

# Optimized approach of analysing the data loads through process chains in SAP Business Ware house

M. Victoria Hebseeba<sup>#1</sup>, Dr. T. Bhaskara Reddy<sup>#2</sup>, Prof. B. Lakshma Reddy<sup>#3</sup>, Dr. T. Venkateswarulu<sup>#4</sup>

<sup>#1</sup>Department of Computer Science, Andra Pradesh, India

<sup>#2</sup>Department of Computer Science, S.K University, Andra Pradesh, India

<sup>#3</sup>Director, Garden City college, Bangalore, Karnataka, India

<sup>#4</sup>Department of Computer Science, Gate College, Tirupathi, India

**Abstract**— This paper describes a unique technique to display the data loading status of all the process chains occurring in the system. The objective is to increase the speed on correcting the data load failures and at the same time provide the capacity to analyse the data load status from a technical point of view. All the process chains can be part of this including any other created after the implementation. This technique can be customized as per the requirement as and when chains are to be included in the system. The approach presented in this paper is best suitable in production support environment in any large scale industries.

**Keywords**— process type, process chain, variant, ETL, business ware house

## I. INTRODUCTION

A process chain is a BW objects which has a scheduled sequence of processes linked together and are executed in a specified order. Process chain waits in the back ground for an event. Some of these processes trigger a separate event that can start other processes in turn. In an operating system there are a multitude of processes in addition to the loading process that occur regularly [1] [2]. Process chain is a mechanism that controls extraction, transformation, and loading (ETL) activities execution at a scheduled time in a defined manner. Process chains are used to automate the complex structures with the help of the controlled processing, visualize the schedule by using network applications and centrally control and monitor the processes [3].

Openness is applied to the theory behind process chains, in that both user-defined programs and processes can be implemented. In addition, you can include process chains in other process chains, so-called Meta chains. In doing so you have the option of integrating process chains from the BW system, in which the Meta chain is found, or from other BW systems [4]. In this context, local or remote process chains are concerned. Processes are scheduled before your run and can be monitored with the standard batch monitor. You start subsequent processes

by using background events. Short dumps and terminations are recognized and handled respectively.

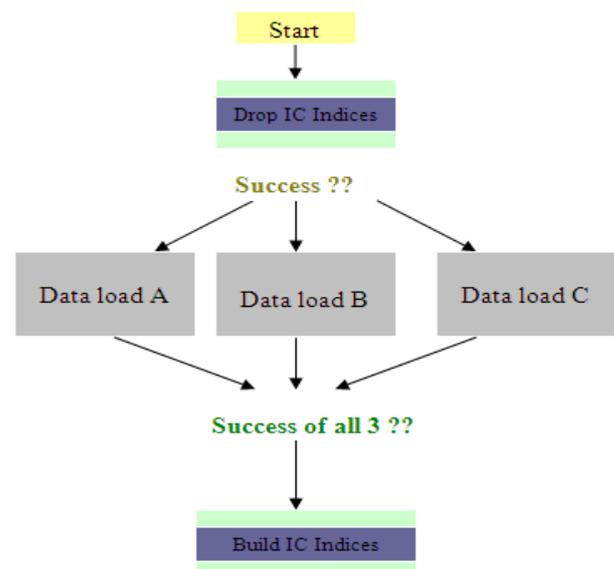


Fig. 1 Process Flow: Data loading from Info packages in parallel to 1 info cube

## II. CURRENT CHALLENGE

A process chain is a mechanism that controls extraction, transformation, and loading (ETL) activities execution at a scheduled time in a defined Sequence. Since in many support project environments in IT industry multiple process chains are run you come across some issues with the system executing them at appropriate times and becomes difficult to analyze the data load status from technical point of view and takes time in correcting the issue.

## III. PROPOSED SOLUTION

The proposed solution will be delivered to automate the ETL team data load monitoring to increase the speed on correcting data load failures at the same time it provides the capacity to analyze the data load status from a

technical point of view and from update frequencies. This solution consists of a unique report for status check which displays the final status for all process chains occurring in the system, with the purpose of identifying delays/issues and solving these ones as soon as possible.

All the process chains can be part of this report including any other created after the program implementation. We can choose the selection for the date range and objects that the user wants to monitor. The objects can be process chain, aggregates, ODS and cubes. Once the selection is made push the execute button to generate the report with the status of each load. With the intention of facilitate the comprehension of process chain status on monitoring task, two traffic lights were added to the report.

#### A. Technical Status

Green represents that status can be basically finished, Red for error and yellow represents the process chain is still running. To implement these flags, a table was created to relate all the possible chains with two times defined as maximum\_period(technical) and maximum\_waiting (Data Quality). The logic is applied to the program.

If process chain finish → display green light 

Else if not finish and process chain time < maximum\_period → display yellow light 

Else → display red light 

#### B. Data Quality Status

Data Quality is the maximum reasonable time for each process chain. This logic is applied to the program so that it can represent the real status of each load into BW system. The report can be executed when ever required to check the status which helps to increase the speed on correcting data load failures and monitoring the executions. Every new user who needs to access the report can be easily customized.

If process Chain finish and process chain Time <

Maximum\_waiting → Display Green light 

Else if not finish, next Process Chain finish and process chain time < maximum\_waiting → Display green light



Else if not finish, next Process Chain not finish and process chain time < maximum\_waiting → display yellow light 

Else if not finish, next Process Chain not finish and process chains time > maximum\_waiting → display red light . End if.

#### C. Logic Decomposition

Copy the table to find the next process chain  
if tb\_loads\_auxi-scheduled\_starting\_date is initial move corresponding loads to loads\_next  
Collect loads\_next.

End if.

Sort next\_loads by chain\_name, actual\_starting\_date and actual\_starting\_time.

To find the Number of records  
Describe table loads\_next lines wc\_lines.  
Loop at loads\_auxi.

To check the time of the Process Chain  
If not loads\_auxi-actual\_start\_date is initial.  
Read table tb\_chain\_config with key  
chain\_id = loads\_auxi-chain\_id.

Perform calculate\_maximum\_waiting using actual start date, actual start time, actual end date, actual end time.

Operation to move the finish into maximum waiting. If load is completed successfully moves green light to load status. If load is not completed modifies the status as per the current status and continues.

If maximum\_waiting\_ok = 'X' and wc\_end\_ok = 'X'.  
move green\_light to loads\_status.  
modify loads\_auxi  
continue.

Else if if loads\_auxi-actual\_end\_time is initial.

This operation checks if process chain is not finished and next Process chain finish into max\_wait for the current process chain date and Time.

If not tb\_loads\_aux-act\_end\_time is initial clear:  
wf\_tzone.  
Convert date tb\_loads\_aux-act\_end\_date  
time tb\_loads\_aux-act\_end\_time  
into time stamp wc\_time time zone wf\_tzone.  
End if.

Move Status Red or Yellow to QM status based on Finish after max\_wait

If wc\_time >= maximum\_waiting.  
Move icon\_red\_light to loads\_auxi-qmstatus.  
modify loads\_auxi  
End if

else "=> act\_start is initial  
move icon\_yellow\_light TO tb\_loads\_aux-qmstatus.  
modify tb\_loads\_aux INDEX wf\_tabix.  
End if.

#### D. Test Results

Performed several lab tests with real SAP Business ware house implementation.

- Developed process chains for Query Run time Statistics-Delta and loads from Operational Data

Store (ODS) and Info cube considering the dependency condition for business critical Assets data. Customised the report for the Process chains developed

- Executed the report by giving date range in the selection criteria
- Compared the result of the report developed with that of the info packages inside the process chain.
- The status displayed was exactly matching in both the scenarios.

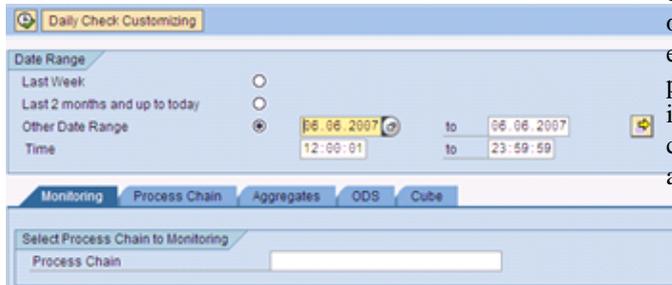


Fig. 2 Date range in selection screen

Result is displayed in the format: Process chain description, scheduled date, scheduled time, actual start, actual end, technical status, update frequency, comments.

Process Chain	Schedule...	Schedule...	Actual Start	Actual Start	Actual End	Actual End	Technical...	Upd Freq
Query Runtime Statistics - Delta	12.12.2011	08:39:13	12.12.2011	08:39:13	00:00:0000	00:00:00	000	000
Query Runtime Statistics - Delta	13.12.2011	08:39:09	13.12.2011	08:39:09	00:00:0000	00:00:00	000	000
Query Runtime Statistics - Delta	14.12.2011	08:39:24	14.12.2011	08:39:24	00:00:0000	00:00:00	000	000

Fig. 3 Status check program displaying the failed status of the process chains

Process Chain	Schedule...	Schedule...	Actual Start	Actual Start	Actual End	Actual End	Technical...	Upd Freq
PCA ODS and Cubes	06.06.2007	12:55:33	06.06.2007	12:55:34	06.06.2007	13:23:47	000	000
PCA ODS and Cubes	06.06.2007	13:55:33	06.06.2007	13:55:37	06.06.2007	14:19:17	000	000
PCA ODS and Cubes	06.06.2007	14:55:33	06.06.2007	14:55:34	06.06.2007	15:53:34	000	000
PCA ODS and Cubes	06.06.2007	15:55:32	06.06.2007	15:55:33	06.06.2007	16:10:13	000	000

Fig. 4 Displaying success status for the successful completions.

#### IV. CONCLUSIONS

Process Chains are the new generation of job scheduling and monitoring tools for tasks such as data loads, reporting agent jobs, or index rebuilds. The proposed approach is very beneficial and best suited to be executed daily or when ever required to increase the speed on correcting data load failures and monitoring the executions. All the process chains in the system can be part of this report including any other newly created. This is very helpful to the ETL team to increase the speed on correcting data load failures at the same time helps to analyze the loads status.

#### REFERENCES

- [1] Thomas Schroder, "SAP BW Performance Optimization Guide", 1st ed., Galileo Press, Germany, (2006)
- [2] Andrew J. Ross, "SAP Net weaver BI Accelerator", 1st ed., Galileo Press, Boston (MA), (2009)
- [3] Christian Merwald, Sabine Morlock, "Data ware Housing with SAP BW7, BI in SAP Net Weaver 2004s, 1st ed., Rocky Nook Inc., (2009)
- [4] Biao Fu, Henry Fu. SAP BW a step by step Guide, 2002
- [5] Arshad Khan, "SAP and BW Data Warehousing. How to plan and implement (2005)
- [6] Danniell Knapp, "SAP Netweaver BI 7.0 Migration guide"
- [7] [http://www.sap.com/platform/netweaver/pdf/BWP\\_AR\\_IDC\\_BI\\_Accelerator.pdf](http://www.sap.com/platform/netweaver/pdf/BWP_AR_IDC_BI_Accelerator.pdf)
- [8] Kevin McDonald, Andreas Wilmsmeier, David C. Dixon, W.H Inmon, "Mastering the SAP Business Information Ware house" Wiley Publishing Inc, USA (2002)
- [9] Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification, IEEE Std. 802.11, 1997.
- [10] Elizabeth Vitt, Michael Luckevich, Stacia Misner, "Making better Business Intelligence Decisions faster", Microsoft Press, USA
- [11] Gary Nolan, Debasish Khaitan, "Efficient SAP Netweaver BW implementation and Upgrade Guide.
- [12] <http://airccse.org/journal/ijngn/papers/3111ijngn04.pdf>
- [13] <http://www.ijmo.org/papers/13-P30001.pdf>
- [14] <http://airccse.org/journal/acij/papers/1111acij08.pdf>