Evaluating the Reliability of PwCOV: A Loosely Coupled Software as a Service for COVID-19 Data Processing System

A Bora*, T. Bezboruah

Department of Electronics & Communication Technology, Gauhati University, Guwahati, Assam, India

1. INTRODUCTION

The worldwide pandemic issues of coronavirus diseases 2019 (COVID-19) have created a typical vision of life for human beings and societies across the globe. The experiences of symptoms with different side effects in the human body have broadened the history of COVID-19 [1]. As such, the pharmaceutical and clinical entities have become the key player for managing and handling emergency issues of health management cares [2]. According to Cortegiani et al. (2020), the clinical test cases of COVID-19 are generating regional-based different data sets based on blood samples, symptoms, population, histopathology, clinical images, and affected organs [3]. The clinical test reports of drug repositioning, antibody agents, an antimalarial drug, chemical structure, and the effect of hydroxyl chloroquine are observed to be important resources for the development of vaccines and antibody injections [4–6]. Dhouha et al. [7] had elaborated on the image data set for clinical investigation over diseases. They had discussed the feasibility of deploying a multi-tenant software environment for processing such data sets. However, in the medical industry, the deployment of software as a service (SaaS) for generating clinical reports is getting popularity [8]. With the increase in users over web-based systems, the execution of a multi-tenant environment for disease data set and observing its service reliability has become an important measure. In the health care units, the integration of data set from different geographical locations and observing the reliability of the report has also become a crucial factor for medical industries. The core problem of SaaS during COVID-19 can be categorized as follows: (a) Finding the reliability of the service, (b) Evaluating the operational limit during high usage of the service, (c) Assessing the reliability beyond the capacity of the service, and (d) Observing the nature of the distribution of service failure. As such, a proper...
methodology is required to be developed to evaluate the reliability of service, so that the medical units can obtain an assessment of SaaS before their projection in the society. In this work, a novel SaaS oriented model as a prototype web service (WS) for COVID-19 (PwCOV) is proposed for processing disease data set. A novel reliability assessment cycle for PwCOV is also discussed to observe the reliability of such deployment during high stresses of usage. The novelty of the PwCOV from the viewpoint of service-oriented computing (SOC) can also be observed in the way of deployment methodology where the roles of SOC are segregated among WSs for consumer, parent, and service layer. This work also emphasizes the reliability evaluation of the PwCOV while executing through load balancing clustered (LBC) webserver against the massive growth of users.

2. RELATED WORK

Many authors have discussed the importance of processing the COVID-19 dataset and the role of software agents for the clinical domain [3, 7]. In the year 2017, Medhi et al. [9] had discussed a novel model to evaluate the data set processing system through the paradigm of SOC. In the year 2019, Azab et al. [10] had emphasized the study of clinical features for muscles, eye and ear, oral cavity, upper and lower airway, gut, cardiovascular, nervous system, urinary tract, skeleton, and skin. The properties of such a dataset are elaborated through a data model. Lai et al. [11] had introduced an index parameter over the cumulative measure of COVID-19 impact in the societies. An infection control management technique was introduced that can be followed to restrict the spread of the diseases. In the year 2020, Benjamin et al. had introduced a novel software agent called ‘ChatBot’ for processing interactive medical data set [8]. In the same year, Jorge et al. had introduced a data model that can be followed to evaluate the probability of the spread of the COVID-19 in native communities and societies [12]. Juan et al. [13] emphasized the deployment methodology and tool while considering the importance of mechanical ventilation management, biosafety measures, patient route management, management strategies, and patient wise ventilation mode. A novel model was introduced to implement a decision-making system based on available data reports. Gilmiyarova et al. [14] had proposed a diagnostic tool that can be deployed to evaluate the clinical remarks of oral fluid. They had introduced an empirical model to study various characteristics of data set over blood groups. In the same year, Lee et al. had carried out a study over the test case results and outcomes of treatment for COVID-19 patients [15]. The study highlighted the pros and cons of the applied treatment measure for COVID-19 infected patients. Bora and Bezboruah [16] had proposed a WS oriented model for processing clinical service. They proposed the importance of reliability and performance metrics for implementing WS for medical industries. In the same year, Lai et al. [17] had evaluated a statistical model to study the daily cumulative index for COVID-19 in different regions of the world. The analysis was carried out by preparing a data set of cumulative cases, incidence per population, number of death, death per population, mortality rate, local transmission, and days since last reported. In the research community of biomedical engineering units, medical practitioners, and pharma industries, the processing and evaluating the geographically separated data model is getting popularity, as it can provide better outcomes for the collaborative fight against COVID-19 [12]. Many authors have emphasized the importance of deploying SaaS for medical services. During the outbreak of COVID-19, the clinical remarks and instruction based on different patients asymptomatic and symptomatic are generated at different locations. As such, the clinical experiments are generating medical data set for better medical service delivery to society. In the domain of COVID-19, the delivery of clinical remarks over the web even during the high usage of the service has become the primary role. As such, the deployment of SaaS with LBC based web server can enhance the reliability of the service. With the rapid growth of medical infrastructure, the deployment of SaaS and evaluating its reliability against massive users has become a critical call for processing the COVID-19 data set. Hence, deployment of a collective support system and generation of clinical remarks among health care units and service consumers of the COVID-19 management system has become a demanding concern. This work proposes a novel reliability assessment methodology for the deployment of SaaS in LBC based web server while processing clinical remarks against the asymptomatic and symptomatic nature of COVID-19.

3. MATERIALS AND METHOD

The features of SaaS are deployed through the paradigm of service-oriented architecture (SOA) [18, 19]. The programmable interface can be developed for establishing tightly and loosely coupled SaaS modules [20, 21]. The SOA provides different communication hierarchies through WSs [22, 23]. Figure 1 shows the architecture of SaaS for the deployment of PwCOV. The hardware and software configuration for deployment is discussed elsewhere [24]. The load balancing web server is clustered into two working nodes. Each working node contains the module of PwCOV architecture. The PwCOV consists of three WS. They are: (a) consumer layer, (b) parent layer and (c) service layer, respectively for functional execution of business logic (BL). It also contains the database layer for the clinical mapping of the
disease data set. The consumer layer contains a WS for handling the graphical user interface (GUI) of the PwCOV. The presentation code including server-side Java class files, Java Server Pages (JSP), and Javascript form controls are included in this layer. The objective of this layer is to capture end-user data and send it to the parent layer. The parent layer contains a WS for capturing, validating and forwarding the end user parameter as received from the consumer layer. The layer also responds back to the consumer layer of the query results of the service layer. The service layer contains a WS for executing functions and querying the database layer. The primary role of this layer is the establishment of the database connection, generating results set, and sending back the report to the parent layer. The database layer contains the database engine for clinical instructions of the symptoms as per COVID-19. The tenant entity is the set of system generated users that can invoke PwCOV through the cluster-based load balancing web server. The load balancing of the webserver is configured as discussed elsewhere [25]. The necessary system variables are configured to support the simultaneous execution of the working node. As such, the set of tenant will get a response of PwCOV even during the failure or high load in a particular working node. Figure 2 shows the cycle of the reliability evaluation model for evaluation of SaaS execution for PwCOV. It contains 8 steps. They are: (a) Deployment of PwCOV, (b) Set test case, (c) Environment value, (d) Execute system, (e) Observe failure rate data sample, (f) Identify the distribution of data sample, (g) Evaluating reliability, and (h) Overall assessment. In step (a), the SaaS for processing COVID-19 disease data is developed and deployed in LBC based web server after functional testing. In step (b), the test case is created. Here, the test case is set by using the Mercury LoadRunner load testing tool [26]. The test case contains the necessary path of the hosted SaaS along with instruction for fetching records through PwCOV. In step (c), the environmental value is assigned. It contains instructions for the ramp-up and ramps down of all tenants that executes the SaaS under the same testbed environment. The ramp down parameter is set to release the load gradually from the SaaS. In step (d), the system is executed. Here, the load of the tenant is generated gradually. Once, all the tenants enter the system, the SaaS is executed simultaneously by the set of enants that are active. The system is executed for a specific turnaround time. In this case, the turnaround time is set to be 15 minutes. In step (e), the failure rate data sample is observed. Here, each run of the test case against each set of tenants is observed. If failure records are observed, a data sample of 30 repetitive test is collected for further study of the data points. In step (f), the distribution of the recorded data points is identified. This step helps to identify the validation and analysis of the recorded data sample. The data sample is recorded for each turnaround system execution. The PwCOV is monitored for failure records. In step (g), the reliability metric is evaluated. The reliability metric is evaluated for the observed failure records against the stress of a set of tenants. In step (h), the overall assessment of the study is carried out. This step contains the interpretation of the data sample and concluding the overall assessment.

4. EXPERIMENTAL RESULTS

In this study, the test case is set for fetching data from the clinical data set. The test case is prepared as discussed elsewhere [27]. The PwCOV is executed for the set of tenant 50, 100, 200, 400, 600, 800, 1000, 1200, 1400, 1700 and 1800. Table 1 shows the observed records of SaaS transaction pass, fail and failure rate for the execution of PwCOV. Many studies described that “With an increase in users the connection refusal increases. This behavior is due to the garbage collected heap in the server-side” [9, 24, 27]. In this work, from Table 1, it is
observed that the SaaS along with LBC based web server can manage the heap up to 1700. The role of the heap at server-side is to manage the correlation of Java class, objects, and member variables for each Hyper Text Transfer Protocol (HTTP) request. However, beyond this capacity limit, the heap error occurs, due to which the failure of HTTP processing generates. At the set of 1800 tenants, the SaaS transaction failure record of 40002 is observed out of 123800 HTTP requests. To further study the failure distribution, the data sample of 30 repeated executions is generated. However, to study the reliability metric (RM), the recorded failure record is evaluated through Equation (1) \[28, 29\]. The RM value is evaluated over a service time \( t \) and failure rate \( f \) \[30, 31\]. In this study, \( t \) is set to be day one as each data sample record is collected over 30 calendar days.

\[
RM = e^{-ft}
\]  

(1)

The RM is evaluated to be in the range of 0 to 1. The RM estimated value nearer to 1 reveals the strong reliability of the system. Otherwise, moderate reliable service can be observed. To study the distribution of the recorded failure rate and RM, the interpretation of histogram and normal probability plot (NPP) is followed \[31\]. The SaaS is stable up to 1700 tenants. That means the SaaS will process the HTTP request properly without any failure. However, the SaaS is generating a failure rate for the recorded HTTP request of 1800 tenants. To further study the failure record, the distribution of the data is evaluated through a histogram. Figure 3 and Figure 4 show the histogram of the data sample of 30 failure rate and RM against 1800 tenants. In Figure 3, it is observed that 1 data count is in the range of 0 to <=30%, 6 data count is in the range of >30% to <=32%, 11 data count is in the range of >34% to <=37%, 2 data count is in the range of >37% to <=39% and 3 data count is in the range of >39%, respectively for failure rate. In Figure 4, it is observed that 1 data count is in the range of 0 to <=0.65, 2 data count is in the range of >0.65 to <=0.67, 3 data count is in the range of >0.67 to <=0.69, 6 data count is in the range of >0.69 to <=0.71, 11 data count is in the range of >0.71 to <=0.72% and 7 data count is in the range of >0.72%, respectively for RM. Here, for each case, a single peak value is observed. For the failure rate data sample, the highest recorded data points lie within the range of >0.32 to <= 0.34. For the RM data sample, the highest recorded data points lies within the range of >0.71 to <= 0.72. Figure 3 concludes normal distribution and Figure 4 concludes left-skewed distribution. However, based on the data range, we may evaluate different observations. As such, we evaluate further through NPP. The NPP states the normality of observed data points for each data sample of failure rate and RM. Figure 5 and Figure 6 show the NPP of the data points recorded for failure rate and RM, respectively. The data samples are linear and following the mean of the recorded parameters.

### TABLE 1. SaaS transaction records for PwCOV against different stress of tenant sets

| Set of tenant for PwCOV | SaaS transaction pass for PwCOV | SaaS transaction fail for PwCOV | SaaS transaction failure rate for PwCOV |
|-------------------------|---------------------------------|-------------------------------|----------------------------------------|
| 50                      | 235                             | 0                             | 0                                      |
| 100                     | 405                             | 0                             | 0                                      |
| 200                     | 819                             | 0                             | 0                                      |
| 400                     | 16152                           | 0                             | 0                                      |
| 600                     | 18974                           | 0                             | 0                                      |
| 800                     | 23547                           | 0                             | 0                                      |
| 1000                    | 26785                           | 0                             | 0                                      |
| 1200                    | 36251                           | 0                             | 0                                      |
| 1400                    | 41054                           | 0                             | 0                                      |
| 1700                    | 62982                           | 0                             | 0                                      |
| 1800                    | 83798                           | 0                             | 0                                      |
|                         |                                  |                               |                                         |

>34% to <=37%, 2 data count is in the range of >37% to <=39% and 3 data count is in the range of >39%, respectively for failure rate. In Figure 4, it is observed that 1 data count is in the range of 0 to <=0.65, 2 data count is in the range of >0.65 to <=0.67, 3 data count is in the range of >0.67 to <=0.69, 6 data count is in the range of >0.69 to <=0.71, 11 data count is in the range of >0.71 to <=0.72% and 7 data count is in the range of >0.72%, respectively for RM. Here, for each case, a single peak value is observed. For the failure rate data sample, the highest recorded data points lie within the range of >0.32 to <= 0.34. For the RM data sample, the highest recorded data points lies within the range of >0.71 to <= 0.72. Figure 3 concludes normal distribution and Figure 4 concludes left-skewed distribution. However, based on the data range, we may evaluate different observations. As such, we evaluate further through NPP. The NPP states the normality of observed data points for each data sample of failure rate and RM. Figure 5 and Figure 6 show the NPP of the data points recorded for failure rate and RM, respectively. The data samples are linear and following the mean of the recorded parameters.
PwCOV is introduced that can be executed for different stress of usages. The contribution of the work highlights the reliability evaluation of SaaS while processing clinical remarks of COVID-19 against the massive growth of multi-tenant set in LBC based web server. The work emphasizes the deployment of SaaS through the segregation of SOC roles among WSs. The reliability of SaaS is recorded to be strong while executing PwCOV up to 1700 sets of tenants. Beyond that capacity, the PwCOV is generating failure records. As such, the reliability degrades up to a range of 0.71 to 0.72. For recorded stress of 1800 tenants set, the reliability lies within 71% to 72%. The applicability of the cycle of the reliability evaluation model for SaaS with PwCOV is observed. The proposed SaaS architecture for the execution of PwCOV can deliver reliable service for processing COVID-19 disease data sets. The scalability of SaaS can be achieved for the massive growth of a set of tenants. The study can help the medical industries, software practitioners, and other clinical entities to gain an in-depth idea about reliability for the deployment of SaaS and PwCOV for processing COVID-19 data sets.

7. FUTURE PROSPECTS

The development of a model for the performance and scalability study of SaaS while processing the COVID-19 data set can mimic a scenario for any size of consumers. Undoubtedly, future work will focus on optimizing the performance aspects of SaaS for COVID-19, as it can provide effective support for the treatment delivery and reliable service to the society.

8. REFERENCES

1. Leite, H., Hodgkinson, I. R., and Gruber, T. “New development: ‘Healing at a distance’—telemedicine and COVID-19.” Public Money and Management. Vol. 40, No. 6, (2020), 483–485. https://doi.org/10.1080/09540962.2020.1748855
2. Wang, L., Wang, Y., Ye, D., and Liu, Q. “Review of the 2019 novel coronavirus (SARS-CoV-2) based on current evidence.” International Journal of Antimicrobial Agents. Vol. 55, No. 6, (2020), 105948. https://doi.org/10.1016/j.ijantimicag.2020.105948
3. Cortegiani, A., Ingoglia, G., Ippolito, M., Giarratano, A., and Einav, S. “A systematic review on the efficacy and safety of chloroquine for the treatment of COVID-19.” Journal of Critical Care. Vol. 57, (2020), 279–283. https://doi.org/10.1016/j.jcrc.2020.03.005
4. Mullard, A. "Drug repurposing programmes get lift off." Nature Reviews Drug Discovery. Vol. 11, (2012), 505–506. https://doi.org/https://doi.org/10.1038/nrd3776
5. Zhou, P., Yang, X. Lou, Wang, X. G., Hu, B., Zhang, L., Zhang, W., Si, H. R., Zhu, Y., Li, B., Huang, C. L., Shi, Z. L. “A pneumonia outbreak associated with a new coronavirus of probable bat origin.” Nature. Vol. 579, No. 7798, (2020), 270–273. https://doi.org/10.1038/s41586-020-2012-7
6. Polz-Dacewicz, M. “Novel coronavirus – SARS CoV-2.” Polish
چکیده
بیماری همه گیر ویروس کرونا 2019 در مناطق مختلف جهان مجموعه داده‌های مختلفی تولید می‌کند. مجموعه داده‌ها مشاهده می‌شود که در نهادهای پزشکی استخراج شده از نظر جغرافیایی موجود است. با این حال، نقاشی برای دسترسی و تحلیل مطمئن این مجموعه داده‌ها از طریق یک مزرعه تحت وب به تدریج در حال افزایش است. در این پژوهش، یک چرخه جدید از مدل ارزیابی قابلیت اطمینان برای استقرار نرم‌افزار در ناحیه نیروگاه‌ها ارائه می‌شود که در این مقاله PwCOV نامیده می‌گردد. نمونه اولیه اطمینان از طریق الگوی محاسبات سرویس گرای و مدیریت بار خوشه‌ای و اصول PwCOV برای پردازش مجموعه داده‌های بیماری‌های COVID-19 در برابر فشارهای مختلف مجموعه موجودات کاربر مورد بحث قرار گرفته است. اعتبار و کاربرد مدل در تجزیه و تحلیل آماری ارزیابی می‌شود. قابلیت اطمینان PwCOV با ارزیابی وضعیت مباشته معنی‌دار و کاربردی است. این مطالعه نشان می‌دهد که PwCOV برای پردازش نیروگاه‌های بیماری COVID-19 با استفاده از یک وب سرویس پایه خوشه‌ای مشترک Multi-Tenant نیز مورد بحث قرار گرفته است.