METRICS AND TOOLS THAT ARE AVAILABLE FOR TESTING SCALABILITY

Anila Umar, Muhammad Abbas and Saad Rehman

National University of Sciences & Technology (NUST), Islamabad, Pakistan

ABSTRACT

Scalability covers the ability of an application to cope up with heavier processing load as desired. Other than that, it also covers how small one can start a system. Testing scalability is a very challenging task among all the other quality attributes i.e. performance, reliability, usability, accessibility and security. However empirically there is not much research work done on this specific domain. The focus of this study is to identify the all the metrics and tools that are in practice for measuring the scalability of a specific system. Moreover, finding out the challenges that are being faced while performing these measurement practices.

To collect information regarding the identification of metrics and tools for testing scalability, a systematic mapping study was conducted. Through this method we tried to get a generalized data from many resources over the internet.

KEYWORDS

Scalability, Metrics, Tools, Testing, Systematic Mapping Study

1. INTRODUCTION

Scalability plays a very vital role in the design of parallel and distributed systems. Generally, scalability is the measure of a system that keeps its performance when its size is scaled up. In proper terms it is defined as “Scalability is a desirable property of a system, a network or a process which indicates its ability to either handle growing amounts of work in a graceful manner or to be enlarged” (Wikipedia).

In other terms “Scalability is a characteristic of a system, model or function that describes its capability to cope and perform under an increased or expanding workload. A system that scales well will be able to maintain or even increase its level of performance or efficiency when tested by larger operational demands.” (Investopedia). Testing scalability is a very challenging task among all the other quality attributes i.e. performance, reliability, usability, accessibility and security. In research domain several research work has been carried out and a number of metrics and tools are proposed so far for measuring scalability. However, there is still lack of an effective mapping that can help the testers to find out which are the best metrics and tools so far which can be used for measuring scalability to get better results.

2. LITERATURE REVIEW

(Serdar Dogan, 2014) tried to identify the tools and metrics available for testing different applications specifically web applications. He also identified these metrics based on the criteria of cost effectiveness. (Xiaokai Xia, 2010) focused on explaining some of the metrics and tools considered while testing. He mainly proposed to analyze and evaluate different attributes of web applications. He proposed in his model the issues that left unidentified during the phase of testing. (Gupta., 2015.) worked mainly on addressing the tools, types and methods used for testing different kinds of applications. In his research he also focused on explaining some metrics related to performance. (J. Križanic, 2010) attempted to conduct a systematic study to evaluate the testing tool Grinder. Other than that, he also finds out some of the other tools related to performance testing. (Wang., 2010) worked on a framework by integrating the basic functionalities of the tools like Selenium and JMeter together. This work concentrates mainly on different sort of testing Backend

testing, load testing and UI testing etc. (N, 2009) worked on comparing web application for two different technologies .NET and Java by using different testing tools and metrics. In their work both the technologies are evaluated in terms of response time and throughput. They find out in their research that the Java technology performs better than that of the .Net in web applications. (Arora., 2015) performed a case study on different automated tools available for testing performance and other attributes. As a part of their research they also focused on literary survey and find out 18 different automated tools. (Vahid Garousi, 2013) performed a systematic mapping study to find out the tools available for web testing. They have also identified few tools along with the factor availability of the tool. (Rina., A Comparative Study of Performance Testing Tools. , 2013) worked on comparison of some of the testing tools in terms of some metrics available. They have also evaluated in their research work the testing tools by conducting experiments to make selection for suitable tools for conducting testing.

The metrics used for testing are not fixed for a application, but they vary from one organization to the other. Some organizations are small while others are large. These organizations face severe problems in selecting metrics for their testing activities due to constrains they have regarding resources like cost and time etc. There is a very little amount of research work done in this specific domain i.e. the selection of metric in different organizations by the software testing team (Tanjila Kanij, Performance assessment metrics for software testers, 2012). As in literature these drawbacks are also not particularly provided. Also, the new set of features that the testers want to have in their tools. All these identified issues represent the gap in the research work so far, hence providing basis and motivation for our problem domain. We in our work focused on a systematic mapping of the metrics and tools available and tested so far for measurement of scalability by the testers.

3. METHODOLOGY

A systematic mapping study is conducted to identify the metrics and tools available for the measurement of scalability. A systematic mapping is a type of study in which an overview of the topic is provided to the researchers for their research study. It also gives the research the number of publications available in their specific research area. As in our case we used systematic mapping study to get the data regarding the tools and metrics for the measurement of scalability. To conduct systematic mapping study the guidelines provided by (Kai Petersen, 2008) are followed.

3.1 Systematic Mapping Study Protocol

To conduct systematic study a protocol is designed and made selection of the key words related to our domain:

3.1.1 Selection of Keywords

Table 1. Keywords Selection

Category ID	Category	Keywords
C1	Quality Attributes	Scalability, Scalable, Scale up, Scale down.
C2	Testing	Verify, verification, validation, validate, testing
C3	Tool	Tool, tools, framework
C4	Metric	Metric, Metrics, Measure, Measures, evaluation, evaluate.

3.2 Formation of the Search String

To form search string, we have used 2 Boolean operators, i.e. AND & OR operators. The search string is formed as:

((Scalability) OR (scalable)) AND ((Assess) OR (Assessment) OR (Evaluate) OR (Evaluating) OR (Measure) OR (Measurement) OR (Metric) OR (Metrics)) AND ((Quality) OR (Attributes))

After having a search string, we have further two important steps.

3.2.1 Selection of Scientific Library or Database

We have selected 5 scientific libraries over the web. These includes: INSPEC, IEEE, SCOPUS, ACM, WILEY.

3.2.2 Study Selection Criteria

Literature regarding scalability, metrics and tools. Literature that is in English language Literature from year 2000 to 2017.

3.3 Execution of the Search String

The search string is then executed in all the 5 databases mentioned above keeping in view the selection criteria given. All the results that are duplicated or are irrelevant to our selection criteria is excluded from the final gathered results. The following table shows our findings

Table 2. Results after String Execution

Database	Tools		Metrics	
	Total Searched Articles	Selected Articles	Total Searched Articles	Selected Articles
INSPEC	133	43	120	35
IEEE	140	47	135	32
SCOPUS	97	29	87	34
ACM	66	37	59	10
WILEY	40	2	92	5

In total 969 articles were obtained from all the databases. In which only 274 were selected keeping in view the selection criterion. Similarly, some of the articles were part of more than one database which were also merged to get the selected total.

3.4 Selected Literature for our Study

After studying the abstract of the 274 articles we have selected only 25 articles which were more specific and relevant to our study. During screening of the articles, the relevancy of the articles to our problem is thoroughly examined.

3.5 Studying and Assessing the Articles

The selected 25 articles are carefully studied to identify the metrics and tools that are used to measure the scalability of different applications. In all these articles important points are highlighted for further analysis.

3.6 Data Extraction Strategy

From all the highlighted points' data is extracted in the form of a table. This table gives the overview of our systematic study.

Table 3. Extracted Data

Author Name	Quality Attr.	Tools	Metrics
Fei Wang, Wencai Du (Du., 2012)	Scalability	Provided	Not Provided
Osama Hamed, Nedat Kafri (N, 2009)	Scalability	Provided	Provided
Qinglin Wu, Yan Wang (Wang., 2010)	Scalability	Provided	Provided
Zao-Bin GAN, Varadharajan (Varadharajan., 2005)	Scalability	Provided	Provided
G. Ruffo, R. Schifanella, & M. Sereno (G. Ruffo, 2004)	Scalability	Provided	Provided
Jordi Guitart, Vicenç Beltran, (Jordi Guitart, 2005)	Scalability	Provided	Provided
Tiantian Gao, Gongxin Wu (Wu Gongxin Jinlong Gao, 2010)	Scalability	Provided	Provided
Lakshmi S.Iyer, B. Gupta, N. Johri (Gupta., 2015.)	Scalability	Not Provided	Not Provided
Niclas Snellman , Adnan Ashraf(Niclas Snellman, 2011)	Scalability	Provided	Provided
R. Aganwal, B. Ghosh, S. Banerjee (R. Aganwal, 2000)	Scalability	Provided	Provided
Sandhya,Kiran,AkshyansunMohapatra (S. Kiran, 2015)	Scalability	Provided	Provided
Jörg Zinke, Jan Habenschuss, (J Zinke, 2012)	Scalability	Provided	Provided
B.M.Subraya, C.Ravi (B. M. Subraya, 2001)	Scalability	Provided	Provided
M. Kalita, S. Khanikar, T. Bezboruah (M. Kalita, 2011)	Scalability	Provided	Provided
Shraddha Mungekar, Dhanashree (taas., 2015)	Scalability	Provided	Provided
Arabi Keshk, Amal Ibrahim (Ibrahim, 2007)	Scalability	Provided	Provided
Hidam Kumarjit, Tulshi Bezboruah (Bezboruah., 2015)	Scalability	Provided	Provided
Manish Rajendra Dhote, G.G. Sarate (Sarate., 2013)	Scalability	Not Provided	Provided
Ana Cavalli, Gerardo Morales (Ana Cavalli, 2007)	Scalability	Provided	Provided
Hashemian, Martin Arlitt (Arlitt Martin Hashemian, 2012)	Scalability	Provided	Provided
Kai Lei1, Yining Ma, Zhi Tan (Kai Lei, 2014)	Scalability	Provided	Provided
Martti Vasar, Marlon Dumas (Martti Vasar, 2012)	Scalability	Provided	Provided
Breno Lisi, Romano (Breno Lisi Romano, 2009)	Scalability	Provided	Provided
Kalita1, T. Bezboruah (M. Kalita, 2011)	Scalability	Provided	Provided
Sebastian Lebrig (Sebastian Lebrig, 2015)	Scalability	Provided	Not Provided

4. RESULTS AND DISCUSSION

After we are done with all the steps of our systematic mapping study we resulted in two tables. One for the determined metrics for measurement of scalability in testing an application while the other for the tools for measuring scalability. In both tables not only, the names are listed but also their frequency count in the articles taken under consideration.

Table 4. Results for Scalability Metric

Scalability Metric	Frequency Count	Scalability Tool	Frequency Count
Response time	22	Tsung	5
Throughput	20	WebLOAD	5
Number of concurrent users	18	Silk performer	4
CPU utilization	16	QA load	3
Number of hits per second	14	CloudTest	3
Memory utilization	13	TestComplete	3
Disk I/O (access)	8	MBPeT	2
Latency	8	Pylot	2
Number of connections per second (user)	6	Testmaker	2
Disk queue length (request)	5	Application Center Test	1
Number of transactions per second (http)	4	Selenium WebDriver	1
Disk space	1	BrowserMob	1
CPU model	1	Zabbix	1
CPU clock	1	Nagios	1

5. CONCLUSION

This study was conducted to identify all the metrics and tools available and are in practice by the testers for the measurement of scalability of a system. We accomplished our objective by following steps of systematic mapping study. To have the required information for making our task done we have used 5 scientific libraries databases and through these databases we were able to get our desired articles that resulted in the final analysis and extraction of data.

Based on these results obtained we identified all the metrics that are available and are in practice for measuring scalability. We observed that the metrics response time and throughput are most commonly in practice. We also observed that there are very limited number of tools available for testing scalability as compared to other quality attributes like performance and reliability. Though we succeeded in refining a few of them that are successfully in practice, i.e. Tsung, Webloader and Silk Performer. Through this study we were successfully able to generate 2 lists, one for tools and the other for metrics. These lists can be helpful for the software testers to gain knowledge regarding testing scalability of different applications.

REFERENCES

- Ana Cavalli, S. M. (2007). Regression and performance testing of an e-learning web application. *Dotlrn Proceedings-International Conference on Signal Image Technologies and Internet Based Systems, SITIS*.
- Arlitt Martin Hashemian, K. (2012). Overcoming web server benchmarking challenges in the multi-core era. *Proceedings - IEEE 5th International Conference on Software Testing, Verification and Validation*.
- Arora., I. (2015). A Brief Survey on Web Application Performance Testing Tools Literature Review. *International Journal of Latest Trends in Engineering and Technology*.
- B. M. Subraya, S. V. (2001). A new model to bridge the gap between user and designer perceptions. *25 th Annual International Confernece on Computer Software and Applications, COMPSAC*.
- Bezboruah., H. K. (2015). Performance metrics of a customized web application developed for monitoring sensor data. *IEEE 2nd International Conference on Recent Trends in Information Systems (ReTIS)*.
- Breno Lisi Romano, G. B. (2009). Software testing for web applications for non-functional requirements. *6th Intenational Conference on Information Technology ITNG*.
- Du., J. F. (2012). A test automation framework based on WEB. *IEEE/ACIS 11th International conference on Computer and Information Science, ICIS*.

- G. Ruffo, R. S. (2004). A tool for evaluating web application performance in Quantitative Evaluation of Systems. *First International Conference QEST 2004*.
- Gupta., D. D. (2015.). Performance testing and evaluation of web applications using wapt pro. . *International Journal of Innovative Research in Computer and Communication Engineering*.
- H.M.a Aguiar, J. S. (2010). Profiling of real-world web applications. *International Workshop on Parallel and Distributed Systems: Testing, Analysis, and Debugging*.
- Ibrahim, A. K. (2007). Ensuring the Quality Testing of Web Using a New Methodology. *IEEE International Symposium on Signal Processing and Information technology*.
- J Zinke, J. H. (2012). Generating representative workloads for web server benchmarking. *Simulation Series*, 44.
- J. Križanec, A. G. (2010). Load testing and performance monitoring tools in use with ajax based web applications. *MIPRO*.
- Jordi Guitart, V. B. (2005). Characterizing secure dynamic web applications scalability. *19th IEEE International Parallel and Distributed Processing Symposium*.
- Kai Lei, Y. M. (2014). Performance comparison and evaluation of web development technologies in php, python and node. js. *IEEE 17th International Conference on Computational Science and Engineering (CSE)*.
- Kai Petersen, R. F. (2008). Systematic mapping studies in software engineering. *12th international conference on evaluation and assessment in software engineering*.
- M. Kalita, S. K. (2011). Investigation on performance testing and evaluation of PReWebN: A Java Technique for Implementing Web Application. . *IET*.
- Martti Vasar, S. N. (2012). Framework for monitoring and testing web application scalability on the cloud. *Proceedings of the WICSA/ECSA*.
- N, H. O. (2009). Performance testing for web based application architectures (.NET vs. Java EE). *1st International Conference on Networked Digital Technologies, NDT 2009*.
- Niclas Snellman, A. A. (2011). Towards Automatic Performance and Scalability Testing of Rich Internet Applications in the Cloud. *37th EUROMICRO Conference on Software Engineering and Advanced Applications*.
- R. Aganwal, B. G. (2000). Ensuring website quality: a case study. In *Management of Innovation and Technology. IEEE International Conference on ICMIT*.
- Rina., T. (2013). A Comparative Study of Performance Testing Tools. *International Journal of Advanced Research in Computer Science and Software Engineering*.
- S. Kiran, A. M. (2015). Experiences in performance testing of web applications with unified authentication platform using jmeter. *Technology Management and Emerging Technologies (ISTMET), An International Symposium*.
- Sarate., M. R. (2013). Performance testing complexity analysis on ajax-based web applications. *IEEE Conference*.
- Sebastian Lehrig, H. E. (2015). Scalability, Elasticity, and Efficiency in Cloud Computing: A Systematic Literature Review of Definitions and Metrics. *Proceedings of the 11th International ACM SIGSOFT Conference on Quality of Software Architectures*.
- Serdar Dogan, A. B.-C. (2014). Web application testing: A systematic literature review. *Journal of Systems and Software*.
- taas., S. M. (2015). An architecture of website analysis in a cloud environment. *1st International Conference on Next Generation Technologies (NGCT)*.
- Tanjila Kani, R. M. (2012). Performance assessment metrics for software testers. *5th International Workshop on Cooperative and Human Aspects of Software Engineering, CHASE*.
- Vahid Garousi, A. M. (2013). A systematic mapping study of web application testing. *international journal of Information and Software Technology*.
- Varadharajan., V. (2005). Evaluating the Performance and Scalability of Web Application Systems. *Third International Conference on Information and Technology and Applications*.
- Wang., Q. W. (2010). Performance testing and optimization of J2EEbased web applications. *2nd International Workshop on Education Technology and Computer Science*.
- Wu Gongxin Jinlong Gao, T. (2010). A Reactivity-based Framework of Automated Performance Testing for Web Applications. *Ninth International Symposium on Distributed Computing and Applications to Business, Engineering and Science*.
- Xiaokai Xia, Q. P. (2010). Multi-level logs based web performance evaluation and analysis. *International Conference on Computer Application and System Modeling, Proceedings, 4(Iccasm)*.