IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 17, Issue 6, Ver. III (Nov – Dec. 2015), PP 44-52
www.iosrjournals.org

Role of Operational System Design in Data Warehouse Implementation: Identifying the Operational System Design Gaps and techniques to handle them

Deepak Asrani1, Renu Jain2, Usha Saxena3
[1]*(Department of Computer Science Engineering, Teerthanker Mahaveer University, Moradabad, U.P (India))*
[2]*(Department of Computer Science Engineering, University Institute of Engineering & Technology, Kanpur, U.P (India))*
[3]*(State Institute of Health and Family Welfare, Indira Nagar, Lucknow, U.P (India))*

Abstract: *Data warehouse designing process takes input from operational system of the organization. Quality of data warehousing solution depends on design of operational system. Often, operational system implementations of organizations have some limitations. Thus, we cannot proceed for data warehouse designing so easily. In this paper, we have tried to investigate operational system of the organization for identifying such limitations and determine role of operational system design in the process of data warehouse design and implementation. We have worked out to find possible methods to handle such limitations and have proposed techniques to get a quality data warehousing solution under such limitations. To make the work based on live example, National Rural Health Mission (NRHM) Project has been taken. It is a national project of health sector, managed by Indian Government across the country. The complex structure and high volume of data makes it an ideal case for data warehouse implementation.*

Keywords: *Data warehouse, Operational System, Dimensional Model, Atomic Data, NRHM*

I. Introduction

With the exponential growth of data volumes in business organizations, and highly competitive business environment, business organizations are looking for solutions that could provide them effective business reporting not only for their short term business goals but also for long term strategic decision making. Data warehousing solutions are highly useful in this direction and have been taken by the researchers as an important area of research. Data warehouse designing process takes input from operational system of the organization. For smooth and effective implementation of data warehousing solution it is necessary to have an operational system in place. Quality of data warehousing solution depends on design of operational system. Operational systems are capable of generating reports for monitoring and control of routine business operations, but are not able to produce reports for strategic decision making. Often, operational system implementations of organizations have some limitations. Thus, we cannot proceed for data warehouse designing so easily. We have worked out to find possible methods to handle such limitations and have proposed techniques to get a quality data warehousing solution under such limitations. We have taken National Rural Health Mission (NRHM) Project for data warehousing solution implementation. It is a national project of health sector, managed by Indian Government across the country. In the NRHM project, Business managers in different states across the country want to analyze all the key operations, which require flexible reporting capabilities. When we need to analyze a system for its long term performance and we also want to identify the favoring and non-favoring factors in the fulfillment of organizational mission and objectives, we require a system which is capable of generating reports useful for long term data analysis and decision making.

Operational system is a key component required for data warehouse implementation. Unlike operational systems, data warehousing solutions are not mandatory but are useful in decision making process of the organization. Standard data warehouse design techniques exist. If applied effectively, they result in quality solutions. But, often the approach followed by the designers is not as per standards. It results in solutions that have several drawbacks. It is one of the critical factors responsible for data warehousing system failure. Our major focus is upon improving the methods of data warehouse development process. We have tried to explore the measures required to be taken during data warehouse design and development. Using the techniques identified here, one gets solution of the mentioned limitations and an effective data warehousing solution.

Organization of the paper is as follows: Section 2 gives literature review. Section 3 covers design techniques for operational system and data warehousing system, followed by section 4 that describes NRHM project, its goals and objective, operations, and key performance indicators. In section 5, we have discussed limitations of NRHM operational systems. In section 6 we have proposed design techniques and practices to

deal with limitations of operational system and achieve effective data warehousing solution. Section 7 concludes and specifies future scope.

II. Literature Review

There are many research areas in data warehousing where substantial research work has been done. Still, further research is required because of continuous technological evolutions and changing user expectations from these technologies. Major areas of research in the past on various aspects of data warehousing include materialized view maintenance, query execution performance, modeling and design, integrated approach of data warehouse design, Schema design, Data warehouse development process, effective maintenance, standardization of data warehousing techniques, technical and managerial issues in data warehousing and many others. Some of the areas that have received attention of research are clear from the given extracts.

Querying over Data Warehouse is much faster than traditional Relational DBMS. The performance improvement is expected to be much higher as the size of the Dataset increases [1]. [2] has implemented Real Time Data Warehouse framework. Their experimental analysis shows that if the delay comes from the incremental maintenance of DMM, no ETL technology (full-reloading or incremental-loading) can help in real-time business intelligence. Among all the I/O schedulers, the anticipatory scheduler should be avoided [3]. The DW process, though supporting bottom-up extraction of information from data, fails in top-down enforcing the company strategy. A new approach to BI, called Business Performance Management (BPM), is emerging from this framework: it includes DW but it also requires a reactive component capable of monitoring the time-critical operational processes to allow tactical and operational decision-makers to tune their actions according to the company strategy. [4]. [5] discusses issues of slow-evolving fact, transaction-oriented fact table, large dimensions and have proposed the concept of virtual data cubes and shown its usefulness. Active Data Warehousing refers to a new trend where data warehouses are updated as frequently as possible, to accommodate the high demands of users for fresh data [6]. Peer Data Management Systems (PDMSs) have been proposed as architectures to support sharing of operational data across networks of peers while guaranteeing peers' autonomy, based on semantic mappings that mediate between the heterogeneous schemata exposed by peers [8]. Data warehousing is an essential element of decision support. It aims at enabling the knowledge user to make better and faster daily business decisions [7]. The process of data warehousing starts by analyzing operational system of the organization. Incomplete and poor operational system design is bound to put hurdles in smooth design and implementation of data warehousing solution [9].

Failure rate of data warehousing projects has been a major concern in data warehousing industry. Prediction about data warehouse failure rate is clear from the statement of the Gartner's research published in a news article. According to STAMFORD, CONN February 24, 2005 — "Data warehouses play a crucial role in the success of a business intelligence (BI) program. However, through 2007, more than 50 percent of data warehouse projects will have limited acceptance, or will be outright failures, as a result of a lack of attention to data quality issues, according to Gartner, Inc." [10]. Fewer than 30% of business intelligence projects meet the objectives of the business, according to research from analyst Gartner [11]. To cater with such issues, we need to investigate probable reasons of data warehousing project failure. Some of the apparent reasons for their failure are that Data warehousing projects are costly and require long time periods for design and implementations. Often data warehousing projects are neither given due weightage by all levels of management nor are effectively utilized post implementation. Lack of user participation during design and development is one of the major reasons for their failure. It is evident from the included references that usually the focus of researchers has been on design and implementation of data warehousing solution without considering the operational system design. Moreover, the high failure rate of data warehousing projects has also been a major concern for the researchers. We have performed in depth investigation of operational system design and its limitations and have focused mainly on the interaction with the client to determine their expectations from the system to be developed.

III. Operational Systems

Operational systems are developed to smoothly run all business operations of the organization on regular basis. These systems automate functioning of the organization. Objective of these systems is to help in effective handling of daily business activities and ensure proper functioning of all the organizational operations. Operational systems are mainly used by junior management to record details of daily business operations and to enquire for information about a particular business operation held in the organization. Operational systems usually handle one record at a time to extract its details. Major operations in operational system are regular insertion/addition and updating of records. Reports are usually daily, weekly, monthly or quarterly that are used to visualize various figures related to operations of the organization and are used to monitor execution of business activities. Reports in such systems help in decision making about corrective measures for short term to improve operational efficiency of the organization. Operational systems maintain atomic data for every business

transaction. Queries are mostly for individual records or list of records. Often, moderate aggregation is performed on data to generate reports. Operational system usually maintains data for a period of one or two years and transfers the older data to a separate data archival in order to save storage space and improve performance of query execution. Operational system generates reports, based on only current data. It does not generate historical reports.

3.1 Operational System Design

Operational systems are designed by taking inputs from different organizational vouchers, and interacting with different user groups about data items they need to record for proper functioning of the organization. Database schema for operational system is designed either by using ER- Modeling or Normalization or combination of both. ER-modeling follows top down approach. In this approach the first step is to identify all possible entity sets in the organization followed by determination of attributes of each entity set in turn and declaring prime attribute and foreign keys on the basis of uniqueness property and cardinality of relationship between any two entity sets respectively. ER modeling is effective in situations where organizational documents and vouchers are not available and we need to design the schema by analyzing user requirements from scratch. Design using normalization is done in the presence of some structured data of the organization's business operations. Normalization is based on the concept of functional dependencies and suggests including only those dependencies in the tables where non prime attributes are dependent on prime attribute. It also helps in enabling removal of multi valued dependencies from the tables. Fig. 1 shows a sample ER-Diagram for one of the identified segment of NRHM project.

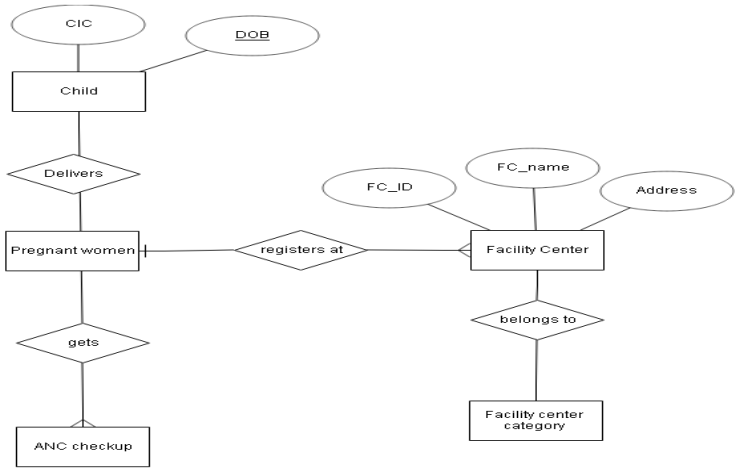

Fig1. ER Diagram for ANC Registration Process

3.2 Role of Operational System in data warehousing

Data warehouse schema is populated with the data from operational system of the organization and also with the domain specific data from external sources. Data warehouse design involves cleaning, integration, and transformation of operational system data. Complexity of these steps depends on quality of operational system.

Because of differences in objectives, and query requirements the data structure for both type of systems is highly different. While operational systems are generally exposed to queries asking for non aggregated figures for small time span for selected records, the nature of queries for data warehousing solutions are based on asking highly aggregated figures for a large time span based on multiple parameters. Since inception of data takes place with the transactions occurring in the organization that are recorded in the operational system. It is the same data that needs to be transferred to data warehouse after cleaning and transforming into the form suitable for data warehouse schema. If data quality of operational system is good it makes the population process of data warehouse simple and fast otherwise the process of data warehouse population becomes complex.

Following are quality features identified in the operational system of the organization for implementation of data warehousing.

i. Completeness: All required data items must be present in the operational system. Partial operational system implementation would not facilitate effective data warehouse implementation.

ii. Minimum heterogeneity among multiple operational systems or their versions: Higher the heterogeneity among multiple operational systems, higher the complexity of the cleaning process.

iii. Descriptive names of data items in the operational system, containing expected type and format of values.

iv. Availability of continuous data for past multiple years.

3.3 Data warehousing systems

Data warehousing is a process of accumulating vast amount of data from operational system of the organization and some external sources depending on requirements and nature of the organization to create huge data repository exposed to ad hoc business queries that produce required information within acceptable response time [12]. Data warehouse systems are designed with an objective to analyze performance of the organization required for strategic decision making. It includes high level data summarization and aggregation. These systems are used to store and maintain data for years in a structure that is suitable for strategic reporting. These systems have to deal with very large volume of data. Hence the issue of query execution performance arises, which is resolved by keeping pre-aggregated intermediate query results. Data warehouse environment needs extraction, transportation, transformation, and loading (ETL) solution, online analytical processing (OLAP), data mining capabilities, client side analysis tools, and other data gathering tools to deliver useful reports to business users.

Effectively designed, implemented and utilized data warehousing solutions would result in the growth of the business organization backed up by informed decision making capabilities.

3.4 Data warehouse system design

Data warehouse designing is generally based on dimensional modeling which includes identification of key business processes, facts to be measured, dimensions of analysis, and grain for measurement of facts followed by designing multi dimensional star or snowflake schema depending on the nature of data and user requirements. Fig. 2 shows example dimensional model for ANC registration process.

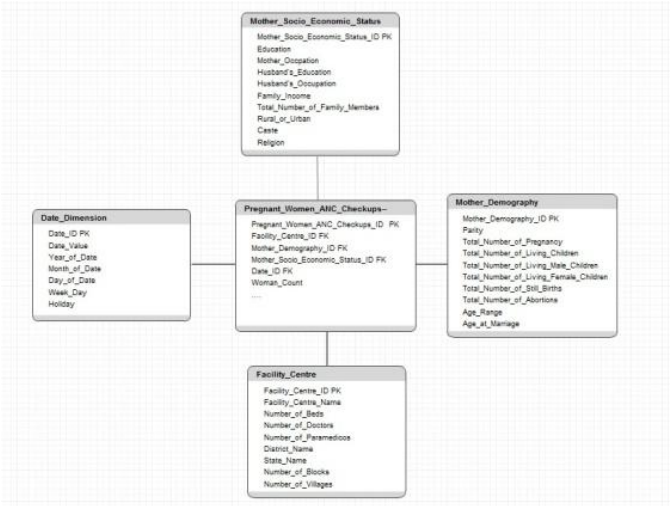

Fig2. Dimensional model for ANC checkup process of NRHM project

IV. National Rural Health Mission (Nrhm) Project

National Rural Health Mission (NRHM) is a project of Indian government that started in the year 2005 to provide quality health services throughout the nation under ministry of health and family welfare. Though, the project runs with 17 different health programmes, major focus of NRHM has been on health care of mother and child under Reproductive and Child Health (RCH). Due to lack of facilities in the large area of the country, especially in the rural areas, government took initiative to establish infrastructure in the form of setting up primary health centers and community health centers for providing health services at a very large scale. Major objective of the RCH programme has been to improve statistics of mother mortality rate (MMR), infant mortality rate (IMR), and total fertility rate (TFR). To monitor the progress and effectiveness of various health programme, concept of hierarchical reporting from Primary Health Centre (PHC), Community Health Centre (CHC), to District head quarters, state head quarters to state government and central government was formalized specifying various monthly, quarterly, and annual reporting formats under Health management Information System. Later for monitoring and management of Mother and Child health an online portal has been launched called MCTS (Mother and Child Tracking System) that accepts inputs from PHCs, CHCs, District Hospitals, and Sub District Hospitals at transaction level for the services these facility centres are providing.

To effectively analyze performance of various health facilities provided under NRHM project, with an objective to get in depth analysis of factors responsible for lower performance as compared to the desired level.

The organization needs to develop a system that could generate various reports as per user expectations. Process wise details of expected reports are given in the tables from 1 to 6.Many research activities and surveys are conducted by programme managers and decision makers of NRHM project to find out

performance of programme execution, reasons for lesser performance in specific areas, and measures required to be taken for improvement. Carrying out such surveys are costly affairs when conducted at national level. Instead what we need is presence of such an operational system implementation and data warehousing solution at national level so that need of such research activities are eliminated and replaced by an automated application implementation that is capable for reporting such reports from within the application. It will save huge costs of separate surveys/researches, time and efforts. The decision makers/ Programme managers will be benefited, if they find the causes of lesser performance in order to provide type of intervention required in the project.

Similar data is being currently collected by various research activities conducted separately. Many times the data is based on respondent's interviews and is not so huge. Often data collected is based on various sampling procedures. Huge data require financial burden and its collection is time taking also. Therefore introducing above indicators in the data warehousing solution will give us facts about such a huge data and will save from additional financial burden and will save time also.

4.1 Key Performance Indicators

Performance indicators are means of measuring quality of services provided and level of achievements of a programme. NRHM project has taken infant mortality rate (IMR), maternal mortality rate (MMR), total fertility rate (TFR) and morbidity and mortality rates as key performance indicators.The objective of NRHM is to reduce maternal and infant mortality rates, total fertility rate (TFR) and morbidity and mortality rates. There are many other identified performance indicators under various programmes but we have mentioned the performance indicators related to the mother and child health care as it is highly focused segment of the NRHM project.

Table1: Business Processes from RCH Programme of NRHM

Business Process	**Sub Processes**
Ante Natal Care (ANC)	Registration, ANC Checkups, Immunization
Delivery	Home Delivery, Institutional Delivery, Child Health
Post Natal Care (PNC)	Folloup Visits, Child Immunization

Table 2: Measurements for Ante Natal Care (ANC): Registration

Facts to be measured	**Dimensions of Analysis**
Total number of registered women	Education, Occupation, Per capita income
Total number of registered women	Education, Occupation, Parity
Total number of registered women	No. of times received JSY benefit, Education, Per capita income
Total number of registered women	No. of times received JSY benefit, facility centre
Total number of registered women	Family type, Education, Occupation, Per capita income
Total number of registered women	Total no. of pregnancy
Total number of registered women	Outcome of pregnancy like live birth, still birth, abortion, Education, Occupation, Family type
Total number of registered women	No. of surviving children
Total number of registered women	Ever born, Live birth, Education, Occupation, Family type, Facility type

Table 3: Measurements for Ante Natal Care (ANC): ANC Checkups

Facts to be measured	**Dimensions of Analysis**
Total number Women completed 3 ANC checkups	Source of information like ASHA or ANM, Facility centre type, Education, Occupation, Per capita income, Rrural/urban settings, Distance from home to facility centre, Transport facility available, Family type, Parity of children
Total number Women not completed 3 ANC checkups	Source of information like ASHA or ANM, Facility centre type, Education, Occupation, Per capita income, Rrural/urban settings, Distance from home to facility centre, Transport facility available, Family type, Parity of children

Table 4: Measurements for ANC Immunization

Facts to be measured	**Dimensions of Analysis**
No of women who received TT dose	Source of information like ASHA or ANM, Facility type like PHC/CHC/SC, Provider, Distance from home to facility centre, Transport facility, Family type, Parity of children
No. of women registered who have high BP	Source of information like ASHA or ANM, Facility type like PHC/CHC/SC, Provider, Distance from home to facility centre, Transport facility, Family type, Parity of children
No. of women registered who have hemoglobin less than 11	Source of information like ASHA or ANM, Facility type like PHC/CHC/SC, Provider, Distance from home to facility centre, Transport facility, Family type, Parity of children, Education, Occupation, Per capita income, Total no of pregnancies, Ever born, Surviving children, Availability of service provider

Table 5: Measurements for Delivery

Facts to be measured	**Dimensions of Analysis**
Total number of registered women	Home Delivery, Facility Centre
Total number of registered women	Institutional delivery, Facility Centre
Total number of registered women	Home delivery, No. of times JSY benefit received
Total number of registered women	Institutional delivery,No. of times JSY benefit received
Total number of registered women	Institutional delivery, Duration of stay in hospital
Total number of registered women	Home delivery, No. of pregnancy, Education, Occupation, Per capita income, Rural/urban settings, Distance from home to facility centre, Transport facility available, Family type, Parity of children, No. of children ever born, No. of children born alive
Total number of registered women	Institutional delivery, No. of pregnancy, Education, Occupation, Per capita income, Rural/urban settings, Distance from home to facility centre, Transport facility available, Family type, Parity of children, No. of children ever born, No. of children born alive
Total number of registered women	Home delivery, Total no of children ever born
Total number of registered women	Institutional delivery, Total no. of children born alive
Total no. of children born	No. of still birth, No. of abortions
Total no. of children born	No. of surviving children
Count of delivery	Trained birth attendant, Education, Occupation
Count of delivery	Trained birth attendant, Family type
Count of delivery	Trained birth attendant, Transportation facility
Count of delivery	Untrained birth attendant, Education, Occupation, Availability of doctor, Distance from home to facility centre
Count of delivery	Untrained birth attendant, Family type, Availability of doctor, Distance from home to facility centre
Count of delivery	Untrained birth attendant, Transportation, Facility centre, Availability of doctor, Distance from home to facility centre
Result of pregnancy	ANC checkups, Availability of staff at facility centre, Education, Occupation, Family structure
Result of pregnancy	No. of children ever born, No. of surviving children
Duration of stay in the hospital	No. of surviving children, Family type, Per capita income
Weight of new born	Level of hemoglobin, Family type, Per capita income
Weight of new born	3 ANC checkups
Weight of new born	Education, Occupation
Children not breast fed within one hour	Weight of new born, Availability of staff
Weight of pregnant women	Weight of new born, Income status
No. of Delivery conducted by staff nurse	Weight of new born, Income status
No. of Delivery conducted by Doctor	Weight of new born, Income status
No. of Delivery conducted by ANM LHV	Weight of new born, Income status

Table 6: Measurements for Post Natal Care (PNC)

Facts to be measured	**Dimensions of Analysis**
no. of follow up visits	availability of staff, Occupation, Per capita income, Weight of new born
Follow up visit 1	Availability of staff, Occupation, Per capita income, weight of new born
Follow up visit 2	Availability of staff, Occupation, Per capita income, weight of new born
Follow up visit 3	Availability of staff, Occupation, Per capita income, weight of new born
Follow up visit 4	Availability of staff, Occupation, Per capita income, weight of new born
Danger sign in child	Danger sign in mother, No. of follow up visits
Danger sign in mother	Education, Occupation, Per capita income, No. of surviving children
Weight of child	Danger sign in mother, Education, Occupation, Per capita income, No. of surviving children
Immunization up to one year	Education, Occupation, Per capita income, No. of surviving children
No. of women distributed IFA tablets	Education, Occupation, Per capita income, No. of surviving children
Danger sign in mother	PPH, fever sepses
Danger sign in child like jaundice hypotherma diarrhea vomiting fever	Danger sign in mother, Education, Occupation, Per capita income, No. of surviving children
Weight of child	Danger sign in mother, Education, Occupation, Per capita income, No. of surviving children
Outcome of delivery still births	Education, Occupation, Per capita income, No. of surviving children
Outcome of delivery Live births	Education, Occupation, Per capita income, No. of surviving children

4.2 Data warehouse Design Methodology for NRHM Project

NRHM programme implementation is nationwide covering 29 states 7 union territories with a population of about 1.25 billion. To ensure true representation of the whole population in terms of various health facilities provided across the country, three different states namely Uttar Pradesh, Rajasthan, and Madhya

Pradesh were chosen for interaction with the business managers/ user groups to ensure comprehensive investigation and getting truly representing requirements. Methods applied for analysis of user requirements and finding out user expectations from the data warehousing solutions include:

1. Reference of organizational documents like NRHM Journey so far, Monitoring and Evaluation, Family welfare statics 2011, Directory of Innovations Implemented in the Health Sector, Health programme manager's manual, Integrated MIS reporting etc. to know organizational goals and objectives, mode of functioning and current reporting abilities and user's expectations from data warehousing solution.
2. Personal interaction with the business managers of the organizations at various levels in the organizational hierarchy like CMOs, Dy. CMOs, DPMs, DEO, HEOs, Nodal Officers, BAO, Statistical officer, Data Entry Operator.
3. Circulation of questionnaire and collection of response from various user groups in the organization.
4. Exploration of national and state level websites of NRHM maintained by central and state governments respectively.

Following steps were taken to start data warehouse designing for NRHM project and identify designing techniques required to overcome limitations of the existing operational system.

1. Investigation of operational system to assess its design quality.
2. Identification of operational system design limitations.
3. Investigation of reporting capabilities of operational system.
4. Determination of level of dependability on these reports
5. Identification of design techniques to overcome limitations of operational system.

4.2.1 Design of Questionnaire

For identification of user requirements and the kind of problems they have been facing with the current system, a Questionnaire was designed and circulated among various user groups of the system. The questionnaire covers questions mainly of the following categories:

What measurements you want to make? What are parameters on which you want to analyze these measurements? What strategic decisions you are required to make on the basis of above measurements? How would these measurements help in taking strategic decision making?

Which is most important time dimension in your domain of analysis? What are other important time dimensions? What type of reporting formats required?

V. Limitations Of Nrhm Operational System

By applying the steps and methods of investigation during analysis of operational system design with an objective to design data warehousing solution, some limitations were found in the operational system of the NRHM project. We have briefly summarized the identified limitations point wise.

i. All data items required for performance evaluation are not present in the operational system. ii. Presence of multiple isolated applications across different states resulting in heterogeneity.
iii. Many applications are partially implemented.
iv. Data submission from facility centres is not regular.
v. Some of the centres never submit data resulting incomplete data.
vi. Aggregated Data submission formats diminishing scope for drill down details.
vii. Non-uniform grain of data submission from different centres up in the hierarchy.
viii. The cases not availing the facilities from NRHM are not considered

VI. Techniques For Dealing With Operational System Design Limitations

Presence of identified limitations in the operational system hampers design and implementation of data warehousing solution. It is important to find out methods to deal with such limitations so that one can proceed for data warehousing system design and implementation with the assurance of quality solution even in the presence of these limitations. In this section we have discussed each of the identified limitation and its solution.

i. Non comprehensive consideration of data items in the operational system is a critical limitation. A simple technique to deal with such problem includes identification of data items that are important but have not been considered in the operational system. We then explore and find out the reason for its non consideration. If there is a justified valid reason then we need to find an alternate solution. Assessment of its requirement for data warehousing solution is to be done. If it is found to be an important and required data item for data warehousing solution, we need to make provision for its entry into data warehousing system. A practical approach of investigating all data entry forms in the operational system was followed, with an objective to determine list of data items that have provision for input and also the list of data items that are not entered into the system, but are required for generating reports as per user's expectations. Problem occurs when a particular data item has not

been considered just due to casual design approach during the operational system design. We can represent the solution for the problem of unavailability of required data items in the following pseudo code solution:

For all forms i= 1 to n
Check forms[i]
Make a list of data items on the form and name it list 1
Check the users reporting expectations related to the form under consideration
Make a list of data items required and name it List2
If (List2 is subset of list1)
Form[i] is complete
Else
(list2-list1) items are missing and need to be considered.

ii. Presence of multiple isolated applications running in the organization that are not integrated with each other: The approach used that could identify such problem in the operation system was interaction with the client organization by means of questionnaire and personal interaction. Business users revealed that there is not a single integrated system across all states, all states have got different operational systems developed tailor made to their local needs. First step of solution for such problem is to plan and execute migration of multiple operational systems of the organization to a single integrated application. Next, the ETL procedure needs to be written to transform data from all the existing operational systems in the data warehouse then design ETL for the integrated operational system.

iii. Incomplete or partial implementation of the functionalities: upon investigation of different operational systems in use, it was found that all operational systems were not having all the required functionalities.

Solution for such problem is to first identify the functional modules that have not been implemented but are actually required. We also need to investigate functionality of the existing functional modules to determine if they are complete. Accordingly new modules or additional code in the existing modules can be added in the operational system.

iv. Incomplete or partial submission of data in the operational system from different facility centers: Identification of such problems in the system is possible by investigating the data submitted by different centres with the objective to check their completeness. Two fold solution for this problem needs to be implemented.

We need to incorporate feature in operational system to report at pre-specified time intervals about details of non execution of data submission by different facility centers. We also need to develop and communicate among the concerned staff, about the business process protocols to ensure that action is taken on defaults on part of data submission and compulsion is enforced for execution of submission of data.

v. Non submission of data by the centres is similar to partial submission of data and could be resolved in the same way.

vi. Aggregated data submission formats have been taken in HMIS based formats which is not an online application but is a reporting facility for the work executed during a specified time interval like month, quarter, or a year. If aggregated data figures are reported, details cannot be extracted from the data because of loss of details due to aggregation. The solution for this problem is possible by way of implementation of an online application with the provisions of recording atomic level data items against facilities being provided at each of the facility centre.

vii. Variable grain across different reporting centres up in the hierarchy: The solution to the problem of varying level of grain is similar to the solution for aggregated data reporting. Solution to this problem is possible by implementation of an online application with the provisions of recording atomic level data items against facilities being provided at each of the facility centre.

viii. For cases not availing the facilities from the NRHM affiliated centres, a separate survey needs to be conducted for collection of required data. Such data then needs to be fed into the system and them loaded into data warehousing solution. An interface could be developed in the operational system to facilitate the centres not affiliated to NRHM to directly upload data eliminating need of additional surveys.

VII. Conclusion

Operational systems provide input to data warehousing solutions. They need to be designed and implemented well both for the purpose of operational efficiency of the organization and also for effective data warehousing. Often we find some limitations in the organizational operational system that affects design and implementation of data warehousing solution. We have followed a systematic practical oriented approach of operational system quality assessment to identify such limitations. Each of the identified limitation has been analyzed thoroughly. We have designed techniques to handle such limitations of operational system, so that the data warehouse system development can be done smoothly. If limitations of operational system are handled effectively, they would result in the effective data warehousing solutions. In future, we plan to determine how effective are the identified techniques in designing data warehouse schema and the overall data warehousing solution for NRHM project to overcome such limitations of its operational system. It was seen that most of the limitations of the operational system were attributed to large size of the organization, geographic distribution of services, poor network connectivity, lack of systematic application architecture design, poor analysis of requirements to be implemented in the operational system and lack of vision regarding long term data analysis needs. All above factors combined together, resulted in the limitations in operational system. Once these limitations and issues are attended as per the given methods and techniques, it will not only improve the operational system efficiency but would also result in an effective and useful data warehousing solution. Future plans of this work include design of data warehousing solution for NRHM project under existing limitations

References

[1] Shahina Ferdous, Leonidas Fegaras, Fillia Makedon "Applying Data Warehousing Technique in Pervasive Assistive Environment" PETRA'10, June 23–25, 2010, Samos, Greece. Copyright © 2010 ACM ISBN 978-1-4503-0071-1/10/06... $10.00

[2] Farrah Farooq, Syed Mansoor Sarwar "Real-Time Data Warehousing For Business Intelligence" FIT '10, 21-DEC-2010, Islamabad, Pakistan Copyright © 2010 ACM 978-1-4503-0342

[3] Peter Wai Yee Wong*, Ric Hendrickson*, Haider Rizvi**, Steve Pratt* "Performance Evaluation of Linux File Systems for Data Warehousing Workloads" INFOSCALE '06. Proceedings of the First International Conference on Scalable Information Systems, May 29-June 1 2006

[4] Matteo Golfarelli, Stefano Rizzi, Iuris Cella "Beyond Data Warehousing:What's Next in Business Intelligence?" DOLAP'04, November 12–13, 2004, Washington, DC, USA.

[5] Chung-Min Chen Munir Cochinwala Elsa Yueh" Dealing with Slow-Evolving Fact: A Case Study on Inventory Data Warehousing" DOLAP 99 Kanas city MO USA Copyright 1999 ACM

[6] Alexandros Karakasidi, Panos Vassiliadis, Evaggelia Pitoura "ETL Queues for Active Data Warehousing" IQIS 2005, June 17, 2005, Baltimore, MD, USA. Copyright 2005 ACM 1-59593-160-0/05/06

[7] Rami Rifaieh, Nabila Aïcha Benharkat "Query-based Data Warehousing Tool" DOLAP'02, November 8, 2002, McLean, Virginia, USA. Copyright 2002 ACM 1-58113-590-4/02/0011

[8] Matteo Golfarelli, Federica Mandreoli, Wilma Penzo "Towards OLAP Query Reformulation in Peer-to-Peer Data Warehousing" DOLAP'10, October 30, 2010, Toronto, Ontario, Canada. Copyright 2010 ACM 978-1-4503-0383-5/10/10

[9] Deepak Asrani, Renu Jain, & Usha Saxena "Limitations of Operational System Design and Requirements of a Successful Data Warehousing Solution: A Gap Analysis for NRHM" 4th International Conference on Emerging Trends in Engineering and Technology College Of Engineering ,Teerthanker Mahaveer University, 24-25 Apr, 2015

[10] STAMFORD, CONN February 24, 2005

[11] Predicts2012:business_intelligence_still_subject_to_non-technical_challenges http://www.saiindia.gov.in/english/home/our_products/Audit_Report/Government_Wise/union_audit/recent_reports/union_performance/2009_2010/Civil_%20Performance_Audits/Report_no_8/chap_9.pdf

[12] Deepak Asrani, & Renu Jain "Review of techniques used in data warehouse implementation: An initiative towards designing a frame work for effective data warehousing" IEEE International Conference on Advances in Engineering and Technology Research-ICAETR-2014, August 1, 2 - 2014