This is a postprint of an article published in the Proceedings of the International Summer School on Search- and Machine Learning-based Software Engineering (SMILESENG) by UCOPress. The final article is available at: https://www.uco.es/ucopress/index.php/es/2018-07-26-11-32-47/e-books/product/970-ebook-proceedings-of-the-international-summer-school-on-search-and-machine-learning-based-software-engineering

Resource optimization in End-to-End Testing

Cristian Augusto
Department of Computing
University of Oviedo
Gijón, Spain
augustocristian@uniovi.es

Jesus Morán
Department of Computing
University of Oviedo
Gijón, Spain
moranjesus@uniovi.es

Claudio de la Riva
Department of Computing
University of Oviedo
Gijón, Spain
claudio@uniovi.es

Javier Tuya
Department of Computing
University of Oviedo
Gijón, Spain
tuya@uniovi.es

***Abstract*—End-to-end (E2E) test suite execution is expensive due to the number of complex resources required. When E2E test suites are executed frequently into a continuous integration system, the total amount of resources required may be prohibitive, moreover when the tests are run in the Cloud with different billing strategies. Traditional techniques to optimize the test suites, such as test suite reduction, minimization, or prioritization, are limited in E2E due to the fact that reordering or selecting a subset of test cases also requires deploying the same expensive system. The current Ph.D. thesis aims to achieve an efficient E2E test execution for large systems in the Cloud. This is done through a smart characterization of the resources required by the test cases, grouping and scheduling them according to their resource usage to avoid unnecessary redeployments and reduce execution time, and finally, executing them into a combination of Cloud infrastructure (i.e., containers, virtual machines, and services) to optimize the costs employed in executing the test suite. Based on the scheduled test cases, we elaborate a cost model for selecting the most cost-effective infrastructure of those available in the Cloud, considering both the cost of the resources required by the test cases and the oversubscription cost (cost incurred in infrastructure contracted and not used during the test suite execution).**

I. Introduction

The execution of End-to-end (E2E) test suites is costly in resource terms due to the expensive and complex systems required. When the test suites are put into a continuous integration environment executed with every repository change, the huge number of resources required can be prohibitive for some companies, particularly when they are run in the Cloud with pay-per-usage billing. The traditional approaches used for optimizing test suites, such as test prioritization, minimization, or reduction, are limited in E2E because selecting a subset or executing the test cases in a fixed order to preserve the effectiveness, may require fewer but also expensive resources. Moving E2E testing to the Cloud brings the opportunity for a better cost but is challenging because of the number of different infrastructures with different costs available to deploy the same resource. Often, the selected infrastructure combination does not match the test suite requirements, which leads to oversubscription. Oversubscription is the difference between the infrastructure contracted and used and causes that part of the cost is not invested in testing tasks, incurring in an extra cost budget [1]. In this thesis, we aim to achieve efficient E2E test suite execution of large systems in the Cloud. For this purpose, we focus on the resources: the physical, logical, or computational entities that are required by the test cases during their execution. We elaborate an orchestration process that, through the information provided by the characterization, schedules the test cases to reduce time and unnecessary resource redeployments and execute the test suite into a combination of Cloud infrastructure. To choose the most cost-effective combination we elaborate a model which considers the cost of the (contracted) infrastructure, as well as the cost of the resources required by the test cases and the cost of oversubscription during the execution.

II. Research Hypothesis

Working in the Cloud has become complex and heterogeneous and with it, selecting the most suitable alternative to carrying out the E2E testing, has become increasingly more challenging. The organization and management of the E2E test cases could derive from different execution times and, depending on how they are deployed in the Cloud, the costs might change. To determine the specific objectives of the research we consider the following four hypotheses:

H1: The execution of End-to-End (E2E) test suites can require large amounts of physical-logical resources, limiting the effectiveness of state-of-the-art techniques such as test case minimization, reduction, or parallelization.

H2: The scheduling of E2E test suites can be optimized by analyzing the dependencies between test cases and looking for concurrency in the use of resources (e.g., web servers, databases, hardware devices, and files, among others).

H3: A smart resource characterization, complemented by grouping and scheduling according to the resources used by the test cases, could save time and avoid unnecessary redeployments in the execution of the test suite.

H4: Considering the cost of the resources and oversubscription during the deployment of the test suite in the Cloud could help to select the most cost-effective combination among the available infrastructure.

III. Objectives of the work

The general objective of this thesis is to optimize end-to-end test suite execution. This objective is particularized in the following sub-objectives:

O1 Analyze the challenges and issues arising from the execution of the E2E Test Suites on large software developments in the Cloud.

- **O2** Characterize the resources employed during the E2E testing by considering different attributes and how the resources are used by the test cases.
- **O3** Improve the E2E test suite execution by considering: (1) relationships between resources, (2) how test cases access the resources, and (3) the resource attributes.
- **O4** Select a cost-effective infrastructure to execute the test suite in the Cloud, considering both the cost of the resources employed in E2E and oversubscription.
- **O5** Experimentation and validation of the approach, executing test suites of real-world scenarios.

IV. Proposed Methodology

The methodologies used to achieve the prior objectives are the following:

1) **Review the state of the art and literature:** we intend to review the state-of-the-art works such as test suite minimization, reduction, prioritization, cloud economics, or cost models in the cloud, among others.
2) **Research-action:** we are collaborating with the *Institute of Information Science and Technologies "Alessandro Faedo"* which provides us with real-world problems faced in European projects such as *Elastest* [2]. Specifically, we intend to validate the research results with several test suites in the industrial field.
3) **Incremental development:** we develop the support tool based on agile methodologies [3] . Each new feature added is a research result according to its relevance

V. First Results

To accomplish the second objective (**O2**), we carried out a characterization of the resources employed in the E2E testing, consisting of several attributes that represent the different resource features and their relationship with the test cases (e.g., available resources, access mode performed by test cases, or if the resource is shared). We propose a four-phase orchestration process called **RETORCH** (acronym of ***R**esource aware **E**nd-to-end **T**est **ORCH**estation*) to accomplish the first three objectives (**O1**, **O2**, and **O3**). **RETORCH** was published at the *QUATIC19* conference [4] and extended to the *Software Quality Journal* [5]. The work was also presented in the *ACM Research competition* in the *ICSE20* [6] and won an award in the *5th edition of the SISTEDES-Everis Awards* [7]. **RETORCH** uses the information from the characterization (***resource identification***) to generate sets of test cases with compatible usage of resources (***grouping***). The sets are split into subsets and scheduled sequentially or in parallel to reduce the execution time and avoid unnecessary redeployments (***scheduling***) and are finally deployed into a continuous integration environment (***orchestration***). The approach was validated (**O5**) with a real-world example of an educational application called *FullTeaching* [8], achieving reductions in the execution time (*61% less than the non-orchestrated test suite*) and fewer resources employed (*due to resource sharing between test cases*). To accomplish the first and fourth objectives (**O1** and **O4**), we develop a cost model focused on the cost of the resources employed in E2E test suites. The model estimates the cost invested in executing the test suite and the oversubscription cost. These two costs with the infrastructure cost (contracted) support the tester's decision-making to choose the most cost-effective infrastructure among those available in the Cloud. This model has also been submitted to the *JISBD22* [9] conference.

VI. Conclusions and Future Work

The current thesis addresses the upcoming issue of optimizing the execution of E2E test suites in the Cloud. Its first results have proven that using a smart characterization of the resources employed on end-to-end test suites, grouping the test cases according to the resource usage, scheduling, and orchestrating them, savings in terms of resources and time can be achieved. The cost model that considers both the cost of executing the test suite and the cost incurred in oversubscription is a work in progress and we expect that it will help in selecting the most cost-effective Cloud infrastructure and obtain a better execution cost. As future work, we want to validate **RETORCH** in more real-world end-to-end test suites. We are exploring how a smarter cost model could improve the efficient E2E test execution. Specifically, we aim to integrate the cost model in an infrastructure advisor engine that analyses all three costs. The purpose of the advisor is to make suggestions about Cloud infrastructure changes that lead to a more cost-effective test suite execution (e.g., new infrastructures available, or changes in those selected that reduce one of the costs).

Ackowledgements

The authors would like to thank **Dra. Antonia Bertolino** (*ISTI-CNR, Pisa Italia*) for her contributions to the line of research. This work was supported in part by the *Spanish Ministry of Economy and digital transformation under **TestBUS (PID2019-105455GB-C32)*** and *SEBASENet 2.0 **(RED2018-102472-T)***".

References

[1] K. Inçki, I. Ari, and H. Sözer, "A survey of software testing in the cloud," *Proceedings of the 2012 IEEE 6th International Conference on Software Security and Reliability Companion, SERE-C 2012*, pp. 18–23, 2012.

[2] URJC, FOKUS, TUB, INSTI-CNR, IMDEA, and ATOS, "Elastest." [Online]. Available: https://elastest.eu/

[3] P. Deemer, G. Benefield, C. Larman, and B. Vodde, *The Scrum Primer: A Lightweight Guide to the Theory and Practice of Scrum*. InfoQ, 2012. [Online]. Available: www.odd-e.com

[4] C. Augusto, J. Morán, A. Bertolino, C. de la Riva, and J. Tuya, "Retorch: Resource-aware end-to-end test orchestration," *Communications in Computer and Information Science*, vol. 1010, pp. 297–310, 9 2019.

[5] ——, "Retorch: an approach for resource-aware orchestration of end-to-end test cases," *Software Quality Journal*, 2020.

[6] C. Augusto, "Efficient test execution in end to end testing," in *Proceedings - 2020 ACM/IEEE 42nd International Conference on Software Engineering: Companion, ICSE-Companion 2020*, 2020, pp. 152–154.

[7] C. Augusto and C. de la Riva, "Optimización de recursos en pruebas de sistema," *5th edition of the SISTEDES-Everis Award*, 2021.

[8] URJC and P. F. Pérez, "Fullteaching - elastest repository," 2019. [Online]. Available: https://github.com/elastest/full-teaching

[9] C. Augusto, J. Morán, C. de la Riva, and J. Tuya, "Modelo de costes para el despliegue de pruebas e2e en entornos cloud," 2022, unpublished Manuscript.