DATA BACKUP AND RESTORING USING RPA

Revathi.N[1], Piraivendhan.K[2], Sreerag.R[3], Nandakumar.S.D[4]

UG Students[1,2,3], Assistant Professor[4]
Department of Computer Science, KGiSL Institute of Technology, Coimbatore, TamilNadu, India

ABSTRACT

Data backup is the process of backing up the data, relate to copying the data from one file location and storing the data in some other location namely an archived file. Backups have two distinct purposes. The primary purpose is to recover the data after its loss, it can be data deletion or data corruption. Data loss is a common experience of computer users while transferring the data. The secondary purpose of data backup is to restore data from an early time, according to the users demand and need, typically organized within a backup application for how long copies of data are required. These processes can be fully automated with the RPA solutions by providing them with the required credentials, initial and terminal details for the whole task to be automated. These process can be fully monitored with the RPA solution

Keywords- Data collection, separating the data, data backup, data restoring, Orchestrator, bot connection.

I. AREA INTRODUCTION

RPA stands for Robotic Process Automation is the application of technology that allows employees in a company to configure computer software or a "robot" to capture and interpret existing applications for transaction processing, data manipulating, response triggering and act with other digital systems.

In traditional workflow automation tools, a software developer produces a list of actions to automate a task and interface to the back-end system using internal application programming interfaces (APIs) or dedicated scripting language. In contrast, RPA systems develop the action list by watching the user perform that task in the application's graphical user interface (GUI), and then perform the automation by repeating those tasks directly in the GUI. This can lower the barrier to use of automation in products that might not otherwise feature API's for this purpose.

II. LITERATURE REVIEW

Current backup methods (i.e. taking a copy of data outside of the system it resides on) involves the data transfer to storage media, mostly tape which undergo certain problems. Most business processes know of tape method of backup of computer data and it is the most common method of data backup today. Data Backup using tape technology requires a connection to a media rotation and replacement plan - neither simple nor quick. Inserting tapes, rotating tapes, labeling tapes, storing or offsiting tapes, finding tapes for recovery purposes all rely on manual processes which introduces errors. Instead someone needs to be given the work of managing the system - easy in a large company perhaps but still the job is ordinary. But in a smaller company, there are no obvious choices of person, but when it comes to restoring a critical lost file, it is usually chaos and panic.

Data Backups may have been carried out, but they may not have worked successfully or may not have completed and rarely does anybody check whether the backup has been successful until it's too late, e.g. when a real disaster has occurred. Due to the nature of tape backups, the backup process need to be carried out once a week and it will allow only a small amount of data to the backup process. Most tape backups suffer from slow restore speeds, the recovery process are tested very rarely without any restore.

International Journal of Technical Innovation in Modern Engineering & Science (IJTIMES)
International Conference on Recent Explorations in Science, Engineering And Technology (ICRESET'19)
Volume- 5, Special Issue- March, 2019

In order to accommodate the ever increasing data volumes, existing backup systems and tape libraries need to be expanded which is costly, backup windows may run into the business day which impacts on system availability and the recovery time objective (i.e. the time it takes to restore all data) cannot be met.

Tapes can become old, unreadable, break, crack or be lost. Technology advances mean that the system used to backup your data could easily be lost in the event of theft, disaster, etc. Thus it needs to be replaced with a current system. The problem being that the new system is likely to be incompatible with the tapes that hold your vital data.

III.   PROPOSED SYSTEM

In our proposed system, the data from the different sources are collected on a daily basis. The bot will be trained to the process of extraction of data based on condition. The process will be idle until the admin triggers the bot using orchestrator. After the triggering factor, the files from the different sources are collected as a master file. The backup process is done for the master file.

Once the backup process is done, the data are stored in the orchestrator and it will generate a downloadable file. This file contains the output of the backup process. After this the restoring process is done through the orchestrator and the output is generated as a downloadable file. The file is in the form of excel sheet.

The Fig.III.1 is the system design of the proposed system. It indicates that the input is received as a business file and the data gets extracted based on condition. After that the backup and the restoring process is performed through the orchestrator. Once the admin triggers the bot, the process gets executed and the output is generated.
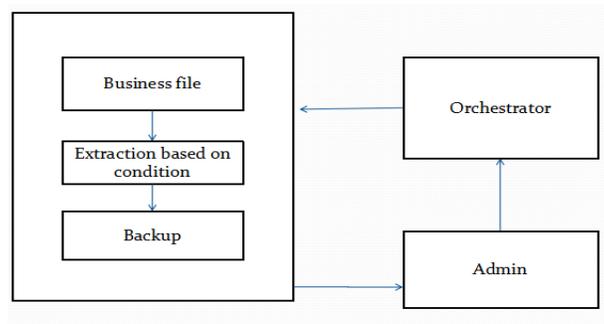
Fig.1.System Design

IV.   PROCESS FLOW

This is the basic process flow of the data backup and restoring process. Data from the different sources are collected and saved as a master file. The data backup and restoring process will be done for this master file through the orchestrator and the output is generated.

Fig.2.Process Flow

International Journal of Technical Innovation in Modern Engineering & Science (IJTIMES)
International Conference on Recent Explorations in Science, Engineering And Technology (ICRESET'19)
Volume- 5, Special Issue- March, 2019
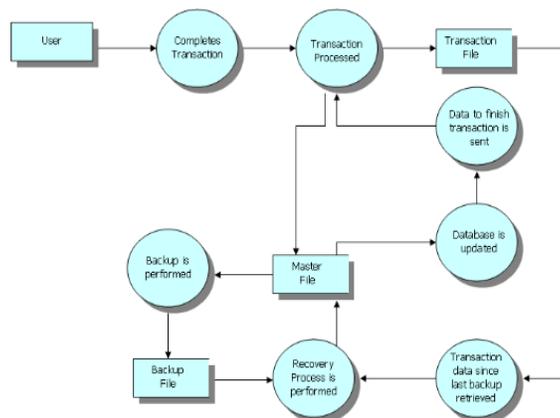
The process flow of the data backup and restoring process has been modularized as,

A: Reading the input file and separating the data

This is the initial process of the data backup and the restoring process. The input file is received from the mail and the data are written into the excel sheet. Once the input file has been read, the data gets separated from the input file based on the given condition. This process is shown in Figure 3 and Figure 4.
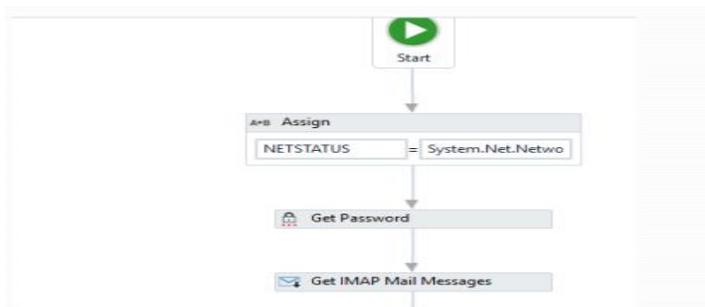
Fig 3. Reading the input file from the mail.
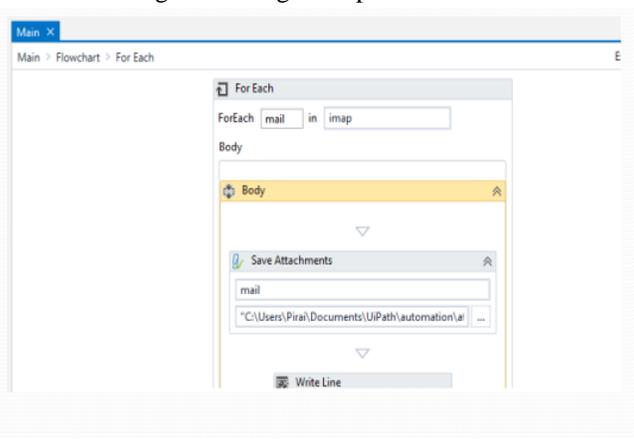
Fig 4. Separating the data

B: Backup process

After separating the data, the backup process has to be performed. In order to perform this process, the queue activities must be used. Queues are used to store the large amount of data. These queues are present in the UiPath Orchestrator. Initially we have to create a queue in the orchestrator and that queue name has to be used in the UiPath studio so that the data can be stored in the queue. Once the backup process gets completed, it will generate the output in the form of an excel sheet. This process is shown in figure 5.

Fig 5. Backup process

International Journal of Technical Innovation in Modern Engineering & Science (IJTIMES)
International Conference on Recent Explorations in Science, Engineering And Technology (ICRESET'19)
Volume- 5, Special Issue- March, 2019
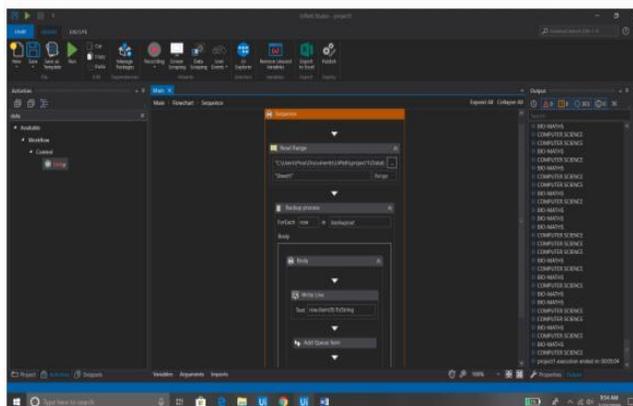
C: Orchestrator connection

The orchestrator is used to deploy the project that has been created in the UiPath studio. The developed workflow in the Uipath studio has to be published to the Orchestrator. Once it is done, a key will be generated which is mandatory to connect the robot to the Orchestrator. After the connection has been established, the output is store in the queue which is present in the Orchestrator.

D: Restoring process

After the backup process, the data restoring process will be performed. The restoring process will be done for the data that has been backed up. It is also done by the queue activity and the output data gets stored in the Orchestrator queue. Finally, it will generate the output file in the form of an excel sheet. This is shown in Figure.6
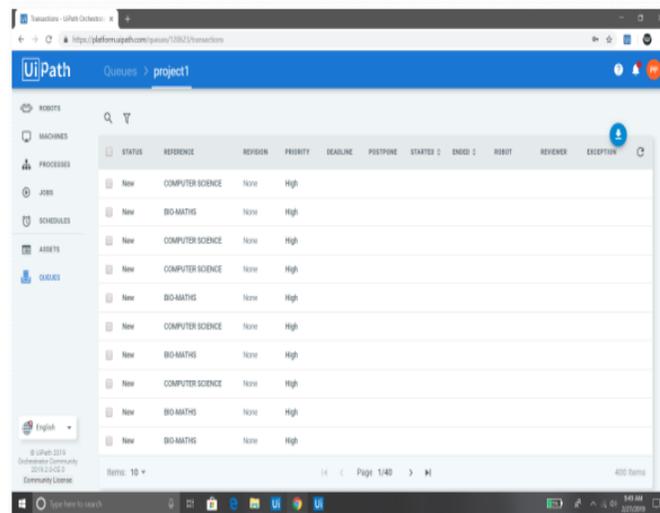
Fig 6. Restoring Process

V. CONCLUSION

The main aim of this paper is to provide the essential information about the data backup and the recovery process. The key idea we have applied here is the automation. Automation we have applied does not require any manual work in order to manage the data during the process. It supports large volume of data to backup and because of the automation process, the performance of the system does not affected by the internal process. It helps in scaling the resources and meet the need of growing business data. It requires only short time for the whole automation process which is very much beneficial when compared to the manual process. The security is enhanced so that the data cannot be lost during the process. The inter operability speed of the system gets improved by our proposed system, so the efficiency gets benefited.

VI. REFERENCES

[1] S.Suguna and A. Suhasini, "Overview of data backup and disaster recovery in cloud," *International Conference on Information Communication and Embedded Systems (ICICES2014)*, Chennai, 2014, pp. 1-7. doi: 10.1109/ICICES.2014.7033804

[2] S.Q. Yao and C. Gu, "Research and Implementation of Data Recovery Technology Based on WINDOWS FAT," *2010 International Conference on Machine Vision and Human-machine Interface*, Kaifeng, 2010, pp. 549-552. doi: 10.1109/MVHI.2010.214

International Journal of Technical Innovation in Modern Engineering & Science (IJTIMES)
International Conference on Recent Explorations in Science, Engineering And Technology (ICRESET'19)
Volume- 5, Special Issue- March, 2019

[3] Travis S. Luckey Director of Information Technology VTM Group, "The Key Stages of Disaster Recovery Planning" 2009 Continuing Education 1277 University of Oregon Eugene, OR 97403-1277 (800) 824-2714

[4] R.D. McDowal, "Computer (In)security - 2: computer system backup and recovery", *The Quality Assurance Journal*, vol. 5, pp. 149-155, Sep. 2001.

[5] Jun Liu, Shuyu Chen, "HVF: An optimized data backup and recovery system for hard disk based consumer electronics", *Optik*, vol. 126, pp. 251-257, Jan. 2015.

[6] M. Lillibridge, K. Eshghi, D. Bhagwat, "Improving restore speed for backup systems that use inline chunk-based deduplication", *Proc. 11th USENIX Conference on File and Storage Technologies*, pp. 183-197, 2013.