A Usage Analytics Model for Analysing User Behaviour in IBM Academic Cloud

Manoj Kesavulu, Duc-Tien Dang-Nguyen, Marija Bezbradica and Markus Helfert

Lero - Irish Software Research Institute, Dublin City University, Dublin, Ireland
manoj.kesavulu2@mail.dcu.ie, duc-tien.dang-nguyen@dcu.ie, marija.bezbradica@dcu.ie, markus.helfert@dcu.ie

Abstract: Usage in the software domain refers to the knowledge about how end-users use the application and how the application responds to the user's actions. Usage can be revealed by monitoring the user's interaction with the application. However, in the cloud environment, it is non-trivial to understand the interactions of the users by using only the monitoring solutions. For example, user's behaviour, user's usage pattern, which features of a cloud application are critical for a user, to name a few, cannot be extracted using only the existing monitoring tools. Understanding these information require additional analysis, which can be done by using usage analytics. For this purpose, in this paper, we propose a novel process model design for incorporating Usage Analytics in a cloud environment. We evaluate this proposed process model in the context of academic applications and services in the cloud, with the focus on IBM Academic Cloud.

Keywords: Usage, Analytics, IBM Academic Cloud, Process Model, End-user, Application, Cloud, Agile, Software

1 Introduction

Cloud computing is an emerging paradigm, bringing many advantages to both users and services providers while making dynamic platforms, cost-effective, flexible, on-demand resource provisioning, and many others. Consequently, there are significantly increasing developments for cloud computing infrastructures and platforms, from large high-tech enterprises such as Microsoft (with Microsoft Azure), Amazon (with Amazon EC2), and IBM (with IBM Cloud), to name a few. For example, IBM Cloud has seen around 70% increase in popularity in the last decade, becoming the first enterprise-tech vendor to surpass $15 billion in cloud revenue annual, according to Forbes (Technology column, July 28, 2017). Coming together with the benefits, unfortunately, is the increasing difficulties in understanding the way different users using the applications and the provided resources for the cloud services. Normally, cloud providers and users use several monitoring and analytics tools to ensure the cloud services and the applications deployed on them execute as intended. Traditionally, the cloud provider (vendor) provides application performance management

tools to monitor the status of the deployed applications. These tools provide mainly a vast amount of usage data of the resources used which can be turn into some knowledge for resource provisioning or error diagnostics in the cloud systems at the infrastructure and application level (Kesavulu et al. 2018). However, it is non-trivial to obtain user-related information, for example, user's behaviour, user's usage pattern, which features of the application are critical for a user, from the data extracted using existing APM tools. In order to analyse these information, advanced data analytics on extracted usage data is required (Dang-Nguyen et al. 2018). By using usage analytics we found that user behaviour, user's usage pattern can be understood by analysing user's interaction with the cloud applications. This helps the developers of the application to understand which features of the application are critical to users.

In this study, we focus the research fields to investigate on how to analyse user behaviour in cloud services and applications in the academic environment only. The reason of this focus is governed with the importance of the impact (positive/negative) of the behaviour of the students as well as lecturers, as discussed in (Chandra & Malaya 2012). For this reason, we decide to propose our study on IBM academic cloud a true cloud computing solution that has been optimized for the educational and research needs of the academic community (Rindos et al. 2014). The need for frequent updates, constantly changing requirements in the cloud environment encouraged us to adopt Agile model. In this paper we provide a model for Usage Analytics based on Agile model and show how the model can be implemented in IBM Academic Cloud while explaining the challenges involved and how these challenges could be addressed using Usage Analytics.

The contributions of this study are as follows: (i) propose a method that optimise the IBM Academic Cloud resources to best suit to lecturers and administrators to understand students' behaviour; and (ii) propose solutions for developers to understand which features of the cloud application are critical for users.

The remainder of this paper is followed by some related work in Sec. 2. We then introduce the usage analytics process model (Sec. 3) and discuss how to implement it in IBM Academic Cloud (Sec 4). Sec. 5 provides conclusions and some future directions.

2 Related Work

The usage data in the context of this paper are the data generated because of user interactions with the cloud application using which we can determine which features are critical and important for the end-user. Usage data can be classified into the following seven categories: (1) who is using the application; (2) where the application being hosted; (3) what the end user does; (4) when the user performs the operation; (5) how long it takes to complete the operation; (6) other operation details such as errors, background tasks; and (7) user behaviour (Kesavulu et al. 2017). Understanding usage data of an application has various uses such as to personalise the application according to the end-user's preferences (Yang et al. 2017), profiling users for security (Al-Bayati et al. 2016), improvement in marketing of software products (Bucklin & Sismeiro 2009) and so on. Many cloud monitoring solutions exist currently that extract usage data but focus on infrastructure level (De Chaves et al. 2011, Suneja et al. 2016, Montes et al. 2013) and application/service level (Zhao et al. 2015, Anithakumari & Chandrasekaran 2017). However, user-level usage data and its implications on the service and infrastructural resource usage in the cloud has not been deeply considered.

3 Usage Analytics Process Model

Applications deployed over cloud and the services provided by cloud have requirements that change and update frequently. In IBM Academic Cloud, lecturers are required to frequently update and handle changing requirements such as new tools, methods, techniques to improve the quality of course and modules offered to students. Agile development in opposition to traditional model is characterized by a focus on small teams of skilled individual developers, changing requirements, frequent version deliveries, daily contact with stakeholders makes it suitable candidate for cloud environments.

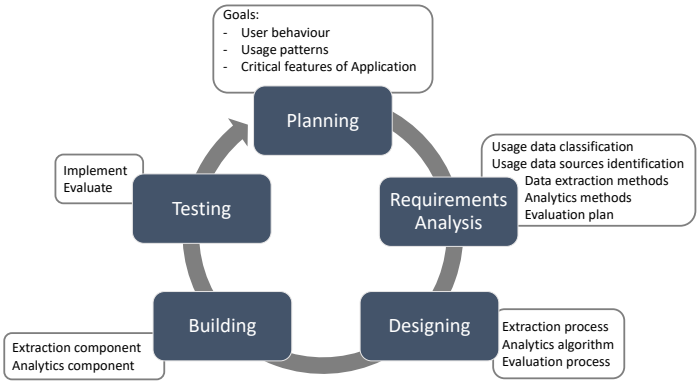

Figure 1 Usage Analytics Process Model

We propose the usage analytics process model (as shown in Figure 1) based on agile model that consists of the following five phases in each iteration:

- Planning: includes planning on the following goals of the usage analytics process model: *Understanding user's usage Pattern, Understanding user behaviour, Identifying Critical features of the application from the user's perspective*.
- Requirements Analysis: involves project managers and software architects dealing with elicitation of requirements such as *Usage data classification* - refer Sec. 2, *Usage data sources identification*, *Data extraction methods* or *Analytics methods*.
- Designing: includes the software architects dealing with the design of *usage data extraction process, Analytics algorithm and evaluation process*.
- Building: involves developers building the *usage data extraction and analytics component*.
- Testing: involves *implementing* the usage data extraction and analytics components and *evaluation*.

4 Implementations

IBM Academic Cloud provides virtual resources (VMs) on-demand basis for universities to help run classes and labs. Users (lecturers, academic administrators and students) access the VMs through RDP connection with any desktop, laptop with basic configuration. Lecturers and administrators can request and change the configurations of the VMs while students can only access the VMs provided to them.

We designed a scenario to incorporate Usage Analytics process model as shown in Figure 2. The first phase is to decide the goals, meaning what to achieve from analysing the user's usage. For example, user behaviour, user's usage pattern can be analysed and understood. This helps the developers of the application to understand which features of the application are critical to users. Consequently, lecturers and administrators can see how students use the applications hosted in VMs provided to them and how changes made to the VM configuration impact (positive/negative) the behaviour of the students. Also, best suited configurations for each student or group of students can be identified. The second phase is to analyse the requirements, for example, application logs can be used to reveal non-trivial information such as the event data, user ID, access time and so on. System logs reveal information such as workload, processing time, errors and so on. Cross matching information from both type of logs can help understand student's usage pattern, impact on student's behaviour (for example, deviation from typical interactions) as a result of configuration changes made by lecturer/administrator to the VMs and so on. Here, application and system log files can be treated as usage data sources.

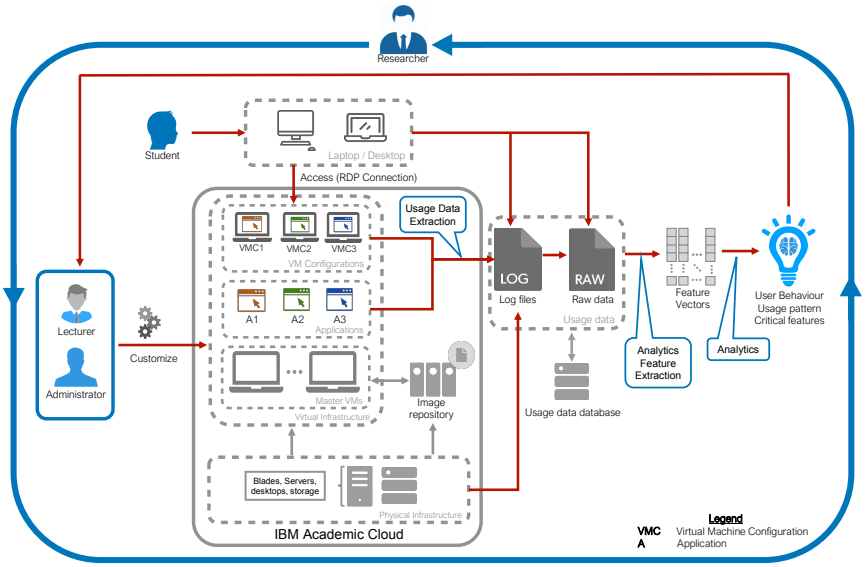

Figure 2 Usage Analytics in IBM Academic Cloud

The third and fourth phases deals with carefully examining and designing suitable usage data extraction process (e.g., log extraction and analysis) and tools (e.g., logstash, ELK stack), analytics methods (e.g., regular expressions, machine learning, statistical analysis). However, tools vary for different scenarios. The fifth stage deals with evaluation of the model. Questionnaires, surveys and interviews are suitable for the initial iterations to validate the extraction and analytics components designed and developed in the previous phases.

5 Conclusion and Future work

We presented a usage analytics model based on agile software model for cloud applications and presented an implementation plan for IBM Academic Cloud. We investigated the benefits of incorporating the usage analytics model with the aim to understand user's (student) behaviour which helps lecturers/administrators make an informed decision while

requesting/assigning resources for/to students. Our future work aims (i) to design and develop the usage analytics model to suit specific use cases; (ii) analyse existing tools and develop strategies for incorporating the model in IBM Academic Cloud.

Acknowledgement

This work was supported with the financial support of the Science Foundation Ireland grant 13/RC/2094 and co-funded under the European Regional Development Fund through the Southern & Eastern Regional Operational Programme to Lero - the Irish Software Research Centre (www.lero.ie).

We are thankful for Dr. Minh-Son Dao from Center for Innovative Engineering, Universiti Teknologi Brunei for the discussions and inputs on Cloud technologies.

References

Al-Bayati, B., Clarke, N. & Dowland, P. (2016), 'Adaptive behavioral profiling for identity verification in cloud computing: A model and preliminary analysis', *GSTF Journal on Computing (JoC)* **5**(1), 21.

Anithakumari, S. & Chandrasekaran, K. (2017), Negotiation and monitoring of service level agreements in cloud computing services, *in* 'Proceedings of the International Conference on Data Engineering and Communication Technology', Springer, pp. 651–659.

Bucklin, R. E. & Sismeiro, C. (2009), 'Click here for internet insight: Advances in clickstream data analysis in marketing', *Journal of Interactive Marketing* **23**(1), 35–48.

Chandra, D. G. & Malaya, D. B. (2012), Role of cloud computing in education, *in* 'Computing, Electronics and Electrical Technologies (ICCEET), 2012 International Conference on', IEEE, pp. 832–836.

Dang-Nguyen, D.-T., Kesavulu, M. & Helfert, M. (2018), Usage Analytics: Research Directions to Discover Insights from Cloud-based Applications, *in* 'International Conference on Smart Cities and Green ICT Systems (SMARTGREENS)'. Accepted.

De Chaves, S. A., Uriarte, R. B. & Westphall, C. B. (2011), 'Toward an architecture for monitoring private clouds', *IEEE Communications Magazine* **49**(12), 130–137.

Kesavulu, M., Dang-Nguyen, D.-T., Helfert, M. & Bezbradica, M. (2018), An Overview of User-level Usage Monitoring in Cloud Environment, *in* 'The UK Academy for Information Systems (UKAIS)'.

Kesavulu, M., Helfert, M. & Bezbradica, M. (2017), A Usage-based Data Extraction Framework for Cloud-Based Application - An Human-Computer Interaction approach, *in* 'International Conference on Computer-Human Interaction Research and Applications (CHIRA)', Madeira, Portugal.

Montes, J., Sánchez, A., Memishi, B., Pérez, M. S. & Antoniu, G. (2013), 'Gmone: A complete approach to cloud monitoring', *Future Generation Computer Systems* **29**(8), 2026–2040.

Rindos, A., Vouk, M. & Jararweh, Y. (2014), 'The virtual computing lab (vcl): an open source cloud computing solution designed specifically for education and research', *International Journal of Service Science, Management, Engineering, and Technology (IJSSMET)* **5**(2), 51–63.

Suneja, S., Isci, C., Koller, R. & de Lara, E. (2016), 'Touchless and always-on cloud analytics as a service', *IBM Journal of Research and Development* **60**(2-3), 11–1.

Yang, J., Wang, H., Lv, Z., Wei, W., Song, H., Erol-Kantarci, M., Kantarci, B. & He, S. (2017), 'Multimedia recommendation and transmission system based on cloud platform', *Future Generation Computer Systems* **70**, 94–103.

Zhao, L., Sakr, S. & Liu, A. (2015), 'A framework for consumer-centric sla management of cloud-hosted databases', *IEEE Transactions on Services Computing* **8**(4), 534–549.