance and evolution activity in MSA projects than the No MSA style group. Moreover, the MSA style group is also more likely to use Unified modeling language and Domain specific languages for describing the architecture of microservices systems than the MSA style group. *Employees size* ≤ 100 vs. *Employees size* > 100 : We did not find any survey question statement with statistically significant difference between *Employees size* ≤ 100 and *Employees size* > 100 groups.

Quality attributes: *Experience* ≤ 2 years vs. *Experience* > 2 years (see Table 11): There are three survey question statements (i.e., **Usability**, **Resilience**, **Interoperability**) with statistically significant differences between *Experience* ≤ 2 years and *Experience* > 2 years groups. The *Experience* > 2 years group is more likely to agree on importance of Usability and Resilience for designing MSA than the *Experience* ≤ 2 years group. On the other hand, the *Experience* ≤ 2 years group is more likely to agree on importance of Interoperability for designing MSA than the *Experience* > 2 years group. *MSA style* vs. *No MSA style*: There are seven survey question statements (i.e., **Security**, **Availability**, **Scalability**, **Reliability**, **Maintainability**, **Compatibility**, **Resilience**) with statistically significant differences between MSA style and No MSA style groups. The MSA style group is more likely to agree on importance of Security, Availability, Scalability, Reliability, Maintainability, and Compatibility for designing MSA than the No MSA style group. On the other hand, the No MSA style group is more likely to agree on importance of Resilience for designing MSA than the MSA style group. *Employees size* ≤ 100 vs. *Employees size* > 100 : We identified two survey question statements (i.e., **Functional suitability**, **Resilience**) with statistically significant differences between *Employees size* ≤ 100 and *Employees size* > 100 groups. The *Employees size* ≤ 100 group is more likely to agree on the importance of the Functional suitability and Resilience quality attributes for designing MSA than the *Employees size* > 100 group. On the other hand, we did not find any survey question statement about the importance of quality attributes with statistically significant difference for the *Employees size* > 100 group.

MSA design patterns: *Experience* ≤ 2 years vs. *Experience* > 2 years (see Table 12): There are four survey question statements (i.e., **Service registry**, **Serverless deployment**, **Polling publisher**, **Shared database**) with statistically significant differences between *Experience* ≤ 2 years and *Experience* > 2 years groups. The *Experience* > 2 years group is more likely to use Service registry, Serverless deployment, Polling publisher, and Shared database patterns for designing MSA than the *Experience* ≤ 2 years group. On the other hand, we did not find any survey question statement about the use of MSA design patterns with statistically significant differences for *Experience* ≤ 2 years group. *MSA style* vs. *No MSA style*: There are sixteen survey question statements (e.g., **Backend for frontend**, **Access token**, **Service registry**) with statistically significant differences between MSA style and No MSA style groups. We identified that the No MSA style group is more likely to use patterns (e.g., Backend for frontend, Access token, Service registry) for designing MSA than the MSA style group. On the other hand, we did not find any survey question statement about using MSA design patterns with statistically significant differences for the MSA style group. *Employees size* ≤ 100 vs. *Employees size* > 100 : We identified three survey question statements (i.e., **Remote procedure invocation**, **Polling publisher**, **Multiple service instance per host**) with statistically significant differences between *Employees size* ≤ 100 and *Employees size* > 100 groups. The *Employees size* > 100 group is more likely to use Polling publisher and Multiple service instance per host patterns for designing MSA than the *Employees size* ≤ 100 group. Moreover, the *Employees size* ≤ 100 group is more likely to use Remote procedure invocation pattern for designing MSA than the *Employees size* > 100 .

4.3. Monitoring of microservices systems (RQ2)

4.3.1. Monitoring metrics

We asked about thirteen monitoring metrics (SQ28) identified from the grey literature, e.g., Fowler (2017), Swersky (2020) and Dave and Degioanni (2020). According to participants' responses, the most popular monitoring metric is resource usage (e.g., CPU and memory usage) (71 out of 106, 66.9%), followed by load balancing (57 out of 106, 53.7%) and availability (56 out of 106, 52.8%).

To understand the reasons behind using monitoring metrics, we asked IQ3.1.1 (see Table 31). All the interviewees were familiar with the monitoring metrics. The interviewees have acknowledged that they mainly use resource usage, load balancing, and availability as the monitoring metrics. They argued that monitoring metrics are used to (i) provide a better customer experience (e.g., avoiding failure and improving performance), (ii) estimate

Table 10

Statistically significant differences on the survey results about MSA architecting activities and description approaches.

Survey question statement	SQ#	Likert Distro.	Exper. ≤ 2 vs. > 2 years		MSA style vs. No MSA style		Empl. ≤ 100 vs. > 100	
		Mean value	P-value	Effect size	P-value	Effect size	P-value	Effect size
MSA architecting activities (Strongly agree (5), Agree (4), Neutral (3), Disagree (2), Strongly disagree (1))								
Architectural analysis	SQ14	3.36	0.55	0.29	0.12	−0.19	0.93	−0.13
Architectural synthesis	SQ15	3.26	0.63	0.31	0.23	−0.03	0.96	0.01
Architectural evaluation	SQ16	3.39	0.59	−0.05	0.23	0.13	0.96	−0.23
Architectural implementation	SQ17	3.38	0.52	−0.16	0.13	−0.32	0.95	0.09
Architectural maintenance and evolution	SQ18	3.40	0.03	−0.11	0.04	0.21	0.89	−0.17
Approaches used for describing the architecture of microservices systems (Very often (5), Often (4), Sometimes (3), Rarely (2), Never (1))								
Boxes and lines	SQ19	3.35	0.33	−1.05	0.12	1.10	0.91	−0.91
Unified modeling language		3.12	0.15	0.13	0.01	0.42	0.90	0.27
Domain specific languages		2.57	0.04	0.70	0.00	0.09	0.91	0.50
Architectural description languages		2.22	0.46	0.58	0.06	−0.59	0.94	0.45

Table 11

Statistically significant differences on the survey results about quality attributes in MSA design.

Survey question statement	SQ#	Likert Distro.	Exper. ≤ 2 vs. > 2 years		MSA style vs. No MSA style		Empl. ≤ 100 vs. > 100	
			Mean value	P-value	Effect size	P-value	Effect size	P-value
Quality attribute when designing MSA-based systems (Very important (5), Important(4), Somewhat important (3), Important (2), Not sure (1))								
Security	SQ22	3.87	0.50	−0.17	0.14	0.10	0.80	0.00
Availability		3.87	0.14	−0.40	0.15	0.50	0.87	−0.12
Performance		3.80	0.41	−0.36	0.14	0.59	0.54	−0.23
Scalability		3.69	0.38	−0.21	0.09	0.25	0.93	−0.07
Reliability		3.52	0.29	−0.45	0.05	0.50	0.92	−0.36
Maintainability		3.51	0.87	−0.03	0.03	0.18	0.78	0.00
Usability		3.43	0.03	−0.05	0.15	0.09	0.78	−0.01
Compatibility		3.42	0.75	0.36	0.01	0.36	0.88	−0.05
Testability		3.42	0.90	−0.20	0.05	0.53	0.97	0.06
Monitorability		3.40	0.32	−0.25	0.06	0.02	0.62	−0.11
Functional suitability		3.36	0.34	0.15	0.12	−0.05	0.04	0.14
Reusability		3.41	0.19	−0.09	0.07	0.36	0.72	−0.07
Resilience	3.20	0.04	−0.04	0.00	−0.07	0.03	0.17	
Portability	3.28	0.18	0.14	0.09	0.47	0.56	0.16	
Interoperability	3.13	0.03	0.35	0.05	0.32	0.59	0.04	

the budget (e.g., payment according to the use of cloud infrastructure, such as Microsoft Azure, Amazon AWS), (iii) assess application behavior on high traffic time (e.g., Christmas), (iv) monitor the overall health of host infrastructure (e.g., memory/CPU utilization), (v) count the number of hosts and pods, (vi) detect problems with multi-threading, (vii) identify degraded states or failure points before the system entirely fails, (viii) improve the performance of microservices, and (ix) monitor the availability state (percentage) of a system for end-users. Following is one representative quotation about the monitoring metrics for microservices systems.

“The use of monitoring metrics allows us to get first-hand information about each microservice's behavior. Monitoring metrics can help to understand, visualize, and isolate failures points. We can also use monitoring metrics to check the efficiency of microservices systems” **Application developer (P5).**

The other popular monitoring metrics reported by the participants are database connections (54 out of 106, 50.9%), threads (47 out of 106, 44.3%), errors and exceptions (46 out of 106, 43.3%), and endpoints success (44 out of 106, 41.5%) (see Fig. 15 for more details).

4.3.2. Monitoring practices

We asked the participants which practices they use to monitor microservices systems (SQ29) about six monitoring practices that have been identified from grey literature (e.g., Carstensen,

2020; Swersky, 2020; Rajput, 2018), (see Fig. 16). The participants mentioned the usage of the following three monitoring practices most often: log management (33.0% very often, 35.8% often, 18.9% sometimes), exception tracking (25.5% very often, 34.9% often, 25.5% sometimes), and health check API (23.6% very often, 34.0% often, 28.3% sometimes).

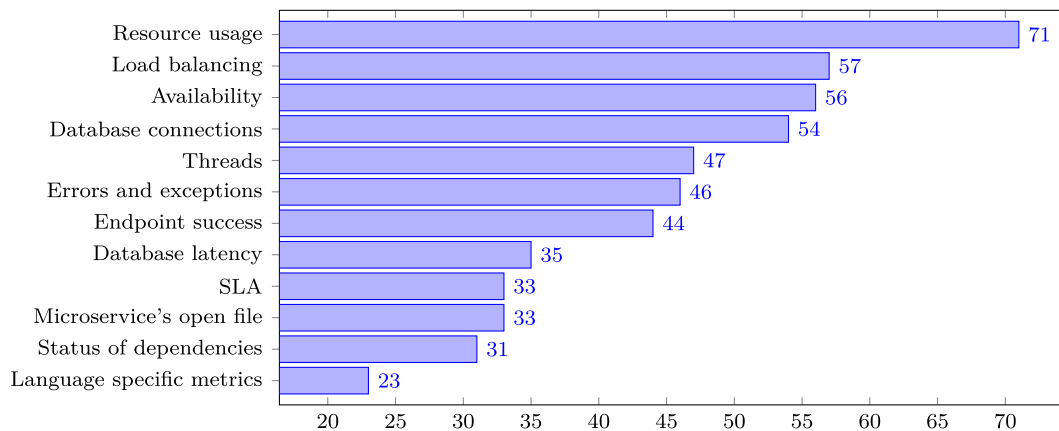
To understand the reasons behind using the top three monitoring practices, we asked IQ3.2.1. We summarized the collected reasons for using the top three monitoring practices according to the responses from five out of six interviewees (i.e., P1, P2, P3, P4, P5): monitoring practices (i) provide the ability to track user activities for preventing, detecting, or minimizing the impact of a data breach, (ii) help to continue analyzing the stability of applications, (iii) help to perform the root cause analysis for specific issues, (iv) help to track the causes of errors and analyze security issues, (v) are used to track and notify the exceptions for developers, and (vi) show the operational status of the service and its interconnection with other services. Following is one representative quotation about the reason of using the top three monitoring practices.

“In my opinion, these three practices complement each other to efficiently manage the quality of each microservice in different operational environments. Besides that, the Health Check API shows the operational status of the service and its interconnection with other microservices.” **Application developer (P5).**

Table 12

Statistically significant differences on the survey results about MSA design patterns.

Survey question statement	SQ#	Likert Distro.	Exper. ≤ 2 vs. > 2 years		MSA style vs. No MSA style		Empl. ≤ 100 vs. > 100	
		Mean value	P-value	Effect size	P-value	Effect size	P-value	Effect size
MSA design patterns when designing MSA-based systems (Very important (5), Important (4), Somewhat important (3), Important (2), Not sure (1))								
API gateway	SQ23	3.67	0.48	−0.37	0.07	0.26	0.94	−0.53
Backend for frontend		3.29	0.23	0.02	0.00	−0.54	0.90	0.05
Access token		3.26	0.31	0.27	0.00	−0.22	0.90	−0.41
Service instance per VM		2.75	0.21	0.03	0.06	−0.60	0.95	−0.12
Service registry		3.13	0.01	−0.15	0.00	−0.37	0.95	−0.60
API composition		3.42	0.45	−0.11	0.06	−0.03	0.93	−0.37
Messaging		3.21	0.32	−0.20	0.06	−0.19	0.90	−0.07
Remote procedure invocation		3.03	0.19	0.12	0.00	−0.31	0.02	0.06
Database per service		3.26	0.40	0.06	0.01	−0.09	0.96	−0.13
Circuit breaker		2.97	0.17	−0.60	0.06	−0.40	0.95	−0.82
Client-side discovery		3.07	0.29	0.12	0.01	−0.12	0.91	−0.21
Server-side discovery		2.92	0.21	−0.42	0.00	−0.21	0.94	−0.20
Serverless deployment		2.70	0.01	−0.16	0.00	−0.52	0.93	0.21
Service instance per host		2.94	0.24	0.01	0.05	0.09	0.89	−0.20
Self-registration		2.79	0.14	0.28	0.00	−0.53	0.94	0.02
Polling publisher		2.64	0.02	−0.29	0.24	−0.68	0.02	−0.38
Application events		2.92	0.30	0.10	0.01	−0.13	0.91	−0.21
Sagas		2.60	0.29	0.07	0.22	−0.81	0.96	−0.22
Service deployment platform		2.88	0.40	0.22	0.01	−0.15	0.95	0.06
Multiple service instance per host		2.90	0.22	−0.06	0.32	−0.17	0.03	−0.43
Shared database		2.93	0.03	−0.24	0.01	−0.13	0.95	0.12
Event sourcing		2.65	0.26	0.18	0.01	−0.32	0.89	0.12
Command query responsibility segregation		2.60	0.25	−0.02	0.20	−0.48	0.92	−0.38
Transaction log tailing	2.75	0.26	0.28	0.01	−0.55	0.95	−0.16	
Service instance per container	3.22	0.20	−0.36	0.00	−0.03	0.95	−0.15	
Domain-specific protocol	2.80	0.33	0.10	0.01	−0.49	0.92	−0.45	

**Fig. 15.** Metrics used for monitoring microservices systems.

On the other hand, distributed tracking (17.0% rarely, 11.3% never) and audit logging (14.2% rarely, 8.5% never) are less used (see Fig. 16 for more details).

4.3.3. Tools

A vast majority of the participants indicated Jira (42 out of 106, 39.6%) and Datadog Kafka Dashboard (41 out of 106, 38.6%) as their monitoring tools, which are followed by Spring Boot Actuator (24 out of 106, 22.6%), Grafana (19 out of 106, 17.9%), and Zipkin (18 out of 106, 16.9%) (see Fig. 17 for details). These tools are mainly used to monitor network performance (e.g., speed,

bandwidth), resource usage (e.g., CPU, memory), and errors and exceptions for microservices systems. We also received 17 short answers in the free text field of survey question SQ30, where the participants provided 11 monitoring tools that they used to monitor microservices systems. These tools are Apache skywalking¹⁴ (3 responses), Azure portal¹⁵ (3 responses), Stackdriver¹⁶

¹⁴ <https://skywalking.apache.org/>.

¹⁵ <https://azure.microsoft.com/en-us/features/azure-portal/>.

¹⁶ <https://cloud.google.com/stackdriver/>.

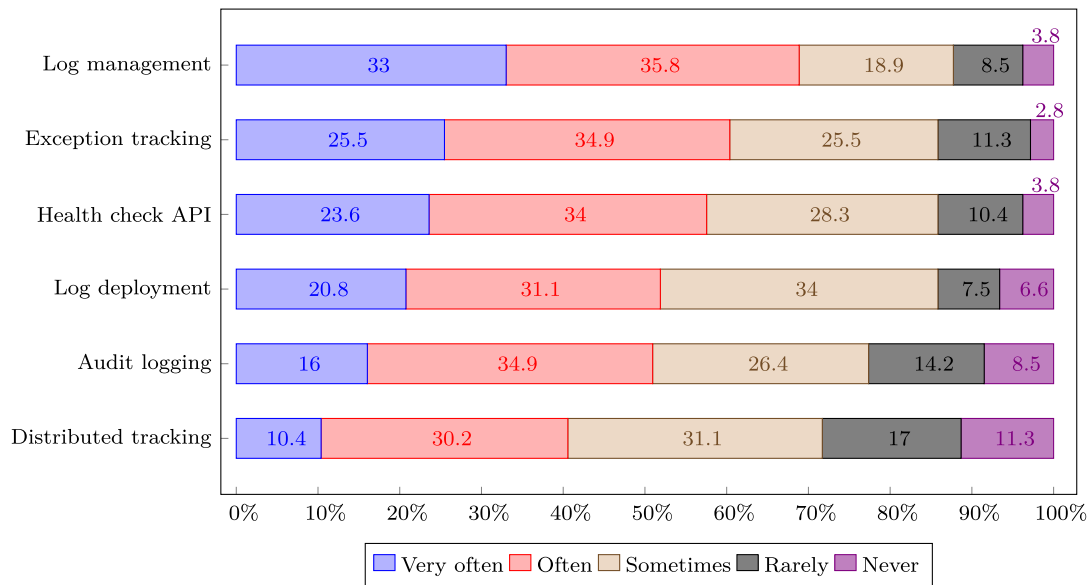


Fig. 16. Practices used for monitoring microservices systems.

(2 responses), Dynatrace¹⁷ (2 responses), [Jaeger,¹⁸ DaoCloud microservice platform¹⁹ (1 response), Heapster²⁰ (1 response), Sensu²¹ (1 response), Red Hat Jboss fuse²² (1 response), KintoHub²³ (1 response), and self-developed RPC tracking tool (1 response).

Besides that, we also asked IQ3.3.1 about the use of Jira as a monitoring tools. Out of the six interviewees, three interviewees stated that Jira could be integrated with Integrated Development Environment (IDE) to monitor microservices systems. Two other interviewees considered issue tracking as part of monitoring of microservices systems and other types of systems, and they said that Jira as an issue tracking tool can be considered as a monitoring tool. One interviewee was not sure about the use of Jira for monitoring purposes, and he mentioned that “we are using Dynatrace for monitoring microservices systems in our company”.

Key Findings of RQ2

Finding 13. Resource usage, load balancing, and availability are the most frequently used monitoring metrics for microservices systems.

Finding 14. Log management, exception tracking, and health check API are the most frequently employed monitoring practices for microservices systems.

4.3.4. Challenges and solutions

Challenges: The respondents were asked to indicate what monitoring challenges they faced (i.e., SQ31). The monitoring challenges listed as answers of SQ31 were collected from the academic and grey literature (Waseem et al., 2020b; Khare, 2020; Heinrich et al., 2017). The third column of Table 13 shows the percentage of the practitioners' responses for each challenge.

Table 13

Challenges faced during the monitoring of microservices systems (in %).

ID	Monitoring challenges	Responses
MC1	Collection of monitoring metrics data and logs from containers	54.7
MC2	Distributed tracing	45.3
MC3	Having many components to monitor (complexity)	41.5
MC4	Performance monitoring	40.6
MC5	Analyzing the collected data	34.9
MC6	Failure zone detection	34.0
MC7	Availability of the monitoring tools	21.7
MC8	Monitoring of application running inside containers	21.7
MC9	Maintaining monitoring infrastructures	17.0

The three main challenges that the participants faced during the monitoring of microservices systems are *collection of monitoring metrics data and logs from containers* (58 out of 106, 54.7%), *distributed tracing* (48 out of 106, 45.2%), and *having many components to monitor (complexity)* (44 out of 106, 41.5%). Other highly ranked challenges are *performance monitoring* (43 out of 106, 40.5%), *analyzing the collected data* (37 out of 106, 34.9%), and *failure zone detection* (36 out of 106, 33.9%).

We asked IQ3.4.1 about the reasons for the top three monitoring challenges. Four out of six interviewees (i.e., P1, P2, P3, P5) provided the following reasons about the monitoring challenges: (i) communication between hundreds of microservices, (ii) run-time monitoring of microservices systems often does not have a standardized infrastructure, (iii) different languages, databases, and frameworks for developing microservices, and (iv) logs and dataflows in different formats. Following is one representative quotation about the reason for the monitoring challenges.

“Indeed, these are significant challenges. However, I think this is due to the complex nature of microservices systems” **Architect and Application developer (P2).**

Solutions for monitoring challenges – from interview and survey results: We received 24 distinct answers as the solutions for addressing monitoring challenges (SQ33). After checking the answers, we found only 11 valid solutions. We excluded those solutions that were incomplete, inconsistent, or did not adequately address the challenges of monitoring microservices systems. The participants recommended ten tools and one in-house developed framework to address the challenges related to monitoring of microservices systems.

¹⁷ <https://www.dynatrace.com/>.

¹⁸ <https://www.jaegertracing.io/> (1 response).

¹⁹ <http://blog.daocloud.io/dmp/>.

²⁰ <https://github.com/kubernetes-retired/heapster>.

²¹ <https://sensu.io/>.

²² <https://www.redhat.com/en/technologies/jboss-middleware/fuse>.

²³ <https://www.kintoHub.com/>.

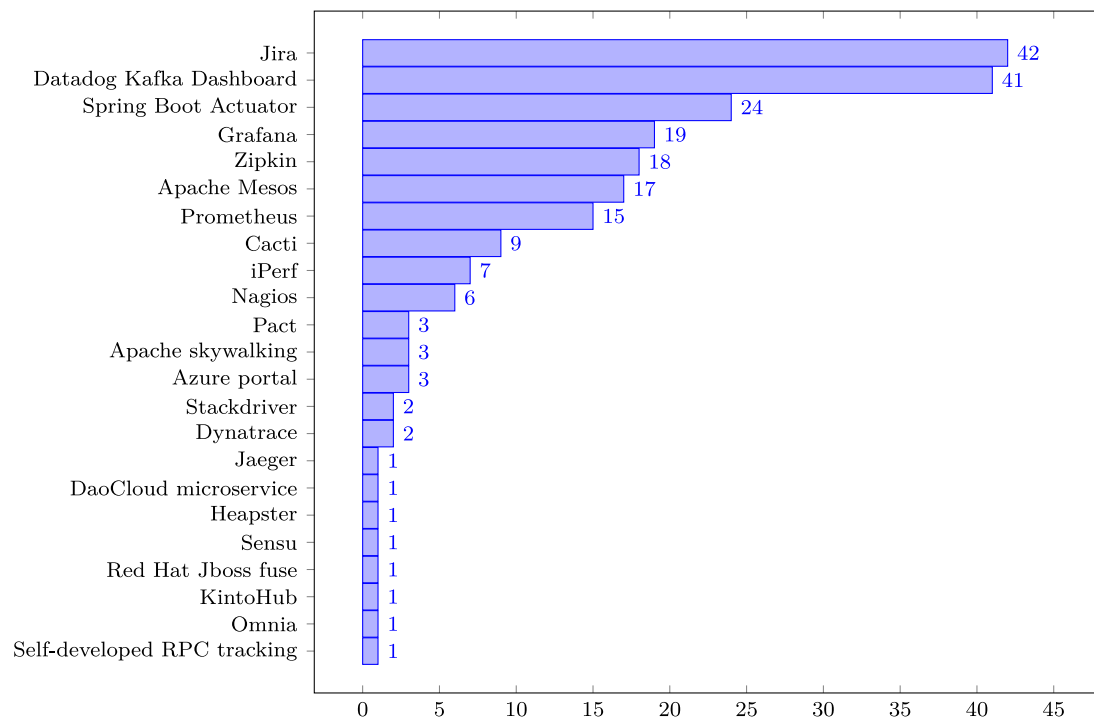


Fig. 17. Tools used for monitoring microservices systems.

Table 14

The severity of the challenges on the monitoring of microservices systems (in %).

ID	Severity scale				
	Catastrophic	Major	Moderate	Minor	Insignificant
MC1	3.8	17.9	48.1	23.6	6.6
MC2	3.8	25.5	38.7	24.5	7.5
MC3	5.7	34.0	33.0	17.0	10.4
MC4	6.6	35.8	32.1	20.8	4.7
MC5	6.6	38.7	38.7	12.3	3.8
MC6	9.4	31.1	36.8	16.0	6.6
MC7	3.8	35.8	35.8	18.9	5.7
MC8	5.7	19.8	43.4	24.5	6.6
MC9	7.5	24.5	44.3	17.0	6.6

Table 15

Proposed solutions to address the monitoring challenges.

Type	Proposed solution	Challenge ID addressed	Count
Monitoring tools and frameworks	ELK (Elasticsearch, Logstash, and Kibana) Stack	MC1, MC4, MC5	11
	Prometheus and Spring Boot Actuator	MC1	
	Grafana	MC1	
	Zipkin	MC2	
	Raygun APM	MC4	
	Health check API tools	MC6	
	In-house framework	MC6	

- Monitoring tools and a self-developed framework:** We received five responses in which the participants named the tools and provided the purposes of using tools. An respondent said that “we use Prometheus and Spring Actuator for collecting, aggregates, and visualize the monitoring data” **Operational staff (R6)**. Similarly, another participant shared that “our team uses ELK (Elasticsearch, Logstash, and Kibana) Stack to address the challenges related to a different kind of monitoring, for instance, application performance and log monitoring” **DevOps engineer (R11)**. In another example, a participant pointed out that “we use health check API tools to detect the health of microservices (e.g., failure zone detection)” **DevOps engineer (R91)**. Some participants only recommended a few tools without explaining the reasons behind using them in monitoring microservices systems. For example, Grafana²⁴ (1 response), Zipkin²⁵ (2 responses), and Raygun APM²⁶ (2 responses). However, we identified

that these tools could be used to address several challenges. For instance, MC1, MC2, MC4, and MC5 (see Table 15). On the other hand, only one participant referred to their in-house developed framework to address the challenges of MC1: collection of monitoring metrics data, MC6: failure zone detection, and MC8: monitoring of application running inside containers. However, this participant did not provide any reference for his developed framework. We also asked IQ3.4.2 about the solutions of monitoring challenges. Three out of six interviewees (i.e., P1, P3, P4) only suggested the names of several tools that they were using in their organizations and these tools were successfully used to address most of the monitoring challenges. The tool names are (i) Dynatrace, (ii) Elastic Stack, (iii) Amazon CloudWatch, (iv) Jaeger, (v) Graphana Dashboard, and (vi) Prometheus.

²⁴ <https://grafana.com/>.

²⁵ <https://zipkin.io/>.

²⁶ <https://raygun.com/platform/apm>.

Key Findings of RQ2

Finding 15. The most prominent challenges of monitoring microservices systems are collection of monitoring metrics data and logs from containers, and distributed tracing.

4.3.5. Differences between the subpopulations of the study for monitoring of microservices systems

We analyzed the answers to Likert scale survey questions related to monitoring microservices in Tables 16 and 17. In the following, we report the statistical comparison for monitoring practices and challenges of microservices systems.

Monitoring practices (see Table 16): *Experience* ≤ 2 years vs. *Experience* > 2 years: There are two survey question statements (i.e., 🛠️ **Health check API**, 🛠️ **Log deployment and changes**) with statistically significant differences between *Experience* ≤ 2 years and *Experience* > 2 years groups. The *Experience* > 2 years group is more likely to use Health check API and Log deployment and changes for monitoring microservices systems than the *Experience* ≤ 2 years group. On the other hand, we did not find any survey question statement about using monitoring practices with statistically significant differences for *Experience* ≤ 2 years group. *MSA style* vs. *No MSA style*: There are two survey question statements (i.e., 🐞 **Audit logging**, 🐞 **Distributed tracking**) with statistically significant differences between *MSA style* and *No MSA style* groups. The *No MSA style* group is more likely to use Audit logging and Distributed tracking for monitoring microservices systems than the *MSA style* group. However, we did not find any survey question statement about using monitoring practices with statistically significant differences for the *MSA style* group. *Employees size* ≤ 100 vs. *Employees size* > 100 : We identified one survey question statement (i.e., 🐞 **Distributed tracking**) with statistically significant difference between *Employees size* ≤ 100 and *Employees size* > 100 groups. The *Employees size* > 100 group is more likely to use Distributed tracking for monitoring microservices systems than the *Employees size* ≤ 100 group. Moreover, we did not find any survey question statement about using monitoring practices with statistically significant differences for the *Employees size* ≤ 100 group.

Monitoring challenges (see Table 17): *Experience* ≤ 2 years vs. *Experience* > 2 years: There are two survey question statements (i.e., 🛠️ **MC3**, 🛠️ **MC8**) with statistically significant differences between *Experience* ≤ 2 years and *Experience* > 2 years groups. The *Experience* ≤ 2 years group is more likely to agree that MC3 (i.e., having many components to monitor) challenge has a severe impact on monitoring of microservices systems than the *Experience* > 2 years group. On the other hand, the *Experience* > 2 years group is more likely to agree that MC8 (i.e., monitoring of application running inside containers) challenge has a critical impact on monitoring of microservices systems than *Experience* ≤ 2 years group. *MSA style* vs. *No MSA style*: There are two survey question statements (i.e., 🐞 **MC4**, 🐞 **MC6**) with statistically significant differences between *MSA style* and *No MSA style* groups. The *MSA style* group is more likely to agree that MC4 (i.e., performance monitoring) and MC6 (i.e., failure zone detection) challenges have severe impact on monitoring of microservices systems than the *No MSA style* group. However, we did not find any survey question statement about monitoring challenges with statistically significant differences for the *No MSA style* group. *Employees size* ≤ 100 vs. *Employees size* > 100 : We identified one survey question statement (i.e., 🐞 **MC5**) with statistically significant difference between *Employees size* ≤ 100 and *Employees size* > 100 groups. The *Employees size*

> 100 group is more likely to agree that MC5 (i.e., analyzing the collected data) challenge has a severe impact on monitoring of microservices systems than the *Employees size* ≤ 100 group. Moreover, we did not find any survey question statement about monitoring challenges with statistically significant difference for the *Employees size* ≤ 100 group.

4.4. Testing of microservices systems (RQ3)

4.4.1. Testing strategies

We asked SQ34 to collect information relevant to the use of testing strategies for microservices systems (Waseem et al., 2020a; Renz et al., 2016; Lumetta, 2019; Clemson, 2020). The top three most frequently used testing strategies are: Unit testing (34.0% very often, 29.2% often, and 10.4% sometimes), E2E testing (26.4% very often, and 29.2% often, 10.4% sometimes), and Integration testing (23.6% very often, 35.8% often, and 6.6% sometimes).

We asked IQ4.1.1 (see Table 32) about the reasons for using the top three testing strategies. All of the interviewees were familiar with these testing strategies. They also acknowledged that they frequently use these testing strategies to test microservices systems. According to the interviewees, these are fundamental testing strategies that can be used for all type of systems (e.g., MSA-based systems, SOA-based systems). In the following, we summarized the collected reasons. Concerning the reasons of using Unit testing, all the interviewees mentioned that they use this strategy because Unit testing (i) allows isolated testing of the design components of microservices systems (e.g., services, domains, service interfaces), (ii) makes sure each microservice working correctly and delivering the required functionalities, (iii) helps to identify the problems with individual microservice in the early stage of system development. About the reasons of using integration testing, four out of six interviewees (i.e., P2, P3, P5, P6) stated that this strategy integrates microservices together to verify that (i) microservices are working collaboratively to achieve the business goals, (ii) communication paths between microservices are working correctly, and (iii) third-party APIs and microservices are working together as expected. About the reasons of using E2E testing, three out of six interviewees (i.e., P2, P5, P6) stated that they typically perform E2E testing for testing the whole microservices systems, and the purposes of using E2E testing are to (i) ensure the correctness of the whole microservices systems and (ii) understand how the customers will use the developed systems. Following is one representative quotation about the reason of using testing strategies for microservices systems.

🗨️ “The reasons for using these testing strategies are the same for microservices and monolithic systems. To the best of my knowledge. There is no specific testing technique that is designed or used in the industry specifically to test microservices systems” **Architect and Application developer (P1)**.

On the other hand, the following three testing strategies are used sparingly: A/B testing (26.4% rarely, 27.4% never), Consumer-driven contract testing (29.2% rarely, 16.0% never), and Service component testing (33.0% rarely, 5.7% never) (see Fig. 18 for details).

4.4.2. Tools

We also asked SQ36 about the tools used for testing microservices systems. A large number of participants responded that they used JUnit (55 out of 106, 51.9%), JMeter (29 out of 106, 27.3%), and Mocha (23 out of 106, 21.7%) (see Fig. 19 for details).

We also asked IQ4.2.1 to understand the reasons for using the top three testing tools. Regarding JUnit and JMeter, all the interviewees agreed that these are fundamental testing tools that






Table 16

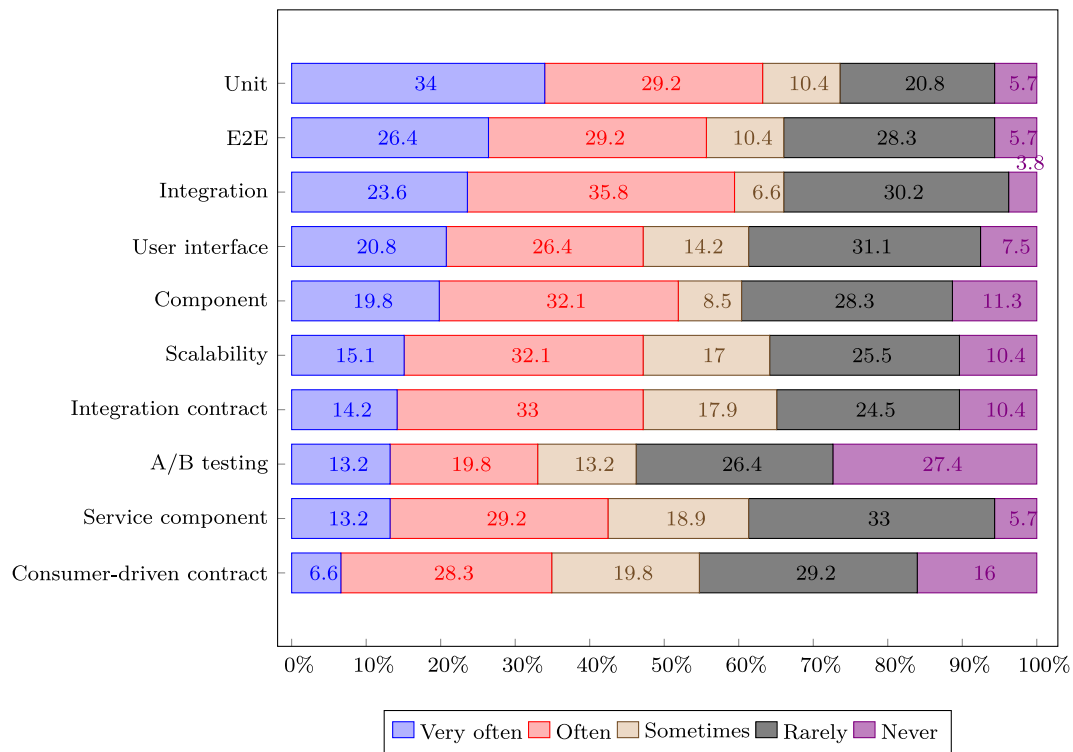
Statistically significant differences on the survey results about monitoring practices.

Survey question statement	SQ#	Likert Distro.	Exper. ≤ 2 vs. > 2 years		MSA style vs. No MSA style		Empl. ≤ 100 vs. > 100	
		Mean value	P-value	Effect size	P-value	Effect size	P-value	Effect size
Practices to monitor microservices systems (Very important (5), Important (4), Somewhat important (3), Important (2), Not sure (1))								
Log management	SQ29	3.89	0.51	−0.42	0.10	0.93	0.94	−0.54
Audit logging		3.42	0.43	−0.42	0.02	−0.09	0.92	−0.54
Distributed tracking		3.13	0.53	−0.47	0.02	−0.18	0.02	−0.38
Exception tracking		3.73	0.60	−0.31	0.07	0.41	0.93	−0.01
Health check API		3.64	0.02	−0.40	0.05	0.24	0.89	0.00
Log deployment and changes		3.54	0.03	−0.17	0.05	0.17	0.94	0.37

Table 17

Statistically significant differences on the survey results about monitoring challenges.

Survey question statement	SQ#	Likert Distro.	Exper. ≤ 2 vs. > 2 years		MSA style vs. No MSA style		Empl. ≤ 100 vs. > 100	
		Mean value	P-value	Effect size	P-value	Effect size	P-value	Effect size
Severity of each challenge related to the monitoring of MSA-based systems (Catastrophic (5), Major (4), Moderate (3), Minor (2), Insignificant (1))								
MC1	SQ32	3.15	0.16	−0.03	0.09	−0.02	0.94	−0.13
MC2		2.94	0.44	−0.21	0.06	0.03	0.94	0.23
MC3		3.09	 0.04	0.21	0.08	0.99	0.94	−0.25
MC4		3.20	0.14	0.04	 0.03	−0.43	0.94	−0.20
MC5		3.33	0.19	−0.09	0.14	−0.06	 0.04	−0.01
MC6		3.24	0.12	−0.12	 0.04	−0.08	0.97	0.10
MC7		2.89	0.19	−0.04	0.12	0.26	0.95	−0.16
MC8		2.94	 0.02	−0.06	0.07	−0.04	0.94	0.04
MC9		3.09	0.15	0.08	0.08	0.41	0.94	0.04

**Fig. 18.** Testing strategies used to test microservices systems.

are required to test Java-based systems (including microservices systems). Moreover, three out of six interviewees (i.e., P2, P3, P5) mentioned that these tools also provide efficiency, simplicity, and agility in testing of microservices systems.

In the “Other” field of this survey question, the participants also mentioned several other tools. We classified these tools into the category of “Others” in Fig. 19, including Nock (2 responses),

Nunit (2 responses), Xunit (2 responses), Moq (2 responses), Jest (2 responses), Fluent assert (1 response), and Pact (1 response). We also received one response in which the participant mentioned that “We manually test microservices systems”.

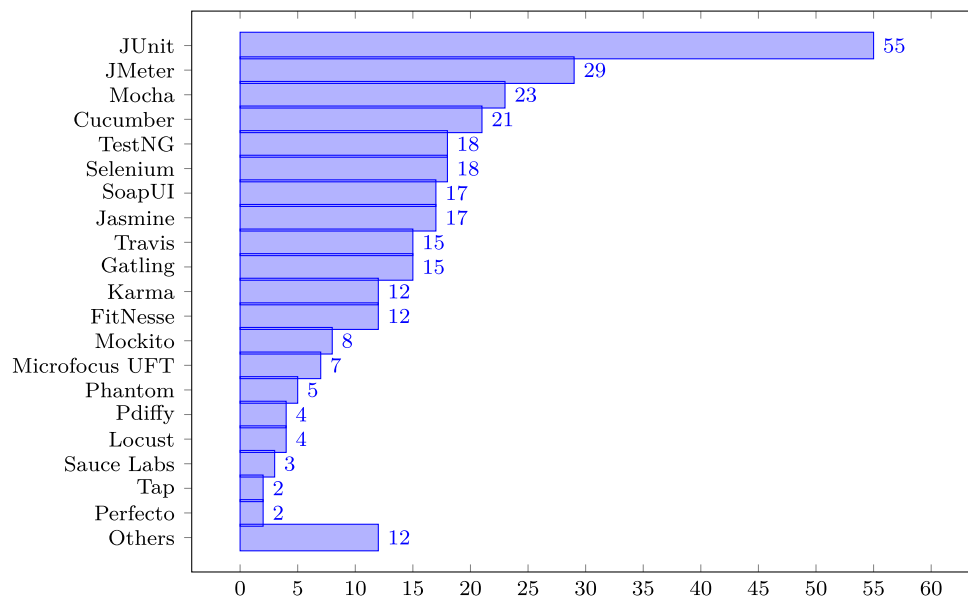


Fig. 19. Tools used to test microservices systems.

Key Findings of RQ3

Finding 16. The most commonly used testing strategies to test microservices systems are unit testing, E2E testing, and integration testing. These testing strategies are also used to test monolith and SOA-based systems.

Finding 17. Our results show that the most of the practitioners used Junit, JMeter, and Mocha tools to test microservices systems.

Table 18

Challenges faced during the testing of microservices systems (in %).

ID	Testing challenges	Responses
TC1	Creating and implementing manual tests	50.0
TC2	Integration testing of microservices	43.4
TC3	Debugging of microservices on container platforms	42.5
TC4	Creating and implementing automated tests	38.7
TC5	Independent microservices	32.1
TC6	Complexity and size of microservices systems	31.1
TC7	Understanding about each testable microservice	30.2
TC8	Performance testing of microservices	30.2
TC9	Many independent teams	20.8
TC10	Polyglot technologies in microservices systems	15.1

4.4.3. Skills

This survey also reports the skills required to test microservices systems properly (SQ38). Seven skills related to testing of microservices systems were collected after reviewing the relevant literature, e.g., Xia et al. (2019), Jędrzejewski (2020) and Tejaswini (2019). According to results of the survey, the top three skills that are required to test microservices systems are *writing good integration test cases* (70 out of 106, 66%), *writing good unit test cases* (63 out of 106, 59.4%), and *having knowledge about multiple databases* (61 out of 106, 57.5%). Other highly-ranked skills are *having analytically and logically thinking* (58 out of 106, 54.7%) and *knowledge about test automation tools* (54 out of 106, 50.9%) (see Fig. 20 for more details).

4.4.4. Challenges and solutions

Challenges: Similar to SQ24 and SQ31, we asked the participants to select the challenges of testing microservices systems from a given list in SQ36. The list of the testing microservices challenges was prepared by reviewing the literature, e.g., Rajput (2018), Waseem et al. (2020a), Lumetta (2019) and Sotomayor et al. (2019). The third column of Table 18 shows percentage of responses for each challenge. Exactly half (53 out of 106, 50.0%) of the participants reported that TC1: *creating and implementing manual tests for microservices is challenging*. Other prominent challenges are TC2: *Integration testing* (46 out of 106, 43.3%), TC3: *Debugging of microservices on container platforms* (45 out of 106, 42.4%), and TC4: *Creating and implementing automated tests* (41 out of 106, 38.6%). Moreover, some of the participants (16 out of 106, 15.0%) felt that TC10: *Polyglot technologies* could pose a challenge for testing of microservices systems.

Furthermore, we investigated the severity of each challenge on testing of microservices systems by asking the participants to rate them through a five-point Likert-scale: “catastrophic”, “major”, “moderate”, “minor”, and “insignificant” (see Table 19). The Likert-scale points “catastrophic” and “major” are used to record the “major long-term impact” and “major short-term impact” of the challenges on testing of microservices systems. The Likert-scale point “moderate” is used to record “significant impact” of the challenges. In contrast, the Likert-scale points “minor” and “insignificant” are used to record the “short-term impact” and “minimal impact” of the challenges. According to the results, all the challenges have “significant” or “major short-term impact” on testing of microservices systems, for instance, more than 56.6% of the participants rated all the testing challenges as “major” or “moderate” on the severity scale, and a few (less than 10%) of the participants reported that all the challenges have a “major long-term impact” on testing of microservices systems (see Table 19).

We asked IQ4.3.1 about the reasons for the top three testing challenges. All the interviewees were agreed that these are valid challenges. Regarding the reasons of the challenge TC1, the interviewees stated that manual testing could become a problem due to the number of microservices and communication between them. Regarding the reasons of the challenge TC2, three interviewees (i.e., P2, P3, P5) mentioned that integration testing is challenging because of (i) multiple connecting points, (ii) limited knowledge of testers about all the microservices, and (iii) manually analyzing logs across multiple microservices. Regarding the reasons of the challenge TC3, two interviewees (i.e., P1,

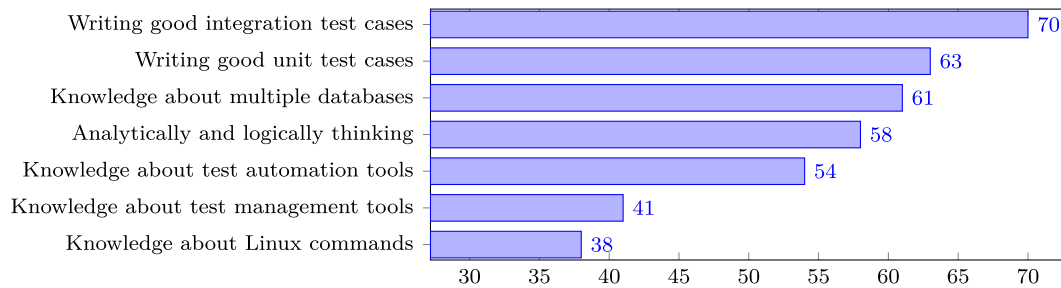


Fig. 20. Skills required to test microservices systems.

Table 19

The severity of the challenges on testing of microservices systems (in %).

ID	Severity scale				
	Catastrophic	Major	Moderate	Minor	Insignificant
TC1	5.7	20.8	35.8	26.4	11.3
TC2	2.8	31.1	39.6	17.9	8.5
TC3	6.6	34.0	32.1	18.9	8.5
TC4	1.9	30.2	37.7	22.6	7.5
TC5	4.7	21.7	37.7	26.4	9.4
TC6	7.5	34.0	36.8	16.0	5.7
TC7	4.7	22.6	42.5	21.7	8.5
TC8	6.6	24.5	46.2	17.9	4.7
TC9	9.4	23.6	37.7	20.8	8.5
TC10	4.7	25.5	43.4	19.8	6.6

P2) pointed out that debugging of microservices on container platforms is challenging due to the lack of automatic debugging tool support. Following is one representative quotation about the reason for the testing challenges.

“Our major challenge is to write effective test cases for the integration testing of microservices systems, which requires extensive knowledge about each microservice and application domain.” **Architect and Application developer (P1).**

On the other hand, many participants reported that the testing challenges have “short-term” or “minimal” impact on testing of microservices systems. For example, TC1: creating and implementing manual tests (26.4% minor, 11.3% insignificant), TC5: independent microservices (26.4% minor, 9.4% insignificant), and TC7: understanding about each testable microservice (21.7% minor, 8.5% insignificant). Overall, we received contradictory results about severity of the challenges on testing of microservices systems.

Solutions for testing challenges – from interview and survey results: We received 25 responses from practitioners for addressing testing challenges, in which only 13 responses are valid after our checking (see Table 20). The participants recommended the following testing strategies, and approaches and guidelines to address the testing challenges.

- **Testing strategies:** Two participants (**DevOps engineer (R69)** and **Software quality engineer (R89)**) suggested that there should be test automation for alleviating the challenge of TC1: creating and implementing manual tests in microservices systems. According to the participants, testing the whole microservices system is tricky and difficult in a manual way. They extended their answers by suggesting automation for testing strategies (e.g., unit, integration, E2E, and consumer-driven contract testing) with the help of DevOps. One **DevOps engineer** also said that “manual testing has several issues with distributed systems like microservices, for example, test case coverage, testing the whole system, domain knowledge about the system under test, and many others. I must suggest enterprises should use test case and test

Table 20

Proposed solutions to address the testing challenges.

Type	Proposed solution	Challenge ID addressed	Count
Testing strategies	Unit, Integration, E2E, and Consumer-driven contract testing	TC1, TC2	13
Approaches and guidelines	Behavior-driven development	TC2, TC4	
	Test cases should have enough coverage rate.	TC7	

management software tool like Testrail²⁷ **DevOps engineer (R69)**”.

- **Approaches and guidelines:** We received three answers in which the participants (**System analyst (R24)**, **Architect (R54)**, and **Software quality engineer (R19)**) stressed upon the development of microservices systems through the Behavior-Driven Development (BDD) approach. According to the participants, the BDD approach can be used to address the challenges like implementing automated tests and integration testing. Concerning to BDD, one **Software quality engineer** said that “we have focused very much on BDD and want to explore this even further. Right now, we are very much doing unit level testing through BDD but want to grow out of it to other testing strategies (e.g., integration, end to end) with DevOps infrastructure (R19)”. About addressing the testing challenge of TC7: understanding about each testable microservice, we received several answers. For example, one other participant suggested that “you must ensure the project has an explicit requirement, required test cases are consistent and have enough coverage rate” **Software quality engineer (R31)**. We asked IQ4.3.2 about the solutions for testing challenges. In the following, we report the summary of the solutions received from the interviewees: (i) start testing from individual microservice with technically sound staff, (ii) each microservice should have its own database, (iii) introduce DevOps or CI/CD for continuous testing of microservices systems, (iv) write effective test cases according to the granularity of each microservice, and (v) thoroughly test microservices systems through unit, integration, component, and E2E testing strategies.











Key Findings of RQ3






Finding 18. Most of the practitioners considered creating and implementing manual tests, integration testing, and debugging of the microservices deployed on container platforms a challenging job.

²⁷ <https://www.gurock.com/testrail/>.

4.4.5. Differences between the subpopulations of the study testing of microservices systems

We analyzed the answers to Likert scale survey questions related to testing of microservices systems in Tables 21 and 22. In the following, we report the statistical comparison for testing strategies and testing challenges of microservices systems.

Testing strategies (see Table 21): *Experience* ≤ 2 years vs. *Experience* > 2 years: There are three survey question statements (i.e.,  **Component testing**,  **User interface testing**,  **A/B testing**) with statistically significant differences between Experience ≤ 2 years and Experience > 2 years groups. The Experience ≤ 2 years group is more likely to use User interface and A/B testing strategies to test microservices systems than the Experience > 2 years group. On the other hand, the Experience > 2 years group is more likely to use Component testing strategy to test microservices systems than the Experience ≤ 2 years group. *MSA style* vs. *No MSA style*: There are six survey question statements (i.e.,  **Integration contract testing**,  **Consumer-driven contract testing**,  **Component testing**,  **Scalability testing**,  **End to End testing**,  **A/B testing**) with statistically significant differences between MSA style and No MSA style groups. The MSA style group is more likely to use Integration contract, Component, and End to End testing strategies to test microservices systems than the No MSA style group. On the other hand, the No MSA style group is more likely to use Consumer-driven contract, Scalability, and A/B testing strategies to test microservices systems than the MSA style group. *Employees size* ≤ 100 vs. *Employees size* > 100 : We identified only one survey question statement (i.e.,  **End to End testing**) with statistically significant difference between Employees size ≤ 100 and Employees size > 100 groups. The Employees size ≤ 100 group is more likely to use End to End testing strategy to test microservices systems than the Employees size > 100 group. Moreover, we did not find any survey question statement about using testing strategies with statistically significant difference for the Employees size > 100 group.

Testing challenges (see Table 22): *Experience* ≤ 2 years vs. *Experience* > 2 years: There are two survey question statements (i.e.,  **TC4**,  **TC7**) with statistically significant differences between Experience ≤ 2 years and Experience > 2 years groups. The Experience ≤ 2 years group is more likely to agree that TC4 (i.e., creating and implementing automated tests) and TC7 (i.e., understanding about each testable microservice) challenges have a critical impact on testing of microservices systems than the Experience > 2 years group. However, we did not find any survey question statement about testing challenges with statistically significant difference for the Experience > 2 years group. *MSA style* vs. *No MSA style*: There are two survey question statements (i.e.,  **TC1**,  **TC5**) with statistically significant differences between MSA style and No MSA style groups. The MSA style group is more likely to agree that TC1 (i.e., creating and implementing manual tests) challenge has a severe impact on testing of microservices systems than the No MSA style group. On the other hand, the No MSA style group is more likely to agree that TC5 (i.e., independent microservices) challenge has a critical impact on testing of microservices systems than the MSA style group. *Employees size* ≤ 100 vs. *Employees size* > 100 : We identified only one survey question statement (i.e.,  **TC6**) with statistically significant difference between Employees size ≤ 100 and Employees size > 100 groups. The Employees size ≤ 100 group is more likely to agree that TC6 (i.e., complexity and size of microservices systems) challenge has a severe impact on testing of microservices systems than the Employees size > 100 group. Moreover, we did not find any survey question statement about testing challenges of microservices systems with statistically significant difference for the Employees size > 100 group.

5. Discussion

Our empirical study sought to unveil the state of practice of designing, testing, and monitoring microservices systems. In particular, we conducted an online survey with 106 practitioners from 29 countries and interviewed 6 practitioners from four countries. In the following, Section 5.1 discusses the main findings and implications for researchers. Section 5.2 compares the key findings of our survey with other surveys conducted in the context of non-MSA-based systems, and Section 5.3 discusses the implications for practitioners.

5.1. Analysis of the results

5.1.1. Design of microservices systems

Application decomposition strategies: An application can be divided into small-scale microservices using various decomposition strategies. DDD and business capability strategies define services corresponding to subdomains and business objects, respectively (Richardson, 2018). Our survey participants indicated that the applications can be decomposed using DDD, business capability, and a combination of DDD and business capability, in which the combination of DDD and business capability is identified as the leading strategy (Finding 4). The interview participants also acknowledged this finding that adopting a combination of DDD and business capability (i) provides a better opportunity to achieve decomposition goals, (ii) is suitable when using microservices with DevOps or agile practices, (iii) helps to organize the development and operations teams, and (iv) helps to establish a flexible and scalable design for microservices systems. We have found several responses in which both the interviewees and the survey participants have confirmed that DDD and business capability strategies are complementary to each other and can also help to address the challenges of DC1: *defining boundaries of microservices* and DC3: *managing the complexity of microservices at the design level*. Overall, concerning application decomposition strategies, the result of our study partially confirms the finding of several existing studies (e.g., Zhang et al., 2019; Wang et al., 2021; Richardson, 2018; de la Torre et al., 2018) and the personal experience of the microservices practitioners (e.g., Brown, 2021; Ramalingam, 2021; Norelus, 2021). We also found several application decomposition strategies for microservices (e.g., decomposed by verbs or use cases Taibi and Systä, 2019, data flow-driven approach Li et al., 2019, interface analysis Baresi et al., 2017, service cutter Gysel et al., 2016) in the literature. However, none of the survey and interview participants mentioned these or other strategies for decomposing applications into microservices in their responses.

Research Implication 1

We found that many organizations use the combination of DDD and business capability to decompose monolithic applications into microservices. We assert that future empirical studies should seek the guidelines that detail how to break a monolith into microservices using the DDD and business capability strategies together. Such guidelines would be beneficial for software teams (particularly inexperienced ones) to smoothly migrate from a monolithic system to a microservices system.

MSA architecting activities: The results of this study indicate that practitioners are more likely to agree on using Architectural Evaluation (AE) and Architectural Implementation (AI) activities when designing microservices systems (Finding 6). This finding was also confirmed by the interviewees, who mentioned several reasons for using AE and AI. These results are expected because

Table 21

Statistically significant differences on the survey results about testing strategies.











Survey question statement	SQ#	Likert Distro.	Exper. ≤ 2 vs. > 2 years		MSA style vs. No MSA style		Empl. ≤ 100 vs. > 100	
		Mean value	P-value	Effect size	P-value	Effect size	P-value	Effect size
Strategies to test MSA-based systems (Very important (5), Important (4), Somewhat important (3), Important (2), Not sure (1))								
Unit testing	SQ34	3.82	0.26	−0.64	0.06	0.75	0.93	−0.26
Integration contract testing		3.27	0.57	−0.33	 0.01	0.37	0.90	−0.16
Consumer-driven contract testing		2.92	0.23	0.13	 0.01	−0.03	0.95	0.06
Integration testing		3.75	0.27	−0.13	0.09	0.45	0.94	−0.20
Component testing		3.42	 0.03	−0.16	 0.02	0.14	0.92	−0.25
Scalability testing		3.29	0.25	0.16	 0.001	−0.13	0.90	−0.12
Service component testing		3.26	0.25	−0.01	0.12	0.33	0.92	−0.26
User interface testing		3.41	 0.04	0.42	0.08	0.02	0.96	0.38
End to end testing		3.63	0.41	0.08	 0.02	0.12	 0.03	0.22
A/B testing		2.80	 0.02	0.20	 0.00	−0.42	0.89	0.51

Table 22

Statistically significant differences on the survey results about testing challenges.

Survey question statement	SQ#	Likert Distro.	Exper. ≤ 2 vs. > 2 years		MSA style vs. No MSA style		Empl. ≤ 100 vs. > 100	
		Mean value	P-value	Effect size	P-value	Effect size	P-value	Effect size
Severity (impact) of each challenge related to testing of MSA-based systems? (Catastrophic (5), Major (4), Moderate (3), Minor (2), Insignificant (1))								
TC1	SQ37	2.87	0.45	−0.16	⚙️ 0.03	0.32	0.92	0.09
TC2		3.06	0.27	0.00	0.09	0.23	0.94	−0.11
TC3		3.13	0.26	−0.01	0.05	0.33	0.93	0.08
TC4		2.95	🔧 0.02	0.22	0.08	−0.02	0.94	0.00
TC5		2.89	0.25	−0.08	⚙️ 0.04	−0.05	0.93	0.13
TC6		3.21	0.25	−0.03	0.07	0.16	👥 0.01	0.09
TC7		2.95	🔧 0.04	0.06	0.07	−0.12	0.94	0.09
TC8		3.10	0.55	0.00	0.13	−0.30	0.94	0.06
TC9		3.08	0.24	0.07	0.05	−0.18	0.92	0.23
TC10		3.03	0.26	0.14	0.09	−0.13	0.94	0.15

Table 23

A comparison of the results between this survey and the existing studies.

This survey results	Existing studies									
	Knoche and Has-selbring (2019)	Ghofrani and Lübke (2018)	Viggiano et al. (2018)	Hasel-böck et al. (2018)	Başkarada et al. (2018)	Zhang et al. (2019)	Zhou et al. (2021)	Wang et al. (2021)	Taibi et al. (2017)	Di Francesco et al. (2018)
Decomposition strategies	–	✓	–	–	–	✓	–	✓	–	–
Implementation strategies	–	–	–	–	–	–	–	–	–	–
Architecting activities	–	–	–	–	–	–	–	–	–	–
Description methods	–	✓	–	–	–	–	–	–	–	✓
Diagrams	–	–	–	–	–	–	–	–	–	–
MSA components	–	–	–	–	–	–	–	–	–	–
QAs	–	–	–	–	–	–	–	–	✓	–
MSA patterns	–	–	–	–	–	–	–	–	–	–
Design skills	–	–	–	–	–	–	–	–	–	–
Monitoring metrics	–	–	–	–	–	–	–	✓	–	–
Monitoring practice	–	–	–	–	–	–	–	✓	–	–
Monitoring tools	–	–	–	–	–	–	–	✓	–	–
Testing methodologies	–	–	–	–	–	–	–	–	–	–
Testing tools	–	–	–	–	–	–	–	–	–	–
Testing skills	–	–	–	–	–	–	–	–	–	–
Design challenges	✓	–	✓	–	✓	✓	–	✓	✓	✓
Monitoring challenges	–	–	–	–	–	–	✓	✓	–	–
Testing challenges	–	–	–	–	–	–	–	–	–	–

practitioners focus on identifying, evaluating, and implementing QAs while designing the architecture of microservices systems. According to [Cojocaru et al. \(2019\)](#), the minimum quality requirements to be evaluated during the design of microservices systems are granularity, cohesion, coupling, scalability, response time, security, health management, execution cost, and reusability.

Moreover, many participants confirmed the use of architectural implementation, in which a coarse-grained MSA design is refined into a detailed design. Detailed design (1) provides a clear vision for microservices implementation, such as each microservice's functionality and interface for communicating with

other microservices and (2) allows developers to evaluate design options before implementing them in microservices systems. Furthermore, a good number of participants also considered Architectural Maintenance and Evolution (AME) during the design of microservices systems. This is because generally, microservices systems are developed using iterative development methods (e.g., SCRUM and DevOps, see **Finding 3**) in which changes (e.g., new requirements) are continuously accommodated, and AME helps to ensure the consistency and integrity of the MSA during the iterations. The results regarding using AE, AI, and AME indicate that practitioners pay more attention to the architecting activities close to the solutions space (i.e., AE, AI, and AME) instead of the problem space (i.e., AA and AS).

On the other hand, we received a significant amount of “neutral” responses from the participants about using AA and AS. This result indicates that many participants were undecided about using these two activities for architecting microservices systems. One potential reason is that microservices practitioners are not fully aware of the concepts or terminologies (e.g., AA, AS) used to describe architecting activities. Our results also show that about 40% to 50% of the participants either “strongly agree” or “agree” with using various architecting activities when designing a microservices system; on the other side, around 10% to 20% of the participants “strongly disagree” or “disagree” with using the architecting activities (see Fig. 10). Therefore, the opinions of the participants on using MSA architecting activities are contradictory.

Research Implication 2

We found a lack of understating about MSA architecting activities among the practitioners and some contradictory responses from the survey practitioners. Research is needed to empirically investigate the necessity, understandability, and actual use of the MSA architecting activities in the industry.

MSA description methods: Architecture description methods are used to document architecture of software systems using specific notations (e.g., UML diagrams) in architecture views (Bachmann et al., 2010). The results show that formal approaches are not widely used in the industry to describe MSA, in contrast, the practitioners prefer to use informal and semi-formal methods for describing MSA (**Finding 7**). We identified several reasons about why the practitioners do not prefer formal approaches but tend to use informal and semi-formal methods for describing MSA (see Section 4.2.2). This finding has also been confirmed in several review studies on microservices architecture (e.g., Di Francesco et al., 2019; Alshuqayran et al., 2016; Waseem et al., 2020b).

On the other hand, the results related to diagrams representing the design and architecture of MSA-based systems (see Fig. 12) reflect that practitioners describe MSA in multiple views, which is common practice in documenting architecture (Bachmann et al., 2010). For instance, our results show that the leading diagrams for representing the design and architecture of microservices systems are flowchart, use case diagram, data flow diagram, activity diagram, and sequence diagram (**Finding 8**), which are used to describe the process (dynamic) view of microservices systems (Fowler, 2004). According to Gerber et al. (2017), microservices systems could have many dynamic operations, such as scaling up and scaling down services and deployment of services across data centers or cloud platforms. The behavioral diagrams (e.g., activity diagram, use case diagram) represent the dynamic operations of microservices systems by showing collaborations between microservices and changes to their internal states. We also asked about diagrams that can be used to illustrate the behavioral view of the microservice

system. For instance, communication, state machine, timing, and interaction overview diagrams. We also noticed that a considerable number of the participants reported class, deployment, component, package, and object diagrams for representing the design and architecture of microservices systems, which are used to describe the structural (or Static) view of microservices systems. Moreover, our results indicate that the diagrams used by practitioners for describing MSA also follow the 4+1 architectural view model (Kruchten, 1995), a commonly-used architecture framework for describing the architecture of traditional systems. For instance, use case diagram represents the scenario view; flowchart, activity, functional flow block, communication, and timing diagram represent the process view; class diagram, object diagram, package diagram, a composite structure diagram, and microservice circuits represent the logical view; component diagram represents the implementation view, and deployment diagram represents the deployment view.

Research Implication 3

Regarding MSA description methods, the survey results confirm the findings of our recent SMS on microservices architecture in DevOps (Waseem et al., 2020b), in which we identified that MSA is mostly described by using informal description methods (e.g., Boxes and Lines). Therefore, we suggest further research about the pros and cons of using informal description methods for describing MSA. Moreover, it is also interesting to investigate how formal description methods (e.g., DSLs, ADLs) can be adapted to describe MSA in industry.

Quality attributes: Our study results show that security, availability, performance, and scalability are the most important QAs considered in MSA design (**Finding 9**). Previous research (e.g., Li et al., 2020; Waseem et al., 2020b) also highlights the importance of these QAs. For instance, **Finding 9** partially confirms the study results by Li et al. (2020) that scalability, performance, availability, monitorability, and security are identified as the most concerned QAs in MSA. We reported several reasons collected from the interviews in the Results section regarding **Finding 9** (see Section 4.2.3). Here, we further analyzed the top four QAs: (1) About security, unlike traditional monolithic applications, MSA-based systems are composed of independent services that work together to fulfill the business goals. Microservices provide public interfaces, use network-exposed APIs, and develop applications using polyglot technologies, thus open the doors to intruders, and consequently the security of messages, queues, and API endpoints requires advanced strategies to protect at the infrastructure and code level. According to the literature review conducted by Pereira-Vale et al. (2019), authentication, authorization, and credentials are the most widely used security strategies by microservices developers. (2) Regarding availability, one of the main benefits of MSA (and also one of the reasons that the industry adopts MSA) is that it provides a fault-tolerant architectural style (Newman, 2015). It also implies that practitioners must use virtualized resources (e.g., containers, Virtual Machines (VMs), data storage) to ensure a high availability of microservices systems (Toffetti et al., 2015). (3) Performance is a major concern for all types of systems. With MSA, an application is decomposed into microservices based on functionality. The count of integration points could increase in dozens or even hundreds while integrating microservices, thus, in turn, leads to a performance challenge due to frequent inter-process communication of microservices. Therefore, performance is essential to be addressed while designing MSA. (4) Scalability is not straightforward to achieve. In microservices systems, business functionalities are written in different programming languages, loaded on different

hardware, executed on different servers, and deployed across multiple clouds. Therefore, architects should devise strategies that allow microservices to be coordinated to scale or eventually identify which microservices must be scaled to meet the increased demand (Wolff, 2016). The survey conducted by Dragoni et al. (2017b) also highlights the importance of scalability in MSA design.

Overall, we received a significant number of responses in which practitioners consider all of the remaining QAs as “important” or “very important”, such as reliability, usability, maintainability, and compatibility. It is because, MSA has a significant influence (e.g., positive impact) on most of QAs (Waseem et al., 2020b). For example, MSA can help to improve reliability through scalability and redundancy with the flexibility of increasing the instances of services (Singleton, 2016).

Research Implication 4

We found that security is one of the important QAs during the design of microservices systems. We also identified the design patterns (e.g., Access token) that support security in microservices systems. However, there is still a lack of knowledge of state-of-the-practice security techniques and tools as well as their potential security vulnerability, which deserves an in-depth empirical investigation.

MSA design patterns: A pattern is a reusable solution to problems recurring in a particular context. Our results show that the most often used MSA patterns are API gateway, Backend for frontend, and Access token (Finding 10). API gateway and Backend for frontend patterns are also known as external API patterns used to address how different clients (e.g., mobile and desktop version) of MSA-based systems access the individual services. In contrast, Access token pattern is employed to address the problem of microservices security. We identified several reasons for using these patterns in the Results section (see Section 4.2.3). The result regarding the frequent use of security pattern (Access token) is partially aligned with the result that security is considered as an important QA in MSA design. The aggregated responses regarding QAs (e.g., security), patterns (e.g., Access token), and design challenges (e.g., addressing security concerns) suggest that security requires more attention during the design of microservices systems.

On the other hand, practitioners reported the frequent use of deployment patterns, such as Service instance per container, Service instance per host, and Service instance per VM, which suggests that practitioners prefer to deploy service instances in isolation. The reasons are that (i) it is easy to monitor, manage, and redeploy each service instance, (ii) deployment of service instances in isolation helps to achieve scalability, performance, and high cohesion by increasing the number of e.g., containers, hosts, VMs, and (iii) deployment of service instances in isolation facilitates the development of microservices by using multiple languages. Another challenging aspect of MSA-based systems is the communication between microservices (Chelliah et al., 2017), and several design patterns provide efficient communication mechanisms (e.g., synchronous and asynchronous communication) between microservices. According to the results, many practitioners use Messaging, Remote procedure invocation, and Domain-specific protocol patterns for microservices communication. It is because, these communication patterns (i) provide loose runtime coupling, (ii) increase microservices availability, and (iii)

use popular communication technologies (e.g., RESTful APIs,²⁸ gRPC,²⁹ Apache Thrift³⁰).

Moreover, our survey results cover the data management patterns, such as API composition, Database per service, Sagas, Shared database, and CQRS, which are also used frequently in the industry. Benefits like a simple method to query data, supporting polyglot database strategies, promoting data consistency, providing ACID transactions, and supporting multiple denormalization could be the possible reasons for frequent use of these data management patterns in the industry for microservices systems (Richardson, 2019). For instance, API composition pattern provides a simple method to query data in microservices systems, Database per service pattern allows the use of polyglot database strategies according to service needs, Sagas pattern ensures the data consistency across multiple microservices services, Shared database pattern provides straightforward ACID transactions to enforce data consistency. In contrast, the CQRS pattern improves the separation of concerns and supports multiple denormalization of databases for microservices systems. Furthermore, we found in the literature (e.g., Richardson, 2020) that service discovery is a key but challenging issue in cloud-based microservices systems due to the containerized and virtualized environment and dynamically assigned network locations. Our survey results indicate that a fear number of practitioners frequently employ several patterns for service discovery. For instance, Service registry, Client-side discovery, Server-side discovery, and Self-registration. The reasons for using these patterns are (i) identification of the dynamically assigned locations to service instances for service clients (e.g., mobile or desktop browsers) and (ii) ability to manage the load balancing requests.

Research Implication 5

Our study results highlight the importance of quality attributes and the use of design patterns for microservices systems. Future studies can present decision models for selecting MSA design patterns, which can assist developers and architects in choosing appropriate patterns by considering the positive and negative effect of QAs for diverse design areas (e.g., communication, service discovery, monitoring and logging) of microservices systems.

Major MSA design challenges and their solutions: The survey results show that a large number of participants responded that DC1: *clearly defining the boundaries of microservices*, and DC2: *addressing security concerns* as two major design challenges (Finding 11). Some other survey and interview-based studies (e.g., Zhang et al., 2019; Di Francesco et al., 2018; Knoche and Hasselbring, 2019) also reported the empirically identified challenges related to MSA design. Compared with the existing work of Zhang et al. (2019), Di Francesco et al. (2018), and Knoche and Hasselbring (2019), our study has independently identified several unique challenges, reasons, and solutions. Defining the boundaries of microservices is deciding about microservices functionality according to business capabilities or subdomains. This challenge mainly occurs due to a lack of business knowledge to establish the right boundaries, for example, in a new application domain. Our survey results also indicate that unclear boundaries have a negative impact on the monitoring and testing of microservices systems (see Table 8).

To address DC1, we have collected several suggestions from the interviews (see Section 4.2.5). For instance, gaining a deeper

²⁸ <https://restfulapi.net/>.

²⁹ <https://grpc.io/>.

³⁰ <http://thrift.apache.org/>.

understanding of business requirements, introducing and implementing miniservices (Gupta, 2020), using DDD, and asking experts for their judgments. Another frequently reported challenge is addressing security concerns (i.e., DC2). Microservices systems are vulnerable to various security attacks. Therefore, the security of microservices systems demands serious attention. In our study and several other studies, the security of microservices is recognized as one of the leading challenges faced by many practitioners. For instance, Ponce et al. (2021) developed a taxonomy of smells and refactorings for microservices security by reviewing the gray and peer review literature. In our recent work (i.e., Waseem et al., 2021), security issues are also identified as one of the top issues (10.18%) in microservices systems. The study result indicates that practitioners use patterns (e.g., Access token), cloud solutions (e.g., Role-Based Access Control), and development practices (e.g., DevSecOps) as solutions to address security challenges (see Section 4.2.5). We also found that some review studies (e.g., Waseem et al., 2020b; Pereira-Vale et al., 2019; Mao et al., 2020; Pereira-Vale et al., 2021) collected a comprehensive set of standards, strategies, and mechanisms from gray and peer review literature that can be used as solutions to secure microservices systems.

Research Implication 6

We collected several strategies, design patterns, DSLs, and development methods to address MSA design challenges (see Table 9). However, further research is needed to develop and empirically evaluate the solutions for addressing the challenges in MSA design as listed in Table 8.

5.1.2. Monitoring of microservices systems

Monitoring metrics: Infrastructure monitoring metrics involve the status of the infrastructure and the servers on which the microservices are running, while microservice metrics concern the metrics of a specific microservice. As reported in Fig. 15, most of the respondents focus on resource usage, database connection, load balancing, and threads that serve to monitor the infrastructure. It is not surprising as the infrastructure monitoring metrics require less development, maintenance, and operation effort (Fowler, 2017). Moreover, the infrastructure software (e.g., Docker, Kubernetes) generally provides tools to collect such metrics. The microservices metrics, such as availability, errors and exceptions, and endpoint success, are more complex to implement because they depend on the languages, libraries, and frameworks (Fowler, 2017). Therefore, it is not unusual that the participants declared to use less of these application metrics (e.g., availability, errors and exceptions, and endpoint success). Furthermore, infrastructure and application metrics are used to monitor quality aspects, such as performance (e.g., resource usage, threads, database connections, and latency), availability (e.g., status of service), scalability (e.g., load balancing), and reliability (e.g., errors and exceptions). These QAs assume particular importance when building microservices systems. For example, more than 85% of the respondents considered performance a very important or essential QA when designing microservices systems (see Table 6). These quality aspects are in line with companies' motivation when adopting the MSA style to develop or modernize their applications, i.e., to have scalable, high-performance, and resilient systems (e.g., Bertram, 2020).

Research Implication 7

Our study results highlight that practitioners prefer to use infrastructure monitoring metrics (e.g., resource usage, database connections) over application monitoring metrics (e.g., availability, errors and exceptions). The latter, although more challenging to implement and provides better information for monitoring microservices systems. Therefore, it would be interesting to propose empirically evaluated monitoring practices and tools for application metrics.

Monitoring practices: The participants indicated log management, exception tracking, and health check API as the frequently used monitoring practices as shown in Fig. 16. Microservices and their instances run on multiple machines (e.g., VMs, containers), and these instances are dynamically created and destroyed during the execution of the microservices. Thus, an adequately managed log data (e.g., centralized logging) helps to identify the service failure, their causes, and unexpected behavior of microservices systems. In contrast, poorly managed log data can affect the performance of microservices systems adversely. Therefore, it is not surprising that 68.8% (33.0% very often + 35.8% often) of the participants use log management very often or often (see Fig. 16). Exceptions may occur when some services handle the requests of other services or instances. Typically, monitoring tools, e.g., Grafana, Zipkin (see Fig. 17), report exceptions to a centralized exception tracking service that aggregates and tracks exceptions and notifies them to developers and debugging professionals for further actions (e.g., addressing the underlying issues) (Richardson, 2018). Another monitoring practice, frequently mentioned by the participants, is health check API (see Fig. 16). From the survey results, we observed that practitioners use monitoring practices that allow developing production-ready microservices as also highlighted in Richardson (2018).

Research Implication 8

Concerning the monitoring practices, we found that many practitioners frequently use log management, exception tracking, and health check API to better trace issues and identify causes of issues in microservices systems. The recognized issues and their causes, through these monitoring practices, can be used to propose empirically validated taxonomies for microservices issues.

Monitoring tools: The participants mentioned that the most commonly used monitoring tools are Jira, Datadog Kafka Dashboard, Spring Boot Actuator, Grafana, and Zipkin. Jira is developed as an issue tracking tool, but now it can be integrated with tools for monitoring microservices systems, such as Dynatrace (Grabner, 2021), Prometheus (Yenin, 2021), and Grafana (Yenin, 2021). To understand the opinions of the microservices practitioners about Jira, we also asked the interview question IQ3.3.1, and the results indicate that the practitioners have a different understanding of Jira as a monitoring tool for microservices systems, for instance, three interviewees suggested Jira integrated with IDE to monitor microservices systems, two interviewees considered issue tracking (e.g., Jira) as part of monitoring of microservices systems, and one interviewee was not sure about the use of Jira for monitoring purposes. On the other hand, Datadog Kafka Dashboard and Spring Boot Actuator are mainly used to gather monitoring data. One reason that the participants prefer to use such tools could be that most of them are application developers (65.1%), therefore, they mainly focus on using monitoring tools for collecting application data instead of the ones more oriented to analyze or visualize data. The results show that practitioners do not prefer to use visualizing and alerting data tools for monitoring

microservices systems concerning the dashboard and alerting components. For example, only 17.9% (19 out of 106) of the participants use Grafana, and 16.9% (18 out of 106) use Zipkin. One reason could be that 59.4% of the participants declared that they have 0 to 2 years of experience in microservice development, therefore, they may not yet have well experienced the added value of the monitoring activities resulting from these tools' use.

Research Implication 9

Our survey reports that the practitioners frequently use tools to monitor network performance, resource usage, and errors and exceptions for microservices systems. Therefore, future studies can empirically investigate the need, understandability, and actual use of monitoring tools in the industry.

Monitoring challenges and solutions: Most of the challenges faced by the participants during the monitoring activities focus on several aspects, such as collecting monitoring metrics data and logs from containers, and distributed tracing (see Table 13). These results show that microservice monitoring challenges are mostly related to their highly distributed nature, autonomy, independence, and self-containment. Concerning the challenge of collecting monitoring data and container records, the practitioners treat it as moderate (48.1%) or minor (23.6%) severity (see Table 14). It is because there are solutions to deal with these challenges. For example, the practitioners recommend monitoring tools, such as Prometheus, Spring Boot Actuator, and Grafana, to search, analyze, and visualize monitoring metrics. On the other hand, the respondents pointed out that distributed tracing of microservices systems is challenging because of the complexity and dynamicity originating from traditional distributed systems. However, only a few participants declared to use monitoring tools, such as Zipkin and Jaeger, to support distributed tracing. It could be the reason why distributed tracing raises several challenges, such as distributed state, concurrency, and heterogeneity of microservices (Spoonhower, 2020).

Research Implication 10

Our research indicates that many organizations use tools and frameworks to address monitoring challenges (see Table 15). However, further research can be explored to develop and empirically evaluate the solutions for addressing the challenges, such as having many components to monitor (complexity), performance monitoring, and distributed tracing.

5.1.3. Testing of microservices systems

Testing strategies: Concerning testing microservices systems in the industry (RQ3), the results indicate that the most widely used testing strategies are unit, E2E, and integration testing (see Fig. 18). Whether in microservices or monolithic systems, unit testing provides a type of analysis that allows examining the functionality of methods or code modules. However, unit testing evaluates components in isolation, consequently, it is also necessary to test the communication paths and interactions between microservices. This is why integration testing is also prevalent for testing of microservices systems. Unit and integration testing are robust testing strategies, but the complexity of microservices systems demands more systematic testing strategies. That is why E2E testing emerges as one of the preferred testing strategies for microservices systems in the industry. E2E testing tests whether a microservices system satisfies the business goals, i.e., verifies that the microservices system meets external requirements and achieves its goals from an end user perspective (Clemson, 2020). Fig. 18 also shows that A/B testing and consumer-driven contract

testing are not frequently used to test microservices systems. In this regard, we have identified the following reasons for both testing strategies. About A/B testing, 33.0% of the participants use this strategy very often or often, however, 53.8% of the participants rarely or never use A/B testing, which shows a contradictory preference of practitioners. We identified the following reasons: (1) there are not enough libraries or tools to use A/B testing (Renz et al., 2016), (2) there are not enough users to perform A/B testing (BrillMark, 2020), (3) A/B testing takes a long time (e.g., minimum seven days) to get results (BrillMark, 2020), and (4) for some domains (e.g., healthcare, transportation) A/B testing does not produce significant results (Leonardi, 2020). About consumer-driven contract testing, we identified two reasons: (1) development teams have communication obstacles when several people working on one microservice (e.g., over 8 people, see Fig. 6 (Right)) and (2) microservices systems extensively use third party resources (Lehvä et al., 2019).

Additionally, concerning the testing strategies (see Fig. 18), the survey results from practitioners partially coincide with the results of our recent mapping study from academia regarding testing of MSA-based applications (Waseem et al., 2020a), in which we identified 39 testing techniques used to test MSA-based applications. Among the testing strategies collected in our survey, nine strategies (i.e., unit testing, E2E testing, integration testing, user interface testing, component testing, scalability testing, A/B testing, service component testing, and consumer-driven contract testing) (see Fig. 18) can be found in the results of our recent SMS (Waseem et al., 2020a). Nevertheless, we found no evidence about integration contract testing in the SMS, despite its popularity (47.2% of the participants very often or often use this testing strategy) in the industry.

Research Implication 11

End-to-End testing is a practical solution for testing each microservice or the whole microservices system. Nevertheless, using this critical testing strategy can cause several challenges in the testing process, such as delivery delays, generating bugs that are difficult to isolate, and liability conflicts. Future work should address the challenges related to E2E testing to make it more beneficial for testing microservices systems.

Testing tools: Our survey results indicate that the respondents use a variety of tools to test microservices systems. The most used testing tool is JUnit (51.9% of preferences) that is used for unit testing. Unit testing is one of the most straightforward but useful testing strategies. It can be used for testing microservices systems and other types of applications or services. In monolithic systems, applications structure all their functionalities into a single set (Dragoni et al., 2017a). Nevertheless, in microservices systems, functionalities are isolated, which makes testing of each microservice possible. Although other tools (e.g., Cucumber, Sauce Labs, Travis) offer a more complete and automated suite of tests for microservices systems, the respondents still chose traditional testing strategies (e.g., unit testing, integration testing) as shown in Fig. 18 and eventually selected popular tools (e.g., JUnit, JMeter) as listed in Fig. 19. The respondents also highlighted that unit testing is part of the core testing strategies for microservices systems (Clemson also agrees with this observation (Clemson, 2020)). Other testing strategies, such as integration, component, consumer-driven contract, and E2E testing are also part of these core testing strategies. Some of the tools (e.g., JUnit, Selenium, Mockito) presented in Fig. 19 partially support these core testing strategies (e.g., integration and E2E testing). However, most of the mentioned tools (e.g., JUnit, JMeter) are not dedicatedly developed to test microservices systems.

Research Implication 12

The survey and interview results indicate a need for more investments in testing tools specific to microservices systems. Such tools can systematically and automatically test all components and elements of microservices systems (e.g., microservices, microservice collaboration, communication, data processing, networking, databases, CI/CD pipelines).

Testing challenges and solutions: Table 18 shows 10 testing challenges identified in this survey. These challenges range from manual testing to polyglot technologies for testing microservices systems. The results indicate that the microservices testing challenges are mostly related to creating and *TC1: Creating and implementing manual tests*, *TC2: Integration testing of microservices*, *TC3: Debugging of microservices on container platforms*, and *TC4: implementing automated testing*. Compared to monolithic systems, practitioners need more expertise about the system design and the flow between microservices to successfully test microservices systems. One reason is that it is necessary to control and verify all communication points, as well as the functionality of all microservices. Therefore, more than one testing strategies are required to test all the components (e.g., APIs, messaging formats, services, databases) that characterize microservices systems.

On the other hand, the responses from the participants show that performing manual testing of microservices systems is challenging because of the complexity of microservices systems. This challenge can be addressed by using automated testing tools, such as Cucumber, Sauce Labs, and Travis. These tools provide comprehensive control over the test cases that need to be executed on each microservice. At this point, a useful approach to manage the complexity of microservices systems is Behavior-Driven Development (BDD). As suggested by a respondent (see Section 4.4.4), BDD provides a common vocabulary among stakeholders (such as domain experts and developers). The vocabulary of the BDD approach can be used to establish test cases for testing microservices systems, such as automated acceptance testing (e.g., Rahman and Gao, 2015).

Research Implication 13

We collected several testing strategies, approaches, and guidelines as solutions to address the respective testing challenges (see Table 20). However, these solutions are not enough to address all microservices testing challenges. Further research is needed to develop and empirically evaluate the solutions for addressing the testing challenges in microservices systems as listed in Table 18.

5.2. Comparing the survey results with non-MSA-based systems

In this section, we compare the key findings of our survey with the surveys that explore the design, monitoring, and testing of non-MSA-based systems.

5.2.1. Design of microservices systems

Although we explored studies related to application decomposition, approaches to high-level system design, architecting activities, architecture description, and design challenges reported in the industry, we did not find sufficient evidence to be compared with our results. Nevertheless, we realized that addressing quality attributes is a major concern in the design of both non-MSA-based and MSA-based systems. The results of the surveys conducted by Kassab et al. (2018), Ameller et al. (2016) and our study show that some quality attributes (e.g., performance, availability and security) are the common interest to practitioners

of both MSA-based and non-MSA-based systems. Kassab et al. (2018) found that the most relevant quality attributes from an architectural design point of view in practice are usability, performance, modifiability, and availability. Ameller et al. (2016) identified that the most important quality attributes for practitioners in service-based systems are dependability, performance, security, and usability.

Although these studies mention other quality attributes, performance appears to be the quality attribute of common interest between MSA-based and non-MSA-based systems. Aspects related to a system's responses to perform actions for a certain period of time are concerns that must be addressed in any system. Other quality attributes, such as availability and security are also of interest to practitioners. Both attributes address recovery concerns from failures and security incidents, which are relevant in MSA-based systems. Scalability, on the other hand, is also relevant because the design of MSA-based systems must consider aspects of concurrency and partitioning (Fowler, 2016). Concurrency allows each task to be broken down into smaller pieces, while partitioning is essential to allow these smaller pieces to be processed in parallel. Therefore, scalability becomes more relevant when designing MSA-based systems, as these systems often address requirements that involve constant traffic growth. In these cases, each microservice must be able to scale with the entire MSA-based system without suffering from performance degradation.

5.2.2. Monitoring of microservices systems

To compare our results concerning the monitoring of MSA-based systems with non-MSA-based systems, we explored studies related to monitoring metrics, practices, tools, challenges, and solutions reported in the industry. However, we did not find sufficient evidence to be compared with our results. We only identified one literature survey in which Cassar et al. reviewed runtime monitoring instrumentation techniques for non-MSA-based systems (Cassar et al., 2017). The results of this survey and our study show that only one monitoring practice (i.e., log management) is commonly used by practitioners of both MSA-based and non-MSA-based systems. Log management practice facilitates the scanning and monitoring of log files generated by servers, applications, and networks. The generated log files are used to detect and solve various types of problems (e.g., performance and security issues) of MSA-based and non-MSA-based systems.

5.2.3. Testing of microservices systems

We explored several survey studies on testing conducted in the context of non-MSA-based systems (e.g., Lee et al., 2012; Raulamo-Jurvanen et al., 2019; Kochhar et al., 2019; Garousi and Zhi, 2013; Quesada-López et al., 2019; Jahan et al., 2019; Garousi et al., 2015; Dias-Neto et al., 2017; Garousi et al., 2017, 2020), and we compared our key findings on testing strategies, tools, and challenges with these surveys. Regarding testing strategies, the results of our survey and the surveys related to non-MSA-based systems (e.g., Lee et al., 2012; Raulamo-Jurvanen et al., 2019; Kochhar et al., 2019; Garousi and Zhi, 2013; Quesada-López et al., 2019) show that unit testing is commonly used as a *de facto* strategy for testing all types of systems. Meanwhile, integration testing is also a commonly used testing strategy for both MSA-based and non-MSA-based systems (e.g., Quesada-López et al., 2019; Raulamo-Jurvanen et al., 2019; Garousi et al., 2015). To our surprise, we did not find evidence regarding the use of scalability, integration contract, and consumer-driven contract testing strategies in the context of non-MSA-based systems. About testing tools, JUnit, JMeter, and Selenium are commonly used by practitioners of both MSA-based and non-MSA-based systems (e.g., Kochhar et al., 2019; Garousi and Zhi, 2013; Quesada-López

Table 24
Abbreviations used in this study.

AA	Architectural Analysis
ACID	Atomicity, Consistency, Isolation, Durability
ADL	Architectural Description Language
AS	Architectural Synthesis
AE	Architectural Evaluation
AI	Architectural Implementation
AME	Architectural Maintenance and Evolution
API	Application Program Interface
BDD	Behavior-Driven Development
CQRS	Command Query Responsibility Segregation
DevOps	Development and Operations
DDD	Domain-Driven Design
DQ	Demographic Question
DSDM	Dynamic Systems Development Model
DSL	Domain-Specific Language
E2E	End-to-End
FDD	Feature Driven Development
IDC	International Data Corporation
IDE	Integrated Development Environment
JAD	Joint Application Development
MDD	Model-Driven Development
MSA	Microservices Architecture
QA	Quality Attribute
RAD	Rapid Application Development
RQ	Research Question
RUP	Rational Unified Process
SOA	Service Oriented Architecture
SLA	Service-Level Agreement
SMS	Systematic Mapping Study
SQ	Specific Question
UML	Unified Modeling Language
VM	Virtual Machine
XP	Extreme Programming

et al., 2019; Jahan et al., 2019), for example, Selenium is used as a tool to perform testing on MSA-based systems (see Fig. 19), and is also employed for testing non-MSA-based systems due to its test automation capabilities. Regarding testing challenges, the results of our survey and other surveys (e.g., Quesada-López et al., 2019; Garousi et al., 2017, 2020) show that two testing challenges (i.e., integration testing, creating and implementing automated tests) are commonly faced by practitioners of both MSA-based and non-MSA-based systems.

5.3. Implications for practitioners

1. As shown in Fig. 7 (Right), the majority of the participants use Scrum and DevOps as the development methods for microservices systems. It could be interesting to further explore what Scrum and DevOps practices are used to design, develop, and maintain microservices systems, the challenges that arise when MSA is used with Scrum and DevOps practices, and how practitioners address these challenges.
2. The results of SQ11 indicate that 36.7% (39 out of 106) of the participants' organizations do not have a dedicated architecture team or an architect who is responsible for the architecture of microservices systems. It is then meaningful to explore the impact of the absence of architecture teams or architects on the quality of microservices systems as well as the countermeasures employed in practices.
3. Our survey results show that many organizations (42.3%) use the combination of DDD and business capability to decompose monolithic applications into microservices. We assert that software organizations and practitioners develop guidelines that detail how to break a monolith into microservices using the DDD and business capability strategies together. Such guidelines would be beneficial for software teams (particularly inexperienced ones) to smoothly

migrate from a monolithic system to a microservices system.

4. Our survey results highlight that the complexity of microservices systems poses several challenges to the design, monitoring, and testing of microservices systems. However, we did not get any satisfactory solutions from our survey that address such challenges. In this regard, a comprehensive investigation with practitioners is required to understand the nature of complexity and its mitigating strategies in the context of microservices systems.
5. We asked two survey questions (i.e., SQ30, SQ35) regarding monitoring and testing tools for microservices systems. However, most of the tools are general purpose tools that are not dedicatedly developed for microservices systems. The complex nature of microservices systems further exposes the weakness of the existing tooling, specifically in the context of intelligent monitoring and testing of microservices systems. Therefore, practitioners can put effort into developing and using such tools for improved productivity.

6. Threats to validity

This section outlines the validity threats to our survey by following the guidelines described by Easterbrook et al. (2008).

Internal validity refers to the factors that could negatively affect the data collection and analysis process. We considered the following threats to the internal validity of this survey:

- **Participants selection:** When the survey invitations were sent through emails, those who responded may lack the required expertise to participate in our survey. We tried to minimize this issue by searching microservices practitioners through personal contacts and relevant platforms (see detailed approaches in Section 3.2.2). Additionally, the characteristics of the prospective participants were made explicit in the survey preamble.
- **Closed-ended survey questions:** Most of the survey questions are closed-ended. The primary threat to this type of questions is an insufficient number of answer options. In this regard, the answer options of multiple-choice questions were defined by taking into account the existing literature. Moreover, we allowed the participants to answer the survey questions according to his/her experience in the free text area. As a result, we received several other answers besides the already provided answer options.
- **Length of the survey:** Our survey is composed of 39 survey questions, which may have affected the response rate of our survey. While almost 9000 emails were sent over three months to the potential practitioners, we only got 135 responses (i.e., 1.5%), out of which 106 were valid. We tried to mitigate this threat, to some extent, by reducing the number of open-ended questions as much as possible to avoid memory bias (Murphy-Hill et al., 2015). Second, in the survey preamble, we mentioned that completing the survey would take 20–25 min, which might help practitioners find their best available time to fill out the survey.

Construct validity in a survey focuses on whether the survey constructs are defined and interpreted correctly (Easterbrook et al., 2008). Design, monitoring, and testing of microservices systems are the core constructs of this survey. Having said this, we identified the following threats:

- **Inadequate explanation of the constructs:** This threat refers to the fact that the constructs are not sufficiently described before they are translated into measurements and treatments. To deal with this threat, we designed the survey

Table 25
Questions on demographics of the survey.

ID	Survey questions	Type of answers
DQ1	Which country are you working in?	Selection from countries list
DQ2	(Multiple Choice) What are your major responsibilities in your company?	Application Developer/Database Developer/ Database Administrator/Architect/System Analyst/Tester/DevOps Engineer/Operations Staff/Business Analyst/Consultant/C-level Executive (CIO, CTO, etc.)
DQ3	Regarding the usage of Microservices Architecture (MSA) for developing applications	(1) We use MSA style for building all applications in our organization/team. (2) The MSA style is only used for building some specific applications in our organization/team. (3) Our organization/team does not use MSA style at all
DQ4	How many years have you been working in the IT industry?	0–2 Years/3–5 Years/6–10 Years/> 10 Years
DQ5	How many years have you been working with MSA-based systems?	0–2 Years/3–5 Years/6–10 Years/> 10 Years
DQ6	Have you ever received any professional training (i.e., excluding higher education) related to MSA?	Yes/No
DQ7	(Multiple Choice) What is the work domain of your organization?	E-commerce/Education/Embedded systems/Financial/Healthcare/Insurance/Internet/Manufacturing/Professional services/Real estate/Telecommunication/Transportation and Warehousing/Other
DQ8	How large is your organization?	1–20 employees/21–100 employees/101–500 employees/501–1000 employees/ >1000 employees
DQ9	On average, how many people work (i.e., design, develop, test, deploy, etc.) on one microservice in your organization?	1/2–3/4–5/6–7/8–9/> 10
DQ10	(Multiple Choice) Which of the following software development methods is (commonly) used in your MSA projects?	DevOps/SCRUM/Crystal Methods/DSDM/XP/FDD/JAD/Lean Development/RAD/Rational Unified Process/Iterative Development/Spiral/Waterfall (a.k.a. Traditional)/We do not use any defined method/Other

based on our recent SMSs on MSA in DevOps (Waseem et al., 2020b) and testing of microservices (Waseem et al., 2020a) as well as the existing literature on microservices systems (e.g., Rajput, 2018; Di Francesco et al., 2019; Alshuqayran et al., 2016; Swersky, 2020; Lumetta, 2019; Clemson, 2020). Additionally, the authors had several internal meetings to identify issues regarding question format, understandability, and consistency. Moreover, we invited two experts who had experience in survey-based research to check the validity and integrity of the survey questions. The feedback collected from the experts helped us improve several survey questions that were found confusing or unclear. Finally, we evaluated the question format, understandability, and consistency of the survey questions by conducting a pilot survey, which was completed by ten microservices practitioners. Based on the feedback received from the pilot survey participants, we also rephrased a few survey questions (i.e., SQ21, SQ26, SQ33, SQ39).

- **Fear of being evaluated:** This threat is related to the fact that some people, due to work restrictions, anxiety, or other reasons, are afraid of being evaluated. To mitigate this threat, at the beginning of the survey, we clearly mentioned that all the responses gathered through this survey will be treated with confidentiality and privacy and that the data (responses) gathered will be used for academic purposes only.
- **Survey dissemination platforms:** The survey was disseminated on several social media and professional networking groups (see Table 2). The main threat to survey dissemination platforms is the identification of the relevant groups. To address this threat, we read the existing discussion in the groups. After ensuring that the group members frequently discuss different aspects (e.g., techniques, tools, technology issues) of microservices, we posted our survey invitation.
- **Inclusion of valid responses:** To identify valid answers, we asked three survey questions about the responsibilities of practitioners (DQ2), usage of MSA for developing applications (DQ3), and experience in the development of microservices systems (DQ5), and applied inclusion and exclusion

criteria on received answers (see Table 3). We removed those answers that were “meaningless”, “inconsistent”, and “logically senseless” (see the details in Section 3.2.3). We also excluded the entire response from a respondent if more than 5 survey questions have inconsistent, randomly filled, or meaningless answers.

External validity is related to the extent to which the study results can be generalized to other cases. We sent almost 9000 emails to microservices practitioners and made dozens of posts in relevant groups available at social and professional networking platforms to reach the practitioners (see Table 2). We get 106 valid responses from 29 countries (see Fig. 1). The respondents' demographics show that they had various responsibilities (see Fig. 2 and worked in a wide range of teams and organizations in terms of size (see Fig. 6) and development method (see Fig. 7). Our participants had varying experiences on the development of microservices systems, from less than one year to more than ten years, in different domains (see Section 4.1). We also interviewed 6 respondents from four countries to understand the reasons behind the key survey findings. Although we have a relatively large sample size from the survey and interviews and variations in participants' demographics in terms of professional experience, company size, and work domains. Still, our findings may not generalize or represent all microservices practitioners' perspectives. For example, most of the survey respondents (i.e., 56.6%) are from Pakistan and China. It would be interesting to perform another study with diverse practitioners in order to increase the external validity of the study results.

Conclusion validity is related to the issues that affect the right conclusion in empirical studies. To mitigate this threat, the first author analyzed the survey data, and the second and third authors reviewed the survey data analysis results. Conflicts on data analysis results were resolved through mutual discussions and brainstorming among all the authors. Furthermore, the conclusion can be influenced as different researchers may interpret the inclusion and exclusion criteria for selecting responses differently. To minimize the effect of this issue, we defined and applied

Table 26

Questions on design of MSA-based systems.

ID	Survey questions	Type of answers
SQ11	Does your organization have a dedicated individual or team for creating and managing architecture for MSA-based systems?	Yes/No/Yes, but with some other responsibilities
SQ12	(Multiple Choice) How do you decompose an application into microservices?	We define microservices corresponding to business capabilities./We define microservices corresponding to DDD sub-domains./We define microservices by using the combination of business capability and Domain-Driven Design (DDD)./ Other (please specify)
SQ13	Which strategy do you follow from design to implementation of MSA-based systems?	We model the high-level design for MSA-based systems that represents only major components (e.g., main microservices, databases, messaging systems), and then write the code based on that design model./We model the detailed design for MSA-based systems that represents all components (e.g., design for individual microservice), and then write the code based on that design model./We directly write the code for MSA-based systems without any design model./We use a Model Driven Development (MDD) approach to generate the code from the design model./Other (please specify)
SQ14	We consider Architecturally Significant Requirements (ASRs) when designing MSA-based systems (Architectural Analysis).	Strongly disagree/Disagree/Neutral/Agree/Strongly agree
SQ15	We design the candidate architecture solutions to address ASRs when designing MSA-based systems (Architectural Synthesis).	Strongly disagree/Disagree/Neutral/Agree/Strongly agree
SQ16	We evaluate the candidate architecture solutions when designing MSA-based systems to ensure design decisions are made right (Architectural Evaluation).	Strongly disagree/Disagree/Neutral/Agree/Strongly agree
SQ17	We refine the coarse-grained architecture design into the fine-grained detailed design so that developers can write the code based on the MSA (Architectural Implementation).	Strongly disagree/Disagree/Neutral/Agree/Strongly agree
SQ18	To ensure the consistency and integrity of the MSA, we regularly fix issues in or update the MSA according to changed operational environments or new requirements (Architectural Maintenance and Evolution).	Strongly disagree/Disagree/Neutral/Agree/Strongly agree
SQ19	(Multiple Choice) Which of the following methods do you use for describing the architecture of MSA-based systems?	UML/Architectural description language/Domain-specific language/Boxes and Lines
SQ20	(Multiple Choice) Which of the following diagrams do you use for representing the design and architecture of MSA-based systems?	Composite structure diagram/Interaction overview diagram/Object diagram/Timing diagram/State machine diagram/Package diagram/Communication diagram/Functional flow block diagram/BPMN/Component diagram/Deployment diagram/Class diagram/Sequence diagram/Activity diagram/Data flow diagram/Use case diagram/Flowchart diagram
SQ21	(Multiple Choice) Which of the following entities can be considered as architectural design components of MSA-based systems?	Service interface (e.g., Pages, Grids/Panels)/Process (e.g., Use cases)/Domain (e.g., Domain objects, Factories/Repositories)/Service (e.g., Service gateways, Service clients)/Storage (e.g., Relations, MongoDB) /All the above/Other (please specify)
SQ22	(Multiple Choice) Please rate the importance of the following quality attribute when designing MSA-based systems. Please use the Likert-scale points "Very Important, Important, Somewhat Important, Not Important, and Not Sure" to indicate the importance of QAs.	Availability/Compatibility/Functional suitability/Maintainability/Portability/Performance/Reliability/Scalability/Usability/Security/Resilience/Testability/Monitorability/Reusability/Interoperability
SQ23	(Multiple Choice) Please indicate how often you use the following MSA design patterns when designing MSA-based systems. Please use the Likert-scale points "Very Often, Often, Sometimes, Rarely, and Never" to indicate the use of MSA patterns.	Access token/API composition/API gateway/Application events/Backend for frontend/Circuit breaker/ Client-side discovery/Command query/CQRS/Database per service/Domain-specific protocol/Event sourcing/Messaging/Sidecar/3rd party registration/ Transactional outbox/Language-specific packaging format/Polling publisher/Remote procedure invocation/Sagas/Self-registration/Serverless deployment/Server-side discovery/deployment platform/Deploy a service as container/Service mesh/Deploy a service as a VM/Service registry/Shared database/Transaction log tailing/Other (please specify)
SQ24	(Multiple Choice) What challenges do you face when you design the architecture for MSA-based systems?	Clearly defining the boundaries of microservices/Finding appropriate modeling abstractions for microservices/Recovering MSA from existing code/Managing microservices complexity at the design level/Reaching scalability/Separating functional and operational concerns/Addressing data management at the design level/Addressing microservices communication at the design level/Addressing security concerns/Aligning team structure with the architecture of MSA-based system/Other (please specify)

(continued on next page)

the inclusion and exclusion criteria on all the received responses (see Section 3.2.3). We also confirmed the interpretations of the results and conclusions by arranging several brainstorming sessions among all the authors.

7. Related work

In this section, we report the survey studies that investigated microservices systems from different perspectives as well as the difference between our survey and these existing surveys. Knoch

Table 26 (continued).

ID	Survey questions	Type of answers
SQ25	(Multiple Choice) Which of the following architectural challenges may encounter when monitoring and testing MSA-based systems? <ul style="list-style-type: none"> Monitoring of MSA-based systems Testing of MSA-based systems 	Clearly defining the boundaries of microservices/Finding appropriate modeling abstractions for microservices/Recovering MSA from existing code/Managing microservices complexity at the design level/Reaching scalability/Separating functional and operational concerns/Addressing data management at the design level/Addressing microservices communication at the design level/Addressing security concerns/Aligning team structure with the architecture of MSA-based system/Other (please specify)
SQ26	What solutions (e.g., practices, tools, team structures) can be employed to address or mitigate the challenges related to architecting for MSA-based systems?	Free text
SQ27	(Multiple Choice) What skills (including hard and soft skills) are required to properly design and implement MSA?	A practitioner should have the skill to break down a complex coding task into smaller tasks that can be implemented separately/A practitioner should have the skill to implement the functionality of a microservice by using suitable MSA patterns (e.g., circuit breaker)/A practitioner should have architectural knowledge of technology decisions/A Practitioner should have knowledge about the basic concepts of DevOps/A practitioner should have knowledge about container technologies/A practitioner should have knowledge of how to secure microservices/A practitioner should be able to work in a dynamic team structure/Other (please specify)

Table 27

Questions on monitoring of MSA-based systems.

ID	Survey questions	Type of answers
SQ28	(Multiple Choice) Which metrics do you collect for monitoring MSA-based systems?	Resource usage (e.g., CPU, memory)/Threads/Microservice's open File Descriptors (FD)/Database connections/Load balancing/Language-specific metrics (e.g., garbage collector behavior, RPC, and database latency)/Status of service availability/Service Level Agreement (SLA)/Latency (of both the service as a whole and its API endpoints)/Endpoint success/Endpoint response time/Errors and exceptions/Status of dependencies/Other (please specify)
SQ29	(Multiple Choice) Which of the following practices or techniques do you use to monitor MSA-based systems?	Log management/Audit logging/Distributed tracking/Exception tracking/Health check API/Log deployment and changes/Other (please specify)
SQ30	(Multiple Choice) What tools do you use to monitor MSA-based systems?	Datadog Kafka Dashboard/Apache Mesos/Cacti/Grafana/iPerf/Jira/Nagios/Omnia/Pact/Prometheus/Raygun APM/Spring Boot Actuator/Zipkin/Other (please specify)
SQ31	(Multiple Choice) What challenges do you usually face when monitoring MSA-based systems?	Having many components to monitor (complexity)/Collection of monitoring metrics data and logs from containers/ Monitoring of application running inside the container/Distributed tracing/Failures zone detection/Performance monitoring/Availability of the monitoring tools/Maintaining monitoring infrastructures/Analyzing the collected data/Other (please specify)
SQ32	(Multiple Choice) How do you rate the severity (impact) of each challenge related to the monitoring of MSA-based systems? Please use the Likert-scale points "Catastrophic, Major, Moderate, Minor, and Insignificant" to rate the severity of monitoring challenges.	Having many components to monitor (complexity)/Collection of monitoring metrics data and logs from containers/ Monitoring of application running inside the container/Distributed tracing/Failures zone detection/Performance monitoring/Availability of the monitoring tools/Maintaining monitoring infrastructures/Analyzing the collected data/Other (please specify)
SQ33	What practices, techniques or tools do you use to improve the monitorability of MSA-based systems (i.e., making MSA-based systems easy to monitor)?	Free text

and Hasselbring surveyed the drivers, barriers, and goals of using microservices to modernize legacy systems along with the impact of microservices on runtime performance and transactionality (Knoche and Hasselbring, 2019). The number of participants who are from Germany is 71 in this survey. The key results regarding the drivers are high scalability and elasticity, high maintainability, and short time to market. The key findings regarding the barriers that prevent the adoption of microservices are insufficient operation skills, resistance from operational staff, and insufficient development skills. The identified goals in the context of microservices are improve maintainability, time to market, and scalability. Regarding the impact of microservices on performance and transactionality, they found that the respondents considered performance degradation as a minor issue, whereas transactionality (e.g., cross-service ACID transactionality) is regarded as a major issue. Ghofrani and Lübke reported the challenges related

to the design and development of microservices systems through a survey study with 25 practitioners (Ghofrani and Lübke, 2018). The reported challenges are distributed nature of MSA, management of multiple repositories, issue tracking in MSA, sharing data among multiple services, dockers networking, and debugging of dependent microservices. Their survey also reported the goals for adopting microservices (e.g., scalability, agility, expendability), approaches for deriving services boundary (e.g., DDD, formal methods), and notations used for describing MSA.

A survey with 122 practitioners was conducted to understand how microservices are practiced in industry by Viggiano et al. (2018). A major focus of their survey is on identification of the programming languages, technologies, challenges, and advantages associated with microservices. The results of their survey indicated that popular programming languages used in microservices systems are Java, JavaScript, Node.js, C#, and PHP, and the

Table 28

Questions on testing of MSA-based systems.

ID	Survey questions	Type of answers
SQ34	(Multiple Choice) Which of the followings testing methodologies do you use to test MSA-based systems?	Unit testing/Integration contract testing/Consumer-driven contract testing/Integration testing/Component testing/Scalability testing/Service component testing/User Interface Testing/End to End Testing/A-B Testing/Other (please specify)
SQ35	(Multiple Choice) What tools do you use to test MSA-based systems?	Cucumber/FitNesse/Gatling/Jasmine/JMeter/Junit/Karma/Locust. IO/Microfocus UFT/Mocha/MockLito/Pdify/Perfecto/Phantom/Sauce Labs/Selenium/SoapUI/Tap/TestNG/Travis/Other (please specify)
SQ36	(Multiple Choice) What challenges do you usually face when testing MSA-based systems?	Creating and implementing manual tests/ Creating and implementing automated tests/Integration testing of microservices/Independent microservices/Debugging of microservices on container platforms/Understanding about each testable microservice/Many independent teams/Performance testing of microservices/Polyglot technologies in MSA-based systems/Complexity and size of MSA-based systems/Other (please specify)
SQ37	(Multiple Choice) How do you rate the severity (impact) of each challenge related to testing of MSA-based systems? Please use the Likert-scale points "Catastrophic, Major, Moderate, Minor, and Insignificant" to rate the severity of testing challenges.	Creating and implementing manual tests can pose a challenge/Creating and implementing automated tests can pose a challenge/Integration testing of microservices can pose a challenge/Independent microservices can pose a challenge for testing/Debugging of microservices on container platforms/Understanding about each testable microservice/Many independent teams can pose a challenge for testing/Performance testing of microservices is challenging/Implementation of polyglot technologies in MSA-based systems can pose a challenge for testing/Complexity/size of the MSA-based system can pose the challenge for testing/Other (please specify)
SQ38	(Multiple Choice) What skills (both soft and hard skills) are required to test MSA-based systems properly?	A tester has the ability to write good unit test cases to detect potential bugs/A tester has the ability to write good integration test cases to detect potential bugs/A tester should have basic knowledge about multiple databases (e.g., SQL, NoSQL, MongoDB)/A tester should have basic knowledge of Linux commands/A tester should have knowledge about test management tools (e.g., a tool for tracking test cases written by your team)/A tester should have knowledge about the test automation tools/A tester should have analytically and logically thinking/Other (please specify)
SQ39	What practices, techniques, or tools do you use to make MSA-based systems more testable (i.e., making MSA-based systems easy to test)?	Free text

Table 29

Interview questions on demographics information.

ID	Interview questions
IQ1.1	Which country are you working in?
IQ1.2	What are your major responsibilities in your company?
IQ1.3	How many years have you been working in the IT industry?
IQ1.4	What is the work domain of your organization?
IQ1.5	How many years have you been working with microservices systems?

common Database Management Systems (DBMS) are PostgreSQL, MySQL, MongoDB, SQL Server, and Oracle. Regarding the communication protocols, 62% of the participants in their survey declared that they used REST over HTTP. The key identified challenges are distributed transactions, microservices testing, faults diagnosing, and slow Remote Procedures Calls (RPCs). The major advantages that can be gained by employing microservices are independent deployment, scalability, maintainability, and flexibility to use several technologies for implementing microservices.

Haselböck et al. conducted interviews with 10 experts to investigate the design areas and their importance in microservices systems (Haselböck et al., 2018). The design areas that they identified are service design (e.g., service partitioning), organization structure design (e.g., communication structures), monitoring and logging infrastructure, system design (e.g., integration, fault tolerance, service discovery, versioning, scalability, security), and automated testing. The participants of this survey assessed that monitoring and logging is the most important microservices design area. Baskarada et al. explored the opportunities and challenges related to MSA through in-depth interviews (Baškarada

et al., 2018). The survey results show that MSA provides better development and deployment agility along with operational scalability over the monolithic architecture. Some of the identified challenges are skills required for working on a different set of technologies, organizational culture, application decomposition into microservices, deciding about when to use orchestration and choreography in MSA, microservices testing, and performance degradation. Zhang et al. interviewed 13 practitioners from China to investigate the gap between literature and practice about microservices (Zhang et al., 2019). They took the nine characteristics of MSA (e.g., componentization via services, organization around business capabilities, smart endpoints, and dumb pipes) proposed by Fowler and Lewis (Lewis and Fowler, 2019) and characterized the gaps between vision and the reality of microservices. They also identified several benefits (e.g., independent upgrade, independent scale (up/down), independent development), and challenges (e.g., chaotic independence, unguided organizational transformation, complexity of API management) related to microservices.

Zhou et al. conducted an industrial survey with 16 practitioners to investigate the effectiveness of existing debugging practices for microservices systems (Zhou et al., 2021). They specifically investigated faults of microservices systems, debugging practices, and the related challenges faced by practitioners. They reported 22 faults with their root causes, and also developed a benchmark system to replicate the identified debugging faults of microservices systems. The results of their study show that proper tracing and visualization analysis can help to locate various kinds of faults, and there is a need for more intelligent trace analysis and visualization techniques. Wang et al. conducted interviews with 21 practitioners and an online follow-up survey completed

Table 30

Interview questions on design of microservices systems.

ID	Questions context	Interview questions
IQ2.1	Our survey results show that most of the respondents (42.3%) used a combination of DDD and business capability to decompose applications into microservices. Regarding this finding, we would like to ask the following three questions:	<p>IQ2.1.1: Do you think that this finding reflects the actual use of application decomposition strategies (i.e., a combination of DDD and business capability) for microservices (Yes OR No)? Please give a reason(s) to support your answer.</p> <p>IQ2.1.2: What could be the reasons for the frequent use of a combination of DDD and business capability strategy to decompose applications into microservices?</p> <p>IQ2.1.3: Are you familiar with any other strategy that can be used to decompose applications into microservices? If yes, please provide the name of the strategy.</p>
IQ2.2	Our survey results show that 46.7% of the respondents reported that they only created a high-level design for major microservices before implementing them. Regarding this finding, we would like to ask the following two questions:	<p>IQ2.2.1: What could be the reasons for creating the only high-level design for major microservices?</p> <p>IQ2.2.2: What could be the possible risks if practitioners do not create a design or detailed design for microservices systems?</p>
IQ2.3	Regarding MSA architecting activities when designing microservices systems investigated in this survey (Architectural Analysis (AA), Architectural Synthesis (AS), Architectural Evaluation (AE), Architectural Implementation (AI), and Architectural Maintenance and Evolution (AME)), we would like to ask the following two questions:	<p>IQ2.3.2: What could be the reasons for the frequent use of Architecture Evaluation (AE) activity when designing microservices systems?</p> <p>IQ2.3.3: What are the reasons that practitioners pay more attention to architecting activities close to the solutions space (i.e., AE, AI, AME) instead of the problem space (i.e., AA, AS)?</p>
IQ2.4	Our survey results show that a large number of participants did not use architectural description languages (84.7%) and domain-specific languages (72.7%) for describing the architecture of microservices systems. Regarding this finding, we would like to ask the following question:	<p>IQ2.4.1: What are the reasons that practitioners do not use architectural description languages and domain-specific languages for describing the architecture of microservices systems?</p> <p>IQ2.4.2: What are the reasons that practitioners prefer to use informal (e.g., Boxes and Lines) and semi-formal (e.g., UML) approaches for describing MSA?</p>
IQ2.5	Our survey results show that security, availability, performance, and scalability are the most important quality attributes for microservices systems. Regarding this finding, we would like to ask the following question:	IQ2.5.1: What could be the reasons that practitioners consider security, availability, performance, and scalability the most important quality attributes for microservices systems?
IQ2.6	Our survey results show that API gateway and Backend for frontend as communication patterns, Database per service as data management pattern, and Access token as security pattern are reported as the most often used MSA design patterns in practice. Regarding this finding, we would like to ask the following question:	IQ2.6.1: What could be the reasons that these patterns are frequently used for designing microservices systems?
IQ2.7	Our survey results show that the following are the most prominent challenges of designing microservices systems <ul style="list-style-type: none"> Clearly defining the boundaries of microservices Addressing security concerns Managing microservices complexity at the design level 	<p>IQ2.7.1: What could be the reasons that these design challenges happen?</p> <p>IQ2.7.2: What could be the potential solutions to address these design challenges?</p>

Table 31

Interview questions on monitoring of microservices systems.

ID	Questions context	Interview questions
IQ3.1	Our survey results show that the resource usage (66.9%), load balancing (53.7%), and availability (52.8%) are the most frequently used monitoring metrics for microservices systems. Regarding this finding, we would like to ask the following question:	IQ3.1.1: What could be the reasons that these monitoring metrics are frequently used?
IQ3.2	Our survey results show that the log management (68.8%), exception tracking (60.4%), and health check API (57.6%) are the most frequently used monitoring practices for microservices systems	IQ3.2.1: What could be the reasons that these monitoring practices are frequently used?
IQ3.3	Our survey results show that the majority of practitioners used Jira, Datadog Kafka Dashboard, Spring Boot Actuator, and Grafana to monitor microservices systems. However, not all the tools are dedicated for monitoring microservices systems, for example, Jira is a bug tracking tool. Regarding this finding, we would like to ask the following question:	IQ3.3.1: Do you agree that Jira can be used to monitor microservices systems? If so, please explain how Jira supports the monitoring of microservices systems?
IQ3.4	Our survey results show that the following are the most prominent challenges of monitoring microservices systems. <ul style="list-style-type: none"> Collection of monitoring metrics data and logs from containers Distributed tracing Having many components to monitor (complexity) 	<p>IQ3.4.1: What could be the reasons that these monitoring challenges happen?</p> <p>IQ3.4.2: What could be the potential solutions to address these monitoring challenges?</p>

by 37 practitioners to investigate the best practices, challenges, and solutions concerning successfully developing microservices systems (Wang et al., 2021). Specifically, they investigated three aspects of microservices systems: architecture (e.g., microservice granularity), infrastructure (e.g., logging and monitoring), and code management (e.g., managing API change). They found

that a clear sense of microservices ownership, controlled API management, process automation, and investment in robust logging and monitoring infrastructure are the best practices for the successful development of microservices systems. In addition, the study (Wang et al., 2021) indicates that it is always not beneficial to use a plurality of languages (multiple languages)

Table 32

Interview questions on testing of microservices systems.

ID	Questions context	Interview questions
IQ4.1	Our results show that the most commonly used testing strategies to test microservices systems are unit testing (63.3%), integration testing (59.4%), and E2E testing (55.6%). Regarding this finding, we would like to ask the following question:	IQ4.1.1: What could be the reasons that these testing strategies are frequently used?
IQ4.2	Our results show that most of the practitioners used Junit (51.8%), JMeter (26.3%), and Mocha (21.7%) to test microservices systems. Regarding this finding, we would like to ask the following question:	IQ4.2.1: What could be the reasons that these testing tools are frequently used?
IQ4.3	Our results show that the following are the most prominent challenges of testing microservices systems. <ul style="list-style-type: none"> • Creating and implementing manual tests • Integration testing of microservices • Debugging of microservices on container platforms 	IQ4.3.1: What could be the reasons that these testing challenges happen? IQ4.3.2: What could be the potential solutions to address these testing challenges?

and utilize business capabilities to decompose applications into microservices. Finally, they revealed that managing the shared code between microservices is challenging.

Taibi et al. conducted an empirical study by interviewing 21 practitioners who migrated their monoliths to microservices. In their work, they identified the motivations (e.g., maintainability, scalability), issues (e.g., monolith decoupling, database migration and data splitting), and benefits (e.g., maintainability improvement, scalability improvement) for the migration from monoliths to microservices. They also proposed a migration process framework after comparing three migration processes adopted by the interviewed practitioners (Taibi et al., 2017). Regarding the migration from monoliths to microservices, Di Francesco et al. also conducted interviews with 18 practitioners to report the migration activities and challenges. Domain decomposition, identification of the new system's services, and application of DDD practices are recognized as leading migration activities. In contrast, high coupling, identification of the microservice boundaries, and decomposition of the pre-existing systems are identified as leading migration challenges. Besides that, we also identified an important body of knowledge that reports dozens of patterns for implementing different aspects (e.g., data management, deployment) of microservices systems (Alshuqayran et al., 2016; Richardson, 2018; Márquez and Astudillo, 2018; Marquez et al., 2018; Taibi et al., 2018a). However, none of these studies report how often practitioners use these patterns when designing microservices systems.

Table 23 shows the comparison between the results of our survey and the existing studies. It is demonstrated in Table 23 that the results of our survey and the existing studies are significantly different. For instance, the existing studies do not report MSA architecting activities from practice, and do not discuss MSA patterns and QAs of microservices systems as well. Also, we did not find any study that reports the evidence from practice about monitoring and testing of microservices systems. Note that the symbol “✓” in Table 23 indicates that our survey results are similar to some extent with the results of related work. In contrast, the symbol “-” indicates that our survey results are different from the results of existing studies.

8. Conclusions

Through this mixed-methods study, we gained insight into microservices systems regarding how practitioners design, monitor, and test microservices systems in the industry. The study results are meaningful for both researchers and practitioners who intend to understand: (1) team structure, application decomposition and implementation strategies, MSA architecting activities, MSA design description methods, QAs of microservices systems, MSA patterns, and skills required to design and implement MSA; (2) microservices monitoring metrics, practices, and tools, (3)

microservices testing strategies, tools, and skills, (4) the challenges and solutions related to design, monitoring, and testing of microservices systems, as well as (5) reasons behind key findings of the study. The main findings of this survey are summarized as follows:

1. From 29 countries, 106 practitioners participated in our survey. 65% of the participants have a role of application developers, and 59.5% of the participants indicated that their organizations use the MSA style for building some applications. In most of the organizations, 2 to 3 people are working on one microservice. Regarding application domains, almost one-third (33.1%) of the participants' organizations develop E-commerce applications. Moreover, Scrum (59.4%) and DevOps (50%) are the leading development methods employed in the MSA projects in the industry.
2. For the interviews, we recruited 6 practitioners from four countries. Mainly they have design and development responsibilities. Their average work experience in the IT industry is 11.7 years, and their average work experience in microservices systems is 3.5 years.
3. The majority of the organizations (63.2%) have a team or person for designing the MSA. Applications are mainly decomposed by using the combination of DDD and business capability strategies, responded by 42.3% of the participants. 46.7% of the participants indicated that they only modeled the high-level design of microservices systems. Concerning MSA architecting activities, practitioners are more likely to focus on architecture evaluation, implementation, and maintenance and evolution. More than half of the participants (50.9%) used informal approaches (i.e., Boxes and Lines) for describing MSA. Flowchart, use case diagram, data flow diagram, activity diagram, and sequence diagram are the most popular diagrams for representing the design and architecture of microservices systems. Security, availability, performance, and scalability are ranked as most important QAs by the participants for designing microservices systems. Regarding MSA design patterns used in practices, API gateway, Database per service, Backend for frontend, and Access token are reported as the most often used patterns.
4. Resource usage, load balancing, availability, and database connections are most frequently used monitoring metrics. About monitoring practices, more than 50% of the participants pointed out that they used log management, exception tracking, and health check API most often. Jira (39.6%) and Datadog Kafka Dashboard (38.6%) are identified as the popular tool for monitoring microservices systems.
5. Unit testing, E2E, and integration testing are recognized as the top three strategies to test microservices systems. The most popular tools among the participants for testing microservices systems are Junit, Jmeter, and Mocha.

6. The most frequently reported challenges regarding the design of microservices systems are *Clearly defining the boundaries of microservices* (69.8%), *addressing the security concerns* (46.2%), and *managing microservices complexity at the design level* (34.9%). Regarding monitoring challenges, we found that *collection of monitoring metrics and logs from containers* (54.7%), *distributed tracing* (45.3%), and *having many components to monitor (complexity)* (41.5%) are the leading challenges. Concerning testing challenges, practitioners highlighted *creating and implementing manual tests* (50%), *integration testing of microservices* (43.4%), and *debugging of the microservices on a container platform* (42.5%) as critical challenges.
7. We received several recommendations (e.g., design patterns, approaches and tools) from practitioners to address some of the design, monitoring, and testing challenges. However, there are still several challenges that do not have dedicated solutions. For instance, *recovering MSA from existing code*, *failure zone detection*, *performance testing of microservices*.
8. We also compared the key findings of our survey with the surveys conducted with non-MSA-based systems (see Section 5.2).

We have also identified several promising research directions to improve the practices of the design, monitoring, and testing of microservices systems (see the research implications in Section 5.1). We argue that researchers and practitioners should work together to bring more dedicated solutions to address the challenges reported in this study.

CRediT authorship contribution statement

Muhammad Waseem: Conceptualization, Investigation, Writing - original draft, Visualization. **Peng Liang:** Conceptualization, Investigation, Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition. **Mojtaba Shahin:** Conceptualization, Investigation, Writing - review & editing, Visualization, Supervision. **Amleto Di Salle:** Investigation, Writing - review & editing, Visualization. **Gastón Márquez:** Investigation, Writing - review & editing, Visualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is partially sponsored by the National Key R&D Program of China with Grant No. 2018YFB1402800 and the NSFC with Grant No. 62172311, the Italian Ministry of Economy and Finance, Cipe resolution n. 135/2012 (project INCIPICT - INnovating City Planning through Information and Communication Technologies), the Department of Electronics and Informatics of the Federico Santa María Technical University (Concepción, Chile) and SISMA national research project funded by the MIUR under the PRIN 2017 program (Contract 201752ENYB). We would also like to thank all the participants who contributed their knowledge and insights to this industrial survey.

Appendix A. Abbreviations used in this study

See Table 24.

Appendix B. Survey and interview questionnaires

The Welcome Page of the Survey

Microservices Architecture (MSA) is a cloud-native architectural style that is inspired from Service-Oriented Architecture (SOA). In general, microservices are organized as a suite of small granular services that can be implemented through different technological stacks.

This survey aims at understanding how MSA-based systems are designed, tested, and monitored in the software industry. The results of this survey will be used to develop an evidence-based body of knowledge to better and further support of design, monitoring, and testing of MSA-based systems.

The survey takes approximately **20–25** minutes to complete. This survey is comprised of 4 sections with the questions about:

1. Background of the survey participants (e.g., their role in software project)
2. The design aspect of MSA-based systems
3. Monitoring of MSA-based systems
4. Testing of MSA-based systems

We are keen to invite you to participate, if

1. you are professionally working in the software industry and your organization has adopted/is adopting or is planning to adopt MSA; or
2. you are an independent practitioner and have experience and/or expertise with MSA-based systems.

Please note, all of the responses gathered through this survey will be treated with confidentiality and privacy. The results of this survey will be used for academic purposes only. We would very much appreciate your participation in this research project

See Tables 25–32.

References

- Alshuqayran, N., Ali, N., Evans, R., 2016. A systematic mapping study in microservice architecture. In: Proceedings of the 9th IEEE International Conference on Service-Oriented Computing and Applications (SOCA). IEEE, pp. 44–51.
- Ameller, D., Galster, M., Avgeriou, P., Franch, X., 2016. A survey on quality attributes in service-based systems. *Softw. Qual. J.* 24 (2), 271–299.
- Arijs, P., 2020. Comparing microservices and monolithic applications from the perspective of monitoring. p. 4. <https://dzone.com/articles/comparing-microservices-and-monolithic-application> (accessed on 2020-04-29).
- Bachmann, F., Bass, L., Garlan, D., Ivers, J., Little, R., Merson, P., Nord, R., Stafford, J., 2010. Documenting Software Architectures: Views and Beyond, second ed. Addison-Wesley Professional.
- Balalaie, A., Heydarnoori, A., Jamshidi, P., 2016. Microservices architecture enables devops: Migration to a cloud-native architecture. *IEEE Softw.* 33 (3), 42–52.
- Baresi, L., Garriga, M., De Renzis, A., 2017. Microservices identification through interface analysis. In: Proceedings of the 6th European Conference on Service-Oriented and Cloud Computing (ESOCC). Springer LNCS, pp. 19–33.
- Başkarada, S., Nguyen, V., Koronios, A., 2018. Architecting microservices: Practical opportunities and challenges. *J. Comput. Inf. Syst.* 58 (1), 1–9.
- Bass, L., Clements, P., Kazman, R., 2012. Software Architecture in Practice, third ed. Addison-Wesley Professional.
- Bertram, A., 2020. Seven reasons to switch to microservices and five reasons you might not succeed. <https://www.cio.com/article/3201193/7-reasons-to-switch-to-microservices-and-5-reasons-you-might-not-succeed.html> (accessed on 2020-01-21).

- Brillmark, 2020. Five common A/B testing problems and how to overcome them. <https://www.brillmark.com/5-common-a-b-testing-problems-and-how-to-overcome-them/> (accessed on 2020-08-10).
- Brown, K., 2021. Apply domain-driven design to microservices architecture. <https://www.ibm.com/garage/method/practices/code/domain-driven-design/> (accessed on 2021-01-25).
- Carstensen, N., 2020. What is log management? A complete logging guide. <https://www.graylog.org/post/what-is-log-management-a-complete-logging-guide> (accessed on 2020-01-19).
- Cassar, I., Francalanza, A., Aceto, L., Ingólfssdóttir, A., 2017. A survey of runtime monitoring instrumentation techniques. In: Proceedings of the 2nd International Workshop on Pre- and Post-Deployment Verification Techniques (PrePost). EPTCS, pp. 15–28.
- Cerny, T., Donahoo, M.J., Trnka, M., 2018. Contextual understanding of microservice architecture: current and future directions. ACM SIGAPP Appl. Comput. Rev. 17 (4), 29–45.
- Chelliah, P.R., Subramanian, H., Murali, A., 2017. Architectural Patterns: Uncover Essential Patterns in the Most Indispensable Realm of Enterprise Architecture. Packt Publishing Ltd..
- Clemson, T., 2020. Testing strategies in a microservice architecture. <https://martinfowler.com/articles/microservice-testing/> (accessed on 2020-01-10).
- Cojocaru, M.-D., Oprescu, A., Uta, A., 2019. Attributes assessing the quality of microservices automatically decomposed from monolithic applications. In: Proceedings of the 18th International Symposium on Parallel and Distributed Computing (ISPD). IEEE, pp. 84–93.
- Comunytex, 2020. Githubanalysis2018. <https://comunytex.com/en/introduction-to-microservices/> (accessed on 2020-3-11).
- Dave, A., Degioanni, L., 2020. The five principles of monitoring microservices. <https://thenewstack.io/five-principles-monitoring-microservices/> (accessed on 2020-01-19).
- Dawid, A., 2020. Microservices through the eye of the QA expert – testing microservices. <https://tsh.io/blog/testing-microservices-strategy-and-tools/> (accessed on 2020-05-25).
- Di Francesco, P., Lago, P., Malavolta, I., 2018. Migrating towards microservice architectures: An industrial survey. In: Proceedings of the 15th International Conference on Software Architecture (ICSA). IEEE, pp. 29–38.
- Di Francesco, P., Lago, P., Malavolta, I., 2019. Architecting with microservices: A systematic mapping study. J. Syst. Softw. 150, 77–97.
- Dias-Neto, A.C., Matalonga, S., Solari, M., Robiolo, G., Travassos, G.H., 2017. Toward the characterization of software testing practices in south america: Looking at Brazil and Uruguay. Softw. Qual. J. 25 (4), 1145–1183.
- Dragoni, N., Giallorenzo, S., Lafuente, A.L., Mazzara, M., Montesi, F., Mustafin, R., Safina, L., 2017a. Microservices: Yesterday, today, and tomorrow. In: Mazzara, M., Meyer, B. (Eds.), Present and Ulterior Software Engineering. Springer, pp. 195–216.
- Dragoni, N., Lanese, I., Larsen, S.T., Mazzara, M., Mustafin, R., Safina, L., 2017b. Microservices: How to make your application scale. In: Proceedings of the 11th International Andrei Ershov Memorial Conference on Perspectives of System Informatics (PSI). Springer LNCS, pp. 95–104.
- Easterbrook, S., Singer, J., Storey, M.-A., Damian, D., 2008. Selecting empirical methods for software engineering research. In: Shull, F., Singer, J., Sjberg, D.I.K. (Eds.), Guide to Advanced Empirical Software Engineering. Springer, pp. 285–311.
- EBSE, 2019. Template for a survey protocol. <https://github.com/wasimsse/EBSE/blob/main/SurveyTemplate.pdf> (accessed on 2019-02-20).
- Esposito, C., Castiglione, A., Choo, K.-K.R., 2016. Challenges in delivering software in the cloud as microservices. IEEE Cloud Comput. 3 (5), 10–14.
- Fan, C.Y., Ma, S.P., 2017. Migrating monolithic mobile application to microservice architecture: An experiment report. In: Proceedings of the 1st International Conference on AI & Mobile Services (AIMS). IEEE, pp. 109–112.
- Fowler, M., 2004. UML Distilled: A Brief Guide to the Standard Object Modeling Language. Addison-Wesley Professional.
- Fowler, S.J., 2016. Production-Ready Microservices: Building Standardized Systems Across an Engineering Organization. O'Reilly Media, Inc..
- Fowler, S.J., 2017. Production-Ready Microservices. O'Reilly Media.
- Gamma, E., Helm, R., Johnson, R., Vlissides, J.M., 1994. Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley Professional.
- Garousi, V., Coşkunçay, A., Betin-Can, A., Demirörs, O., 2015. A survey of software engineering practices in Turkey. J. Syst. Softw. 108, 148–177.
- Garousi, V., Felderer, M., Kuhrmann, M., Herkilöglu, K., 2017. What industry wants from academia in software testing?: Hearing practitioners' opinions. In: Proceedings of the 21st International Conference on Evaluation and Assessment in Software Engineering (EASE). ACM, pp. 65–69.
- Garousi, V., Felderer, M., Kuhrmann, M., Herkilöglu, K., Eldh, S., 2020. Exploring the industry's challenges in software testing: An empirical study. J. Softw.: Evol. Process 32 (8), e2251.
- Garousi, V., Zhi, J., 2013. A survey of software testing practices in Canada. J. Syst. Softw. 86 (5), 1354–1376.
- Gerber, A., Mayr, A., Putz, P., Wallerstorfer, D., 2017. Dynamic microservices. In: Gerber, A., Mayr, A., Putz, P., Wallerstorfer, D. (Eds.), Cloud-Native Evolution. O'Reilly Media.
- Ghofrani, J., Lübke, D., 2018. Challenges of microservices architecture: A survey on the state of the practice. In: Proceedings of the 10th Central European Workshop on Services and their Composition (ZEUS). CEUR-WS, pp. 1–8.
- Glaser, B.G., Strauss, A.L., 2017. Discovery of Grounded Theory: Strategies for Qualitative Research. Routledge.
- Grabner, A., 2021. Optimizing atlassian JIRA and confluence productivity with dynatrace. <https://www.dynatrace.com/news/blog/optimizing-atlassian-jira-and-confluence-productivity-with-dynatrace/> (accessed on 2021-01-21).
- Gupta, N., 2020. Microservice, miniservice, and macroservice. <https://dzone.com/articles/micro-service-mini-service-and-macro-service> (accessed on 2021-01-22).
- Gysel, M., Kölbener, L., Giersche, W., Zimmermann, O., 2016. Service cutter: A systematic approach to service decomposition. In: Proceedings of the 5th European Conference on Service-Oriented and Cloud Computing (ESOCC). Springer LNCS, pp. 185–200.
- Haselböck, S., Weinreich, R., Buchgeher, G., 2018. An expert interview study on areas of microservice design. In: Proceedings of the 11th International Conference on Service-Oriented Computing and Applications (SOCA). IEEE, pp. 137–144.
- Hasselbring, W., 2016. Microservices for scalability: Keynote talk abstract. In: Proceedings of the 7th ACM/SPEC on International Conference on Performance Engineering (ICPE). ACM, pp. 133–134.
- Hasselbring, W., Steinacker, G., 2017. Microservice architectures for scalability, agility and reliability in e-commerce. In: Proceedings of the 1st IEEE International Conference on Software Architecture Workshops (ICSAW). IEEE, pp. 243–246.
- Heinrich, R., van Hoorn, A., Knoche, H., Li, F., Lwakatare, L.E., Pahl, C., Schulte, S., Wettinger, J., 2017. Performance engineering for microservices: research challenges and directions. In: Proceedings of the 8th ACM/SPEC on International Conference on Performance Engineering (ICPE). ACM, pp. 223–226.
- Heorhiadi, V., Rajagopalan, S., Jamjoom, H., Reiter, M.K., Sekar, V., 2016. Gremlin: Systematic resilience testing of microservices. In: Proceedings of the 36th IEEE International Conference on Distributed Computing Systems (ICDCS). IEEE, pp. 57–66.
- Hoda, R., Noble, J., 2017. Becoming agile: A grounded theory of agile transitions in practice. In: Proceedings of the 39th IEEE/ACM International Conference on Software Engineering (ICSE). IEEE, pp. 141–151.
- Hofmeister, C., Kruchten, P., Nord, R.L., Obbink, H., Ran, A., America, P., 2007. A general model of software architecture design derived from five industrial approaches. J. Syst. Softw. 80 (1), 106–126.
- Hoogendoorn, S., 2020. Modeling microservices. https://www.slideshare.net/aahogendoorn/modeling-microservices?from_action=save (accessed on 2020-03-18).
- ISO/IEC, 2016. Systems and Software Engineering: Systems and Software Quality Requirements and Evaluation (SQuARE): Measurement of System and Software Product Quality. Standard ISO 25023:2016.
- ISO/IEC/IEEE, 2011. Systems and Software Engineering – Architecture Description. Standard ISO 42010:2011.
- Jahan, M.S., Riaz, M.T., Abbas, M., 2019. Software testing practices in IT industry of Pakistan. In: Proceedings of the 6th Conference on the Engineering of Computer Based Systems (ECBS). ACM, pp. 1–10.
- Jamshidi, P., Pahl, C., Mendonça, N.C., Lewis, J., Tilkov, S., 2018. Microservices: The journey so far and challenges ahead. IEEE Softw. 35 (3), 24–35.
- Jedrzejewski, B., 2020. Seven essential skills for microservices developers. <https://www.e4developer.com/2018/01/25/seven-essential-skills-for-microservices-developers> (accessed on 2020-01-25).
- Kang, H., Le, M., Tao, S., 2016. Container and microservice driven design for cloud infrastructure devops. In: Proceedings of the 8th International Conference on Cloud Engineering (IC2E). IEEE, pp. 202–211.
- Kassab, M., Mazzara, M., Lee, J., Succi, G., 2018. Software architectural patterns in practice: An empirical study. Innov. Syst. Softw. Eng. 14 (4), 263–271.
- Khare, N., 2020. Monitoring challenges with containers and microservices. <https://hasgeek.com/rootconf/2017/proposals/monitoring-challenges-with-containers-and-microser-LyqwsFpJtJQ3gyMrSXj7YP> (accessed on 2020-01-21).
- Kitchenham, B.A., Pfleeger, S.L., 2008. Personal opinion surveys. In: Shull, F., Singer, J., Sjberg, D.I.K. (Eds.), Guide To Advanced Empirical Software Engineering. Springer, pp. 63–92.
- Knoche, H., Hasselbring, W., 2019. Drivers and barriers for microservice adoption – a survey among professionals in Germany. Enterpr. Model. Inf. Syst. Archit. - Int. J. Concept. Model. 14 (1), 1–35.
- Kochhar, P.S., Xia, X., Lo, D., 2019. Practitioners' views on good software testing practices. In: Proceedings of the 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP). IEEE, pp. 61–70.
- Kruchten, P., 1995. The 4+1 view model of architecture. IEEE Softw. 12 (6), 42–50.
- Lee, J., Kang, S., Lee, D., 2012. Survey on software testing practices. IET Softw. 6 (3), 275–282.

- Lehvä, J., Mäkitalo, N., Mikkonen, T., 2019. Consumer-driven contract tests for microservices: A case study. In: Proceedings of the 20th International Conference on Product-Focused Software Process Improvement (PROFES). Springer LNCS, pp. 497–512.
- Leonardi, L., 2020. When A/B testing isn't worth it. <https://www.braze.com/blog/not-worth-ab-testing/> (accessed on 2020-08-10).
- Lethbridge, T.C., Sim, S.E., Singer, J., 2005. Studying software engineers: Data collection techniques for software field studies. *Empir. Softw. Eng.* 10 (3), 311–341.
- Lewis, J., Fowler, M., 2019. Microservices: A definition of this new architectural term. <http://martinfowler.com/articles/microservices.html> (accessed on 2019-10-11).
- Li, P.L., Ko, A.J., Zhu, J., 2015. What makes a great software engineer?. In: Proceedings of the 37th IEEE/ACM International Conference on Software Engineering (ICSE). IEEE, pp. 700–710.
- Li, Z., Liang, P., Avgeriou, P., 2013. Application of knowledge-based approaches in software architecture: A systematic mapping study. *Inf. Softw. Technol.* 55 (5), 777–794.
- Li, S., Zhang, H., Jia, Z., Li, Z., Zhang, C., Li, J., Gao, Q., Ge, J., Shan, Z., 2019. A dataflow-driven approach to identifying microservices from monolithic applications. *J. Syst. Softw.* 157, 110380.
- Li, S., Zhang, H., Jia, Z., Zhong, C., Zhang, C., Shan, Z., Shen, J., Babar, M.A., 2020. Understanding and addressing quality attributes of microservices architecture: A systematic literature review. *Inf. Softw. Technol.* 131, 106449.
- Lumetta, J., 2019. These are the most effective microservice testing strategies, according to the experts. <https://www.freecodecamp.org/news/these-are-the-most-effective-microservice-testing-strategies-according-to-the-experts-6fb584f2edde/> (accessed on 2019-10-12).
- Malavolta, I., Lago, P., Muccini, H., Pelliccione, P., Tang, A., 2012. What industry needs from architectural languages: A survey. *IEEE Trans. Softw. Eng.* 39 (6), 869–891.
- Mao, R., Zhang, H., Dai, Q., Huang, H., Rong, G., Shen, H., Chen, L., Lu, K., 2020. Preliminary findings about DevSecOps from grey literature. In: Proceedings of the 20th International Conference on Software Quality, Reliability and Security (QRS). IEEE, pp. 450–457.
- Márquez, G., Astudillo, H., 2018. Actual use of architectural patterns in microservices-based open source projects. In: Proceedings of the 25th Asia-Pacific Software Engineering Conference (APSEC). IEEE, pp. 31–40.
- Marquez, G., Osses, F., Astudillo, H., 2018. Review of architectural patterns and tactics for microservices in academic and industrial literature. *IEEE Lat. Amer. Trans.* 16 (9), 2321–2327.
- Mauersberger, L., 2020. Microservices: What they are and why use them. <https://www.softwaretestinghelp.com/why-documentation-is-important-in-software-testing/> (accessed on 2020-3-11).
- Mayer, B., Weinreich, R., 2017. A dashboard for microservice monitoring and management. In: Proceedings of the 1st International Conference on Software Architecture Workshops (ICSAW). IEEE, pp. 66–69.
- Meinke, K., Nycander, P., 2015. Learning-based testing of distributed microservice architectures: Correctness and fault injection. In: Proceedings of the 13th International Conference on Software Engineering and Formal Methods Workshops (SEFMW). Springer LNCS, pp. 3–10.
- Murphy-Hill, E., Zimmermann, T., Bird, C., Nagappan, N., 2015. The design space of bug fixes and how developers navigate it. *IEEE Trans. Softw. Eng.* 41 (1), 65–81.
- Nadareishvili, I., Mitra, R., McLarty, M., Amundsen, M., 2016. Microservice Architecture: Aligning Principles, Practices, and Culture. O'Reilly Media.
- Nakagawa, S., Cuthill, I.C., 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol. Rev.* 82 (4), 591–605.
- Newman, S., 2015. Building Microservices: Designing Fine-Grained Systems. O'Reilly Media.
- Newman, S., 2020. Monolith to Microservices: Evolutionary Patterns to Transform Your Monolith. O'Reilly Media.
- Norelus, E., 2021. Implementing domain-driven design for microservice architecture. <https://medium.com/design-and-tech-co/implementing-domain-driven-design-for-microservice-architecture-26eb0333d72e> (accessed on 2021-01-25).
- O'Brien, L., Merson, P., Bass, L., 2007. Quality attributes for service-oriented architectures. In: Proceedings of the 1st International Workshop on Systems Development in SOA Environments (SDSOA). IEEE, pp. 3–9.
- Pahl, C., Jamshidi, P., 2016. Microservices: A systematic mapping study. In: Proceedings of the 8th International Conference on Cloud Computing and Services Science (CLOSER). Springer CCIS, pp. 137–146.
- Pahl, C., Jamshidi, P., Zimmermann, O., 2018. Architectural principles for cloud software. *ACM Trans. Internet Technol.* 18 (2), 1–23.
- Pereira-Vale, A., Fernandez, E.B., Monge, R., Astudillo, H., Márquez, G., 2021. Security in microservice-based systems: A multivocal literature review. *Comput. Secur.* 103, 102200.
- Pereira-Vale, A., Márquez, G., Astudillo, H., Fernandez, E.B., 2019. Security mechanisms used in microservices-based systems: A systematic mapping. In: Proceedings of the 45th Latin American Computing Conference (CLEI). IEEE, pp. 1–10.
- Ponce, F., Márquez, G., Astudillo, H., 2019. Migrating from monolithic architecture to microservices: A rapid review. In: Proceedings of the 38th International Conference of the Chilean Computer Science Society (SCCC). IEEE, pp. 1–7.
- Ponce, F., Soldani, J., Astudillo, H., Brogi, A., 2021. Smells and refactorings for microservices security: A multivocal literature review. [arXiv:2104.13303](https://arxiv.org/abs/2104.13303).
- Quesada-López, C., Hernandez-Agüero, E., Jenkins, M., 2019. Characterization of software testing practices: A replicated survey in costa rica. *J. Softw. Eng. Res. Dev.* 7 (6), 1–20.
- Rahman, M., Gao, J., 2015. A reusable automated acceptance testing architecture for microservices in behavior-driven development. In: Proceedings of the 9th IEEE International Symposium on Service-Oriented System Engineering (SOSE). IEEE, pp. 321–325.
- Rajput, D., 2018. Hands-on Microservices—Monitoring and Testing: A Performance Engineer's Guide to the Continuous Testing and Monitoring of Microservices. Packt Publishing Ltd..
- Ramalingam, C., 2021. Building domain driven microservices. <https://medium.com/walmartglobaltech/building-domain-driven-microservices-af688aa1b1b8> (accessed on 2021-01-25).
- Raulamo-Jurvanen, P., Hosio, S., Mäntylä, M.V., 2019. Practitioner evaluations on software testing tools. In: Proceedings of the 23rd Evaluation and Assessment on Software Engineering (EASE). ACM, pp. 57–66.
- Renz, J., Hoffmann, D., Staubitz, T., Meinel, C., 2016. Using A/B testing in mooc environments. In: Proceedings of the 6th International Conference on Learning Analytics and Knowledge (LAK). ACM, pp. 304–313.
- Richards, M., 2015. Microservices vs. Service-Oriented Architecture. O'Reilly Media.
- Richardson, C., 2018. Microservices Patterns. Manning Publications.
- Richardson, C., 2019. Microservices Patterns: With Examples in Java. Manning Publications.
- Richardson, C., 2020. Service discovery in a microservices architecture. <https://www.nginx.com/blog/service-discovery-in-a-microservices-architecture/> (accessed on 2020-6-25).
- Richardson, C., Smith, F., 2016. Microservices: From Design to Deployment. Nginx Inc..
- Saha, G.K., 2008. Understanding software testing concepts. *Ubiquity* 9 (6), 1.
- Singleton, A., 2016. The economics of microservices. *IEEE Cloud Comput.* 3 (5), 16–20.
- Soldani, J., Tamburri, D.A., Van Den Heuvel, W.-J., 2018. The pains and gains of microservices: A systematic grey literature review. *J. Syst. Softw.* 146, 215–232.
- Sotomayor, J.P., Allala, S.C., Alt, P., Phillips, J., King, T.M., Clarke, P.J., 2019. Comparison of runtime testing tools for microservices. In: Proceedings of the 43rd Annual IEEE Computer Software and Applications Conference (COMPSAC). IEEE, pp. 356–361.
- Spoonhower, D., 2020. The complete guide to distributed tracing. <https://lightstep.com/distributed-tracing/> (accessed on 2020-08-24).
- Sridharan, C., 2018. Distributed Systems Observability: A Guide to Building Robust Systems. O'Reilly Media.
- Stenberg, J., 2020. The long history of microservices. <https://www.infoq.com/news/2016/11/microservices-history/> (accessed on 2020-3-11).
- Swersky, D., 2020. The hows, whys and whats of monitoring microservices. p. 22. <https://thenewstack.io/the-hows-whys-and-whats-of-monitoring-microservices/> (accessed on 2020-04-29).
- Taibi, D., Auer, F., Lenarduzzi, V., Felderer, M., 2021. From monolithic systems to microservices: An assessment framework. *Inf. Softw. Technol.* 137, 106600.
- Taibi, D., Lenarduzzi, V., Pahl, C., 2017. Processes, motivations, and issues for migrating to microservices architectures: An empirical investigation. *IEEE Cloud Comput.* 4 (5), 22–32.
- Taibi, D., Lenarduzzi, V., Pahl, C., 2018a. Architectural patterns for microservices: A systematic mapping study. In: Proceedings of the 8th International Conference on Cloud Computing and Services Science (CLOSER). Springer CCIS, pp. 221–232.
- Taibi, D., Lenarduzzi, V., Pahl, C., 2018b. Continuous architecting with microservices and DevOps: A systematic mapping study. In: Proceedings of the 8th International Conference on Cloud Computing and Services Science (CLOSER). Springer CCIS, pp. 126–151.
- Taibi, D., Systä, K., 2019. From monolithic systems to microservices: A decomposition framework based on process mining. In: Proceedings of the 9th International Conference on Cloud Computing and Services Science (CLOSER). Springer CCIS, pp. 153–164.
- Tang, A., Avgeriou, P., Jansen, A., Capilla, R., Ali Babar, M., 2010. A comparative study of architecture knowledge management tools. *J. Syst. Softw.* 83 (3), 352–370.
- Tejaswini, 2019. Software testing documentation guide (why it's important). <https://www.softwaretestinghelp.com/why-documentation-is-important-in-software-testing/> (accessed on 2019-9-10).
- Toffetti, G., Brunner, S., Blöchliger, M., Dudouet, F., Edmonds, A., 2015. An architecture for self-managing microservices. In: Proceedings of the 1st International Workshop on Automated Incident Management in Cloud (AIMS). ACM, pp. 19–24.

- de la Torre, C., Wagner, B., Rousos, M., 2018. .NET Microservices: Architecture for Containerized .NET Applications. Microsoft Press.
- Tulie, S., 2020. Microservices vs SOA: What's the difference?. <https://dzone.com/articles/microservices-vs-soa-whats-the-difference> (accessed on 2020-01-20).
- Viggiato, M., Terra, R., Rocha, H., Valente, M.T., Figueiredo, E., 2018. Microservices in practice: A survey study. In: Proceedings of the 6th Brazilian Workshop on Software Visualization, Evolution and Maintenance (VEM). CBSoft, pp. 1–8.
- Wang, Y., Kadyala, H., Rubin, J., 2021. Promises and challenges of microservices: An exploratory study. *Empir. Softw. Eng.* 26 (63), 1–44.
- Waseem, M., Liang, P., Márquez, G., Di Salle, A., 2020a. Testing microservices architecture-based applications: A systematic mapping study. In: Proceedings of the 27th Asia-Pacific Software Engineering Conference (APSEC). IEEE, pp. 119–128.
- Waseem, M., Liang, P., Shahin, M., 2020b. A systematic mapping study on microservices architecture in DevOps. *J. Syst. Softw.* 170, 110798.
- Waseem, M., Liang, P., Shahin, M., Ahmad, A., Nasab, A.R., 2021. On the nature of issues in five open source microservices systems: An empirical study. In: Proceedings of the 25th International Conference on Evaluation and Assessment in Software Engineering (EASE). ACM, pp. 201–210.
- Wolff, E., 2016. Microservices: Flexible Software Architecture. Addison-Wesley Professional.
- Xia, X., Wan, Z., Kochhar, P.S., Lo, D., 2019. How practitioners perceive coding proficiency. In: Proceedings of the 41st IEEE/ACM International Conference on Software Engineering (ICSE). IEEE, pp. 924–935.
- Yarygina, T., Bagge, A.H., 2018. Overcoming security challenges in microservice architectures. In: Proceedings of the 12th Symposium on Service-Oriented System Engineering (SOSE). IEEE, pp. 11–20.
- Yenin, A., 2021. Jira active monitoring 24/7 with prometheus + grafana: use case. <https://polontech.com/blog/how-to-monitor-your-jira-with-prometheus-grafana/> (accessed on 2021-01-21).
- Zhang, H., Li, S., Jia, Z., Zhong, C., Zhang, C., 2019. Microservice architecture in reality: An industrial inquiry. In: Proceedings of the 3rd International Conference on Software Architecture (ICSA). IEEE, pp. 51–60.
- Zhou, X., Peng, X., Xie, T., Sun, J., Ji, C., Li, W., Ding, D., 2021. Fault analysis and debugging of microservice systems: Industrial survey, benchmark system, and empirical study. *IEEE Trans. Softw. Eng.* (2018) 47 (2), 243–260.

Muhammad Waseem is a Ph.D. candidate in the School of Computer Science, Wuhan University, China. Before pursuing his Ph.D., he had worked as a lecturer and teaching research associate in the Department of Computer Science and Software Engineering, International Islamic University, Pakistan. He

received his master degree in software engineering from International Islamic University, Pakistan. His current research interests include software architecture, microservices architecture, and DevOps.

Peng Liang is Professor of Software Engineering in the School of Computer Science, Wuhan University, China. His research interests concern the areas of software architecture and requirements engineering. He is a Young Associate Editor of *Frontiers of Computer Science*, Springer. He has published more than 100 articles in peer-reviewed international journals, conference and workshop proceedings, and books.

Mojtaba Shahin is a Research Fellow at Monash University, Australia. He completed his Ph.D. in Computer Science at The University of Adelaide, Australia in 2018. Previously, he was a lecturer and a researcher with the Department of Computer Engineering, Neyriz Branch, Islamic Azad University, Iran. His current research mainly focuses on Software Architecture, Empirical Software Engineering, DevOps, Continuous Delivery and Deployment, and Human Factors in Software Engineering. He has authored/co-authored +25 peer-reviewed articles in scientific journals and conferences in the field of software engineering such as EMSE, JSS, IJIM, ESEM, WICSA, ECSA, and EASE. Further details can be found on <https://mojtashahin.wordpress.com>.

Amleto Di Salle is a Research Fellow at the Department of Information Engineering, Computer Science, and Mathematics - University of L'Aquila, Italy. His research focuses on the application of model-based software engineering methods, and on the methodologies, formalisms for the analysis, modeling, and automatic generation of (micro)service-oriented distributed systems. He received a Ph.D. in computer science from the University of L'Aquila discussing the thesis entitled "Automated Synthesis of Service Choreographies". He has been involved in several European and national research projects such as CHOREOS, CHOREVOLUTION, INCIPIC, Territori Aperti.

Gastón Márquez is a researcher fellow at the Department of Electronics and Informatics in Federico Santa María Technical University, Chile. His research is focused on architectural tactics, architectural patterns, microservice architectures, clinical software, and telehealth systems. He has published in several international conferences, journals and has participated in international software architecture schools. He also participated as a Research Visitor at the Rochester Institute of Technology (RIT), Rochester, NY, USA, and the Université de Technologie de Compiègne (UTC), Compiègne, France. Before becoming a researcher, he worked in financial companies for five years.