THE INTERNATIONAL JOURNAL OF SCIENCE & TECHNOLEDGE

Optimising University Intranet Services Using a Data Warehouse

Adekunle Joshua Akinjobi

Lecturer / Head of Department, Department of ICT/ Computer, Science, Crawford University, Igbesa, Ogun State, Nigeria

Abstract:

Valuable and fast intra-communication has been a problem among the University academic and non-academic staff, by not being kept updated regularly about events, inability to have standard information sharing and other related non-knowledge of students and staff activities. There are also problems emanating from lack of information integration that sometimes render intranet services ineffective. This paper discusses and recommends solutions to the aforementioned problems.

Diverse operational systems exist in the various departments of a University in different legacy formats. Utilising the Extraction, Transformation and Loading (ETL) facilities of SQL Server 2008 R2, data is extracted from the individual operational systems into a staging area, cleansed (removing errors and inconsistencies), transformed and integrated using a uniform format, and finally loading the integrated data into a data warehouse. The University intranet is optimised for efficiency by linking with the data warehouse as the user's interface for easy and speedier access and queries only by the University community. The intranet works on the existing University Local Area Network (LAN) infrastructure.

The data in the warehouse is subject oriented (stored by departments), nonvolatile (cannot be changed or modified), timevariant (stored according to the period of several years transactions) and up to date (also contains the most recent). Ensuring an optimised intranet service and having an integrated data warehouse would reduce the costs of data access, provide a uniform and timely information dissemination for collaboration and an overall effective University governance.

Keywords: ETL, Intranet, Information integration, LAN, Data warehouse, SQL Server 2008

1. Introduction

Most academic institutions lack a coordinated structure for information dissemination and delivery. Information dissemination such as invitation for academic meetings, bulletin of events, promotions, retirements, news of organisation restructuring, information on staff movements and new appointments, academic achievements, alerts and cautions on campus security, students admissions, budget development and requests and others are carried out by a dispatch clerk moving from one office to the other. When a staff particularly one without an assigned secretary is not on seat and lock his office, that information is not delivered; meetings are missed and urgent attention on critical issues become volatile. Important information from the top University to students and staff is presently done by congregating the entire staff for such to be disseminated.

Institutional learning requires that students collaborate to carry out assignments, exchange academic information, view lectures and examination timetable, view and retrieve results of semester examinations, view and print hand books and lecture notes, monitor meal schedules of the institution canteens, carry out on line course registration, get updated on campus events, know individual hostel allocation, evaluate lectures, get regular orientation and learning, among others.

University management can presently obtain vital information from operational units by calling or visiting the individual operators of such units. Such information could be duplicated and inconsistent.

1.1. Information Integration for the Intranet

The intranet is a way to access all University operational systems through the data warehouse which conceptually is an information integration repository.

All the students, colleges and other employees of the University are entitled to make use of its services and facilities. For example, in a typical relatively young University, operational systems include the library, exams and records, and ICT Resource center as major units in the university that can use the Intranet services. The library keeps tracks of book collections, reference works, technical reports etc. The exams and records stores student data including past exam questions, answer booklet and marking schemes. The ICT Resource center ensures that staff and students have access to a particular university hotspot over the same access point.

Others are bursary, departments and colleges, school of postgraduate and part time studies, chaplaincy, security, works and services, student's affairs, university farm and ventures, sports, quality control and academic planning, university clinic, hostel and canteen services, entrepreneurship unit, JUPEG, and the University Cooperative. All of these systems could be integrated.

An integrated information system contains information from all the operational units in a single repository such as a data warehouse that can be accessed by authorised personnel to promote knowledge sharing and productivity for staff and solve the communication

problems. Information integration enables queries and analysis of the data warehouse by management to discover trends and patterns of events. Since the data warehouse holds archival data of several years, discovered trends and patterns can reveal knowledge valuable for effective decision making. The dual functions of an intranet and data warehouse provide a reliable decision tool for management.

While data tables are built from some of the existing operational systems to build the data warehouse, interface programs are incorporated in the intranet structure to be able to access the data warehouse for information viewing, downloading, updating, and data querying to discover trends and patterns. The available intranet made to access the University data warehouse should be effectively managed through enforced procedures and routines.

An Intranet provide a controlled access and security to the institutional confidential data that is only meant for members of the institutional community. Intranet refers to a network that is accessible only by the organization's employees or others with authorization. A firewall surrounding the Intranet also prevents unauthorized access.

Using Intranet technologies linked to data warehouses significantly reduces the cost associated information deployment. Security and administration are centralized along with the elimination of the need to maintain client code on the end-user's computer. New users can be granted access more quickly and application updates are automatic and contribute to the reduction in deployment and maintenance costs. The compelling advantages in using the intranet for access are magnified even further in a data warehouse (Berson).

1.2. Optimization of the Intranet

Optimization of the intranet in the University under case study starts with the physical inspection of the network installation round the institution to determine total network coverage. This involves measurements from the main gate to the administrative block, to the library and to the hostel, taking note of the internet facilities available and the present internet status.

The population of staff and students are obtained in order to determine bandwidth usage and management.

Patch panels should be fixed in the administrative buildings to serve the offices and another in the college buildings to serve the classrooms and academic staff offices, Racks should be fixed across the buildings to implement a star network topology. The fiber optics system (for a fast-medium transfer of data) should be introduced later to extend the network to the hostels so as to implement a Wireless LAN network. The entire networks are expected to also have communication facilities like Switches, Routers, Modem, PoE and antennas. All these are enough to support the intranet to make communication easier.

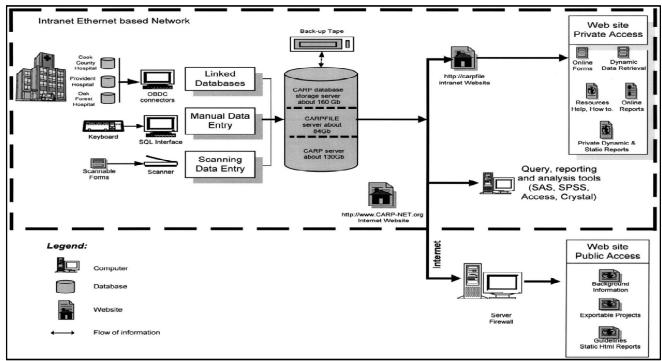


Figure 1: Data bases integration and Intranet access
Source: Mary F. Wisniewski, MSNet.all in Development of a Clinical Data Warehouse......

2. Literature Review

Intranets have become an integrated part of almost every organization's information infrastructure. When intranets were introduced a little more than a decade ago, they were hailed as the ultimate solution to almost any organizational issue; including anything from dissemination of management vision to integration of seemingly incompatible computer systems. William (Butchanan2000). Several works have been carried out on the subjects relating the intranets with a data repository like the data warehouse. Some few of them are cited as follow.

In a paper titled 'A Data Warehouse Design for a Typical University Information System' Bassil (2011) posited that in order to perform data mining, regular databases must be converted into data warehouse. The paper presents a design model for building data warehouse for a typical university information system based on transforming an operational database into an informational warehouse useful for decision makers to conduct data analysis, prediction, and forecasting. The proposed model considers data migration as comprising: Data extraction, data cleansing, data transforming, and data indexing and loading. The system is implemented under MS Access 2010 and is meant to serve as a repository of data for data mining operations.

Providing Access to an Academic Data Warehouse Over the Web by Rhodes& Kincheloe ()

Use the SAS System®, to develop a web-based Enterprise Information System for University decision maker to evaluate various programs and academic departments over time. The system serves as the access point into the University data warehouse, and includes graphical and tabular

reports with drilldown and reach through capability into individual student records. The only requirement on the client side is a web browser, with Java required to view the interactive graphics.

Data Warehousing and Web Enablement: Opportunities, Issues, and Trends, a paper of Nemati, Cannoy & Delk (2014) consider opportunities, issues, and trends in Data Warehousing and Web Enablement by looking at the web becoming the biggest contributor towards the growth of business intelligence and data warehousing. They propose that using Internet/Intranet technologies to implement data warehouses is significantly reducing the cost associated with information deployment. This is because security and administration are centralized along with the elimination of the need to maintain client code on the end-user's computer. New users can be granted access more quickly and application updates are automatic and contribute to the reduction in deployment and maintenance costs.

Data Warehousing and Data Mining - A Case Study by Suknović, Čupić, Martić& Krulj (2005) shows the design and implementation of data warehouse as well as the use of data mining algorithms for the purpose of knowledge discovery as the basic resource of adequate business decision making process. Realized for the needs of Student's Service Department of the Faculty of Organizational Sciences (FOS), University of Belgrade, Serbia and Montenegro, the system represents a good base for analysis and predictions in the following time period for the purpose of quality business decision-making by top management. The first part of the paper shows the steps in designing and development of data warehouse of the mentioned business system. The second part of the paper shows the implementation of data mining algorithms for the purpose of deducting rules, patterns and knowledge as a resource for support in the process of decision making.

Azeta, Omoregbe, Oluranti &Bello (2015) in their paper titled 'An Intranet Portal for a Learning Institution', develop an intranet portal that handles communication and interaction for a learning institution and also help students improve academically from their comfort zone and outside the lecture hall. This application enables lecturers and students to interact better. The system serves as a medium on which information is communicated and shared by keeping members of the institution regularly updated with current information on happenings and events in the institution. The application will also facilitate collaboration among learners.

Interactive Intranet Portal for effective Management in Tertiary Institution by Philipa, Kenneth & Agajo (2015) shows an enhanced and interactive method of analysing problems of result processing, tuition fee payment, library resources management by generating an interface. They use a software containing several modules such as Library Console, Admin, Staff, Course Registration, Checking of Results and E-News. The server shall be accessible by client computers on the intranet via thin-client browsers through a well-secured authentication system.

3. Method

3.1. Identification and Collection of Data Sources

The following data sources are considered and selected to build the data warehouse.

A typical young university shall have data from the following selected operational units:

- i. Registry: Students course and admissions data.
- ii. Bursary: Salary management, General ledger (containing financial balances, actual and estimated budget), Journal detail information, Billing and Receivables, Students Financial data, Property management asset (building, fixed, movable and land assets).
- iii. College (academic and nonacademic): College staff personal data plus the Payroll and Salary Management data from bursary.
- iv. All data collection contains current and historic information.

The Warehouse is organized by subject areas by data collections and tables from various collections are joined together to provide integrated information about the University.

3.2. Tables Creation

Because the data is stored in tables, it is easy to access just the data needed, rather than having to plow through all the data in the data collection. A table may be a part of more than one data collection.

Some data elements in tables may be indexed. Indexing enables the system to execute queries faster

- Assets Data Collection: Asset name, Asset type, Location, Asset value, Date acquired
- Student data collection: Name, Matric Number, Birth Date, Address, Date enrolled, Course Enrolled, Requirements fulfilled?, Department, Head of Dept, Course level adviser,
- > College Data Collection: Staff Name, Gender, Address, Birth date, Employment Date, Committee memberships, Date of last

- leave, Highest qualification, Last promotion date
- General Ledger Data Collection: Account Name, Account Number, Fiscal year, Fiscal month, Budget actual, budgets estimate, Year to date GL balance,
- > Salary Management Data Collection: Employee name, Employee number/ID, Job status, Post held, Department, Salary grade level, Name of bank,
- Admissions Data Collection: Applicant name, Applicant Number/ID, Gender, Date of birth, Program Applied for, JAMB score, UTME score, Last school attended,
- iv. University Clinic Data Collection: Student Name, Home address, Parents' Phone Number, Gender, Ailment Type, Date of diagnosis, Prescribed drug, Hostel Number, Fitness Date
- v. University Canteen: Canteen Name, Location on Campus, Breakfast Type, Lunch Type, Dinner Type, Breakfast Average Price, Lunch Average Price, Dinner Average Price, Meal Day

Input Variables	Data Type	Content Type	Usage
Name	Alphabetic	Discrete	
Matric Number	Numeric	Discrete	
Birth Date	Alphanumeric	Discrete	
Address	Alphanumeric	Discrete	
Date Enrolled	Alphanumeric	Discrete	
Course Enrolled	Alphanumeric	Discrete	
Requirements fulfilled	Alphabetic	Discrete	
Department	Alphabetic	Discrete	
Head of Dept.	Alphabetic	Discrete	·
Course Level Adviser	Alphabetic	Discrete	

Table 1: Student Table

Input Variables	Data Type	Content Type	Usage
Name	Alphabetic	Discrete	
Gender	Alphabetic	Discrete	
Address	Alphanumeric	Discrete	
Birth Date	Alphanumeric	Discrete	
Employment Date	Alphanumeric	Discrete	
Committee Membership	Alphabetic	Discrete	
Date of last Leave	Alphanumeric	Discrete	
Highest Qualification	Alphabetic	Discrete	
Last promotion Date	Alphanumeric	Discrete	

Table 2: College Table

Input Variables	Data Type	Content Type	Usage
Account Name	Alphabetic	Discrete	
Account Number	Alphanumeric	Discrete	
Fiscal Year	Alphanumeric	Discrete	
Fiscal Month	Alphanumeric	Discrete	
Budget Actual	Currency	Continuous	
Budget Estimate	Currency	Continuous	
Year to date Balance	Currency	Continuous	
Journal Entries	Currency	Continuous	

Table 3: General Ledger Table

Input Variables	Data Type	Content Type	Usage
Employee Name	Alphabetic	Discrete	
Employee Number/ID	Numeric	Discrete	
Employment Category	Alphabetic	Discrete	
Post held	Alphabetic	Discrete	
Department	Alphabetic	Discrete	
Salary Grade Level	Alphanumeric	Discrete	
Bank Name	Alphabetic	Discrete	

Table 4: Salary Table

Input Variables	Data Type	Content Type	Usage
Applicant's Name	Alphabetic	Discrete	
Applicant Number/ID	Numeric	Discrete	
Gender	Alphabetic	Discrete	
Date of Birth	Alphanumeric	Discrete	
Program Applied for	Alphabetic	Discrete	
JAMB Score	Numeric	Discrete	
UTME Score	Numeric	Discrete	
Last School Attended	Alphabetic	Discrete	

Table 5: Admissions Table

Input Variables	Data Type	Content Type	Usage
Asset Name	Alphabetic	Discrete	
Asset Type	Alphanumeric	Discrete	
Asset code	Alphanumeric	Discrete	
Asset Location	Alphanumeric	Discrete	
Asset Value	Currency	Continuous	
Date Acquired	Alphanumeric	Discrete	

Table 6: Asset Table

Input Variables	Data Type	Content Type	Usage
Student's Name	Alphabetic	Discrete	
Home Address	Alphabetic	Discrete	
Parents' Phone Number	Numeric	Discrete	
Gender	Alphabetic	Discrete	
Ailment Type	Alphanumeric	Discrete	
Date of Diagnosis	Alphanumeric	Discrete	
Prescribed Drug	Alphabetic	Discrete	
Hostel Number	Alphanumeric	Discrete	
Fitness Date	Alphanumeric	Discrete	

Table 7: Clinic Diagnosis and Prescription Table

Input Variables	Data Type	ContentType	Usage
Canteen Name	Alphabetic	Discrete	
Location on Campus	Alphabetic	Discrete	
Breakfast Type	Alphabetic	Discrete	
Dinner Type	Alphabetic	Discrete	
Breakfast Average Price	Currency	Continuous	
Lunch Average Price	Currency	Continuous	
Dinner Average Price	Currency	Continuous	
Meal Day	Alphabetic	Discrete	

Table 8: Canteen Table

Meta Data

This describes the basic contents of the table. It also gives the following information:

- → Explanation. Describes the physical makeup or content of the table.
- → Primary key. Lists the data elements that are the primary keys in the table.
- → Indexed data elements. Lists the data elements that are indexed in the table. Since tables can consist of many rows, queries that include record selection conditions based on indexed data elements provide faster results.
- → Related tables. Identifies other tables that may be used for query. For example, using the Enrollment table to list students in a specific course section, the Person table may be used to get the students' names.

4. Finding and Discussion

4.1. Building the Data Warehouse

4.1.1. Data Collections

The Warehouse contains the archived databases built from relational tables of the operational transactions data. The information from

each transaction system is referred to as a data collection. For example, information from Student Records System is referred to as the Student Data Collection, while account balance information from the General Ledger is referred to as the General Ledger Data Collection

Access to a specific data collection depends on individual unit authorization level. If there are questions about a specific data collection or want access to a collection, a contact is made to the Data Administration of the operational unit. The final data warehouse is an SQL Server 2008 R2 database, which uses SQL Server Integration Services(SSIS) as the ETL tool and SQL Server Reporting Services as the reporting tool. It is used for the creation of the dimensional and fact tables with their related structures from the data base tables, followed by the creation of integration service project such as data-source and data-source views using the Business Intelligence package in the SQL server.

The University server uses the Microsoft Windows NT® operating system and is used to store the large amounts of data required, and through the SSIS of the SQL Server, also aids in transporting data from the University operation systems databases to the Warehouse. The access to the warehouse is delivered over the University intranet,

Since application in each operating unit runs on a Windows-based intranet, the Windows integrated authentication is used for database access. Integrated security uses the current Windows identity established on the operating system thread to access the SQL Server database. The Windows identity is then mapped to an SQL Server database and permissions.

To connect to SQL Server using Windows integrated authentication, the Windows identity is done under which the ASP.NET application runs. The identity would have been granted access to the SQL Server database.

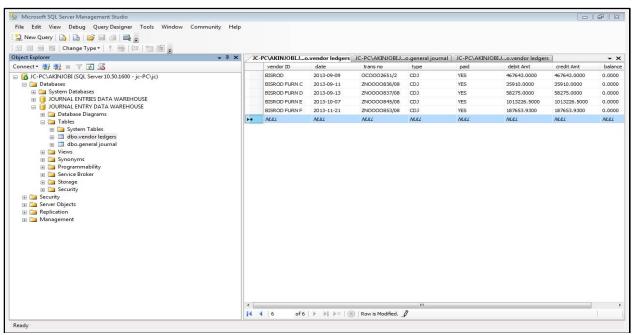


Figure 2: Using SQL Analysis Services to build the Data warehouse

4.2. Building a General Ledger Database Example

The Journal Entry data in the general ledger is extracted to an SQL Server database using SQL Server Integration Services (SSIS). The SQL Server Analysis Services use the tables to create the schema for the journal entries database.

4.2.1. Creating Content for the Vendor Ledgers Table

This process enables populating the table structure with the relevant data. For example, the vendors' ledger table contains the vendor identification (vendor ID), date of posting, transaction

number, type of transaction, etc. This is required for the data warehouse creation.

4.2.2. Connecting to the SQL Server

Since the SQL Server is on the University server, the Windows identity must be able to flow across the network to the remote instance of SQL Server.

Storage components are the databases that contain configuration information and jobs (msdb) as well as the collected performance data (MDW).

Execution components manage the data collection and storage operations.

API components enable interaction between the user interfaces and the data collector These components include the SQL Server metadata.

Client components are the user interfaces for managing the data collection and presenting the reports.

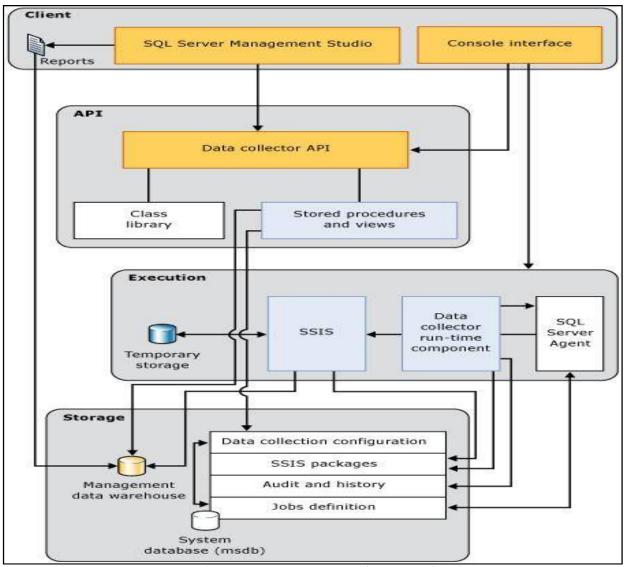


Figure 3: Connecting SQL Server and Data Warehouse Source: SQL Server 2008 Management Data Warehouse

4.2.3. The Intranet implementation

The following are materials required in each operating systems of the University for the intranet implementation:

- i. Laptops with WIFI and WLAN capabilities, with Operating system of windows 2000 or higher, also with windows server 2012, RAM 2GB and HDD Memory of 80GB to enable the installation of the intranet collaboration tools.
- ii. A switch and crimped twisted pair cables.
- iii. Installed intranet collaboration tools i.e. Team-Viewer, LAN talk and network chat.
- iv. If the intranet is desired to be used with a switch such as a D-Link 24 Port Switch, the following steps are followed:

Plug the switch to a power outlet and connect, then switch to a configured TP-LINK router via the Cat 5e Ethernet cable connection.

v. Open the network sharing centre from the control panel accessible from the windows menu.

Select, the connect to a network option and a list of the available networks would be displayed.

Select the network named intranet and the laptop would automatically connect.

vi. IP Address should be issued to computers connected to the network for identification; this is done by selecting the network properties of the network in use and inserting an Internet Protocol address issued to you by the network administrator.

vii. Network Properties (TCP/IP)

Select properties for the internet protocol version 4(TCP/IPV4) and insert the right IP address.

Open the required intranet collaboration tool i.e. either network chat, team viewer.

The University chosen ICT Resource Centre technical personnel is responsible for configuring Computer workstations in each operating unit to access the Data Warehouse.

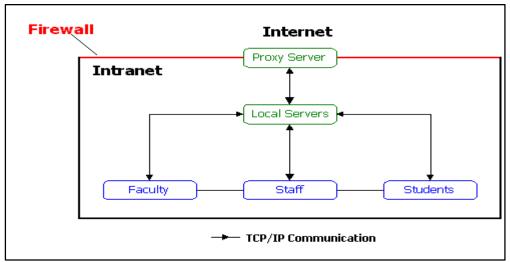


Figure 4: Model of a University Intranet implementation

4.2.4. Systems Administration

The System administrator uses the SQL Server facilities through the Intranet Interface to monitor the operating unit's activities. Some of the activities are illustrated below:

4.3. Viewing the Integrated Data or Data Collection

This shows the graphical reports available to view the integrated data. The administrator can look at the reports across the operating unit's data base, by right clicking.

Figure shows you an example of an overview report spanning the entire operating units. The Management Data Warehouse Overview report displays the last upload time so that you can get a quick status of all the upload activity to the central MDW. The reports are also accessible via the right click menu on the Data Collection node for the target instance.

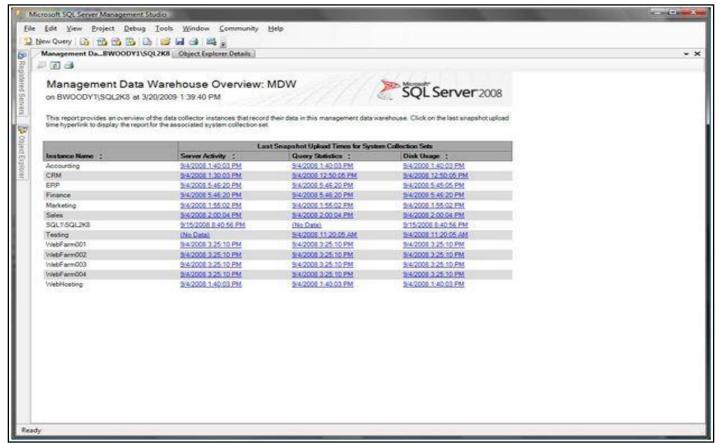


Figure 5: Management Data Warehouse reports, spanning SQL Server instances Source: SQL Server 2008 Management Data Warehouse

4.4. Server Query Statistics

This shows the various queries made to the data warehouse. It assists the systems to create the log file and used for the audit trail.

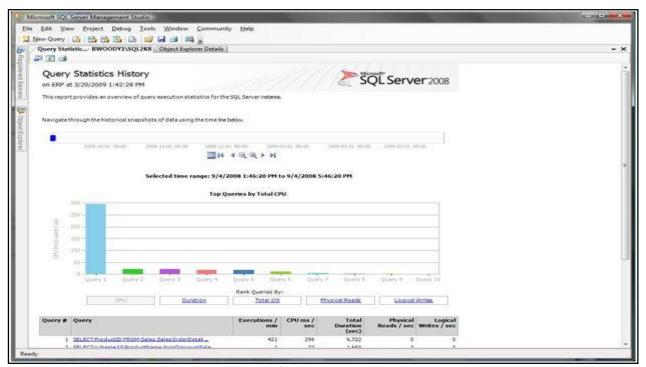


Figure 6: Server Query Statistics Source: SQL Server 2008 Management Data Warehouse

4.5. Server Usage and Activity

The Server Activity report shows the system resource usage information and SQL Server wait statistics.

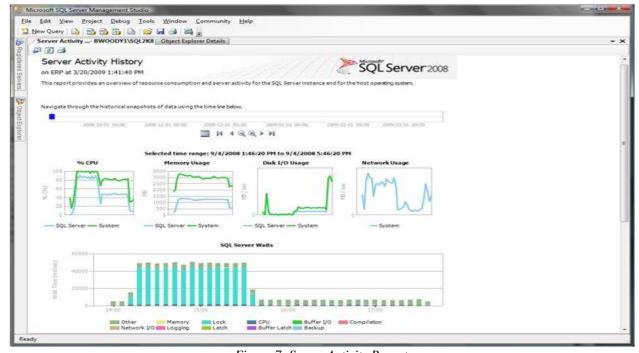


Figure 7: Server Activity Report Source: SQL Server 2008 Management Data Warehouse

By selecting one of the system resource reports, it enables a drill down into details of CPU, Memory, Disk I/O, and Network Usage. iv. Disk Space Usage

Clicking on the last Disk Usage upload time for any of the server instances, a report that indicates the data file size and log size for

each database on the instance at the time the last upload made is analysed.

All of these reports can also be accessed from the Data Collection Node of the Management folder in the SQL Server Management Studio by: Right clicking on Data Collection, selecting Reports, clicking on Management Data Warehouse, and then selecting one of the Collection Sets. Custom reports can similarly be created using SQL Server Reporting Services and querying the Management Data Warehouse directly.

Data Warehouse activities by the operating unit systems users

There are numerous activities that can be carried out in each of the operating units using the intranet user interface with the SQL Server. Only three of them are cited here.

i. Identifying Source components

Source Components at the operating units consist of the source database, table name, column name, and data type. The source database is a name of the instance of the database where data is located and it is a connection string necessary to connect to the database. The table and column names are required to populate the target in the target data warehouse

ii. Connecting to the Data base

A connection is set up between the intranet interacting tools and the SQL Server database before beginning to define, adding columns to and creating the model. The Intranet Tool uses

the server and database information provided to create a connection to SQL Server and Analysis Services database.

4.6. Invoking the SQL Server Management Studio

- Launch the windows start button
- Click All Programs on the task bar
- Click Microsoft SQL Server 2008 R2
- Click SQL Server Management Studio
- Connect to server

4.7. Connection of Intranet to the SQL Server Database



Figure 8: Connecting to a Data base server.

Valuable and fast intra-communication has been a problem among the University academic and nonacademic staff, by not being kept updated regularly about events, inability to have standard information sharing and other related non-knowledge of students and staff activities. This paper also discusses and recommends solutions to problems emanating from lack of information integration that sometimes render intranet services ineffective in higher institutions.

5. Conclusion and Suggestion

The paper has discussed and recommended solutions to problems existing among the University academic and non academic staff of not being kept updated regularly about events, inability to have standard information sharing and other related non-knowledge of

students and staff activities. It also discussed the lack of information integration that sometimes encourage increased information costs and information duplication thereby rendering intranet services where they exist, ineffective. Because diverse operational systems exist in the various departments of the University in different legacy formats, facilities embedded in SQL Server 2008 R2 were used to build a data warehouse as a data integration repository. Methods of optimising the University intranet for efficiency were identified by linking the intranet with the data warehouse. The intranet as the user's interface to the data warehouse, works on the existing University Local Area Network (LAN) infrastructure.

6. References

- i. Bassil, Y (2012). A Data Warehouse Design for A Typical University Information System
- Youssef Bassil. Journal of Computer Science & Research (JCSCR) ISSN 2227-328X http://www.jcscr.com Vol. 1, No. 6, Pages. 12-17.
- iii. Rhodes, P., & Kincheloe, F. (). Providing Access to an Academic Data Warehouse Over the Web
- iv. Baylor University, Waco, TX. Paper 126-25
- v. Nemati, H.R., Cannoy, S., Delk, R., (2014). Data Warehousing and Web Enablement: Opportunities, Issues, and Trends. The University of North Carolina, Greensboro, NC 27014
- vi. Mehedintu, A.,Buligiu, I.,Pîrvu, C (2008). Web-enabled Data Warehouse and Data Webhouse. Revista Informatica Economică nr. 1(45)/2008
- vii. Philipa, I.O., Akpado, K., James Agajo, J (2011). Interactive Intranet Portal for effective Management in Tertiary Institution. International Journal of Advanced Computer Science and Applications(IJACSA), Vol. 2, No. 6.
- viii. Azeta, A.A., Omoregbe, N. A., Oluranti, J., Bello, O. A(2015). An Intranet Portal For A Learning Institution. Proceedings of ICERI2015 Conference, 16th-18th November 2015, Seville, Spain, ISBN: 978-84-608-2657-6
- ix. Suknović, M., Čupić, M., Martić, M., Krulj, D (2005). Data Warehousing and Data Mining A Case Study. Yugoslav Journal of Operations Research. 15 (2005), Number 1, 125-145
- x. Mirabedini,S(2014).The Role of Data warehousing in Educational Data Analysis.Journal of Novel Applied Sciences Available online at www.jnasci.org ©2014 JNAS Journal-2014-3-S1/1439-1445 ISSN 2322-5149 ©2014 JNAS