

DevOps in support of Big Data Applications and Analytics

Juan Ni

Bloomington, Indiana 47401

nijuan@iu.edu

ABSTRACT

This paper discuss the relationship between devops and big data applications and analytic. The investigation will focus on cloud platform devops , because the carrier for most of big data application and analytic is cloud platform. Also, define the difficulty of implement devops is important for designing a big data application.

KEYWORDS

big data,devops, cloud platform

1 INTRODUCTION

Following with cloud computing and big data analytic prosperous, devops began getting more and more attention. In the traditional development group, we usually have six processes which are "Analyze", "Design", "Code", "Test", "Deploy", and "Maintain". And each of the traditional process is isolate with other, once a group finish their part, they will shift the work to next group for the next level processing. But once the client change their needs, it will take really long time to track back the work and making adjustment at traditional development team. Especially for big data application, due to the complex of big data application, the development period will be longer than usual, therefore, we need devops to help us satisfy our client needs at long period development and keep efficiency processing speed. Devops is not a set of tool, it is a methodology that include basic principle and practical. According to the devops concept, the processes of devops are "Code", "Build", "Test", "Package", "Release", "Configure", and "Monitor" [3]. Unlike the traditional development processes, all processes of devops are connecting inside a loop. So devops is the integration of Development and Operations,

2 THE NEED OF DEVOPS FOR BIG DATA APPLICATION

The gap of communication between development and operations on big data application is the main issue. In big data application and analytic project, analyst change their algorithm ceaselessly, and the change of analytically model will make the infrastructure and resource demeaned become much different with the original one. In the traditional development team, developers are not involve into the analyst activity, because the big data analyst and application developer is isolated. Once the developer get the changing requirement from analytic team, they need to take time to understand the adjustment and reorganized the manpower inside the group. This communication delay will lower the entire processing speed, decreasing the competitiveness of big data analysis. Big data is timeliness, if the processing take to much time, the outcome will be less value, so big data application need devops to prevent data losing value by fall behind.

3 THE VALUES OF DEVOPS

The main value of devops is to break down the "Wall of confusion" between developer and operator. According to Jerome's idea, devops have two main values which are "Continuous Delivery" and "Benefits". [2] The "Benefits" of devops include but not limit to "Repeatability and Reliability", "Productivity", "Time to recovery", "Guarantee that infrastructure is homogeneous", "Make sure standards are respected", and "Allow developer to do lots of tasks themselves". Those benefit allow developer and operator working better as a team, and understnad each others work. "Continuous Delivery" help project team decrease the application delivery period by having faster application development, high frequency update can reduce the risk and cost of changing demand during the delivery. This is extremely useful for big data application because it always requirement lot of change during delivery. The increasing of delivery frequency can let the project team more familiar with the processes of application deployment, also will getting more feedback from the user.

IBM organizes the values of devops into three domains, "increase customer experience", "improve innovation ability", and "faster achieve value". [1]. Devops is not the goal of application development, but it can let development team reach their goal. It increase customer experience via faster update, and having faster response to customer's feedback. Then we using devops to avoid rework cause by misunderstanding the demand, so the project team have time and energy to investigate new technology. Finally, once the delivery period is shorter and shorter, user can actually use the application early before the content inside the application are out date, this is important for big data application because the replacement of big data analysis algorithm is changing all the time. Devops is endless, it will continually develop the entire project team include the technology, processes flow, teamwork, and team culture.

4 CLOUD PLATFORM AND DEVOPS

The expectation of cloud platform is change the capital cost to operate cost, but the true value of using cloud platform is decrease the barrier which slow down the development speed and developer time. Most of work on the cloud platform can be automatic, so development team can be free from the heavy lifting of daily work such as management hardware and patch installing. So using cloud platform mean company from providing product change to providing service, which fit the big data application concept. Customer don't buy the application if the content doesn't make sense, so let cloud platform as the big data application carrier is good for the project team founs on the content inside the applicaton and provide better service.

The SaaS significance is mixing operation and function, so user experience on SaaS will feel like they are integral. So user expect having high quality function and continuous update in the same

time. Unlike the traditional waterfall team provide "Waterfall methodology" rare release event, devops provide frequent release events can satisfy customer need and maximize utilize the automation of cloud platform. Devops make operate and development keep almost the same speed and flexibility which is suitable for big data business need in cloud platform.

5 THE WEAKNESS OF DEVOPS

The difficulty of implement is the biggest weakness of devops. Each company have their own development tool and process, they all have their own feature, so it is really hard to have a template that fit to every company and situation. For team leader, they only able to organized team and develop processes under limit source. So it is hard to having a optimal plan at most of time, for example to let each member inside the team have same understanding and sense of duty is really hard. The implement of devops having many requirements, clarify the current situation and set up phased goal is basic requirement for early devops.

6 CONCLUSIONS

The main purpose of doing devops in big data applications and analytic is to eliminate the isolate situation between Solution Architect(big data analyst) and programmer. This can be achieve by cross education training for both architect and programmer, so they can understand the basic concept and terminology from both domain. Once they finish the training, they will suppose to have a better understanding of each others' idea, and prevent the process off track. Furthermore, devops can let those two group testing the application environment and adjusting the foundation framework to meet the new needs, it represent faster fixing and update capacity.

REFERENCES

- [1] Gina Poole. 2017. DevOps delivers real value and successful business outcomes. (2017). <https://www.ibm.com/developerworks/community/blogs/invisiblethread/entry/devops.delivers.real.value?lang=en> [Online; accessed 08-October-2017].
- [2] Jerome Kehrli. 2017. DevOps explained. (2017). <https://www.niceideas.ch/roller2/badtrash/entry/devops-explained> [Online; accessed 08-October-2017].
- [3] Wikipedia. 2017. Devops — Wikipedia, The Free Encyclopedia. (2017). <https://en.wikipedia.org/wiki/DevOps> [Online; accessed 08-October-2017].

7 BIBTEX ISSUES

8 ISSUES

DONE:

Example of done item: Once you fix an item, change TODO to DONE

8.1 Formatting

Incorrect number of keywords or HID and i523 not included in the keywords

Other formatting issues

8.2 Writing Errors

Errors in title, e.g. capitalization

Spelling and grammatical errors

Are you using *a* and *the* properly?

Do not use the phrase *In this paper/report we show* instead use *We show*. It is not important if this is a paper or a report and does not need to be mentioned

8.3 Citation Issues and Plagiarism

It is your responsibility to make sure no plagiarism occurs. The instructions and resources were given in the class

Claims made without citations provided

8.4 Character Errors

Erroneous use of quotation marks, i.e. use "quotes", instead of " "

To emphasize a word, use *emphasize* and not "quote"

8.5 Structural Issues

Acknowledgement section missing

The paper has less than 2 pages of text, i.e. excluding images, tables and figures