ISSN 2347 - 3983
Volume 8. No. 5, May 2020
International Journal of Emerging Trends in Engineering Research
Available Online at http://www.warse.org/IJETER/static/pdf/file/ijeter37852020.pdf
https://doi.org/10.30534/ijeter/2020/37852020

Designing Automation System Based on Log Management for Bank XYZ's Data Center

Adrian Kosasih[1], Kevin Angriawan[2], Rahmat Ivan Aziz[3], Sfenrianto Sfenrianto[4]
[1234]Information Systems Management Department, BINUS Graduate Program-Master of Information Systems Management, Bina Nusantara University, Jakarta, Indonesia 11480.
[1]adrian010@binus.ac.id; [2]kevin.angriawan@binus.ac.id; [3]rahmat.aziz001@binus.ac.id; [4]sfenrianto@binus.edu

ABSTRACT

Bank XYZ is a banking company that has a Data Center department for the core of data and information distributions. Data Center department supports data storage, and providing the platform for the deployment of business applications. But some downtime issues still could emerge and require human resources to solve those issues. In this paper, we define some major issues those are connection down, server down, and broken data in UAT environment. Therefore, Bank XYZ's data center needs an automation system for accurate and immediate solution. Automation system will need information from the log management application to do responses. Thus, this research's objective is to have an automation design for operating automation in the Data Center department based on rules in log management. Research data are obtained from observation and interview methods. This research has discovered an automation design to help the Data Center department for helping them fixing their downtime issues.

Key words : Data Center, downtime issues, log management, automation system.

1. INTRODUCTION

Bank XYZ is one of the biggest banking companies in Indonesia. Their branch offices spread almost in every area of Indonesia. For getting all data and information from their branches, Bank XYZ's head office relies on their Data Center department. The Data Center department controls all of the data and information flows, between head office and branch offices. It is very important to make sure the flows of data and information are working well. Somehow it is not that simple.

The Data Center department is currently having some downtime problems. Disconnected internet often makes connections go down. It is not a good sign, because internet plays a very important role in communication these days [1]. Data servers also often go down and are needed to be turned on by the data center team manually. And because of the combination of those problems, data transfer processes are often disrupted and makes the data corrupted or broken in their UAT environment. If not solved immediately, those downtime issues can threaten the business of Bank XYZ.

Data center is a physical environment facility intended for housing computer systems and associated components [2]. It means the data center is the core for the distributions and flows of all data and information in the organizations. Those downtime problems often disrupt the flows of data and information of Bank XYZ. If those problems happen in working hours, the Data Center team can immediately solve the problems. But the biggest issue is when they happen outside of working hours and no one in the head office building can fix the happening problem [3].

From the problems described above, Bank XYZ needs solutions to recover their data and information flow if something bad unexpectedly happens, especially outside the working hours. Bank XYZ needs an automation system to handle all of those problems. Automation systems play a huge role in the successions of data centers [4]. Even the system cannot handle those problems fully, it can send notifications to the data center team in real-time to solve the happening problem. Data from logs are needed to build an automation system [5].

With logs collected and logic of processing defined, the automation system will do the procedure step by step [6]. If the happening problems are not defined, then the automation system can send alerts or notifications to devices of the data center team. They can immediately solve the happening problems, without waiting until working hours. The automation system can help the data center team to solve their downtime issues in faster ways. In result, downtime issues will be much reduced in Bank XYZ's data center using the automation system based on log management.

2. LITERATURE REVIEW

2.1 Data Center

Currently, the data center is one very important part of large business companies. The Datacenter can be interpreted as a facility that consists of computer networks and storage used by other businesses or organizations to organize, process,

store, and disseminate large amounts of data [7]. In business, it usually depends on the applications, services, and data contained in the data center, where the data center is used as a focal point and an important asset in daily needs. Data centers have large-scale, critical-scale computing infrastructure that operates all the time to drive the rapid growth of the IT industry and transform the economy at large [8].

Besides being an architectural part, understanding the data center can be interpreted as a facility for placing computer systems and related equipment, such as data communication systems and data storage. These facilities include redundant power supplies, excessive data communication connections, environmental control, fire prevention, and physical security devices [9]. So the data center can be interpreted as a combination of computer network infrastructure, supporting electrical devices, and building infrastructure that is adequate for the management and monitoring of data properly [10]. Based on the type of service, in general, data center development is grouped into two namely, First, the internet data center, only to support internet-related applications, usually built and operated by service providers or companies that have a business model based on internet commerce. Second, business data centers to support all functions that enable various business models to run on internet services, intranets, or both [11].

2.2 Downtime Issues

In the current era of globalization, increasing competition, and dynamic changes in the business environment forces companies to take action to gain a competitive advantage [12]. To achieve this requires constant development and the use of the best resources. For example, the data center sector is very dynamic. Datacenter equipment can be upgraded frequently, new equipment can be added, obsolete equipment can be removed, and simultaneously old and new systems can be used [13].

The problem of data center downtime that often occurs in some companies is indeed caused by many factors such as equipment failure, poor datacenter design, or human error. Downtime is a period when IT is only available in part because of planned maintenance or unplanned events [14]. To reduce "downtime" to increase efficiency is a well-known concept used in many industries [15]. In the measurement of effectiveness depends on the particular case and a series of methods used. Some situations are more frightening for companies than data center outages. Server downtime carries huge financial costs, potentially costing the company millions of dollars for every minute of the network and their data remains unavailable. For example, from direct financial costs, the impact of decreased productivity, lost opportunities, brand damage, and the potential loss of data can have a trickle-down effect that can affect business for years to come.

2.3 Log Management

Log Management is the process of handling every log event generated by all applications and software infrastructure where they run it. This process involves the collection, decomposition, storage, analysis, search, archiving, and disposal, with the ultimate goal of using data to solve problems and gain business insight, while also ensuring compliance and security of applications and infrastructure. According to the National Institute for Standards and Technology (NIST), Special Publication the definition of a log management system is the process of generating, transmitting, storing, analyzing, and managing data security on a computer [16].This logging process is usually recorded in one or more log storage files. So log management makes it possible to collect data in one place and see it as part of a whole, not a separate entity. Thus allowing analysis of log data collected, identifying problems and patterns formed to provide a clear and visual picture of how all systems work at any given moment.

Log management is a solution for handling large amounts of logs. Good log management can improve the quality of service, security and efficiency of the devices being monitored, by defining the correct rules. By using log management, the administrator gets a notification when a problem occurs, such as an error report from the hard disk drive [17]. Some of the objectives of log management are to collect logs from all log sources such as syslog, win log, and others. Make log searches and restore archived logs fast and flexible and identify irregularities in applications, databases, systems and devices in real time [18].

2.4 Automation for Data Center

Automation is a recent technology development that make improvement of the quality of living by automating the process in environments needed [19]. The current computing center has become increasingly complex and is adjusting to the need for new methods to support automation, analytics, and control so that it is currently receiving a lot of attention from the wider community [4]. All of these techniques exploit low-level data center infrastructure hardware, for example, such as using applications and systems, and related energy power consumption. Specifically, depending on the usage problem needed, some features that can be used to reveal more information than others, very flexible movements must collect as many metrics as possible. Fully automated data centers are now possible thanks to several powerful technologies such as advanced AI and machine learning solutions [20]. But, for any data center to be truly autonomous, a great deal of information needs to be collected about its operations and infrastructure first. Particularly, the visibility of an involved network plays a huge role in just how automated an operation can be.

There are three forces shaping up to help make the automated data center a reality. First is real-time infrastructure visibility has become a critical element of automating the data center [21]. Because an issue can occur anywhere on the virtualized,

hybrid infrastructure, and the issue can ripple to create problems elsewhere, making issue resolution a perennial challenge. Second is real-time analytics in short, it is because real-time responsiveness is critical for supporting anomaly detection, root cause analysis, remediation, prevention, and planning use cases, where processing and correlating massive amounts of information is critical. Lastly closed-loop systems, IT organizations are acutely aware of the perils of introducing changes to critical production environments [21]. In fact, they take great pains to establish a governance of change. Closing the loop is about humans delegating decision-making to technology.

3. METHODOLOGY

Researchers use observation methods and interview the manager of Bank XYZ Data Center department to obtain all the data and information needed. From Bank XYZ's data center downtime issues, researchers find their needs and requirements. Researchers then capture all the user requirements to define the rules in log management and design an automation system based by those rules.

For the first step, researchers collect the existing business processes data. The existing business processes data will be captured in activity diagrams. By knowing the details of existing business processes, researchers can figure out all of the main problems that cause downtime issues in Bank XYZ's data center.

The second step, after figuring out all of the issues, researchers find the needs and requirements of users based on those issues found in the existing business process. Those data center team's requirements will be mapped for defining the rules in log management and developing the automation system.

Next in third step, after all needs collected, researchers define and describe rules that are gathered based on user requirements use. The purpose of defining rules in log management, is to provide the accuracy of the automation system based on the happening problems. Rules in log management will be the decider of what automation system will do next.

And for the fourth or last step, researchers design the automation system for Bank XYZ's data center. All of the rules defined in log management will be implemented in the automation system. Most of the problems when the server is down, connection down, and data broken in UAT environment, will be automatically solved by the system. But if there are errors or bugs that are unknown by the system, it will send notifications to the data center team for further investigation. Researchers also propose new business processes for handling those issues, using the automation system designed.

4. RESULT AND DISCUSSION

4.1 Problem in Existing Business Process

Researchers have obtained the existing business process. The data center team always gets to know the server enter downtime based on notification from the people who use the server, or as in the bank called it as an IT General Person in Charge (PIC). The IT General PIC has responsibilities to make sure the server is usable for their purpose, as an example if the server is a web server that means the server will be used by many users. Since the server must be ready to be used by those users, therefore the availability and health of the server is a priority. The PIC must contact the Data Center department, which is the team that will be responsible for all components in the data center environment, from server, network device, scanner, etc.

This data center team is also responsible to resolve for any kind of issues that emerge from day to day. The report mechanism for PIC of the server to raise a ticket issue to the data center team as follows. Firstly, the PIC of the server will send an email to the data center team to troubleshoot their server to figure out the problem and fix it. From there, the data center team will analyze the issues. If the severity of issues are high, then the data center team has to open a new ticket and to be approved by the CTO of Bank XYZ as a work permit to resolve the issues.

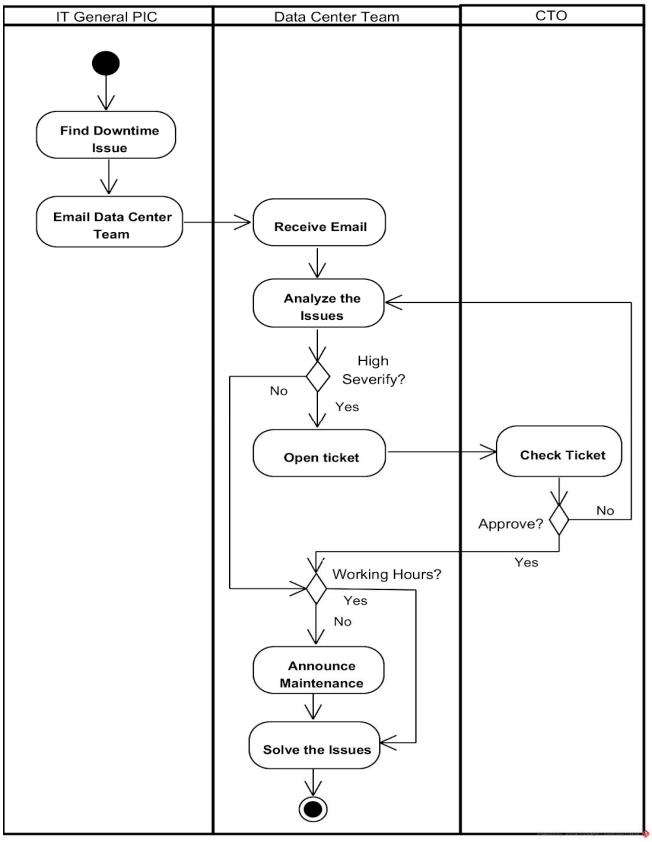

Figure 1: Activity diagram of the existing business process

Usually time to resolve the ticket could vary, depending on the availability of the data center team and the number of tickets that need to be resolved. Sometimes it also depends on the availability of the CTO for approving the open tickets. This flow is considered an old way or manual for a data center team to respond to every incident that emerges in a data center environment.

Because of this existing business process, the downtime issues often showed up. And sometimes if urgency does matter, those issues are still hard to solve because of this long and complicated workflow. Most technical issues found are when the server down and connection timed out**.** Server and connection are the critical tools for the continuance of the data and information distributions.

Broken data often found when there are issues with the server and connection**.** When transferring data, server, or connection down unexpectedly, in result the data in need is often corrupted or broken. Data center team must restore data manually and surely it wastes their time. Those are the main issues that cause downtime, and need to be solved immediately.

4.2 User Requirements

Every data center demand skilled operations team with ownership of the maintenance and lifecycle strategy, this combination is core to a data center's critical systems infrastructure's ability to continuously provide high-availability service delivery and uptime over a long amount of time for every component in the data center environment. In order to do that, Bank XYZ is required to address some of the issues that will be managed automatically by the system.

In this paper we figure out that bank XYZ has some issues that require automation to solve those servers down, connection down, and broken data issues. Not all of the issues can be solved by an automation system. Some of them need to be solved by the data center team. The issues mentioned are just three, but the variation of these issues often happen, and new kinds of sub-issues will be found. The automation system cannot fix a problem unless the problem has been fixed before.

As described above, the data center team of Bank XYZ needs automation solutions with high flexibility. High flexibility here means, when the issues found are the new one that the automation system cannot fix, then the system must send push notifications to the data center team directly to their mobile application. In result, either the issues are already scripted or not yet scripted in the automation system, they will be delivered faster to the data center team. The accuracy of information from the automation system is the most critical point for this workflow. That is the main reason why the automation system needs support from the log management.

4.3 Rules in Log Management

Rules about alerting when the server is down will be triggered when the system log capturing the event of the system shutdown unexpectedly with level error. This condition will trigger the log management application to trigger an alert to send a notification to the data center team with all related events for this incident attached in the notification and will automatically reboot the server. The sample of the log can be seen in figure 2 below.

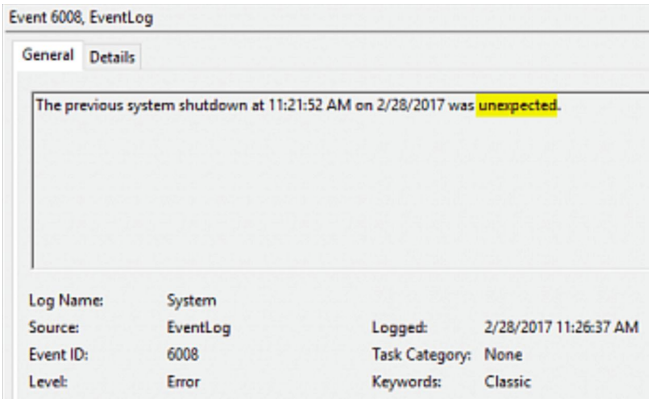

Figure 2: System log for unexpectedly shutdown

Next one if the application captured the log behavior shutting down on schedule with level info, then it will trigger an alert to send a notification to the data center team for notification of the incident. This sample of condition can be seen in figure 3.

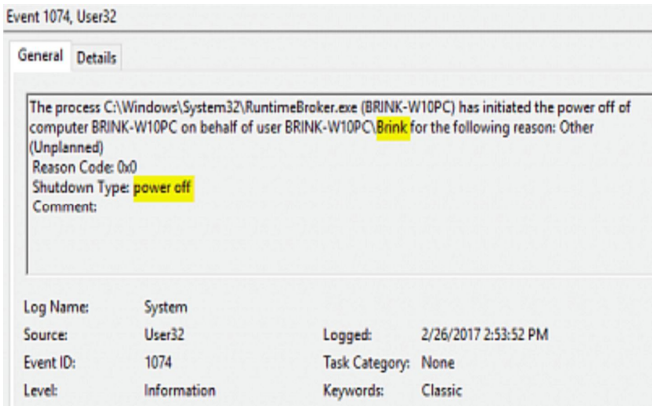

Figure 3: System log for shutdown on schedule

Rules about alerting when server connection is down will be triggered when the log management application captures a condition in the data. By monitoring the server uptime log, the log management application can monitor the behavior of the connection and if the uptime log is empty and status is failed, then this incident will trigger an alert to send a notification to the data center team to notify and execute a script to reboot the connection. The sample of the condition can be seen in figure 4 below.

Server Name	IP Address	Status	Uptime
Server01	0.0.0.0	Success	Days: 17; Hours: 14; Minutes: 5; Seconds: 38
Server02	1.1.1.1	Success	Days: 31; Hours: 14; Minutes: 10; Seconds: 10
XYZ		Failed	
Server03	2.2.2.2	Success	Days: 13; Hours: 14; Minutes: 39; Seconds: 9
Server04	3.3.3.3	Success	Days: 6; Hours: 20; Minutes: 14; Seconds: 57
Server05	4.4.4.4	Success	Days: 11; Hours: 9; Minutes: 57; Seconds: 34
Server06	5.5.5.5	Success	Days: 19; Hours: 14; Minutes: 15; Seconds: 39
Server07	6.6.6.6	Success	Days: 0; Hours: 1; Minutes: 10; Seconds: 49

Figure 4: Uptime log sample

For the broken data issues, the automation system also needs some rules. The broken data issues are often caused by the server and connection problems. This broken data only happens in the User Acceptance Test (UAT) environment, so the folders of data used are fixed. For the first rule, the automation system will read the log data center in a looping mode until it finds one or some folders with incomplete data, and get the logs of the folder (folder name, location, date of transfer data, and the reasons). Then the second rule, the automation will run the script to restore the broken data until it is completed. And the final or third rule, the automation system will send notifications to the data center apps. In result, the data center team can get information immediately after the auto-restore data completed.

As described before, researchers create a table to capture all the rulesets and rules needed for log management in the automation system.

Table 1: Rules in log management

Rulesets	Rules	Actions
1 Handling server down issues	1.1	Read log data center for server down issues until meet condition: there are staging apps cannot be accessed
	1.2	If condition in rule 1.1 met, then: • If expected (scheduled), then auto-restart server • Else if unexpected, then send notification to data center apps the detail of problems
	1.3	If server already ON, send notification to data center apps
2 Handling connection down issues	2.1	Read log data center for connection down issues until meet condition: latency > 999 ms or RTO (request timed out)
	2.2	If condition in rule 2.1 met, then fix the problems as scripted and restart connection
	2.3	If rule 2.2 done (connection restored), then send notification to data center apps
3 Handling broken data issues	3.1	Read log data center for every folder in UAT environments until meet condition: data broken or corrupted
	3.2	If condition in rule 3.1 met, then auto-run the scripted rules to restore the data
	3.3	If rule 3.2 done (all data restored), then send notification to data center apps

4.4 Design of Automation System

In order to boost operational efficiency for the data center team, it will require some sort of automation system that will help the data center team to resolve some issues automatically with less human interference. The methods that will be used in this paper will be using log approachment by collecting the log into a log management application. In order to do that, it will require installing an agent in each of the servers that will be monitored and start collecting the data by reading the logs and sending it to the log management application.

After all onboarding data has already successfully been set up, from there all rules that have already been defined need to be created in the log management application and fine tune the scheduler for the application to capture the incident for triggering the alert when the rules are met.

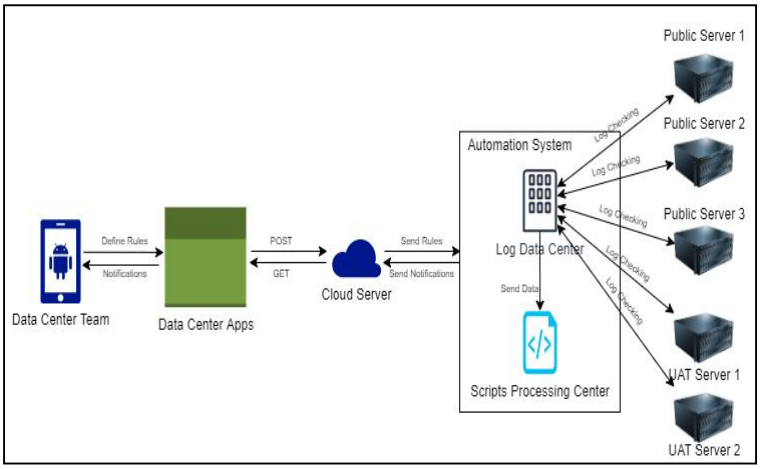

Figure 5: Architecture of the automation system

In figure 5 above, it can be seen that the log data center will always check the logs from Bank XYZ's servers. From the log data center, if there are downtime issues, the rules in the automation system will be checked and sent to the scripts processing center. It will automate the rules that have been assigned by the data center team. Several issues that arise but have not been defined in the rules then the automation system will send notifications to the data center application that connected with the data center team mobile application. The automation system will always check the log all the time,

since banking operational time is 24 hours. The automation system is always on standby mode to get the logs, and process them, when capturing downtime events based on the scripted rules.

When implementing the automation system, there will be some changes in the business process. In this automation system data center team acts as a main role, but they do not have to check the servers all the time. They just need to assign or make scripts to run the rules needed to solve the downtime issues. Each issue will have different ways to be solved. Researchers capture the new business processes of handling issues in activity diagrams.

For server down issues, the automation system will always loop-check and read the log data center. If there is no server down, the system will always continue logging. But if there is a server down, the system will try to match into two conditions. First condition is when the issue has been saved in the rules (scheduled). The automation will run the scripted rules and restart the server. The second condition, the issue is new and has not been saved in the rules (unscheduled). It will send notifications to the data center team in real-time, so that the data center team can solve the issue immediately after receiving notifications.

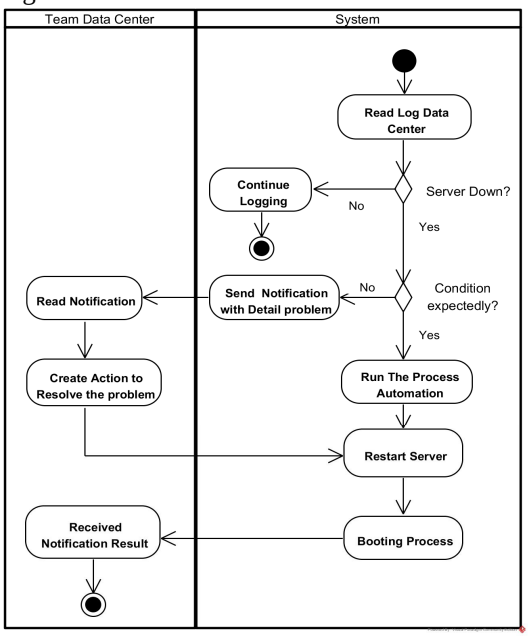

Figure 6: Activity diagram of handling server down issues

Similar to handling server down issues, the automation system will always check and read the log data center. If there is no connection timed out or down, the log checking of Bank XYZ's servers will keep logging. But if there is a connection down issue, the scripts will be run based by rules defined. Because there are some types of connection timed out problems and each problem has their own solving-rules. After the booting process is done, the automation system will send the notifications to the data center team. By default, servers' internet connection must be on all the time.

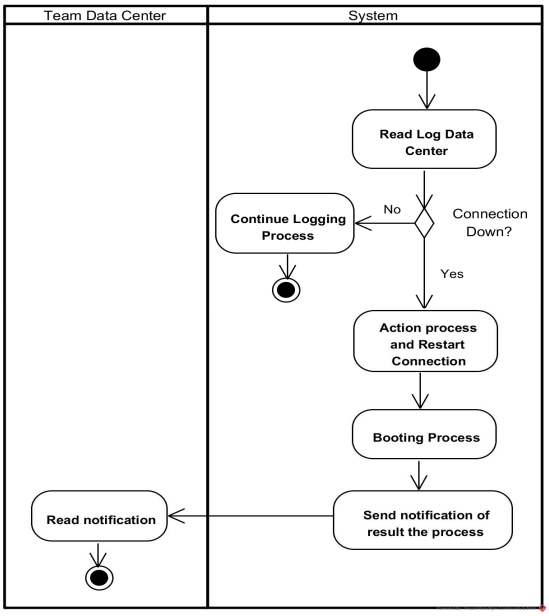

Figure 7: Activity diagram of handling connection down issues

For broken data issues, the automation system, at the beginning, reads the log for the UAT environment. The rules implemented in the automation system and it only defined for the folders in the UAT environment. After that, the automation system will run the scripted rules to auto-restore the broken data. After the data restoration process is completed, the data center team will get notified. The UAT data can be used in a proper way again.

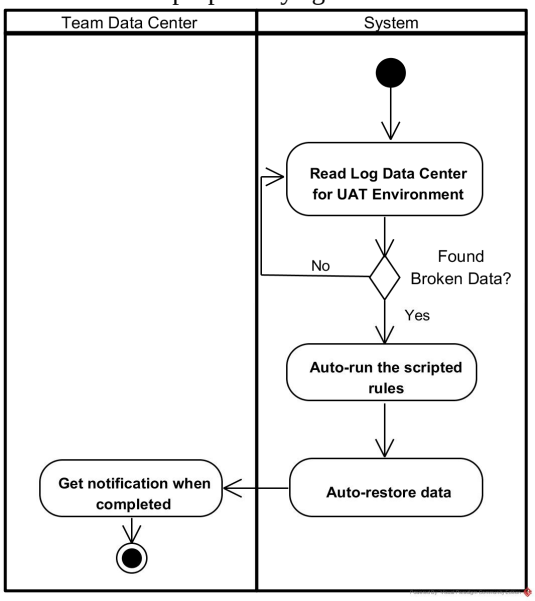

Figure 8: Activity diagram of handling broken data issues

With implementations of the automation systems based by log management, the new business processes of handling server down, connection down, and broken data in UAT environment issues, are found. **The key process is when the automation system read the data log center and found downtime issues.** And if the issues are not being "learned" yet by the automation system, it can send notifications to the

data center team. After that, the data center can add new rules in log management based on new issues faced.

5. CONCLUSION

The design of the automation system for the data center team gives help by handling various troubleshooting and recovery tasks based on the knowledge held by the system. The automation system can solve issues with less human interference. This research has discovered three rulesets including three rules in each ruleset in log management to be run in the automation system. Those rules will be scripted in the automation system for handling server down, connection down, and broken data in UAT environment issues.

The benefit of using the automation system can be implemented in Bank XYZ's data center to reduce downtime issues (server down, connection down, and broken data in UAT environment). Rules in the log management make the data center team easier to find downtime problems and systematically solve them. Issues can be solved immediately, even outside working hours. Using the automation system, data and information distributions in Bank XYZ will run much faster and smoother.

REFERENCES

1. U. Akshita, R.C. Venna, K. Niharika, K. R. Vidya Prasad & G. K. Yogeshwari. **Voice-Based E-mail system for Visually Impaired People.** *International Journal of Emerging Trends in Engineering Research (IJETER)*, 8(4), April 2020, pp. 1315-1318.
2. Rashid, Md. Abdur. **Data Center Architecture Overview**. 2019
3. T. Callahan. **Doing Security Orchestration, Automation, and Response Before It was Born**. 2019
4. A. Libri, A. Bartolini, & L. Benini. **Dig: Enabling out-of-band scalable high-resolution monitoring for data-center analytics, automation and control.** *In 2nd International Industry/University Workshop on Data-center Automation, Analytics, and Control (DAAC)*, 2018.
5. S. M. McCoy, S. M. Gydesen, & C. M. Markus, **U.S. Patent No. 8,180,824.** *Washington, DC: U.S. Patent and Trademark Office*, 2012.
6. A. D. M. Africa, T. K. Dolores, M. C. Lim, L. S. San Miguel & V. R. Sayoc. **Understanding Logical Reasoning Through Computer Systems.** *International Journal of Emerging Trends in Engineering Research (IJETER)*, 8(4), April 2020, pp. 1187-1191.
7. H. Zhang, K. Chen, W. Bai, D. Han, C. Tian, H. Wang & M. Zhang. **Guaranteeing deadlines for inter-data center transfers**. *IEEE/ACM transactions on networking*, 2016, 25(1), pp. 579-595. https://doi.org/10.1109/TNET.2016.2594235
8. Dayarathna, Miyuru, Y. Wen, and R. Fan. **Data center energy consumption modeling: A survey**. *IEEE Communications Surveys & Tutorials 18.1*, (2015), pp.732-794.
9. Dewanannta. **Pengantar Jaringan Komputer - Data Center**. *ilmukomputer.com*, April 2017.
10. F. Fernando Asali, I. Afrianto. **Rekomendasi Data Center Menggunakan Pendekatan Standarisasi TIA-942 di Puslitbang XYZ**. *Jurnal CoreIT*, Vol.3, No.1, 2017. https://doi.org/10.24014/coreit.v3i1.3532
11. Henriyadi. **Data Center dan Implementasinya Pada Perpustakaan**. *Jurnal Perpustakaan Pertanian*, Vol. 17, No.2, 2008,
12. K. Stecuła, and J. Brodny. **Generating knowledge about the downtime of the machines in the example of mining enterprise**. *International Multidisciplinary Scientific GeoConference: SGEM: Surveying Geology & mining Ecology Management*, 17(1.3), 2017, pp.359-366.
13. M. Levy, and D. Raviv. **A novel framework for data center metrics using a multidimensional approach**. *In 15th LACCEI International Multi-Conference for Engineering, Education, and Technology: Global Partnerships for Development and Engineering Education*. 2017 https://doi.org/10.18687/LACCEI2017.1.1.387
14. Y. Wang, E. Coiera, B. Gallego, O.P. Concha, M.S. Ong, G. Tsafnat, D. Roffe, G. Jones, and F. Magrabi. **Measuring the effects of computer downtime on hospital pathology processes.** *Journal of biomedical informatics*, 59, 2016, pp.308-315.
15. R. Wolniak. **Downtime in the Automotive Industry Production Process–Cause Analysis**. *Quality Innovation Prosperity*, 23(2), 2019, pp.101-118. https://doi.org/10.12776/qip.v23i2.1259
16. C. Phillips, S. Kevin, et al. **Logging and Log Management**. *Waltham: Elsevier*, 2013, pp. 267-303.
17. B. Sudha, & S. A. Kumar. **Service Delegating Log Management-For Secure Logging In Cloud Environment.** *International Journal of Computer Techniques*, 2(2) 2015, pp. 23-29.
18. K. Agrawal, & R. H. Makwana. **Data Analysis and Reporting using Different Log Management Tools.** *International Journal of Computer Science and Mobile Computing*, 4(7), 2015, pp. 224-229.
19. Bestley Joe S., R. Ramadevi, V. Amala Rani & G. Rajalakshmi. **Automatic Cooking Machine using Arduino.** *International Journal of Emerging Trends in Engineering Research (IJETER)*, 8(1), January 2020, pp. 35-40. https://doi.org/10.30534/ijeter/2020/07812020
20. D. Acemoglu and P. Restrepo. **Artificial intelligence, automation and work (No. w24196)**. *National Bureau of Economic Research*, 2018.
21. P. Desnoyers, T. Wood, P. Shenoy, R. Singh, S. Patil and H. Vin. **Modellus: Automated modeling of complex internet data center applications**. *ACM Transactions on the Web (TWEB)*, 6(2), 2012, pp.1-29.