

A Study on File Processing System and Pivot Automation Process

T. Murali Krishna¹, Ravi Kumar Lingala², Desta Dana³

^{1*} Asst. Prof in IT, School of Informatics, Wolaita Sodo University, Ethiopia (murali2007tel@gmail.com)

² CSE Department, Best management Consulting Solutions Ltd, Hyderabad. (ravikumar.lingalal@gmail.com)

³ Lecturer in IT, School of Informatics, Wolaita Sodo University, Ethiopia (destaju@yahoo.com)

ABSTRACT:

This paper illustrates the file processing system in the form of templates, flat files, documents and all other kind of source files, various methods and their inter-relationships. This is accomplished through the staging or bridging of different methods and structures, and then blending to the distinct process. Here, the research has abstracted all the good features of the three paradigms with application of good database design principles. The Pivot automation file process system is the example to illustrate more of the automation of existing files or reports or templates or pivot files which are coming from various large source systems. The schemas which all are interconnected through centralized distributed operating database processing system.

KEY TERMS

Data Source, Schemas, Data Mart, Files, Templates and Reports.

INTRODUCTION:

For many years computer scientists spent most of their work on research of software development technologies, while less effort was spent to make the use of the already existing and well developed software more convenient. In part, this problem can be explained by software developers' concerns about software sales leaving software usage, utility problems into users' hands. The complexity of the whole system is increased when one company or organization acquires software from more than one vendor and software is introduced with significant time span. This way a complex heterogeneous system environment is formed.

In this environment, there are multiple groups involved, among system users: those who are end users of the system's business functionality and system administrators. The responsibilities include system security as well as technical configuration of system environment including end to end process of the system. By its complexity, the area of business system administration is comparable to network administration. There are plentiful tools for network administration, maintenance and monitoring in the market; however, the authors of this paper could find no conventional solutions for this mixed system administration process execution governing. It can be explained by the diversity of the systems including the custom based nature of legacy system communication.

Understanding the current state of requirements engineering practice in software industry is an important step for research practitioners in order to improve software quality. In spite of attempts to practice requirements engineering in small, medium-sized software enterprises in order to improve software quality. The significant issues relating to this area such as inconsistent, incomplete requirements, and inadequate requirements management tools still remain. Several requirement techniques have been proposed in the last decade in order to reduce requirement engineering challenges. However, Large number of requirement engineering techniques causes confusion while technique selection process. Thus, there is a need to investigate requirement techniques in the best interest of the quality. The proposed methodology overcomes all the short falls, absence of correctness, completeness authentication which includes limitation in the number of view elements abstracted through existing approaches. The proposed method identifies the functional dependencies on control flow system which in turn, provide the best solution with more accuracy.

Automation Process or controlling system is a well-known problem. There have been many attempts to solve it in software history. In the era of mainframes, process management was partly delegated to the operating system and Job Control mechanism. As a significant tool of this area, Description Language must be mentioned (1, Version control, Control -M referring staging process 2, Jams scheduling). It is a language targeted current areas of application include automation process and real-time applications in general; however, it requires a very

detailed process description. It's not suitable for high-level business process description. Therefore, authors introduce a simple domain-specific language for business process control. In the first part, problems are identified and brief solutions are explained. The second part describes the architecture of the process control mechanism, its domain-specific language used for process control description.

PROBLEM STATEMENT

The current engineering practice in software industry needed various advanced, comprehensive mechanisms make available for best solutions to reduce the high manual intervention in the day to day activities completion in the industry.

To promote qualitative deliverable in the specified time line along with service level agreements and also to deliver accurate information in the least time bound. In order to perform the listed above objects, introducing of Automation method within the industry provided the best solution to the long time pending issues. In addition to that (i) *Use of heterogeneous systems* (ii) *Framework prototype and it's technology* (iii) *Automation Process*.

(I) Practice of Heterogeneous Systems:

The aim of this research paper is to describe the practice of heterogeneous software in large companies, where many different software platforms are used. The formation of heterogeneous environment in long running large companies is unavoidable, if the necessary software is acquired gradually; the size and functions of the companies are changing over time. The most serious problems are caused by distributed environment where many source systems are running simultaneously on different platforms across the place and communicating with each other. Typically, operators and administrators of each of these systems have to have specific management skills. As a rule, file processing system process carried out by the systems are done correctly; and there is no process control built in systems. If one process is carried out by two or more independent systems, each system can control execution as far as it is in its scope, but the whole process typically is controlled manually. Therefore, the usage of systems depends on the qualification of the number files which are processing at a given point of time and precise execution of operations by the scheduling part. Automated process is proposed to solve the problems described above and to reduce the dependency of system usage on the subjective factor of the supporting staff. In brief, the proposed solution contains two components: a description of controlled processes like Control M/JAMS and a methodology for automating the process execution according to process.

(II) Framework/Prototype:

The process of automation is purely based on the technology which is needed to provide the similar kind of features which can react sufficiently changes of existing domain as well as other internal configuration and environment. A framework is the step by step process to execute the specified function and which can be useful to introduce the new method or process to reduce the complexity. The concept of designing the framework and relevant technology includes internal, external environment testing, version upgrade and self-testing and others. The concept of pivot automation includes Pivot Table Toolkit and Tools for Advanced Industrial Solution Applications(AISA), volume 4. References mentioned in the excel reports and dash boards by John and various researches, basic automation control concepts.

In the design phase, the triggering processing plays vital role to make the various operations like copying the data and refreshing of the data with help of designed technology or software macro's and template data copy to specified destination Cogappl.

(iii) Automation Process or Process Control:

The suggested automated process concept works over heterogeneous systems within the organization and which is the combination two steps data extraction from various systems or sources and keeping the data in the staging place and moving into destination path or location. In the initial part, there are numerous stages like query generation, invoking the template macro and batch collection. In the middle stage, jobs status changed to next level and checks the each of the pending job which are associated to multiple data sources and all templates moved to staging Job Access Management System(JAMS). It is an enterprise Job Scheduler tackles one of the greatest challenges facing businesses today – the ability to integrate complex business processes across multiple applications and platforms place at the end, another new trigger invokes to move the file to destination path or location. In this ending steps, all the files which are in the staging place which consist of updated information and it's populates the template data tab from the query result set for each of its process log's the each step of the process.

METHODOLOGY:

As per as implementation concern, it contains more technical and administration process to determine the actual process being executed within the time and it's synchronized and tracking of order of all the heterogeneous systems. In this implementation process consist of (i) Development process and (ii) Automation Implementation along with Archive. In the development process follows the below steps

Development Process:

- Source file or template basically having multiple tabs, there are two tabs generally uses this automation process those are data tab and pivot tab and it requires data source like any kind of raw sources materialized views, tables to maintain information and (MVW, materialized view is a database object that contains the results of a query. The FROM clause of the query can name tables, views, and other materialized views. Collectively these objects are called master tables (a replication term) or detail tables, then its copy the data and paste it onto data tab.
- In each template has macro which was developed as per the need, which when triggered by automation process, copies the data from data source, copies it onto data tab and zips the template and copies onto destination path (naming as given as COGAPP1, It is a Pivot application production server which has live data for all the listed files which are coming from various sources and scheduling in Job access management system).
- Data sources are created under any database or a particular schema (like Pivot team schema) under production environment/destination path like enterprise data ware house production EDWP1, Intermediate place called enterprise data ware house release test EDWRT and create the authentication setup credentials like User Name: pivot team, Password.
- List of pivot templates and concern reports are available in table Pivot_List_Excel10.
- Each pivot table/ template has four process flag.
 JAMS currently refreshing data source set by JAMS
 data source updated, pivot ready for processing set by JAMS & ETL
 pivot currently being processed set by Pivot Automation
 pivot processing complete set by Pivot Automation
- For any new pivot table request, we create the excel template with data tab, pivot tab and macros and add an entry into Pivot_List_Excel10 table in both staging and live environment and test them on sandbox server(orld853).Once, they are tested successfully and approved, we move these templates to subversion.(Version control). From Subversion, through configuration management process all the files /templates moves to destination server (cdc69).

```
<?xml version="1.0" encoding="utf-8" ?>
```

```
<Configuration>
```

```
<ConfigSections>
```

```
<section name="databaseEnvironments type="System.Configuration.DictionarySectionHandler" />
<section name="databases" type="System.Configuration.DictionarySectionHandler" />
<section name="directoryPaths" type="System.Configuration.DictionarySectionHandler" />
<section name="sharePointPaths" type="System.Configuration.DictionarySectionHandler" />
<section name="testingContext" type="System.Configuration.DictionarySectionHandler" />
<section name="notifications" type="System.Configuration.DictionarySectionHandler" />
<section name="archiveRetension" type="System.Configuration.DictionarySectionHandler"/>
<section name="envBackup" type="System.Configuration.DictionarySectionHandler"/>
<sectionname="adminEmail" type="System.Configuration.DictionarySectionHandler" />
<sectionname="autoPromote" type="System.Configuration.DictionarySectionHandler" />
</configSections>
```

Processing – Continued:

This is the actual place where the entire automation process can take place with supporting of scheduling software applications or tools like Control-M and JAMS, In this process invokes Pivot Automation is the first step, In this stage, all pivot Automation queries (generated / pre-defined code) pending pivot ids with a status of 4 (pending) into a batch collectionUpdates those statuses to 6 (in-process) through automation code script and Processes each template in the batchQueries the associated datasource then populates the template data tab from the query result set and it Invokes the template macro to create the corresponding pivot tabbed output file zipped then deploys the resulting pivot or template (email, SharePoint, other shared folders etc.) and updates status to 8 (completed) finally, In each stage there would be logs recorded to provide the status of the each process.

Processing Failures

All thrown exception details are emailed to Prod Support for action

Typically, the situation is corrected and the processing flag is reset and awaits JAMS

Common examples failing the template's macro:

Empty datasource, unexpected data values

Deployment

Pivots deploy to both internal and external clients via:

Email	internal, external
SharePoint	internal
Cognos	internal

Archiving

All pivot output is archived for 90 (as per the requirement) days then rolled off unless identified on the exceptions list

Exceptions are held indefinitely (selected MTD and YTD pivots)

Backups

Backed-up nightly (templates and application)

Disaster Recovery is formally outlined and practiced (DRE)

Development & testing

Pivot automation follows our Release Mgmt process using ITEMS and Subversion to control changes

Nothing is committed to Subversion without an ITEM number.

Test environment parallels the production environment -- orld53

Pivot processing is invoked manually, not by JAMS.

Test server also serves as our production failover server regarding Disaster Recovery

Automated process for Deploying template changes to production

Production pivot changes typically take effect on specified date

Staged template files are auto-promoted to production on scheduled time.

When necessary, templates are promoted to production manually.

<Databases>

```
<add key="EDWRD" value="Data Source=EDWRD.WORLD;User id=pivotteam;Password=team2pivot;" />
```

```

<add key="EDWRT" value="Data Source=EDWRT.WORLD;User id=pivotteam;Password=team2pivot;" />
<add key="EDWP" value="Data Source=EDWP1.WORLD;User id=pivotteam;Password=team2pivot;" />
<add key="UDWD" value="Data Source=UDWD.WORLD;User id=pivotteam;Password=team2pivot;" />
<add key="UDWT" value="Data Source=UDWT.WORLD;User id=pivotteam;Password=team2pivot;" />
<add key="UDWP" value="Data Source=UDWP.WORLD;User id=pivotteam;Password=team2pivot;" />
</databases>
</archiveRetension>
<envBackup>
<add key="backupMgmtEnabled" value="Y" />
<add key="backupRetensionCount_Daily" value="10" />
<add key="backupRetensionCount_DisasterRecovery" value="1" />
<add key="backupDirectories_Daily" value="..\admin,
        ..\batch,
        ..\bin
        ..\etc,
        " />
<add key="backupDirectories_DisasterRecovery" value="..\admin,
        ..\batch,
        ..\bin,
        ..\build,
        ..\deployment,
        ..\etc,
        ..\log,
        ..\template,
        " />
</envBackup>

```

RESULTS:

There are numerous advantages of this heterogeneous file processing automation system. This can be helps to organization at various levels and mainly this will be helpful to maintain the history / archive information (depends on the configuration and internal structure of the organization you can retrieve the one year history data) and also this will be helpful to save the huge amount of time and to reduce the manual process or intervention. In this automation process, batch of jobs or schedules generate as per the specified time and will send the updated information to the destination place and this also maintain the log history to keep the status of the files/ jobs and also to tack the failures and error details.

- Primary focus is to avoid manual intervention.
- Log history and error handling process and triggering alerts.
- Archive/History
- 24/7 Support.

CONCLUSION:

Scheduled tasks are best executed from a server, usually from a system account with administrator privileges. The only drawback to executing tasks from a system account is that a user interface is not available; however, automated tasks typically are intended to run in the background, outside a user's notice.

Reduction in production time – having a machine that is automated definitely speeds up the production time since no thinking is needed by the machine; there is better repeatability and less human error.

Increase in accuracy and repeatability – when an automated machine is programmed to perform a task over and over again, the accuracy and repeatability compared to an employee is far greater.

Less human error – no one is perfect, and we are all prone to making mistakes. This is why a machine that performs repeated tasks is less likely to make mistakes than an employee.

Less employee costs – by adding automated machines to an operation, means less employees are needed to get the job done. It also indicates less safety issues, which leads to financial savings. With having fewer employees, there are numerous costs that are diminished or reduced such as payroll, benefits, sick days etc.

Increased safety – having automated machines means having fewer employees who perform tasks that can be dangerous and prone to injury, which can make the work environment safer.

Higher volume production – investing in automated equipment creates a valuable resource for large production volumes, which in turn, will increase profitability.

RECOMMENDATION:

The Automation process is an advanced system and it can be recommended as a commercially viable production technique given its superiority in terms of quality, cost and production rate as compared to the conventional approach. This report has only investigated the commercial viability of the production method. It has not made an assessment of the market demand or marketability of this product. It is recommended that these factors be assessed. Market demand should be quantified and analyzed against production and implementation costs in order to establish the process. Further research is required to establish the validity of the smaller grain sizes resulting from the short. In addition, this report recommends further work to make an assessment of the demand or marketability, quantify the market and establish the optimal scale by analyzing market demand against production and set-up costs; The reason for recommending one of the actions is pointed out.

REFERENCES:

- [1] Excel Pivot Tables That Automate Tasks You No Longer Have Time for (How to Get the Most Done In the Least Time Book 3)19 June 2013 by Erik Kopp
- [2] Information processing systems - Open Systems Interconnection - File Transfer, Access and Management...23 August 2007 by ISO/IEC JTC 1
- [3] <http://www.excelforum.com/member.php?action=getinfo&userid=20970>
- [4] Process Pivot Control Instrumentation Technology” by Curtis Johnson.
- [5] <http://scitechcommentary.blogspot.com/2012/08/guide-to-automating-excel-reports.html#!/2012/08/guide-to-automating-excel-reports.html>
- [6] Database System Concepts and file processing system A. Silberschatz, H.F. Korth, S. Sudarshan McGraw-Hill, 2006.
- [7] <http://www.i-programmer.info/ebooks/automating-excel.html>
- [8] <http://scitechcommentary.blogspot.com/2012/08/guide-to-automating-excel-reports.html#!/2012/08/guide-to-automating-excel-reports.html>